Journal of Safety Research 83 (2022) 57–65

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

An analysis of the New York City traffic volume, vehicle collisions, and safety under COVID-19

Paolo Cappellari, Bryan S. Weber *

City University of New York, College of Staten Island, 2800 Victory Blvd, Staten Island 10316, NY, USA

ARTICLE INFO

Article history:
Received 26 May 2021
Received in revised form 25 October 2021
Accepted 4 August 2022
Available online 10 August 2022

Keywords:
COVID-19
Traffic
Collisions
Smart city

ABSTRACT

Introduction and Method: We use the arguably exogenous intensity of COVID-19 as an instrument in order to study the relationship between traffic volume and vehicle collisions in a large metropolitan area. We correlate data from multiple sources and consider a time interval ranging from about one year before to one year after the pandemic breakout, which allows to account for preexisting seasonal patterns as well as the disruption brought by the pandemic. *Results:* We identify that increased traffic volume is associated with significantly more collisions with a robust elasticity varying between 1.2 and 1.7. At the same time, higher traffic volumes are associated with a significant reduction in casualties. Conversely, low traffic volumes are associated with high speeds and with particularly dangerous collisions. In terms of social cost, we separately calculated the cost of property damage and casualties. We measured that the reduction in the per-day social cost of collisions during the COVID-19 period is approximately $453,000 in property damage. However, the increase in casualties from collisions at lower traffic volumes are worth approximately $2.6 million in injuries and fatalities, entirely offsetting any benefit from reduced collisions. *Practical Applications:* This research provides valuable insights that policy makers may take into consideration when shifting traffic volume in relation to social cost and safety, such as congestion taxes.

1. Introduction

The analysis of traffic congestion and associated collisions has been the interest of research for a number of decades (Cui, Henrickson, Ke, & Wang, 2018; De Fabritiis, Ragona, & Valenti, 2008; Jain, Sharma, & Subramanian, 2012), and continues to be an area of great interest (Faghih-Imani, Anowar, Miller, & Eluru, 2017; Mammen, Shim, & Weber, 2019). Studies focus on a number of aspects such as traffic patterns predictions (Jain et al., 2012), traffic flow speed estimation (Cui et al., 2020; De Fabritiis et al., 2008), relationship between traffic volume and collisions (Noland & Quddus, 2004; Shefer & Rietveld, 1997; Wang, Quddus, & Ison, 2009), and more (Baghestani, Tayarani, Allahviranloo, & Gao, 2020; Liu, Zheng, Chawla, Yuan, & Xing, 2011).

Many of these studies rely on potentially endogenous factors (Cullinane, 2004), such as the introduction of car sharing services (Dills & Mulholland, 2018), taxi ridership (Faghih-Imani et al., 2017), public transit usage (Iyer, Boxer, & Subramanian, 2018), policy making (Abouk & Adams, 2013; Mammen et al., 2019), or eco nomic impact (Parry & Bento, 2002), whose effects are usually visible over a relatively long period of time, making it hard to understand if, and which, other factors may contribute to the changes.

In this paper, we use the data associated with the COVID-19 pandemic as an exogenous factor to analyze the traffic volume and vehicle collisions in the boroughs of New York City. This is not because we are interested in pandemic traffic patterns per se, but because the pandemic provides a substantive adjustment to traffic volume without other associated traffic rules changing, allowing for a consistent estimation of the association between traffic volume and vehicle collisions (Cullinane, 2004). The New York area was hit with the pandemic starting March 2020. The COVID-19 pandemic and the associated health and public policy implications created an unprecedented scenario that gives us the ability to observe the largest shift in vehicle traffic that is due to an exogenous factor. This is novel in a literature where changes to traffic volume are mostly linked with new transportation or economic policies (Green, Heywood, & Navarro, 2016), or the intertemporal shifts in volume that may be accompanied by other traffic changes (Shefer & Rietveld, 1997; Zhou & Sisiopiku, 1997). In particular, we analyze the relationship between the number of COVID-19 cases, its impact on traffic, and the associated collision between vehicles. The source data for our analysis are data sets from both New York city and state. Specifically, we consider the following distinct data sets: traffic volumes, vehicle collisions, traf-

* Corresponding author.
E-mail addresses: paolo.cappellari@csi.cuny.edu (P. Cappellari), bryan.weber@csi.cuny.edu (B.S. Weber).

https://doi.org/10.1016/j.jsr.2022.08.004
0022-4375/

fic speeds, traffic summons issued, and COVID-19 testing reports. These data sets are combined together to generate a comprehensive and unique view of traffic changes around the time period the exogenous factor is observable. All data sets are publicly available to enable full verifiability and reproducibility of our study.

For this study, we used an instrumental variable (IV) approach, where we infer the traffic volume by the number of COVID-19 cases on the week prior (Cullinane, 2004). We control for standard variables, including time based-fixed effects, the number of traffic summons, as well as for several possible systemic breaks at the time of the first COVID-19 case and the stay-at-home order. We also consider additional robustness checks such as changing the data window, additional leads of potentially obscuring events, and variations on the functional form of the instrument.

Among the findings, we observed that the reduction in traffic volume is associated with a significant reduction in collisions, at a rate of approximately 1.7% fewer collisions for each 1% reduction in traffic volume. Conversely, we measure an increase in fatalities and injuries as traffic volume decreases, suggesting higher volumes of traffic has a dampening effect on collisions. We also find the suggestion that the casualty/collision rate has increased in the lower volume traffic. Corroborating on this finding, we point out that speed is generally higher in lower-volume traffic, and significantly increased speeds in each borough are associated with declines in traffic volume. The simple elasticity of speed with respect to volume is approximately equal to -0.13, suggesting that the policies still may be double-edged. Lastly, we estimate that the estimated monetary value in collisions, injuries, and casualties from the COVID-19 traffic declines across all four boroughs (excluding Staten Island).

Despite the reduction in collisions, we find an increase in social costs during the period. We calculate a reduction of \$453,000 in property costs (excluding casualties) and the counter-veiling value of increased fatalities and injuries during this period are approximately \$2.6 million. This highlights the complex relationship between traffic volume, speed, and safety, but indicates that the overriding effect of substantive volume changes does not guarantee improved safety.

The paper is organized as follows: in Section 2 we discuss the state of the art on the research in this area; in Section 3 we present the data sets used in this study; in Section 4 we illustrate the methodology to analyze the data; in Section 5 we discuss the findings and observations; finally, in Section 6 we draw our conclusions.

2. Literature review

Many works focus on the relationship between traffic, safety, and economics. In Shefer and Rietveld (1997), authors highlight the complexity of the relationship between congestion and collisions. The authors highlight three stages of congestion, named Stage 1, Stage 2, and Stage 3. Stage 1 describes low density traffic at presumably high speed with substantial variance in speed. In this stage, there are so few cars that fatalities are rare, but additional cars greatly contribute to the fatality rate. Stage 2 describes moderately dense traffic, where additional cars lead to additional fatalities, but mitigate the speed (and speed variance). Stage 3 describes gridlocked traffic, such that fatalities rarely occur since the speed (and its variance) is essentially zero. Shefer and Rietveld appeal to several national hour-of-day traffic decompositions to make their case; and highlight the fundamental work of Vickrey (1969) and Pigou (1920) on appropriate approaches to obtain an optimal level of congestion. The general association between traffic density and collisions was identified independently around that time period by Zhou and Sisiopiku (1997), who analyzed traffic patterns on the interstate I-94 in Detroit. Still, this leaves a great deal of empirical work to do in identifying the marginal effect of additional traffic volume.

More recent work, Green et al. (2016) uses a difference in differences approach against a public congestion tax in order to highlight the broad safety improvements of the congestion-reducing tax. Several other researchers investigate the role of additional public transit services in safety improvements (Anderson, 2014; Bauernschuster, Hener, & Rainer, 2017; Jackson & Owens, 2011; Lichtman-Sadot, 2019), and have managed to find that public transit services simultaneously reduce traffic volume and improve safety. Edlin and Karaca-Mandic (2006) further builds upon the relationship between congestion and traffic safety to calculate the marginal insurance premium for an additional automobile in the state of California. These successes do not mean that the relationship is easily estimable. Noland and Quddus (2004) performed a spatial analysis of London and found suggestive evidence of casualties (fatalities or injuries) being associated with congestion, with the latter inferred through regional population and employment. Finally, Wang et al. (2009) found no evidence of an association of traffic congestion on fatalities in the M25 motorway in England.

Compared to the state of the art, our work differs from the existing research for the following main factors: first, our between-day variation in vehicle traffic is substantially larger than the previous papers we mention (we see a decline in traffic that falls to nearly 16% of the original peak), giving us a broad base for estimation; second, our traffic variation comes not from public policy targeting traffic or the potentially endogenous traffic density, but instead from the entirely uncorrelated daily variation in the severity of the COVID-19 crisis the day prior.

3. Data sets

In this section, we present the data sources used in this study. We considered a total of five different data sets, specifically: motor vehicle collision, number of vehicles on the road, vehicle speeds, traffic related summons, and COVID-19 test reporting. In total, we collect these data for the period between 2019/01/01 and 2021/02/21, approximately 1 year before and after the first case of COVID-19. In the reminder of this section, we briefly discuss each data set and its nature. The Motor Vehicle Collision data set contains the records of collisions occurring between vehicles within the city of New York (NYPD, 2020). The data set is maintained and provided by New York Police Department under the Open-Data (2020) initiative. The NYC Open Data initiative is meant to provide free and transparent access to data from the city and the administration to residents and beyond. By New York City law it is mandatory to report collisions where someone is injured or killed, or where there is at least \$1000 worth of damage, which makes this data set a fairly complete records of all collisions. Each record in the data set reports a single collision, specifically: the date and time of when it occurred; its location; the number of people injured or killed, broken down into motorists, cyclists, and pedestrians; the factors that contributed to the accident; and, the type of vehicles involved. Note that the data set does not include sensitive information that would allow one to trace back the people or cars involved in the collision or the report.

The second data set, the number of vehicles on the road, is also provided under the NYC Open Data initiative and it is maintained by the Metropolitan Transportation Authority (2020) of New York. This data set tracks the vehicles passing through the city bridges and tunnels on an hourly basis. Specifically, the data set reports on the number of vehicles that go through individual toll plazas, every hour, and broken down by: the number of vehicles using an electronic toll collection system; and, the number of vehicles not using the electronic toll collection system or for which such system malfunctioned. These vehicle counts are also associated with the direction of the traffic, which allows us to aggregate the counts to the individual boroughs, which is the geographical aggregation level of reference in our study. Note that this data set does not carry the exact number of vehicles on every single road: we use the counts in this data set to evaluate the relative change in the traffic volume in the boroughs.

The third data set is the traffic speed, provided by the City of New York Department of Transportation (2020a). The data set is a collection of vehicle speed records observed from a multitude of sensors disseminated throughout the boroughs of the city, mostly on major arterials and highways. Each record in the data set contains the following information: the date of the observation, the identifier of the sensor detecting the speed, the speed of the vehicle (measured as the ratio of the segment of road observed over the travel time to cover it), and the location of the sensor. Speed sensors only retain in memory the speed percentiles of vehicles, not the count of vehicles, therefore speed and volume must be measured at different places.

The fourth data set is the traffic summons issued, provided by the City of New York Department of Transportation (2020b). This data set is a record of all traffic summons issued by the city, organized by borough on a daily basis. This is intended to be a proxy for measures of enforcement, since some might be concerned that enforcement has declined during this period.

Finally, the COVID-19 testing data set contains the historical records of the results from the COVID-19 testing campaign. The data set is maintained and provided by New York State Department of Health (2020). The data set begins on March 1st, and it is currently ongoing. Each record in the data set carries data for a specific date and New York State county, and contains the following information: the number of new positive cases discovered during the past 24 hours, with the time cut off set at 12am of the day for which the report is provided; the number of tests performed in past 24 hours, including positive, negative, and inconclusive; and, the cumulative number of new positive cases and of tests performed since the beginning of the testing campaign.

3.1. Data summary

A summary of the data is presented in Table 1. Data are separated into two coarse groups based on the number of new COVID-19 cases in the *week prior*; we will later instrument based on the number of new cases. We can see a clear difference in days preceded by no new COVID-19 cases, Table 1a, with a much larger number of vehicles on the road and a much larger number of collisions, injuries, and fatalities. There does not appear to be a large difference in the number of traffic summons on these days (13.7 vs 13.5), which suggests traffic enforcement is relatively unchanged. On days following no new COVID-19 cases, we reason that drivers believe there is low risk associated with leaving the home, since there are fewer infections and the pandemic spread appears to have slowed in the short run. Conversely, Table 1b, on days following many new infections, drivers tend to avoid leaving their homes, since the spread appears to be relatively rapid. At the same time, collisions drop precipitously on days after new COVID-19 cases, though no direct traffic safety regulations have been passed (only an indirect change in the daily traffic volume of automobiles). Finally, every day has at least one collision, regardless of circumstance, suggesting safety is a major problem.

Fig. 1 shows that the number of vehicles on the road (as measured by tolls) dramatically drop around the time period of the first COVID-19 case, and again during the second wave (though the drop is less dramatic). We also indicate the stay-at-home order issued by the New York State Governor on March 22, 2020, by the dotted vertical line in the figure. We observe that the stay-at-home order was preceded, to a large extent, by individuals choosing to stay home ahead of the legal order as COVID-19 initially began its onset in the region. This reduction in mobility preceding state and local stay-at-home policies matches the observations of others using cell phone records (Badr et al., 2020). We later take the possibility of a systemic break in time trends around the stay-at-home order into careful consideration (Section 5.2). One might also anticipate that the level of collisions afterwards varies because of policies altering the type of individuals on the road. To mitigate this we include fixed effects for each day, which capture changes in the level of collisions before and after the policy. A quadratic spline fits the collisions in each borough before and after the first NYC COVID-19 case. Speed-measuring observations that have been flagged with an "error" by the NYC Department of Transportation have been omitted, but there are still errors clearly visible in the simple mean speed, such as the spike in speed for Brooklyn around 2019-07, so we consider these data somewhat tenuous.

Fig. 1 suggests that the COVID-19 cases caused a sudden and exogenous shock to the number of vehicles passing through tolls in the NYC area. A similar but smaller decline follows the so-called "second wave" during the winter of 2021. We observe that this precipitous drop follows the first NYC COVID-19 case (vertical black line) but precedes the stay-at-home order: individuals seemed to have already stopped driving prior to the announcement to a large extent. We assume the decline in vehicles passing through tolls serves as a proportional measure of the daily traffic volume.

'We also observe an increase in speed in Fig. 1 coinciding with COVID-19 cases, which conforms with the thesis of Shefer and Rietveld (1997). Lower densities of traffic will have higher speeds (and higher variances between the speed of vehicles), and therefore one might anticipate more collisions or more dangerous collisions from the countervailing consequences of speed. We highlight that there has been no major change in the posted speed in NYC in the recent time frame, though there have been major adjustments 6 years ago, in 2014 (Mammen et al., 2019).

In the following sections we use these data to measure the relationship between daily traffic volume, measured by the volume of cars passing through tolls, and the number of collisions. For our experimental variation, we use the number of COVID-19 cases on the week prior in each borough in order to instrument for the current day's traffic.[1] Note that in analyzing and merging the data sets, we realized that there are no data on traffic volume going to Staten Island, which we then removed from our analysis.

4. Methodology

In order to identify the relationship between daily traffic volume and collisions, we use the standard instrumental variables (IV) technique (Cullinane, 2004) with the primary specification:

$$\begin{aligned} ln(Collisions)_{i,t} &= \beta_0 + \beta_1 ln(VehicleCount)_{i,t} + \beta_2 boroughFE_i \\ &+ \beta_3 timeFE_i + \beta_4 boroughFE_i * trend_t \\ &+ \beta_5 boroughFE_i * trend_t * PostCovid_t \\ &+ \beta_6 ln(SummonsCount)_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{1}$$

[1] We consider in the robustness section several other specifications rather than simply the 7-day response and find it does not meaningfully alter the estimations.

Table 1
Comparison between days one week after new COVID-19 cases and after no new cases.

a. Characteristics of days 1 week after a COVID-19 case

Variables	Mean	Std. Dev.	Min	Max
Count Vehicles*	168194.6	82775.35	33,194	348,303
Count Collisions	88.75228	30.24864	19	183
Count Injured	23.95947	11.43227	2	77
Count Fatalities	0.0901826	0.30767	0	3
Total Fatal/Injured	24.04966	11.48352	2	77
Simple Mean Speed**	36.33974	6.221759	18.91236	49.33846
New Cases	0.2111872	3.549588	0	123
Count Traffic Summons	13.53499	23.65592	0	171

N:1752, i = 4 boroughs, t = 438 days, *Missing 16 observations, **Missing 184 observations.

b. Characteristics of days 1 week after no COVID-19 cases

Variables	Mean	Std. Dev.	Min	Max
Count Vehicles	121933.3	69626.33	10,891	303,348
Count Collisions	41.11735	19.03718	3	125
Count Injured	16.0809	9.792077	0	54
Count Fatalities	0.0867347	0.3225972	0	41
Total Fatal/Injured	16.16764	9.855463	0	55
Simple Mean Speed*	40.91199	6.389576	26.74412	55.46777
New Cases	472.8761	490.0707	2	2722
Count Traffic Summons	13.73174	23.6801	0	190

N:260, i = 4 boroughs, t = 65 days, *Missing 12 observations.

With a first stage of:

$$\begin{aligned} ln(VehicleCount)_{\mathrm{i,t}} &= \beta'_0 + \beta'_1 z_{\mathrm{i,t-7}} + \beta'_2 boroughFE_i + \beta'_3 timeFE_i \\ &+\beta'_4 boroughFE_i * trend_t \\ &+\beta'_5 boroughFE_i * trend_t * PostCovid_t \\ &+\beta'_6 ln(SummonsCount)_{\mathrm{i,t}} + \varepsilon_{\mathrm{i,t}} \end{aligned} \tag{2}$$

We have data measured daily (t) for each (and all) borough (i) with complete data. Our interest is in the consistent estimation of the parameter β'_1, which represents the elasticity between vehicle counts and collisions, the percentage change in vehicle collisions for each 1% increase in vehicle counts (Cullinane, 2004). We use the IV technique because we are concerned there could be omitted factors that affect both the count of vehicles and collisions. It could even be the case that collisions lead to changes in the volume of traffic, resulting in inconsistent estimation of the β'_1 parameter.

To resolve this issue, we exploit plausibly random variation in the count of vehicles, where fewer cars are on the road the week following high-infection days. This change in vehicle counts is uncorrelated with other potentially problematic factors - the presence of the disease does not alter driving behavior, except as mediated through the volume of traffic. We operate with the understanding that drivers do drive differently during the pandemic, but this is derived from the fact that the density of the traffic is lower and there is additional space (Tucker & Marsh, 2021), not because drivers have fundamentally changed during the period. Therefore, COVID-19 serves as an important source of experimental variation for studying the traffic collisions in NYC, and allows for the derivation of consistent parameter estimates for β'_1 through this technique.

The vector of controls $borough_i$ contains simple fixed effects for each borough, controlling for the fact that each borough has different average populations and commuting patterns. Similarly, the vector $timeFE_t$ contains fixed effects for each day, since certain days may have unusual weather or be prone to more collisions (new years). These individual and time fixed effects will be partialled out by the fixed effect estimation. In our robustness checks, we also consider that the borough FE undergo some systemic shift during the stay at home period in order to account for possible changes in the constitution of drivers during that period. In our preferred specifications, we include adding borough-specific time trends ($trend_t$), and further break these trends along important dates like the first COVID-19 case (or the stay-at-home order), in case there was a systemic break in commuting patterns during those periods. We also include the log of traffic summons for each day, since one might believe police are altering their enforcement or presence along this data window.[2]

The instrument $z_{i,t-7}$ is a simple count of the number of COVID-19 cases that occurred exactly-one week prior (a lag) as a measure of the severity of the pandemic, which exploits the fact that individuals appear to have stayed at home of their own volition in response to the pandemic (Badr et al., 2020). We note that the lag is appropriate because new cases on day t do not alter driving patterns on day t (they have already left the home), but cases from the week prior (t-7) appear to be strongly associated with changes in automobile traffic.[3] We anticipate that while additional COVID-19 cases alters the number of cars on the road, the virus itself does not directly alter how people drive (e.g., individuals do not drive slower or more cautiously because of the existence of COVID-19). Any remaining change in driving habits during this period is likely, therefore, a result of the altered volume of traffic on the road. Measuring the association between traffic volume and collisions is critical for forecasting the consequences of such policies like congestion taxes in urban areas, particularly if they are severe (Green et al., 2016).

5. Results

In this section we present the results of our work. First, we discuss the changes on the collision patterns, then we elaborate on the robustness checks built in the methodology of the data analysis.

5.1. Primary results: change in collisions

In Table 2 we show the association between the count of collisions and traffic as we sequentially increase the controls:

[2] We use log(SummonsCount + 1) since there are several days with no summons in the data window.

[3] As a robustness check in Table 3 we consider both additional lags up to t-4 and a polynomial lag structure and the results remain similar.

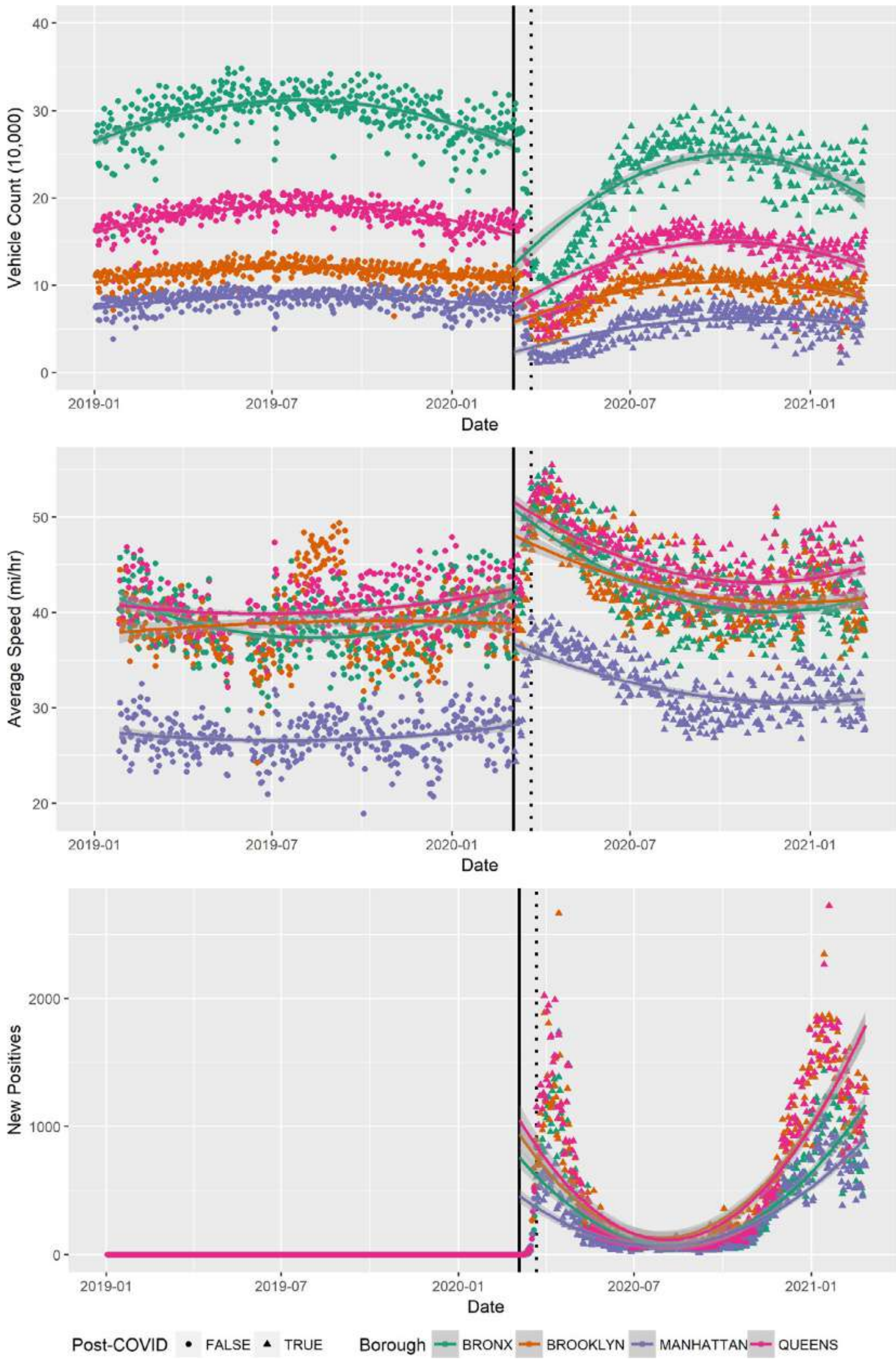

Fig. 1. Traffic volume and speed following the first COVID-19 cases (solid line) and the stay-at-home order (dotted line) in New York City. The trend lines are second-degree polynomial fits of the variable with respect to time. The periods before and after the first COVID-19 case, and for each borough are fit independently.

individual and time fixed effects, individual borough time trends, and systemic breaks in time trends at the day of the first case. The first stage is extremely strong, with an F-statistic of over 75 in all specifications, suggesting that our instrument, the number of cases, is an effective predictor of the number of cars passing through tolls the following day (Stock & Yogo, 2005). It suggests, as one might expect, that people have a strong reaction to protect their health and safety from COVID-19 and, as a result, fewer peo-

Table 2
Estimated elasticity between collisions and vehicle counts.

Variables	(1)	(2)	(3)	(4)
ln(Vehicle Count)	1.480*	1.415***	1.695**	1.661**
	(0.531)	(0.165)	(0.362)	(0.354)
Daily Fixed Effects	YES	YES	YES	YES
Borough Fixed Effects	YES	YES	YES	YES
Borough Trends		YES	YES	YES
Break After First Case			YES	YES
Enforcement Control				YES
Observations	3,108	3,108	3,108	3,108
R-squared (Centered)	0.226	0.337	0.323	0.332
Log Likelihood	1272	1513	1479	1501
Number of Boroughs	4	4	4	4

Standard errors are clustered by borough.
* p < 0.1.
** p < 0.05.
*** p < 0.01.

Table 3
Robustness checks.

Variables	(1) Smaller Data Window	(2) Stay At Home Systemic Break	(3) Polynomial Instrument	(4) COVID-19 Lags
ln(Vehicle Count)	1.379***	1.228***	1.339**	1.619**
	(0.194)	(0.207)	(0.329)	(0.390)
Daily Fixed Effects	YES	YES	YES	YES
Borough Fixed Effects	YES	YES	YES	YES
Borough Trends	YES	YES	YES	YES
Break After First Case	YES	YES	YES	YES
Enforcement Control	YES	YES	YES	YES
Observations	2,500	3,108	3,108	2,940
R-squared (Centered)	0.362	0.310	0.390	0.347
Number of Boroughs	4	4	4	4

Standard errors are clustered by borough.
* p < 0.1.
** p < 0.05.
*** p < 0.01.

ple drive on days after many people have tested positive. Since our specification is log–log, the coefficients represent estimations of the elasticity of the number of collisions relative to the volume of traffic (i.e., they are the percentage change in collisions over the percentage change in traffic volume).

In Table 2, column 1 reports on the values of the elasticity between collisions and vehicle traffic when only fixed effects for borough and time periods are included, for which we observe a value of 1.5, roughly. These time fixed effects take into account day of week and city-wide weather effects, which accounts for the highly cyclical nature of traffic: day of week traffic patterns, and seasonal commuting patterns. The estimated elasticity remains similar at 1.4 when adding individual time trends, (see column 2). These individual time trends take into account to the fact that each borough may be trending upward or downward in traffic safety independently, due to factors such as gentrification. In column 3, we add a systemic break for the first COVID-19 case, since the borough-specific traffic trends may have altered at that time, perhaps as individuals work from home at different rates in different regions. The estimated elasticity between vehicle count and collisions increases to 1.7, though the standard error increases. This is robust to the addition of ln(*SummonsCount*) in column 4. Critically, this inclusion does not substantively alter the coefficient of interest. Our final estimation measures that an increase in the number of vehicles by 1% is associated with a significant 1.7% increase in the number of collisions, and the final R^2 is around 30%. Using a short calculation, the nearly 28% reduction in vehicle traffic in the Covid-19 period we see in Fig. 1, is associated with nearly a 37% reduction in collisions.[4]

5.2. Robustness checks

In this section we run a number of tests to highlight that the estimation in Table 2, column 4, is robust to a variety of specification changes. Results of the robustness checks are reported in Table 3. The changes that we take into account are: we modify the data window's size; we add a second systemic break in trends on the day of the stay-at-home-order; and, we explored several variations of lags for the COVID-19 case instrument. The general pattern of the result remains consistently positive and significant, and the estimated elasticity between traffic volume and collisions falls typically between 1.2 and 1.7, and none are more than 3 standard deviations away from our preferred specification of 1.661.

In Table 3 column 1, we reduce the size of the data window by about 6 months, such that it begins on June 1st 2019 instead of January 1st 2019. We note that the estimated elasticity only modestly increases in size and it is still significant at about 1.4. In column 2 we consider that perhaps the breaks in time trends were insufficient, and includes another systemic break in trends at the time of the governments stay-at-home order. We further allow each borough to have an additional break in the value of their associated fixed effect within the stay-at-home period, in case the rule trig-

[4] $1-(1-0.01660956)^{28} \approx 0.37$.

gered particular changes in the constitution of travelers from each borough. While the stay-at-home order appears to have been preceded to large extent by New Yorkers already staying home (see Fig. 1), it is possible that there were still transitions to the driving (beyond simply daily traffic volume) during the post-announcement period. Jointly, these breaks are significant, reinforcing that this is another plausible specification. Still, including this second systemic break only decreases the estimated coefficient slightly to 1.2, while the significance remains. In column 3, we consider alternatives to individuals waiting 7 days to respond to new COVID-19 cases. Instead, we consider a fourth-degree polynomial of the 7-day instrument. Our point with this specification, and the next, is that individuals have a generalized pattern of reducing driving in response to past cases of COVID-19 and that our results do not appear to be dependent on a particular functional form in the first stage. The instruments remain strongly significant, with an F-statistic of 52. The estimated elasticity increases to a significant 1.3, still in the same general direction and within one SD of the primary estimates in Table 2. In column 4 we test another set of instruments, exploring the idea that individuals respond to 7, 14, and 21 day lags of COVID-19 cases, that is they have a memory of just under a month before they return to normal habits. The instruments remain strongly significant, with an F-statistic of 25. The estimated elasticity remains at a significant 1.62, within 0.05 of the original estimates in Table 2, column 4.

5.3. Discussion and further elaboration

So far, we have focused heavily on the percentage change in the number of collisions as a consequence of daily traffic volume. However, not all collisions are equally dangerous. In this section, we consider how the daily traffic volume is associated with the total count of fatalities and injuries. Previous research highlights that the direction of the effect appears to be complicated by potentially reducing the distance between cars, therefore softening the collisions (Shefer & Rietveld, 1997; Wang et al., 2009). We examine these concerns, particularly the notion that denser traffic leads to "cushioning" and safer collisions. We find significant support for the idea that higher traffic volumes are associated with safer collisions.

In Table 4, we continue to use the same IV specification as in column 3 of Table 2, though in each of the 4 columns we change the independent variable in each of the estimations to the count of: injuries, fatalities, the sum of both (casualties), and casualties per collision. In each of the 3 rows, we consider a different subset of victims: total, pedestrians, or cyclists, which are contained as subsets of the crash data. This leaves a table of 12 coefficients. Note that in Table 4 we do *not* use logarithms on the dependent variables, since numerous days have zero fatalities, and the number of injuries is occasionally zero for some subsets. Omitting these periods would bias the estimates (Wooldridge, 2010).

Each coefficient in Table 2 then represents a semi-elasticity: the change in injury (fatality, casualty, or casualties per collision) with respect to the percentage change in traffic volume. In the first row, column 1, we look first at overall fatalities, and the coefficient suggests that a 1% increase in traffic volume is associated with a large but insignificant decrease of 0.003 deaths per day per borough. This suggests that there is a substantive cushioning effect, since additional automobiles are associated with more collisions but fewer fatalities. A quick calculation suggests that the elasticity of fatalities with respect to traffic volume is approximately -3.4.[5] This is roughly the same as the decline in fatalities for low density neighborhoods (-3.75% fatalities when density is $<1500/\text{km}^2$) in Noland and Quddus (2004), which used congestion implied from population and employment. It is also entirely the opposite direction suggested by the congestion tax studied in Green et al. (2016), highlighting the importance of the magnitude of the change in traffic. This reiterates that the relationship between traffic volume and collisions is nonlinear – when the changes in traffic volume are particularly large, the change in collisions may be noticeably different from smaller changes.

Our point estimates also show a significant decrease in overall injuries in row 1, column 2, Table 4. Each 1% additional vehicle traffic is associated with a significant decrease of about 0.44 injuries per day, an approximate 2.2% decrease.[6] In Table 4, row 1, column 4, we look at the casualties, the sum of fatalities and injuries: we find essentially the same coefficient and significance as in the injuries (column 2). This seems reasonable since injuries are the vast majority of casualties. Finally, in row 1, column 4, we look at the rate of casualties/collision. We find a negative association between the rate of casualties/collision and traffic volume, though the coefficient is not significant. This suggests that denser traffic may have fewer injury occurrences. The pattern in the first row persists for pedestrian casualties (row 2), but the coefficients are insignificant with large standard errors. For cyclists the standard errors are even larger, however, the rate of casualties/collision is found to be negative for all categories. This suggests that collisions may be generally safer in high-volume roadways. We conclude, therefore, that higher volumes of vehicle traffic are more likely to have collisions, but those collisions tend to have fewer injuries and potentially fewer deaths than their low-volume collision counterparts.

We find that collisions have declined during this COVID-19 period, but the safety of those collisions also declined, matching the mixed effects of Shefer and Rietveld (1997). As a result, we would like to investigate if the other elements of the "cushioning" hypotheses hold true in the relevant range, in particular that vehicles will slow down under higher densities of traffic, which leads to safer accidents.

We note it is visually apparent that the simple mean speed of traffic has drastically increased in all boroughs post-COVID-19, see Fig. 1, which according to Shefer and Rietveld (1997) is a consequence of the reduced density. To corroborate on this observation, we run a simple regression of log vehicle volume against log speed (controlling only for time and borough fixed effects). We find a significant estimated elasticity between volume and speed of about -13%. This component of the evidence seems to reinforce the plausibility of traffic volume as a double-edged sword: while the higher traffic volumes lead to more collisions, sudden decreases in traffic volumes are associated with higher speeds. This suggests a rich, complex relationship, despite this factor being outweighed in the relevant range by the safety benefits we measure in Table 4.

In total, we calculate that there were nearly 356 collisions each day prior to COVID-19, combined across all four boroughs. Prior to COVID-19, the value of these collisions, injuries, and deaths each day are approximately \$4.2 million per day.[7] In the post-COVID-19 period, we calculate that the 28% reduced traffic volume leads to about a 37% reduction in collisions, resulting in approximately \$453,000 in social property costs avoided daily, excluding any injuries or fatalities prevented.[8]

[5] -0.003 change in fatalities for each 1% change in traffic volume / 0.089 average fatalities over data window $* 100\% \approx -3.4\%$ in fatalities for each 1% change in traffic volume.

[6] –0.4367 change in injuries for each 1% change in traffic volume / 20.45 average injuries over data window $* 100\% \approx -2.1$ injuries per day for each 1% change in traffic volume.

[7] 88.75 collisions/borough*4*\$3, 447 + 23.95 injuries/borough*4 *\$28, 299 + 0.0901 fatalities/borough*4*\$3,186,408 [approximately equal symbol goes here] \$4.2 million/day.

[8] We use the value of \$3,447 for property damage-only collisions from (Parry, 2004), so 37% reduction in collisions 88.75 collisions per day 4 boroughs \$3, 447 \$452, 763.

Table 4
Changes in casualties by type and severity.

Total	Fatalities	Injuries	Casualties	Casualties per Collision
ln(Vehicle Count)	−0.343	−43.67*	−44.01*	−0.114
	(0.183)	(14.81)	(14.80)	(0.183)
PEDESTRIAN	Fatalities	Injuries	Casualties	Casualties per Collision
ln(Vehicle Count)	−0.276	−4.994	−5.270	−0.0605
	(0.173)	(2.462)	(2.537)	(0.126)
CYCLIST	Fatalities	Injuries	Casualties	Casualties per Collision
ln(Vehicle Count)	0.146	−6.665	0.105	−6.519
	(0.124)	(7.691)	(0.0617)	(7.645)

Standard errors are clustered by borough.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

However, the increased injury rate of those remaining collisions more than offsets the benefits of the reduced collisions. We further calculate that the increase in fatalities and injuries are worth approximately $2.6 million in social cost a day, though this will vary substantially depending on the statistical value of a human life (Parry, 2004), and any variation in the fatalities estimate.[9] This suggests that policymakers may want to be careful in considering policies like congestion taxes that may adjust the traffic volume, since there may be unintended consequences for safety in terms of speed. These consequences may not appear when adjustments to traffic volume are modest, but are visible when there are sudden and substantive traffic volume changes, such as in the case of COVID-19.

6. Conclusion

In this paper, we utilize a unique instrumental variable: the number of COVID-19 cases, to instrument for the traffic volume on the following day. The use of the instrumental variable helps address the concerns of Wang et al. (2009), Shefer and Rietveld (1997), and Noland and Quddus (2004), about the challenges of identifying a relationship between congestion and traffic safety (Cullinane, 2004). We find that the exogenous shocks to traffic volume substantively reduce the number of collisions at a rate of roughly 1.7% fewer collisions for every 1% reduction in traffic volume. This finding is robust to various specifications, including different time windows, borough-specific time trends, alternative specifications of the instrument itself, and a structural break in time trends at the time of the stay-at-home order. On the other hand, our estimates suggest lower volumes of traffic are associated with net increase in injuries and fatalities over the relevant range, specifically an increase of 0.003 fatalities for each 1% decrease in traffic volume, and an increase of 0.44 for injuries for each 1% decrease in traffic volume.

Both of these findings suggest that the remaining collisions are of a more dangerous type when traffic volume is reduced, a concern highlighted by Shefer and Rietveld (1997) and Zhou and Sisiopiku (1997). We find other corroborating evidence in our data: we find an increase in casualties per collision as traffic volume decreases. Though this association is not significant, the pattern remains persistently negative for pedestrians. Accordingly, we estimate a simple elasticity of speed/volume approximately equal to −0.13.

We also provide an estimates of the value of this transition. Using a back-of-the-envelope calculation, we estimate that the approximate 28% decline in traffic volume during the post-COVID-19 period is associated with about $453,000 per day in savings from the property damage of collisions (Parry, 2004). However, the increased bodily harm from collisions is estimated to be worth nearly $2.6 million a day across the boroughs, though this estimate may vary dramatically depending on the difficult valuation of human life and the effects on external participants (Parry, 2004). Despite variations in valuation, the estimated value of the injuries and fatalities are an order of magnitude larger than the measured reduction in property damage. This highlights a complex relationship between traffic volumes, speeds, and safety that is worth further study.

[9] 28% reduction in traffic volume*4 boroughs * (0.01 * 0.343 fatalities *$3,186,408 + 0.01*43.67 injuries*$28, 299) [approximately equal symbol goes here]$2, 6082, 005. We also note that these estimates exclude any change that may occur in Staten Island.

References

Abouk, R., & Adams, S. (2013). Texting bans and fatal accidents on roadways: Do they work? Or do drivers just react to announcements of bans? *American Economic Journal: Applied Economics, 5*, 179–199.
Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review, 104*, 2763–2796.
Badr, H., Du, H., Marshall, M., Dong, E., Squire, M., & Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study Online. *The Lancet Infectious Diseases*. https://doi.org/10.1016/S1473-3099(20)30553-3.
Baghestani, A., Tayarani, M., Allahviranloo, M., & Gao, H. O. (2020). Evaluating the traffic and emissions impacts of congestion pricing in New York City. *Sustainability, 12*, 3655.
Bauernschuster, S., Hener, T., & Rainer, H. (2017). When labor disputes bring cities to a standstill: The impact of public transit strikes on traffic, accidents, air pollution, and health. *American Economic Journal: Economic Policy, 9*, 1–37.
City of New York Department of Transportation (2020a). Traffic speeds. https://data.cityofnewyork.us/Transportation/ DOT-Traffic-Speeds-NBE/i4gi-tjb9 (accessed May, 2021).
City of New York Department of Transportation (2020b). Traffic summons. https://data.cityofnewyork.us/Public-Safety/NYPD-B-Summons-Historic-/bme5-7ty4 (accessed August, 2021).
Cui, Z., Henrickson, K., Ke, R., & Wang, Y. (2020). Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. *IEEE Transactions on Intelligent Transportation Systems, 21*(11), 4883–4894. https://doi.org/10.1109/TITS.2019.2950416.
Cullinane, K. (2004). *Statistical and econometric methods for transportation data analysis*. Springer.
De Fabritiis, C., Ragona, R., & Valenti, G. (2008). Traffic estimation and prediction based on real time floating car data. In *2008 11th international IEEE conference on intelligent transportation systems* (pp. 197–203). IEEE.
Dills, A. K., & Mulholland, S. E. (2018). Ride-sharing, fatal crashes, and crime. *Southern Economic Journal, 84*, 965–991.
Edlin, A. S., & Karaca-Mandic, P. (2006). The accident externality from driving. *Journal of Political Economy, 114*, 931–955.
Faghih-Imani, A., Anowar, S., Miller, E. J., & Eluru, N. (2017). Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City. *Transportation Research Part A: Policy and Practice, 101*, 11–21.
Green, C. P., Heywood, J. S., & Navarro, M. (2016). Traffic accidents and the London congestion charge. *Journal of public Economics, 133*, 11–22.
Iyer, S., Boxer, K., & Subramanian, L. (2018). Urban traffic congestion mapping using bus mobility data. In *CEUR workshop proceedings* (Vol. 2227, pp. 7–13). CEUR-WS.
Jackson, C. K., & Owens, E. G. (2011). One for the road: Public transportation, alcohol consumption, and intoxicated driving. *Journal of Public Economics, 95*, 106–121.
Jain, V., Sharma, A., & Subramanian, L. (2012). Road traffic congestion in the developing world. *Proceedings of the 2nd ACM symposium on computing for*

development ACM DEV '12. New York, NY, USA: Association for Computing Machinery.
Lichtman-Sadot, S. (2019). Can public transportation reduce accidents? Evidence from the introduction of late-night buses in israeli cities. *Regional Science and Urban Economics, 74*, 99–117.
Liu, W., Zheng, Y., Chawla, S., Yuan, J., & Xing, X. (2011). Discovering spatio-temporal causal interactions in traffic data streams. In *Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining KDD 2011* (pp. 1010–1018). New York, NY, USA: Association for Computing Machinery.
Mammen, K., Shim, H. S., & Weber, B. S. (2019). Vision Zero: Speed limit reduction and traffic injury prevention in New York City. *Eastern Economic Journal*, 1–19.
Metropolitan Transportation Authority (2020). Hourly traffic on Metropolitan Transportation Authority (MTA) bridges and tunnels. https://data.ny.gov/Transportation/ Hourly-Traffic-on-Metropolitan-Transportation-Auth/qzve-kjga (accessed May, 2021).
New York State Department of Health (2020). New York state statewide COVID-19 testing. https://health.data.ny.gov/Health/ New-York-State-Statewide-COVID-19-Testing/xdss-u53e (accessed May, 2021).
Noland, R. B., & Quddus, M. A. (2004). A spatially disaggregate analysis of road casualties in england. *Accident Analysis & Prevention, 36*, 973–984.
NYPD (2020). Motor vehicle collision. https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95 (accessed May, 2021).
OpenData, N. (2020). New York City OpenData initiative. https://opendata.cityofnewyork.us/ (accessed May, 2021).
Parry, I. W. (2004). Comparing alternative policies to reduce traffic accidents. *Journal of Urban Economics, 56*, 346–368.
Parry, I. W., & Bento, A. (2002). Estimating the welfare effect of congestion taxes: The critical importance of other distortions within the transport system. *Journal of Urban Economics, 51*, 339–365.
Pigou, A. C. (1920). *The economics of welfare*. Palgrave Macmillan. Reprinted in 2013.
Shefer, D., & Rietveld, P. (1997). Congestion and safety on highways: Towards an analytical model. *Urban Studies, 34*, 679–692.
Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and Inference for Econometric models: Essays in Honor of Thomas Rothenberg*, 109–120.
Tucker, A., & Marsh, K. (2021). Speeding through the pandemic: Perceptual and psychological factors associated with speeding during the covid-19 stay-at-home period. *Accident Analysis & Prevention, 159*. https://doi.org/10.1016/j.aap.2021.106225 https://www.sciencedirect.com/science/article/pii/S0001457521002566.
Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review, 59*, 251–260.
Wang, C., Quddus, M. A., & Ison, S. G. (2009). Impact of traffic congestion on road accidents: A spatial analysis of the m25 motorway in England. *Accident Analysis & Prevention, 41*, 798–808.
Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.
Zhou, M., & Sisiopiku, V. P. (1997). Relationship between volume-to-capacity ratios and accident rates. *Transportation Research Record, 1581*, 47–52.

Paolo Cappellari is an Associate Professor at the College of Staten Island, CUNY. His research focuses the effective management, interpretation, and utilization of data, with specific attention to the study and transformation of schema and data between models, keyword search over semantic datasets and data management in sensor networks. He has published his research results in the major journals and conferences, including ACM-SIGMOD, VLDB, EDBT, ER, ICDE, FoIKS.

Bryan S. Weber is an Assistant Professor of Economics at the College of Staten Island, CUNY. His research focuses on transportation, crime, and statistical methods. Currently, his work explores the introduction of new and novel transportation systems into urban areas, helping enumerate the costs and benefits to residents. His published research results include Transportation Research Part A: Policy and Practice, Regional Science and Urban Economics, and Economic Inquiry.

Journal of Safety Research 83 (2022) 195–203

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

A new alcohol-related traffic law, a further reduction in traffic fatalities? Analyzing the case of Turkey

José Ignacio Nazif-Munoz [a,*], Gül Anıl Anakök [b,c], Junon Joseph [a], Santosh Kumar Uprajhiya [a], Marie Claude Ouimet [a]

[a] Université de Sherbrooke, Quebec, Canada
[b] Kocaeli University, Kocaeli, Turkey
[c] Kartepe District Health Directorate, Kocaeli, Turkey

ARTICLE INFO

Article history:
Received 26 July 2021
Received in revised form 4 May 2022
Accepted 23 August 2022
Available online 1 September 2022

Keywords:
Policy evaluation
Turkey
Fixed effects
Alcohol related laws
Time-series

ABSTRACT

Background: In June 2013, an alcohol-related traffic law took effect in Turkey. The law 6487 introduced administrative fines for not respecting blood alcohol concentration limits, health warning messages on alcohol containers (bottles, cans), and prohibited the sale of alcohol beverages in retail facilities between 10 p.m. and 6 a.m.. This article examines how this law is associated with traffic fatality variation. *Methods:* Data from the Turkish Statistical Institute for the 2008–2019 period were analyzed. Outcomes were traffic fatality rates per 100,000 population and 10,000 motor vehicles. Exposure variable was the presence of law 6487. Alcohol, tobacco, and related beverages' household expenditure, unemployment rate, number of health professionals, number of crashes, and lags of the outcomes represented control variables. A time-series cross-regional fixed effect model was applied. *Results:* Empirical estimates suggest that the law 6487 was associated with a reduction of 15% (Incidence Rate Ratio (IRR) 0.85, 95% Confidence Interval (CI): 082, 0.94) in the traffic fatality per population rate and with a reduction of 14% (IRR: 0.86 (95% CI: 0.78, 0.92) in the traffic fatality per motor-vehicle rate. After 6 years of its implementation, this intervention was associated with an absolute reduction of 1519 (95% reduction interval: 1177, 1810) traffic fatalities. *Conclusions:* Our research emphasizes that legislation with direct and indirect measures targeting driving under the influence of alcohol (DUIA) may be related to traffic fatalities reduction. *Practical applications:* This finding has important implications for policy and future research in contexts in which alcohol consumption is low such is in Turkey. Future research should seek to identify mechanisms that explain how laws are ultimately associated with DUIA variation.

1. Introduction

Driving under the influence of alcohol (DUIA) is a major public health challenge. It has been estimated that globally more than 270,000 people die yearly in alcohol-related crashes (Vissers and Houwing, 2017). Countries have introduced direct and indirect policies to reduce this burden. Direct policies include, for instance, legislation that targets the reduction of DUIA through setting *per se* blood alcohol concentration (BAC) limits for the general population (Castillo-Manzano et al., 2017), and/or specific subgroups, such as younger drivers (Lovenheim and Slemrod, 2010) or professional drivers (Smailović et al., 2020). They also comprise deterrence measures like increases in the penalties associated with DUIA (Miller et al., 2018). Indirect policies may reduce DUIA by decreasing the opportunities to drink alcohol (Babor et al., 2010). Examples of these measures are: limiting the hours and places of when and where alcohol could be sold or consumed (Wagenaar et al., 2015; Sanchez-Ramirez and Voaklander, 2018), increasing alcohol-related taxes (Lavoie et al., 2017), and regulating alcohol advertising through the provision of health warnings (Rehm et al., 2020).

After the publication of the first world road safety report in 2004 (World Health Organization, 2004), countries have converged to adopt multiple road safety policies (Nazif-Muñoz, 2015), including lowering legal BAC limits (Fell and Scherer, 2017). Furthermore, in the last decade both the World Health Organization (WHO) (Fell and Voas, 2014) and the European Transport Safety Council (1) have explicitly suggested that to reduce DUIA and its consequences countries should lower the legal BAC limit to 0.05 g/dL. A lowering

* Corresponding author.
E-mail address: jose.ignacio.nazif-munoz@usherbrooke.ca (J.I. Nazif-Munoz).

https://doi.org/10.1016/j.jsr.2022.08.015
0022-4375/© 2022 The Author(s). Published by the National Safety Council and Elsevier Ltd.

of the *per se* BAC limit from 0.10 g/dL to 0.05 g/dL has been associated with a decrease of 48% in road traffic fatalities in Australia (Homel et al., 1995), a decrease of 25% in alcohol-related fatal crashes in Japan (Deshapriya and Iwase, 1970), and a decrease in alcohol-related crashes of 9% in Austria (Bartl and Esberger, 2000) and 36% in France (Mercier-Guyon, 1998). A more comprehensive study, across 27 European countries that restricted BACs from 0.10 g/dL or 0.08 g/dL to 0.05 g/dL, reported an 11% decrease in all road traffic fatalities (Castillo-Manzano et al., 2014). A recent study in Scotland suggested, however, that road traffic crashes did not vary after a reduction from 0.08 g/dL to 0.05 g/dL was enacted (Haghpanahan et al., 2019). In sum, multiple studies have suggested that laws with 0.05 g/dL restrictions may be an effective strategy in reducing both traffic fatalities and traffic alcohol-related fatalities.

While these studies have provided us with important conclusions regarding DUIA, policies, and the decreasing of road traffic fatalities, two important challenges should be acknowledged. First, most of this research, with the exception of Castillo-Manzano et al. (2014) has not considered explicitly how both direct measures (e.g., introducing administrative fines for not respecting the BAC legal limit of 0.05 g/dL, refusing to provide a breath sample to measure BAC (Voas et al., 2009) and indirect measures (e.g., regulating the hours and premises of alcohol sales, promotion of health warnings associated with alcohol consumption) may have jointly contributed to reductions in DUIA, when BAC legal limits have been set in place. Indeed, research has indicated that a USD$100 increase in the minimum fine for not respecting a BAC legal limit of 0.08 g/dL was associated with decreases between 2% and 8% in all alcohol-related traffic fatalities in the United States (Wright and Lee, 2021). Introduction of administrative fines for not respecting a BAC legal limit of 0.05 g/dL was associated with a decrease of 3% in drivers' alcohol-related fatalities in Canada (Blais et al., 2015), while an increase of administrative fines was associated with a decrease of 9% in drivers' alcohol-related fatalities in Slovenia (Kralj et al., 2009). Research conducted in Texas compared counties that implemented "no refusal" search warrant programs with counties that did not implement programs; the implementation of the program was associated with a 2% decrease in alcohol-related traffic crashes (Ames et al., 2016). On the other hand, research on indirect measures have suggested that increasing alcohol sales hours was associated with an increase in traffic collisions in both Australia and Canada, and paradoxically with a decrease in the United Kingdom (Sanchez-Ramirez and Voaklander, 2018; Popova et al., 2009). Reviews regarding health warning labels suggest that this policy could be associated with mild reductions in DUIA in Canada and the United States (Greenfield et al., 1999; Martin-Moreno et al., 2013).

Second, an important common element of all of the precited studies is that they have been carried out in high-income countries, where alcohol consumption is culturally accepted and where traffic fatalities rates are relatively low (World Health Organization, 2018). As such, results from these studies may not be generalizable to settings where alcohol consumption is low and traffic fatality rates are relatively high. For instance, in 2015 alcohol consumption per capita was 1.4 l in Turkey and 10.8 l in Australia (World Bank, 2021). In the same year, there were 9.9 road fatalities per 100,000 population in Turkey (World Health Organization, 2018), but 5.4 in Australia (World Health Organization, 2018). Studies conducted in lower- and middle-income countries are warranted.

2. Study context

On May 24, 2013, Turkey enacted law 6487, which took effect on June 11, 2013. In Table 1, we classify the direct and indirect policies associated with DUIA in relation to norms before and after the introduction of law 6487. This law reinforced the BAC legal limit of 0.05 g/dL by including new administrative fines for not respecting the existing BAC legal limit, driving under the influence of drugs, or refusing to provide a breath sample to measure BAC. Administrative sanctions in Turkey can be imposed only by administrative bodies and not by courts or judicial institutions, and administrative bodies are not entitled to impose sanctions that result in imprisonment (Administrative and Turkey, 2005). The law also introduced mandatory health warning labels on alcohol containers (cans, bottles) and prohibited the sale alcohol beverages in retail facilities between 10 p.m. and 6 a.m.

In Turkey, the implementation of the law 6487 could be associated with a decrease in traffic fatalities by directly changing transport preferences (using public transport instead of private vehicles once individuals have consumed alcohol) and/or indirectly by reducing alcohol consumption. However, this law may be confounded with other factors or simply not stringent enough to influence variation in traffic outcomes. Thus, attention needs to be given to at least three alternative explanations. First, health systems may be more effective in responding to the occurrence of traffic crashes and therefore one could observe reduction in traffic fatalities (Castillo-Manzano et al., 2014). In cross-national studies a proxy to measure health system performance in timely responses to road crashes injuries has been number of health professionals (Ali et al., 2019). Second, unemployment could also be associated with traffic fatality reductions (Wegman et al., 2017). Studies have

Table 1
Direct and indirect policies before and after the Law 6487.

Policy	Sub-policy	Before law 6487	With law 6487
Direct	Administrative sanctions involving fines for driving over the BAC legal limit of 0.05 g/dL	No administrative sanctions (Law 2918)	Administrative sanction for drivers not respecting the BAC limit: – 1st time: 700 Turkish liras (USD$ 83) – 2nd time: 877 Turkish liras (USD$ 102) – 3rd time: 1407 Turkish liras (USD$ 159)
	Administrative sanctions involving fines for driving under the influence of drugs	No administrative sanctions (Law 2918)	Administrative sanction of: - 3600 Turkish liras (USD$ 423) for individuals driving under the influence of drugs
	Administrative sanctions related to refusal of provision of biological samples to assess BAC and drug consumption	No administrative sanctions (Law 2918)	Administrative sanction of: - 2000 Turkish liras (USD$ 230) for refusing to provide biological samples to measure BAC levels or drug consumption
Indirect	Health warnings associated with alcohol consumption	Alcohol producers were not mandated to introduce health warning labels on alcohol containers (Law 4250)	Alcohol producers must include warning messages associated with the risks of alcohol consumption on alcohol containers (bottles, cans)
	Restricted hours for selling alcohol	Alcohol beverages in retail facilities could be sold without restrictions of hours (Law 4250)	Alcohol beverages in retail facilities cannot be sold between 10:00 pm and 6:00 am

suggested that when individuals face economic restrictions such as being unemployed, they are less likely to either go to bars or restaurants, which reduces the opportunity to drink alcohol and then drive, and therefore one can expect declines in traffic fatalities. Third, the number of vehicles has also been associated with traffic fatalities' variation. Studies suggest that when jurisdictions increase their motor-vehicle fleet, this can be associated with increases in traffic congestion, which in turn is linked with decreases in speed, and then with reductions in traffic fatalities (Cabrera-Arnau et al., 2020). Each of these factors could be influencing traffic fatality variation in Turkey. Furthermore, studies in Turkey have indicated that important differences across regions are found in terms of road safety behavior and alcohol consumption. For instance, while in 2010 in Ankara and Afyon the percentage of seat belt use by drivers was 21% and 4%, respectively, in 2014 these numbers were 37% and 28% (Gupta et al., 2017). Studies monitoring geographical differences in alcohol consumption via wastewater-based epidemiology suggest that alcohol consumption in Aegean (32 l/1000 population/day), is much higher than Akdeniz (25 l/1000 population/day) or Bati Karadeniz (18 l/1000 population/day) (Kuloglu Genc et al., 2021). (Similar studies in western countries have reported for Copenhagen more than 40 l/1000 population/day and in Barcelona more than 12 l/1000 population/day (Cabrera-Arnau et al., 2020). More information regarding Turkey's regional differences can be found in the supplementary material in Table S1.

In this study, we contribute to the body of research interested in assessing how a policy, which includes both direct and indirect measures to reduce DUIA, may be associated with traffic fatalities' reduction. For this we specifically conducted a time-series cross-regional fixed-effect analysis in Turkey between 2008 and 2019 to empirically assess, at the national level, the association of a law that targets DUIA through direct and indirect interventions with traffic crash fatality rates (Republic of Turkey, 2013).

3. Methods

3.1. Study design

Given the important differences across and within regions in Turkey over time, we applied a time-series cross-regional study design to evaluate the impact the of law 6487 from 2008 to 2019 on two outcomes: total number of traffic fatalities per 100,000 population and total number of traffic fatalities per 10,000 motor vehicles (Wegman and Oppe, 2010). By using two denominators (population and motor vehicles), we facilitate international comparisons of the results and test the robustness of results across models (Elvik et al., 2009).

3.2. Data

3.2.1. Outcome variables

Data on traffic fatalities were obtained from Turkish Statistical Institute (TSI) (Turkish Statistical Institute, 2019). This organization has compiled, from police reports and hospitals, an extensive dataset of traffic fatalities in Turkey from 1998 to the present. Data are available per year and per region. We restrict our analysis to 2008 and 2019 since Turkey changed its entire mortality classification system, improving its collection practices substantially from 2008 (Özdemir et al., 2015). Complete annual data for all 12 regions are available for this period. This provides us with a balanced sample with $N = 144$ region-year observations. Data on road traffic fatalities are compiled from Record of Traffic Crashes forms filled out for every crash and prepared for both judicial and statistical purposes in accordance with the Highway Traffic Law No. 2918. The TSI reports two type of road fatalities: (1) number of persons killed in a road crash as deaths occurred at the crash scene (2008–2019), and (2) number of persons who were victims of a road crash and died within 30 days after this event (2015–2019). The main analysis is conducted with deaths occurring at crash scenes as data are available for the full study period. The number of fatalities having occurred within 30 days of a crash is also used to test the robustness of our results, but it is only available since 2015. For the 2008–2014 period, we multiplied the number of deaths occurring at crash scene by a correction factor of 2.06 to obtain estimates of deaths occurring within 30 days. We obtained this factor from averaging differences between both reported fatalities for the 2015–2019 period (see supplementary material Formula S1 and results in Table S2). The selected correction factor in the current study is slightly larger than the factor proposed by Cetin et al. (2018) for Turkey (i.e., 1.60) as their work only considered urban zones. Data on vehicle fleet and population were both obtained from the TSI. Compiled annual fatality data do not include information on drivers' alcohol consumption nor the time when crashes occurred.

3.2.2. Exposure

We defined the pre-intervention period from 2008 to 2013, and the post intervention period from 2014 to 2019. Since monthly data were not available, we were not able to define pre- and post-intervention periods using the month in which the law was implemented, June 2013. (In a complementary analysis, we also defined the pre-intervention and post-intervention periods without the year 2013, pre-intervention period 2008–2012 and post-intervention period 2014–2019. This analytical strategy facilitates comparison with models in which the year 2013 was assumed to be absent, even though the law had taken effect in June 2013.) Since the effect of the intervention may vary with time since exposure, we also considered adding a linear term following the intervention. This allows us to detect whether the intervention's effect may have varied over time, rather than assuming the law had an immediate and permanent change.

3.2.3. Control variables

In consideration of previous literature (Castillo-Manzano et al., 2017; Wegman et al., 2017; Cabrera-Arnau et al., 2020), several control variables were introduced: (1) *Alcohol, tobacco, and related beverages households' expenditure*. This is the average of households' income share of alcohol, tobacco, and related beverages per region. Data come from the Household Budget Survey, which is administrated by the TSI. (2) *Unemployment rate*. Following the TSI definition, this variable includes the non-institutional working age population who are not employed, have utilized at least one of the channels for seeking job, and are available to start a job within two weeks. This considers the total number of individuals present under this characterization divided by the total of working age population in each region. (3) *Number of health professionals per capita*. It measures the total number of physicians, health officers, and nurses in each region. A physician is a person who has completed six years of higher training in medicine; a health officer is a person with a four-year bachelor's degree trained in health; and a nurse is a person with a four-year bachelor's degree from high school trained in health and with a two-year nursing education; (4) *Number of crashes* is defined as an event in which a collision between one or more moving vehicles in a road or a highway results in the death or injury of one or more individuals, and/or material loss. This variable is considered since traffic fatalities are highly dependent on the number of crashes. All data are taken from TSI.

3.2.4. Lags, year effects, and region time trends

Since traffic mortality rates can be autocorrelated, we considered a one-year lag for both outcomes to correct for autocorrelation (Freivalds and Johnson, 1990). We also explored whether controlling for year effects, assuming unobserved systematic changes happen in each region for a given year (Wooldridge, 2002), could also improve our estimations. We also estimated region-specific linear time trends, since time trends of traffic fatalities could differ between regions.

3.3. Model selection and statistical approach

Due to the non-negative integer nature of road traffic fatality count data, we used generalized linear models, particularly Negative binomial regression. These models assume that the conditional distribution of the count of traffic fatalities in region r in year y is negative binomial (NB):

$$y_{ry} \sim \mathrm{NB}(\mu_{ry}) \tag{1}$$

Secondly, since autocorrelated data may bias estimates, we applied the Born and Breitung test to detect the presence of autocorrelation in these regional time-series data, *p* values lower than 0.05 suggest the time-series is autocorrelated (Wursten, 2018). Then to systematically detect if autocorrelation was corrected, we built models in which lags (Eq. (2)), year fixed effects (Eq. (3)) and linear trends per region were introduced (Eq. (4)) sequentially.

$$y_{ry} = \exp[\zeta \mathrm{Lag} * y_{ry} + \varepsilon_{ry}], \tag{2}$$

$$y_{ry} = \exp[\lambda \mathrm{Year}_r + \varepsilon_{ry}], \tag{3}$$

$$y_{ry} = \exp[\kappa \mathrm{Trend}_r + \varepsilon_{ry}], \tag{4}$$

where

"Lag*y" represents the dependent variables lagged by one year.
"Year" represents the yearly fixed effect variable.
"Trend" represents the linear trend per region. The linear trend is a continuous variable from 1 to 12, associated to each year of the analysis, that is multiplied by each region.
"ε_{ry}" is a gamma error term with mean 1.0 and variance α^2.

We use two offsets: population size and total number of vehicles.

Based on our model comparisons, we selected the best fitting model in which the errors were not autocorrelated. For traffic fatality population rates the best model was:

$$y_{ry} = \exp[\zeta \mathrm{LAG}^* y_{ry} + \gamma \mathrm{Law}_{ry} + \theta \mathrm{Controls}_{ry} + \varepsilon_{ry}], \tag{5}$$

For traffic fatality motor-vehicle rates the best model was:

$$y_{ry} = \exp[\zeta \mathrm{LAG}^* y_{ry} + \gamma \mathrm{Law}_{ry} + \gamma \mathrm{Law\text{-}trend}_{ry} + \theta \mathrm{Controls}_{ry} + \varepsilon_{ry}], \tag{6}$$

where

"Law" is the presence and absence of the law in the period of analysis (0 = pre-intervention; 1 = post-intervention).
"Law-trend" is an increment over time one year after the law took effect.
"Controls" represent 'alcohol, tobacco, and related beverages expenditure,' 'unemployment rate,' 'number of health professionals per capita,' and 'number of crashes.'.

We report results from negative binomial models (results from Poisson models are available in supplementary material Tables S3–S6), with estimates presented on the incidence rate ratio (IRR) scale with robust standard errors. To suggest best fit models we include Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) factors. The lowest values with a distance of 10 points suggest the best specified model (Raftery, 1995). All analyses are performed using Stata 16 (StataCorp, 2017).

3.4. Formulas to calculate number of fatalities prevented by the law 6487

Based on previous work (Wagner et al., 2002), we present three formulas derived from equation 5 to compute the rate of traffic fatalities per population or motor vehicles with the law 6487 versus the counterfactual—the law had not ever been enacted. Computing differences for each period between the rates with and without the law 6487 provides estimated absolute numbers of prevented traffic fatalities in Turkey.

Formula 1 "With law: $Y_{ry} = \exp[\zeta \mathrm{LAG}^* Y_{ry} + \gamma \mathrm{Law}_{ry} {}^*6 + \theta \mathrm{Controls}_{ry} {}^*12 + \varepsilon_{ry}]$.

Formula 2 "Without law: $Y_{ry} = \exp[\zeta \mathrm{LAG}^* Y_{ry} + \theta \mathrm{Controls}_{ry} {}^*12 + \varepsilon_{ry}]$.

Formula 3 Change in relative percentage: (Formula 2 —Formula 1) /Formula 2 * 100.

In Formula 1, number 6 corresponds to 6 years after the law, and number 12 for the total number of years of the period of analysis. In Formula 2, number 12 indicates the total number of years for the period of analysis.

4. Results

4.1. Descriptives

Fig. 1 shows a steady decline in Turkey's traffic fatalities (per 100,000 population or 10,000 motor vehicles) from 2008 to 2019. Specifically, the rate of traffic fatalities per population dropped by 34% ((13.6–8.93)/13.6), and the rate of traffic fatalities per motor vehicles by 59% ((9.1–3.7))/9.1). Fig. 2 depicts cross regional variation in these two outcomes for the same period in all 12 administrative regions of Turkey. These numbers suggest considerable cross-regional variation in 2008. Whereas in Bati Karadeniz, Bati Marmara and Orta Anadolu had an average 20.0 traffic fatalities per 100,000 population, Ortadogu Anadolu, Güneydogu Anadolu and Istanbul had lower than 10.0. A very similar pattern was observed in traffic fatalities per 10,000 vehicles. In 2008 Bati Karadeniz, Orta Anadolu, Dogu Karadeniz and Ortadogu Anadolu had more than 10.0 traffic fatalities in this outcome, whereas in Istanbul and Aegean this rate was lower than 9.0. When attention is focused on the year 2019, we observe that variation between and within regions decreased overall.

4.2. Autocorrelation results

In Tables S7–S8 (supplementary material), we report results for Born and Breitung tests using both denominators and traffic fatalities that occurred at the crash scene, with the correction factor for traffic fatalities as if they had occurred 30 days after the crash for the 2008–2014 period. We observe that the only model that successfully corrects for autocorrelation is when a lag for one year in traffic fatality is introduced (Lag model (Eq. (2))). Every other model has Born and Breitung values with $p < 0.001$, suggesting that autocorrelation is present, and therefore results from these models may be biased.

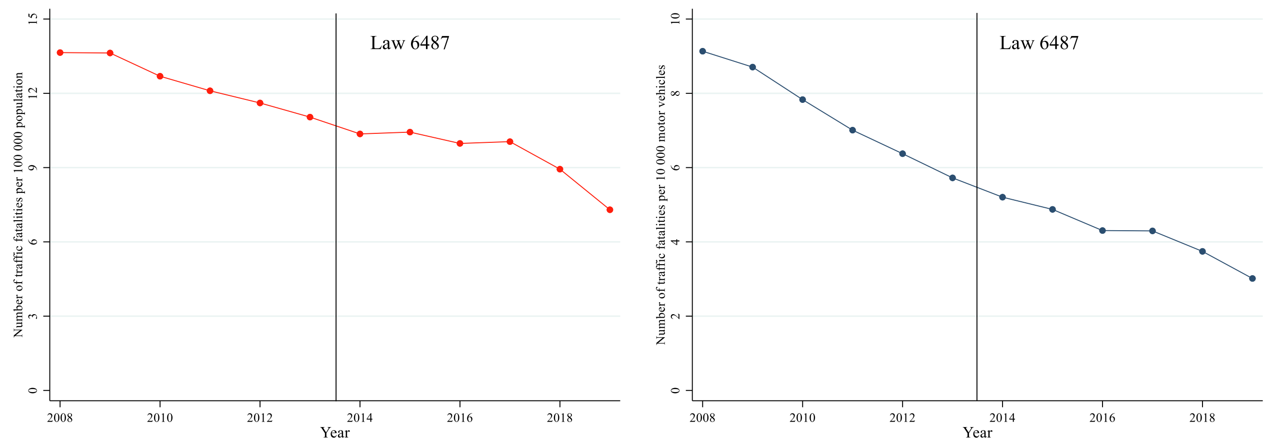

Fig. 1. Traffic fatalities per 100 000 population (in red) and 10 000 vehicles (in blue), 2008–2019. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

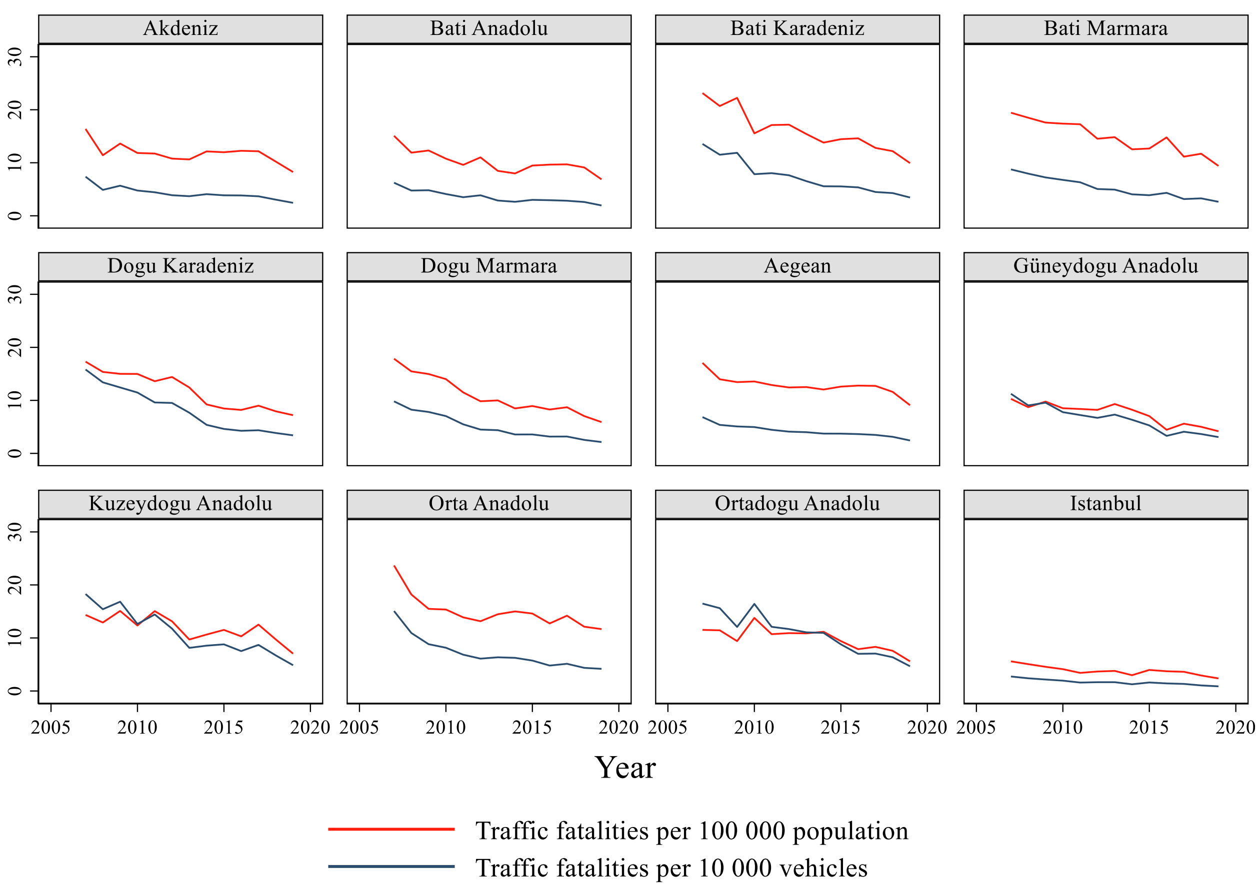

Fig. 2. Traffic fatalities per 100 000 population and 10 000 vehicles, 2008–2019, per each administrative Turkish region.

4.3. Fixed effects results with and without year 2013

In Tables 2 and 3 we report the results of traffic fatalities per 100,000 population and per 10,000 motor vehicles, respectively. Models in which we exclude year 2013 are the ones that better specify our multivariable results in both rates (all AIC and BIC results are the lowest and larger than 10 points). In terms of the law, we observe concomitant variations in every model and the estimator is stable. Model 1 suggests a 12% (IRR 0.88, 95% CI: 0.82, 0.94) decrease in traffic fatalities per 100,000 population. A similar decrease is observed in Model 2 when the number of crashes variable is considered (IRR 0.86, 95% CI: 0.78, 0.92). Model 5 suggests a 15% (IRR 0.85, 95% CI: 0.79, 0.91) decrease in traffic fatalities per 10,000 motor vehicles. Further, a linear association is also observed with the law trend variable. Every additional year of the law is associated with a 7% (IRR 0.94, 95% CI: 0.91, 0.96) decrease in the same outcome. In terms of the control variables, no concomitant associations regarding traffic fatality per popula-

Table 2
Fixed effects negative binomial models with traffic fatalities at the crash site per 100 000 population in Turkey 2008–2019.

	Traffic fatalities per 100 000 population											
	Model 1 Without year 2013 and without number of crashes			Model 2 Without year 2013 and with number of crashes			Model 3 With year 2013 and without number of crashes			Model 4 With year 2013 and with number of crashes		
Variable	**IRR**	**95% CI**		**IRR**	**95% CI**		**IRR**	**95% CI**		**IRR**	**95% CI**	
Law	**0.88**	**0.82**	**0.94**	**0.85**	**0.78**	**0.92**	**0.90**	**0.84**	**0.94**	**0.88**	**0.82**	**0.94**
Unemployment rate	1.00	0.99	1.01	1.00	0.99	1.02	1.00	0.99	1.01	1.00	0.99	1.02
Number of health professionals per capita	0.99	1.00	1.00	0.99	1.00	1.00	0.99	1.00	1.00	0.99	1.00	1.00
Alcohol, tobacco, and related beverages expenditure	0.97	0.87	1.07	0.96	0.87	1.07	0.95	0.86	1.04	0.95	0.86	1.05
Number of crashes				1.10	0.97	1.24				1.07	0.94	1.19
Lag of traffic fatalities	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Born and Breitung test	1.40*p* = 0.161			0.62*p* = 0.534			0.89*p* = 0.372			0.72*p* = 0.471		
AIC	1074.093			1073.797			1190.889			1191.776		
BIC	1090.818			1093.309			1208.186			1211.956		
Number of year-region	120			120			132			132		

Note. AIC: Akaike information criterion; BIC: Bayesian Information Criterion. CI: Confidence interval; IRR: Incidence rate ratio. Bold indicates $p < 0.05$. To obtain 95% CI standard errors were clustered at the regional level.

Table 3
Fixed effects negative binomial models with traffic fatalities at the crash site per 10 000 motor vehicles in Turkey 2008–2019.

	Traffic fatalities per 10 000 motor vehicles											
	Model 5 Without year 2013 and without number of crashes			Model 6 Without year 2013 and with number of crashes			Model 7 With year 2013 and without number of crashes			Model 8 With year 2013 and with number of crashes		
Variable	**IRR**	**95% CI**		**IRR**	**95% CI**		**IRR**	**95% CI**		**IRR**	**95% CI**	
Law	**0.86**	**0.79**	**0.91**	**0.85**	**0.78**	**0.92**	**0.88**	**0.83**	**0.94**	**0.89**	**0.83**	**0.96**
Law trend	**0.93**	**0.91**	**0.95**	**0.93**	**0.91**	**0.95**	**0.93**	**0.91**	**0.96**	**0.94**	**0.91**	**0.96**
Unemployment rate	**1.02**	**1.01**	**1.04**	**1.03**	**1.01**	**1.04**	**1.02**	**1.01**	**1.04**	**1.02**	**1.00**	**1.04**
Number of health professionals per capita	0.99	099	1.00	0.99	0.99	1.00	**0.99**	**0.99**	**1.00**	0.99	1.00	1.00
Alcohol, tobacco, and related beverages expenditure	0.97	0.87	1.08	0.97	0.87	1.08	0.96	0.86	1.06	0.96	0.86	1.07
Number of crashes				1.02	0.90	1.15				0.96	0.85	1.08
Lag of traffic fatalities	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Born and Breitung test	1.32*p* = 0.187			1.37*p* = 0.171			0.89*p* = 0.372			0.72*p* = 0.471		
AIC	1093.183			1095.062			1216.694			1218.23		
BIC	1112.696			1117.362			1236.873			1241.293		
Number of year-region	120			120			132			132		

Note. AIC: Akaike information criterion; BIC: Bayesian Information Criterion; CI: Confidence interval; IRR: Incidence rate ratio. Bold indicates $p < 0.05$. To obtain 95% CI standard errors were clustered at the regional level.

tion rate variation are observed. In traffic fatality per motor-vehicles models, we observe that 'number of health professionals per capita' is associated with this rate only in Model 7. One unit increase in the rate of health professionals per population is associated with 1% decrease of traffic fatalities per year. Lastly, we observe that 'unemployment' captures concomitant variation in traffic fatalities per 10,000 motor vehicles. A one-unit increase in the unemployment rate is associated with a 2% increase in the traffic fatality per 10,000 motor-vehicle rate per year.

4.4. Traffic fatalities prevented by the law 6487

Results from Tables 2 and 3 can be used to derive the number of traffic fatalities prevented as of 2019. We use Formulas 1, 2 and 3 to compute the change in relative percentage in year rates of traffic fatalities "with" the law. These results are reported in Table 4. Based on Model 2, which has the lowest AIC, a total of 1,519 fatalities were avoided in Turkey in the post law period (2014–2019), which represents a reduction of 6.56%. Our estimates from Model 5, with the lowest AIC value, suggests that 1,655 fatalities were prevented, representing a relative reduction of 7.14%.

Table 4
Traffic fatalities prevented by the law 6487.

Intervention	Total traffic fatalities prevented (95 CI%)	Change in relative percentage (95 CI %)	Chosen model
Traffic fatalities at the crash site per population*	1519 (1177, 1810)	6.56% (5.08, 7.81)	2
Traffic fatalities at the crash site per motor vehicles*	1655 (1521, 1767)	7.15% (6.57, 7.63)	5

Note. Results in brackets represent the lower and the upper bound of the estimates (95% Confidence Interval (CI)).
[*] Subtraction symbols mean that there were fewer fatalities with than without the law.

5. Discussion

Our study suggests that an alcohol-related traffic law (i.e., law 6487) is associated with concomitant changes in the rates of traffic fatalities per population and motor vehicles in Turkey. Between January 2008 and December 2019 our estimates of the law point to reductions of 15% and 14% in the population and motor-vehicle rates, respectively. We observe that 1,519 and 1,655 fatalities were prevented if we consider models for traffic fatality per population rates and motor-vehicles rates, respectively.

We suggest three interrelated hypotheses to understand why the law 6487 could be associated with reductions in traffic fatality outcomes. As we mentioned, this law considered both direct and indirect interventions to target DUIA. Thus a combination of factors may have been set in place to deter individuals from first drinking alcohol, and then if they drank alcohol to avoid DUIA. First, while our variable to measure alcohol consumption at the regional level was not associated with traffic fatalities in our study, other studies have empirically confirmed that in Turkey tax increases and hours of alcohol sales restrictions of this law have been associated with reductions at the national level in alcohol consumption (Alkan et al., 2021; Koç and Koç, 2020). Second, in terms of its direct intervention characteristics, we also propose that individuals who may have consumed alcohol were increasingly able to avoid DUIA because alternative means of transportation were available as other studies have suggested in other jurisdictions (Fell and Scherer, 2017). Another complementary possibility is that effective sanctions associated with this law, particularly the introduction of administrative sanctions, may have been successful in moving away risky drivers from the roads (Fell et al., 2015).

Our results indicate decreases of 15% and 14% in traffic fatalities per population and motor vehicle, respectively. These reductions observed in our study deserve the discussion of an alternative hypothesis. While the direction and magnitude of the law is similar in these two outcomes, the model for motor-vehicle rates considered the inclusion of a post-trend (i.e., a linear estimation that fatalities were decreasing over time after the introduction of the law). Further, when analyses to obtain the number of lives saved considered the post-trend estimator, considerable differences between traffic fatalities per population and per motor-vehicle fleet were found (see Table S9). Previous studies of low- and middle-income countries have indicated that important increases in motor vehicles are associated with traffic fatality reductions. Two explanations have been advanced to understand this paradox. On the one hand, road vehicle saturation is associated with slower speeds, or with the other presence of a "substitute effect," whereby road users shift from walking or using the public transportation system to driving private vehicles (Bhalla et al., 2007). In our case, even if we controlled for several factors, the magnitude of the association of the law with traffic fatalities per motor vehicle could be confounded with the speed of the number of vehicles introduced in Turkey in the period in which the law was enacted as the post-trend suggests. In fact, the annual average growth in motor vehicle was 4.8% between 2008 and 2019, whereas the annual average growth for the population has been 1.3% (Turkish Statistical Institute (TURKSTAT), 2021). In short it may not be the law capturing a post period reduction change, but rather people opting for private transportation use in which the number of pedestrians exposed to crashes is by default reduced. Overall, these results may be taken with extreme caution because a previous study (Castillo-Manzano et al., 2014), applying a similar methodology as ours to 27 European countries and using traffic fatalities per population, found an 11% decrease in this rate, whereas results in Canada (Blais et al., 2015) or Slovenia (Kralj et al., 2009), which focused on the effect of administrative changes for not respecting the BAC limit of 0.05 g/dL on alcohol-related driver fatality rates, more conservative changes of 22 and 9%, respectively.

Our evaluation of the general effect of law 6487 in Turkey has the following limitations. First, our main outcome did not consider whether alcohol consumption was associated with the recorded traffic fatalities analyzed. Efforts in Turkey should be moved forward to report this type of information as currently the TSI may not have access to it. Forensic studies of Turkey from 2005 to 2018 suggest, however, that between 10% and 17% of individuals attended in emergency services due to a crash event had BAC levels over 0.05 g/dL (Aygencel et al., 2008; Demirel et al., 2018). Analyses of alcohol tests of individuals who died in a motor-vehicle crash should be carried out to confirm or reject results presented in this study. Second, our proxy to capture alcohol consumption considered regional averages of how households allocated budgets to buy alcohol and tobacco simultaneously. While there is an important overlap between tobacco and alcohol purchase in Turkey since single males are more likely to purchase alcohol and tobacco (Aksoy et al., 2019), other characteristics such as level of education, have opposite directions. In Turkey higher educated people are more likely to buy alcohol and less educated people more likely to spend their budget on tobacco. The heterogeneity of our proxy may explain its null effect on this study, further signaling the need of finding adequate information regarding alcohol consumption when studying DUIA in Turkey. Third, as is common in road safety studies, the inspection of variables representing how traffic police reinforce these policies remains an important task. The effects of the law could indeed be mediated by different police monitoring strategies. Importantly, not having access to information regarding DUIA arrests limits our understanding of an important mechanism that could explain the observed variation. Last, our law estimates may be confounded by other unmeasured time-varying variables, such as increases in seat belt use or road infrastructure improvements, which may have occurred in Turkey (Miller et al., 2018). Studies have reported that in Turkey in 2010, 20% of drivers wore seat belts (Bilgic et al., 2011) whereas in 2018 this rate increased to 50% (Global status report on road safety, 2018).

Practical applications

Our research emphasizes that legislation targeting DUIA may be related to traffic fatalities, but reductions may be a function of policies considering direct and indirect interventions. This finding has important implications for policy and future research in contexts in which alcohol consumption is relatively low, as the case of Turkey suggest. Authorities should thus consider direct and indirect interventions before moving forward measures that only target DUIA with increases of penalties. It is also noteworthy that DUIA proven countermeasures, such as administration license suspension, publicized sobriety checkpoints, and alcohol ignition interlocks for convicted DUIA offenders, which were not part of Turkey's reform, should also explicitly guide these efforts. Last, future research should seek to identify how public transportation can support the decision of avoiding DUIA when individuals have consumed alcohol. Rigorous scientific study will inform policies that maximize the potential road safety and public health benefits in places where alcohol consumption has unique patterns.

Funding

Nazif-Munoz was supported through a career award from the Fonds de recherche du Québec–Santé [269540] and Ouimet was also supported through a career award from the Fonds de recherche du Québec–Santé [252261].

Conflict of interest

The authors declare that they have no competing interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsr.2022.08.015.

References

Aksoy, A., Bilgic, A., Yen, S. T., & Urak, F. (2019). Determinants of household alcohol and tobacco expenditures in Turkey. *Journal of Family and Economic Issues, 40*(4), 609–622. https://doi.org/10.1007/s10834-019-09619-1.

Ali, Q., Yaseen, M. R., & Khan, M. T. I. (2019). Road traffic fatalities and its determinants in high-income countries: a continent-wise comparison. *Environmental Science and Pollution Research, 26*(19), 19915–19929. https://doi.org/10.1007/s11356-019-05410-9.

Alkan, Ö., Abar, H., & Gençer, Ö. (2021). Analysis of factors affecting alcohol and tobacco concurrent use by bivariate probit model in Turkey. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-021-12849-2.

Ames, C. L., Lu, Y., & Weaver, R. (2016). Evaluating the effectiveness of "no refusal" blood draw search warrant programs to reduce alcohol-related traffic crashes in central Texas. *Papers in Applied Geography, 2*(1), 113–120. https://doi.org/10.1080/23754931.2015.1114513.

Aygencel, G., Karamercan, M., Ergin, M., & Telatar, G. (2008). Review of traffic accident cases presenting to an adult emergency service in Turkey. *Journal of Forensic and Legal Medicine, 15*(1), 1–6 https://www.sciencedirect.com/science/article/pii/S1752928X07000868.

Babor, T. F., Caetano, R., Casswell, S., Edwards, G., Giesbrecht, N., Graham, K., et al. (2010). Alcohol: No ordinary commodity: Research and public policy [Internet]. 2.a ed. Oxford: Oxford University Press; 2010 [citado 17 de septiembre de 2021]. pp. 384. Disponible en: https://oxford.universitypressscholarship.com/10.1093/acprof:oso/9780199551149.001.0001/acprof-9780199551149.

Bartl, G., Esberger, R. (2000). Effects of lowering the legal BAC limit in Austria. En 2000 [citado 27 de mayo de 2021]. Disponible en: https://trid.trb.org/view/674163.

Bhalla, K., Ezzati, M., Mahal, A., Salomon, J., & Reich, M. (2007). A risk-based method for modeling traffic fatalities. *Risk Analysis: An International Journal, 27*(1), 125–136 https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.2006.00864.x.

Bilgic, S., Barut, H. B., Karacasu, M., Er, A., & Yaliniz, P. (2011). The changes in usage of seat belts in Antalya, Turkey. *Procedia-social and Behavioral Sciences, 20*, 588–593 https://www.sciencedirect.com/science/article/pii/S1877042811014455.

Blais, É., Bellavance, F., Marcil, A., & Carnis, L. (2015). Effects of introducing an administrative 05% blood alcohol concentration limit on law enforcement patterns and alcohol-related collisions in Canada. *Accident Analysis & Prevention, 82*, 101–111 https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2015-33605-013&site=ehost-live.

Cabrera-Arnau, C., Prieto Curiel, R., & Bishop, S. R. (2020). Uncovering the behaviour of road accidents in urban areas. *Royal Society Open Science, 7*(4) https://royalsocietypublishing.org/doi/full/10.1098/rsos.191739.

Castillo-Manzano, J. I., Castro-Nuño, M., & Fageda, X. (2014). Can health public expenditure reduce the tragic consequences of road traffic accidents? The EU-27 experience. *The European Journal of Health Economics, 15*(6), 645–652 http://link.springer.com/10.1007/s10198-013-0512-1.

Castillo-Manzano, J. I., Castro-Nuño, M., Fageda, X., & López-Valpuesta, L. (2017). An assessment of the effects of alcohol consumption and prevention policies on traffic fatality rates in the enlarged EU. Time for zero alcohol tolerance? 1 de octubre de 2017 [citado 27 de mayo de 2021] *Transportation Research Part F: Traffic Psychology and Behaviour, 50*, 38–49 https://www.sciencedirect.com/science/article/pii/S1369847816301838.

Cetin, V. R., Yilmaz, H. H., & Erkan, V. (2018). The impact of increasing speed limit in Turkey: The case of Ankara-Sivrihisar road section. *Case Studies on Transport Policy, 6*(1), 72–80 https://www.sciencedirect.com/science/article/pii/S2213624X17301621.

Demirel, B., Ergin, M., Özhasenekler, A., Teke, C., Tanrıverdi, F., Çelik, G. K., et al. (2018). The influence of alcohol related accidents on health care cost. *Disaster and Emergency Medicine Journal, 3*(3), 71–74 https://journals.viamedica.pl/disaster_and_emergency_medicine/article/view/DEMJ.2018.0017.

Deshapriya, E. B. R., & Iwase, N. (1970). Impact of the 1970 legal BAC 0.05 mg% limit legislation on drunk-driver-involved traffic fatalities, accidents, and DWI in Japan. *Substance Use & Misuse, 33*(14), 2757–2788 http://www.tandfonline.com/doi/full/10.3109/10826089809059349.

Elvik, R., Høye, A., Vaa, T., Sørensen, M. (Editors), (2009). The Handbook of Road Safety Measures [Internet]. Emerald Group Publishing Limited; 2009 [citado 14 de junio de 2021]. Disponible en: http://www.emeraldinsight.com/doi/book/10.1108/9781848552517.

Fell, J. C., & Scherer, M. (2017). Estimation of the Potential Effectiveness of Lowering the Blood Alcohol Concentration (BAC) Limit for Driving from.08 to.05 grams per deciliter in the United States diciembre de 2017 [citado 27 de mayo de 2021]. *Alcoholism: Clinical and Experimental Research, 41*(12), 2128–2139 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5790204/.

Fell, J. C., & Voas, R. B. (2014). The effectiveness of a 0.05 blood alcohol concentration (BAC) limit for driving in the United States. *Addiction, 109*(6), 869–874.

Fell, J. C., Waehrer, G., Voas, R. B., Auld-Owens, A., Carr, K., & Pell, K. (2015). Relationship of impaired-driving enforcement intensity to drinking and driving on the roads. *Alcoholism: Clinical and Experimental Research, 39*(1), 84–92 https://onlinelibrary.wiley.com/doi/abs/10.1111/acer.12598.

Freivalds, A., & Johnson, A. B. (1990). Time-series analysis of industrial accident data. *Journal of Occupational Accidents, 13*(3), 179–193 https://www.sciencedirect.com/science/article/pii/037663499090020V.

Global status report on road safety 2018 [Internet]. [citado 24 de septiembre de 2021]. Disponible en: https://www.who.int/publications-detail-redirect/9789241565684.

Greenfield, T. K., Graves, K. L., & Kaskutas, L. A. (1999). Long-term effects of alcohol warning labels: Findings from a comparison of the United States and Ontario. *Canada. Psychology & Marketing, 16*(3), 261–282 https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291520-6793%28199905%2916%3A3%3C261%3A%3AAID-MAR5%3E3.0.CO%3B2-Z.

Gupta, S., Hoe, C., Özkan, T., Lajunen, T. J., Vursavas, F., Sener, S., et al. (2017). Evaluation of a five-year bloomberg global road safety program in Turkey. *Public Health, 144*, S45–S56 https://www.sciencedirect.com/science/article/pii/S0033350617300136.

Haghpanahan, H., Lewsey, J., Mackay, D. F., McIntosh, E., Pell, J., Jones, A., et al. (2019). An evaluation of the effects of lowering blood alcohol concentration limits for drivers on the rates of road traffic accidents and alcohol consumption: a natural experiment. *The Lancet, 393*(10169), 321–329 https://www.sciencedirect.com/science/article/pii/S0140673618328502.

Homel, R., McKay, P., Henstridge, J. (1995). The impact on accidents of random breath testing in new south wales: 1982–1992. undefined [Internet]. 1995 [citado 27 de mayo de 2021]; Disponible en: https://paper/The-impact-on-accidents-of-random-breath-testing-in-Homel-McKay/1b322169594560ba6c60a457b7b76b8b654efaeb.

Koç, Ö. E., Koç, N. (2020). ALKOL TÜKETİMİ İLE ALKOL ÜZERİNDEN ALINAN ÖZEL TÜKETİM VERGİSİ İLİŞKİSİNİN İNCELENMESİ. Önetim Ve Ekon Araştırmaları Derg [İnternet]. 30 de septiembre de 2020 [citado 17 de junio de 2021]; vol. 18, no. 3, pp. 212–226. Disponible en: https://dergipark.org.tr/tr/pub/yead/676661.

Kralj, E., Pezdir, T., & Balažic, J. (2009). Post-mortem blood alcohol concentration of the traffic accident victims and changes in DUI legislation in Slovenia 1980–2006. *Forensic Science International Supplement Series, 1*(1), 46–51 https://www.sciencedirect.com/science/article/pii/S1875174109000317.

Kuloglu Genc, M., Mercan, S., Yayla, M., Tekin Bulbul, T., Adioren, C., Simsek, S. Z., et al. (2021). Monitoring geographical differences in illicit drugs, alcohol, and tobacco consumption via wastewater-based epidemiology: Six major cities in Turkey. *Science of the Total Environment, 797* https://www.sciencedirect.com/science/article/pii/S0048969721042297.

Lavoie, M. C., Langenberg, P., Villaveces, A., Dischinger, P. C., Simoni-Wastila, L., Hoke, K., et al. (2017). Effect of Maryland's 2011 alcohol sales tax increase on alcohol-positive driving. *American Journal of Preventive Medicine, 53*(1), 17–24 https://search.ebscohost.com/login.aspx?direct=true&db=rzh&AN=123529630&site=ehost-live.

Lovenheim, M. F., & Slemrod, J. (2010). The fatal toll of driving to drink: The effect of minimum legal drinking age evasion on traffic fatalities. *Journal of Health Economics, 29*(1), 62–77.

Martin-Moreno, J. M., Harris, M. E., Breda, J., Møller, L., Alfonso-Sanchez, J. L., & Gorgojo, L. (2013). Enhanced labelling on alcoholic drinks: Reviewing the evidence to guide alcohol policy. *European Journal of Public Health, 23*(6), 1082–1087. https://doi.org/10.1093/eurpub/ckt046.

Mercier-Guyon, C. (1998). Lowering the BAC limit to 0.05: Results of the French experience. En Washington, DC.

Miller, T. R., Levy, D. T., & Swedler, D. I. (2018). Lives saved by laws and regulations that resulted from the Bloomberg road safety program. *Accident Analysis & Prevention, 113*, 131–136.

Nazif-Muñoz, J. I. (2015). Did child restraint laws globally converge? Examining 40 years of policy diffusion. *Traffic Injury Prevention, 16*(Suppl 2), S32–S40.

Oğurlu, Y. (2005). Administrative Sanctioning system in Turkey. *Erzincan Binali Ildırım Üniversitesi Hukuk Fakültesi Derg, 9*(3–4), 507–514.

Özdemir, R., Rao, C., Öcek, Z., & Dinç, Horasan G. (2015). Reliable mortality statistics for Turkey: Are we there yet? *BMC Public Health, 15*(1), 545. https://doi.org/10.1186/s12889-015-1904-1.

Popova, S., Giesbrecht, N., Bekmuradov, D., & Patra, J. (2009). Hours and days of sale and density of alcohol outlets: impacts on alcohol consumption and damage: A systematic review. *Alcohol & Alcoholism, 44*(5), 500–516. https://doi.org/10.1093/alcalc/agp054.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163 https://www.jstor.org/stable/271063.

Rehm, J., Manthey, J., Lange, S., Badaras, R., Zurlyte, I., Passmore, J., et al. (2020). Alcohol control policy and changes in alcohol-related traffic harm. *Addiction, 115*(4), 655–665.

Republic of Turkey (2013). Law Number 6487. Official Gazette.

Sanchez-Ramirez, D. C., & Voaklander, D. (2018). The impact of policies regulating alcohol trading hours and days on specific alcohol-related harms: a systematic review. *Injury Prevention, 24*(1), 94–100 https://search.ebscohost.com/login.aspx?direct=true&db=mnh&AN=28647704&site=ehost-live.

Smailović, E., Pešić, D., Marković, N., Antić, B., & Lipovac, K. (2020). The effects of a change in permissible blood alcohol concentration limit on involving drink-driving in road accidents 1 de enero de 2020 [citado 27 de mayo de 2021]. *Transportation Research Procedia, 45*, 201–209 https://www.sciencedirect.com/science/article/pii/S2352146520301708.

StataCorp (2019). Stata statistical software: Release 16. College Station, TX: StataCorp LLC.
Turkish Statistical Institute (TURKSTAT) [Internet]. [citado 15 de junio de 2021]. Disponible en: https://www.tuik.gov.tr/Home/Index.
Turkish Statistical Institute (2019). Road Traffic Accident Statistics, 2019 [Internet]. 2019. Disponible en: https://turkstatweb.tuik.gov.tr/PreHaberBultenleri.do?id=33628.
Vissers, L., & Houwing, W. (2017). *Alcohol-related road casualties in official crash statistics* (p. 55). Paris: International Transport Forum ITF.
Voas, R. B., Kelley-Baker, T., Romano, E., & Vishnuvajjala, R. (2009). Implied-consent laws: A review of the literature and examination of current problems and related statutes. *Journal of Safety Research, 40*(2), 77–83 https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2009-07209-003&site=ehost-live.
Wagenaar, A. C., Livingston, M. D., & Staras, S. S. (2015). Effects of a 2009 illinois alcohol tax increase on fatal motor vehicle crashes. *American Journal of Public Health, 105*(9), 1880–1885.
Wagner, A. K., Soumerai, S. B., Zhang, F., & Ross-Degnan, D. (2002). Segmented regression analysis of interrupted time series studies in medication use research. *Journal of Clinical Pharmacy and Therapeutics, 27*(4), 299–309 https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2710.2002.00430.x.
Wegman, F., Allsop, R., Antoniou, C., Bergel-Hayat, R., Elvik, R., Lassarre, S., et al. (2017). How did the economic recession (2008–2010) influence traffic fatalities in OECD-countries? *Accident Analysis & Prevention, 102*, 51–59 https://www.sciencedirect.com/science/article/pii/S0001457517300490.
Wegman, F., & Oppe, S. (2010). Benchmarking road safety performances of countries. *Safety Science, 48*(9), 1203–1211 https://www.sciencedirect.com/science/article/pii/S0925753510000378.
Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data [Internet]. Cambridge, Mass.: MIT Press; 2002 [citado 31 de mayo de 2021]. 794 p. Disponible en: http://archive.org/details/econometricanaly0000wool.
World Bank (2021). Total alcohol consumption per capita (liters of pure alcohol, projected estimates, 15+ years of age) – Turkey, Australia, Brazil [Internet]. Total alcohol consumption per capita (liters of pure alcohol, projected estimates, 15+ years of age) – Turkey, Australia, Brazil. 2021. Disponible en: https://data.worldbank.org/indicator/SH.ALC.PCAP.LI?locations=TR-AU-BR.
World Health Organization (2004). World report on road traffic injury prevention [Internet]. 2004 [citado 14 de junio de 2021]. Disponible en: https://www.who.int/publications-detail-redirect/world-report-on-road-traffic-injury-prevention.
World Health Organization (2018). Global status report on alcohol and health 2018 [Internet]. 2019 [citado 27 de mayo de 2021]. Disponible en: https://www.who.int/publications-detail-redirect/9789241565639.
Wright, N. A., & Lee, L. T. (2021). Alcohol-related traffic laws and drunk-driving fatal accidents. *Accident Analysis & Prevention, 161* https://www.sciencedirect.com/science/article/pii/S0001457521003894.
Wursten, J. (2018). XTHRTEST: Stata module to perform Born & Breitung Bias-corrected HR-test for first order panel serial correlation [Internet]. 2018 [citado 31 de mayo de 2021]. (Statistical Software Components). Disponible en: https://econpapers.repec.org/software/bocbocode/S458235.htm.

José Ignacio Nazif-Munoz is assistant professor at fellow at Les programmes d'études et de recherche en toxicomanie, in the Faculté de médecine et des sciences de la santé, Université de Sherbrooke. He completed his BA Sociology (1999) at Universidad de Chile, and MA Sociology (2004) and PhD in Sociology (2016) at McGill University. He did his postdoctoral training at the Institute of Health and Policy at McGill University (2017) and Harvard University, at the T. H. Chan School of Public Health (2019). His key research interest area is injury prevention currently focused on the determinants of road safety policy across the world. He focuses on multiple countries in order to understand the implications of global promoted road safety policies.

Dr. Gül Anıl Anakök is a physician and researcher at the Kartepe District Health Directorate Kocaeli, Turkey. She received her medicine degree from Kocaeli University (2012). Her research interests are injury epidemiology with attention to gender differences within Turkey.

Junon Joshep is a physician and master student fellow at Les programmes d'études et de recherche en toxicomanie, in the Faculté de médecine et des sciences de la santé, Université de Sherbrooke. She completed her BA (Hons) (2015) and Medicine (2017) specializing in prevention medicine and at the Université d'État d'Haïti. Her research interests include the identification of road safety policies in Quebec and Haiti.

Santosh Kumar Uprajhiya is a master student fellow at Les programmes d'études et de recherche en toxicomanie the Faculté de médecine et des sciences de la santé, Université de Sherbrooke. He completed his BA (Hons) (2020) in prevention injuries Tribhuvan University, Nepal. His research interests include the identification of road safety policies in Alberta with attention to child injuries.

Marie Claude Ouimet is full professor at Les programmes d'études et de recherche en toxicomanie the Faculté de médecine et des sciences de la santé, Université de Sherbrooke. She has a B.Sc., Psychologie, Université du Québec à Montréal, a M.Sc., Psychologie (psychophysiologie et ergonomie), Université de Montréal, a Certificat, Informatique appliquée, Université de Montréal, a Ph.D., Psychologie (profil recherche), Université de Montréal, a Post-doctorat, Sciences comportementales et addictions, Institut universitaire en santé mentale Douglas, and a Post-doctorat, Sciences comportementales et prévention des blessures, National Institutes of Child Health & Human Development, Bethesda, MD, USA. She is the Director at Les programmes d'études et de recherche en toxicomanie and Director of the Réseau de recherche en sécurité routière du Québec . Her research interests include understanding and prevention of risky road behaviors, injury prevention, effectiveness of brief interventions in reducing risk behaviours, causal mechanisms and neurobiological markers, young drivers and passengers, and parenting and the prevention of risky behaviour

Journal of Safety Research 83 (2022) 232–237

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

A pilot study evaluating the effectiveness of preventing railway suicides by mid-track fencing, which restrict easy access to high-speed train tracks

Johan Fredin-Knutzén [a,*], Gergö Hadlaczky [a], Anna-Lena Andersson [b,c], Marcus Sokolowski [a]

[a] National Centre for Suicide Research and Prevention of Mental Ill-Health (NASP), Karolinska Institute, Stockholm, Sweden
[b] Swedish Transport Administration (STA), Kungsgatan 32, SE-461 30 Trollhättan, Sweden
[c] Institute of Clinical Sciences, Sahlgrenska Academy, Department of Orthopedics, University of Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 19 January 2022
Received in revised form 2 May 2022
Accepted 29 August 2022
Available online 8 September 2022

Keywords:
Suicide
Accident
Railway
Mid-track fence
Means restriction

ABSTRACT

Background: Suicides in the railway system is a serious health, societal, and transportation concern. Restriction of the access to suicide methods in the form of different physical barriers is a promising approach for suicide prevention. *Method:* Mid-track fencing, which is fencing placed in-between the high-speed and commuter train tracks, was installed at one out of seven stations along a train line outside of Stockholm in the years 2013/2014. The number of suicides at the intervention station was compared to six other stations used as controls, over a total period of 20 years (2002–2021). *Results:* Suicides at high-speed tracks occurring at stations was the major cause of death on the investigated railway line. Prior to the year 2014, the intervention and control stations displayed similar time trends in the number of suicides. After installation of the mid-track fencing in 2014, there was a 62.5% reduction in the rate of suicides occurring at the intervention station. Compared to the six other control stations, the intervention station displayed a significant reduction in the number of suicides during the years 2014–2021 (OR = 0.14, 95%CI 0.013–0.95). Suicides at the railway lines in-between stations were not increased post-intervention. However, nearby control stations showed a 162% increase in suicides after the intervention, suggesting the induction of transfer effects. *Conclusion*: Mid-track fences restricting access to high-speed train tracks may have a large effect on reducing the number of railway suicides at intervention stations, but may also induce an increase in suicides at nearby stations without mid-track fences. *Practical applications:* Partial physical barriers such as mid-track fencing is deemed to be relatively easy and cheap to install (as compared to full barriers; e.g., full height platform screen doors) and should be considered at all stations on railway lines that have high-speed trains passing by.

1. Introduction

Each year there are approximately 1,500 suicide deaths in Sweden (National Centre for Suicide Research and Prevention of Mental ill health (NASP)', 2020). Of these, approximately 130 suicides occur in the transport system overall and ~80 occur in the railway system specifically (Fredin-Knutzén et al., 2020). Aside from the tragic deaths, this phenomena has a negative impact on personnel working at the railway (Giupponi et al., 2019; Tranah & Farmer, 1994) and on passengers witnessing these tragic events. These suicides also affect the efficiency of public and cargo transportation.

In general, there is good evidence for reducing suicides by restrictions of suicide methods (Pirkis et al., 2015) and method substitution rarely counteracts this reduction in full (Zalsman et al., 2016). For this reason, means restriction is an important and recommended strategy for suicide prevention (Mann et al., 2021; Zalsman et al., 2016). In the railway system, means restriction by using full or half height platform doors have been shown to be effective, even though the effect for the latter appears to be smaller (Chung et al., 2016; Law et al., 2009; Ueda et al., 2015; Xing et al., 2019). Some studies also showed an effect regarding trenches located in-between the tracks, or so called "suicide pits" (Barker et al., 2017; Coats & Walter, 1999; O'Donnell & Farmer, 1994). However, there is currently limited evidence concerning other types of prevention measures through means restriction in the railway system (Ryan et al., 2018) and there is a need to

* Corresponding author at: National Centre for Suicide Research and Prevention of Mental Ill-Health (NASP), Karolinska Institute, S-171 77 Stockholm, Sweden.

E-mail address: johan.fredin.2@ki.se (J. Fredin-Knutzén).

https://doi.org/10.1016/j.jsr.2022.08.019
0022-4375/© 2022 The Author(s). Published by the National Safety Council and Elsevier Ltd.

develop better evidence for suicide prevention strategies and methods (Fredin-Knutzén et al., 2020; Mishara & Bardon, 2016).

In a previous study, Rådbo and Andersson (2012) investigated suicides and other trespass fatalities in the railway system of greater Stockholm and its urban areas. Their main findings showed that most fatalities in Stockholm occurred at station areas and that most victims entered the tracks from platforms. Passing express trains were overrepresented compared to commuter trains. Similarly, a recent Japanese study also found that suicides are more frequent at stations with passing trains (Sueki, 2021). Another study from Stockholm area showed that suicide rates tend to increase near high-speed trains and decrease where fences and noise-barrier walls are installed along tracks (Ceccato & Uittenbogaard, 2016).

Rådbo and Andersson (2012) suggested that preventive measures installed on station areas should be prioritized, at least in the Stockholm region. Indeed, the Swedish Transport Administration (STA) installed fences in 2014 between the tracks used by commuter trains and the tracks used by high-speed trains at one of the stations in northern Stockholm (Fig. 1; refer also to Fig. S1 and S2 in the supplement). This installation, here on referred to as "mid-track fencing," limited easy access to the high-speed trains but not to the lower speed commuter trains. These fences obstruct access to the more lethal high-speed track, even though access to the less lethal commuter tracks remains the same. This type of measure is similar to those mentioned in the Restrail toolbox as "Intermediate fencing between tracks" (Restrail-project, 2015a) or "mid-platform fencing" (Restrail-project, 2015b), thus referring to where the fence is installed. In a questionnaire study (Restrail-project, 2015b), these "mid-platform fences" were rated by traffic safety experts to have a high likelihood of preventing suicide.

However, to the best of our knowledge, mid-track fences have never been evaluated with actual suicides as an outcome. It is reasonable to hypothesize that the mid-track fence installed at one of the stations in Stockholm had a preventive effect. Therefore, the present study investigated the hypothesis that fewer suicides would be observed after installation of mid-track fencing at the intervention station.

2. Data and methods

2.1. Data

Data about suicide and non-suicidal accidents involving persons struck by a train were taken from the register of the STA. This database contains a summary of information regarding suicide and accidents and is based on case reports written by STA-investigators, who carry out extensive investigations for each fatal accident. The case-reports often contain the exact location of the accident, attached photos, tracking of the individuals' movement preceding the accident, and information from the police. The classifications of suicides used since 2015 were also improved, by use of extended psychosocial investigations to resolve more unclear cases (Andersson & Sokolowski, 2021; Fredin-Knutzén et al., 2020).

The data used in this study were extracted from the STA, but were also validated against the information in the case reports for each incident. All events were extracted involving persons that were hit by a train between the years 2002–2021 (n = 65 events in total), from eight consecutive stations at the same line in northern Stockholm area (see also Fig. S1 in the supplement). For example, the data included date of event, station name, location, whether the event was classified as suicide or accident, and outcome of the accident (death, injury, or no injury). We classified events as being at or in the immediate vicinity of the platforms if they occurred up to approximately 25 meters away from the platform area, as this corresponded to the coverage of the mid-track fence at the intervention station.

Fig. 1. Photo of the installed mid-track fence and a high-speed train passing by the intervened station. The lower speed commuter train track which has trains stopping at the station, is to the right of the high-speed train track in the picture. Also refer to the Supplement for a brief discussion about the fence design.

2.1.1. Intervention station

The inclusion criteria was that the station had 1 m high mid-track fences (Fig. 1) partially restricting access to all high speed train tracks at the station and a known date for the installation. The mid-track fences at the included intervention station were installed for both north- and southbound tracks, began to be installed during/at the end of 2013 and was finished at the beginning of 2014 (confirmed by using time-stamped photos). Another station was excluded from the analysis due to uncertainties regarding the installation date of a wider 80 cm high concrete barrier in the mid-track (see Fig. S3 in the supplement). The excluded station had 4 suicides observed during the years 2002–2021, which was relatively less compared to the 10 suicides observed at the included intervention station.

2.1.2. Control stations

The control consisted of six other stations on the same line as the intervention station (see Fig. S1 in the supplement), but having no mid-track-fence or other obstacle separating the commuter train tracks from the high-speed tracks. The inclusion criteria were that they should be as similar as the intervention station as possible. Therefore, all stations had tracks dedicated to the same high-speed trains, the same train frequency, and trains passing at approximately the same speeds. The accessibility to the high-speed tracks at these six stations was the same as for the intervention station prior to 2014, when there were no mid-track fences installed at any of the stations. Suicidal persons were thus able to step out in the same high speed tracks that are passing by the stations, with the same degree of accessibility for all included stations prior to year 2014 (Rådbo & Andersson, 2012).

2.2. Analyses

We hypothesized that the mid-track fence resulted in fewer suicides during the years 2014–2021 at the intervention station, compared to the prior years 2002–2013 and compared to the other six control stations without any mid-track fence (i.e., a one-tailed hypothesis). To assess the relative change in average number of suicides per year between intervention and control stations, a standard difference-in-differences (DD) regression analysis with the following model and dummy variables [coded 1/0] was used: # suicides = $\beta_0 + \beta_1$ * [time$_{2014\text{-}2021}$] + β_2 * [intervention station] + β_3 * [time$_{2014\text{-}2021}$ × intervention station]. The DD effect is the interaction-term and regression was conducted by using the *reg* command in Stata v.9.2 (Columbia University Mailman School of Public Health, 2019; StataCorp., 2005). To test the effect with statistical inferences about the count data (number of suicides), we calculated odds ratios and one-tailed Fisher's exact tests for the number of suicides occurring at the intervention station during years 2014–2021 versus 2002–2013, compared to the control stations during the same years, respectively. In addition, we plotted the simple moving averages of the number of suicides per year occurring at the intervention and control stations during 2002–2021 (Fig. 2). Finally, we also conducted a series of secondary analyses to investigate impacts on the results of putative uncertainties.

3. Results

3.1. Majority of rail traffic deaths were suicides at the high-speed train track at stations

There was a total of n = 65 events during the years 2002–2021 at the investigated rail line, wherein both suicides and accidents occurred more frequently at the stations (n = 40 and n = 10; see Table 1), in comparison to the line sections in-between stations (n = 11 and n = 4; data not shown). Furthermore, station suicides occurred mainly on the high-speed tracks rather than on the slower on commuter train tracks (Table 1). For example, among the six control stations there was a total of n = 30 suicides, of which n = 24 (80%) occurred in the high-speed track; this was more than n = 3 fatal high-speed rail accidents observed in total (Table 1). Together, these observations emphasized the importance of preventing station suicides specifically and to consider the high-speed train track in particular.

3.2. Mid-track fencing prevented suicides at the intervention station

In the years prior to the installation of the mid-track fences (2002–2013), the number of suicides occurring at the intervention station and surrounding control stations displayed similar levels (Table 1 and Fig. 2). However, after the installation of mid-track fences in 2014, the levels of suicides displayed a notable difference between the intervention and control stations (Table 1 and Fig. 2). The number of yearly suicides decreased by 62.5% at the intervention station (down from an average of 0.66 to 0.25 suicides/year), while increasing by 162% at the control stations (up from an average of 0.91 to 2.38 suicides/year), which corresponds to a relative reduction at the intervention station of −1.875 suicides/year (Table 1 and Fig. 2). The odds ratio (OR) for the number of suicides occurring at the intervention station during years 2014–2019 versus 2002–2013, compared to the control stations during the same years (2 and 8 *vs* 19 and 11; Table 1), showed a significant reduction in suicides occurring at the intervention station (*OR* = 0.14, 95%CI 0.013–0.95; Fisher's $p = 0.021$). Finally, restricting analysis to the subset of suicides occurring only on the high-speed track at stations (Table 1), also showed a relative reduction of −1.125 suicides/year and a similar OR effect size (*OR* = 0.18, 95%CI 0.003–2.94; Fisher's $p = 0.14$). However, it should also be noted, that half of the suicides on the intervened station occurred in the lower speed commuter train track, which were reduced to a similar extent as the suicides on the high-speed track (Table 1). However, analyzing only suicides at the lower-speed commuter track was not feasible, due to the low counts in the control group (Table 1). Together, results suggested that the mid-track fencing had a suicide preventive effect at the intervened station overall, including an effect on suicides at the high-speed track. However, there was also an increase of suicides at the control stations, which suggest that the mid-track fencing intervention may have induced a transfer of suicides away from the intervention station to the nearby control stations.

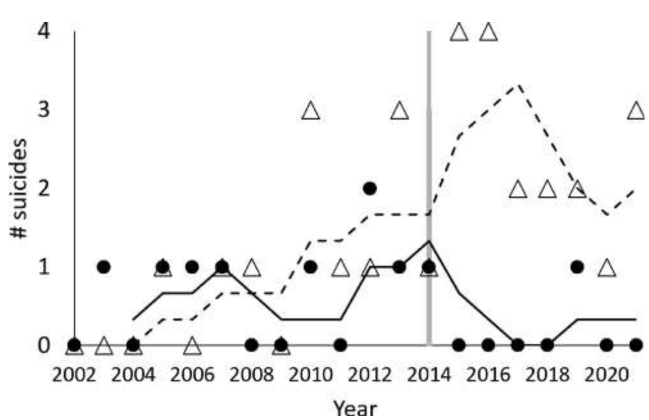

Fig. 2. Suicides at the intervention station (black dots) were apparently reduced after installation of a mid-track fence in year 2014 (vertical grey line), whereas among the six control stations (open triangles) suicides remained high. Solid and dashed lines depict the respective 3-year simple moving averages.

3.3. Secondary analyses supported a suicide preventive effect of the mid-track fencing at the intervention station

First, one of the high-speed track suicides at the intervention station occurred near a wide breach in the mid-track fence (Table 1 and Fig. S4 in the supplement), indicating the absence of a mid-track physical barrier in this case. Repeating OR analysis without this case showed a significant effect also for the high-speed track suicides *per se* (*OR* = 0, 95%CI 0–0.80; Fisher's $p = 0.049$). Thus, we observed a significant effect above despite this flaw in the mid-track fence design. Secondly, during the pandemic years 2020–2021, high-speed trains were periodically cancelled. Nevertheless, repeating analysis without inclusion of these years also showed a relative reduction of −1.92 suicides/year and a maintained OR effect size (*OR* = 0.18, 95%CI 0.002–1.23; Fisher's $p = 0.047$). Thirdly, the installation period of the mid-track fences was ongoing from the last half of 2013 until the first half of 2014. No suicides occurred at the intervention station during this period, but three suicides occurred at the control stations during the autumn of 2013. Nevertheless, repeating analysis without inclusion of this period also showed a relative reduction of −2.13 suicides/year and an improved OR effect size (*OR* = 0.11, 95%CI 0.01–0.73; Fisher's $p = 0.009$). Fourth and finally, there was no increase in the number of suicides occurring along the lines in-between stations in the years 2014–2021 (4 suicides, 0.5 suicides/year), compared to the prior years 2002–2013 (7 suicides, 0.58 suicides/year), suggesting that suicides were not transferred to non-station locations. Together, none of these secondary analyses convincingly negated, but rather supported the observed suicide preventive effect of the mid-track fences at the intervention station.

4. Discussion

The mid-track fences at the intervention station had the effect of reducing suicides, as there was a 62.5% decrease in suicides at the intervention station (from 0.66 to 0.25 suicides/year) which sharply interrupted the slightly increasing pre-2014 trend (Fig. 2). This was not the case for suicides at the control stations

Table 1
Summary of suicides and accidents occurring at the stations.

Year	All suicides		Only suicides in the high-speed tracks		All accidents	
	Intervention station	Control group[a]	Intervention station	Control group[a]	Intervention station	Control group[a]
Before mid-track fencing						
2002	0	0	0	0	1[c]	1[c,d]
2003	1	0	0	0	0	0
2004	0	0	0	0	0	0
2005	1	1	0	1	0	0
2006	1	0	1	0	1	0
2007	1	1	1	1	1[c,d]	2[c,d]
2008	0	1	0	1	0	1[c,d]
2009	0	0	0	0	0	0
2010	1	3	1	2	0	0
2011	0	1	0	1	0	0
2012	2	1	1	1	0	0
2013	1	3	0	3	0	0
Total:	8	11	4	10	3	4
After mid-track fencing						
2014	1	1	0	1	0	0
2015	0	4	0	3	1[c]	0
2016	0	4	0	3	0	0
2017	0	2	0	1	0	1
2018	0	2	0	2	1[c,d]	0
2019	1	2	1[b]	2	0	0
2020	0	1	0	1	0	0
2021	0	3	0	1	0	0
Total:	2	19	1	14	2	1

[a] Sums for the six control stations are displayed.
[b] Occurred near a wide breach in the mid-track fence (Fig. S4 in the supplement).
[c] One (1) fatal accident.
[d] One (1) accident in high-speed tracks.

(Fig. 2) or for accidents at any stations (Table 1). Nevertheless, two suicides still occurred during the post-period (years 2014–2021) at the intervention station. One of these occurred in front of a commuter train as the train started to accelerate from the station, which is an unusual behavior (Ceccato et al., 2021) and did not invoke the mid-track fence acting as a physical barrier. The other suicide occurred at the high-speed train track near a wide breach in the mid-track fence (Fig. S4), which likely prevented the functioning of the mid-track fence as a physical barrier for this case. The overall results of this pilot study suggested that mid-track fencing has a preventive effect on deaths by suicide (which occurred most frequently at the high-speed train track).

The decrease in suicides was not followed by an increase in suicides at the train tracks in-between stations, suggesting that suicides were not simply transferred to those locations. However, there was noteworthy 162% increase in suicides among the control stations between the pre- and post-period years, raising the possibility that some suicides had been transferred to the control stations. Results here suggested that mid-track fencing resulted in 3–4 fewer suicides than expected at the intervention station during 2014–2021, which could have contributed, via a transfer effect, to some of the additional 11–12 suicides occurring at control stations during the same period. This indicates the importance of installing mid-track fencing at all stations along the same commuter rail line, rather than only at one station having the most suicides. Future studies of mid-track fencing at further stations and locations will help to resolve the overall and long-term effect on suicide reduction in the wider railway system.

That mid-track fencing may cause a 62.5% (i.e., 2.7-fold) reduction in suicide occurrence may be regarded as surprising, since access to the commuter trains remains the same and the height of the fence is only one meter (a partial physical barrier). We speculate that the mid-track fence may interfere with the cognitive process during the suicidal act *per se*. Indeed, suicidal subjects have been shown to have a number of different cognitive deficits in, for example, decision making and impulsivity (Deisenhammer et al., 2009; Giner et al., 2016; Gvion et al., 2015; Hadlaczky et al., 2018). The results here are in line with a previous study showing that a minimal structured intervention, which only partly restricts access to the lethal means, had the possibility to prevent suicide (Mohl et al., 2012). The time component is of importance in an acute suicidal crisis, as it has been reported that approximately 50% of suicide attempters make their attempt 10 minutes or less after the first current thought of suicide (Deisenhammer et al., 2009). Furthermore, the fence increases the likelihood of being discovered by other individuals, who may in turn intervene by making contact or calling emergency services. Indeed, precautions against discovery was shown to be a predictor of future suicide risk (Beck & Steer, 1989). These putative psychological effects may be shared with mid-platform barriers, which were ranked as having high likelihood to prevent suicide (Restrail-project, 2015b).

There are several limitations with this study. The most severe limitation is that we only investigated one station and have a small sample of suicides in the study. This exposes our findings to the possibility that the effect was caused by other local changes (e.g., other suicide preventive interventions outside of the railway system). But as far as we know there have not been any such interventions (e.g., changes in access to psychiatric care, or other similar activities in the surroundings of the intervention station). The results should thus be interpreted with caution and future studies involving more stations with mid-track fencing will enable to test if our findings are generalizable and provide better estimates of the effect sizes.

Nevertheless, our study suggests a potential new approach to prevent railway suicides. Mid-track fences are likely to be more technologically simple and cost-effective compared to full physical barriers (e.g., full height platform screen doors), although this was not studied here. It could thus be a suitable alternative for railway stations having passing high-speed trains, for example, at the main lines near larger cities where there is a mixture of faster and slower train tracks. In the commuter train network of Stockholm, we estimate that it could be a suitable method for ~20 additional stations (~50% of the stations in the network). However, due to the possibility of transfer effects being induced by the intervention, fencing should also be installed at nearby stations as well. If future studies confirm the preliminary results reported here, we believe that mid-track fencing (preferably having no breaches and if possible, having a higher height; e.g., 1.5–2 meters) should be considered when designing railways where high-speed-trains pass by stations at separate tracks, and likely also at stations where the platforms are located on either side of the tracks.

5. Conclusion

Mid-track fencing to restrict easy access to high-speed tracks has not been evaluated previously in the scientific literature using suicide as main outcome. The results of this pilot study suggest that mid-track fencing may have a high effect on the prevention of suicides at stations in the railway-system. Physical barriers such as mid-track fencing appear to be an effective and low-cost approach to preventing suicides at train stations, although findings suggest it may be important to also have them at nearby stations to prevent transfer effects. We believe that further research on more alternative uses of various types of fencing should be highly prioritized within the railway industry.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgement

The authors would like to thank STA and SL. Especially following people contributing with data and practical help: Iman Arasteh-Khouy, Örjan Eriksson, Arne Grundberg, Britt-Marie Olsson, Dag Lokrantz-Bernitz, Hans Höwitz, Matts Jenelius, Stefan Jonsson, Tobias Lindberg, Anna-Lena Lindström Olsson, Per-Olof Lysell, Mikael Pettersson, Helena Rådbo, Christer Röndum, Anna-Sofia Welander and Anders Wigren. This research has partially been funded by STA.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsr.2022.08.019.

References

Andersson, A.-L., & Sokolowski, M. (2021). Accident or suicide? Improvement in the classification of suicides among road traffic fatalities in Sweden by extended psychosocial investigations, during the years 2010–2019. *Journal of Safety Research*. https://doi.org/10.1016/j.jsr.2021.11.004.
Barker, E., Kolves, K., & De Leo, D. (2017). Rail-suicide prevention: Systematic literature review of evidence-based activities. In *Asia-Pacific Psychiatry* (Vol. 9, No. 3). Wiley-Blackwell. https://doi.org/10.1111/appy.12246.
Beck, A. T., & Steer, R. A. (1989). Clinical predictors of eventual suicide: A 5- to 10-year prospective study of suicide attempters. *Journal of Affective Disorders, 17*(3), 203–209. https://doi.org/10.1016/0165-0327(89)90001-3.
Ceccato, V., & Uittenbogaard, A. (2016). Suicides in commuting railway systems: The case of Stockholm county, Sweden. *Journal of Affective Disorders, 198*, 206–221. https://doi.org/10.1016/j.jad.2016.02.051.
Ceccato, V., Wiebe, D. J., Vrotsou, K., Nyberg, U., & Grundberg, A. (2021). The situational conditions of suicide in transit environments: An analysis using CCTV footage. *Journal of Transport and Health, 20*(December 2020). https://doi.org/10.1016/j.jth.2020.100976.
Chung, Y. W., Kang, S. J., Matsubayashi, T., Sawada, Y., & Ueda, M. (2016). The effectiveness of platform screen doors for the prevention of subway suicides in South Korea. *Journal of Affective Disorders, 194*, 80–83. https://doi.org/10.1016/j.jad.2016.01.026.
Coats, T. J., & Walter, D. P. (1999). Effect of station design on death in the London Underground: Observational study. *British Medical Journal, 319*(7215), 957. https://doi.org/10.1136/bmj.319.7215.957.
Columbia University Mailman School of Public Health (2019). Difference-in-difference estimation. https://www.publichealth.columbia.edu/research/population-health-methods/difference-difference-estimation.
Deisenhammer, E. A., Ing, C. M., Strauss, R., Kemmler, G., Hinterhuber, H., & Weiss, E. M. (2009). The duration of the suicidal process: How much time is left for intervention between consideration and accomplishment of a suicide attempt? *Journal of Clinical Psychiatry, 70*(1), 19–24. https://doi.org/10.4088/JCP.07m03904.
Fredin-Knutzén, J., Andersson, A., Hadlaczky, G., & Sokolowski, M. (2020). Scoping study—Suicide in the transport system. https://ki.se/en/nasp/suicide-in-the-transport-system.
Giner, L., Blasco-Fontecilla, H., De La Vega, D., & Courtet, P. (2016). Cognitive, emotional, temperament, and personality trait correlates of suicidal behavior. *Current Psychiatry Reports, 18*(11). https://doi.org/10.1007/s11920-016-0742-x.
Giupponi, G., Thoma, H., Lamis, D., Forte, A., Pompili, M., Kapfhammer, H. P., Tranah, T., & Farmer, R. D. T. (2019). Posttraumatic stress reactions of underground drivers after suicides by jumping to arriving trains; feasibility of an early stepped care outpatient intervention. *Journal of Trauma and Dissociation, 20*(3), 495–510. https://doi.org/10.1016/0277-9536(94)90448-0.
Gvion, Y., Levi-Belz, Y., Hadlaczky, G., & Apter, A. (2015). On the role of impulsivity and decision-making in suicidal behavior. *World Journal of Psychiatry, 5*(3), 255. https://doi.org/10.5498/wjp.v5.i3.255.
Hadlaczky, G., Hökby, S., Mkrtchian, A., Wasserman, D., Balazs, J., Machín, N., Sarchiapone, M., Sisask, M., & Carli, V. (2018). Decision-making in suicidal behavior: The protective role of loss aversion. *Frontiers Psychiatry, 9*(APR). https://doi.org/10.3389/fpsyt.2018.00116.
Law, C. K., Yip, P. S. F., Chan, W. S. C., Fu, K. W., Wong, P. W. C., & Law, Y. W. (2009). Evaluating the effectiveness of barrier installation for preventing railway suicides in Hong Kong. *Journal of Affective Disorders, 114*(1–3), 254–262. https://doi.org/10.1016/j.jad.2008.07.021.
Mann, J. J., Michel, C. A., & Auerbach, R. P. (2021). Improving Suicide Prevention Through Evidence-Based Strategies: A Systematic Review appi.ajp.2020.2. *American Journal of Psychiatry, 7*. https://doi.org/10.1176/appi.ajp.2020.20060864.
Mishara, B. L., & Bardon, C. (2016). Systematic review of research on railway and urban transit system suicides. *Journal of Affective Disorders, 193*, 215–226. https://doi.org/10.1016/j.jad.2015.12.042.
Mohl, A., Stulz, N., Martin, A., Eigenmann, F., Hepp, U., Hüsler, J., & Beer, J. H. (2012). The 'suicide Guard Rail': A minimal structural intervention in hospitals reduces suicide jumps. *BMC Research Notes, 5*, 2–5. https://doi.org/10.1186/1756-0500-5-408.
National Centre for Suicide Research and Prevention of Mental ill health (NASP) (2020). Självmord i Sverige 'Suicides in Sweden'. https://ki.se/nasp/sjalvmord-i-sverige.
O'Donnell, I., & Farmer, E. D. T. (1994). The epidemiology of suicide on the London underground. *Social Science and Medicine, 38*(3), 409–418. https://doi.org/10.1016/0277-9536(94)90440-5.
Pirkis, J., Too, L. S., Spittal, M. J., Krysinska, K., Robinson, J., & Cheung, Y. T. D. (2015). Interventions to reduce suicides at suicide hotspots: A systematic review and meta-analysis. *The Lancet Psychiatry, 2*(11), 994–1001. https://doi.org/10.1016/S2215-0366(15)00266-7.
Rådbo, H., & Andersson, R. (2012). Patterns of suicide and other trespassing fatalities on state-owned railways in greater Stockholm; implications for prevention. *International Journal of Environmental Research and Public Health, 9*(3), 772–780. https://doi.org/10.3390/ijerph9030772.
Restrail-project (2015a). Restrail toolbox—Intermediate fencing between track. 2015-01-20. http://restrail.eu/toolbox/spip.php?article102.
Restrail-project (2015b). Restrail toolbox—Mid platform fencing. 2015-01-09. http://restrail.eu/toolbox/spip.php?article103#collapse5.
Ryan, B., Kallberg, V. P., Rådbo, H., Havârneanu, G. M., Silla, A., Lukaschek, K., Burkhardt, J. M., Bruyelle, J. L., El-Koursi, E. M., Beurskens, E., & Hedqvist, M. (2018). Collecting evidence from distributed sources to evaluate railway suicide and trespass prevention measures. *Ergonomics, 61*(11), 1433–1453. https://doi.org/10.1080/00140139.2018.1485970.
StataCorp. (2005). *Stata statistical software: Release 9*. College Station, TX: StataCorp LP.
Sueki, H. (2021). Characteristics of train stations where railway suicides have occurred and locations within the stations. *Crisis, 1–6*. https://doi.org/10.1027/0227-5910/a000761.
Tranah, T., & Farmer, R. D. T. (1994). Psychological reactions of drivers to railway suicide. *Social Science and Medicine, 38*(3), 459–469. https://doi.org/10.1016/0277-9536(94)90448-0.
Ueda, M., Sawada, Y., & Matsubayashi, T. (2015). The effectiveness of installing physical barriers for preventing railway suicides and accidents: Evidence from Japan. *Journal of Affective Disorders, 178*, 1–4. https://doi.org/10.1016/j.jad.2015.02.017.
Xing, Y., Lu, J., & Chen, S. (2019). Evaluating the effectiveness of platform screen doors for preventing metro suicides in China. *Journal of Affective Disorders, 253* (February), 63–68. https://doi.org/10.1016/j.jad.2019.04.014.
Zalsman, G., Hawton, K., Wasserman, D., van Heeringen, K., Arensman, E., Sarchiapone, M., Carli, V., Höschl, C., Barzilay, R., Balazs, J., Purebl, G., Kahn, J.

P., Sáiz, P. A., Lipsicas, C. B., Bobes, J., Cozman, D., Hegerl, U., & Zohar, J.. Suicide prevention strategies revisited: 10-year systematic review. *The Lancet Psychiatry, 3*(7), 646–659. https://doi.org/10.1016/S2215-0366(16)30030-X.

Johan Fredin-Knutzén, M.Sc. in psychology, works at the National Centre for Suicide Research and Prevention of Mental Ill-Health (NASP), Karolinska Institute and Stockholm County Council with suicide in the transport system. Is also as a pediatric clinical psychologist.

Gergö Hadlaczky, PhD in psychology, is the head of department at the National Centre for Suicide Research and Prevention of Mental Ill-Health (NASP), Stockholm County Council, and a suicide prevention researcher affiliated to the Karolinska Institute.

Anna-Lena Andersson, PhD, is a special adviser and the psychosocial investigator at the Swedish Transport Administration.

Marcus Sokolowski, PhD, is an associate professor in public health sciences and genetics, whom is conducting various research with a particular focus on suicidal behaviours.

Journal of Safety Research 83 (2022) 8–25

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Applications of wireless sensor networks to improve occupational safety and health in underground mines

Sanaz Sadeghi [a], Nazi Soltanmohammadlou [a,*], Farnad Nasirzadeh [b]

[a] Faculty of Conservation and Restoration, University of Art, Tehran, Iran
[b] School of Architecture and Built Environment, Deakin University, Geelong, VIC 3220, Australia

ARTICLE INFO

Article history:
Received 13 June 2021
Received in revised form 22 December 2021
Accepted 29 July 2022
Available online 8 August 2022

Keywords:
Wireless sensor network (WSN)
Wireless underground sensor network (WUSN)
Occupational safety and health
Underground mine

ABSTRACT

Introduction: The very complex and hazardous environment of underground mines may significantly contribute to occupational fatalities and injuries. Deploying wireless sensor network (WSN) technology has the potential to improve safety and health monitoring of miners and operators. However, the application of WSN in the industry is not fully understood and current research themes in this area are fragmented. Thus, there is a need for a comprehensive review that directly explores the contribution of WSNs to occupational safety and health (OSH) in underground mines. *Method:* This study aims to conduct a systematic literature review on the existing applications of WSNs for improving OSH in the underground mining industry to pinpoint innovative research themes and their main achievements, reveal gaps and shortcomings in the literature, recommend avenues for future scholarly works, and propose potential safety interventions. The major contribution of this review is to provide researchers and practitioners with a holistic understanding of the integration of WSN applications into underground mine safety and health management. *Results:* The review results have been categorized and discussed under three predominant categories including location monitoring and tracking, physiological and body kinematics monitoring, and environmental monitoring. Finally, seven major directions for future research and practical interventions have been identified based on the existing research gaps including: (1) further applications of WSNs for underground mining OSH management; (2) application of WSNs from research to real-world practice; (3) big data analytics and management; (4) deploying multiple WSNs-based monitoring systems; (5) integration of WSNs with other communication systems; (6) adapting WSNs to the Internet of Things (IoT) infrastructure; and (7) autonomous WSNs.

1. Introduction

Occupational Safety and Health (OSH) has become a priority in the workplace in different industries. Due to the dynamic and hazardous working environment of mines, OSH is of great importance for the industry in order to reduce work-related deaths and injuries (Outalha, Le, & Tardif, 2000; Stanek, 1988). Statistics demonstrate the high rate of fatalities and injuries in the mining industry across the world (Moridi, 2015; Nie, 2016; Winn, Biersner, & Morrissey, 1996; Wu, 2011). For example, the Australian mining industry as one of the world's largest producers of minerals had a fatality rate of 2.9 per 100,000 workers in 2019, the fourth-highest fatality rate of any industry with an average of seven workers dying every year (Safe Work Australia, 2020). Additionally, occupational injuries and casualties can result in project delays, cost overruns, and the burden of health disorders (Moridi, 2014; Umer, 2012).

The high complexity and perilous features of underground mine activities have made the everyday monitoring and management of safety and emergent rescue responses very challenging (Akkaş, 2018; Liu, 2010; Qiuping, Shunbing, & Chunquan, 2011). According to previous studies, some of the main risk factors that contribute to lethal and non-lethal accidents involved in mine operations are poor visibility and unsafe behavior of miners as well as jobsite circumstances and environmental conditions (e.g., poor lighting, poor ventilation, gas emission, wet conditions, confined space, rock falls, structural failures and complexity, communication restrictions) (Amponsah-Tawiah & Dartey-Baah, 2011; Gyekye, 2003; Mahdevari, Shahriar, & Esfahanipour, 2014; Moridi, 2014; Sun, 2010). To protect the safety and health of the underground workforce, it is crucial for safety managers to be able to continuously monitor the physiological status of workers and the environmental

* Corresponding author at: 56, Sakhai St., Hafez Ave., Tehran 1136813518, Iran.
E-mail addresses: snzsadeghi.1988@gmail.com (S. Sadeghi), n.soltanmohamadlou@gmail.com (N. Soltanmohammadlou), farnad.nasirzadeh@deakin.edu.au (F. Nasirzadeh).

https://doi.org/10.1016/j.jsr.2022.07.016
0022-4375/

parameters of their work zones (Akkaş, 2018; MSHA, 2011; Sunderman & Waynert, 2012). This necessitates the collection of large amounts of data in real-time. Due to the dynamic and complex working environment of underground mines, the traditional mine safety system is insufficient (Maity et al.; Reyes, 2014; Zeng, Tam, & Tam, 2008), and utilizing advanced technologies to improve OSH in underground mines is required (Bo, 2014; Sunderman & Waynert, 2012). Although some mine safety technologies are currently commercially available and some are under investigation, the research in this area is not mature yet (Misra, 2010; Sun, 2010; Sunderman & Waynert, 2012).

Wireless Sensor Network (WSN) technology has been increasingly adopted in recent years to monitor environmental parameters of underground mines and trace the location and health conditions of miners. A WSN involves a group of sensor nodes arrayed compactly that are capable of data sensing, processing, and communication (Ranjan, Sahu, & Misra, 2016). Compared to wired sensor networks, WSNs offer benefits such as small size, low cost, improved coverage, and operational simplicity in the restricted and harsh environment of underground mines (Chehri, 2010). With the development of WSNs, an increasing number of researchers and practitioners have agreed on their applications to achieving reliable solutions to enhance OSH in the mining industry (Akkaş, 2018; Chehri, 2011; Henriques & Malekian, 2016; Milenković, Otto, & Jovanov, 2006; Ranjan et al., 2016). This technology is capable of online monitoring of underground mine working environment/condition and organizing the collected sensed data at a central station aboveground (Ranjan et al., 2016) to get an insight into the workplace safety and health conditions and possibility of accidents on a real-time basis (Soltanmohammadlou, 2019). Wireless underground sensor networks have emerged as a feasible technology for constant monitoring of underground mines (Akyildiz & Stuntebeck, 2006), which ultimately contributes to the reduction of fatal and nonfatal occupational injuries in the industry.

Existing studies have illustrated the significance of WSN technology for OSH in underground mines. However, current research themes in this area seem to be fragmented and a holistic approach is required to integrate this emerging technology into wider underground mine safety and health management. Recently, two review papers have been developed by Muduli, Mishra, and Jana (2018) and Dohare (2015) on the use of WSNs in underground coal mines. However, the theme of these articles has not addressed WSN applications specifically for underground mining safety and health management. As they primarily focus on environmental monitoring and do not cover different applications of WSN-based monitoring systems in underground mining, including location monitoring and tracking and physiological monitoring. Therefore, a comprehensive review of the applications of WSNs for improving OSH in underground mines should be conducted. The present review will investigate how WSNs have been or can be applied to successfully monitor and manage different causes of accidents and injuries involved in the underground mining industry. This study aims to conduct a systematic literature review on the existing trends and applications of WSNs technology to enhance OSH in the underground mining industry. The objectives are to share innovative research themes and their main achievements, reveal gaps and shortcomings in the current literature, recommend directions for future scholarly works, and offer potential safety interventions. This review paper provides researchers and practitioners with a deep insight into the various aiding functions of WSNs technology to improve OSH in underground mines.

2. Research method

Considering the aim and objectives of the present study, a systematic literature review (SLR) method was employed, which facilitates the recognition, evaluation, and interpretation of existing sources of literature in a specific area of study, through scrutinizing and systemizing key concepts and current evidence (Rowley & Slack, 2004). Among various methods for conducting a systematic literature review, this paper opts for taking a realist approach in order to clarify how WSNs have been or can be applied for improving OSH in underground mines. It is mainly because the realist review method is devised to present an explanatory analysis of what works for what, under what circumstances, in what respects and how, by indicating processes/interventions that lead to effects in complex contexts (Golizadeh, 2018; Pawson, 2005). According to the model designed by Pawson (2005), the realist review consists of five distinct stages, which are followed in this research including scoping, searching, screening, data extraction and synthesis, and reporting. At the primary stage, the research scope is clarified and the review question(s) are formulated; the second stage entails searching databases for an initial list of studies that contain defined search keywords; at the third stage, the most relevant studies are selected based on a set of defined eligibility criteria; the fourth stage aims to extract data from full text of selected studies; at the last stage, the main findings are presented and discussed. The review process is summarized in Fig. 1.

2.1. Search strategy

In this review, the search string consisted of the following four categories of keywords:

- 'wireless sensor network' OR 'wireless underground sensor network' OR 'WSN' OR 'WUSN' OR 'wearable' OR 'wearable sensor'; AND
- 'occupational' OR 'safety' OR 'health' OR 'accident' OR 'incident'; AND
- 'underground'; AND
- 'mine' OR 'miner' OR 'mining'.

Using the keywords, the search was conducted in electronic data-bases including Elsevier (sciencedirect.com), Springer (springerlink.com), IEEE Xplore Digital Library (ieeexplore.ieee.org), Taylor & Francis (T&F) (tandfonline.com), Wiley (onlinelibrary.wiley.com), Emerald (emeraldinsight.com), and ACM Digital Library (dl.acm.org). Then, we extracted each peer-reviewed article published by a reputable journal or international conference, which contained the aforementioned search terms in its title/abstract/keywords. The eligibility period for the inclusion of articles was set by the end of July 2020 when the search was conducted.

2.2. Selection of relevant studies

The screening process involved three phases. In the first phase, duplicates, non-English papers, ineligible article types (e.g., editorials, editor's notes, book reviews) and review articles were removed from the initial publication list. In the second phase, content analysis was carried out on the title and abstract of each article to make sure that it was empirical with a substantive focus on the review questions; so that the article covered at least one application of WSN technology for improving OSH in underground mining. Thus, papers that contained some search terms in their titles or abstracts, but were irrelevant to the scope of this review were excluded by the authors. For instance, articles referring to the use of WSNs for OSH in surface mining were removed. In the last phase, full-text screening was conducted to extract the most relevant publications. This phase was accomplished to make sure that the full text of the selected articles fully addressed the issues raised in their abstracts. Moreover, the reference lists of the selected articles were also considered to ensure the completeness of the

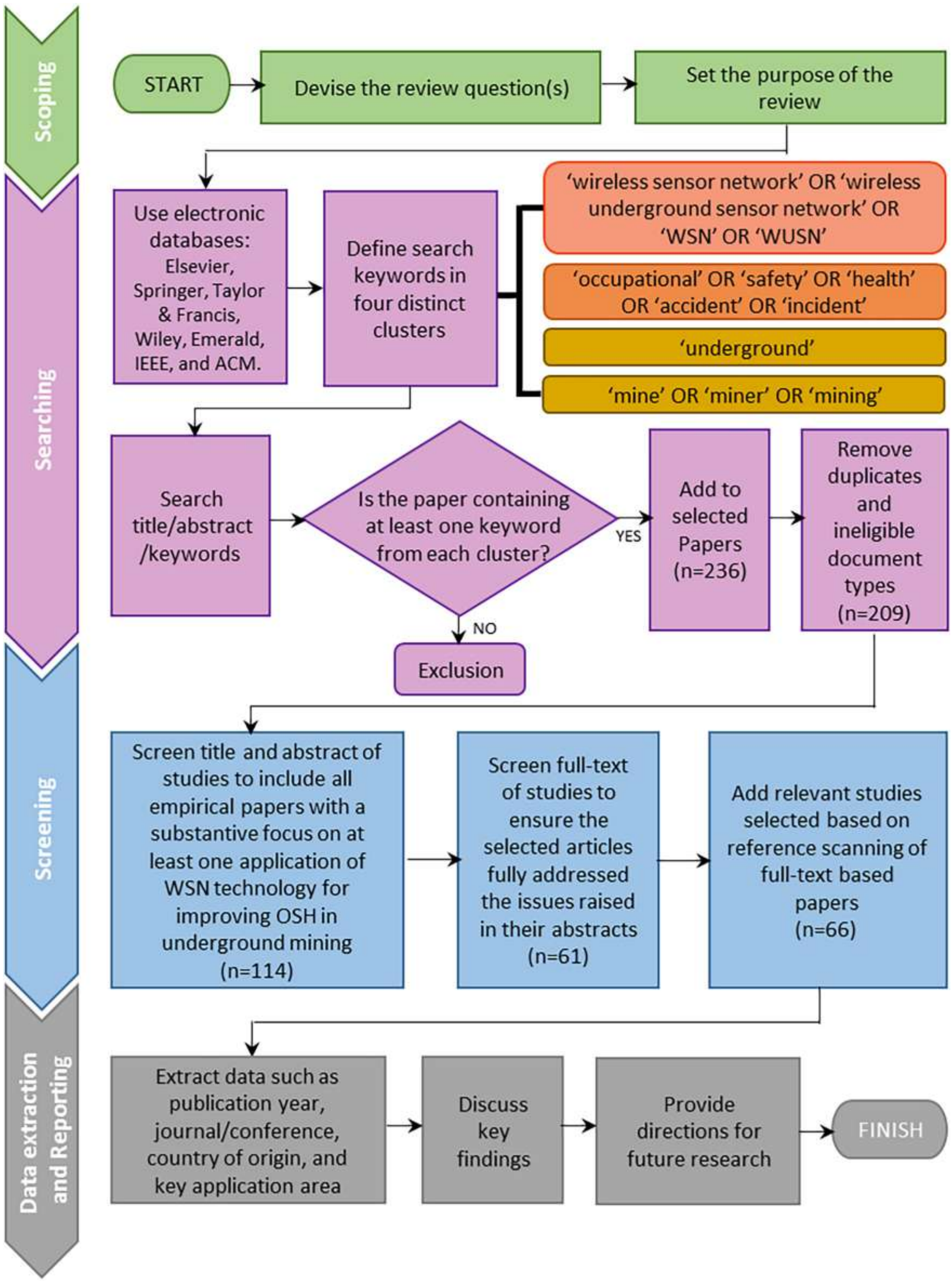

Fig. 1. Systematic literature review stages.

review. In total, 66 articles were shortlisted at this stage of our systematic literature review.

3. Findings

The selected articles were coded and the key data were extracted and tabulated in tables and graphs as shown in the following subsections.

3.1. Publication distribution by year

As illustrated by Fig. 2, the subject area has been of constant interest to researchers since 2005. The distribution of published papers fluctuates from 2005 to July 2020. The highest rate of relevant articles published in a year belongs to 2018 with nine papers, which is followed by 2009 and 2019 with eight and seven papers, respectively.

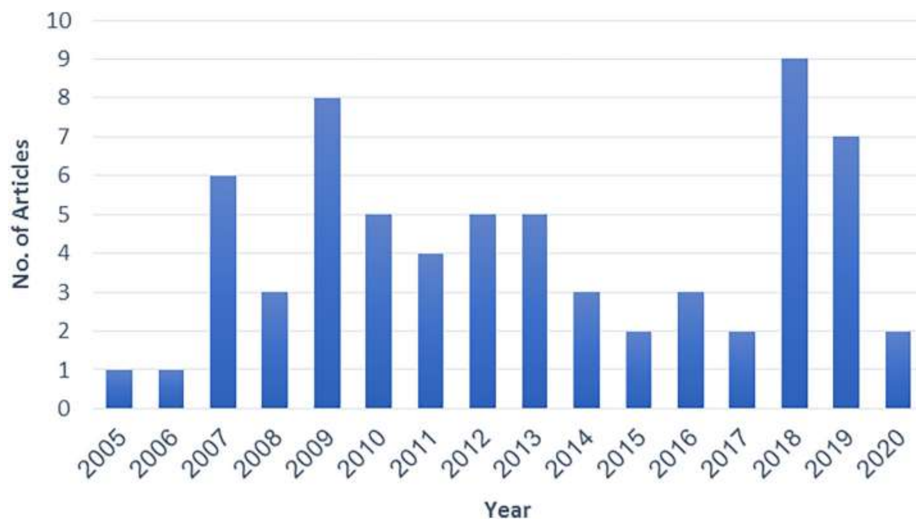

Fig. 2. Publication distribution by year.

3.2. Publication distribution by database

According to Table 1, the total number of selected articles consists of 34 journal papers and 32 conference papers. Above 35% of journal papers and 84% of conference papers were published by Elsevier and IEEE Xplore Digital Library, respectively. In total, IEEE Xplore Digital Library, Elsevier and Springer belong the largest number of published articles in this area with 33, 12 and 9 papers, respectively.

3.3. Publication distribution by journal

Five papers out of 34 journal articles were published by the International Journal of Distributed Sensor Networks as shown in Table 2. Safety Science, IEEE Access, IEEE Sensors Journal, and Sensors each published two articles. Only one paper was published by each of the remaining 21 journals in the list.

3.4. Geographical distribution of publications

The origin of the selected publications was determined based on the location of the organization the authors are affiliated with. Only one country is decided for each article. Among these countries, China played the leading role in this area of study with 32 articles. India and Australia follow China, with 17 and 6 articles, respectively. These three countries are among the largest mineral producers in the world (Blondeel & Van de Graaf, 2018; Fong-Sam, 2008). The number of hits in the other countries listed in Fig. 3 is three or less, which shows that they may require additional research boost to fill the current gap between themselves and the leading countries.

Table 1
Publication distribution by electronic database.

Database	Journal papers	Conference papers	Total
IEEE	6	27	33
Elsevier	12	0	12
Springer	5	4	9
Taylor & Francis	5	0	5
Others	5	0	5
ACM	1	1	2
Total	34	32	66

Table 2
Publication distribution by journal.

Item.	Journal publication	Number of articles
1.	International Journal of Distributed Sensor Networks	5
2.	Safety Science	2
3.	IEEE Access	2
4.	IEEE Sensors Journal	2
5.	Sensors	2
6.	Procedia Engineering	1
7.	American Journal of Operations Management and Information Systems	1
8.	IEEE Communications Magazine	1
9.	Ad Hoc Networks	1
10.	Journal of The Institution of Engineers (India)	1
11.	IEEE Systems Journal	1
12.	Eurasip Journal on Wireless Communications and Networking	1
13.	Physics Procedia	1
14.	Computer Networks	1
15.	Wireless Networks	1
16.	Computer Standards & Interfaces	1
17.	Middle-East Journal of Scientific Research (MEJSR)	1
18.	Wireless Personal Communications	1
19.	International Journal of Mining Science and Technology	1
20.	Tunnelling and Underground Space Technology	1
21.	International Journal of Future Generation Communication and Networking	1
22.	Process Safety and Environmental Protection	1
23.	ACM Transactions on Sensor Networks (TOSN)	1
24.	Measurement	1
25.	International Journal of Coal Science & Technology	1
26.	Journal of Systems and Software	1
Total		34

3.5. Publication distribution by key application areas

Taking the systems theory into account, any circumstance under which an accident might happen at a workplace is a system composed of three components, including environment, human, and machine (Li & Guldenmund, 2018). With this said, to investigate WSN potential in addressing OSH hazards in underground mining, this research considered three main elements of this complex system: working environment, work team, and physical equipment (Dhillon, 2010). Adopting this theoretical lens, WSN technology can apply to improve OSH in underground mines

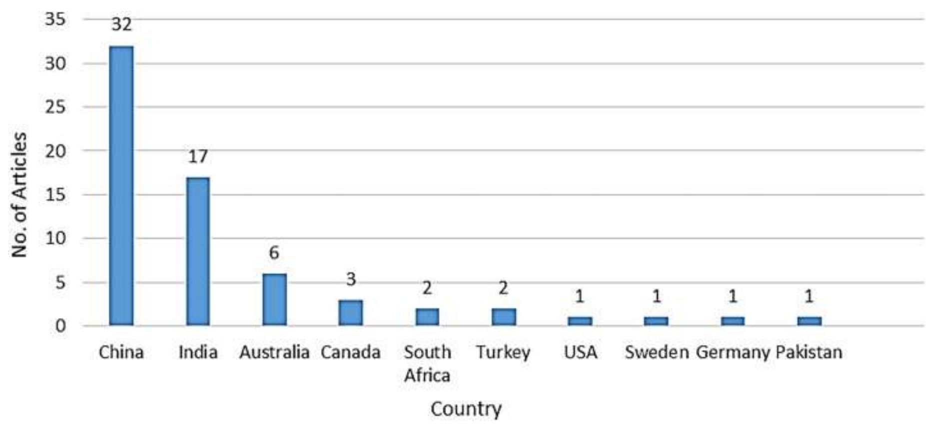

Fig. 3. Geographical distribution of publications.

through location monitoring and tracking of both equipment and workforce, physiological and body kinematics monitoring of on-site personnel, and environmental monitoring. Fig. 4 displays the publication distribution regarding their stance on the applications of WSN technology for improving OSH in underground mining. More than half of the articles addressed environmental safety and health issues, 39% of articles covered location monitoring and tracking of mining equipment and workforce, and 8% of papers dealt with physiological health risks and hazards. These key application areas are discussed in the following section.

4. Discussion

The following sections summarize the key findings from the SLR.

4.1. Location monitoring and tracking

Underground mine safety management systems can be enriched by fusing real-time location monitoring and tracking data of mining assets (i.e., workers and equipment). This facilitates recognition of interference between various tasks in complex job sites, workers' unsafe behaviors, and unauthorized activities or entry of workforces inside a pre-defined risk zone of equipment. It also provides basic information for detection and visualization of labors in dangerous situations and promotion of equipment operators' visibility on-site (Soltanmohammadlou, 2019).

Location and tracking technologies can be attached to a wireless network to get extra data from the working environment, enabling further evaluation on safety, particularly in high-risk workplaces (Soltanmohammadlou, 2019). WSN is an emerging and self-organized technology composed of a large number of small sensor nodes that are capable of collecting, processing, and transmitting information about recognized objects in the realm of the area where it is monitored by the network (Baek, 2017; Karl & Willig, 2007; Li & Liu, 2007). The invulnerable nature and other characteristics of WSN, including self-management, quick set up nature, multipath routing and dynamic topology, make it a very effective and efficient technology for collecting and communicating location monitoring data in complex underground environments (Chehri, 2010; Misra, 2010; Wu, 2019). Thus, a location monitoring and tracking system with WSN offers more flexible, adaptable, and reliable network communication for underground mine environments compared with wired and fixed monitoring tools (Swain, 2018; Wang, 2007, 2011). To contribute with safety monitoring on site, earlier research in this scope used WSN based on a real-time system to continuously track the location and position of either underground workforces or equipment (Chen & Zhao, 2007).

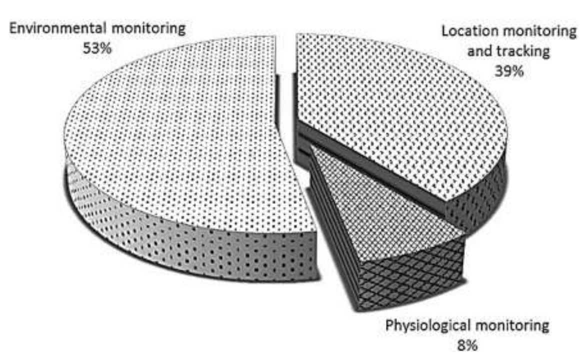

Fig. 4. Publication distribution by key application areas.

4.1.1. Improving transport safety in underground mines

Underground mine transportation represents one of the highest categories of all accidents, especially in some fast-developing countries such as China and South Africa (Jiang, 2009; Rupprecht, 2011). Therefore, constant monitoring of mine transport is very important in assessing the safety of mine tunnels, and hence safety of workforces. This requires localization and tracking of vehicles as well as monitoring the status of exploiting paths. In this regard, Jiang (2009) targeted establishing transport safety in underground mines by proposing a wireless video sensor network platform to be implemented in underground mine tunnels for effective remote safety monitoring. Due to the possibility of multipath effect and path loss in WSN, a braided cooperative reliable transport (BCRT) algorithm was developed that robustly operated in frequent path break and repair environments. Thus, a reliable video transmission within mine tunnels was provided by this system.

4.1.2. Estimating position of workforces in blind areas

When establishing a WSN system in an underground mine, sensor nodes are only placed in the key spots within the underground workings; so, the complex underground network involves many "blind areas" of monitoring and early warning. This way, the central ground station can only get the location information of the underground personnel located in the positioning areas, while their global positions within the blind areas need to be numerically estimated throughout the underground network. By estimating the position of workforces in blind zones and communicating the data to supervisors, the risk of accidents can be more adequately appreciated (Chen, 2020; Liu, 2010). To this end, a low-cost positioning system for estimating coal miners' location in blind areas was proposed by Liu (2010) based on WSN using global positioning technology. This system gathered real-time location data of miners through sensors located in the tunnels' key spots, sent this information to the central computer placed on the surface, corrected errors and faults of the positioning data due to mine environmental interference and, finally, estimated the global three-dimensional location of workers of the network in blind zones numerically. Furthermore, a recent novel approach was used byChen (2020) to provide coal miners working in blind areas with accurate early warnings. They presented an intelligent safety monitoring and early warning system based on two improved Distance Vector-Hop (DV-Hop) localization algorithms that enhance the location accuracy for randomly distributed WSNs. These algorithms are capable of precisely positioning and tracking in underground environments through picking multiple anchor nodes and calculating the average per-hop distance between the deployed sensor nodes.

4.1.3. Improving rescue operations in underground mines

In the case of a disaster in an underground mine, it is very challenging for mine management to detect injured or trapped miners along with their actual location and exact numbers (Bandyopadhyay, Chaulya, & Mishra, 2010). This issue delays emergency rescue operations for the affected miners and will ultimately adversely affect the safety of miners. Thus, location monitoring data are important for localizing miners, assessing damages, and attempting urgent rescue (Wang, Huang, & Yang, 2010; Wu, 2019). In this regard, in a novel approach followed by Qinghua (2009); an automatic underground personnel position tracking system was developed that used radio frequency identification (RFID)-based WSN enriched by Geographic Information System (GIS). The real-time location of every moving object was captured by RFID tags and displayed on virtual maps with GIS technology. Therefore, this system provides an efficient safety management information system for underground coal mines that facilitates precise and timely detection and rescue of endangered and trapped underground coal miners. However, some scholars declared that GPS is not a suitable location tracking technology for underground mines due to the unavailability of its needed signals in indoor places (Kumar, 2016; Mardonova & Choi, 2018). So, another approach followed by Bandyopadhyay et al. (2010), developed a novel wireless information and safety system for underground coal mines. The core system is comprised of ZigBee compliant active RFID devices. The RFID tags were programmed to act as end devices, while the routers were applied as coordinator devices to create an IEEE 802.15.4-based mesh network. Using the RFID ZigBee sensor nodes involved in the wireless mesh network, this system created the possibility to track the location and motion of mine workers and equipment, monitor environmental conditions, fatal incidents, and struck-by accidents in coal mines. In addition, the system could preserve a database consisting of the computerized records of miners' working hours to provide the ground control center with a warning message in emergency situations, which determined the location and numbers of trapped mine personnel.

In the work of Wang et al. (2010), the authors proposed a prototype system for real-time localization and tracking of underground coal miners based on wireless self-organized sensor networks. Self-organization is the ability of sensor nodes to build a network topology without the need for human intervention and prior topology knowledge. Using this technique, WSNs can be established in the harsh and extreme environment of underground mines where on-site technical service is infeasible (Diaz, Mendez, & Kraemer, 2019). In order to improve the environmental adaptability, robustness, and suitability of this system, the authors developed three leading localization technologies. Firstly, a received signal strength indication- (RSSI-) based localization algorithm was proposed to decrease the impact of the severe and complex environment in coal mines. RSSI-based localization algorithm is one of the representatives of range-based localization algorithms that utilizes the received signal power measurements to estimate the distance between two nodes, independent of the environment (Huang, 2015). Secondly, a fault-tolerant localization mechanism was proposed to improve the inherent defect of instability of RSSI localization. Finally, an accurate localization algorithm based on Monte Carlo Localization (MCL) was proposed to adapt to the underground tunnel structure. Localization algorithms based on MCL are frequently used in mobile target positioning (Thrun, 2001) and are suitable for node positioning in underground mobile networks (Baggio & Langendoen, 2008). Their proposed system can undertake four main functions, including the dynamic display of each miner's position, the moving path and working status in real-time, recording the attendance information of every miner, sending an alarm to warn the corresponding manager about a miner entering a designated forbidden area, and giving the location data of any event or disaster and every trapped miner to help emergent rescuing. Moreover, Minhas (2017) proposed a comprehensive WSN-based control and monitoring mechanism, which is also capable of detecting and identifying events and localization of endangered miners. This system covered three design considerations of WSN in underground mines, including radiofrequency propagation modeling, energy-efficient communication protocol, and autonomous event detection and reporting. Another WSN-based safety monitoring system to detect and warn of different events in underground coal mines was presented by Bo (2012). They developed a complex event detection algorithm based on the state automata/machine model to detect the alarming events and generate a corresponding disposal process in real-time. Their proposed algorithm can predict the event occurrences, called states, from the raw data sequence pattern acquired from the underground WSN, through matching the event pattern with the predefined rules from the continuous events streaming. In the paper presented by Wu (2019); WSN was adapted to the Internet of Things (IoT) infrastructure to provide a novel technique for safety to mine activities through establishing a dynamic information platform. This platform contains six practical layers, including the supporting layer, perception layer, WSN-based transmission layer, service layer, data extraction layer, and application layer. The platform is able to monitor and register the information related to working conditions in coal mine production systems as well as position data of underground equipment and personnel. Then, the big data are quickly and precisely analyzed in order to present a three-dimensional computer-generated mine system, safety identification system, safety inspection system, and urgent rescue system for coal mines. Moreover, IoT technology can be applied to the personnel clothing and helmets carrying wireless environment, location, and motion sensors for having effective communication with other employees and equipment in different parts of a mining site in order to make real-time decisions for a high level of safety (Dehran, Agrawal, & Midha, 2018).

4.1.4. Enhancing the localization accuracy and performance of off-the-shelf tracking sensors

Localization accuracy refers to the success rate of a positioning system in identifying the precise areas at which the target is located. The performance of WSN-based safety management systems is highly influenced by the accuracy of the positioning data because these systems use the output information for further applications (Li, Cheng, & Chen, 2020). Enhancing the localization accuracy and performance of a set of commercially available electronic miner tracking sensors (i.e., RFID, UWB and Zigbee), which can communicate wirelessly in a network (Sunderman & Waynert, 2012), was discussed in some previous studies. Kloos

(2006) proposed the application of a range-based localization system using radio frequency (RF) signals for detection of underground mine personnel working in close proximity to heavy equipment such as haul trucks. While such range-based solutions could provide more accurate localization, their implementation imposes higher costs and energy usage. Therefore, a novel range-free localization scheme was presented by Guo, Jiang, and Zhang (2012) in which underground miners were monitored using duty-cycled sensor nodes distributed in a wireless linear network. The system obtains a low-cost, low-energy, and reliable localization system for safety monitoring of underground coal mines. Moreover, Zigbee-based wireless networks were offered by Maity et al. and other similar studies (Xu, 2012; Yin, 2011) for a cost-effective, flexible, constant monitoring of underground miners' safety and security. The system developed by Maity et al. was capable of automatically capturing the chain of site information through a digital wireless communication system to promote meticulousness, soft control, and reliability of the position monitoring process. The mine labors were then informed by various alarming as well as voice systems produced after transmission of data to the central control computer system on the ground.

The studies presented by Aktaş (2019) and Chehri, Fortier, and Tardif (2009) proposed wireless infrastructures using Ultra-wideband (UWB) sensors to collect positioning data of miners and Wi-Fi based communication modules to help send all the collected data to the main server. They tried to hinder the probability of work-related incidents through safety monitoring of workers on-site and delivering immediate feedback from managers. Besides, in these investigations, the accuracy and pace of localization were improved by designing location control algorithms. In the proposed algorithms, it was assumed that the precise position of sensor nodes was already known, which is not necessarily applicable to real-life situations in which sensor nodes are randomly positioned and/or may be dropped. Therefore, a novel WSN-based algorithm was presented by Savic, Wymeersch, and Larsson (2013) for simultaneous improvement of sensor localization and target tracking in the underground tunnels. The algorithm applied the discrete variants of real-time belief propagation to estimate the position of sensor nodes and mobile targets as well as handling all the non-Gaussian uncertainty distributions within mine tunnels. Also, increasing the reliability and robustness of a newly deployed WSN system for coal mine monitoring and positioning of miners was the main parameter that Yan, Ya-ru, and Yong (2008) took into account in their work. The approach followed by them was applying WSN to the mine personnel body to automatically monitor the real-time position of miners.

4.2. Physiological and body kinematics monitoring

Inexpensive, small, lightweight and smart wearable sensor nodes have been recently designed to improve workplace health and safety of mining staff. Generally, there are two different types of wearable sensors, including kinematic and physiological sensors, which can be used to collect data from miners for safety and health applications. The first type of sensors (i.e., kinematic sensors) include wearable inertial measurement units (IMUs) that collect workers' dynamic motions. IMUs will enable users to monitor the kinematic movement of objects by providing the velocity, displacement, and orientation data. On the other hand, wearable physiological sensors provide data about human health measures such as heart rate, respiratory rate, blood oxygen, blood pressure, and body temperature (Awolusi, Marks, & Hallowell, 2018) for further health and safety analysis of workers individually. The application of wearable sensors for physiological monitoring and body kinematics monitoring provides an avenue for ease of creating personalized health care.

Using a WSN platform, the wearable sensor nodes can be interconnected to each other and a statically placed station. Health monitoring systems of this type also allow for preventive diagnosis of miners' health issues and timely rescue or medical responses (Milenković et al., 2006). The applications of wearable wireless monitoring systems in the mining industry are reviewed below:

4.2.1. Using physiological sensors for timely rescue responses and monitoring of emergency state

Wearable physiological sensors can be used to improve the underground mine rescue system and detect hazardous conditions. In the previous studies, physiological responses of workers such as heart rate, respiratory rate, and body temperature have been monitored to assess the health status of workers after underground incidents. In one of these attempts towards an integrated health sensing platform, Ranjan (2019) explored the feasibility of an integrated wearable sensor unit equipped with different biosensors (e.g., temperature sensor, blood pressure sensor, heart rate monitoring sensor). The sensed physiological parameters of each miner were then relayed to the central base station via radio frequency-based WSN for data visualization and analytics. The information of this type improves emergent rescue efficiency. Also, Cicioğlu and Çalhan (2019) proposed a wireless network in which each miner was equipped with a variety of heterogeneous sensor nodes and one coordinator. The coordinator nodes were able to send the monitored data of vital signs of underground workers (e.g., heart rate, respiration rate, EEG, ECG, blood pressure, temperature, and glucose level) to the other coordinators wirelessly as long as they were in range. In this way, miners communicated with each other in the network based on a peer-to-peer framework. When a hazardous condition was detected by one of the sensor nodes, the node, with the help of its coordinator node, used the Ad hoc On-Demand Distance Vector (AODV) routing algorithm to determine the most appropriate route to the destination and to transmit the required data through this route. This novel algorithm is intended for use by mobile nodes in a wireless ad hoc network -a decentralized type of wireless network- and enables routing with continuously changing topologies (Perkins & Royer, 1999). Thus, the system can quickly and accurately deliver the relevant information about the mineworkers to the monitoring center to start rescue responses.

Moreover, wearable health monitoring devices have the potential to be equipped with additional sensors such as position- and location-tracking sensors and environmental sensors. This promising design approach would reduce the implementation and preservation cost of independent monitoring infrastructures without any interference with the routine work tasks of the miners (Ranjan, 2019; Schalk Wilhelm & Reza, 2019). The physiological sensors combined with the location monitoring system and environmental monitoring system will also enable better-informed rescue decisions in emergency situations. Deploying a wearable sensor network that is combined with a locating system has been targeted by Yan et al. (2008) and Yong (2009). The prototypes developed by these studies were able to detect not only personnel position using location monitoring system, but also their health condition when an accident happened underground through capturing important physiological data such as heart rate and body temperature. Being provided with enhanced positioning information, supervisors were able to make the most rational rescue decisions. Besides, a similar approach was also taken by Nie (2011) for smart health monitoring. In this study, positioning of coal miners and monitoring of emergency state were accomplished through IoT-based wireless communications, sensing, and computation. The network used Zigbee protocol and portable wireless sensors embedded in the miners' cap lamp to realize underground personnel location and physiological status (e.g., heart rate, blood pressure and body temperature) to provide timely rescue responses. Such a system can be further augmented through testing the feasibility of other technologies to be embedded in the IoT network (i.e., accompanying sensors, electronic devices, actuators, software, communication sensor modules, network connectivity solutions, etc.). For instance, Revindran, Vijayaraghavan, and Huang (2018) proposed a WSN-based Distance-vector routing (DVR) protocol for finding the shortest route between miner's helmet sending a distress message to the health and safety team to decrease delays. Each mobile node operates as a specialized router (Perkins & Royer, 1999) containing the distance between itself and all possible destination nodes. By using the DVR protocols, each router over internetwork sends the neighboring routers, the information about the destination that it knows how to reach and maintains a list of all destinations that only contains the cost of getting to that destination, and the next node to send the messages to Barma (2014). In this protocol, each router is also required to inform its neighboring routers of topology changes periodically (Mannan & Jayavignesh, 2016). On the other hand, in one of the recent efforts, both location monitoring and environmental monitoring systems were integrated into the physiological monitoring system. The research team integrated physiological, acceleration, location tracking, and environmental sensors to design a smart health monitoring system based on context awareness. Their proposed device was attached to the hard-hat of a miner to sense and capture their contextual data. The collected information was then disseminated amongst other miners' devices and a ground base station using a low-power wireless mesh network. The main part of this research was devoted to the evaluation and selection of feasible and suitable technologies (Schalk Wilhelm & Reza, 2019). Taking this approach, risk assessment and awareness are offered to underground mining staff at three levels including: (1) personal or body risk awareness; (2) risk awareness in the working zone; and (3) risk awareness of neighboring zones; this enables a greater monitoring and communication network and more efficient rescue responses (Mardonova & Choi, 2018). In another approach followed by Adjiski (2019), wireless integrated physiological and environmental sensor nodes including heart rate, body temperature, accelerometer, gyroscope, magnetometer, sound, smoke, dust and gas sensors were attached to the miner's personal protective equipment (PPE) (i.e., helmet, safety glasses or smartwatch). The system was then connected to the miners' smartphone via Bluetooth low-energy. This way, the sensed data were automatically arranged, locally stored, and viewed by a mobile App. The smartphone could also enable the identification of the location of each miner. This prototype network was capable of enhancing the wearer's situation awareness. Once emergency conditions were detected, the smartphone app warned miners of danger by sending them timely audible and visual alerts. The information could then be exchanged with the monitoring center through Wi-Fi technology to notify supervisors of potential risks threatening every single miner. The smartwatch also offered similar capabilities to the mobile phone in terms of warnings and supervision. A similar study suggested for sending the notifications from smartphone to smart eyewear of mine employees over Bluetooth. A visual notification was displayed to the miner wearing smart glasses when being exposed to an emergency state. This system sends safety alerts and emergency rescue messages to miners' smart glasses to assist those who have difficulty with loud noise and low-visibility conditions in their working zone (Mardonova & Choi, 2018).

4.2.2. Using physiological sensors for fatigue monitoring and control

Mine labors are exposed to both physical and mental fatigue in their workplace, which threatens their health and wellness (Butlewski, 2015). Underground miners have to work in an extreme environment that is by no means favorable for prolonged physically demanding workloads. Also, many of them are in charge of monotonous and repetitive tasks. As a result, severe muscle fatigue and stress are very common among this working group (Bauerle, Dugdale, & Poplin, 2018). Moreover, mental fatigue can be caused as a result of brain over-activity, especially when there is a long gap between rest times of mine workers. Shift work is known as a major contributor to mental fatigue among miners (Yu, Chen, & Long, 2017). Due to the decline of concentration ability, this disorder makes work-related unsafe behaviors and wrong actions more likely (Chen, 2021). Constant measurement of the individual's health parameters such as heart rate, respiration rate, and blood pressure can assist in timely detection of fatigue state, preventing the associated risks, and optimizing the underground working environment (Chen, 2020; Meng, 2014). However, to date, the application of wearable wireless physiological sensor systems for real-time monitoring and early warning of underground miners' fatigue is an overlooked area of study.

4.2.3. Using physiological sensors for stress monitoring and control

People working in the mining sector can be exposed to unanticipated disasters and accidents (e.g., being trapped in a confined space), which can result in severe stress reactions (Nie, 2016). Job strain exposure is another risk factor contributing to job stress problems among miners (Hodgskiss & Edwards, 2013). These stressors can bring quick and complex physiological changes to the victims, which potentially lead to many health problems (Nie, 2016; Varga, 2016). Physiological and mental responses to stress can undermine the victims' power for accurate decision-making and performance (Keitel, 2011; Kowalski-Trakofler, Vaught, & Scharf, 2003). Therefore, the use of wearable wireless physiological sensor systems for stress monitoring and control is an open area for further research in the mining industry, which helps promote employees' wellness and well-being.

4.2.4. Using kinematic sensors for mitigation of musculoskeletal disorders

Laborers in mining sites are faced with many physically demanding and manual tasks. They are at high risk of musculoskeletal disorders due to the nature of their physical activities. Whole-body vibration and awkward posture are among the main cause of musculoskeletal disorders (Raffler, 2018). An awkward posture is resulted from an excessive twist or bend of body joints beyond (Wang, Dai, & Ning, 2015). Correcting a bad posture can mitigate the risk of musculoskeletal disorders. Tracking IMU sensor data can provide information about the physical posture and motions of workers while working through capturing real-time data. This will enable safety supervisors to check whether or not the physical pose of a worker on-site is physiologically safe (Liu, Han, & Lee, 2016; Ray & Teizer, 2012). Moreover, the vibration of the whole body is also one of the main risk factors of low back pain for heavy equipment operators in mine activities, contributing to muscle fatigue, intervertebral discs, early spinal degeneration, and herniated lumbar disc (Blood, Ploger, & Johnson, 2010; Boshuizen, Bongers, & Hulshof, 1992; Bovenzi, 2009; Chen, 2003; Sadeghi, Soltanmohammadlou, & Rahnamayiezekavat, 2021; Xu, 2017; Yassierli, 2017). Despite the importance of the issue, to date, the applications of wearable wireless kinematic sensor systems for real-time monitoring and early detection and warning of underground miners' awkward posture and whole-body vibration are the overlooked areas of study.

4.3. Environmental monitoring

Continuous, precise, and reliable monitoring of environmental conditions such as temperature, gas, noise, and dust is an urgent need to achieve better OSH in underground mines. Many calamitous mine accidents have been rooted in out-of-date monitoring and warning system for environmental parameters. With this said, WSNs apply for evaluating, monitoring, and recording various environmental parameters in real-time (Mardonova & Choi, 2018).

4.3.1. Gas

Emission of toxic gas and concentration of non-toxic but explosive gas in the underground mine environment are two major causes of fatalities among miners (Niu, 2007; Osunmakinde, 2013). These gases are colorless, odorless, and cannot easily be detected by human senses (Hazarika, 2016). To perform remote monitoring of the poisonous gas and safety status of underground coal miners, Osunmakinde (Osunmakinde, 2013) developed a reactive autonomous WSN. The proposed system used ambient intelligence for real-time decision-making. The framework deployed by Niu (2007) was a WSN-based distributed heterogeneous hierarchical safety system for monitoring the methane concentration as well as tracking the location of underground miners. This prototype system applied a novel overhearing-based adaptive data collecting scheme to remove the correlation and redundancy of the sampling sensor readings in terms of both time and space, which ultimately optimized the performance of large-scale WSNs. In the work of Hazarika (2016), the author followed the idea of attaching methane and carbon monoxide gas sensors to safety helmets of coal mine workers as a cost-effective, reliable, and sufficient method for monitoring underground mine gas. The collected gas information was sent wirelessly to the controller, where the alarm sounded if the gas concentration went above the life-threatening level.

4.3.2. Temperature, humidity and gas

Several studies suggested prototypes capable of gathering richer monitored data by applying temperature and humidity sensors as well as gas sensors. These parameters have a considerable impact on the environmental conditions of the underground mine and consequently influence the comfort, safety and well-being of mine workforces (Li, 2019). In this regard, Zhu and You (2019) projected a real-time environmental monitoring system for coal mine safety based on ZigBee WSN, which could monitor the gas density, temperature and humidity parameters, send the data to the upper computer and, give early warning when the parameters go beyond a specific level. According to the trials, their proposed system was consistent in performance, precise in evaluation, and efficient in refining mine safety and declining accidents. Also, Wei and Li-Li (2009) presented a multi-parameter WSN monitoring system based on Zigbee protocol for real-time monitoring of the coal mine underground environmental factors (e.g., gas, temperature, humidity) and production parameters. The system intelligently sent early warning signals to a control room on the surface. This way, the staff on the ground become aware of the real-time condition of the underground environment and potential hazards (Qing-liang, Zhi-xian, & Zhen-chuan, 2008). Chehri (2011) assessed the performance and reliability of a Zigbee-based WSN for mine's safety monitoring in terms of its lagging time, throughput and package error rate through quantifying underground mine parameters such as temperature, humidity, level of carbon monoxides, fire exposure, and so forth. The initial results confirmed the possibility of monitoring mine from distant places using the WSN without rational delay. Li-min (2008) deployed a similar monitoring system collecting temperature, humidity, and methane values of underground coal mines in order to cover the real-time monitoring of working surface. The system transferred the captured data to the information processing terminal, which guided the data to the ground through Ethernet. The surface monitoring center then propagated the data by dint of a Local Area Network (LAN) to notify remote users through sending messages. Henriques and Malekian (2016) proposed another WSN-based system that monitored the ambient characteristics inside the mine environment, such as temperature, humidity, and gas. The system communicated the sensor data measured by the measurement node to the data collection node via the ZigBee wireless protocol. The data collection node then transferred the sensor information to the laptop computer/graphical user interface (GUI). The GUI was used as the visual output of the measurement parameters, which displayed the particular zones of the mine where the parameters exceeded their standard level. This helps the mining engineer to make more accurate safety decisions. These papers proposed the design of a WSN monitoring system based on ZigBee technology mainly because this technology promises flexible, small scale, low cost, low data rate, and low power consumption wireless networking (Chehri, 2011).

Some research in this area examined the integration of WSNs with other data communication/transmission systems to offer the benefits of both technologies supplement each other. An automatic underground mine monitoring and communication system based on the integration of Zigbee-based WSN aided Geographic Information System (GIS) was designed in the work of Moridi (2015). The system allowed near real-time monitoring and ventilation system control of underground mining activities from surface agency to provide emergency communication. Environmental features such as temperature, humidity, and gas concentration were sensed, ON and OFF ventilation fans were switched, and emergency messages were texted if monitored features exceeded normal values. Besides, this system achieved the possibility of multiuser surface operation and three-dimensional visualization to gain a realistic insight into the conditions of the underground environment and workers (Moridi, 2019). In the work of Feng, ShengYu, and Qi (2010), RFID was fused to Zigbee-based WSN in order to access the network for more effective data transmission and deal with the problem of poor anti-interference resulting in a better understanding of the existing coal mine safety risks. In this system, the former is applied to recognize target objects and, the latter is applied to monitor the target environment conditions. Taking advantage of WSN merged with the controller area network (CAN) bus technique and Web of Things (WoT) technology, comprehensive and well-timed remote monitoring and smart early-warning in the underground environment was made possible by Bo (2014). The CAN bus technology was applied for real-time monitoring of the operating state of ventilation and air conditioning equipment (e.g., blowers, ventilators, and air doors), while the underground environmental parameters were collected by WSN technology. Based on the WoT technology, all collected information was transmitted to the remote monitoring center for analysis to provide decision-making information for clients. If there was a parameter anomaly, the sound and light alarm was displayed on site and simultaneously in the remote monitoring center window. Another novel early-warning system was developed in the work of Pudke, Bhagat, and Nalbalwar (2017) for monitoring unsafe conditions based on applying wireless technologies, including ZigBee and GSM (Global System for Mobile Communications). Various types of sensors were used for measuring environmental conditions of mine such as temperature, gas, humidity, fire, and so forth. In this system, the constantly collected data from the underground station was transmitted to the ground station through a ZigBee wireless communication system. Then, the signal processor compared the sensed value with the predefined critical safety values and in a dangerous situation gave a warning alert to the GSM system through the micro-controller for calling and sending the message to the safety department. As evidenced above, establishing a reliable two-way communication system between employees working in the harsh environment of underground mines to a fixed ground station and vice versa is of great importance for safety purposes (Maity, Das, & Mukherjee, 2012; Moridi, 2018).

To establish more effective communications between the underground labors and the ground staff, some scholars deployed their monitoring systems by attaching temperature, humidity and gas sensors to the miners' PPE clothing, such as their helmets (Dohare, 2014; Geetha, 2013; Maity et al., 2012; Mishra, Malhotra, & Singh, 2018; Qiang, 2009; Sharma & Maity, 2018). The controller data were then transmitted to a ground control center through the low-power, low-cost Zigbee network. This process enables serving quicker and smarter rescue actions. Moreover, the mobility of WSNs offered by such systems is useful for addressing underground mine dynamic disasters. Designing wearable environmental monitoring systems has been suggested as a very affordable and appropriate technique for safety and health monitoring in the underground mine working environment (Adjiski, 2019; Cicioğlu & Çalhan, 2019; Schalk Wilhelm & Reza, 2019; Ziętek, 2020).

4.3.3. Noise

Noise exposure is one of the most important and highly prevalent occupational hazards in the mining industry (Lutz, 2015). It has been reported that noise levels at mine sites range from 85 to 140 dBA depending on the type of activities that can negatively impact the workers' health (McBride, 2004). The most negative effects caused by noise exposure are related to the hearing system and may produce professional deafness or even permanent deafness (Lu & Davis, 2016). High levels of noise may also result in physical and mental health issues such as elevated blood pressure and heart rate, anxiety, and distraction. Noise can also mask speech and warning signals that make communication difficult and increases the risk of accidents in a noisy environment (Jabłoński, Szer, & Szer, 2018).

To protect mine workers from excessive noise levels, the first step is to measure the noise level to understand its magnitude and take proper managerial actions (Kwon, 2018; Umar, 2018). WSN-based systems will enable mining companies to measure noise levels continuously using wearable devices. In a study conducted in mining, Henriques and Malekian (2016) developed a WSN-based system that was able to measure noise. In the proposed system, the sensed data were communicated to the data collection node via the ZigBee wireless protocol. This hardware then transferred the sensor data to the laptop computer/GUI. Based on the extracted information, a noise protection strategy was ultimately devised. This strategy provides mine engineers with a visual output of the noise levels in particular zones of the mine and helps them make more accurate decisions on whether or not a particular working zone is safe for mine employees. While in this paper, noise measurement devices are located at a specific distance from the source to measure the noise level, wearable devices can also be attached to the body of workers to measure noise exposure (Ali, 2011; Rempel, 2019).

4.3.4. Dust

Mine employees are at high risk for respiratory and lung diseases caused by mine dust (Laney & Weissman, 2014). Exposure to dust from mining may result in many pathological effects depending on composition, dimension, form, and levels of particles and duration of exposure (Utembe, 2015). To address this issue, a prototype wireless system was implemented in the work of Mahdavipour (2015) for continuous monitoring and automatic measurement of the total incombustible content (TIC) of the deposited dust in underground coal mines. The network was comprised of numerous low-cost/low-power optical and microfabricated sensors distributed throughout the mine. The distributed sensors employed continuous optical, dielectrometry and gravimetric methods to calculate the TIC of the deposited stack of float dust/rock dust as well as the moisture content and the mass of the deposited stack, respectively.

4.3.5. Airflow pressure

The volume of air flowing through the underground mine is a significant parameter influencing mine safety. The change in the volume of air and the pressure generated by airflow can contribute to the creation of hazardous situations (Trutwin, 1988; Wu & Gillies, 2005). For example, regarding some of the disasters of mine, reducing the airflow pressure results in increasing the concentration of toxic or explosive gas (Hongqing, 2011). Therefore, airflow monitoring and control within underground mines are of great importance for maintaining the health and safety of miners. Wu and Gillies (2005) approached a computerized method for real-time monitoring and control of airflow over the mine ventilation network. Their system was able to connect sensors into the ventilation network simulation software to report real-time information on variations in airflow and pressure throughout the underground mine. Another research conducted by Henriques and Malekian (2016) proposed a WSN-based system that was able to measure the ambient airflow inside the mine. The system communicated the sensed data to the data collection node via the ZigBee wireless protocol. This hardware then transferred the sensor data to the laptop computer/GUI. Based on the conditions measured by the airflow and gas sensors, a ventilation switching strategy was ultimately devised.

4.3.6. Fire

WSNs can also be applied for early detection and warning of underground mine fire since it is one of the foremost concerns of safety in such environments (Tan, 2007). Edwards (2000) conducted a laboratory experiment to measure the reaction time of fire sensors (e.g., semiconductor metal oxide, carbon monoxide, and smoke fire sensors) to different in-mine flammable materials. They aimed to compare and discriminate the detection and warning capability of fire sensors based on a neural network program. While this study generalized mine fire safety and ignored the variety of structures across different types of mine, Bhattacharjee (2012) calculated the response time of fire sensors specifically in a bord-and-pillar coal mine. The authors used WSN to design a fire detection, alarming, monitoring and prevention system for bord-and-pillar coal mines. Temperature and gas sensor nodes were distributed over the mine galleries for analyzing, storing, and sharing the gathered data in real-time. Thus, this system was able to accurately detect where the fire had occurred and in what direction it was spreading. The performance of the suggested system was assessed through rough simulations. The results revealed that the average network delay differed nearly linearly with the growing number of hops. The objective was to increase coverage of sensor nodes to the optimum level and reducing their fire detection delay to the lowest level. In order to monitor fire hazards in underground mines using WSN-based systems, a variety of environmental data collected by the sensor nodes are forwarded to a sink node for processing. The sink node is directly linked to a central monitoring station aboveground that makes decisions based on processed data (Muduli, Jana, & Mishra, 2018). To enhance the reliability and accuracy of the decision-making process, Muduli et al. (2018) and Basu (2019) developed novel WSN-based fire monitoring systems using the fuzzy logic approach. Their proposed systems would assist in making real-time decisions on intrinsically uncertain and imprecise monitoring data. Through simulation experiments, it was proved that the application of the fuzzy logic-based monitoring system increases the validity and effectiveness of fire safety risk assessment in comparison with the wired and offline monitoring systems already applied in underground coal mines. Another technique offered by Muduli, Mishra, and Jana (2019) to address the problem of imprecise and vague information was based on learning from the former experiences, named data mining. In this system, machine learning was incorporated with the wireless underground sensor network for monitoring the environmental factors and prediction of coal mine fire risk. A supervised learning algorithm was proposed and implemented at the sink nodes instead of a monitoring center for taking prompt real-time decisions on sensed information in case of any fire risk. The application of machine learning methods was also targeted in the work of Zhao, Liu, and Hai (2013) for the classification of fire safety status in underground mines based on sensed data by WSN and the use of an unsupervised neural network model called Self-Organizing Map (SOM). Remotely sensed environmental data (e.g., temperature, gas density, dust density, wind speed) were sent to the sink node to be processed with information fusion technology. Using the SOM, the authors developed an information fusion model, which was effectively able to make high-dimensional environmental data to a low-dimensional one. This model allowed for classifying coal mine safety status into four clusters, including safe, generally safe, abnormal, and dangerous. Therefore, the application of data fusion techniques was introduced as a useful strategy for improving the efficiency of WSN-based environment monitoring and the accuracy of safety predictions and decisions.

4.3.7. Roof fall

The structure variations arising from the unstable nature of underground mining geology are another threat endangering the lives of miners. Li and Liu (2009) deployed a Structure-Aware Self-Adaptive WSN system, SASA, which was able to quickly detect roof fall risk in underground coal mines. A wireless mesh sensor network was deployed on the roof and wall of the mine galleries that assisted in early detection of collapse cavity zones and precisely location monitoring of personnel. When a roof collapse occurs in the underground mine, vulnerable monitoring systems may expose many damages that are infrequent in other WSN systems. So, SASA was designed based on a sound and strong mechanism for controlling inquiries under unstable conditions and was capable of detection and reconfiguration of displaced sensor nodes. The scalability of the SASA prototype was then evaluated by proceeding with a large-scale trace-based simulation using the actual data collected from field testing. Hu, Shu, and Song (2013) proposed a hierarchy WSN topology for identification and localization of coal mine collapse hole underground, where a group of sensor nodes damaged by collapse creates a hole in the wireless network. They followed the connectivity-based localization method for measuring the connectivity of sensor nodes in WSN, which was then calculated for accuracy of detecting collapse hole using Fisher information. The Fisher information measures the amount of information that an observable random variable carries about an unknown parameter. This system was able to enhance safety by delivering accurate location information of a collapse hole to the coal miners.

4.4. Performance evaluation of the communication system in WSNs

As identified by this review paper, WSNs have aided the safety and health management process in underground mines through three major research streams, including location monitoring and tracking, physiological and body kinematics monitoring, and environmental monitoring. However, another research stream has also been followed by some existing studies in this area focusing on performance evaluation of the communication system in WSNs, which is discussed in the following paragraphs.

When designing a wireless communication system, predicting the propagation behavior of radio waves in the hostile and unstable environment of underground mines will be a major challenge. The radio propagation in mines may be affected by either system parameters (e.g., radio signal frequency, antenna radiation pattern, position of transceivers) or distinctive features of different mines (mine geometry, miming methods, obstacles and their positions, humidity, temperature, noise, wall roughness, electromagnetic properties of walls, etc.) (Hrovat, Kandus, & Javornik, 2013; Luomala & Hakala, 2015; Ranjan et al., 2018, 2019, 2020; Zhou, 2015). This prevents the adoption of standard communication systems designed for a normal milieu (Ranjan et al., 2016). In this regard, different propagation models developed by previous studies aim to measure the accurate number and location of base stations and predict high quality and reliable radio coverage for such complex environments. On the basis of its proposed model, each paper then critically analyzes the signal propagation characteristics (e.g., strength, attenuation) across underground mines considering the impacts of the above-mentioned variables. To validate these propagation models for real-life scenarios, experimental measurements are needed to be compared to the theoretically predicted values (Bandyopadhyay, 2007; Bedford, Kennedy, & Foster, 2017; Ranjan, Sahu, & Misra, 2020; Ranjan, 2017; Ranjany, 2016; Sun & Akyildiz, 2010). In addition, performance evaluation of both currently available and novel wireless communication standards for various underground mines is required. For example, Kennedy and Bedford (Kennedy & Bedford, 2014) carried out an experimental analysis to characterize the subterranean performance of IEEE 802.15.1 Bluetooth and IEEE 802.11x Wi-Fi communication standards functioning in the 2.4 GHz and 5 GHz frequency bands. Test results were shown as plots of throughput versus distance across a mostly line-of-sight path. Since the loss of line-of-sight is highly probable in WSN (Nerguizian, 2005); further tests calculated the impact of multipath propagation. The outcomes can assist in network planning by providing insights into the operation of wireless communication networks in underground mines.

On the other hand, taking into account the extreme conditions threatening underground mine safety, the wireless communication systems need to be very reliable and durable. For instance, Fiscor (2008), Yang et al. (2010) discussed the idea of using the Mesh network configuration for mining installations. This delivers a self-organizing and self-healing wireless network in emergencies, which means the network is able to make a connection with the alternative routes for data transmission if the predefined route is damaged. To design an effective communication link, identifying the optimal location to place the wireless mesh access point in the mining gallery is of importance. As an example, Moutairou, Aniss, and Delisle (2006) chose the optimization algorithm that allowed detecting the best neighboring access point and the shortest route in the wireless mesh network.

5. Research gaps and future studies

In this review paper, various applications of WSNs to improve OSH in the mining industry were investigated. However, a set of challenges and issues still exist that should be addressed for the proposed WSN-based systems to improve their effectiveness, reliability, and security. Seven main research gaps that deserve the attention of future studies are highlighted in this section.

5.1. Further applications of WSNs for underground mining OSH management

After reviewing the current applications of WSN for improving OSH in underground mines, this study reveals the gray areas that received limited or no attention from scholars. These are potential areas for future research towards the wide and efficient application of WSN technology into the various scopes of underground mining OSH management systems.

In the scope of location monitoring and tracking, WSN has been widely used for improving rescue operations in underground mines through detection and trajectory of mine employees on a real-time basis. Moreover, enhancing the localization accuracy and performance of off-the-shelf tracking sensors has been also targeted by a number of studies. However, improving transport safety in underground mines and estimating the position of workforces in blind zones are two areas that have not been sufficiently investigated. Besides, future studies should also focus on the utilization of WSN technique for monitoring and reporting some of the important hazardous situations, such as unauthorized action or entry of workforces inside a pre-defined hazardous zone, collision accidents between workers and/or heavy equipment, proximity events resulting from movements of dynamic objects in close proximity to each other, backover accidents, interference between tasks of different working groups, and workers' unsafe behaviors and risky activities.

In the scope of physiological monitoring, although wearable wireless physiological sensor systems have been used for timely rescue responses and monitoring of emergency states, they have not been adopted for real-time monitoring and early warning of some of the most important and highly prevalent occupational hazards for underground miners' including fatigue and stress. Moreover, in the scope of body kinematics monitoring, the applications of wearable wireless kinematics sensor systems for monitoring the awkward posture and whole-body vibration of miners are the overlooked areas of study.

In the scope of environmental monitoring, although WSN has been applied for continuous monitoring and recording of various environmental parameters in underground mines, it's application for addressing different calamitous conditions, such as gas concentration, dust concentration, high temperature and humidity, noise exposure, airflow pressure, fire, and roof fall is not mature enough.

5.2. Application of WSNs from research to real-world practice

The majority of selected articles evaluated WSNs proposed solution through simulation or laboratory experiments. Thus, the performance of these systems has not been validated in real underground mine environments.

Taking into account the extreme and dynamic environment of underground mines, wireless communication and networking through them face unique challenges (e.g., the possibility of roof/wall collapse, spatial disorientations, irregular distributions of minerals, gas concentration, humidity, temperature, and dust) (Akyildiz & Stuntebeck, 2006; Ranjan et al., 2016). These challenges substantially affect the real-time propagation characteristics and strength of wireless signals (Ranjan, 2019). In addition, constant changes of mine infrastructure create the demands for reconfiguration of wireless access points to preserve the network integrity. Sensor blind spots are also produced frequently, which influence the network performance seriously (Dohare, 2015).

To fill the gap between research and practice, the existing systems should be thoroughly experimented by future researchers considering real underground mine propagation structures and environments. Sensor deployment techniques, data transmission approaches and, propagation channel characterization and modeling for WSNs within various underground mines of different types are few areas that deserve more attention of future studies (Forooshani, 2013; Ranjan et al., 2016).

To date, deploying the generic, singular models of WSNs–based communication systems in underground mines has been criticized (Ranjan et al., 2018, 2019). In this regard, enhancing the scalability of WSNs is needed to fulfill the specific requirements of underground mine communications in large-scale real-life scenarios (Ranjan et al., 2016). Moreover, these networks are expected to adopt continuous variations by covering the newly developed working areas (Ranjan, 2019).

Since underground mines are usually in transition between normal and disaster environments (Qaraqe, 2013), future research should have visions of designing intelligent emergency response systems to be able to detect any sudden change or catastrophic event on a real-time basis and recover from it by generating a reliable wireless communication between inside and outside of the underground mine (Kumar, 2016; Ranjan et al., 2016; Yang, Zhang, & Liu, 2010).

Performance evaluation of wireless underground sensor networks is another open area for further investigations. Identification of operating frequencies, the available radio bandwidth, energy consumption, topology division, path loss and fading for underground wireless communications are among the research priorities (Akyildiz & Stuntebeck, 2006; Ranjan et al., 2016). Additionally, this could help design novel communication protocols for underground WSNs as well (Akyildiz & Stuntebeck, 2006; Ranjan, 2019). Understanding the long-term behavior of the evaluated systems is also necessary (Ruiz-Garcia, 2009). Energy constraint has been considered as the major performance bottleneck in the current applications of WSNs (Pantelopoulos & Bourbakis, 2009). Implementation of WSNs in underground mines for a long period necessitates the recognition and availability of the most appropriate methods for powering wireless sensor nodes, such as energy-harvesting techniques and battery technologies (Kiziroglou, 2016). Also, scheduling mechanisms for energy-efficient communication in underground WSNs are other possible directions for future research (Ranjan, 2019).

5.3. Big data analytics and management

The application of WSNs for monitoring underground mine safety creates large volumes of data on a daily basis. One concern with the huge data is difficulties in information management and processing. This challenge disrupts the applications of WSNs to make real-time decisions. With this said, more scholarly works are required on developing novel data mining techniques, which address knowledge extraction from constantly receiving big data from WSNs (Mahmood, 2013). On the other hand, there is also a necessity to further explore the methods to use the collected data for predicting future health and safety risks. In this regard, the research community should focus more on deploying intelligent machine learning techniques to perform big data analytics for the extreme work environment of underground mines (Muduli et al., 2019; Ranjan, 2019).

Another challenge faced by mine management is the privacy and security of big data collected by WSNs.

This is a novel area of research that should gain momentum in the research community, especially with regard to the emergence of technologies such as Cloud Computing, analytics engines, and social networks. Developing privacy and security of big data models, techniques and algorithms are some open areas for further investigations (Cuzzocrea, 2014). For example, Kumari and Om (2016) developed a user authentication protocol for wireless underground sensor networks to minimize the security concerns in coal mines. The protocol mainly aimed to prevent unauthorized access to the information in WSNs and to resist the system attackers.

5.4. Deploying multiple WSNs-based monitoring systems

To date, some studies have been presented with the aim to integrate personal health monitoring devices into multiple monitoring wireless underground sensor networks (Kiziroglou, 2016). According to their results, there is a need to deploy an integrated WSN for underground mines that is capable of localization and tracking, environmental monitoring, structural health monitoring and personnel health condition monitoring simultaneously and continuously (Dohare, 2015). Despite the existing theoretical approaches, a lot of work remains to be done in this area to reach an efficient customized design approach (Ranjan, 2019). In further detail, future researchers should evaluate the strengths and weaknesses of various sensing technologies applied in OSH of the mining industry in order to integrate some of them for the purposes of multi-sensor platforms and multi-parameter monitoring (Antwi-Afari, 2019; Awolusi et al., 2018). This way, researchers and practitioners take maximum advantage of the complementary capabilities of these sensing technologies and increase the precision of health and safety risk assessments via using multisensory data fusion techniques (Ahn, 2019).

5.5. Integration of WSNs with other communication systems

The integration of WSNs and other communication systems offers the benefits of both technologies supplementing each other. For this purpose, the integration of WSNs with other systems such as RFID, GIS, CAN and WoT within underground mines have been proposed by few past studies (Bo, 2014; Feng et al., 2010; Moridi, 2015). In addition, the integration of WSN technology into a fiber-optic platform is also another open area for further investigations (Niu, 2007; Wang, 2009; Yang & Huang, 2007; Zhang, 2014). More effective data transmission, comprehensive and well-timed remote monitoring and smart early warning, a better understanding of the existing mine safety risks, and emergency communication are some benefits that have been pointed out by previous studies. However, there is still a great potential for future research in this area.

5.6. Adapting WSNs to the IoT infrastructure

It is envisioned that the IoT paradigm will be incorporated into underground mines (Muduli et al., 2018); as an enabling technology for protecting underground mine safety such as early warning of different OSH risk factors and major disasters, and safety improvement systems (Qiuping et al., 2011; Singh, Kumar, & Hötzel, 2018). With this said, since WSN is considered as a foundation technology for IoT, adapting WSNs to the IoT infrastructure is of great importance in the scope of mining safety and health research (Singh et al., 2018). This way, through connecting the actual work environments (the sensed data in WSNs) to the digital world (Internet, databases), the underground mines move towards being smarter and digital (Muduli et al., 2018; Singh et al., 2018). The notion of smart mine using IoT can enhance the effectiveness of WSN-based monitoring systems in underground mines. It is because IoT provides each object of the underground mine with a unique identity and connects it to the other objects and the internet. Also, IoT enables the research community, practitioners, mining organizations, and so forth, to reach large volumes of monitoring data, as well as being connected with one another (Muduli et al., 2018).

5.7. Autonomous WSNs

The collaboration between WSNs and robotics is another emerging research area in OSH of underground mines, which obtains benefits from the integrated utilization of both technologies (Alam, Eyers, & Huang, 2015; Kumar, 2016).

One potential application of this contribution can be seen in the movement towards autonomous mining equipment. Robotic WSNs with the ability to sense and map the mine structure can provide the operators with the vision to accomplish mine operations with fewer risks and automated algorithms can keep the mining process online even in harsh working conditions (Kiziroglou, 2016; Kumar, 2016).

Another promising application of robots in WSNs is the entry of mobile robots to the underground mines. While carrying sensors, these robots can explore, characterize, and report the real-time physical and environmental conditions as well as performing search and rescue operations (Alzaq & Kabadayi, 2013; Kumar, 2016).

6. Conclusion

Due to the highly complex and hazardous underground mine working environments, on-site management of OSH in this industry is very challenging. WSN technology is a reliable and promising approach for regular monitoring of mining operations as well as safety and health conditions of underground miners. In the present study, a systematic scoping review of the existing sources of literature has been undertaken to examine the current applications of WSN technology to the safety and health of mining sites. The review has covered 66 relevant peer-reviewed publications, including 34 journal and 32 conference articles. The results have been categorized and discussed under three predominant categories including: (1) location monitoring and tracking; (2) physiological and body kinematics monitoring; and (3) environmental monitoring. Also, seven major directions for future research and practical interventions have been identified based on the existing research gaps including: (1) further applications of WSNs for underground mining OSH management; (2) application of WSNs from research to real-world practice; (3) big data analytics and management; (4) deploying multiple WSNs-based monitoring systems; (5) integration of WSNs with other communication systems; (6) adapting WSNs to the Internet of Things (IoT) infrastructure; and (7) autonomous WSNs. Investigation of these future areas of research could help improve the performance and applicability of WSN technology for OSH in the mining industry to reach a holistic OSH management system for underground miners using WSNs. In conclusion, this research emphasizes the need for taking concrete actions to place the development of WSN technology and the use of information provided via this technology as a top priority of OSH management systems to create healthier and safer mining workplaces in the near future.

Declaration of interest

All writers confirm that there is no known conflict of interest associated with this publication.

Appendix

List of articles reviewed in this research:

M. Li, Y. Liu, Underground structure monitoring with wireless sensor networks, Proceedings of the 6th international conference on Information processing in sensor networks, ACM, 2007, pp. 69–78.

Chehri, H. Mouftah, P. Fortier, H. Aniss, Experimental testing of IEEE801. 15.4/ZigBee sensor networks in confined area, 2010 8th Annual Communication Networks and Services Research Conference, IEEE, 2010, pp. 244–247.

P. Misra, S. Kanhere, D. Ostry, S. Jha, Safety assurance and rescue communication systems in high-stress environments: A mining case study, IEEE Communications Magazine 48 (4) (2010) 66–73.

Y. Wu, M. Chen, K. Wang, G. Fu, A dynamic information platform for underground coal mine safety based on internet of things, Safety science 113 (2019) 9–18.
Swain, M. Bepari, N. Patnaik, H.B. Sahu, Zigbee Based Wireless Data Acquisition System for Underground Mines—A Feasibility Study, 2018 IEEE 21st International Symposium on Real-Time Distributed Computing (ISORC), IEEE, 2018, pp. 152–155.
H. Wang, F. Luo, K. Wang, Y. Xu, An emergency communication architecture based on ethernet and wsn for coal mines, Procedia Engineering 23 (2011) 403–407.
X Wang, X. Zhao, Z. Liang, M. Tan, Deploying a wireless sensor network on the coal mines, 2007 IEEE international conference on networking, sensing and control, IEEE, 2007, pp. 324–328.
P. Chen, C. Zhao, Area wireless sensor networks for personnel location under coalmine, 2007 2nd IEEE Conference on Industrial Electronics and Applications, IEEE, 2007, pp. 2882–2885.
K. Chen, C. Wang, L. Chen, X. Niu, Y. Zhang, J. Wan, Smart safety early warning system of coal mine production based on WSNs, Safety science 124 (2020) 104609.
S. Dehran, P. Agrawal, P. Midha, Digital Applications in Metals and Mining Industry, American Journal of Operations Management and Information Systems 3 (1) (2018) 33–37.
Bo, Z. Peng, Z. Da, C. Junliang, The complex alarming event detecting and disposal processing approach for coal mine safety using wireless sensor network, International Journal of Distributed Sensor Networks 8 (11) (2012) 280576.
Chehri, P. Fortier, P.M. Tardif, UWB-based sensor networks for localization in mining environments, Ad Hoc Networks 7 (5) (2009) 987–1000.
P. Guo, T. Jiang, K. Zhang, Novel energy-efficient miner monitoring system with duty-cycled wireless sensor networks, International Journal of Distributed Sensor Networks 8 (1) (2012) 975082.
H. Jiang, L. Chen, J. Wu, S. Chen, H. Leung, A reliable and high-bandwidth multihop wireless sensor network for mine tunnel monitoring, IEEE Sensors journal 9 (11) (2009) 1511–1517.
Z. Liu, C. Li, D. Wu, W. Dai, S. Geng, Q. Ding, A wireless sensor network based personnel positioning scheme in coal mines with blind areas, Sensors 10 (11) (2010) 9891–9918.
G. Kloos, J.E. Guivant, E.M. Nebot, F. Masson, Range based localisation using RF and the application to mining safety, 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2006, pp. 1304–1311.
Kumar, Application of modern tools and techniques for mine safety & disaster management, Journal of The Institution of Engineers (India): Series D 97 (1) (2016) 77–85.
U.I. Minhas, I.H. Naqvi, S. Qaisar, K. Ali, S. Shahid, M.A. Aslam, A WSN for monitoring and event reporting in underground mine environments, IEEE Systems Journal 12 (1) (2017) 485–496.
Z. Qinghua, W. Bo, C. Guoquan, W. Zhuan, W. Jun, Y. Dawei, Object position tracking based on e-map and RFID in coal mine, 2009 4th IEEE Conference on Industrial Electronics and Applications, IEEE, 2009, pp. 880–885.
V Savic, H. Wymeersch, E.G. Larsson, Simultaneous sensor localization and target tracking in mine tunnels, Proceedings of the 16th International Conference on Information Fusion, IEEE, 2013, pp. 1427–1433.
Y. Wang, L. Huang, W. Yang, A novel real-time coal miner localization and tracking system based on self-organized sensor networks, EURASIP Journal on Wireless Communications and Networking 2010 (2010) 1–14.
X Xu, P. Zheng, L. Li, H. Chen, J. Ye, J. Wang, design of underground miner positioning system based on zigbee technology, International Conference on Web information systems and Mining, Springer, 2012, pp. 342–349.
W. Yan, Z. Ya-ru, M. Yong, Study on the coal mine personnel position system based on wireless body sensor networks, 2008 5th International Summer School and Symposium on Medical Devices and Biosensors, IEEE, 2008, pp. 75–78.
G. Yin, The design and optimization of underground scraper wireless monitoring networks based on ZigBee, Physics Procedia 22 (2011) 270–274.
S. Kumari, H. Om, Authentication protocol for wireless sensor networks applications like safety monitoring in coal mines, Computer Networks 104 (2016) 137–154.
M.A. Akkaş, Using wireless underground sensor networks for mine and miner safety, Wireless Networks 24 (1) (2018) 17–26.
M. Cicioğlu, A. Çalhan, Performance analysis of IEEE 802.15. 6 for underground disaster cases, Computer Standards & Interfaces 66 (2019) 103364.
B.-s. Nie, W.-x. Chen, L.-k. Wang, R.-m. Zhang, C. Wang, Internet of things-based positioning of coalmine personnel and monitoring of emergency state, 2011 Second International Conference on Digital Manufacturing & Automation, IEEE, 2011, pp. 657–660.
Ranjan, Y. Zhao, H.B. Sahu, P. Misra, Opportunities and challenges in health sensing for extreme industrial environment: Perspectives from underground mines, IEEE Access 7 (2019) 139181–139195.
R. Revindran, H. Vijayaraghavan, M.-Y. Huang, Smart helmets for safety in mining industry, 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2018, pp. 217–221.
M. Yong, Y. Yu, W. Yan, L. Huludao, L. Gang, L. Ling, Optimization design of coal mine wireless body sensor network based on Genetic Algorithm, 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing, Vol. 2, IEEE, 2009, pp. 366–369.
I.O. Osunmakinde, Towards safety from toxic gases in underground mines using wireless sensor networks and ambient intelligence, International Journal of Distributed Sensor Networks 9 (2) (2013) 159273.
X. Niu, X. Huang, Z. Zhao, Y. Zhang, C. Huang, L. Cui, The design and evaluation of a wireless sensor network for mine safety monitoring, IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference, IEEE, 2007, pp. 1291–1295.
Y.S. Dohare, T. Maity, P.S. Paul, P.S. Das, Design of surveillance and safety system for underground coal mines based on low power WSN, 2014 International Conference on Signal Propagation and Computer Technology (ICSPCT 2014), IEEE, 2014, pp. 116–119.
Geetha, Intelligent helmet for coal miners with voice over Zigbee and environmental monitoring, Middle-East J. Sci. Res 16 (12) (2013) 1835–1837.
T. Maity, P.S. Das, M. Mukherjee, A wireless surveillance and safety system for mine workers based on Zigbee, 2012 1st International Conference on Recent Advances in Information Technology (RAIT), IEEE, 2012, pp. 148–151.
Mishra, S. Malhotra, H. Singh, Real Time Monitoring & Analyzation Of Hazardous Parameters In Underground Coal Mines Using Intelligent Helmet System, 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT), IEEE, 2018, pp. 1–5.
Qiang, S. Ji-Ping, Z. Zhe, Z. Fan, ZigBee based intelligent helmet for coal miners, 2009 WRI World Congress on Computer Science and Information Engineering, Vol. 3, IEEE, 2009, pp. 433–435.
M. Sharma, T. Maity, Low cost low power smart helmet for real-time remote underground mine environment monitoring, Wireless Personal Communications 102 (1) (2018) 149–162.
Y. Zhu, G. You, Monitoring System for Coal Mine Safety Based on Wireless Sensor Network, 2019 Cross Strait Quad-Regional

Radio Science and Wireless Technology Conference (CSQRWC), IEEE, 2019, pp. 1–2.

S. Wei, L. Li-Li, Multi-parameter monitoring system for coal mine based on wireless sensor network technology, 2009 International Conference on Industrial Mechatronics and Automation, IEEE, 2009, pp. 225–227.

S. Qing-liang, Z. Zhi-xian, Z. Zhen-chuan, Research on safety monitoring system of mine based on ad hoc technology, 2008 International Conference on Information Management, Innovation Management and Industrial Engineering, Vol. 2, IEEE, 2008, pp. 450–453.

A.J. Pudke, S.N. Bhagat, S. Nalbalwar, LabVIEW based coal mine monitoring and alert system with data acquisition, 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, 2017, pp. 1166–1170.

M.A. Moridi, Y. Kawamura, M. Sharifzadeh, E.K. Chanda, M. Wagner, H. Jang, H. Okawa, Development of underground mine monitoring and communication system integrated ZigBee and GIS, International Journal of Mining Science and Technology 25 (5) (2015) 811–818.

M.A. Moridi, M. Sharifzadeh, H. Jang, Y. Kawamura, An Automated Underground Space Monitoring and Communication System Based on Wireless Sensor Networks, International Symposium on Mine Planning & Equipment Selection, Springer, 2019, pp. 255–261.

M.A. Moridi, M. Sharifzadeh, Y. Kawamura, H.D. Jang, Development of wireless sensor networks for underground communication and monitoring systems (the cases of underground mine environments), Tunnelling and Underground Space Technology 73 (2018) 127–138.

Y. Li-min, L. Anqi, S. Zheng, L. Hui, Design of monitoring system for coal mine safety based on wireless sensor network, 2008 IEEE/ASME International Conference on Mechtronic and Embedded Systems and Applications, IEEE, 2008, pp. 409–414.

F. Feng, H. ShengYu, X. Qi, The research of the Zigbee and RFID fusion technology in the coal mine safety, 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering, Vol. 3, IEEE, 2010, pp. 32–36.

Chehri, W. Farjow, H.T. Mouftah, X. Fernando, Design of wireless sensor network for mine safety monitoring, 2011 24th Canadian Conference on Electrical and Computer Engineering (CCECE), IEEE, 2011, pp. 001532–001535.

Bo, C. Xin, Z. Zhongyi, Z. Chengwen, C. Junliang, Web of things-based remote monitoring system for coal mine safety using wireless sensor network, International Journal of Distributed Sensor Networks 10 (8) (2014) 323127.

Q.-F. Wang, S. Zhang, Y. Yang, L. Tang, The application of wireless sensor networks in coal mine, 2009 7th International Conference on Information, Communications and Signal Processing (ICICS), IEEE, 2009, pp. 1–4.

W. Yang, Y. Huang, Wireless sensor network based coal mine wireless and integrated security monitoring information system, Sixth International Conference on Networking (ICN'07), IEEE, 2007, pp. 13–13.

Y. Zhang, W. Yang, D. Han, Y.-I. Kim, An integrated environment monitoring system for underground coal mines—Wireless sensor network subsystem with multi-parameter monitoring, Sensors 14 (7) (2014) 13149–13170.

Zhao, F. Liu, X. Hai, An application of wireless sensor networks in underground coal mine, International Journal of Future Generation Communication and Networking 6 (5) (2013) 117–126.

S. Umar, U. Moulali, T. Shesagiri, M. Venunath, WSN based alter system to save life in mines with high security system, 2018 2nd International Conference on Inventive Systems and Control (ICISC), IEEE, 2018, pp. 1390–1397.

W. Tan, Q. Wang, H. Huang, Y. Guo, G. Zhang, Mine fire detection system based on wireless sensor network, 2007 International Conference on Information Acquisition, IEEE, 2007, pp. 148–151.

H. Wu, A. Gillies, Real-time airflow monitoring and control within the mine production system, Eighth International Mine Ventilation Congress, 2005, pp. 383–389.

L. Muduli, P.K. Jana, D.P. Mishra, Wireless sensor network based fire monitoring in underground coal mines: A fuzzy logic approach, Process Safety and Environmental Protection 113 (2018) 435–447.

L. Muduli, D.P. Mishra, P.K. Jana, Wireless Sensor Network Based Underground Coal Mine Environmental Monitoring Using Machine Learning Approach, Proceedings of the 11th International Mine Ventilation Congress, Springer, 2019, pp. 776–786.

M. Li, Y. Liu, Underground coal mine monitoring with wireless sensor networks, ACM Transactions on Sensor Networks (TOSN) 5 (2) (2009) 1–29.

O. Mahdavipour, T. Mueller-Sim, D. Fahimi, S. Croshere, P. Pillatsch, J. Merukh, V.Z. Baruffa, J. Sabino, K. Tran, G. Alanis, Wireless sensors for automated control of total incombustible content (TIC) of dust deposited in underground coal mines, 2015 IEEE SENSORS, IEEE, 2015, pp. 1–4.

S. Hu, H. Shu, X. Song, Fisher information of mine collapse hole detection based on sensor nodes connectivity, International Journal of Distributed Sensor Networks 9 (9) (2013) 306496.

V. Henriques, R. Malekian, mine safety system using wireless sensor network, IEEE access 4 (2016) 3511–3521.

S. Basu, S. Pramanik, S. Dey, G. Panigrahi, D.K. Jana, Fire monitoring in coal mines using wireless underground sensor network and interval type-2 fuzzy logic controller, International Journal of Coal Science & Technology 6 (2) (2019) 274–285.

S. Bhattacharjee, P. Roy, S. Ghosh, S. Misra, M.S. Obaidat, Wireless sensor network-based fire detection, alarming, monitoring and prevention system for Bord-and-Pillar coal mines, Journal of Systems and Software 85 (3) (2012) 571–581.

Ranjan, H. Sahu, P. Misra, Modeling and measurements for wireless communication networks in underground mine environments, Measurement 149 (2020) 106980.

References

Adjiski, V. et al. (2019). System architecture to bring smart personal protective equipment wearables and sensors to transform safety at work in the underground mining industry. *Rudarsko-geološko-naftni zbornik, 34*(1).

Ahn, C. R. et al. (2019). Wearable sensing technology applications in construction safety and health. *Journal of Construction Engineering and Management, 145*(11), 03119007.

Akkaş, M. A. (2018). Using wireless underground sensor networks for mine and miner safety. *Wireless Networks, 24*(1), 17–26.

Aktaş, A. (2019). Control and application of accuracy positioning estimation based real-time location system. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA), 9*(2), 82–88.

Akyildiz, I. F., & Stuntebeck, E. P. (2006). Wireless underground sensor networks: Research challenges. *Ad Hoc Networks, 4*(6), 669–686.

Alam, A. A., Eyers, D., & Huang, Z. (2015). Helping secure robots in WSN environments by monitoring WSN software updates for intrusions. *2015 6th International Conference on Automation, Robotics and Applications (ICARA)*. IEEE.

Ali, S. A. (2011). A case study of construction noise exposure for preserving worker's hearing in Egypt. *Acoustical Science and Technology, 32*(5), 211–215.

Alzaq, H., & Kabadayi, S. (2013). Mobile robot comes to the rescue in a WSN. *2013 IEEE 24th annual international symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE.

Amponsah-Tawiah, K., & Dartey-Baah, K. (2011). The mining industry in Ghana: A blessing or a curse. *International Journal of Business and Social Science, 2*(12).

Antwi-Afari, M. F. et al. (2019). Sensing and warning-based technology applications to improve occupational health and safety in the construction industry. *Engineering, Construction and Architectural Management*.

Awolusi, I., Marks, E., & Hallowell, M. (2018). Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Automation in Construction, 85*, 96–106.

Baek, J. et al. (2017). BBUNS: Bluetooth beacon-based underground navigation system to support mine haulage operations. *Minerals, 7*(11), 228.

Baggio, A., & Langendoen, K. (2008). Monte Carlo localization for mobile wireless sensor networks. *Ad Hoc Networks, 6*(5), 718–733.
Bandyopadhyay, L., Chaulya, S., & Mishra, P. (2010). Wireless information and safety system for mines. In *Wireless communication in underground mines* (pp. 175–262). Springer.
Bandyopadhyay, L. et al. (2007). Studies on radio frequency propagation characteristics for underground coalmine communications.
Barma, M. K. D. et al. (2014). A new approach to overcome the weakness in DVR protocol based on component neigbhour in MANET. *2014 international conference on soft computing and machine intelligence*. IEEE.
Basu, S. et al. (2019). Fire monitoring in coal mines using wireless underground sensor network and interval type-2 fuzzy logic controller. *International Journal of Coal Science & Technology, 6*(2), 274–285.
Bauerle, T., Dugdale, Z., & Poplin, G. (2018). Mineworker fatigue: A review of what we know and future decisions. *Mining Engineering, 70*(3), 33.
Bedford, M. D., Kennedy, G. A., & Foster, P. J. (2017). Radio transmission characteristics in tunnel environments. *Mining Technology, 126*(2), 77–87.
Bhattacharjee, S. et al. (2012). Wireless sensor network-based fire detection, alarming, monitoring and prevention system for Bord-and-Pillar coal mines. *Journal of Systems and Software, 85*(3), 571–581.
Blondeel, M., & Van de Graaf, T. (2018). Toward a global coal mining moratorium? A comparative analysis of coal mining policies in the USA, China, India and Australia. *Climatic Change, 150*(1), 89–101.
Blood, R. P., Ploger, J. D., & Johnson, P. W. (2010). Whole body vibration exposures in forklift operators: Comparison of a mechanical and air suspension seat. *Ergonomics, 53*(11), 1385–1394.
Bo, C. et al. (2012). The complex alarming event detecting and disposal processing approach for coal mine safety using wireless sensor network. *International Journal of Distributed Sensor Networks, 8*(11), 280576.
Bo, C. et al. (2014). Web of things-based remote monitoring system for coal mine safety using wireless sensor network. *International Journal of Distributed Sensor Networks, 10*(8), 323127.
Boshuizen, H. C., Bongers, P. M., & Hulshof, C. (1992). Self-reported back pain in fork-lift truck and freight-container tractor drivers exposed to whole-body vibration. *Spine, 17*(1), 59–65.
Bovenzi, M. (2009). Metrics of whole-body vibration and exposure–response relationship for low back pain in professional drivers: A prospective cohort study. *International Archives of Occupational and Environmental Health, 82*(7), 893–917.
Butlewski, M. et al. (2015). Fatigue of miners as a key factor in the work safety system. *Procedia Manufacturing, 3*, 4732–4739.
Chehri, A. et al. (2010). Experimental testing of IEEE801. 15.4/ZigBee sensor networks in confined area. *2010 8th annual communication networks and services research conference*. IEEE.
Chehri, A. et al. (2011). Design of wireless sensor network for mine safety monitoring. *2011 24th Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE.
Chehri, A., Fortier, P., & Tardif, P. M. (2009). UWB-based sensor networks for localization in mining environments. *Ad Hoc Networks, 7*(5), 987–1000.
Chen, J. et al. (2003). Predictors of whole-body vibration levels among urban taxi drivers. *Ergonomics, 46*(11), 1075–1090.
Chen, S. et al. (2020). Linear and nonlinear analyses of normal and fatigue heart rate variability signals for miners in high-altitude and cold areas. *Computer Methods and Programs in Biomedicine, 196*, 105667.
Chen, K. et al. (2020). Smart safety early warning system of coal mine production based on WSNs. *Safety Science, 124*, 104609.
Chen, S. et al. (2021). Psychophysiological data-driven multi-feature information fusion and recognition of miner fatigue in high-altitude and cold areas. In *Computers in biology and medicine* (pp. 104413).
Chen, P., & Zhao, C. (2007). Area wireless sensor networks for personnel location under coalmine. *2007 2nd IEEE conference on industrial electronics and applications*. IEEE.
Cicioğlu, M., & Çalhan, A. (2019). Performance analysis of IEEE 802.15. 6 for underground disaster cases. *Computer Standards & Interfaces, 66*, 103364.
Cuzzocrea, A. (2014). Privacy and security of big data: Current challenges and future research perspectives. *Proceedings of the first international workshop on privacy and secuirty of big data*.
Dehran, S., Agrawal, P., & Midha, P. (2018). Digital applications in metals and mining industry. *American Journal of Operations Management and Information Systems, 3*(1), 33–37.
Dhillon, B. S. (2010). Mine safety: A modern approach. Springer Science & Business Media.
Diaz, S., Mendez, D., & Kraemer, R. (2019). A review on self-healing and self-organizing techniques for wireless sensor networks. *Journal of Circuits, Systems and Computers, 28*(05), 1930005.
Dohare, Y. S. et al. (2015). Wireless communication and environment monitoring in underground coal mines–review. *IETE Technical Review, 32*(2), 140–150.
Dohare, Y. S. et al. (2014). Design of surveillance and safety system for underground coal mines based on low power WSN. In *2014 International Conference on Signal Propagation and Computer Technology (ICSPCT 2014)*. IEEE.
Edwards, J., et al. (2000). Mine fire source discrimination using fire sensors and neural network analysis.
Feng, F., ShengYu, H., & Qi, X. (2010). The research of the Zigbee and RFID fusion technology in the coal mine safety. *2010 3rd international conference on information management, innovation management and industrial engineering*. IEEE.
Fiscor, S. (2008). Underground communications & tracking systems update. *Coal Age, 113*(1).
Fong-Sam, Y. et al. (2008). The mineral industries of Asia and the Pacific. *Minerals Yearbook*.
Forooshani, A. E. et al. (2013). A survey of wireless communications and propagation modeling in underground mines. *IEEE Communications Surveys & Tutorials, 15*(4), 1524–1545.
Geetha, A. (2013). Intelligent helmet for coal miners with voice over Zigbee and environmental monitoring. *Middle-East Journal of Scientific Research, 16*(12), 1835–1837.
Golizadeh, H. et al. (2018). Digital engineering potential in addressing causes of construction accidents. *Automation in Construction, 95*, 284–295.
Guo, P., Jiang, T., & Zhang, K. (2012). Novel energy-efficient miner monitoring system with duty-cycled wireless sensor networks. *International Journal of Distributed Sensor Networks, 8*(1), 975082.
Gyekye, S. A. (2003). Causal attributions of Ghanaian industrial workers for accident occurrence: Miners and non-miners perspective. *Journal of Safety Research, 34*(5), 533–538.
Hazarika, P. (2016). Implementation of smart safety helmet for coal mine workers. *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*. IEEE.
Henriques, V., & Malekian, R. (2016). Mine safety system using wireless sensor network. *IEEE Access, 4*, 3511–3521.
Hodgskiss, J., & Edwards, A. (2013). Job strain in the South African mining industry. *Occupational Health Southern Africa, 19*(5), 8–13.
Hongqing, Z. et al. (2011). Theory study on nonlinear control of ventilation network airflow of mine. *Procedia Engineering, 26*, 615–622.
Hrovat, A., Kandus, G., & Javornik, T. (2013). A survey of radio propagation modeling for tunnels. *IEEE Communications Surveys & Tutorials, 16*(2), 658–669.
Hu, S., Shu, H., & Song, X. (2013). Fisher information of mine collapse hole detection based on sensor nodes connectivity. *International Journal of Distributed Sensor Networks, 9*(9), 306496.
Huang, Y. et al. (2015). Robust localization algorithm based on the RSSI ranging scope. *International Journal of Distributed Sensor Networks, 11*(2), 587318.
Jabłoński, M., Szer, I., & Szer, J. (2018). Probability of occurrence of health and safety risks on scaffolding caused by noise exposure. *Journal of Civil Engineering and Management, 24*(6), 437–443.
Jiang, H. et al. (2009). A reliable and high-bandwidth multihop wireless sensor network for mine tunnel monitoring. *IEEE Sensors Journal, 9*(11), 1511–1517.
Karl, H., & Willig, A. (2007). Protocols and architectures for wireless sensor networks. John Wiley & Sons.
Keitel, A. et al. (2011). Endocrine and psychological stress responses in a simulated emergency situation. *Psychoneuroendocrinology, 36*(1), 98–108.
Kennedy, G. A., & Bedford, M. D. (2014). Underground wireless networking: A performance evaluation of communication standards for tunnelling and mining. *Tunnelling and Underground Space Technology, 43*, 157–170.
Kiziroglou, M. E. et al. (2016). Opportunities for sensing systems in mining. *IEEE Transactions on Industrial Informatics, 13*(1), 278–286.
Kloos, G. et al. (2006). Range based localisation using RF and the application to mining safety. *2006 IEEE/RSJ international conference on intelligent robots and systems*. IEEE.
Kowalski-Trakofler, K. M., Vaught, C., & Scharf, T. (2003). Judgment and decision making under stress: An overview for emergency managers. *International Journal of Emergency Management, 1*(3), 278–289.
Kumar, D. (2016). Application of modern tools and techniques for mine safety & disaster management. *Journal of The Institution of Engineers (India): Series D, 97*(1), 77–85.
Kumari, S., & Om, H. (2016). Authentication protocol for wireless sensor networks applications like safety monitoring in coal mines. *Computer Networks, 104*, 137–154.
Kwon, N. et al. (2018). Construction noise risk assessment model focusing on construction equipment. *Journal of Construction Engineering and Management, 144*(6), 04018034.
Laney, A. S., & Weissman, D. N. (2014). *Respiratory diseases caused by coal mine dust. Journal of occupational and environmental medicine/American College of Occupational and Environmental Medicine, 56*(0 10), S18.
Li, J. et al. (2019). Research on the effects of the high temperature and humidity environment on human comfort in coal mine emergency refuge system. *Safety, 5*(2), 28.
Li, M., & Liu, Y. (2007). Underground structure monitoring with wireless sensor networks. In *Proceedings of the 6th international conference on information processing in sensor networks*. ACM.
Li, C. T., Cheng, J. C., & Chen, K. (2020). Top 10 technologies for indoor positioning on construction sites. *Automation in Construction, 118*, 103309.
Li, Y., & Guldenmund, F. W. (2018). Safety management systems: A broad overview of the literature. *Safety Science, 103*, 94–123.
Li, M., & Liu, Y. (2009). Underground coal mine monitoring with wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN), 5*(2), 1–29.
Li-min, Y. et al. (2008). Design of monitoring system for coal mine safety based on wireless sensor network. *2008 IEEE/ASME international conference on mechtronic and embedded systems and applications*. IEEE.
Liu, Z. et al. (2010). A wireless sensor network based personnel positioning scheme in coal mines with blind areas. *Sensors, 10*(11), 9891–9918.
Liu, M., Han, S., & Lee, S. (2016). Tracking-based 3D human skeleton extraction from stereo video camera toward an on-site safety and ergonomic analysis. *Construction Innovation, 16*(3), 348–367.

Lu, X., & Davis, S. (2016). How sounds influence user safety decisions in a virtual construction simulator. *Safety Science, 86*, 184–194.

Luomala, J., & Hakala, I. (2015). Effects of temperature and humidity on radio signal strength in outdoor wireless sensor networks. *2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE.

Lutz, E. A. et al. (2015). Effectiveness evaluation of existing noise controls in a deep shaft underground mine. *Journal of Occupational and Environmental Hygiene, 12*(5), 287–293.

Mahdavipour, O. et al. (2015). Wireless sensors for automated control of total incombustible content (TIC) of dust deposited in underground coal mines. *2015 IEEE sensors*. IEEE.

Mahdevari, S., Shahriar, K., & Esfahanipour, A. (2014). Human health and safety risks management in underground coal mines using fuzzy TOPSIS. *Science of the Total Environment, 488*, 85–99.

Mahmood, A. et al. (2013). Data mining techniques for wireless sensor networks: A survey. *International Journal of Distributed Sensor Networks, 9*(7), 406316.

Maity, T., Das, P. S., & Mukherjee, M. (2012). A wireless surveillance and safety system for mine workers based on Zigbee. *2012 1st international conference on Recent Advances in Information Technology (RAIT)*. IEEE.

Maity, T., Sarathi Das, P., Mukherjee, M. (2012). *Rescue and protection system for underground mine workers based on Zigbee.*

Mannan, P., & Jayavignesh, T. (2016). Alternate simplistic approach to solve count-to-infinity problem by introducing a new flag in the routing table. *2016 second International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE.

Mardonova, M., & Choi, Y. (2018). Review of wearable device technology and its applications to the mining industry. *Energies, 11*(3), 547.

McBride, D. I. (2004). Noise-induced hearing loss and hearing conservation in mining. *Occupational Medicine, 54*(5), 290–296.

Meng, J. et al. (2014). Effects of fatigue on the physiological parameters of labor employees. *Natural Hazards, 74*(2), 1127–1140.

Milenković, A., Otto, C., & Jovanov, E. (2006). Wireless sensor networks for personal health monitoring: Issues and an implementation. *Computer Communications, 29*(13–14), 2521–2533.

Minhas, U. I. et al. (2017). A WSN for monitoring and event reporting in underground mine environments. *IEEE Systems Journal, 12*(1), 485–496.

Mishra, A., Malhotra, S., & Singh, H. (2018). Real time monitoring & analyzation of hazardous parameters in underground coal mines using intelligent helmet system. *2018 4th international conference on Computational Intelligence & Communication Technology (CICT)*. IEEE.

Misra, P. et al. (2010). Safety assurance and rescue communication systems in high-stress environments: A mining case study. *IEEE Communications Magazine, 48*(4), 66–73.

Moridi, M. A. et al. (2014). An investigation of underground monitoring and communication system based on radio waves attenuation using ZigBee. *Tunnelling and Underground Space Technology, 43*, 362–369.

Moridi, M. A. et al. (2015). Development of underground mine monitoring and communication system integrated ZigBee and GIS. *International Journal of Mining Science and Technology, 25*(5), 811–818.

Moridi, M. A. et al. (2018). Development of wireless sensor networks for underground communication and monitoring systems (the cases of underground mine environments). *Tunnelling and Underground Space Technology, 73*, 127–138.

Moridi, M. A. et al. (2019). An automated underground space monitoring and communication system based on wireless sensor networks. *International symposium on mine planning & equipment selection*. Springer.

Moutairou, M., Aniss, H., & Delisle, G. (2006). Wireless mesh access point routing for efficient communication in underground mine. *2006 IEEE antennas and propagation society international symposium*. IEEE.

MSHA (2011). Mine Safety and Health Administration, Program Policy Letter No. P11-V-13, April 28, 2011. MSHA.

Muduli, L., Jana, P. K., & Mishra, D. P. (2018). Wireless sensor network based fire monitoring in underground coal mines: A fuzzy logic approach. *Process Safety and Environmental Protection, 113*, 435–447.

Muduli, L., Mishra, D. P., & Jana, P. K. (2018). Application of wireless sensor network for environmental monitoring in underground coal mines: A systematic review. *Journal of Network and Computer Applications, 106*, 48–67.

Muduli, L., Mishra, D. P., & Jana, P. K. (2019). Wireless sensor network based underground coal mine environmental monitoring using machine learning approach. *Proceedings of the 11th international mine ventilation congress*. Springer.

Nerguizian, C. et al. (2005). Radio-channel characterization of an underground mine at 2.4 GHz. *IEEE Transactions on Wireless Communications, 4*(5), 2441–2453.

Nie, B.-S. et al. (2011). Internet of things-based positioning of coalmine personnel and monitoring of emergency state. *2011 second international conference on digital manufacturing & automation*. IEEE.

Nie, B. et al. (2016). Experimental study on physiological changes of people trapped in coal mine accidents. *Safety Science, 88*, 33–43.

Niu, X. et al. (2007). The design and evaluation of a wireless sensor network for mine safety monitoring. In *IEEE GLOBECOM 2007-IEEE global telecommunications conference*. IEEE.

Osunmakinde, I. O. (2013). Towards safety from toxic gases in underground mines using wireless sensor networks and ambient intelligence. *International Journal of Distributed Sensor Networks, 9*(2), 159273.

Outalha, S., Le, R., & Tardif, P. (2000). Toward a unified and digital communication system for underground mines. *CIM Bulletin*, 100–105.

Pantelopoulos, A., & Bourbakis, N. G. (2009). A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40*(1), 1–12.

Pawson, R. et al. (2005). Realist review-a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy, 10*(1_suppl), 21–34.

Perkins, C. E., & Royer, E. M. (1999). Ad-hoc on-demand distance vector routing. In *Proceedings WMCSA'99. Second IEEE workshop on mobile computing systems and applications*. IEEE.

Pudke, A. J., Bhagat, S. N., & Nalbalwar, S. (2017). LabVIEW based coal mine monitoring and alert system with data acquisition. *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE.

Qaraqe, K. A. et al. (2013). Statistical wireless channel propagation characteristics in underground mines at 900 MHz: A comparative analysis with indoor channels. *Ad Hoc Networks, 11*(4), 1472–1483.

Qiang, C. et al. (2009). ZigBee based intelligent helmet for coal miners. *2009 WRI world congress on computer science and information engineering*. IEEE.

Qinghua, Z. et al. (2009). Object position tracking based on e-map and RFID in coal mine. *2009 4th IEEE conference on industrial electronics and applications*. IEEE.

Qing-liang, S., Zhi-xian, Z., & Zhen-chuan, Z. (2008). Research on safety monitoring system of mine based on ad hoc technology. *2008 international conference on information management, innovation management and industrial engineering*. IEEE.

Qiuping, W., Shunbing, Z., & Chunquan, D. (2011). Study on key technologies of Internet of Things perceiving mine. *Procedia Engineering, 26*, 2326–2333.

Raffler, N. et al. (2018). Musculoskeletal disorders among occupational drivers caused by whole body vibration and awkward posture. *Congress of the international ergonomics association*. Springer.

Ranjan, A. et al. (2017). Studies on propagation characteristics of radio waves for wireless networks in underground coal mines. *Wireless Personal Communications, 97*(2), 2819–2832.

Ranjan, A. et al. (2019). Opportunities and challenges in health sensing for extreme industrial environment: Perspectives from underground mines. *IEEE Access, 7*, 139181–139195.

Ranjan, A., Sahu, H. B., & Misra, P. (2016). Wireless sensor networks: An emerging solution for underground mines. *International Journal of Applied Evolutionary Computation (IJAEC), 7*(4), 1–27.

Ranjan, A., Sahu, H. B., & Misra, P. (2018). DeepSense: Sensing the radio signal behavior in metal and non-metal underground mine workings. *IEEE INFOCOM 2018-IEEE conference on computer communications workshops (INFOCOM WKSHPS)*. IEEE.

Ranjan, A., Sahu, H. B., & Misra, P. (2019). MineSense: Sensing the radio signal behavior in metal and non-metal underground mines. *Wireless Networks, 25*(6), 3643–3655.

Ranjan, A., Sahu, H., & Misra, P. (2020). Modeling and measurements for wireless communication networks in underground mine environments. *Measurement, 149*, 106980.

Ranjany, A. et al. (2016). Channel modeling of wireless communication in underground coal mines. *2016 8th international conference on Communication Systems and Networks (COMSNETS)*. IEEE.

Ray, S. J., & Teizer, J. (2012). Real-time construction worker posture analysis for ergonomics training. *Advanced Engineering Informatics, 26*(2), 439–455.

Rempel, D. et al. (2019). Pneumatic rock drill vs. electric rotary hammer drill: Productivity, vibration, dust, and noise when drilling into concrete. *Applied Ergonomics, 74*, 31–36.

Revindran, R., Vijayaraghavan, H., & Huang, M.-Y. (2018). Smart helmets for safety in mining industry. *2018 international conference on advances in computing, communications and informatics (ICACCI)*. IEEE.

Reyes, J. P. et al. (2014). Health & Safety criteria for determining the sustainable value of construction projects. *Safety Science, 62*, 221–232.

Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News*.

Ruiz-Garcia, L. et al. (2009). A review of wireless sensor technologies and applications in agriculture and food industry: State of the art and current trends. *Sensors, 9*(6), 4728–4750.

Rupprecht, S. (2011). Safety considerations in underground logistics-a look at vertical, horizontal and in-stope transportation systems. *Journal of the Southern African Institute of Mining and Metallurgy, 111*(1), 45–53.

Sadeghi, S., Soltanmohammadlou, N., & Rahnamayiezekavat, P. (2021). A systematic review of scholarly works addressing crane safety requirements. *Safety Science, 133*, 105002.

Safe Work Australia (2020). Key work health and safety statistics Australia 2020: Work-related injury fatalities. Safe Work Australia

Savic, V., Wymeersch, H., & Larsson, E. G. (2013). Simultaneous sensor localization and target tracking in mine tunnels. *Proceedings of the 16th international conference on information fusion*. IEEE.

Schalk Wilhelm, P., & Reza, M. (2019). Survey on a smart health monitoring system based on context awareness sensing. *Communications of the CCISA, 1*, 25.

Sharma, M., & Maity, T. (2018). Low cost low power smart helmet for real-time remote underground mine environment monitoring. *Wireless Personal Communications, 102*(1), 149–162.

Singh, A., Kumar, D., & Hötzel, J. (2018). IoT based information and communication system for enhancing underground mines safety and productivity: Genesis, taxonomy and open issues. *Ad Hoc Networks, 78*, 115–129.

Soltanmohammadlou, N. et al. (2019). Real-time locating systems and safety in construction sites: A literature review. *Safety Science, 117*, 229–242.

Stanek, E. K. (1988). Mine electrotechnology research: The past 17 years. *IEEE transactions on industry applications, 24*(5), 818–819.
Sun, E. et al. (2010). An integrated information technology assisted driving system to improve mine trucks-related safety. *Safety Science, 48*(10), 1490–1497.
Sun, Z., & Akyildiz, I. F. (2010). Channel modeling and analysis for wireless networks in underground mines and road tunnels. *IEEE Transactions on Communications, 58*(6), 1758–1768.
Sunderman, C., & Waynert, J. (2012). An overview of underground coal miner electronic tracking system technologies. *2012 IEEE industry applications society annual meeting*. IEEE.
Swain, A. et al. (2018). Zigbee based wireless data acquisition system for underground mines—A feasibility study. *2018 IEEE 21st International Symposium on Real-Time Distributed Computing (ISORC)*. IEEE.
Tan, W. et al. (2007). Mine fire detection system based on wireless sensor network. *2007 international conference on information acquisition*. IEEE.
Thrun, S. et al. (2001). Robust Monte Carlo localization for mobile robots. *Artificial Intelligence, 128*(1–2), 99–141.
Trutwin, W. (1988). On transients of airflow in mine workings. *Mining Science and Technology (Netherlands), 6*(3).
Umar, S. et al. (2018). WSN based alter system to save life in mines with high security system. *2018 2nd International Conference on Inventive Systems and Control (ICISC)*. IEEE.
Umer, W. et al. (2012). The prevalence of musculoskeletal disorders in the construction industry: A systematic review. *International Archives of Occupational and Environmental Health, 138*(12), 1370–1376.
Utembe, W. et al. (2015). Hazards identified and the need for health risk assessment in the South African mining industry. *Human & Experimental Toxicology, 34*(12), 1212–1221.
Varga, J. et al. (2016). Physiological strain in the Hungarian mining industry: The impact of physical and psychological factors. *International Journal of Occupational Medicine and Environmental Health, 29*(4), 597.
Wang, X. et al. (2007). *Deploying a wireless sensor network on the coal mines2007 IEEE international conference on networking, sensing and control*. IEEE.
Wang, Q.-F. et al. (2009). The application of wireless sensor networks in coal mine. *2009 7th International Conference on Information, Communications and Signal Processing (ICICS)*. IEEE.
Wang, H. et al. (2011). An emergency communication architecture based on ethernet and wsn for coal mines. *Procedia Engineering, 23*, 403–407.
Wang, D., Dai, F., & Ning, X. (2015). Risk assessment of work-related musculoskeletal disorders in construction: State-of-the-art review. *Journal of Construction Engineering and Management, 141*(6), 04015008.
Wang, Y., Huang, L., & Yang, W. (2010). A novel real-time coal miner localization and tracking system based on self-organized sensor networks. *EURASIP Journal on Wireless Communications and Networking, 2010*, 1–14.
Wei, S., & Li-Li, L. (2009). Multi-parameter monitoring system for coal mine based on wireless sensor network technology. *2009 international conference on industrial mechatronics and automation*. IEEE.
Winn, F. J., Jr, Biersner, R. J., & Morrissey, S. (1996). Exposure probabilities to ergonomic hazards among miners. *International Journal of Industrial Ergonomics, 18*(5–6), 417–422.
Wu, L. et al. (2011). Major accident analysis and prevention of coal mines in China from the year of 1949 to 2009. *Mining Science and Technology (China), 21*(5), 693–699.
Wu, Y. et al. (2019). A dynamic information platform for underground coal mine safety based on internet of things. *Safety Science, 113*, 9–18.
Wu, H., & Gillies, A. (2005). Real-time airflow monitoring and control within the mine production system. *Eighth international mine ventilation congress*.
Xu, X. et al. (2012). Design of underground miner positioning system based on zigbee technology. *International conference on web information systems and mining*. Springer.
Xu, X. et al. (2017). Occupational hazards survey among coal workers using hand-held vibrating tools in a northern China coal mine. *International Journal of Industrial Ergonomics, 62*, 21–26.
Yan, W., Ya-ru, Z., & Yong, M. (2008). Study on the coal mine personnel position system based on wireless body sensor networks. *2008 5th International summer school and symposium on medical devices and biosensors*. IEEE.
Yang, W., & Huang, Y. (2007). Wireless sensor network based coal mine wireless and integrated security monitoring information system. In *Sixth international conference on networking (ICN'07)*. IEEE.
Yang, W., Zhang, Y., & Liu, Y. (2010). Constructing of wireless emergency communication system for underground coal mine based on WMN technology. *Journal of Coal Science and Engineering (China), 16*(4), 441–448.
Yassierli, I. (2017). of ergonomic programs to reduce sick leave due to low back pain among nickel mining operators. *International Journal of Industrial Ergonomics, 61*, 81–87.
Yin, G. (2011). The design and optimization of underground scraper wireless monitoring networks based on ZigBee. *Physics Procedia, 22*, 270–274.
Yong, M. et al. (2009). Optimization design of coal mine wireless body sensor network based on Genetic Algorithm. *2009 international conference on networks security, wireless communications and trusted computing*. IEEE.
Yu, H., Chen, H., & Long, R. (2017). Mental fatigue, cognitive bias and safety paradox in chinese coal mines. *Resources Policy, 52*, 165–172.
Zeng, S., Tam, V. W., & Tam, C. M. (2008). Towards occupational health and safety systems in the construction industry of China. *Safety Science, 46*(8), 1155–1168.
Zhang, Y. et al. (2014). An integrated environment monitoring system for underground coal mines—Wireless sensor network subsystem with multi-parameter monitoring. *Sensors, 14*(7), 13149–13170.
Zhao, C., Liu, F., & Hai, X. (2013). An application of wireless sensor networks in underground coal mine. *International Journal of Future Generation Communication and Networking, 6*(5), 117–126.
Zhou, C. et al. (2015). RF Propagation in Mines and Tunnels: Extensive measurements for vertically, horizontally, and cross-polarized signals in mines and tunnels. *IEEE Antennas and Propagation Magazine, 57*(4), 88–102.
Zhu, Y., & You, G. (2019). Monitoring system for coal mine safety based on wireless sensor network. *2019 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC)*. IEEE.
Ziętek, B. et al. (2020). A portable environmental data-monitoring system for air hazard evaluation in deep underground mines. *Energies, 13*(23), 6331.

Sanaz Sadeghi holds a B.Sc. in Architectural Engineering at University of Guilan and M.Sc. in Historic Preservation at University of Art in Iran. She has experience in Occupational Health and Safety, with emphasis on Construction Safety and health Management using sensing technology.

Nazi Soltanmohammadlou holds a B.sc. in Civil Engineering and M.Sc. in Restoration and Rehabilitation of Historical Buildings and Urban Fabrics at University of Art in Iran. She has been working as a licensed civil engineer from 2015 to present. She has also started her investigations in Construction Safety and health Management since 2017 and published some papers in highly-ranked peer reviewed journals.

Dr. Farnad Nasirzadeh is a Senior Lecturer in Construction Management at Deakin University. His main research interest is modelling and simulation of complex systems with a specific focus on workforce performance. He has combined simulation and sensing in his previous research to improve workforce productivity and safety.

Journal of Safety Research 83 (2022) 410–417

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Applying systems thinking to improve the safety of work-related drivers: A systematic review of the literature

Sharon Newnam *, Renee St Louis, Amanda Stephens, Dianne Sheppard

Monash University Accident Research Centre, 21 Alliance Lane, Monash University, VIC 3800, Australia

A R T I C L E I N F O

Article history:
Received 10 March 2022
Received in revised form 14 June 2022
Accepted 22 September 2022
Available online 4 October 2022

Keywords:
Work-related driving
Systems thinking
Systematic review
Risk factors
Crashes

A B S T R A C T

Introduction: Light vehicles (<4.5 tons) driven for work purposes represent a significant proportion of the registered motor vehicles on our roads. Drivers of these vehicles have significant exposure to the dangers of the road transport environment. To optimize safety for these workers, it is critical to understand the factors contributing to risk of being involved in an incident. This information can then be used to inform the review and revision of existing risk controls and the development of targeted prevention activities. *Method:* The aim of the study was to undertake a systematic review of the literature to identify the factors associated with work-related driving incidents. The factors identified in the review were represented within an adapted version of Rasmussen's risk management framework (Rasmussen, 1997). Fifty studies were analyzed following data screening and review of full text. The highest proportion of risk factors were categorized at the lower levels of the system, including the 'Drivers and Other Road Users' level (n = 20, 44.4%) and the 'Equipment, Environment, and Meteorological Surroundings' level (n = 19, 42.2%). There were no risk factors identified at the 'Regulatory and Government Bodies' levels of the framework, confirming the narrow investigative scope of past research and the need to acknowledge a broader range of factors within and across higher levels of the system. *Conclusions:* The findings of this study inform the direction of future research and design of targeted prevention activities capable of creating system change for the safety of work-related drivers.

1. Introduction

Road freight is a safety critical industry and has the highest death rate of its employees compared to that of other industries (Safe Work Australia, 2018). While much research has focused on vehicles over 4.5 tons, smaller vehicles (e.g., passenger vehicles, utility vans) also represent a significant public health issue. However, limited attention has focused in this area given the challenges associated with collecting data on the 'purpose of the journey' of a road traffic incident (i.e., work or personal purposes; Newnam et al., 2014). Regulators do not routinely collect data specifying whether a crash occurred when driving for work or personal purposes.

Despite this, prevention activities are emerging to manage the risks associated with those who drive a light vehicle for work-related purposes. In Australia, vehicles driven for work purposes represent 30% of the registered motor vehicles in Australia, with some drivers reporting travelling over 1,100 kilometers per week (Zurich Insurance, 2015). The risk associated with exposure to the road transport environment is evidenced, globally. To illustrate, a total of 1,270 U.S. workers driving or riding in a motor vehicle for work-related purposes on a public road died in 2019 (representing 24% of all work-related deaths; NIOSH, 2022). Moreover, 56% of these workers who died were not employed in a motor-vehicle operator job; rather, driving was considered a secondary task to their primary job role (e.g., in-home nursing care, sales representatives; Newnam, Lewis, & Watson, 2012). This issue creates some challenges in managing the safety and balancing tensions with competing priorities (i.e., efficiency and productivity).

Managing the safety of these workers is further challenged because, unlike the road freight transport industry, a 'Chain of Responsibility' does not exist for managing the safety of workers who operate a light vehicle. Thus, there is limited guidance in the roles and responsibilities of those responsible for managing the safety of workers that operate a light vehicle, beyond what is specified in Occupational Health and Safety legislation. The complexity of this issue is compounded when there is no single government body or department responsible for managing the allocation of resources for road safety outcomes or are tasked with managing data and monitoring road safety issues (Newnam & Muir, 2021). This is even the case in countries where the national road safety

* Corresponding author.
E-mail address: sharon.newnam@monash.edu (S. Newnam).

https://doi.org/10.1016/j.jsr.2022.09.016
0022-4375/

strategy and associated legislation has adopted the Safe System concept, such as Australia (Muir, Johnston, & Howard, 2018). Given the sheer number of stakeholders capable of influencing change within this dynamic environment, and no central point of responsibility, it is not surprising that limited lessons and evidence-based best practice approaches specific to driving light vehicles for work-related purposes have been established for preventing associated road safety incidents.

Indeed, the lack of ability to learn from crashes or near crashes is a critical barrier to improving the safety of this workforce. Warmerdam et al. (2017) interviewed employees across 79 workplaces that employ individuals to drive a light vehicle for work-related purposes across two states in Australia and identified that few have practices in place for investigating incidents involving a work vehicle. Rather, incidents (e.g., crashes or near crashes) involving a light vehicle are investigated by the companies that insure the vehicle, not the employer. The limitation with this approach is that motor-vehicle insurers use a narrow investigative scope, as a driver interview is used as the primary source of information. This means that investigations are focused mainly on the role of the driver and their actions at the time of the crash. Drivers are often given little reason or opportunity to reflect upon any organizational or external factors that may contribute to crashes, such as vehicle maintenance, scheduling, and regulatory restrictions. Furthermore, there is often limited consultation with other key stakeholders in the system (e.g., fleet managers, supervisors; Newnam, Griffin, & Mason, 2008) that could provide insight into the broader system of factors that contributed to risk in any work-related driving incident.

Historically, crash investigation for heavy vehicle crashes has been described as insufficient for learning and developing appropriate control measures (Newnam & Goode, 2015; Newnam, Warmerdam, Sheppard, Griffin, & Stevenson, 2017), and as such, there is also little substantive learning for light vehicles that can be transferred from investigation processes undertaken in the road freight transport industry. Again, these investigations primarily focus on driver-level factors such as driver characteristics (e.g., age, gender) and behavior (e.g., inappropriate speed, fatigue, and drug use). These types of toolkits imply drivers are to "blame" for crashes, ignoring the broader system of factors influencing crash involvement.

The lack of systematic and rigorous investigation of system and organizational-level circumstances of individual crash incidents involving light or heavy vehicles is an impediment to progressing the safety improvements needed to ensure worker and public safety on roads. Reductionist-focused incident investigation models and methods have also been identified as inadequate across other safety critical industries, including healthcare (Newnam, Goode, Read, & Salmon, 2020; Newnam, Goode, Read, Salmon, & Gembarovski, 2021). More consistent with current thinking, a systems-thinking approach (Rasmussen's risk management framework and the associated Accimap technique; Rasmussen, 1997) is required as a first step to better understand these incidents, followed by a review and revision of existing risk controls to develop feasible, effective, and practicable control measures.

In other high-risk industries (e.g., healthcare), systems-thinking models and analysis methods now represent an accepted approach for optimizing safety activities (Cassano-Piche, Vicente, & Jamieson, 2009; Goode, Salmon, Lenne, & Finch, 2018; Hulme, Stanton, Walker, Waterson, & Salmon, 2019; Newnam et al., 2020; Newnam et al., 2021). These models and methods are underpinned by the idea that incidents occur due the interaction between multiple factors across a system (Leveson, 2011; Rasmussen, 1997). The behavior of the individual-worker, the equipment used to complete the work task, and the safety practices of employers are only some of the factors that need to be considered in an incident investigation. To illustrate this type of investigation tool, Newnam et al. (2020) developed the Patient Handling Injuries Review of Systems (PHIRES) tool to help guide practitioners in the healthcare sector in a system-thinking investigation following the report of a musculoskeletal injury to staff associated with patient handling. The tool is underpinned by the systems-thinking approach, Rasmussen's Risk Management framework, and the associated Accimap technique (Rasmussen, 1997; Svedung & Rasmussen, 2002). The multiple work systems, represented as hierarchical levels, were adapted in the PHIRES tool to represent the healthcare system. A classification scheme was developed to describe the work-related and societal factors, in addition to the physical factors, typically associated with increased risk relating to the work task of patient handling, and subsequently represented at each level of the healthcare system. These factors were identified through a systematic review of the literature and in consultation with key stakeholders in the industry.

Thus, there is much that can be learned from previous research to move toward improved prevention of work-related driving incidents. Systems thinking models (i.e., Rasmussen's risk management framework (1997) are needed to best understand the factors associated with the risk of work-related driving incidents. The first step in creating systemic change in prevention activities is to identify the range of factors contributing to work-related driving incidents. Such an approach is critical to move beyond the current reductionist thinking and towards a more comprehensive understanding of the system of factors contributing to crashes. Improving the capture of data related to risk in work-related driving will inform the development of targeted prevention activities, including creating a culture where responsibility for safety is shared across the system.

The aim of the current study is to undertake a systematic review of the literature to identify the system of factors associated with work-related (light or heavy) vehicle driving incidents. The factors identified in the review will be represented on Rasmussen's risk management framework (Rasmussen, 1997). This framework has been adapted to align with the typical system that employs individuals that operate a vehicle for work-related purposes and has also drawn upon learnings from the road freight transportation system (Newnam, Goode, Salmon, & Stevenson, 2017). The five levels of the system are described in Table 1.

2. Method

A systematic review of the literature was undertaken, guided by PRISMA guidelines, to identify factors contributing to work-related driving incidents, which were defined as crashes and near crashes (i.e., near misses). A comprehensive list of search terms was developed to guide the search using the categories: (i) *primary context*, including workplace (i.e. workplace, work-related, occupation*, vocation*, professional) AND driving (driv*, transport, fleet, vehicle*, commercial), AND injury/incident (injur* (NOT chemical), safety, risk); (ii) *outcome* focused terms (e.g. crash*, accident*, ticket*, fine*, penalty, infringement*, near miss*, loss of control); and (iii) *Other* terms to help to limit/refine the scope of the literature to papers with a focus on factors contributing to such incidents (e.g. caus*, contrib*, predict*, risk factor*, determin*, predict*).

The search was restricted to journal articles published from 2010 through 2021. Six databases were used to conduct the search (Medline, PubMed, AMED, Scopus, PsycINFO and Web of Science). Studies that identified the relationship between work-related driving crashes for *both light and heavy vehicles* were included to expand the scope of knowledge.

Table 1
Hierarchical levels of the system of factors contributing to work-related vehicle incidents (adapted from Rasmussen's Risk Framework, 1997).

Government, Regulators & External Influences	Factors external to the organization relating to laws governing safe working practices. This level also considers factors associated with external influencers (media reporting, social media, community attitudes).
Governance & Administration	Factors associated with personnel working for companies, as well as policies and guidelines that regulate work practices.
Operations Management	Factors associated with the employer and different levels of management personnel (e.g., supervisor, fleet manager). Factors at this level typically occur prior to the incident but can also include decisions and actions made during, or in response to, the incident. Contributory factors related to policy, planning and budgeting typically occur well before the crash itself, and may even exist years before the crash occurred.
Drivers & Other Road Users	Factors contributing to the incident prior to, and during, the crash. This level includes factors related to actors directly involved in the operation of the vehicle (including passengers) as well as other actors at the scene of the crash (e.g. other drivers).
Equipment, Vehicle & Surrounding Environment	Factors associated with the vehicle and equipment (e.g., in-vehicle telemetry), the physical road environment (e.g., road surface conditions), and the ambient and meteorological conditions prior to or during the crash.

The search strategy is outlined in Fig. 1. The initial search resulted in 346 articles that were imported into EndNote. Duplicates were then identified and deleted (n = 183), leaving 163 articles that were examined in the title and abstract screening stage. Two authors (AS, RS) independently screened approximately 60% of the titles and abstracts (n = 98) for potentially relevant articles, and reached 97% agreement. A third author (SN) made the final decision for the remaining three titles and abstracts. One author (RS) completed the title and abstract screening for the remaining 65 articles. The title and abstract screening stage resulted in the exclusion of 100 articles for reasons including that the study was not focused on work-related driving, no risk factors were identified, or the outcome variable was not relevant to crashes or crash risk. Sixty-three articles were retained for the full text stage. These articles were independently reviewed by two authors (AS, RS), resulting in 95% agreement. A third author (SN) made the final decision for the remaining three articles. Thirteen articles were excluded during this stage (see reasons for exclusion in Fig. 1), resulting in 50 articles being retained for the data extraction stage.

Data extracted from each of the 50 articles in the final sample included: industry; country/region in which the study was conducted; employee cohort; outcome variables; and risk factors (mapped onto systems thinking classification scheme). Consistent with the aim of this study, all risk factors identified in the articles were categorized at a level of the system irrespective of the quality of the study or statistical significance with the outcome variable.

3. Results

Of the 50 articles included in this review, the most common industries represented were road freight transportation (n = 18, 36.0%) and farming/agriculture (n = 10, 20.0%). The taxi and bus industries each accounted for 14.0% of the articles included (n = 7 each), followed by emergency services (n = 4, 8.0%), delivery riders (n = 2, 4.0%), and mining (n = 1, 2.0%). There were four articles (8.0%) that did not specify a particular industry. Note that the total sum by industry is greater than 50 due to some articles including more than one industry type. The employee cohort in each study consisted of a combination of light and heavy vehicle employees driving for work-related purposes within the aforementioned

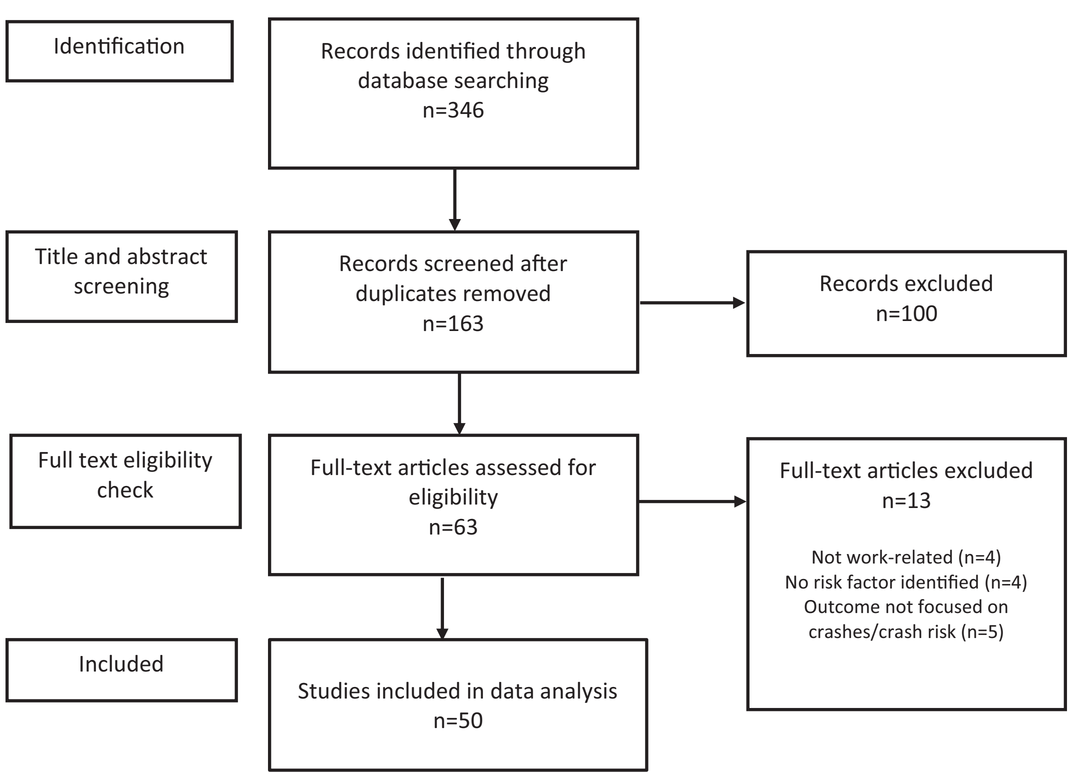

Fig. 1. Flow chart of the systematic search.

industries (e.g., road freight drivers, farmers, bus drivers), and emergency services (i.e., police officers, firefighters, and ambulance drivers).

The largest proportion of studies were from the United States (n = 19, 38.0%). The remaining articles represented diverse countries/regions around the world including Asia (n = 15, 15.0%), Australia/New Zealand (n = 6, 12.0%), Europe (n = 3, 6.0%), Iran (n = 3, 6.0%), Africa (n = 2, 4.0%), and South America (n = 2, 4.0%).

A work-related driving incident was the outcome of interest when conducting the search. As such, dependent variables included crashes (n = 29, 58.0%), injury severity (n = 8, 16.0%), crash risk (n = 6, 12.0%), near crashes (n = 5, 10.0%), loss of control events (n = 3, 6.0%), unsafe driver actions (n = 1, 2.0%) and aberrant driving behavior (n = 1, 2.0%).

A total of 45 risk factors were identified by the systematic review. Each risk factor was mapped onto the relevant level of an adapted version of Rasmussen's risk management framework (Rasmussen, 1997). Table 2 shows that the highest proportion of risk factors were categorized at the Drivers and Other Road Users level (n = 20, 44.4%), followed closely by risk factors at the Equipment, Environment, and Meteorological Surroundings level (n = 19, 42.2%). There were no risk factors identified at the Government, Regulatory and External Influencers level of the framework.

A description of the risk factors identified at the three lower levels of the system follows (with associated reference). The risk factors are sub-categorized and the corresponding articles in which the risk factors were identified are referenced. The number of risk factors identified within each article ranged from one to eight. Across all levels of the system, the most commonly cited risk factors were road design (n = 13), fatigue/sleepiness (n = 11), and traffic violations (previous history) (n = 10).

Table 3 shows the risk factors (n = 19) identified at the Equipment, Environment and Meteorological Surroundings level. This level encompasses a range of factors related to features and design of the vehicle (n = 8 risk factors), the road environment the time of year (n = 8 risk factors) and meteorological conditions (n = 3 risk factors). Risk factors categorized as 'Environment' were identified in the greatest number of articles overall at this level. Road design (n = 13), time of day/week (n = 8), and road surface conditions (n = 6) were the most commonly identified risk factors across the articles at this level.

Table 4 describes the risk factors identified at the Drivers and Other Road Users level. This level encompasses a broad range of factors related to the safe operation of the vehicle by the driver, (n = 17 risk factors) including several factors related to the physical and mental state of the driver (n = 1 risk factors), as well as design of the work environment (n = 2 risk factors). Seventeen risk factors were categorized within the category of Driver and represented the greatest number of articles overall at this level. Fatigue (n = 11), traffic violations (n = 10), and driving behavior (n = 9) were identified as risk factors in the greatest number of articles within this level of the system. Physical/medical condition (n = 8) and driver experience/competence (n = 6) were also frequently cited risk factors contributing to work-related driving incidents.

Table 5 describes the risk factors identified at the Companies and Employers level and encompassed a range of factors related to leadership (n = 2 risk factors) and work scheduling (n = 4 risk factors). Rostering (i.e., assignment of employees to a duty schedule; work scheduling, n = 7) was the most frequently identified risk factor across all articles at this level; however, several articles that identified rostering as a risk factor also identified another risk factor at this level. This manifests as an overlap of articles for these risk factors and demonstrates that these factors are likely closely related to each other. Leadership includes two risk factors that are related to the culture of the workplace, including organiza-

Table 2
Risk factors categorized at the three levels of the framework.

Level	Number of Risk Factors	%
Equipment, Environment and Meteorological Surroundings	19	42.2
Drivers and Other Road Users	20	44.4
Operations Management	6	13.4
Governance and Administration	0	0
Government, Regulators & External Influences	0	0

Table 3
Risk factors identified at the Equipment, Environment, and Meteorological Surroundings level.

Level of system	Risk factors
Equipment	Lack of warning signals (Missikpode, Peek-Asa, Young, & Hamann, 2018; Wang, Zhang, Li, & Liang, 2019)
	In-vehicle technology (Stevenson et al., 2014)
	Vehicle specifications (Chen & Zhang, 2016; Lemp, Kockelman, & Unnikrishnan, 2011)
	Design of vehicle (Haq, Zlatkovic, & Ksaibati, 2020; Milosavljevic et al., 2011)
	Lack of maintenance (Wang & Prato, 2019)
	Road signage (Chu, 2012, 2016; Mehlhorn, Wilkin, Darroch, & D'Antoni, 2015; Ramirez et al., 2016)
	Load/storage (Lemp et al., 2011; Shipp, Vasudeo, Trueblood, & Garcia, 2019; Stevenson et al., 2014)
	Lack of or inappropriate personal protective equipment (Mitchell, Bambach, & Friswell, 2014)
Meteorological surroundings	Lighting (Haq et al., 2020; Lemp et al., 2011; Ramirez et al., 2016; Useche, Cendales, Alonso, & Montoro, 2020)
	Weather conditions (Chen & Zhang, 2016; Chu, 2016; Das, Islam, Dutta, & Shimu, 2020; Haq et al., 2020; Lemp et al., 2011; Mehlhorn et al., 2015; Missikpode et al., 2018; Stevenson et al., 2014; Wang & Prato, 2019; Wang, Zhang, et al., 2019)
	Visibility (Chen & Zhang, 2016)
Environment	Road surface conditions (Besharati & Kashani, 2018; Chen & Zhang, 2016; Milosavljevic et al., 2011; Missikpode et al., 2018; Mitchell et al., 2014; Useche, Cendales, Alonso, & Montoro, 2020)
	Urban/rural (Chu, 2012; Das et al., 2020; Harland, Bedford, Wu, & Ramirez, 2018; Missikpode et al., 2018; Mitchell et al., 2014)
	Road furniture (Chu, 2012; Mehlhorn et al., 2015)
	Time of day/week (Chen & Zhang, 2016; Das et al., 2020; Harland et al., 2018; Mehlhorn et al., 2015; Useche, Cendales, Alonso, & Montoro, 2020; Wang & Prato, 2019; Zhang et al., 2017; Zuzewicz, Konarska, & Luczak, 2010)
	Traffic congestion (Das et al., 2020; Lemp et al., 2011)
	Season of the year (Chen & Zhang, 2016; Zhang et al., 2017)
	Road design (Carman et al., 2010; Chen & Zhang, 2016; Chu, 2012; Das et al., 2020; Gorucu, Murphy, & Kassab, 2017; Haq et al., 2020; Lemp et al., 2011; Mehlhorn et al., 2015; Missikpode et al., 2018; Mitchell et al., 2014; Ranapurwala, Mello, & Ramirez, 2016; Stuckey, Glass, LaMontagne, Wolfe, & Sim, 2010; Wang & Prato, 2019)
	Speed limit (Chu, 2012, 2016; Das et al., 2020)

Table 4
Risk factors identified at the Drivers and Other Road Users level.

Level of system	Risk factors
Work design	Job demands (Mamo, Newnam, & Tulu, 2014; Useche, Cendales, Alonso, & Orozco-Fontalvo, 2020; Zheng, Ma, Guo, Cheng, & Zhang, 2019)
	Safety culture (Mamo et al., 2014)
Drivers	Aggression (Harland et al., 2018; Lemp et al., 2011; Wang, Zhang, et al., 2019)
	Inattention/distractions (Chu, 2016; Harland, Carney, & McGehee, 2016)
	Alcohol/drugs (Haq et al., 2020; Harland et al., 2018; Lemp et al., 2011; Mitchell et al., 2014; Newnam, Blower, Molnar, Eby, & Koppel, 2018)
	Personality traits (Clay, Treharne, Hay-Smith, & Milosavljevic, 2014; Mallia, Lazuras, Violani, & Lucidi, 2015)
	Safety attitudes (Nickenig Vissoci et al., 2020; Sun & Tian, 2018)
	Physical/medical condition (Anderson et al., 2012; Barger et al., 2015; Besharati & Kashani, 2018; Das et al., 2020; Haq et al., 2020; Milosavljevic et al., 2011; Thiese et al., 2017; Zhang et al., 2017)
	Driving behaviour (Ba, Zhou, & Wang, 2018; Chen & Zhang, 2016; Chu, 2012; Nickenig Vissoci et al., 2020; Shams, Mehdizadeh, & Khani Sanij, 2020; Shin, Park, & Jeong, 2018; Useche, Cendales, Alonso, & Orozco-Fontalvo, 2020; Wang, Li, & Prato, 2019; Zuzewicz et al., 2010)
	Experience/competence (Carman et al., 2010; Chen & Zhang, 2016; Stevenson et al., 2014; Wang & Prato, 2019; Zheng et al., 2019; Zuzewicz et al., 2010)
	Hazard perception skill (Besharati & Kashani, 2018; Sun & Tian, 2018)
	Seat belt (Haq et al., 2020; Newnam et al., 2018; Shipp et al., 2019; Stuckey et al., 2010)
	Drugs/medication (Ogeil et al., 2018; Reguly, Dubois, & Bedard, 2014)
	Risk perceptions (Clay et al., 2014; Shams et al., 2020; Zheng et al., 2019)
	Mobile phone use (Ba et al., 2018)
	Fatigue/Sleepiness (Ba et al., 2018; Besharati & Kashani, 2018; Chen & Zhang, 2016; Haq et al., 2020; Kim, Jang, Kim, & Lee, 2018; Mitchell et al., 2014; Shin et al., 2018; Stuckey et al., 2010; Wang, Li, et al., 2019; Wang, Zhang, et al., 2019; Zhang et al., 2017)
	Traffic violations (Chu, 2012, 2016; Mallia et al., 2015; Mehdizadeh, Shariat-Mohaymany, & Nordfjaern, 2019; Nik Mahdi, Bachok, Mohamed, & Shafei, 2014; Reguly et al., 2014; Shams et al., 2020; Shipp et al., 2019; Wang, Zhang, et al., 2019; Zhang et al., 2017)
	Speed (Chu, 2016; Milosavljevic et al., 2011; Mitchell et al., 2014; Newnam et al., 2018; Stuckey et al., 2010)
	Sleep quality (Nik Mahdi et al., 2014; Shams et al., 2020)
Other road users	Behavior: general (Gorucu et al., 2017; Shipp et al., 2019)

Table 5
Risk factors identified at the Operations Management level.

Level of system	Risk factors
Leadership	Mental health/wellbeing/OHS (Baba, Miyama, Sugiyama, & Hitosugi, 2019; Sun & Tian, 2018)
	Safety culture (Sun & Tian, 2018)
Work scheduling	Rostering (Besharati & Kashani, 2018; Kim et al., 2018; Mehdizadeh et al., 2019; Nik Mahdi et al., 2014; Torregroza-Vargas, Bocarejo, & Ramos-Bonilla, 2014; Wang & Wu, 2019; Zheng et al., 2019)
	Shift work (Besharati & Kashani, 2018; Ogeil et al., 2018; Stevenson et al., 2014; Wang & Wu, 2019)
	Breaks (Baba et al., 2019; Chen & Xie, 2014; Stevenson et al., 2014; Torregroza-Vargas et al., 2014)
	Workload (Ba et al., 2018; Wang, Li, et al., 2019; Zheng et al., 2019)

tional policies regarding health, safety, and wellbeing of employees within the company.

4. Discussion

This goal of this study was to establish a systems-perspective evidence-base to better understand the range of factors contributing to work-related vehicle driving incidents. This goal was achieved through undertaking a systematic review of the literature to identify the factors contributing to incidents using a systems perspective. To do this, the factors identified in the systematic review were mapped onto Rasmussen's Risk Management framework (1997). The findings of this study address a gap in current knowledge of the system of factors contributing to work-related driving incidents. This information is important to inform the direction of future research and design of targeted prevention activities.

This study found that most factors were identified at the 'Drivers and Other Road Users' and 'Equipment, Environment and Meteorological Surroundings' levels. This finding is not surprising considering that existing data collection methods use a narrow investigative scope, focusing primarily on the actions of the driver, the vehicle, and the immediate environment surrounding the incident. While it is critical to capture this information, it is equally as important to acknowledge a broader range of factors within and across other levels of the system that have contributed to the likelihood of the crash, potentially in the weeks or months leading up to the work-related driving incident.

To illustrate, there is research to support the argument that a work-related drivers' engagement in inappropriate speed is influenced by higher-level factors such as work pressure (Newnam, Greenslade, Newton, & Watson, 2011), organizational systems and practices (Newnam, Warmerdam, et al., 2017), and the priority and value given to safety in the workplace (or lack thereof; Newnam et al., 2008). Many of these middle-level factors were identified in the results of this study. However, there were no risk factors identified at the Regulatory and Government Bodies levels of the framework. We know that in some countries (i.e., Australia, South Africa, Canada, New Zealand) that responsibility for safety has been allocated to actors at these higher levels for some forms of transportation; for example, Chain of Responsibility legislation in Australia is used to define the roles and responsibilities of actors involved in the heavy vehicle road transport system. Addressing this gap in scientific knowledge presents an opportunity for future research to better understand the influence of regulatory and government bodies in light vehicle work-related driving incidents and areas where they can mitigate risk and improve consultation across levels of the system.

This learning could be achieved through development of a system thinking incident investigation tool designed to guide practitioners in identifying risk factors associated with work-related driving crashes. As established in previous research (Newnam et al., 2020; Newnam et al., 2021), such a tool would provide a comprehensive and standardized approach to identifying targeted prevention activities and creating a shared responsibility for safety across the system; that is, prevention activities focused beyond the lower levels of the system and focused on creating systemic change as opposed to isolated change to individual elements of the system (e.g., speed enforcement). It is also possible that the findings from this tool could be used to develop Chain of Responsibility legislation for the use of light work-related vehicles.

5. Limitations

A potential limitation of the current study is that the systematic review did not include a review of the grey literature. We have learned through the development of system thinking investigation tools (Newnam et al., 2020, 2021) that there are factors at other levels of the system not yet identified in the academic literature due to the historically narrow focus. Thus, future research should ensure that the findings of this study are supplemented with information gained through a scan of the grey literature, as well as knowledge from subject matter experts, to provide a comprehensive understanding of risk factors associated with work-related driving incidents. This information would provide a strong foundation for informing the review and revision of current risk controls and the development of targeted prevention activities focused on creating systemic change.

6. Conclusions

The findings of this study address a gap in current knowledge that has inhibited prevention activities to improve the safety of work-related drivers. Although this study identified that the scope of knowledge on risk factors associated with work-related driving incidents is reductionist, the findings present an avenue for future research to address these gaps. Designing targeted prevention activities focused on sharing the responsibility of safety across the system could be achieved through improving the capture of data. The findings of this study present the first step in development of a system thinking tool that comprehensively captures the range of factors that should be considered in the investigation of light vehicle work-related driving crashes.

Declarations of interest

None.

Acknowledgement

This research was funded through the Institute for Safety, Compensation and Recovery Research development with funding provided by WorkSafe Victoria.

References

Anderson, J. E., Govada, M., Steffen, T. K., Thorne, C. P., Varvarigou, V., Kales, S. N., & Burks, S. V. (2012). Obesity is associated with the future risk of heavy truck crashes among newly recruited commercial drivers. *Accident Analysis and Prevention, 49*, 378–384. https://doi.org/10.1016/j.aap.2012.02.018.

Ba, X. Q., Zhou, F. F., & Wang, Y. G. (2018). Predicting personal injury crash risk through working conditions, job strain, and risky driving behaviors among taxi drivers. *European Transport Research Review, 10*(2). https://doi.org/10.1186/s12544-018-0320-x.

Baba, M., Miyama, G., Sugiyama, D., & Hitosugi, M. (2019). Influence of workplace environment, working conditions and health status of taxi drivers on vehicle collisions or near-miss events. *Industrial Health, 57*(4), 530–536. https://doi.org/10.2486/indhealth.2018-0104.

Barger, L. K., Rajaratnam, S. M. W., Wang, W., O'Brien, C. S., Sullivan, J. P., Qadri, S., ... Hlth Safety, G. (2015). Common Sleep Disorders Increase Risk of Motor Vehicle Crashes and Adverse Health Outcomes in Firefighters. *Journal of Clinical Sleep Medicine, 11*(3), 233–U249. https://doi.org/10.5664/jcsm.4534.

Besharati, M. M., & Kashani, A. T. (2018). Factors contributing to intercity commercial bus drivers' crash involvement risk. *Archives of Environmental & Occupational Health, 73*(4), 243–250. https://doi.org/10.1080/19338244.2017.1306478.

Carman, A. B., Gillespie, S., Jones, K., Mackay, J., Wallis, G., & Milosavljevic, S. (2010). All terrain vehicle loss of control events in agriculture: Contribution of pitch, roll and velocity. *Ergonomics, 53*(1), 18–29. https://doi.org/10.1080/00140130903380919.

Cassano-Piche, A. L., Vicente, K. J., & Jamieson, G. A. (2009). A test of Rasmussen's risk management framework in the food safety domain: BSE in the UK. *Theoretical Issues in Ergonomics Science, 10*(4), 283–304. https://doi.org/10.1080/14639220802059232.

Chen, C. K., & Xie, Y. (2014). The impacts of multiple rest-break periods on commercial truck driver's crash risk. *Journal of Safety Research, 48*, 87–93. https://doi.org/10.1016/j.jsr.2013.12.003.

Chen, C. K., & Zhang, J. (2016). Exploring background risk factors for fatigue crashes involving truck drivers on regional roadway networks: A case control study in Jiangxi and Shaanxi, China. *Springerplus, 5*. https://doi.org/10.1186/s40064-016-2261-y.

Chu, H. C. (2012). An Investigation of the Risk Factors Causing Severe Injuries in Crashes Involving Gravel Trucks. *Traffic Injury Prevention, 13*(4), 355–363. https://doi.org/10.1080/15389588.2012.654545.

Chu, H. C. (2016). Risk factors for the severity of injury incurred in crashes involving on-duty police cars. *Traffic Injury Prevention, 17*(5), 495–501. https://doi.org/10.1080/15389588.2015.1109082.

Clay, L., Treharne, G. J., Hay-Smith, E. J. C., & Milosavljevic, S. (2014). Are agricultural quad bike loss-of-control events driven by unrealistic optimism? *Safety Science, 66*, 54–60. https://doi.org/10.1016/j.ssci.2014.02.002.

Das, S., Islam, M., Dutta, A., & Shimu, T. H. (2020). Uncovering Deep Structure of Determinants in Large Truck Fatal Crashes. *Transportation Research Record, 2674* (9), 742–754. https://doi.org/10.1177/0361198120931507.

Goode, N., Salmon, P. M., Lenne, M. G., & Finch, C. (2018). *Translating systems thinking into practice: A guide to developing incident reporting systems*. CRC Press.

Gorucu, S., Murphy, D. J., & Kassab, C. (2017). Injury risks for on-road farm equipment and horse and buggy crashes in Pennsylvania: 2010–2013. *Traffic Injury Prevention, 18*(3), 286–292. https://doi.org/10.1080/15389588.2016.1198009.

Haq, M. T., Zlatkovic, M., & Ksaibati, K. (2020). Investigating occupant injury severity of truck-involved crashes based on vehicle types on a mountainous freeway: A hierarchical Bayesian random intercept approach. *Accident Analysis and Prevention, 144*. https://doi.org/10.1016/j.aap.2020.105654.

Harland, K. K., Bedford, R., Wu, H., & Ramirez, M. (2018). Prevalence of alcohol impairment and odds of a driver injury or fatality in on-road farm equipment crashes. *Traffic Injury Prevention, 19*(3), 230–234. https://doi.org/10.1080/15389588.2017.1407924.

Harland, K. K., Carney, C., & McGehee, D. (2016). Analysis of naturalistic driving videos of fleet services drivers to estimate driver error and potentially distracting behaviors as risk factors for rear-end versus angle crashes. *Traffic Injury Prevention, 17*(5), 465–471. https://doi.org/10.1080/15389588.2015.1118655.

Hulme, A., Stanton, N. A., Walker, G. H., Waterson, P., & Salmon, P. M. (2019). What do applications of systems thinking accident analysis methods tell us about accident causation? A systematic review of applications between 1990 and 2018. *Safety Science, 117*, 164–183. https://doi.org/10.1016/j.ssci.2019.04.016.

Kim, H., Jang, T. W., Kim, H. R., & Lee, S. (2018). Evaluation for Fatigue and Accident Risk of Korean Commercial Bus Drivers. *Tohoku Journal of Experimental Medicine, 246*(3), 191–197.

Lemp, J. D., Kockelman, K. M., & Unnikrishnan, A. (2011). Analysis of large truck crash severity using heteroskedastic ordered probit models. *Accident Analysis and Prevention, 43*(1), 370–380. https://doi.org/10.1016/j.aap.2010.09.006.

Leveson, N. G. (2011). Applying systems thinking to analyze and learn from events. *Safety Science, 49*(1), 55–64.

Mallia, L., Lazuras, L., Violani, C., & Lucidi, F. (2015). Crash risk and aberrant driving behaviors among bus drivers: The role of personality and attitudes towards traffic safety. *Accident Analysis and Prevention, 79*, 145–151. https://doi.org/10.1016/j.aap.2015.03.034.

Mamo, W. G., Newnam, S., & Tulu, G. S. (2014). Investigating the individual and organisational predictors of work-related driving crash involvement in Ethiopia. *Transportation Research Part F: Traffic Psychology and Behaviour, 23*, 156–164. https://doi.org/10.1016/j.trf.2014.01.001.

Mehdizadeh, M., Shariat-Mohaymany, A., & Nordfjaern, T. (2019). Driver behaviour and crash involvement among professional taxi and truck drivers: Light passenger cars versus heavy goods vehicles. *Transportation Research Part F-Traffic Psychology and Behaviour, 62*, 86–98. https://doi.org/10.1016/j.trf.2018.12.010.

Mehlhorn, S. A., Wilkin, H., Darroch, B., & D'Antoni, J. (2015). Physical characteristics of farm equipment crash locations on public roads in Tennessee. *Journal of*

Agricultural Safety and Health, 21(2), 85–94. https://doi.org/10.13031/jash.21.10607.
Milosavljevic, S., McBride, D. I., Bagheri, N., Vasiljev, R. M., Carman, A. B., Rehn, B., & Moore, D. (2011). Factors associated with quad bike loss of control events in agriculture. *International Journal of Industrial Ergonomics, 41*(3), 317–321. https://doi.org/10.1016/j.ergon.2011.02.010.
Missikpode, C., Peek-Asa, C., Young, T., & Hamann, C. (2018). Does crash risk increase when emergency vehicles are driving with lights and sirens? *Accident Analysis & Prevention, 113*, 257–262. https://doi.org/10.1016/j.aap.2018.02.002.
Mitchell, R. J., Bambach, M. R., & Friswell, R. (2014). Work and non-work-related vehicle crashes: The contribution of risky driving practices. *Safety Science, 68*, 65–72. https://doi.org/10.1016/j.ssci.2014.02.025.
Muir, C., Johnston, I., & Howard, E. (2018). Evolution of a holistic systems approach to planning and managing road safety: the Victorian case study, 1970–2015. *Injury Prevention, 24*. Retrieved from https://doi.org/10.1136/injuryprev-2017-042358.
Newnam, S., Blower, D., Molnar, L., Eby, D., & Koppel, S. (2018). Exploring crash characteristics and injury outcomes among older truck drivers: An analysis of truck-involved crash data in the United States. *Safety Science, 106*, 140–145. https://doi.org/10.1016/j.ssci.2018.03.012.
Newnam, S., & Goode, N. (2015). Do not blame the driver: A systems analysis of the causes of road freight crashes. *Accident Analysis & Prevention, 76*, 141–151. https://doi.org/10.1016/j.aap.2015.01.016.
Newnam, S., Goode, N., Read, G., & Salmon, P. (2020). Closing the research-practice gap in healthcare: The development and usability evaluation of a patient handling incident investigation toolkit. *Safety Science, 129*. https://doi.org/10.1016/j.ssci.2020.104844 104844.
Newnam, S., Goode, N., Read, G., Salmon, P. M., & Gembarovski, A. (2021). Systems-thinking in action: Results from implementation and evaluation of the patient handling injuries review of systems Toolkit. *Safety Science, 134*. https://doi.org/10.1016/j.ssci.2020.105086 105086.
Newnam, S., Goode, N., Salmon, P., & Stevenson, M. (2017). Reforming the road freight transportation system using systems thinking: An investigation of Coronial inquests in Australia. *Accident Analysis & Prevention, 101*, 28–36. https://doi.org/10.1016/j.aap.2017.01.016.
Newnam, S., Greenslade, J., Newton, C., & Watson, B. (2011). Safety in Occupational Driving: Development of a Driver Behavior Scale for the Workplace Context. *Applied Psychology, 60*(4), 576–599. https://doi.org/10.1111/j.1464-0597.2011.00448.x.
Newnam, S., Griffin, M. A., & Mason, C. (2008). Safety in work vehicles: A multilevel study linking safety values and individual predictors to work-related driving crashes. *Journal of Applied Psychology, 93*(3), 632–644. https://doi.org/10.1037/0021-9010.93.3.632.
Newnam, S., Lewis, I., & Watson, B. (2012). Occupational driver safety: Conceptualising a leadership-based intervention to improve safe driving performance. *Accident Analysis & Prevention, 45*, 29–38. https://doi.org/10.1016/j.aap.2011.11.003.
Newnam, S., & Muir, C. (2021). Reforming the future of workplace road safety using systems-thinking workplace road safety surveillance. *Safety Science, 138*. https://doi.org/10.1016/j.ssci.2021.105225 105255.
Newnam, S., Sheppard, D. M., Griffin, M. A., McClure, R. J., Heller, G., Sim, M. R., & Stevenson, M. R. (2014). Work-related road traffic injury: A multilevel systems protocol. *Injury Prevention, 20*(4), 1–6. https://doi.org/10.1136/injuryprev-2013-041132.
Newnam, S., Warmerdam, A., Sheppard, D. M., Griffin, M., & Stevenson, M. (2017). Do management practices support or constrain safe driving behaviour? A multi-level investigation in a sample of occupational drivers. *Accident Analysis & Prevention, 102*, 101–109. https://doi.org/10.1016/j.aap.2017.02.007.
Nickenig Vissoci, J. R., Krebs, E., Meier, B., Vieira, I. F., de Andrade, L., Byiringiro, F., ... Staton, C. A. (2020). Road traffic crash experience among commercial motorcyclists in Kigali, Rwanda. *International Journal of Injury Control and Safety Promotion, 27*(2), 181–187. https://doi.org/10.1080/17457300.2020.1724158.
Nik Mahdi, N. N. R., Bachok, N., Mohamed, N., & Shafei, M. N. (2014). Risk Factors for Near Miss Incident among Long Distance Bus Drivers in Malaysia. *Iranian Journal of Public Health, 43*, 117–124. Retrieved from <Go to ISI>://WOS:000345527300016.
NIOSH (2022). Motor vehicle safety at work. Crash facts. Retrieved from https://www.cdc.gov/NIOSH/.
Ogeil, R. P., Barger, L. K., Lockley, S. W., O'Brien, C. S., Sullivan, J. P., Qadri, S., ... Rajaratnam, S. M. W. (2018). Cross-sectional analysis of sleep-promoting and wake-promoting drug use on health, fatigue-related error, and near-crashes in police officers. *BMJ Open, 8*(9), e022041.
Ramirez, M., Bedford, R., Wu, H., Harland, K., Cavanaugh, J. E., & Peek-Asa, C. (2016). Lighting and marking policies are associated with reduced farm equipment-related crash rates: A policy analysis of nine Midwestern US states. *Occupational and Environmental Medicine, 73*(9), 621–626. https://doi.org/10.1136/oemed-2016-103672.
Ranapurwala, S. I., Mello, E. R., & Ramirez, M. R. (2016). A GIS-based matched case-control study of road characteristics in farm vehicle crashes. *Epidemiology, 27*(6), 827–834. https://doi.org/10.1097/ede.0000000000000542.
Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science, 27*(2), 183–213. https://doi.org/10.1016/S0925-7535(97)00052-0.
Reguly, P., Dubois, S., & Bedard, M. (2014). Examining the impact of opioid analgesics on crash responsibility in truck drivers involved in fatal crashes. *Forensic Science International, 234*, 154–161. https://doi.org/10.1016/j.forsciint.2013.11.005.
Safe Work Australia (2018). Number and Incidence Rate of Work-related Traumatic Injury Fatalities by Industry 2012–2016. Retrieved from Canberra, ACT.
Shams, Z., Mehdizadeh, M., & Khani Sanij, H. (2020). "I neither sleep well nor drive cautiously": How does sleep quality relate to crash involvement directly and indirectly? *Journal of Transport and Health, 18*. https://doi.org/10.1016/j.jth.2020.100907.
Shin, D. S., Park, M. H., & Jeong, B. Y. (2018). Human factors and severity of injury of delivery truck crashes registered for work-related injuries in South Korea. *Industrial Engineering and Management Systems, 17*(2), 302–310. https://doi.org/10.7232/iems.2018.17.2.302.
Shipp, E. M., Vasudeo, S., Trueblood, A. B., & Garcia, T. P. (2019). Single Vehicle Logging-Related Traffic Crashes in Louisiana from 2010–2015. *Journal of Agromedicine, 24*(2), 177–185. https://doi.org/10.1080/1059924x.2019.1567422.
Stevenson, M. R., Elkington, J., Sharwood, L., Meuleners, L., Ivers, R., Boufous, S., ... Wong, K. (2014). The role of sleepiness, sleep disorders, and the work environment on heavy-vehicle crashes in 2 Australian states. *American Journal of Epidemiology, 179*(5), 594–601. https://doi.org/10.1093/aje/kwt305.
Stuckey, R., Glass, D. C., LaMontagne, A. D., Wolfe, R., & Sim, M. R. (2010). Risk Factors for Worker Injury and Death From Occupational Light Vehicles Crashes in New South Wales (Australia). *American Journal of Industrial Medicine, 53*(9), 931–939. https://doi.org/10.1002/ajim.20854.
Sun, Q., & Tian, S. (2018). Research on influential factors of transportation near-misses in coal mine based on structural equation model. *Journal of Mines, Metals and Fuels, 66*(8), 444–448 https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053146371&partnerID=40&md5=8b0c884db6478b0650a301cdd78c58f8.
Svedung, I., & Rasmussen, J. (2002). Graphic representation of accident scenarios: Mapping system structure and the causation of accidents. *Safety Science, 40*(5), 397–417. https://doi.org/10.1016/s0925-7535(00)00036-9.
Thiese, M. S., Hanowski, R. J., Kales, S. N., Porter, R. J., Moffitt, G., Hu, N., & Hegmann, K. T. (2017). Multiple Conditions Increase Preventable Crash Risks among Truck Drivers in a Cohort Study. *Journal of Occupational and Environmental Medicine, 59*(2), 205–211. https://doi.org/10.1097/jom.0000000000000937.
Torregroza-Vargas, N. M., Bocarejo, J. P., & Ramos-Bonilla, J. P. (2014). Fatigue and crashes: The case of freight transport in Colombia. *Accident Analysis and Prevention, 72*, 440–448. https://doi.org/10.1016/j.aap.2014.08.002.
Useche, S. A., Cendales, B., Alonso, F., & Montoro, L. (2020). Multidimensional prediction of work traffic crashes among Spanish professional drivers in cargo and passenger transportation. *International Journal of Occupational Safety and Ergonomics*. https://doi.org/10.1080/10803548.2020.1732102.
Useche, S. A., Cendales, B., Alonso, F., & Orozco-Fontalvo, M. (2020). A matter of style? Testing the moderating effect of driving styles on the relationship between job strain and work-related crashes of professional drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 72*, 307–317. https://doi.org/10.1016/j.trf.2020.05.015.
Wang, S. Y., & Wu, K. F. (2019). Reducing intercity bus crashes through driver rescheduling. *Accident Analysis and Prevention, 122*, 25–35. https://doi.org/10.1016/j.aap.2018.09.019.
Wang, Y., Li, L., & Prato, C. G. (2019). The relation between working conditions, aberrant driving behaviour and crash propensity among taxi drivers in China. *Accident Analysis and Prevention, 126*, 17–24. https://doi.org/10.1016/j.aap.2018.03.028.
Wang, Y., & Prato, C. (2019). Determinants of injury severity for truck crashes on mountain expressways in China: A case-study with a partial proportional odds model. *Safety Science, 117*, 100–107. https://doi.org/10.1016/j.ssci.2019.04.011.
Wang, Y., Zhang, Y., Li, L., & Liang, G. (2019). Self-reports of workloads and aberrant driving behaviors as predictors of crash rate among taxi drivers: A cross-sectional study in China. *Traffic Injury Prevention, 20*(7), 738–743. https://doi.org/10.1080/15389588.2019.1650267.
Warmerdam, A., Newnam, S., Sheppard, D., Griffin, M., & Stevenson, M. (2017). Workplace road safety risk management: An investigation into Australian practices. *Accident Analysis & Prevention, 98*, 64–73. https://doi.org/10.1016/j.aap.2016.09.014.
Zhang, X. J., Yang, Y. M., Chen, Y., Yao, H. Y., Wu, M., Cui, M. J., ... Xiang, H. Y. (2017). Road traffic crashes among farm vehicle drivers in southern China: A cross-sectional survey. *Traffic Injury Prevention, 18*(1), 83–87. https://doi.org/10.1080/15389588.2016.1190840.
Zheng, Y. B., Ma, Y., Guo, L. X., Cheng, J. C., & Zhang, Y. L. (2019). Crash Involvement and Risky Riding Behaviors among Delivery Riders in China: The Role of Working Conditions. *Transportation Research Record, 2673*(4), 1011–1022. https://doi.org/10.1177/0361198119841028.
Zurich Insurance (2015). The Total Cost of Road Risk. Retrieved from https://www.nrspp.org.au/resources/zurich-insurance-the-total-cost-of-road-risk/.
Zuzewicz, K., Konarska, M., & Luczak, A. (2010). Injured professional drivers in Poland - an analysis of the causes and effects in relation to the time of the road accident. *International Journal of Occupational Safety & Ergonomics, 16*(1), 81–91.

Sharon Newnam Associate Professor Sharon Newnam is Associate Director of the Systems Safety team at the Monash University Accident Research Centre. Sharon has published widely in the area of workplace safety from a systems-thinking perspective and has applied this knowledge to improve safety across industries, including transportation and healthcare. Sharon is a prolific author, is an Associate Editor of the leading safety journal Safety Science, and an international member on

a Transportation Research Board committee, which is a division of the National Academies of Science, Engineering and Medicine.

Renée M. St. Louis, PhD Dr. St. Louis is an Assistant Research Scientist in the Behavioral Sciences Group at the University of Michigan Transportation Research Institute. She received her BA and MPH from the University of Michigan, Ann Arbor and PhD from the Monash University Accident Research Centre in Melbourne, Australia. She has managed a variety of projects aimed at enhancing safe mobility throughout the lifespan, with research focused on several areas of transportation safety including commercial motor vehicles, older driver safety and mobility, risk taking behaviour, and occupant protection issues.

Amanda Stephens Amanda is a Senior Research Fellow at Monash University Accident Research Centre, Melbourne Australia. She has been involved in road safety research for almost two decades and her background is in the psychology behind driver behaviour. Amanda's current research focuses on understanding and addressing the mechanisms behind risky driving behaviour. Particular focus is on aggressive, anti-social or non-compliant driving behaviour. She is currently involved in designing and implementing behaviour modification programs to support drivers in managing emotional, fatigued or inattentive driving.

Dianne Sheppard Dianne is a senior research fellow at Monash University Accident Research Centre, Melbourne Australia. Her work focuses on a variety of projects that facilitate social and work-related rehabilitation from injury or post-disease diagnosis.

Journal of Safety Research 83 (2022) 349–356

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Are plumbing apprentice graduates safer than their non-apprentice peers? Workers' compensation claims among journey level plumbers by apprenticeship participation

Sara Wuellner *, David Bonauto

Washington State Dept. of Labor and Industries, Safety and Health Assessment and Research for Prevention (SHARP) Program, PO Box 44330, Olympia, WA 98504, United States

ARTICLE INFO

Article history:
Received 18 October 2021
Received in revised form 27 May 2022
Accepted 13 September 2022
Available online 27 September 2022

Keywords:
Work
Injury
Occupational injury
Apprenticeship
Workers' compensation

ABSTRACT

Introduction: Apprenticeships combine mentored on-the-job training with related instruction to develop a workforce with the skills sought by employers. Workplace safety is an important component of apprenticeship training. Whether that training results in fewer work injuries, however, is largely unknown. *Method:* We linked Washington's registered apprenticeship data, plumber certification (licensing) data, employment data, and workers' compensation claims to compare claim rates among journey level plumbers (JLP) by apprenticeship participation. We used negative binomial regression models to estimate rates of total claims, wage replacement/disability claims, acute injuries, and musculoskeletal disorders (MSD), adjusted for worker characteristics. *Results:* Among JLP certified between 2000 and 2018, rates among JLP with no apprenticeship training were 46% higher for total workers' compensation claims (adjusted Rate Ratio (aRR) = 1.46, 95% CI: 1.26–1.69) and 60% higher for wage replacement/disability claims (aRR = 1.60, 95% CI: 1.22–2.11), compared to rates among JLP who completed a plumbing apprenticeship. Apprentice graduates experienced a greater decline in the rate of total claims between the 5 years preceding JLP certification and the years after certification (55.3% vs. 41.4% among JLP with no apprenticeship training). Greater rate reductions among JLP apprentice graduates were also observed for acute injuries and MSD, although the decline in MSD was not significantly different from the decline among JLP with no apprenticeship training. *Conclusions*: Successful completion of a plumbing apprenticeship program is associated with fewer work injuries throughout the career of a JLP. Apprenticeships appear to play a key role in reducing work injuries among JLP, especially acute injuries. *Practical Applications*: Apprenticeships are an effective model for reducing workplace injuries. The mechanisms by which apprenticeship training improves workplace safety should be identified to better inform injury prevention efforts among apprentices as well as among workers outside of a formal apprenticeship arrangement.

1. Introduction

Apprenticeships are a longstanding model of workforce development. Available for an increasing array of trades and professions, formal apprenticeship programs registered with the federal government train participants through on-the-job training (OJT) supervised by experienced mentors and related/supplemental instruction (RSI), which may include training seminars, online courses, or classroom instruction at community colleges or technical schools. Employers, labor unions, or business associations develop and sponsor apprentice programs, often supported, in part, through federal and state funding. Most apprenticeship programs take multiple years to complete. Participants earn wages during their apprenticeship, receive a nationally recognized credential upon graduation from the program, and most find employment in their trade immediately upon completion of the program. Acceptance into apprentice programs is often competitive, with applicants outnumbering spaces. Apprenticeships are currently experiencing renewed interest as a means to ensure that workers' skills meet the needs of contemporary employers (US Department of Labor, 2021).

State and federal oversight of apprenticeship programs extends back more than 80 years (National Aprenticeship Act, 1937). Occupation-specific program standards define required hours of OJT for specific work processes, ratio of apprentices to journey-

* Corresponding author.
E-mail address: sara.wuellner@lni.wa.gov (S. Wuellner).

https://doi.org/10.1016/j.jsr.2022.09.009
0022-4375/

level employees, hours of RSI, and wages. Federal regulations require that apprenticeship programs include workplace safety training in both the OJT and RSI (Standards of Apprenticeship, 2022).

The many documented benefits to apprenticeships include higher lifetime earnings and stable employment for apprentice graduates, and higher productivity and lower turnover among employers (Lerman, Eyster, & Chambers, 2009; Reed, Liu, Kleinman, Mastri, Reed, Sattar, & Ziegler, 2012). Increased workplace safety is an assumed benefit of apprenticeship training, in part because of the required safety training, as well as the structured learning environment of apprenticeship programs that can be leveraged to promote workplace safety and health competencies (Guerin et al., 2020). However, there is little evidence to either support or refute the assumption that apprentice training reduces work injuries. Generally, research into health outcomes among apprentices has focused on work-related injuries and illnesses among apprentices compared to more experienced workers, and have highlighted the increased risk of injuries among apprentices (Lipscomb, Dement, Nolan, Patterson, & Li, 2003; Sahl, Kelsh, Haines, Sands, & Kraus, 1997). Less is known about the work-related injury experience of workers who participated in apprentice programs compared to peers with no apprenticeship participation.

We linked state databases of licensed certified plumbers, apprenticeship participants, employment history, and workers' compensation claims to compare rates of work-related injuries and illnesses of journey level plumbers (JLP) who completed a plumbing apprenticeship to those who did not complete an apprenticeship. Participation in apprenticeships is not the result of random selection. Factors that lead an individual to choose to enroll in—and complete—an apprentice program may also influence aspects of occupational safety. We used worker and employer characteristics captured in the administrative databases in an attempt to control for some of the underlying differences between apprentices and non-apprentices.

2. Materials and methods

2.1. Journey level plumber data

To work in the plumbing trade in Washington, all workers must be licensed and certified by the Washington State Plumber Certification Program (18.106 RCW). Plumbers achieve journey level certification by fulfilling the state's requirements as a plumber trainee, working 8,000 hours (approximately 4 years) under the supervision of a journey level plumber, completing 8 hours of continuing education annually, and passing the state's plumber exam. Plumber trainees may fulfill the requirements as an apprentice, but apprenticeship is not required to become a certified JLP.

A database of all individuals licensed to perform plumbing work in Washington is maintained by the Washington State Department of Labor & Industries (L&I), and includes the worker's name, date of birth, social security number, trade specialty, certification start and end dates, and active status. It serves as a registry of all workers in the occupation, with entry and exit dates into and out of the profession.

We chose to focus on plumbers rather than other occupations because of the availability of the professional licensing data, and because plumbing apprenticeships are among the occupations with the greatest enrollment (US Department of Labor, 2020). Journey level plumbers were selected over other plumbing specialties (e.g., residential) because the size of the population, stratified by apprenticeship participation status, was sufficient for analysis. The study included JLP certified between January 1, 2000 and December 31, 2018.

2.2. Plumbing apprentice participation

All individuals who enroll in a registered apprentice program in Washington are recorded in Washington's apprentice registry data, also maintained by L&I, which includes the participant's name, social security number, apprenticeship start and end dates, program completion status, and a Standard Occupational Classification code to describe the type of apprentice program.

We identified apprentice program participation among JLP by linking social security numbers across the plumbing licensing data and apprentice registry data. Plumbers were grouped into one of three categories of apprentice program participation: completion of a plumbing apprentice program within four quarters of the JLP certification date; no participation in any apprentice program; or other apprentice program participation – including incomplete plumbing apprenticeship, completion of a plumbing apprentice program more than four quarters from the JLP certification date, or participation in an apprenticeship program for an occupation other than plumbing.

2.3. Work-related injury data

We used Washington workers' compensation claims data to evaluate work injuries among JLP.

In Washington, with few exceptions, employers are required to obtain workers' compensation insurance from the state (referred to as state funded), unless they are approved by the state to self-insure. L&I administers the state funded program, oversees the self-insured program, and maintains records on claims filed through either program. We included both state funded and self-insured workers' compensation claims data, for claims involving wage replacement and/or disability payments and claims limited to medical-aid payments. In Washington, a claimant may be eligible for wage replacement or disability benefits if unable to work following a three-day waiting period.

Workers compensation claims among JLP were identified using the plumber's social security number. We assessed the total number of accepted claims (medical aid-only claims plus wage replacement/disability claims) and wage replacement/disability claims experienced during the JLP's active licensure or through 2019, for plumbers whose licenses were active beyond 2018. We also assessed workers' compensation claims for injuries or illnesses that occurred in the five years prior to the certification start date.

We used Occupational Injury and Illness Classification System (OIICS v1) codes, assigned by L&I staff based on narrative injury descriptions reported on the claim initiation form, to characterize injuries and illnesses. We followed a definition of musculoskeletal disorders based on a combination of OIICS nature of injury and event or exposure codes, developed and validated previously using Washington workers' compensation data (Marcum & Adams, 2017; Silverstein, Viikari-Juntura, & Kalat, 2002; Spector, Adams, & Silverstein, 2011). We defined acute injuries as claims that did not meet the definition of a musculoskeletal disorder and which were assigned an OIICS nature code within the division "Traumatic Injuries and Disorders." Completeness of OIICS codes differs by insurer; nearly all state funded claims in this study were coded while almost 70% of self-insured claims were missing OIICS codes, reflecting the relatively limited injury description data available for self-insured claims. Using the distribution of codes assigned to state funded claims, we randomly assigned OIICS codes to replace the missing values.

To differentiate claimants working as plumbers at the time of injury or illness from those employed in some other occupation, we used the Washington State Risk Classification system (a method for setting insurance premiums by grouping types of work according to risk). To compare claims at the same point in their development, we evaluated claim costs incurred at one-year post-injury, converted to 2018 US dollars using the Consumer Price Index to adjust for inflation. Analysis of costs are limited to state funded claims because of missing or incomplete data among self-insured claims.

2.4. Employment data

To calculate claim rates, hours worked by JLP were extracted from the Washington Unemployment Insurance (UI) database, which includes quarterly records of employer-reported wages and hours worked for each worker in the state covered for unemployment insurance, by worker name and social security number. Using the JLP social security numbers, we extracted quarterly employment data for the five years preceding the license start through the license end date or 2019Q4, whichever occurred first. We excluded from the analysis JLP for whom no wage and hour data was found in UI, as we could not estimate hours-based injury rates among this group. Individuals not found in the UI data may be self-employed, employed out of state, or out of the work force. Self-employed individuals (with no employees) who meet the state's definition of an independent contractor are exempt from both Washington's workers' compensation insurance coverage and unemployment insurance coverage (Revised Code, 2002). JLP with wage and hour data reported in UI were considered to be employed for wages, and included in the final analysis.

2.5. Statistical analysis

Differences in characteristics between the three categories of plumbing apprenticeship status (completed plumbing apprenticeship, never enrolled in apprenticeship, participated in some apprenticeship training) were assessed using chi square tests for homogeneity of proportions. To test for differences in continuous variables following either normal or non-normal distributions, we used analysis of variance and Kruskal-Wallis tests, respectively.

We used negative binomial regression models to estimate rates of accepted claims, rates of wage replacement/disability claims, rates of accepted claims for acute injuries, and rates of accepted claims for musculoskeletal disorders. To account for underlying differences between the apprenticeship groups, the final regression models were adjusted for year of initial JLP certification, worker age at JLP certification, number of employers during JLP license, size of employer, license for plumbing specialty other than JLP, hours worked in plumbing industry prior to JLP certification. Level of significance was chosen as $\alpha = 0.05$. Analyses were performed using SAS 9.4.

3. Results

3.1. Apprenticeship participation among journey level plumbers

There were 4,086 plumbers with initial JLP certification between January 1, 2000 and December 31, 2018. Of those, 18.9% (773) JLP completed a plumbing apprenticeship program within four quarters of their JLP certification, 9.1% (373) had some participation in an apprenticeship program, and 72.0% (2,940) did not appear in the apprenticeship registry data. In total, 72.5% of JLP had wage and hour data identified in UI, although identification in UI differed by apprenticeship status. Nearly all plumbing apprentice graduates had wage and hour data reported in UI (768 out of 773, 99.4%), while a smaller portion of JLP who never enrolled in an apprentice program were reported in UI (2,202 out of 2,940, 74.9%) (Table 1). Plumbers with no wage and hour data reported in UI were excluded from further analysis.

3.2. Worker characteristics by apprenticeship participation

Apprentice graduates differed from those who never enrolled in an apprentice program across several characteristics (Table 2). On average, apprenticeship graduates were younger at the start of their JLP certification, became a JLP later in the study period, worked more hours each quarter as a JLP, worked for larger employers, and worked for more employers ($p < 0.05$ adjusted for post hoc pairwise comparisons). Conversely, JLP who never enrolled in an apprentice program tended to be older at their initial JLP certification, enter the profession earlier in the study period, work fewer hours each quarter, work for a smaller employer, and work for fewer employers during their JLP license.

For most measures, JLP with some apprenticeship participation (the 357 individuals who enrolled in but discontinued a plumbing apprenticeship [48%], completed a plumbing apprenticeship more than four quarters from JLP certification [34%], or completed an apprenticeship other than plumbing [18%]) fell in between apprenticeship graduates and never enrollees. For example, at the time of initial JLP certification, they were older than apprenticeship graduates and younger than never enrollees. One exception was holding another type of plumbing license – of the three groups, JLP with some apprenticeship participation were most likely to have another type of plumbing license in addition to the JLP license.

Regardless of apprenticeship participation, one in three plumbers experienced an injury or illness resulting in a workers' compensation claim accepted for medical aid payments, disability, or wage replacement benefits.

3.3. Workers' compensation claim characteristics by apprenticeship participation

Table 3 presents select characteristics of workers' compensation claims by apprenticeship status, for work-related injuries or illnesses experienced as JLP. Several claim characteristics showed no significant difference by apprenticeship participation. Eighty-three percent of claims occurred among JLP performing plumbing work (based on risk classification) and one quarter of claims were eligible for wage replacement and or disability benefits. Median claim costs of state funded claims did not differ significantly across the three groups.

Claims among JLP with no apprenticeship participation were more likely to be insured by the state fund (90.4% among JLP with no apprentice training compared to 83.9% among JLP apprentice graduates and 84.1% among JLP with some apprentice training, based on post hoc pairwise comparisons). State fund claims among JLP with no apprenticeship participation were more likely to result in days of missed work than JLP with any apprenticeship participation ($p < 0.05$ adjusted for post hoc pairwise comparisons).

Acute injuries were the most common injury type. JLP with any apprenticeship training had a greater percent of claims lacking a nature of injury classification, attributable to an overrepresentation of both non-classified claims and apprenticeship participation among the self-insured. Stratified by insurer, the distribution of the nature of injury did not differ by apprenticeship participation. Additionally, injury type did not differ by apprenticeship participation for state funded and self-insured claims combined, after random reassignment of non-classified values to a nature of injury mirroring the distribution among state funded claims.

Table 1
Study inclusion by plumbing apprenticeship status among journey level plumbers.

	JLP who completed plumbing apprentice	JLP with some apprentice training	JLP with no apprentice training	Total
JLP identified	773 (100%)	373 (100%)	2940 (100%)	4086 (100%)
Excluded: No wage/hour data in UI	5 (0.6%)	16 (4.3%)	738 (25.1%)	759 (18.6%)
Included: Wage/hour data In UI	768 (99.4%)	357 (95.7%)	2202 (74.9%)	3327 (81.4%)

Table 2
Characteristics of Journey Level Plumbers (JLP) employed for wages by plumbing apprenticeship participation. Data for continuous variables are presented as mean (standard deviation) or median (Q1, Q3). Data for categorical variables are presented as number (%).

	JLP who completed plumbing apprentice	JLP with some apprentice training	JLP with no apprentice training	Stat. sig.[a]
Total JLP	768 (100.0)	357 (100.0)	2202 (100.0)	
Age at initial JLP certification, mean yrs (std dev)	32.0 (6.6)	34.8 (8.0)	37.6 (9.1)	p <.0001
Year of initial JLP certification				p <.0001
2000–2004	140 (18.2)	94 (26.3)	741 (33.7)	
2005–2009	227 (29.6)	109 (30.5)	615 (27.9)	
2010–2014	251 (32.7)	83 (23.2)	401 (18.2)	
2015–2018	150 (19.5)	71 (19.9)	445 (20.2)	
Hours worked per quarter as JLP, mean (std dev)	456 (82)	425 (97)	420 (127)	p <.0001
Employer size: average FTE of employer, median (Q1, Q3)	170.9 (56.4, 410.2)	149.4 (34.2, 399.6)	60.8 (17.2, 211.2)	p <.0001
Number of employers a plumber worked for over the course of their JLP license, median (Q1, Q3)	3 (2, 6)	3 (2, 7)	2 (1, 4)	p <.0001
Other plumbing license[b]				p <.0001
No	718 (93.5)	271 (75.9)	1830 (83.1)	
Yes	50 (6.5)	86 (24.1)	372 (16.9)	
Workers' compensation claims accepted for injuries experienced as JLP				NS
0 claims	516 (67.2)	233 (65.3)	1446 (65.7)	
1+ claims	252 (32.8)	124 (34.7)	756 (34.3)	

Std dev = Standard deviation.
FTE = Full time equivalent; calculated as 1 FTE = 2000 hrs.
Q1, Q3 = First quartile, third quartile.
NS = No significant difference between groups.
[a] Stat. sig. = statistical significance of difference among groups of apprenticeship participation (p < 0.05 = statistically significant), based on analysis of variance, Kruskal-Wallis tests, or chi square tests for homogeneity of proportions, for continuous variables with normal distribution, continuous variables with non-normal distribution, and categorical variables, respectively.
[b] Residential, residential service, pump & irrigation and domestic well, backflow.

3.4. Rates of workers' compensation claims by apprenticeship participation

Rates of workers' compensation claims were lowest among JLP apprentice graduates and highest among those with no apprenticeship training, for total claims and wage replacement/disability claims, both unadjusted, and adjusted for underlying differences in the populations (Table 4). The unadjusted rate of total accepted claims was 70% higher among JLP with no apprentice training, compared to apprentice graduates. The adjusted rate of total accepted claims (controlling for differences in year of initial JLP certification, worker age at initial JLP certification, number of employers as JLP, size of employer, license for plumbing specialty other than JLP, and workers' compensation claim rate in the five years prior to JLP certification) was 46% higher among JLP with no apprentice training, compared to apprentice graduates.

Claim rate ratios were greater for wage replacement/disability claims, where unadjusted rates among non-apprenticeship participants were more than twice as high as rates among apprenticeship graduates. After controlling for covariates, rates among never enrollees still exceeded rates among apprenticeship graduates by 60%.

Claim rates among the group of JLP with some apprenticeship training fell in between the rates among non-participants and apprenticeship graduates, although the rates among JLP apprentice graduates and JLP with some apprenticeship participation were not significantly different.

3.5. Workers' compensation claim rates over time by apprenticeship participation

3.5.1. Claim rates by benefit type

In the five years preceding JLP certification, rates of wage replacement/disability claims were lower among apprentice graduates than among JLP with no apprentice training, while rates of total claims did not differ significantly by apprenticeship participation (Table 5).

For all three apprentice groups, rates in total accepted claims declined from the five years preceding the JLP license to the period of the JLP license. Rates of total accepted claims among JLP apprentice graduates declined the most (55.3%, 95% CI: 48.8%–61.0%), exceeding the 41.4% decline (95% CI: 35.9%–46.4%) among JLP with no apprenticeship participation. Declines were smaller for wage replacement/disability claims, estimated at less than 30% among JLP apprenticeship graduates and among JLP with no apprenticeship participation, and did not differ significantly by apprenticeship participation.

3.5.2. Claim rates by injury type

Injury-specific rates based on original injury codes were similar to estimates based on reassignment of missing values; only rates based on reassigned missing values are presented in Table 5. In the five years preceding JLP certification, rates of acute injuries did not differ significantly by apprenticeship participation, while rates of musculoskeletal disorders were marginally lower among

Table 3
Characteristics of workers' compensation claims among Journey Level Plumbers (JLP) employed for wages by plumbing apprenticeship participation. Data presented as number (%), unless otherwise noted.

	Claims among JLP who completed plumbing apprentice	Claims among JLP with some apprentice training	Claims among JLP with no apprentice training	Stat. sig.[a]
Claims	428 (100.0)	220 (100.0)	1530 (100.0)	
Insurer				p <.0001
State Fund	359 (83.9)	185 (84.1)	1383 (90.4)	
Self-insured	69 (16.1)	35 (15.9)	147 (9.6)	
Benefits paid				NS
Medical aid only	335 (78.3)	169 (76.8)	1119 (73.1)	
Wage replacement or disability	93 (21.7)	51 (23.2)	411 (26.9)	
Risk classification				NS
Plumber	359 (83.9)	174 (79.1)	1276 (83.4)	
Other than plumber	69 (16.1)	46 (20.9)	254 (16.6)	
Nature of injury				p < 0.05
Traumatic Injuries and Disorders	264 (61.7)	123 (55.9)	924 (60.4)	
Musculoskeletal disorders	97 (22.7)	57 (25.9)	418 (27.3)	
Diseases and non-traumatic conditions	15 (3.5)	15 (6.8)	76 (5.0)	
Not classified	52 (12.1)	25 (11.4)	112 (7.3)	
Nature of injury, "not classified" re-assigned				NS
Traumatic Injuries and Disorders	297 (69.4)	142 (64.5)	998 (65.2)	
Musculoskeletal disorders	112 (26.2)	61 (27.7)	448 (29.3)	
Diseases and non-traumatic conditions	19 (4.4)	17 (7.7)	81 (5.3)	
Not classified	0 (0)	0 (0)	3 (0.2)	
Injury event or exposure				NS
Contact With Objects And Equipment	189 (44.2)	88 (40.0)	612 (40.0)	
Bodily Reaction And Exertion	129 (30.1)	72 (32.7)	543 (35.5)	
Falls	26 (6.1)	17 (7.7)	131 (8.6)	
Exposure to Harmful Substances, Environments	24 (5.6)	11 (5.0)	60 (3.9)	
Other events or exposures[b]	9 (2.1)	6 (2.7)	45 (2.9)	
Not classified	51 (11.9)	26 (11.8)	139 (9.1)	
Time loss days paid[c]				p < 0.05
0 days	328 (91.4)	161 (87.0)	1146 (82.9)	
1–30 days	9 (2.5)	8 (4.3)	91 (6.6)	
31–100 days	6 (1.7)	5 (2.7)	64 (4.6)	
>100 days	16 (4.5)	11 (5.9)	82 (5.9)	
Claim costs one year post injury[c,d]				
Total claim costs, median (Q1, Q3)	$820 (490, 2510)	$770 (490, 2620)	$910 (460, 3010)	NS
Medical aid costs, median (Q1, Q3)	$820 (480, 2320)	$730 (490, 2380)	$840 (450, 2760)	NS

NS = No significant difference between groups.
Q1, Q3 = First quartile, third quartile.
[a] Stat. sig. = statistical significance of difference among groups of apprenticeship participation ($p < 0.05$ = statistically significant), based on analysis of variance, or chi square tests for homogeneity of proportions, or Kruskal-Wallis tests, for categorical variables and continuous variables with non-normal distribution, respectively.
[b] Includes Transportation Accidents, Fires And Explosions, Assaults And Violent Acts.
[c] Limited to State Fund claims because of missing values among self-insured claims.
[d] Workers' compensation claim costs incurred at 15 months post injury.

Table 4
Workers' compensation claim rates and rate ratios among Journey Level Plumbers (JLP) employed for wages by plumbing apprenticeship participation.

	Unadjusted estimates		Adjusted[a] estimates	
	Claim rate[b] (95% CI)	Rate ratio (95% CI)	Claim rate[b] (95% CI)	Rate ratio (95% CI)
Accepted claims				
Completed plumbing apprentice	74.3 (65.5–84.3)	referent	73.1 (63.4 –84.3)	referent
Some apprentice training	97.9 (82.3–116.5)	1.32 (1.06–1.63)	84.2 (70.5–100.5)	1.15 (0.93–1.43)
No apprentice training	126.4 (117.7–135.9)	1.70 (1.47–1.97)	106.4 (97.8–115.8)	1.46 (1.26–1.69)
Wage replacement/disability claims				
Completed plumbing apprentice	16.4 (12.9–20.9)	referent	16.6 (12.7–21.7)	Referent
Some apprentice training	22.5 (16.1–31.7)	1.37 (0.91–2.08)	18.2 (13.1–25.3)	1.10 (0.73–1.65)
No apprentice training	34.5 (30.8–38.6)	2.10 (1.61–2.73)	26.6 (23.2–30.4)	1.60 (1.22–2.11)

CI = Confidence Interval.
[a] Claim rate adjusted for: year of initial JLP certification, worker age at initial JLP certification, number of employers as JLP, size of employer, license for plumbing specialty other than JLP, and workers' compensation claim rate in the five years prior to JLP certification.
[b] Claims per 1,000 FTE.

JLP apprentice graduates. The change in rates of acute injuries before and after JLP certification differed by apprenticeship participation, with JLP apprenticeship graduates experiencing a greater decline in acute injuries than JLP with no apprenticeship training (59.7% vs 43.1%, respectively). Declines in rates of musculoskeletal disorders were statistically similar, although the rate among JLP

Table 5
Workers' compensation claim rates in the 5 years before initial Journey Level Plumber (JLP) certification and the years during JLP license, among JLP employed for wages by plumbing apprenticeship participation.

	Claim rate before JLP[a,b] (95%CI)	Claim rate during JLP[a,b] (95%CI)	Percent decline (95%CI)
Rates by benefit eligibility			
Accepted claims			
Completed plumbing apprentice	163.6 (145.9–183.3)	73.1 (63.4–84.3)	55.3% (48.8%–61.0%)
Some apprentice training	174.8 (147.5–207.3)	84.2 (70.5–100.5)	51.8% (41.3%–60.5%)
No apprentice training	181.5 (166.5–197.9)	106.4 (97.8–115.8)	41.4% (35.9%–46.4%)
Wage replacement/disability claims			
Completed plumbing apprentice	22.6 (17.3–29.5)	16.6 (12.7–21.7)	26.5% (+2.1%–47.1%)
Some apprentice training	35.1 (25.0–49.4)	18.2 (13.1–25.3)	48.2% (21.5%–65.9%)
No apprentice training	37.5 (31.8–44.2)	26.6 (23.2–30.4)	29.2% (16.0%–40.3%)
Rates by injury type[c]			
Claims for acute injuries			
Completed plumbing apprentice	124.3 (109.5–141.2)	50.2 (42.8–58.9)	59.7% (52.5%–65.7%)
Some apprentice training	112.2 (93.6–134.5)	53.8 (44.0–65.7)	52.1% (39.6%–62.0%)
No apprentice training	124.6 (113.2–137.2)	70.9 (64.4–77.9)	43.1% (37.0%–48.7%)
Claims for musculoskeletal disorders			
Completed plumbing apprentice	31.3 (24.8–39.3)	18.6 (14.6–23.8)	40.4% (22.0%–54.5%)
Some apprentice training	50.6 (38.3–66.7)	22.7 (17.0–30.3)	55.1% (36.0%–68.5%)
No apprentice training	44.5 (38.0–51.9)	29.1 (25.3–33.6)	34.5% (22.4%–44.7%)

CI = Confidence Interval.
[a] Claim rate adjusted for: year of initial JLP certificate on, worker age at initial JLP certification, number of employers as JLP, size of employer, license for plumbing specialty other than JLP, and workers' compensation claim rate in the five years prior to JLP certification.
[b] Claims per 1,000 FTE.
[c] Missing injury classification codes (more common among self-insured claims) were randomly assigned codes to mimic the distribution of codes assigned to state funded claims.

apprentice graduates was lower than the rate among JLP with no apprentice training (18.6 vs 29.1 musculoskeletal disorder claims per 1,000 FTE, respectively).

4. Discussion

This is the first study to our knowledge to compare rates of occupational injuries and illnesses of workers at a similar point in their shared career, differentiated by apprenticeship participation. Our findings suggest that plumbers who complete apprenticeships experience fewer work injuries throughout their career compared with plumbers with no apprenticeship participation. Moreover, apprenticeships appear to play a key role in reducing work injuries, especially acute injuries.

With a focus on workplace safety in both the OJT and RSI, the apprenticeship model provides many opportunities to develop occupational health and safety competencies among workers early in their careers. The details of the safety training are not specified in many program standards (as one plumbing apprentice program simply noted, "Safety instruction is included in every quarter's curriculum of this craft" (Washington State Apprenticeship and Training Council, 2021a), and likely vary by apprenticeship program. But the multi-faceted approach to teaching workplace safety is exemplary of interactive forms of instruction more likely to prevent work injuries and illnesses (Burke et al., 2006). Looking at injury rates across the three study groups, with injury rates among JLP with some apprenticeship participation falling in between the never enrollees and the apprenticeship graduates, suggests that workplace safety increases with higher levels of apprenticeship participation.

Another possible explanation for our findings is that apprenticeships are a proxy indicator for worker safety. Apprenticeship participation is not the result of randomization. Participation in an apprenticeship occurs when an individual elects to apply to a program and an apprenticeship grants admission into the program. We were unable to account for all underlying differences between apprentice graduates and plumbers who never enrolled in an apprenticeship, including whether never enrollees would have met apprenticeship program enrollment criteria (e.g., one Washington plumbing apprenticeship program requires that applicants prove proficiency in high school or college algebra and pass a multi-panel drug test) (Washington State Apprenticeship and Training Council, 2021b). The admission process may differentially favor workers with a predisposition for workplace safety competencies, individuals who experience fewer workplace injuries during the training period, and adopt workplace safety principles more readily.

An alternate explanation for the differences in injury rates, aside from the safety training received as an apprentice, is that work tasks differ by apprenticeship participation, with individuals who never participated in an apprenticeship engaged in work activities that are inherently more dangerous than the work assigned to apprenticeship graduates. Two aspects of this study, however, challenge this hypothesis, suggesting instead that the study groups faced similar occupational hazards. First, restricting the study to the specific worker group of journey level plumbers helps limit the variability of job hazards. Second, characteristics of the workers' compensation claims among JLP, which differed little by apprenticeship participation, supports the assumption that JLP face similar workplace risks, regardless of apprenticeship participation.

Workplace safety is complex, reflecting not only individual worker knowledge and actions, but also organizational factors (DeJoy, Gershon, Murphy, & Wilson, 1996). Apprenticeship programs may be promoting workplace safety as much through the training of individual workers as through the engagement with employers. This is most evident in employer development of training programs, but may also manifest as a talent pipeline by which employers with a strong commitment to safety preferentially hire apprentice graduates or sponsor apprenticeship programs. Additional research is needed to identify the mechanisms by which apprenticeship programs impact workplace injuries and illnesses.

Although not the focus of the study, the high rates of workplace injuries and illnesses during the training period preceding professional certification observed here and noted in other studies of apprentices demonstrate a need for enhanced injury prevention efforts during the apprentice training (Kaskutas et al., 2010; Lipscomb et al., 2003, 2008). The smaller decline in rates of musculoskeletal disorders – often debilitating and expensive injuries, and documented to occur among apprentices (Anton et al., 2020; Rosecrance, Cook, Anton, & Merlino, 2002) – highlights a need for effective prevention efforts throughout a worker's career.

5. Limitations

To measure work-related injuries, we used workers' compensation claims, a data source known to undercount the true occurrence of injuries (Biddle, Roberts, Rosenman, & Welch, 1998; Fan, Bonauto, Foley, & Silverstein, 2006; Shannon & Lowe, 2002). Disincentives for filing a workers' compensation claim, including peer pressure, employer pressure, and unfamiliarity with the system (Azaroff, Levenstein, & Wegman, 2002; Lipscomb, Nolan, Patterson, Sticca, & Myers, 2012; Rosenman et al., 2000) are likely present in all three study groups. Assuming the magnitude of the undercount is similar across the three groups of apprenticeship participation, the claim rate ratios likely would not change if we were able to account for underreporting.

Incomplete injury classification data may have impacted estimates of injury-specific claim rates, which we calculated after randomly assigning OIICS codes to the 6.7% of claims (for injuries before or during JLP license) that were not originally classified. We chose this approach over omitting non-classified claims from the estimates, which differentially would have excluded self-insured claims. To examine the impact of this approach, we modeled injury-specific rates using the original OIICS codes, and found results to be similar to those based on the re-assigned codes. Lack of cost data among self-insured claims limited our ability to describe fully the workers' compensation claim costs. Despite these limitations, more complete data are unlikely to substantially change our findings.

Our findings from this program evaluation are not necessarily generalizable to apprenticeship programs for occupations other than plumbing or in other jurisdictions. Depending on specific occupational hazards, training requirements, and employment opportunities, the magnitude of the association between apprenticeship participation and workplace safety may differ by trade. State-level variation in the administration of apprenticeship or professional licensing programs may lead to dissimilar results in other jurisdictions. Finally, these findings may not hold for self-employed independent contractors (American Community Survey estimates suggest that 8.8% of Washington plumbers are self-employed) and others not reported in the UI or workers' compensation data.

6. Conclusion

Apprenticeships are an effective model for reducing workplace injuries. Identifying the mechanism by which apprenticeship training improves workplace safety (e.g., mentorship by experienced workers, relationships with employers, specific components of the safety training, or some combination) could have implications for decreasing occupational injuries and illnesses not only among apprentices, but also among workers outside of a formal apprenticeship arrangement.

Funding

No outside funding supported this work.

Declarations of Interest

None.

Acknowledgements

The authors would like to thank Jody Robins and Chris Bowe for their thoughtful review and comments on an earlier version of this manuscript.

References

Anton, D., Bray, M., Hess, J. A., Weeks, D. L., Kincl, L. D., & Vaughan, A. (2020). Prevalence of work-related musculoskeletal pain in masonry apprentices. *Ergonomics, 63*(9), 1194–1202.
Azaroff, L. S., Levenstein, C., & Wegman, D. H. (2002). Occupational injury and illness surveillance: Conceptual filters explain underreporting. *American Journal of Public Health, 92*(9), 1421–1429.
Biddle, J., Roberts, K., Rosenman, K. D., & Welch, E. M. (1998). What percentage of workers with work-related illnesses receive workers' compensation benefits? *Journal of Occupational and Environmental Medicine, 40*(4), 325–331.
Burke, M. J., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R. O., & Islam, G. (2006). Relative effectiveness of worker safety and health training methods. *American Journal of Public Health, 96*(2), 315–324.
DeJoy, D. M., Gershon, R. R., Murphy, L. R., & Wilson, M. G. (1996). A work-systems analysis of compliance with universal precautions among health care workers. *Health Education Quarterly, 23*(2), 159–174.
Fan, Z. J., Bonauto, D. K., Foley, M. P., & Silverstein, B. A. (2006). Underreporting of work-related injury or illness to workers' compensation: Individual and industry factors. *Journal of Occupational and Environmental Medicine, 48*(9), 914–922.
Guerin, R. J., Castillo, D., Hendricks, K. J., Howard, J., Piacentino, J., & Okun, A. H. (2020). Preparing the Future Workforce for Safe and Healthy Employment. *American Journal of Public Health, 110*(1), 69–71.
Kaskutas, V., Dale, A. M., Lipscomb, H., Gaal, J., Fuchs, M., & Evanoff, B. (2010). Fall prevention among apprentice carpenters. *Scandinavian Journal of Work, Environment & Health, 36*(3), 258–265.
Lerman, R., Eyster, L., & Chambers, K. (2009). *The Benefits and Challenges of Registered Apprenticeship: The Sponsors' Perspective*. Washington: The Urban Institute.
Lipscomb, H. J., Dale, A. M., Kaskutas, V., Sherman-Voellinger, R., & Evanoff, B. (2008). Challenges in residential fall prevention: Insight from apprentice carpenters. *American Journal of Industrial Medicine, 51*(1), 60–68.
Lipscomb, H. J., Dement, J. M., Nolan, J., Patterson, D., & Li, L. (2003). Nail gun injuries in residential carpentry: Lessons from active injury surveillance. *Injury Prevention, 9*(1), 20–24.
Lipscomb, H. J., Nolan, J., Patterson, D., Sticca, V., & Myers, D. J. (2012). Safety, incentives, and the reporting of work-related injuries among union carpenters: "You're pretty much screwed if you get hurt at work". *American Journal of Industrial Medicine*.
Marcum, J., & Adams, D. (2017). Work-related musculoskeletal disorder surveillance using the Washington state workers' compensation system: Recent declines and patterns by industry, 1999–2013. *American Journal of Industrial Medicine, 60*(5), 457–471.
National Aprenticeship Act of 1937, 29 U.S.C. 50.
Reed, D., Liu, A. Y.-H., Kleinman, R., Mastri, A., Reed, D., Sattar, S., & Ziegler, J. (2012). *An Effectiveness Assessment and Cost-Benefit Analysis of Registered Apprenticeship in 10 States*. Mathematica Policy Research: Washington. Accessed at https://wdr.doleta.gov/research/FullText_Documents/ETAOP_2012_10.pdf on 10/4/2021.
Revised Code of Washington § 50.04.145 *Employment—Exclusions*.
Rosecrance, J. C., Cook, T. M., Anton, D. C., & Merlino, L. A. (2002). Carpal tunnel syndrome among apprentice construction workers. *American Journal of Industrial Medicine, 42*(2), 107–116.
Rosenman, K. D., Gardiner, J. C., Wang, J., Biddle, J., Hogan, A., Reilly, M. J., ... Welch, E. (2000). Why most workers with occupational repetitive trauma do not file for workers' compensation. *Journal of Occupational and Environmental Medicine, 42*(1), 25–34.
Sahl, J. D., Kelsh, M. A., Haines, K. D., Sands, F., & Kraus, J. (1997). Acute work injuries among electric utility linemen. *American Journal of Industrial Medicine, 31*(2), 223–232.
Shannon, H. S., & Lowe, G. S. (2002). How many injured workers do not file claims for workers' compensation benefits? *American Journal of Industrial Medicine, 42*(6), 467–473.
Silverstein, B., Viikari-Juntura, E., & Kalat, J. (2002). Use of a prevention index to identify industries at high risk for work-related musculoskeletal disorders of the neck, back, and upper extremity in Washington state, 1990–1998. *American Journal of Industrial Medicine, 41*(3), 149–169.
Spector, J. T., Adams, D., & Silverstein, B. (2011). Burden of work-related knee disorders in Washington State, 1999 to 2007. *Journal of occupational and environmental medicine, 53*(5), 537–547.
Standards of Apprenticeship, 29 C.F.R. § 29.5 (b)(9).
US Department of Labor. (2020). *Federal Workload Data: Top 30 Apprentice Occupations for Fiscal Year 2020*. Accessed at https://www.dol.gov/agencies/eta/apprenticeship/about/statistics/2020 on 4/6/2021.
US Department of Labor. (2021). US Department of Labor awards more than $130m in grants to support registered apprenticeship programs; increase employment opportunities [Press Release issued June 22, 2021]. Accessed at https://www.dol.gov/newsroom/releases/eta/eta20210622 on 9/27/2021.

Washington State Apprenticeship and Training Council. (2021). *Apprenticeship Program Standards Adopted by Construction Industry Training Council of Washington - Plumber*. 2020. Accessed at https://www.lni.wa.gov/licensing-permits/apprenticeship/_docs/0636.pdf on 4/6/2021.

Washington State Apprenticeship and Training Council. (2021). Apprenticeship Program Standards adopted by Seattle Area Plumbers, Housing Plumbers, Pipefitters, HVAC-Service Controls Technician/Refrigeration Mechanics, and Marine Pipefitters Apprenticeship Committee. Accessed at https://www.lni.wa.gov/licensing-permits/apprenticeship/_docs/0115.pdf on 9/29/2021.

Sara Wuellner is an epidemiologist with the Washington State Department of Labor and Industries. Her work is focused on occupational injury and illness surveillance data and methods.

David Bonauto David Bonauto is the Research Director for the Safety and Health Assessment and Research for Prevention (SHARP) Program and Associate Medical Director for the Washington State Department of Labor and Industries.

Journal of Safety Research 83 (2022) 269–281

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

A spatiotemporal risk prediction of wildlife-vehicle collisions using machine learning for dynamic warnings

Raphaela Pagany

Institute for Applied Informatics, Deggendorf Institute of Technology, Freyung, Germany
Interfaculty Department of Geoinformatics, University of Salzburg, Salzburg, Austria

ARTICLE INFO

Article history:
Received 12 September 2021
Received in revised form 13 May 2022
Accepted 1 September 2022
Available online 16 September 2022

Keywords:
Wildlife-vehicle collision
Machine learning
Spatiotemporal prediction
Random forest
Neural networks
Support vector machine classifier

ABSTRACT

Introduction: The technology in the automotive industry is becoming increasingly safer in the age of automated driving, but the number of accidents is still high, especially in wildlife-vehicle collisions (WVCs). To better avoid these accidents, patterns involved in these accidents must be detected. *Method*: This paper presents a spatiotemporal risk prediction of WVCs, including various road and environmental characteristics. A process of data preparation using GIS automated by Python scripts was developed to enable a spatiotemporal link of diverse features for the subsequent predictive data analysis. Different machine learning (ML) approaches were applied- random forest (RF), feedforward neural networks (FNN), and support vector machine classifier (SVM) - including automated ML to predict the risk of WVCs. Therefore, a dataset of approximately 731,000 accidents reported to the police in Bavaria over a period of 10 years (2010–2019) was used. In addition, non-accidents were randomly generated in Python over time and space for the supervised ML processes. As the actual risk probability for WVCs and non-WVCs is not entirely known, the impact of different training ratios between accidents and non-accidents was tested on the risk prediction quality (RPQ) (25%, 50%, 75%, 90% WVCs) of the double-weighted sensitivity and single-weighted specificity rate. *Results:* The test yielded high mean values of RPQ as an indicator for a suitable WVC prediction. Both RF (86.6%) and FNN (86.7%) were identified as suitable choices for WVC risk prediction in terms of RPQ. The SVM yielded a lower prediction quality, even though acceptable results could be achieved within a shorter runtime. *Practical Applications:* Spatial transferability was verified since the algorithm was trained on the dataset of Bavaria (excluding Upper Bavaria) and successfully tested in Upper Bavaria. WVC forecasts were also proven through training with datasets from 2010-2017 and in prediction for 2018 and 2019.

1. Introduction

Injuries and fatalities due to traffic accidents with wildlife have a great impact on the society and wildlife ecology. Vehicle damages cause significant financial costs, even when the actual numbers of wildlife-vehicle collisions (WVCs) are often unknown. German vehicle insurers registered nearly 300,000 accidents with wildlife in 2019 - more than ever before (GDV, 2020). In total, WVCs cost German insurers around 885 million euros. In other countries, the WVC rates are also estimated to be significantly high, for instance, four million in Belgium (Morelle, Lehaire, & Lejeune, 2013), 1.5 million collisions with deer in the United States (Sullivan, 2011), and five million with amphibians and reptiles in Australia per year (Duellman, 2001). But while the collisions are mostly counted in small regions, nationwide or international numbers are estimated to be several times higher than the reported collision numbers (Gkritza, Baird, & Hans, 2010; Steiner, Leisch, & Hackländer, 2014).

Predictive analysis seems to be a useful approach to analyze WVC risk and its influential conditions in a spatiotemporal manner aiming to transfer the risk prediction model globally, for superregional mitigation possibilities. Subsequently, the transferability of such a model would decimate the need for WVC data as a prerequisite for adequate measures in specific areas. By studying the environmental characteristics of the roadway, topography or weather, patterns can be identified to determine if there is a risk of an accident. By using these patterns, predictive analyses enlarge the knowledge about the accident risk, even if little is known about the accident or non-accident locations.

This study applies three ML algorithms - random forest, feedforward neural networks, and support vector machine classifier - to achieve a spatiotemporal WVC risk prediction model and, hence,

E-mail address: raphaela.pagany@th-deg.de

https://doi.org/10.1016/j.jsr.2022.09.001
0022-4375/

a basis for greater road safety and wildlife conservation. By using GIS tooling automated with Python, heterogeneously structured data regarding the time and the environmental conditions can be brought to one key reference. In this feature engineering process, different spatiotemporal resolutions were harmonized to obtain a pre-processed dataset for further analyzing steps. Road, infrastructure, land use, weather, topographical and further geo-data were identified as potentially influential. A spatiotemporal resolution of one hour and approximately 25-metres road sections was chosen for a detailed WVC analysis. The predictive quality of the algorithms, and their spatial and temporal transferability were examined. For an improved prediction, automated ML (AutoML) was applied. This study addresses three research questions as follows: (1) Are ML methods suitable for an appropriate risk prediction WVCs? (2) How must the training data be structured to enable a high quality of WVC risk classification? (3) Can the approach be transferred temporally and spatially?

The paper is organized as follows: After the literature review (Section 2), the input data for the prediction model and the data pre-processing are described, before the case study region is presented (Section 3). Section 4 describes the analyzing structure including the selection of algorithms and parameters, and evaluation metrics. In Section 5, the results are presented, while in Section 6, the results and advantages of the findings for ecology and traffic management are proposed. Finally, an outlook for future work is given in Section 7.

2. Literature review

Several studies have been employed to address the problem of low data availability since years, either by monitoring carcasses along the road networks (Grilo, Ferreira, & Revilla, 2015; Lutterschmidt, Weidler, & Schalk, 2019), by voluntary citizen science approaches (Heigl et al., 2016; Keken, Sedoník, Kušta, Andrášik, & Bíl, 2019), or by obliging the driver to report the accident to the police or to the local hunter (Snow et al., 2018). Especially in countries with dense road networks, national parks, or high wildlife populations, diverse research teams collected data to study WVCs and their environment (Arevalo, Honda, Arce-Arias, & Häger, 2017; Cherry et al., 2019; Pagany, 2020).

Over the last few decades, various studies constitute that the risk for WVCs is influenced by the drivers' and animals' behavior due to diverse environmental conditions. The temporal factor determines the daily or annual behavior and, thereby, correlates with land use, weather, or light conditions (Garriga, Franch, Santos, Montori, & Llorente, 2017; Kämmerle et al., 2017; Steiner et al., 2014). These dynamical patterns are completed by statistical patterns which together affect the accidents in a spatiotemporal manner. For instance, road conditions, such as the street width, a gentle topography, and the proximity to forest influence the WVC risk, according to a comprehensive review of WVC studies worldwide (Pagany, 2020). In contrast, traffic volume, the distance to urban areas, or road accompanying infrastructure are not clearly assignable influencing or non-influencing factors.

However, most of the studies were undertaken in small areas or considering road sections of only a few kilometers (Gunson, Mountrakis, & Quackenbush, 2011; van der Ree, Jaeger, van der Grift, & Clevenger, 2011), which means that these studies lack information required to mitigate WVCs on a higher level. In addition, the relationship between space and time in WVC is also only sparsely studied. However, this investigation would be essential for dynamically adapted measures in the event of a high WVC risk at a specific time at a specific location, as Pagany (2020) discussed. The review also shows that most studies only analyze a few influencing parameters.

Until now, only a few studies developed a prediction model for WVCs. For instance, Santos, Mota-Ferreira, Aguiar, and Ascensão (2018) developed a Bayesian hierarchical occupancy model that estimated WVC risk, especially in agricultural, open habitats, and within four-lane road sections. Another example is the study of Visintin, van der Ree, and McCarthy (2016), which predicted WVCs with the main determinants being traffic volume, traffic speed, and species occurrence. Both studies explicitly named the advantage of predictive studies, as the analysis decimate the need for a broad data collection. Besides, classic statistics and geo-statistical analyses are chosen in previous studies. They applied hot spot analyses such as the kernel density distribution, as presented in Bíl, Andrášik, Svoboda, and Sedoník (2016), and calculated the density of accidents along certain road section lengths. As further statistical methods for WVC distribution, Malo, Suárez, and Díez (2004) used a Poisson distribution, Valero, Picos, and Álvarez (2015) applied a nearest-neighbor hierarchical clustering, and Tanner, Leroux, and Saunders (2017) used a generalized linear mixed model. Furthermore, Liu, Nieuwenhuis, and McCullagh (2018) and Seo, Thorne, Choi, Kwon, and Park (2015) applied regression analysis to determine the influence of environmental factors such as the roadside land use or seasonal effects.

Studies about other types of traffic accidents used data mining techniques such as machine learning (ML) for identifying the accident risk. Datasets with georeferenced accident occurrences combined with environmental characteristics are often included in the risk analysis. Yang, Chen, and Brown (1999) investigated the severity of road accident injuries in Alabama, United States. A backpropagation neural network was trained including variables such as light conditions and traffic speed to detect safer driving patterns, which may reduce fatalities and injuries by up to 40%. Komol et al. (2021) chose ML-based classification approaches for modeling injury severity of the vulnerable road users: pedestrian, bicyclist, and motorcyclist using k-nearest neighbor, support vector machine, and random forest. They found that motorcyclists have an especially high crash severity. Another study about injury severity is Chen, Song, and Ma (2019). Using a random parameters bivariate ordered probit model, they showed correlations between two drivers' injuries such as driver age, gender, vehicle, airbag or seat belt use or traffic flow. Abdelwahab and Abdel-Aty (2001) applied multilayer perceptron (MLP) and fuzzy adaptive resonance theory (ART) neural networks and included parameters such as the vehicle types and the sex of the driver to predict the severity of the driver's injury in Florida, United States.

In addition, Deublein, Schubert, Bryan, and de Soto Boria (2015) developed an accident risk model using Bayesian Probabilistic Networks including road-related variables, such as the number of lanes, curvature, and traffic characteristics. With a correct prediction of 86.53% of the road segments with a tolerance of 25%, the model could be applied to predict spatial "black spot" locations regarding the accidental injuries on the Swiss highway network. Beshah and Hill (2010) used a rule mining approach to classify road traffic accidents using adaptive regression trees for a case study in Ethiopia. The study included diverse data about the drivers' age and driving experience, the vehicle age, the road surface, as well as light and weather conditions. Besides, recent approaches can be mentioned that are characterized by similar pre-processing steps of GIS data like for accident risk prediction (e.g., the prediction of connected vehicle movement at intersections). Trajectory movement labelling for predicting driver movements and for warning at intersections ensure advanced transportation safety using ML-based approaches to implement these technologies (Komol et al., 2021; Shrivastava, Verma, & Jain, 2021).

Other studies compared ML techniques to test their quality or predictive accuracy. Chang and Chen (2005), for instance, compared a classification and regression tree (CART) with a negative binomial regression model to analyze the relationship between traffic accidents and highway geometric variables, traffic characteristics, and environmental factors in Taiwan. They found that the CART model is an appropriate method for analyzing the frequency of freeway accidents. Precipitation variables and daily traffic volume are the key determinants for the accident risk. Chong, Abraham, and Paprzycki (2005) compared the performance of neural networks, support vector machines, decision trees, and a concurrent hybrid model involving decision trees and neural networks. The severity of an injury resulting from a traffic accident was predicted. The hybrid approach of decision tree and neural networks outperformed the individual approaches.

The applications and the comparisons of diverse ML techniques to model traffic accidents show that these approaches are suitable for accident risk prediction. Several studies are undertaken to understand the characteristics of driver behavior, road conditions, environmental and weather conditions that are related to the accident probability. However, for a risk prediction, it is not decisive to know the impact of an individual factor, as analyzed by classical statistics, but to develop a decision model considering the environmental factors and the learning characteristics of ML techniques. Thus, ML may also enlarge the knowledge about the accidents with wildlife, as already applied for other road accidents types. A supervised ML classification seems to enable this WVC risk prediction.

3. Material and data preparation

3.1. Data description

For this study, a dataset of 730,867 WVCs that were officially registered by the Police between 2010 and 2019 in Bavaria, Germany was used (StMI & Bavarian State Police, 2020). Each of these registered WVCs from the last 10 years contains a georeference and a timestamp. Additional features were used that are identified in previous studies as influential characteristics for road traffic accidents with wildlife (Pagany, 2020) (Table 1): *road-related* features, such as road class, road width, infrastructure elements, (e.g., middle barriers, road barriers, and additional lanes), and topographic features (e.g., ditches, or slopes aside the road network). In addition, *weather* data were used, such as air temperature, precipitation, and cloud coverage, as well as *land use* data in the prediction model. The land use data are classified in 12 classes as follows: artificial land, cropland seasonal, cropland perennial, forest broadleaved, forest coniferous, forest mixed, shrubland, grassland, bare land, water, wetland, and snow and ice areas. Based on the accident data, *time-related* features were derived, namely the day of the year, the weekday, the hour of a day, and the solar altitude.

The road information was obtained from the database BAYSIS (Bavarian road infrastructure system) provided by the Bavarian Ministry of the Interior, for Sport and Integration (StMI, 2017) and the Official Topographical-Cartographic Information System Data - Digital Landscape Models (ATKIS-DLM data) by the Surveying Authorities of the Federal States of Germany (Surveying Authorities of the Federal States of Germany, 2015). Weather data from the German weather service (DWD, 2019) were used. The DWD obtains its data from COSMO-REA6 - a regional reanalysis dataset (product of a numerical data model that assimilates terrestrial and remotely sensed weather data) that covers the EUROCORDEX area with a spatial resolution of 6 × 6 km and uses the ERA-INTERIM global reanalysis dataset for boundary conditions. For the land use classification, PANGAEA land use data with a spatial resolution of 30 × 30 m based on remote sensing Landsat and LUCAS data (Pflugmacher, Rabe, Peters, & Hostert, 2018) from the year 2018 were used in the prediction model.

3.2. Data pre-processing

The different datasets required data preparation and cleaning for the later data mining process. Nearly all data required a trans-

Table 1
Data and derived features as input for the ML predictive analysis (not normalized values).

	Data Source	Feature/Label	Unit	Range
Wildlife-vehicle collisions (WVCs) (spatiotemporal)	StMI & Bavarian State Police, 2020	WVC label	-	1
Non-WVCs (spatiotemporal)	Randomly calculated		-	0
Road-related (spatial)	ATKIS (mainly)	Road class		1–5
	(Surveying Authorities of the Federal States of Germany (2015), 2015)	Road width	metre	2 – 22
	BAYSIS data (StMI, 2017)	Middle barrier	-	0–1
		Road barriers	-	0–1
		Additional lanes	-	0–1
		Ditch	-	0–1
		Slope	-	0–1
Weather (spatiotemporal)	Weather data (DWD, 2019)	Ambient temperature	°C	−22.20 to 39.07
		Precipitation	litre/m^2	0–72.95
		Cloud coverage	%	0–100
Land use (spatial)	PANGAEA land use data (Pflugmacher et al., 2018)	Artificial land	%	0–100
		Cropland, seasonal	%	0–100
		Cropland, perennial	%	0–100
		Forest, broadleaved	%	0–100
		Forest, coniferous	%	0–100
		Forest, mixed	%	0–100
		Shrubland	%	0–100
		Grassland	%	0–100
		Bare land	%	0–100
		Water	%	0–100
		Wetland	%	0–100
		Snow and ice	%	-
Time-related (temporal)	WVC data (StMI & Bavarian State Police, 2020)	Day of year	-	1–366
		Weekday	-	1–7
		Hour of the day	-	0–23
		Solar altitude	°	−64.50–64.50

formation of geographical projections and data pre-processing to obtain a homogenous key component, with which the features WVCs, and non-WVCs (as label) could be correlated. GIS methods automated by Python scripts were used, namely GDAL and OGR from the OSGeo library (OSGeo, 2019) for interacting with geodata file formats, and shapely (Toblerity, 2020) for geometric processing. The road shapefile ver01_l of the ATKIS-DLM data were divided into point vectors with a distance of 25-metres along the road network to define the set of locations to be examined (Fig. 1). These locations were used as a basis layer to extract the feature values from several other datasets.

From the ATKIS-DLM data, the information about the road class and road width was obtained. Missing values of the road classes were derived from the road names, for example, the federal road class (German: B = Bundesstraße) from the street name "B 12," while the missing road width were derived from the road classes using standardized values for road construction in Germany (11.5 m for highways, 7.5 m for federal roads, 6.5 m for state roads, and 5.5 m for municipal road (FGSV, 1996). To avoid duplicate information, road axes of highways were deleted as the lane axes are also registered for highways that run parallel to the road axis. The road information was augmented by the street data of BAYSIS. According to this road database, the road network was classified in sections with middle barriers, additional lanes, ditches, and slopes (Table 1). Details in the attribute column legend of the BAYSIS shapefile Straßenbestand.Deckschichtalter were used for the classification (e.g., hard shoulders or bus bays were classified as roads with an additional lane, while protective stripes with crash barriers were classified as sections with barriers). To obtain all infrastructure elements at both sides of the road, vertical lines at the 25-metre distance points from the ATKIS data were constructed. Along these vertical lines, all elements of the BAYSIS shapefile were identified, and the information were written into the 25-metre point layer as key elements for the later data mining.

The land use classes of the PANGAEA data were identified based on the 25-metre points. The land use types were included on both sides of the road within three different radii to test the impact of the environmental land use of an accident within near, medium, and long distances (Fig. 1, semicircles in the right image). The relative area share of each of the 12 different land use classes was calculated separately for both sides of the street within semicircles of 100, 500, and 200 meter radii. Thus, six different buffers were used for each land use type to model the impact of the land use structure at each potential accident location for ML analysis.

Furthermore, the CSV dataset of WVCs was stored as a shapefile and joined to the nearest of the 25-metre points to create a spatial reference. For the temporal reference, the timestamp of the WVCs were discretised to an hourly resolution. In total, 730,867 WVCs were registered but only 567,233 WVCs (78%) were included in the predictive analysis, since the remaining accidents do not contain any geographic references. A set of non-WVC locations was generated in addition to the WVC registrations that are required for the ML approach. A number of 657 billion spatiotemporal points for potential accident occurrences could be obtained by multiplying the 7.5 million 25-meter road points and 87,600 hours of the 10-year study period, but only 12 million non-WVCs are necessary for the latter analysis (see chapter 4). Hence, spatiotemporal points were randomly generated in an hourly resolution based on the 25-meter distance points of the ATKIS road network using the random library in Python based on the Mersenne Twister algorithm, while spatiotemporal locations with the locations and timestamps of the WVCs were excluded to obtain non-WVCs for supervised ML training. For each of the accidents and non-accidents, the timestamp was used to obtain the features weekday, day of the year, and hour of the day. In addition, the solar altitude was calculated for each spatiotemporal point using the solar position algorithm (Holmgren, Hansen, & Mikofski, 2018). Finally, the features air temperature, precipitation, and cloud coverage were extracted from the weather dataset by determining the fitting tile of the weather data at the specific point and at the specific time.

A SQLite database was used to combine the different generated datasets in one dataset with a label (WVC/non-WVC) as input for the ML process. After the spatiotemporal points of non-WVCs and the time-based features were generated, they were stored directly in tables in the database. The same was applied for the weather data, which was selected according to the spatiotemporal locations. The road infrastructure, land use, and WVC data were stored in shapefiles as an intermediate step and, later, imported into the database as separate tables for each shapefile. Finally, each data row has a reference to one of the 25-m points from the ATKIS dataset using WGS 84 UTM system (World Geographical System 84, Universal Transverse Mercator) as geographical reference, and an hourly resolution (the hour of the day and day of the year) as temporal dimension, which are included in all tables as key elements. The normalization of the data was done as part of the SQL statement that joined the several tables to one source table for ML processing. In this table, each spatiotemporal point is characterized by 88 features, whereby-six classes were deleted due to the nonexistence of snow and ice areas (at both roadsides within three radii).

3.3. Case study area

The data were applied for WVC analysis in Bavaria. The Free State of Bavaria in the southeast of Germany covers 7,548 km^2

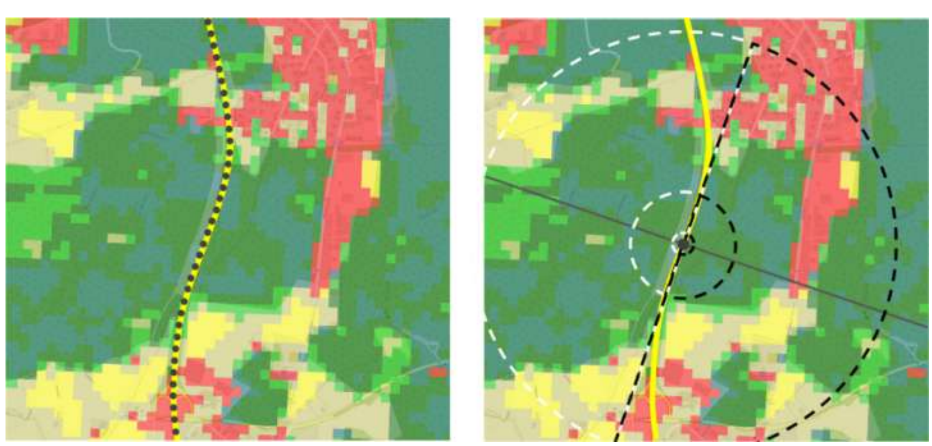

Fig. 1. Data preparation using spatiotemporal points of 25-metres road sections (left image), including, e.g., land use environment within 100-, 500-, and 2000-metres radii (right image).

and seven administrative districts: Lower Bavaria, Upper Bavaria, Lower Franconia, Middle Franconia, Upper Franconia, Upper Palatinate, and Swabia (Fig. 2; Table 2).

Bavaria borders the Czech Republic in the east and Austria in the south-east and south. While a significant share of the around 13 million inhabitants in Bavaria live in a few metropolitan areas (such as Munich or Nuremberg-Fuerth-Erlangen), large parts of Bavaria, such as the Upper Palatinate and Lower Bavaria, are sparsely populated. The northern and eastern parts with the densely forested Bavarian Forest belong to the German low mountain range, while the mostly flat glacial foothills extend from the Alps to the Danube basin. Almost 35% of the Bavarian area is forested. Because of the altitude above sea level and the distance from the sea, Bavaria's climate is rougher and more continental than in most other German federal states. WVCs mainly occur in rural areas, for example, in the county Rottal-Inn in the southeast of Lower Bavaria (Fig. 3). There is a low rate of accidents in the Alpine and forested regions, such as the national parks Bavarian Forest and Berchtesgadener Land, as the road network is not as dense as in other areas. Despite the dense road network, there is a low density of WVCs in urban areas. The WVC density is the highest on state and county roads in relation to the road length (Table 3).

4. Analysis

4.1. Selection of algorithms and parameters

The prediction analysis determines the collision risk in order to warn drivers of WVCs only when necessary. The risk for WVCs was calculated using ML algorithms in Python. The following algorithms were used for the analysis: random forest (Breiman, 2001), feedforward neural networks (McCulloch & Pitts, 1943), and support vector machine classifier (Vapnik, 1963). Since there are no previous studies for risk prediction of WVCs with ML available, several variants of ML algorithms were tested. The algorithms are chosen to test the prediction capability - in contrast to the commonly used statistical hotspot and impact analysis in WVC research (Pagany, 2020). Research on traffic accidents in general already applies the chosen ML algorithms, while there is still a research gap for WVC prediction using such methods. Especially the temporal forecast and the WVC risk prediction in areas without historical datasets about WVCs are an advantage of the ML techniques. In addition, these methods enable a steady training adjustment of the algorithms due to changed environmental conditions. Therewith, dynamic real-time warnings to the driver via a mobile application (wuidi, 2020) is applicable in order to avoid accidents. Hence, the paper focuses on the generation of ML models resulting in high prediction rates through testing.

The impact of different training ratios between accidents and non-accidents on the prediction quality was tested, as the actual risk probability of WVCs and non-WVCs is not entirely known. Only the locations and times with a registered WVC are an indicator for spatiotemporal risk zones, while the remaining road sections and times are assumed to be low risk. While the probability is high that only one, a vehicle, or an animal, or neither of them will cross a certain road at the same time, there is also a possibility that both a vehicle and an animal cross at the same point at the same time, but that the individual reaction of the driver or the animal is such that a collision can be avoided. As the WVC data do not allow any information about these near misses, but the risk is higher than the actual WVC data suggest, higher, and also extreme ratios of WVCs were used for ML training (25%, 50%, 75%, 90% WVCs) (Fig. 4). Afterwards, the models were tested with a uniform ratio of 5% WVCs and 95% non-WVCs to obtain comparable results. Hence, the ML training were conducted with the Bavarian WVC data and different sample sizes of non-WVCs for the predictive risk analysis.

To test the temporal and spatial transferability, WVC data were split in three ways, and the training and test data were enriched with non-WVCs to obtain the abovementioned ratios for training

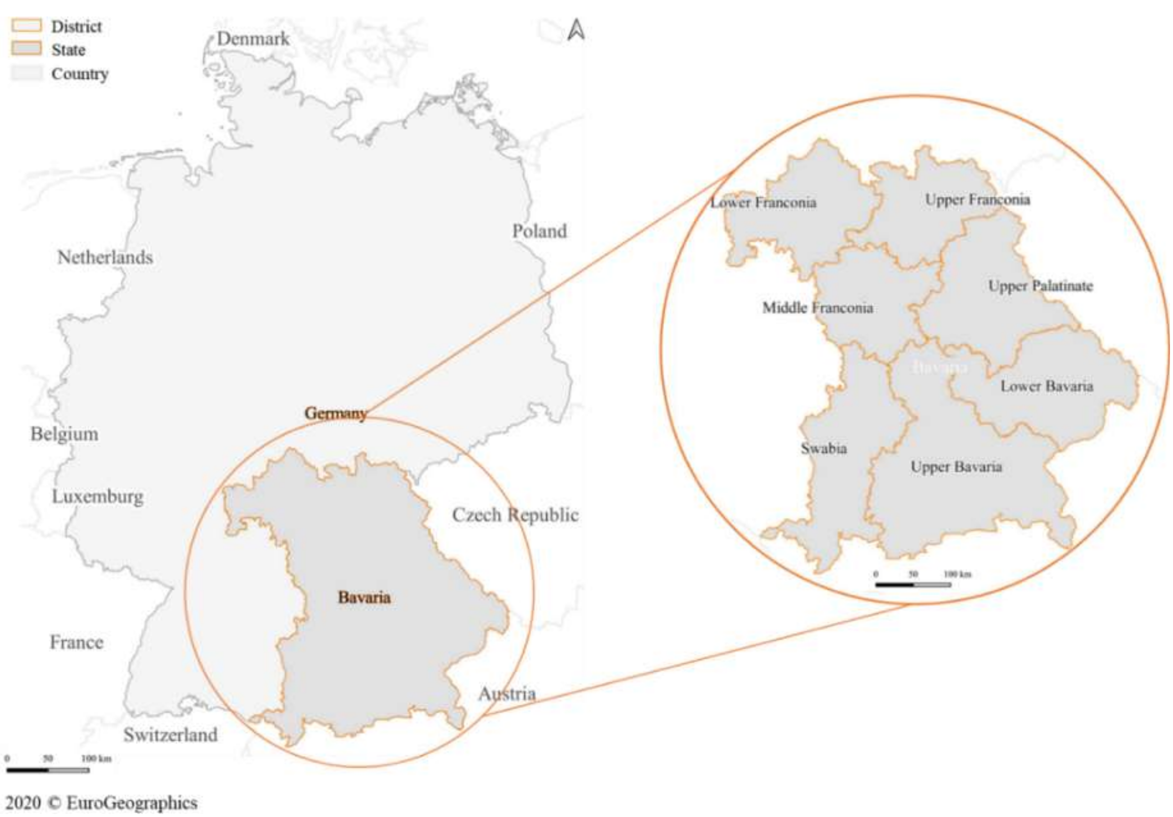

Fig. 2. Case study area of Bavaria (map on the left side) and districts (map on the right side).

Table 2
Population, area, and road length per districts based on ATKIS data and Destatis, (2020).

District	Population		Area		Road length	
			km^2	%	km	%
Lower Bavaria	1,192,641	9.48%	10,330	14.6%	44,436,083	25.96%
Upper Bavaria	4,418,828	35.12%	17,530	24.8%	27,839,851	16.26%
Lower Franconia	1,315,882	10.46%	8,531	12.1%	21,935,317	12.82%
Middle Franconia	1,717,670	13.65%	7,245	10.3%	17,027,147	9.95%
Upper Franconia	1,067,988	8.49%	7,231	10.2%	18,851,596	11.01%
Upper Palatinate	1,081,800	8.60%	9,691	13.7%	17,124,381	10.00%
Swabia	1,788,729	14.21%	9,992	14.2%	23,950,389	13.99%
Total	12,583,538	100.00%	45,476	100.00%	171,164,764	100.00%

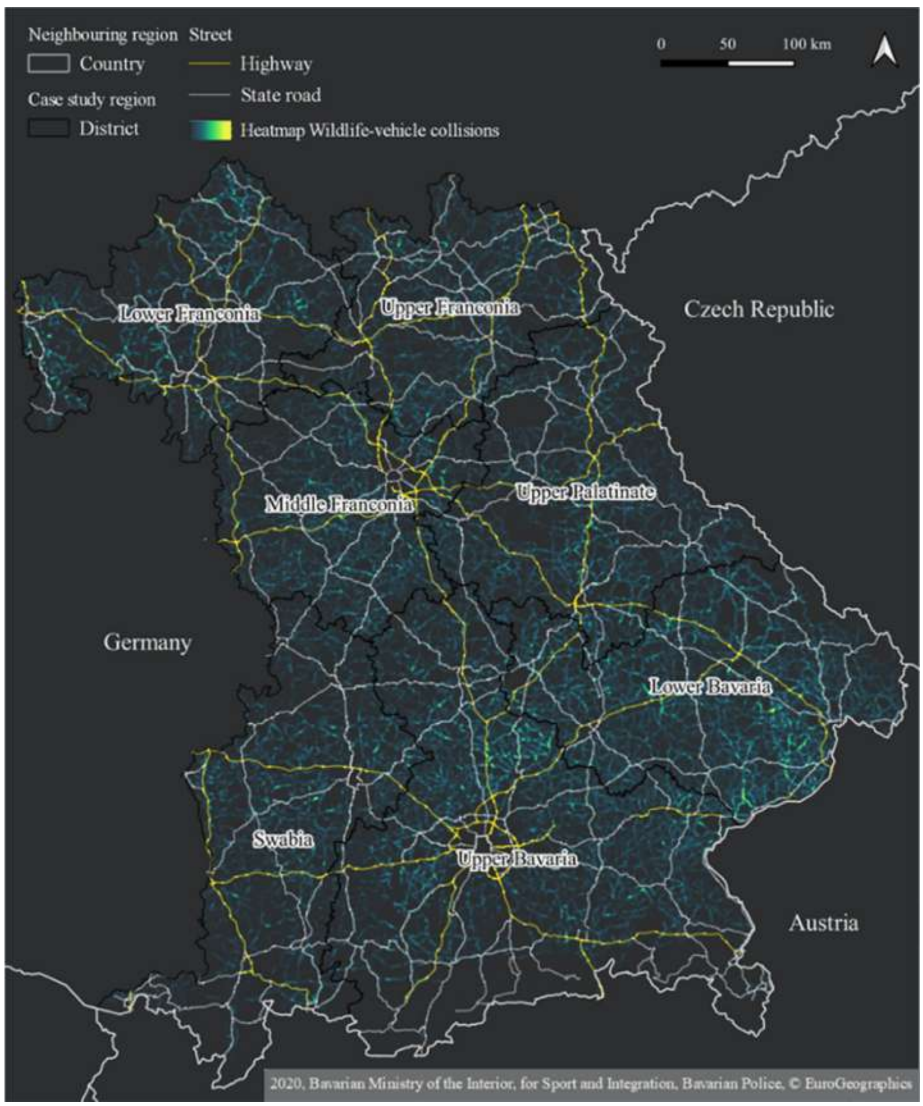

Fig. 3. Heatmap of WVCs in Bavaria.

Table 3
Distribution of WVCs by road class.

Road class	WVC share (%)
Municipal roads	17.09
Country roads	31.43
State roads	33.90
Federal roads	14.88
Highways	2.70

and testing. The random split in a ratio of 80% training data and 20% test data were used as a basis for comparison. Furthermore, the predictive analyses were carried out with a training sample containing WVCs and non-WCVs from between 2010 and 2017, while data from 2018 and 2019 were used for testing (time-based split). Also, the algorithm that was trained on the dataset from Bavaria except Upper Bavaria, was then tested in Upper Bavaria (geographical split) (Fig. 4).

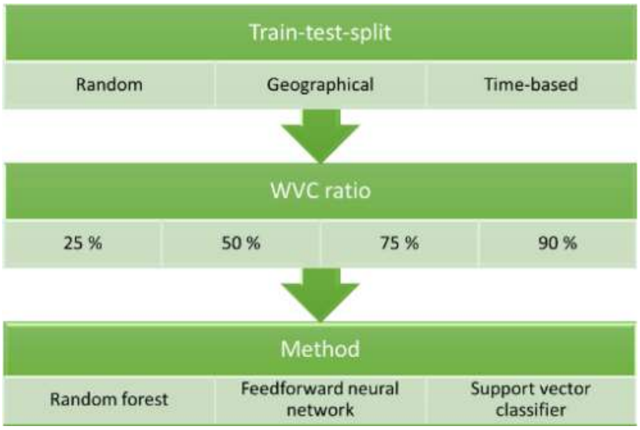

Fig. 4. Analyzing structure for the WVC risk prediction.

From the class of neural networks, the feedforward neural networks performed best in a pre-test, where convolutional, recurrent, and feedforward neural networks were analyzed. Thus, convolutional networks and recurrent networks, which are typically used for image classification and temporal processes, respectively, can be excluded. With support vector machine classifier, a very fundamental algorithm was added. In addition, random forest was chosen as third option to have an algorithm, which yielded good results in previous studies and enables impurity calculation for the input features.

Furthermore, the hyperparameter tuning was done in an automated way using automated ML (AutoML) systems. The StructuredDataClassifier from AutoKeras was used, which is based on the TensorFlow library for the feedforward neural networks (Jin, Song, & Hu, 2019). Moreover, auto-sklearn, which is an AutoML toolkit for the scikit-learn library, was applied for the random forest and the support vector machine classifier (Feurer, Klein, Eggensperger, Springenberg, Blum, & Hutter, 2015). Even though the hyperparameters are chosen by AutoML systems, some specifications needed to be delimited.

In AutoKeras, it is necessary to specify the count of trials and the epochs for the model training. Ten trails and 10 epochs were chosen, as tests with 5 trails were insufficient for adequate results and the training with 15 trails did not improve the results significantly but took an inappropriately long runtime. To determine the number of epochs, several trainings were run with 20 epochs and examined at which number of epochs the accuracy and loss functions did not improve significantly. Using Auto-sklearn, the run time of the entire process and of a single model training was specified, which indirectly defines the number of different hyperparameter settings. In addition, a number of 10 cores were continually used for all model generations. For the further parameters, default values were used.

4.2. Evaluation metrics

To evaluate the results, confusion matrices (Table 4) and further indicators were generated to show the numbers of correct and incorrect matches of the classification. With the accuracy rate, all correctly classified matches are related to the overall number of matches of the confusion matrix (Eq. (1)). The **sensitivity**, also called recall, is an indicator that shows how often WVCs are correctly classified in relation to the real number of WVCs (Eq. (2)). **Specificity** shows how often non-WVCs are correctly classified in relation to the real number of non-WVCs (Eq. (3)). While the **false negative (FN) rate** is the inverse value of the sensitivity rate, the **false positive (FP) rate** is the inverse of the specificity rate.

Table 4
Confusion Matrix.

	WVC (1) predicted	*Non-WVC (0) predicted*
WVC	actual	Actual/true positive (TP), meaning WVC occurred, and is predicted to be positive
False	negative (FN), meaning WVC occurred, and is predicted to be negative	
Non-	WVC actual	False positive (FP), meaning no WVC occurred, but is predicted to be positive
True	negative (TN), meaning no WVC occurred, and is predicted to be negative	

$$accuracy\ \ rate = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$specificity = \frac{TN}{TN + FP} \tag{3}$$

Especially high specificity and sensitivity rates are useful indicators for evaluating the quality of a WVC risk prediction model concerning its application as warning service. Therefore, an own measure, the **risk prediction quality**, is defined as the follows: the mean value of the double-weighted sensitivity and single-weighted specificity rate (Eq. (4)).

$$risk\ \ predictio\ \ nquality(RPQ) = \frac{2 * sensitivity + specificity}{3} \tag{4}$$

With the double-weighted sensitivity, the emphasis is laid on the potential to correctly classified WVC risk in contrast to false classified high WVC risk. This sensitivity value is the basis for a warning in the case of WVC risk and only a few non-warnings even if the risk would be high. Too many warnings due to low specificity rates would annoy the drivers without a reason for an accident risk.

5. Results

With accuracy rates between 87.36% and 92.00%, all three ML algorithms assign most observations correctly as WVCs or non-WVCs (bold values in Table 5), using a ratio of 25% WVCs and 75% non-WVCs in the training dataset. With RPQ values of 86.7% and 86.6%, neural networks and the random forest, respectively, performed slightly better than the support vector machine classifier with its highest value of 84.6% (Fig. 5, black line) (Table 6; bold values > 75th quantile). The training ratio of 25% WVCs might achieve the best accuracy rates, but it is not the best choice according to the risk prediction quality measure (see Section 4). While sensitivity rates of under 70% are calculated, more than 30% of the WVCs are predicted as no risk cases, inversely. Using the measure of RPQ, the ratio of 75% WVCs and 25% non-WVCs has been proven as the best training ratio. The training and test run of the different case studies show that it is possible to forecast WVCs in a temporal and in a spatial manner (bold values > 75th quantile; modes: rnd = random, geo = geographical, and time = time-based; Fig. 4). Hence, the applicability of predictive analysis for WVCs is shown.

Based on the random forest, the importance of the individual factors is calculated. The results also show that temporal, environmental and roadside conditions affect the WVC risk (Fig. 6). Espe-

Table 5
Confusion matrix of ML results due to WVC ratio, algorithm (neural network = NN, support vector machine classifier = SVM, random forest = RF), and mode (geographical split = geo; randomly split = rnd; time-based split = time).

			WVC		non WVC		
			pred. Non-WVC = No Warn.	pred. WVC = Warning	pred. Non-WVC = No Warn.	pred. WVC = Warning	
WVC ratio	Algorithm	Mode	FN	TP	TN	FP	Accur %
25%	NN	Geo	45,027	88,600	2,358,856	180,057	**91.58**
		Rnd	34,067	79,379	1,977,663	177,811	**90.66**
		Time	39,280	84,946	1,902,819	161,669	**90.82**
	SVC	Geo	47,420	86,207	2,294,631	244,282	**89.09**
		Rnd	39,785	73,661	1,919,229	236,245	**87.83**
		Time	42,538	81,688	1,830,269	234,219	**87.36**
	RF	Geo	61,119	72,508	2,386,359	152,554	**92.00**
		Rnd	39,622	73,824	1,953,723	201,751	**89.36**
		Time	41,418	82,808	1,859,100	205,388	**88.72**
50%	NN	Geo	19,563	114,064	2,128,044	410,869	83.89
		Rnd	14,355	99,091	1,743,486	411,988	81.21
		Time	16,158	108,068	1,691,907	372,581	82.24
	SVC	Geo	21,475	112,152	2,047,878	491,035	80.82
		Rnd	16,102	97,344	1,666,100	489,374	77.72
		Time	18,361	105,865	1,603,721	460,767	78.11
	RF	Geo	19,356	114,271	2,087,965	450,948	82.40
		Rnd	12,040	101,406	1,638,286	517,188	76.67
		Time	14,051	110,175	1,622,192	442,296	79.15
75%	NN	Geo	7489	126,138	1,812,158	726,755	72.53
		Rnd	4867	108,579	1,434,604	720,870	68.01
		Time	5503	118,723	1,385,820	678,668	68.74
	SVC	Geo	7867	125,760	1,666,913	872,000	67.08
		Rnd	5766	107,680	1,332,631	822,843	63.48
		Time	6116	118,110	1,283,013	781,475	64.02
	RF	Geo	8546	125,081	1,839,873	699,040	73.52
		Rnd	4821	108,625	1,445,009	710,465	68.47
		Time	5122	119,104	1,388,963	675,525	68.90
90%	NN	Geo	1931	131,696	1,305,375	1,233,538	53.77
		Rnd	2099	111,347	1,173,934	981,540	56.65
		Time	1945	122,281	1,098,623	965,865	55.78
	SVC	Geo	1506	132,121	1,055,130	1,483,783	44.42
		Rnd	1231	112,215	847,635	1,307,839	42.30
		Time	1129	123,097	793,250	1,271,238	41.87
	RF	Geo	2251	131,376	1,348,305	1,190,608	55.37
		Rnd	1753	111,693	1,091,939	1,063,535	53.05
		Time	1763	122,463	1,053,540	1,010,948	53.73

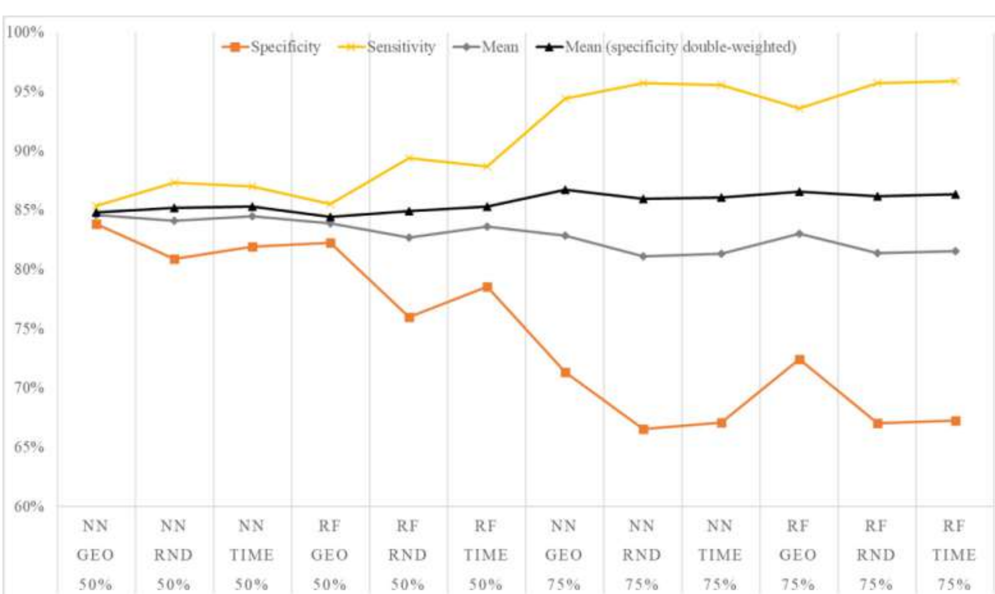

Fig. 5. Sensitivity and specificity rates and their mean values as indicator for WVC risk prediction.

Table 6
Metrics of WVC risk prediction due to WVC ratio, algorithm (neural network = NN, support vector machine classifier = SVM, random forest = RF), and mode (geographical split = geo; randomly split = rnd; time-based split = time).

WVC ratio	Algo-rithm	Mode	WVC		Non-WVC		$\frac{sensitivity + specificity}{2}$	$\frac{FN + FP}{2}$	$\frac{sensitiv.*2 + specific.}{3}$	$\frac{FN*2 + FP}{3}$
			Sensitivity	FN rate	Specificity	FP rate				
			%	%	%	%	%	%	%	%
25%	NN	geo	66.30	33.70	92.91	7.09	79.61	20.39	75.17	24.83
		rnd	69.97	30.03	91.75	8.25	80.86	19.14	77.23	22.77
		time	68.38	31.62	92.17	7.83	80.27	19.73	76.31	23.69
	SVC	geo	64.51	35.49	90.38	9.62	77.45	22.55	73.13	26.87
		rnd	64.93	35.07	89.04	10.96	76.99	23.01	72.97	27.03
		time	65.76	34.24	88.65	11.35	77.21	22.79	73.39	26.61
	RF	geo	54.26	45.74	93.99	6.01	74.13	25.87	67.50	32.50
		rnd	65.07	34.93	90.64	9.36	77.86	22.14	73.60	26.40
		time	66.66	33.34	90.05	9.95	78.36	21.64	74.46	25.54
50%	NN	geo	85.36	14.64	83.82	16.18	**84.59**	15.41	84.85	15.15
		rnd	87.35	12.65	80.89	19.11	**84.12**	15.88	**85.19**	14.81
		time	86.99	13.01	81.95	18.05	**84.47**	15.53	**85.31**	14.69
	SVC	geo	83.93	16.07	80.66	19.34	82.29	17.71	82.84	17.16
		rnd	85.81	14.19	77.30	22.70	81.55	18.45	82.97	17.03
		time	85.22	14.78	77.68	22.32	81.45	18.55	82.71	17.29
	RF	geo	85.51	14.49	82.24	17.76	**83.88**	16.12	84.42	15.58
		rnd	89.39	10.61	76.01	23.99	82.70	17.30	84.93	15.07
		time	88.69	11.31	78.58	21.42	**83.63**	16.37	85.32	14.68
75%	NN	geo	94.40	5.60	71.38	28.62	82.89	17.11	**86.72**	13.28
		rnd	95.71	4.29	66.56	33.44	81.13	18.87	**85.99**	14.01
		time	95.57	4.43	67.13	32.87	81.35	18.65	**86.09**	13.91
	SVC	geo	94.11	5.89	65.65	34.35	79.88	20.12	84.63	15.37
		rnd	94.92	5.08	61.83	38.17	78.37	21.63	83.89	16.11
		time	95.08	4.92	62.15	37.85	78.61	21.39	84.10	15.90
	RF	geo	93.60	6.40	72.47	27.53	**83.04**	16.96	**86.56**	13.44
		rnd	95.75	4.25	67.04	32.96	81.39	18.61	**86.18**	13.82
		time	95.88	4.12	67.28	32.72	81.58	18.42	**86.34**	13.66
90%	NN	geo	98.55	1.45	51.41	48.59	74.98	25.02	82.84	17.16
		rnd	98.15	1.85	54.46	45.54	76.31	23.69	83.59	16.41
		time	98.43	1.57	53.22	46.78	75.82	24.18	83.36	16.64
	SVC	geo	98.87	1.13	41.56	58.44	70.22	29.78	79.77	20.23
		rnd	98.91	1.09	39.32	60.68	69.12	30.88	79.05	20.95
		time	99.09	0.91	38.42	61.58	68.76	31.24	78.87	21.13
	RF	geo	98.32	1.68	53.11	46.89	75.71	24.29	83.25	16.75
		rnd	98.45	1.55	50.66	49.34	74.56	25.44	82.52	17.48
		time	98.58	1.42	51.03	48.97	74.81	25.19	82.73	17.27

cially the road-related factors road width (19%) and road class (14%) are determinants for the WVC prediction. Also, artificial land in the near and middle circle around the potential WVC locations (6% and 4% for each side s1 and s2; see Fig. 6), and time-based factors, such as sun angle (6%), hour (3%), and day of the year (1%), influence the ML prediction. Therein, the importance of the factors should not be understood as individual impact, but as combination of all factors influencing the prediction model. It is also independent of whether the factors increase or reduce the WVC risk. Finally, the ML analysis only determines the collision risk in order to warn drivers of WVCs only when necessary.

6. Discussion

From the large dataset of WVCs, spatiotemporal environmental patterns could be extracted with an ML model that predict the risk for WVCs appropriately. Even when the accident conditions varied case by case, it could be shown that the interaction of different road-. infrastructure-. land use-. weather-. topographic-. and temporal components is a complex cause-effect relationship for actual collisions. With ML and the collected dataset for the last 10 years covering the whole Bavaria, it was possible to predict WVCs spatiotemporally in a resolution and regional extent that had never been achieved before. Especially for dynamic warning systems, highly detailed data mining seems to be a critical advantage for precise warning opportunities. Thus, a detailed space–time-link between the accidents and potential influencing factors was able to be done. To not overlook the broader environment of WVCs, which may have an influence, the proportional shares of each land use type within the near, medium, and long distances, different infrastructure elements, and the topography were considered in the ML analysis.

The case study showed that ML, especially feedforward neural networks and random forests, are suitable for an appropriate risk prediction. Also, the support vector machine classifier achieved a relatively high risk prediction quality (RPQ) – not as high as with the beforementioned algorithms, but within a shorter runtime. As the WVC data does not allow any information about near misses and, hence the actual risk probability is not entirely known, different ratios of WVCs were used for ML training. The models with a training ratio of 75% WVCs and 25% non-WVCs yielded the highest quality of WVC risk classification. AutoML helped to define the best hyperparameters for a high quality of risk prediction.

The results show, on the one hand, the predictability in general, as the RPQ achieved continually values of over 85% using feedforward neural networks and random forests (Table 6). On the other hand, the transferability and applicability of the prediction model is proven, as all modes – randomly split, time-based, and geographical training splits yielded similar high results. The model was successfully tested for the Bavarian region, which shows that a transfer of the risk prediction of WVCs is generally possible.

The developed method can be applied to areas outside the case study region. As WVC data are only available for some areas, having a method that can be applied anywhere may be helpful in tack-

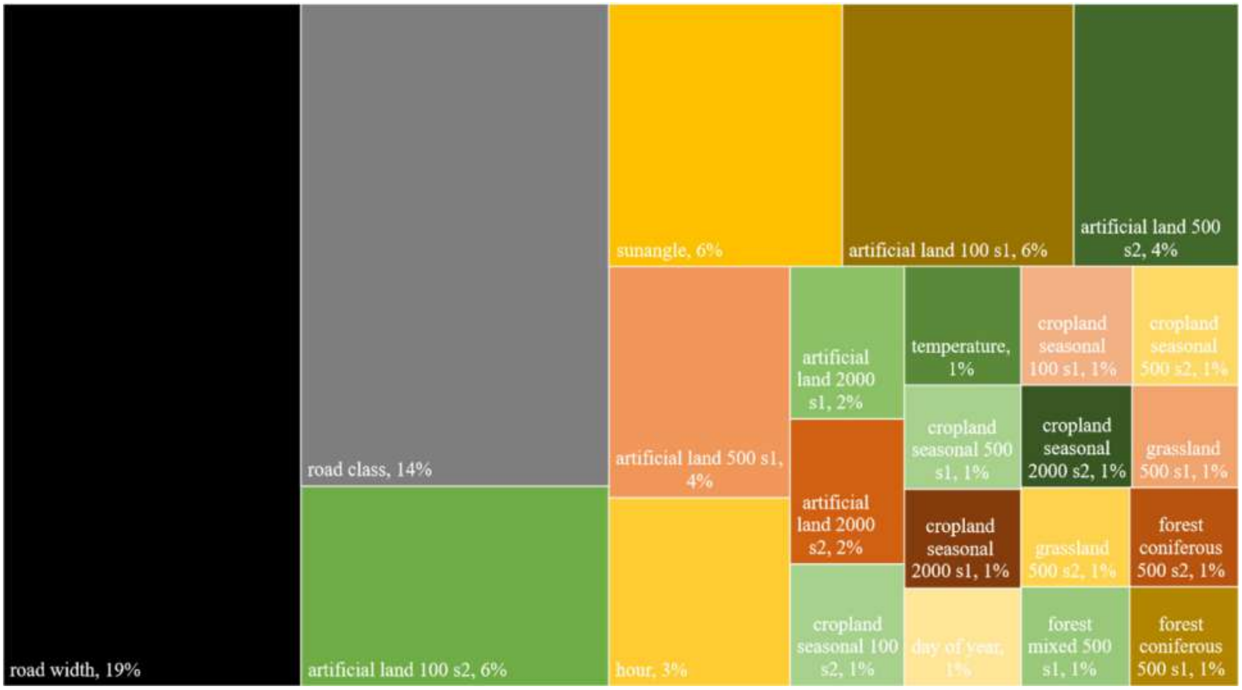

Fig. 6. Impurity results of features used in the random forest model (s1/s2: side 1/2 of the road); impurity ≥ 1% is shown.

ling the problem on a global level. However, the global transferability must be proven in future research and probably adapted to regional or species-specific characteristics. Since the case studies were only tested within Bavaria, this hypothesis is limited because it has not been proven whether the spatial transferability between different test areas in Bavaria also applies to other European or global regions. Since the prediction was made primarily with deer (around 80% of the total sample), the system may be overfit by learning pattern recognitions for collisions between deer and vehicles. Hence, the risk prediction was not applied for species that were not as frequently registered within the test region (e.g., foxes or badgers), nor for other non-native animal species. Also, the nonexistence of land use classes such as snow and ice areas in the Bavarian case study area may generate a bias for other areas during the data pre-processing stage. A limiting factor remains the data availability and quality for other parts of Europe and around the world (Pagany, 2020).

Temporal forecast tests showed that it is possible to apply this prediction analysis for dynamic warnings. As the drivers are more precisely warned when the risk is potentially high and animal movement is more likely, warning applications are implicitly also a mitigation measure in the sense of road ecology and wildlife conservation. Real-time warnings regarding the current place, time, and variable conditions avoid familiarization effect for drivers. For providing real time safety warnings, drivers' reaction time as well as vehicle kinematic reaction must be considered, meaning that the warnings needed to be provided earlier for driving along a road. For WVC warning, the application Wildwarner (wuidi, 2020) is already tested successfully with positive feedback from users (Trager, Kalová, Pagany, & Dorner, 2021).

The impurity calculations confirm that static and dynamic factors influence the WVC risk, which is also confirmed in previous studies (Pagany, 2020). Road width and road class especially determine the risk prediction. The influence of both factors was confirmed in several other studies, such as Clevenger, Barrueto, Gunson, Caryl, and Ford (2015), Pagany, Valdes, and Dorner (2020), Santos, Lourenço, Mira, and Beja (2013), or Williams, Collinson, Patterson-Abrolat, Marneweck, and Swanepoel (2019). The majority of the studies identified more open roads with high visibility through wider lanes, more gentle slopes, and fewer curves as reinforcing for an accident risk. Implicitly, higher road classes such as state and federal roads with a higher number of lanes and broader road constructions are riskier than smaller roads. In Germany, WVCs are the lowest on highways due to the high rate of fenced road sections (Pagany, 2020).

In addition, this study also identified artificial or rather urban areas as highly influential for WVC risk. However, the influence of urban areas and population density is not clearly assignable in scientific literature (Pagany, 2020). While some studies analyzed urban areas as riskier for several species (Visintin, van der Ree, & McCarthy, 2017; Wright, Coomber, Bellamy, Perkins, & Mathews, 2020), other studies identified road sections near urban areas as less risky for wildlife (Jakubas, Ryś, & Lazarus, 2018; Kantola, Tracy, Baum, Quinn, & Coulson, 2019; Zuberogoitia et al., 2014). For the WVCs in Bavaria, the WVCs seem to be generally agglomerated in rural areas with high distance to urban areas (Fig. 3).

According to this study, temporal factors - sun angle, hour of the day, and day of the year – also influence the WVC risk. Especially the time factor seems to be related to wildlife's behavior and activities due to migration (Coe et al., 2015), breeding (Grilo, Bissonette, & Santos-Reis, 2009), or rutting (Zuberogoitia et al., 2014). Several studies found daily peaks around dawn and dusk, which is a peak time of movement for animals and for humans (rush hours) (Bíl, Kubeček, Sedoník, & Andrášik, 2017; Laliberté & St-Laurent, 2020; Visintin et al., 2016). These peak times also correlates with the changes of light conditions. As examples, Pagany and Dorner (2016) found a temporal peak of accidents with wild boar in the night hours in autumn, while Biggs, Sherwood, Michalak, Hansen, and Bare (2004) identified an increased risk for deer during the early evening hours in autumn and winter. Sullivan (2011) sees the human and animals' visibility, Rodríguez-Morales, Díaz-Varela, and Marey-Pérez (2013) the wildlife's activities at twilight as causal factors. According to their study, Visintin et al. (2016) found an influence of higher temperature on the WVC risk.

The presence or absence of forested areas also influence the risk prediction in the own analysis. In previous studies, the proximity of the road network to forests and cropland is assessed as risky for WVCs. Chen and Wu (2014), for instance, found that three-quarters of their investigated WVCs had forest on both sides of the road, while several studies have proven the same increasing effect for cropland (Carvalho-Roel, Iannini-Custódio, & Marçal Júnior, 2019; Cuyckens, Mochi, Vallejos, Perovic, & Biganzoli, 2016; Found & Boyce, 2011).

The importance of the factors should not be understood as individual impact but in the combination of all factors considered into the prediction model. It is not decisive to know the increasing or decreasing impact of an individual factor on the WVC risk, as analyzed by previous WVC research using classical statistics, but to develop a predictive learning model considering the environmental factors altogether.

Targeted traffic and ecological measures are possible through the spatiotemporal data analysis and prediction, for example, to warn drivers dynamically, or to install protective measures only in times of high risk. As the case studies achieved a high quality in the risk prediction of WVCs, a basis is found to apply appropriate measures, on the one hand, for the management of road safety for humans, and, on the other hand, for biological conservation. For humans, the prediction model is able to provide dynamic warnings due to the results with high resolution in space and time. Therefore, impact studies for individual species are not relevant. However, for ecology management, species-specific risk analysis would be interesting, as different species have different activity times. Consequently, measures such as conscious feeding to change the tracks of the animals, should be adapted to reduce road crossings of a species at risky times and zones. As forest and seasonal cropland on both sides of the road led to high WVC risk, measures should be implemented at these particular road sections. Wide state and federal roads should especially be targeted. The effect of light conditions leads us to conclude that temporal warnings could be a suitable method of risk reduction.

Traffic volume and speed were two factors that were not included into the ML analysis, as the available dataset of floating car data does not provide representative information on the actual speed and volume. If the traffic speed and the volume would be available, they might help to improve the prediction performance. Furthermore, it would be interesting to investigate further ML algorithms to improve the WVC risk prediction and, thus, the road safety for human and wildlife. Even when the prediction study tried to include as many conditions as possible, based on extensive data collection and feature engineering, not all eventualities could be considered for the prediction. Even when the dynamic measures, such as the dynamic warning via a mobile application, may mitigate the collision risk as drivers drove more attentively, the remaining risk cannot be ruled out due to negligent driving behavior or unpredictable movements of, for example, startled animals.

7. Conclusions

The paper studied the WVCs of Bavaria from 2010 to 2019 including their environmental conditions, and investigated the performance of a random forest, feedforward neural networks, and a support vector machine for predicting WVC risk in a spatiotemporal manner. A large number of input data concerning WVCs was pre-processed together with a variety of environmental factors and road conditions using GIS. The pre-processing step enabled the researchers to incorporate all components, including the heterogeneous datasets and spatial and temporal dimensions of the chosen criteria into the WVC model. The risk prediction quality - the mean value of the specificity and the double-weighted sensitivity rate – was calculated as valuable indicator for assessing risk prediction in the event of road accidents with wildlife. With values of 86.7% and 86.6%, neural networks and random forests are the best choice for WVC prediction. A ratio of 75% WVCs and 25% non-WVCs was found to be the best choice for ML training. Support vector machine classifiers achieved acceptable results in shorter runtimes but at lower prediction rates.

The 36 test runs have proven that the approach can be transferred temporally and spatially. The spatial transferability was verified since the algorithm was successfully tested in Upper Bavaria with prediction values as high as with randomly split training data. The WVC forecast in a temporal manner was also proven through training with datasets from the first eight years (2010–2017), and in the prediction and testing with WVC data of 2018 and 2019. Hence, this study shows that a high quality of risk prediction for the occurrence of WVCs is possible, not only for real-time warnings but also for regions where accident data are not available. In addition to previous studies, the tests have also shown that it is important to predict WVCs spatiotemporally, as a combination of static and dynamic factors determines the predictive analysis. Other studies on WVCs have mainly focused on impact and hotspot analysis. This paper extended the WVC research by applying data mining techniques for the WVC risk prediction.

This developed prediction model for WVC risk makes a valuable contribution to take appropriate action at the actual spatiotemporal accident hotspots, even when accidents with wildlife are rarely registered. The model has the potential to be transferable to areas without WVC data availability and, thus, can help decision makers to define better traffic safety control policies, such as dynamical warnings on a global scale. Nevertheless, the limitations of this study show that more efforts should be made to improve data quality and availability of the necessary input data for a globally applicable risk-learning system for the reduction of traffic accidents.

Acknowledgments

This research was funded by the German Federal Ministry of Transport and Digital Infrastructure (BMVI) as part of the mFund project "WilDa – Dynamic Wildlife-vehicle collision warning, using heterogeneous traffic, accident and environmental data as well as big data concepts" grant number 19F2014A. I thank the Bavarian Ministry of the Interior, for Sport and Integration & Bavarian State Police for the data provision.

References

Abdelwahab, H. T., & Abdel-Aty, M. A. (2001). Development of Artificial Neural Network Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections. *Transportation Research Record, 1746*. Paper No. 01-2234.

Arevalo, J. E., Honda, W., Arce-Arias, A., & Häger, A. (2017). Spatiotemporal variation of roadkills show mass mortality events for amphibians in a highly trafficked road adjacent to a national park, Costa Rica. *Revista de Biología Tropical, 65*(4), 1261–1276. https://doi.org/10.15517/rbt.v65i4.27903.

Beshah, T. & Hill, S. (2010). Artificial Intelligence for Development, Papers from the *2010 AAAI Spring Symposium, Technical Report* SS-10-01, Stanford, California, USA, March 22-24, 2010. Retrieved March 12, 2020, from http://www.aaai.org/ocs/index.php/SSS/SSS10/paper/view/1173.

Biggs, J., Sherwood, S., Michalak, S., Hansen, L., & Bare, C. (2004). Animal-related vehicle accidents at the Los Alamos National Laboratory, New Mexico. *Southwestern Naturalist, 49*(3), 384–394. https://doi.org/10.1894/0038-4909.

Bíl, M., Andrášik, R., Svoboda, T., & Sedoník, J. (2016). The KDE+ software: A tool for effective identification and ranking of animal-vehicle collision hotspots along networks. *Landscape Ecology, 31*(2), 231–237. https://doi.org/10.1007/s10980-015-0265-6.

Bíl, M., Kubeček, J., Sedoník, J., & Andrášik, R. (2017). Srazenazver.cz: A system for evidence of animal-vehicle collisions along transportation networks. *Biological Conservation, 213*, 167–174. https://doi.org/10.1016/j.biocon.2017.07.012.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324.
Carvalho-Roel, C. F., Iannini-Custódio, A. E., & Marçal Júnior, O. (2019). Do roadkill aggregations of wild and domestic mammals overlap? *Revista de Biologia Tropical, 67*(1), 47–60. https://doi.org/10.15517/RBT.V67I1.33011.
Chang, L. Y., & Chen, W. C. (2005). Data mining of tree-based models to analyze freeway accident frequency. *Journal of Safety Research, 36*(4), 365–375. https://doi.org/10.1016/j.jsr.2005.06.013.
Chen, F., Song, M., & Ma, X. (2019). Investigation on the Injury Severity of Drivers in Rear-End Collisions Between Cars Using a Random Parameters Bivariate Ordered Probit Model. *International Journal of Environmental Research and Public Health, 16*(14). https://doi.org/10.3390/ijerph16142632.
Chen, X., & Wu, S. (2014). Examining patterns of animal-vehicle collisions in Alabama, USA. *Human-Wildlife Interactions, 8*(2), 235–244. https://doi.org/10.26077/p18k-1089.
Cherry, C. C., Dietz, S., Sauber-Schatz, E., Russell, S., Proctor, J., & Buttke, D. (2019). Characteristics of animal-related motor vehicle crashes in select National Park Service units—United States, 1990–2013. *Traffic Injury Prevention, 20*(1), 58–63. https://doi.org/10.1080/15389588.2018.1508835.
Chong, M., Abraham, A., & Paprzycki, M. (2005). Traffic Accident Analysis Using Machine Learning Paradigms. 11. *Informatica (Slovenia).* 29(1), 89–98. Retrieved September 24, 2018, from http://02.softcomputing.net/isda-mam.pdf.
Clevenger, A. P., Barrueto, M., Gunson, K. E., Caryl, F. M., & Ford, A. T. (2015). Context-dependent effects on spatial variation in deer-vehicle collisions. *Ecosphere, 6*(4). https://doi.org/10.1890/ES14-00228.1.
Coe, P. K., Nielson, R. M., Jackson, D. H., Cupples, J. B., Seidel, N. E., Johnson, B. K., et al. (2015). Identifying migration corridors of mule deer threatened by highway development. *Wildlife Society Bulletin, 39*(2), 256–267. https://doi.org/10.1002/wsb.544.
Cuyckens, G. A. E., Mochi, L. S., Vallejos, M., Perovic, P. G., & Biganzoli, F. (2016). Patterns and Composition of Road-Killed *Wildlife in Northwest Argentina. Environmental Management, 58*(5), 810–820. https://doi.org/10.1007/s00267-016-0755-6.
Destatis (2020). Bevölkerungszahlen - Kreise. Retrieved March 16, 2020, from https://www-genesis.destatis.de/.
Deublein, M., Schubert, M., Bryan, A. T., & de Soto Boria, G. (2015). A Bayesian network model to predict accidents on Swiss highways. *Infrastructure Asset Management, 2*(4), 145–158. https://doi.org/10.1680/jinam.15.00008.
Duellman, W. E. (2001). Patterns of Distribution of Amphibians. A Global Perspective. *Journal of Animal Ecology, 70*(3). https://doi.org/10.1046/j.1365-2656.2001.00515-1.x. 536–536.
DWD (German Weather Service) (2019). *Climate and environment data. COSMO-Reanalysis-6 Hourly Data*. Retrieved March 24, 2020, from ftp://opendata.dwd.de/climate_environment/REA/COSMO_REA6/hourly/2D/.
Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., & Hutter, F. (2015). Efficient and Robust Automated Machine Learning. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Hrsg.) *Advances in Neural Information Processing Systems*, 28, 2962–2970, Curran Associates, Inc. https://proceedings.neurips.cc/paper/2015/file/11d0e6287202fced83f79975ec59a3a6-Paper.pdf.
FGSV (Forschungsgesellschaft für Straßen- und Verkehrswesen) (1996). *Richtlinien für die Anlage von Straßen*, FGSV Verlag, Köln, Germany.
Found, R., & Boyce, M. S. (2011). Predicting deer-vehicle collisions in an urban area. *Journal of Environmental Management, 92*(10), 2486–2493. https://doi.org/10.1016/j.jenvman.2011.05.010.
Garriga, N., Franch, M., Santos, X., Montori, A., & Llorente, G. A. (2017). Seasonal variation in vertebrate traffic casualties and its implications for mitigation measures. *Landscape and Urban Planning, 157*, 36–44. https://doi.org/10.1016/j.landurbplan.2016.05.029.
GDV (Gesamtverband der Deutschen Versicherer) (2020). *Wildunfälle erreichen Rekordhoch*. Retrieved 12 November 2020, URL: http://www.gdv.de/de/medien/aktuell/wildunfaelle-erreichen-rekordhoch-63692.
Gkritza, K., Baird, M., & Hans, Z. N. (2010). Deer-vehicle collisions, deer density, and land use in Iowa's urban deer herd management zones. *Accident Analysis and Prevention, 42*(6), 1916–1925. https://doi.org/10.1016/j.aap.2010.05.013.
Grilo, C., Bissonette, J. A., & Santos-Reis, T. M. (2009). Spatial-temporal patterns in Mediterranean carnivore road casualties: Consequences for mitigation. *Biological Conservation., 142*(2), 301–313. https://doi.org/10.1016/j.biocon.2008.10.026.
Grilo, C., Ferreira, F. Z., & Revilla, E. (2015). No evidence of a threshold in traffic volume affecting road-kill mortality at a large spatio-temporal scale. *Environmental Impact Assessment Review, 55*, 54–58. https://doi.org/10.1016/j.eiar.2015.07.003.
Gunson, K. E., Mountrakis, G., & Quackenbush, L. J. (2011). Spatial wildlife-vehicle collision models: A review of current work and its application to transportation mitigation projects. *Journal of Environmental Management, 92*(4), 1074–1082. https://doi.org/10.1016/j.jenvman.2010.11.027.
Heigl, F., Stretz, C. R., Steiner, W., Suppan, F., Bauer, T., Laaha, G., et al. (2016). Comparing road-kill datasets from hunters and citizen scientists in a landscape context. *Remote Sensing, 8*(10). https://doi.org/10.3390/rs8100832.
Holmgren, W. F., Hansen, C. W., & Mikofski, M. A. (2018). pvlib python: A python package for modeling solar energy systems. *Journal of Open Source Software, 3* (29), 884. https://doi.org/10.21105/joss.00884.
Jakubas, D., Ryś, M., & Lazarus, M. (2018). Factors affecting wildlife-vehicle collision on the expressway in a suburban area in northern Poland. *North-Western Journal of Zoology, 14*(1), 107–116.
Jin, H., Song, Q., & Hu, X. (2019). Auto-Keras: An Efficient Neural Architecture Search System. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA* (pp. 1946–1956). https://doi.org/10.1145/3292500.3330648.
Kämmerle, J.-L., Brieger, F., Kröschel, M., Hagen, R., Storch, I., & Suchant, R. (2017). Temporal patterns in road crossing behaviour in roe deer (Capreolus capreolus) at sites with wildlife warning reflectors. *PLoS ONE, 12*(9). https://doi.org/10.1371/journal.pone.0184761.
Kantola, T., Tracy, J. L., Baum, K. A., Quinn, M. A., & Coulson, R. N. (2019). Spatial risk assessment of eastern monarch butterfly road mortality during autumn migration within the southern corridor. *Biological Conservation., 231*, 150–160. https://doi.org/10.1016/j.biocon.2019.01.008.
Keken, Z., Sedoník, J., Kušta, T., Andrášik, R., & Bíl, M. (2019). Roadside vegetation influences clustering of ungulate vehicle collisions. *Transportation Research Part D: Transport and Environment, 73*, 381–390. https://doi.org/10.1016/j.trd.2019.07.013.
Komol, M. M. R., Hasan, M. M., Elhenawy, M., Yasmin, S., Masoud, M., & Rakotonirainy, A. (2021). Crash severity analysis of vulnerable road users using machine learning. *PLoS ONE, 16*(8). https://doi.org/10.1371/journal.pone.0255828.
Laliberté, J., & St-Laurent, M.-H. (2020). In the wrong place at the wrong time: Moose and deer movement patterns influence wildlife-vehicle collision risk. *Accident Analysis and Prevention, 135*. https://doi.org/10.1016/j.aap.2019.105365.
Liu, Y., Nieuwenhuis, M., & McCullagh, A. (2018). The effect of roadside land-use on the occurrence of deer vehicle collisions. *Irish Forestry*, 75(1&2), 8-25. Retrieved October 21, 2019, from https://journal.societyofirishforesters.ie/index.php/forestry/article/view/10908.
Lutterschmidt, W. I., Weidler, J. M., & Schalk, C. M. (2019). Hot moments and hot spots in the bayou: Spatiotemporal patterns of road occurrence in a snake assemblage in Louisiana, USA. *Herpetological Conservation and Biology*, 14(2), 533–545. Retrieved September 24, 2019, from http://www.herpconbio.org/Volume_14/Issue_2/Lutterschmidt_etal_2019.pdf.
Malo, J. E., Suárez, F., & Díez, A. (2004). Can we mitigate animal-vehicle accidents using predictive models? *Journal of Applied Ecology, 41*(4), 701–710. https://doi.org/10.1111/j.0021-8901.2004.00929.x.
McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics, 5*, 115–133. ISSN 0007-4985.
Morelle, K., Lehaire, F., & Lejeune, P. (2013). Spatio-temporal patterns of wildlife-vehicle collisions in a region with a high-density road network. *Nature Conservation, 5*, 53–73. https://doi.org/10.3897/natureconservation.5.4634.
OSGeo (2019). *GDAL/OGR*. Retrieved October 21, 2019, from https://www.osgeo.org/projects/gdal/.
Pagany, R. (2020). Wildlife-vehicle collisions - Influencing factors, data collection and research methods. *Biological Conservation, 251*. https://doi.org/10.1016/j.biocon.2020.108758.
Pagany, R., & Dorner, W. (2016). Spatiotemporal Analysis for Wildlife-Vehicle Collision Based on Accident Statistics of the County Straubing-Bogen in Lower Bavaria. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial, Information Sciences, XLI-B8*, 739–745. https://doi.org/10.5194/isprsarchives-XLI-B8-739-2016.
Pagany, R., Valdes, J., & Dorner, W. (2020). Risk Prediction of Wildlife-vehicle Collisions Comparing Machine Learning Methods and Data Use. In *2020 IEEE 10th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 2020* (pp. 436–440). https://doi.org/10.1109/ACIT49673.2020.9208946.
Pflugmacher, D., Rabe, A., Peters, M., & Hostert, P. (2018). Pan-European land cover map of 2015 based on Landsat and LUCAS data. *PANGAEA*. https://doi.org/10.1594/PANGAEA.896282.
Rodríguez-Morales, B., Díaz-Varela, E. R., & Marey-Pérez, M. F. (2013). Spatiotemporal analysis of vehicle collisions involving wild boar and roe deer in NW Spain. *Accident Analysis and Prevention., 60*, 121–133. https://doi.org/10.1016/j.aap.2013.07.032.
Santos, S. M., Lourenço, R., Mira, A., & Beja, P. (2013). Relative effects of road risk, habitat suitability, and connectivity on wildlife Roadkills: The case of tawny owls (Strix aluco). *PLoS ONE, 8*(11), e79967.
Santos, R. A. L., Mota-Ferreira, M., Aguiar, L. M. S., & Ascensão, F. (2018). Predicting wildlife road-crossing probability from roadkill data using occupancy-detection models. *Science of The Total Environment, 642*, 629–637. https://doi.org/10.1016/j.scitotenv.2018.06.107.
Seo, C., Thorne, J. H., Choi, T., Kwon, H., & Park, C.-H. (2015). Disentangling roadkill: The influence of landscape and season on cumulative vertebrate mortality in South Korea. *Landscape Ecology Engineering, 11*(1), 87–99. https://doi.org/10.1007/s11355-013-0239-2.
Shrivastava, A., Verma, J. P. V., Jain, S., et al. (2021). A deep learning based approach for trajectory estimation using geographically clustered data. *SN Appl. Sci., 3*, 597. https://doi.org/10.1007/s42452-021-04556-x.
Snow, N. P., Zhang, Z., Finley, A. O., Rudolph, B. A., Porter, W. F., Williams, D. M., et al. (2018). Regional-based mitigation to reduce wildlife–vehicle collisions. *Journal of Wildlife Management, 82*(4), 756–765. https://doi.org/10.1002/jwmg.21420.
Steiner, W., Leisch, F., & Hackländer, K. (2014). A review on the temporal pattern of deer-vehicle accidents: Impact of seasonal, diurnal and lunar effects in cervids. *Accident Analysis and Prevention, 66*, 168–181. https://doi.org/10.1016/j.aap.2014.01.020.
StMI (Bavarian State Ministry of the Interior, for Sport and Integration) (2017). Road d*ata from the BAYSIS database* (Bavarian road infrastructure system), Munich, Germany. (not published).

StMI (Bavarian State Ministry of the Interior, for Sport and Integration) (2017). Road data from the BAYSIS database (Bavarian road infrastructure system), Munich, Germany. (not published).

Sullivan, J. M. (2011). Trends and characteristics of animal-vehicle collisions in the United States. *Journal of Safety Research, 42*(1), 9–16. https://doi.org/10.1016/j.jsr.2010.11.002.

Surveying Authorities of the Federal States of Germany (2015). *Official Topographical-Cartographic Information System data – Digital Landscape Models* (ATKIS-DLM Data). (not published).

Tanner, A. L., Leroux, S. J., & Saunders, P. W. (2017). Road characteristics best predict the probability of vehicle collisions with a non-native ungulate. *Ecoscience, 24* (1–2), 1–11. https://doi.org/10.1080/11956860.2017.1292100.

Toblerity (2020). *Shapely*. Retrieved February 19, 2020, from https://github.com/Toblerity/Shapely.

Trager, J., Kalová, L., Pagany, R., & Dorner, W. (2021). Warning Apps for Road Safety: A Technological and Economical Perspective for Autonomous Driving – The Warning Task in the Transition from Human Driver to Automated Driving. *International Journal of Human-Computer Interaction, 37*(4), 363–377. https://doi.org/10.1080/10447318.2020.1860545.

Valero, E., Picos, J., & Álvarez, X. (2015). Road and traffic factors correlated to wildlife-vehicle collisions in Galicia (Spain). *Wildlife Research, 42*(1), 25–34. https://doi.org/10.1071/WR14060.

van der Ree, R., Jaeger, J. A. G., van der Grift, E. A., & Clevenger, A. P. (2011). Effects of roads and traffic on wildlife populations and landscape function: Road ecology is moving toward larger scales. *Ecology and Society, 16*(1). https://doi.org/10.5751/ES-03982-160148.

Vapnik, V. (1963). Pattern recognition using generalized portrait method. *Automation remote control, 24*, 774–780.

Visintin, C., van der Ree, R., & McCarthy, M. A. (2016). A simple framework for a complex problem? Predicting wildlife–vehicle collisions. *Ecology and Evolution, 6*(17), 6409–6421. https://doi.org/10.1002/ece3.2306.

Visintin, C., van der Ree, R., & McCarthy, M. A. (2017). Consistent patterns of vehicle collision risk for six mammal species. *J. Environ. Manag., 201*, 397–406. https://doi.org/10.1016/j.jenvman.2017.05.071.

Williams, S. T., Collinson, W., Patterson-Abrolat, C., Marneweck, D. G., & Swanepoel, L. H. (2019). Using road patrol data to identify factors associated with carnivore roadkill counts. *PeerJ, 7*, e6650.

Wright, P. G. R., Coomber, F. G., Bellamy, C. C., Perkins, S. E., & Mathews, F. (2020). Predicting hedgehog mortality risks on British roads using habitat suitability modelling. *PeerJ, 1*, 2020. https://doi.org/10.7717/peerj.8154.

wuidi (2020). *Gemeinsam Wildunfälle vermeiden. Wuidi - Die Nr. 1. Wildwarner*. Retrieved March 11, 2020, from https://wuidi.com.

Yang, W. T., Chen, H. C., & Brown, D. B. (1999). Detecting Safer Driving Patterns By A Neural Network Approach. In *1999 Proceedings of Smart Engineering System Design Neural Network, Evolutionary Programming, Complex Systems and Data Mining, St. Louis, Missouri, USA* (pp. 839–844).

Zuberogoitia, I., del Real, J., Torres, J. J., Rodríguez, L., Alonso, M., & Zabala, J. (2014). Ungulate vehicle collisions in a peri-urban environment: Consequences of transportation infrastructures planned assuming the absence of ungulates. *PLoS ONE, 9*(9), e107713.

Raphaela Pagany did this research as a senior scientist at the Technische Hochschule Deggendorf. Her research is focused on Geoinformatics, spatio-temporal analysis and risk prediction. In the project WilDa, she developed algorithms for the Wildwarner service. In addition, she did her PhD study at the Paris Lodron University in Salzburg. Currently, she is working in transport planning.

Journal of Safety Research 83 (2022) 96–104

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Association of falls and fear of falling with objectively-measured driving habits among older drivers: LongROAD study

Carolyn G. DiGuiseppi [a,*], Hailey A. Hyde [a], Marian E. Betz [b], Kenneth A. Scott [a], David W. Eby [c,d], Linda L. Hill [e], Vanya C. Jones [f], Thelma J. Mielenz [g], Lisa J. Molnar [c,d], David Strogatz [h], Guohua Li [g,i], on behalf of the AAA LongROAD Research Team

[a] Department of Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, USA
[b] Department of Emergency Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA
[c] University of Michigan Transportation Research Institute, University of Michigan, Ann Arbor, MI, USA
[d] Center for Advancing Transportation Leadership and Safety (ATLAS Center), University of Michigan, Ann Arbor, MI, USA
[e] School of Public Health, University of California San Diego, La Jolla, CA, USA
[f] Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[g] Department of Epidemiology, Mailman School of Public Health, Columbia's Injury Control Research Center, Columbia University, New York, NY, USA
[h] Bassett Research Institute, Bassett Healthcare Network, Cooperstown, NY, USA
[i] Department of Anesthesiology, Columbia University College of Physicians and Surgeons, New York, NY, USA

ARTICLE INFO

Article history:
Received 10 August 2021
Received in revised form 28 February 2022
Accepted 9 August 2022
Available online 25 August 2022

Keywords:
Falls
Fear of falling
Traffic safety
Mobility
Aging
Automobile driving

ABSTRACT

Objective: Falls in older adults are associated with increased motor vehicle crash risk, possibly mediated by driving behavior. We examined the relationship of falls and fear of falling (FOF) with subsequent objectively measured driving habits. *Methods:* This multi-site, prospective cohort study enrolled 2990 active drivers aged 65–79 (53% female). At enrollment, we assessed falls in the past year and FOF (Short Falls Efficacy Scale-International). Driving outcomes included exposure, avoidance of difficult conditions, and unsafe driving during one-year follow-up, using in-vehicle Global Positioning System devices. *Results:* Past-year falls were associated with more hard braking events (HBE). High FOF was associated with driving fewer days, miles, and trips, driving nearer home and more HBE. Differences were attenuated and not significant after accounting for health, function, medications and sociodemographics. *Discussion:* Differences in objectively measured driving habits according to past-year fall history and FOF were largely accounted for by differences in health and medications. Rather than directly affecting driving, falls and FOF may serve as markers for crash risk and reduced community mobility due to age-related changes and poor health.

1. Introduction

Over the past two decades, the number of licensed older drivers and the miles they drive have steadily risen (Pomidor, 2019). In many societies, driving is vital to maintaining mobility -- the ability to move within one's environment – which is essential for healthy aging (Webber et al., 2010). However, fatal motor vehicle crashes per vehicle mile traveled begin to increase at ages 70–74 and continue to increase with age (Cox & Cicchino, 2021), making the prevention of crashes among older drivers an important public health concern.

The risk of crashes for older drivers relates in part to age-related functional changes, medical conditions and medications that affect driving ability (Pomidor, 2019). Falls are one such age-related condition, common among older adults (Bergen et al., 2016), that have been associated with an increased risk of motor vehicle crashes (MVC). In a systematic review of 15 studies of varying designs, sizes and geographic locations, Scott et al. (2017) found a fall history to be associated with a 40% increased risk of subsequent MVC, as well as a higher risk of MVC-related hospitalization and death. The mechanisms underlying the relationship between falls and crashes are unclear but might result from changes in driving behaviors and patterns after a fall. Fall injuries may cause functional impairments (e.g., fracture that reduces range of motion) that directly affect driving ability or behaviors. Falls, regardless of injury, might also lead to changes in driving behaviors through

* Corresponding author at: Department of Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, 13001 East 17th Place, Campus Box B119, Aurora, CO 80045, USA.

E-mail address: Carolyn.DiGuiseppi@cuanschutz.edu (C.G. DiGuiseppi).

https://doi.org/10.1016/j.jsr.2022.08.007
0022-4375/

their psychological impact. Older adults who fall are significantly more likely to develop fear of falling (FOF) (Friedman et al., 2002), a concern that can lead to limitations in physical and social activity (Scheffer et al., 2008; Tinetti & Powell, 1993), self-care and household activities (Liu et al., 2021), and could potentially result in changes in driving habits as well. Finally, falls may be markers for age-related health and functional declines that influence both falls and driving abilities and behaviors, such as cognitive, vision and hearing impairment, gait and balance problems, and use of psychotropic medications (Deandrea et al., 2010; Karthaus & Falkenstein, 2016; Pomidor, 2019).

While changes in driving behavior associated with falls or FOF might explain the observed association between falls and MVC, the relationship between falls, FOF and driving habits has not been well-delineated. Scott et al. (2017) did not find consistent evidence of an association between falls and subsequent driving habits from the seven studies identified that had examined and reported driving habits. These studies were limited by their cross-sectional design (Forrest et al., 1997; Lyman et al., 2001; Vance et al., 2006), small samples (Crizzle et al., 2013; Lyman et al., 2001; Marie Dit Asse et al., 2014) and use of self-reported driving behavior outcomes, which may be subject to bias (Dugan & Lee, 2013; Forrest et al., 1997; Lyman et al., 2001; Marie Dit Asse et al., 2014; MacLeod et al., 2014; Vance et al., 2006). Only one study examined the relationship between falls and objectively measured driving habits, assessing 27 older adults with Parkinson's disease followed for two weeks (Crizzle et al., 2013). This study found that participants who had fallen in the past year exhibited more hard braking and drove more slowly than those without a fall. In addition to the small sample restricted to persons affected by Parkinson's, this study did not account for demographic or health-related differences between individuals who had and had not fallen. The literature examining FOF in relation to driving habits is similarly limited. One cohort study found that older women expressing high FOF were more likely to report driving cessation or reduction in the subsequent six years than women with low FOF, although this pattern was not found for men (Marie Dit Asse et al., 2014). A study with older residents of retirement communities (71% women) found that FOF was negatively associated with the number of objectively measured vehicle trips taken per day over six days' follow-up, but not with daily distance or minutes for vehicle trips (Takemoto et al., 2015). Studies in larger samples examining objectively measured driving habits over a longer period of time could help to establish whether falls or FOF (or both) are associated with changes in driving habits and driving behaviors among older adults.

The current study aimed to examine associations of self-reported history of having fallen, and of fear of falling, with objectively measured driving patterns during one year of follow-up in a large, geographically diverse cohort of older drivers.

2. Materials and methods

2.1. Study design

The AAA LongROAD study is a prospective cohort study designed to examine medical, behavioral, environmental, and other factors associated with safe driving in older adults. LongROAD enrollment occurred at five US sites (Ann Arbor, MI; Baltimore, MD; Cooperstown, NY; Denver, CO; and San Diego, CA). The study design and population have been described in detail previously; the study collects self-reported and objectively measured data on health, functioning, and driving behaviors (Li et al., 2017). A sample of 3000 drivers with average follow-up of 2.5 years was planned to provide study power >80% to detect an age-adjusted risk ratio of 3.0 for crash involvement associated with mild cognitive impairment (see Li et al., 2017). The study was approved by the institutional review board at each site, including Bassett Research Institute, Columbia University, Johns Hopkins University, University of California San Diego, University of Colorado Anschutz Medical Campus, and University of Michigan. All enrolled participants provided written informed consent for participation and received $100 at the baseline visit. The STROBE cohort reporting guidelines were followed for this report (von Elm et al., 2007).

2.2. Sample

Study participants were aged 65–79 years at enrollment, had a valid driver's license, drove on average at least once a week, drove one car (1996 or newer with an accessible OBDII port) at least 80% of the time, spoke English, had no significant cognitive impairment (e.g., Alzheimer's disease) based on medical record review and a Six-Item Screener score ≥ 4 (sensitivity 67.5% and specificity 96.1% for clinically diagnosed dementia) (Callahan et al., 2002), and resided in the catchment area at least 10 months a year with no plans to move away within 5 years (Li et al., 2017). Using electronic medical records from healthcare systems affiliated with study sites, study staff identified potentially eligible patients, sent initial recruitment letters followed by telephone calls for eligibility screening, and scheduled eligible, interested participants for a baseline study visit for enrollment and data collection. Of 40,806 individuals sent recruitment letters, 19.0% could not be contacted by phone, 29.7% declined eligibility screening, 19.0% were ineligible (most often due to no/infrequent driving or ineligible residence), 25.0% were eligible but declined, and 7.3% (range 5.1–18.3% across study sites) (n = 2990 participants) enrolled (Li et al., 2017). Recruitment and enrollment were completed between July 2015 and March 2017.

2.3. Driving outcomes

We objectively measured driving outcomes with a device installed in each participant's vehicle that collected data when the vehicle was turned on. The device could determine if the participant was the driver using a Bluetooth receiver to detect participant codes and signal strengths transmitted by Bluetooth beacons carried by the participant. This study used data recorded during the first 12 months after the baseline fall history. Driving measures are defined in Table 1. Driving habit measures were based on previous work (Molnar et al., 2013a), conceptualized based on three components of the Driving Habits Questionnaire (DHQ) (Owsley et al., 1999): driving space, driving exposure, and driving avoidance. Two driving measures - rapid deceleration ("hard braking") events and speeding events (Table 1) - served as proxies for unsafe driving (Chevalier et al., 2017; Eby et al., 2019; Williams et al., 2006). We excluded participants if they were missing all driving measurements (n = 15) or drove fewer than 14 days or 100 miles during the 12-month period (n = 18). We derived means and standard deviations for each driving habit measure from the full 12 months of data. Two variables with skewed distributions (proportion of trips at night and number of rapid deceleration events) were log-transformed. Speeding events, which were uncommon (median = 1 event/1000 miles driven), were categorized as any versus none during the 12-month period.

2.4. Exposures

At enrollment, research staff administered questionnaires about demographics, health and healthcare utilization. We used the following variables as primary exposures: (1) "In the last 12 months have you fallen down?" (Yes/No), and (2) the 7-item Short Falls

Table 1
Means, standard deviations, definitions, and category for each driving habit measure.

Objective Driving Measure	Mean (SD)	Definition for the Monthly Variable (Trip is defined as ignition on to ignition off)	Category
Average Percent of Trips Within 15 Miles of Home	64.1 (22.4)	Percent of trips traveled in month within 15 miles of home.	Driving Space
Average Number of Miles	736.0 (433.6)	Total number of miles driven in month.	Driving Exposure
Average Number of Days Driving	21.0 (5.5)	Total number of days in month with at least one trip.	Driving Exposure
Average Number of Trips	110.8 (53.3)	Total number of trips in a month.	Driving Exposure
Average Percent of Trips at Night	1.9 (0.7)	Percent of trips in month during which at least 80% of trip was during nighttime, with nighttime defined as end of evening civil twilight to beginning of morning civil twilight or a solar angle greater than 96 degrees.	Driving Avoidance
Average Percent of Trips on High Speed Roads	12.8 (11.0)	Percent of trips in month during which at least 20% of distance travelled was at a speed of 60 MPH or greater.	Driving Avoidance
Average Percent Trips in AM Peak	7.1 (4.9)	Percent of all trips taken in month during 7:00–9:00 AM on weekdays.	Driving Avoidance
Average Percent Trips in PM Peak	9.6 (4.4)	Percent of all trips taken in month during 4:00–6:00 PM on weekdays.	Driving Avoidance
Right to Left Turn Ratio	0.9 (0.1)	Ratio of all right-hand to left-hand turning events identified for driver in month.	Driving Avoidance
Average Speeding Events	7.8 (17.2)	Number of speeding events (speed > 80 MPH sustained for at least 8 seconds) per 1000 miles driven.	Unsafe Driving
Average Rapid Deceleration Events	5.4 (6.4)	Number of events with deceleration greater than or equal to 0.4 g (hard braking, near crash, crash) per 1000 miles driven.	Unsafe Driving

Efficacy Scale-International (Short FES-I), which assesses concerns about falling (categorized as Low [<11]/High [11–28]) (Delbaere et al., 2010). For those who answered "Yes" to having fallen, we also asked, "In the last 12 months, have you fallen down more than one time? (Yes/No)." We excluded participants from all analyses if they did not answer the question about falling in the last 12 months (n = 16) and from analyses of FOF if they were missing Short FES-I scores (n = 3).

2.5. Covariates

Demographic characteristics collected at baseline included age group, gender, race, marital status, educational attainment, household income, work for pay in the past month and urbanicity of the participant's residence. Self-reported health characteristics included: vision with correction (rated as poor to excellent); Patient-Reported Outcomes Measurement Information System short form (PROMIS SF) measures of physical function, cognitive health ("applied cognition – general concerns"), depression and anger (HealthMeasures, 2020) (as detailed in Li et al. (2017)); and self-reported driving reduction due to a health condition in the past 12 months (Yes/No). Moderate and severe categories of physical (dys)function were combined for analysis due to small numbers. None of the participants had more than slight concerns about their cognition; these scores were categorized into tertiles for analysis. We assessed self-reported health-care utilization in the past 12 months (i.e., emergency department visits [None, 1, 2 or More] or hospitalizations [None/Any]). Use of potentially impairing substances included assessment of alcohol and medication use. Alcohol consumption frequency in the past three months was categorized into any versus none, as few reported more than light-to-moderate drinking. Current medication use (prescribed and over-the-counter) was collected and categorized according to the American Hospital Formulary System (AHFS) classification, as described in Hill et al (2020). We examined medications that act on the central nervous and cardiovascular systems, which have been associated with fall risk (Hartikainen et al., 2007; Park et al., 2015). Individuals who reported taking one or more psychotherapeutic, anxiolytic, sedative, hypnotic or anticonvulsant agents were categorized as taking a central nervous system (CNS) medication. Individuals who reported taking one or more antiarrhythmic, cardiotonic, or diuretic agents were categorized as taking a cardiovascular medication.

2.6. Statistical analysis

Chi-Square tests (or Fisher's exact tests) were used to assess each covariate's association with each exposure of interest. Unadjusted associations between each exposure and each driving habit of interest were examined using separate linear or logistic regression models, as appropriate. For each driving outcome, we accounted for potential differences in age group, gender, race and marital status between participants with and without a past-year history of falls (Yes/No) or concern about falling (High/Low) in all multivariable models ("base models"). Participants missing data on any of these four sociodemographic variables were excluded from adjusted analyses (n = 75). Self-reported measures of health, health care utilization, medication or alcohol use, and additional sociodemographic factors were assessed as potential covariates if they were associated with both fall history or concern about falling and the driving outcome measure at $p < 0.20$. As a sensitivity analysis, we also examined exposure to past-year falls categorized as none, one and more than one. Model assumptions and fit were assessed using residuals, probability plots, and Akaike information criteria (AIC) as appropriate. All results are reported as beta estimates or odds ratios, as indicated, with 95% confidence intervals (CI), using an alpha level of 0.05 for testing statistical significance. All analyses were conducted using SAS University Edition software (version 9.04.01, SAS Institute, Inc., Cary, North Carolina).

3. Results

Of the 2990 participants enrolled in LongROAD, 2941 (98.4%) had complete data on both self-reported falls and objective driving measures; 2938 of these (99.9%) also had FOF data. Adjusted base models examining self-reported falls and FOF included 2866 and 2863 participants, respectively.

A substantial proportion (28.2%) reported having fallen at least once in the 12 months prior to enrollment, and 5.1% had fallen more than once, while 18.6% expressed a high FOF (Table 2). Table 2 shows characteristics of the sample by fall history. At baseline,

Table 2
Characteristics of Older Drivers with and without a History of Falls in the Past 12 Months, LongROAD Cohort of Older Drivers.

Characteristics	Past-Year Fall N = 828 (28.2%)	No Past-Year Fall N = 2113 (71.8%)	p-value
Short FES-I			
High Concern (≥11)	238 (28.7)	309 (14.6)	<0.001
Low Concern (<11)	590 (71.3)	1803 (85.4)	
Age Group			
65–69	329 (39.7)	893 (42.3)	0.457
70–74	296 (35.7)	722 (34.2)	
75–79	203 (24.5)	498 (23.6)	
Gender			
Female	493 (59.5)	1064 (50.4)	<0.001
Male	335 (40.5)	1049 (49.6)	
Race			
White	745 (91.3)	1830 (88.2)	0.015
Non-White	71 (8.7)	246 (11.8)	
Marital Status			
Married or Living with Partner	527 (64.1)	1420 (67.9)	0.052
Separated, Divorced, Widowed, Never Married	295 (35.9)	672 (32.1)	
Highest Level of Education			
Less than High School	18 (2.2)	43 (2.0)	0.968
High School, Vocational, Some College, Associate	282 (34.1)	702 (33.3)	
Bachelor Degree	190 (23.0)	494 (23.5)	
Master, Professional, Doctoral Degree	336 (40.7)	867 (41.2)	
Total Household Income			
$100,000 or more	252 (31.3)	693 (34.1)	0.216
$80,000–$99,999	118 (14.7)	304 (14.9)	
$50,000–$79,999	198 (24.6)	514 (25.3)	
Less than $50,000	237 (29.4)	523 (25.7)	
Worked for Pay Last Month			
Yes, Full-Time	66 (8.0)	220 (10.5)	0.084
Yes, Part Time	178 (21.6)	410 (19.5)	
No	579 (70.4)	1470 (70.0)	
Urbanicity of Residence			
Metropolitan Core	595 (71.9)	1545 (73.1)	0.788
Metropolitan Area/Non-Core	120 (14.5)	293 (13.9)	
Micropolitan/Small Town/Rural	113 (13.6)	275 (13.0)	
Eyesight with Correction			
Excellent	186 (22.5)	556 (26.3)	0.031
Very Good	342 (41.3)	888 (42.1)	
Good	264 (31.9)	603 (28.6)	
Fair + Poor	36 (4.4)	64 (3.03)	
Physical Function Limitations			
Moderate to Severe (T-Score ≤ 39.9)	114 (13.9)	124 (5.9)	<0.001
Mild (T- Score 40.0–55.0)	369 (45.0)	741 (35.4)	
None to Slight (T score > 55.0)	337 (41.1)	1230 (58.7)	
Applied Cognition-General Concerns (T-Score Tertiles)			
T-Score > 32.4	320 (38.7)	615 (29.2)	<0.001
T-Score 26.3–32.3	214 (25.9)	548 (26.0)	
T-Score < 26.2	292 (35.4)	941 (44.7)	
Depression (T-scores)			
Moderate to Severe (T-Score ≥ 60.0)	17 (2.1)	23 (1.1)	0.022
Mild (T- Score 55.0–59.9)	50 (6.1)	94 (4.5)	
None to Slight (T score < 55.0)	759 (91.9)	1994 (94.5)	
Anger (T-Scores)			
Moderate to Severe (T-Score ≥ 60.0)	13 (1.6)	18 (0.9)	0.012
Mild (T- Score 55.0–59.9)	39 (4.7)	62 (2.9)	
None to Slight (T score < 55.0)	772 (93.7)	2030 (96.2)	
Decreased Driving Due to Health in Past 12 Months	146 (17.7)	188 (8.9)	<0.001
Emergency Department Visits Past 12 Months			
2+	75 (9.1)	110 (5.2)	<0.001
1	182 (22.0)	310 (14.7)	
At Least One Hospital Stay Past 12 Months	177 (21.4)	275 (13.1)	<0.001
Any Alcohol Consumption Past 3 Months	612 (73.91)	1528 (72.3)	0.391
Current Use Central Nervous System Medications*	348 (42.0)	571 (27.0)	<0.001
Current Use Cardiovascular Medications**	211 (25.5)	532 (25.2)	0.864

Missing data: race (n = 49, 1.7%), marital status (n = 27, 0.9%), total household income (n = 102, 3.5%), worked for pay (n = 18, 0.6%), hospital stay (n = 12, 0.4%), physical function limitations (n = 26, 0.9%); variables with 1–11 missing values (<0.4%) included highest level of education, emergency department visits, hospital stays, alcohol use, decreased driving due to health in past 12 months, depression, anger, applied cognition-general concerns, eyesight with correction and short FES-I. Remaining variables had no missing data.

* Anticonvulsants, Psychotherapeutic Agents, Anxiolytics, Sedatives, Hypnotics.

** Antiarrhythmics, Cardiotonic Agents, Diuretics.

compared to participants who had not fallen in the past year, participants who had fallen were twice as likely to report high FOF (28.7% vs 14.6%, p < 0.001). Those with a past-year fall history were significantly more likely to be female (59.5% vs 50.4%, p < 0.001) and of white race (91.3%% vs 88.2%, p = 0.015). They also perceived their visual acuity, physical and cognitive function to be poorer, described more current symptoms of depression and anger, and were more likely to have reduced their driving due to a health condition and to have had greater healthcare utilization in the past 12 months, and to use central nervous system medications currently (Table 2).

Drivers who did and did not report any past-year fall at baseline had similar driving exposure, driving space and driving avoidance (as defined in Table 1) during the first 12 months after baseline, after accounting for gender, age group, race, and marital status (Table 3). Participants who fell had a median of four rapid deceleration events per 1000 miles driven versus three per 1000 miles among participants who had not fallen. There were significantly more hard-braking events per 1000 miles driven in participants who had fallen in the past year versus participants who had not. Drivers who did and did not report any past-year fall at baseline did not differ significantly in their odds of having had at least one speeding event per 1000 miles driven. The difference in rapid deceleration events was attenuated and no longer statistically significant when group differences in general concerns about cognition and current use of CNS medications were taken into account. Neither cognitive concerns nor central nervous system medication use substantively affected the magnitude, direction or statistical significance of the other results (Table 3). No other differences in participant characteristics between groups were retained in adjusted models.

Sensitivity analysis revealed similar results for persons with more than one fall and persons with only one fall in the past year, compared to participants with no falls, except that persons with more than one fall made a significantly smaller proportion of driving trips during morning rush hour (adjusted beta = −0.79 [95%CI: −1.35, −0.22]).

Nearly all sociodemographic characteristics differed significantly between participants with high versus low concern about falling. Participants with high concern were more likely to be female (63.3% vs 50.6%, $p < 0.001$), ages 75–79 years (30.7% vs 22.2%, $p < 0.001$), and non-white race (14.4% vs 10.0%, $p = 0.002$), and less likely to be married or living with a partner (55.6% vs 68.6%, $p < 0.001$). Further, they were less likely to have a Bachelor's or higher degree (53.6% vs 66.6%, $p < 0.001$) or to work full-time in the last month (5.7% vs 10.6%, $p < 0.001$), and more likely to have a total household income less than $50,000 (38.0% vs 23.0%, $p < 0.001$). Differences between participants with high versus low FOF in perceived vision with correction, concerns about cognition, emergency department (ED) visits, hospitalizations, use of CNS medications and decreased driving due to a health problem were similar to differences observed between participants who had and had not fallen. In addition, participants who expressed high FOF reported poorer physical function (median t-score 43.4 vs 56.9, $p < 0.001$) and more often used cardiovascular medications (32.9% vs 23.9%, $p < 0.001$) compared to participants with low FOF.

Compared to participants with low FOF, participants with high FOF drove significantly fewer miles, days, and trips per month (all $p < 0.01$) and made more trips within 15 miles of home ($p < 0.001$) during the 12 months after baseline (Table 4). Driving avoidance was similar between participants with high versus low FOF, except that the former were significantly more likely to avoid trips during morning rush hour. Participants with high FOF, like participants with a fall history, had significantly more hard braking events but did not differ in speeding events.

After adjusting for additional sociodemographic, cognitive, and health factors, healthcare utilization and CNS medication use, participants with high FOF did not differ significantly from participants with low FOF in their driving exposure, driving space, or unsafe driving (Table 4). However, individuals with high FOF took a significantly smaller percentage of trips in morning rush hour and a significantly greater percentage of trips during evening rush hour. Groups did not differ significantly by any other measure of driving avoidance.

4. Discussion

Older drivers who reported at least one fall in the past year had a modestly higher rate of rapid deceleration events, a potential marker for unsafe driving, compared to adults who did not report falling. The observed difference was attenuated after accounting for differences in cognitive concerns and CNS medication use. Otherwise, driving habits measured objectively over a 12-month period were essentially unrelated to recent fall history. On the other hand, older drivers with a high FOF drove significantly differently than participants with low FOF; that is, they drove shorter distances, less often, and closer to home, and demonstrated more hard braking. Nearly all differences between participants with high

Table 3
Association between Past Year Fall and Objectively-Measured Driving Habits in the Subsequent 12 Months, LongROAD Cohort of Older Drivers.

Driving Outcomes	Past-Year Fall n = 828 Mean (SD)	No Past-Year Fall n = 2113 Mean (SD)	Base Model [a] Beta Estimate (95% CI)	Adjusted Model Beta Estimate (95% CI)
Driving Exposure				
Miles Driven	732.7 (433.8)	737.3 (433.7)	4.16 (−30.73, 39.06)	15.92 (−19.34, 51.17) [bc]
Days Driving	20.9 (5.5)	21.0 (5.5)	−0.14 (−0.59, 0.31)	−0.06 (−0.51, 0.39) [c]
Trips Driven	110.3 (53.5)	111.0 (53.3)	0.65 (−3.70, 5.00)	1.45 (−2.93, 5.83) [c]
Driving Space				
% Trips Within 15 Miles of Home	63.9 (22.2)	64.2 (22.4)	−0.25 (−2.04, 1.53)	−1.01 (−2.81, 0.78) [bc]
Driving Avoidance				
% Trips on High Speed Roads	12.2 (10.6)	13.1 (11.1)	−0.77 (−1.66, 0.11)	−0.77 (−1.66, 0.11)
% Trips in AM Peak	6.8 (4.9)	7.2 (4.9)	−0.32 (−0.72, 0.08)	−0.24 (−0.64, 0.16) [c]
% Trips in PM Peak	9.7 (4.3)	9.5 (4.5)	0.19 (−0.17, 0.55)	0.19 (−0.17, 0.55)
Log % Trips at Night	1.8 (0.7)	1.8 (0.7)	0.01 (−0.04, 0.07)	0.02 (−0.03, 0.08) [c]
Right-to-Left Turn Ratio	0.9 (0.1)	0.9 (0.1)	0.00 (−0.01, 0.01)	0.00 (−0.01, 0.01) [bc]
Unsafe Driving				
Log Rapid Deceleration Events per 1000 Miles Driven	1.6 (0.7)	1.5 (0.7)	**0.08 (0.03, 0.14)**	0.06 (−0.00, 0.12) [bc]
	N (%)	**N (%)**	**Odds Ratio (95% CI)**	**Odds Ratio (95% CI)**
At Least One Speeding Event per 1000 Miles Driven	468.0 (56.5)	1236.0 (58.5)	0.94 (0.80, 1.11)	0.90 (0.76, 1.07) [b]

Bold font indicates statistical significance at $p < 0.05$.
[a] Base model adjusted for gender, age, race, and marital status.
[b] Adjusted for cognitive concerns.
[c] Adjusted for current use of central nervous system medications (psychotherapeutics, anxiolytics, sedatives, hypnotics and anticonvulsants).

Table 4
Association between Fear of Falling and Objectively-Measured Driving Habits in the Subsequent 12 Months, LongROAD Cohort of Older Drivers.

Driving Outcomes	High Fall Concern n = 547 Mean (SD)	Low Fall Concern n = 2391 Mean (SD)	Base Model[a] Beta Estimate (95% CI)	Adjusted Model Beta Estimate (95% CI)
Driving Exposure				
Miles Driven	659.4 (442.5)	752.5 (427.5)	**−60.66 (−101.15, −20.18)**	−10.92 (−57.04, 35.20)[bcdef]
Days Driving	20.3 (6.0)	21.2 (5.4)	**−0.84 (−1.36, −0.31)**	−0.31 (−0.90, 0.27)[cdef]
Trips Driven	104.0 (50.2)	112.4 (54.0)	**−7.48 (−12.56, −2.41)**	−3.00 (−8.69, 2.70)[cdefg]
Driving Space				
% of Trips Within 15 Miles of Home	68.7 (21.7)	63.1 (22.4)	**3.92 (1.84, 5.99)**	1.81 (−0.05, 4.13)[bcdg]
Driving Avoidance				
% Trips on High Speed Roads	11.5 (10.4)	13.1 (11.1)	−0.88 (−1.92, 0.15)	−0.08 (−1.11, 0.95)[efgh]
% Trips in AM Peak	6.1 (4.6)	7.3 (4.9)	**−1.08 (−1.54, −0.62)**	**−0.57 (−1.08, −0.06)**[cde]
% Trips in PM Peak	9.9 (4.5)	9.5 (4.4)	0.41 (−0.01, 0.83)	**0.60 (0.18, 1.02)**[deh]
Log % of Trips at Night	1.8 (0.7)	1.9 (0.7)	−0.01 (−0.07, 0.06)	0.01 (−0.06, 0.07)[bceh]
Right-to-Left Turn Ratio	0.9 (0.1)	0.9 (0.1)	0.00 (−0.01, 0.01)	−0.01 (−0.02. 0.01)[bc]
Unsafe Driving				
Log Rapid Deceleration Events per 1000 Miles Driven	1.7 (0.8)	1.5 (0.7)	**0.10 (0.03, 0.17)**	0.05 (−0.02, 0.12)[bdfi]
	N (%)	**N (%)**	**Odds Ratio (95% CI)**	**Odds Ratio (95% CI)**
At Least One Speeding Event	308 (56.3)	1393 (58.3)	0.95 (0.79, 1.16)	0.98 (0.80, 1.20)[bfh]

Bold font indicates statistical significance at $p < 0.05$.
[a]Base model adjusted for gender, age, race, and marital status. Additional adjustment for general cognitive concerns[b], physical function[c], current use of central nervous system medications (included psychotherapeutics, anxiolytics, sedatives, hypnotics, and anticonvulsants)[d], work for pay in last month[e], total household income[f], alcohol consumption past three months[g], highest level of education[h], hospital stay past 12 months[i].

versus low FOF were accounted for by health and sociodemographic differences between these two groups.

Like our study, the systematic review by Scott et al. (2017) did not find consistent evidence of an association between prior self-reported falls and driving frequency, distance or space. However, all but one of those studies were based on subjectively-measured driving habits. The one included study with objectively-measured naturalistic driving (Crizzle et al., 2013) was limited to a small number of patients with Parkinson's disease followed for only two weeks, but similarly found no significant differences in driving exposure between those who had and had not fallen. Overall, our data suggest that having a past-year fall history does not lead older drivers to substantially reduce where or how often they drive. In contrast, FOF was found to be significantly associated with driving less often and for shorter distances. Previous research has shown that FOF is associated with cognitive decline, frailty, poor health, and gait abnormalities (Scheffer et al., 2008). While actual functional declines influence driving ability, both actual capacity and self-perceived physical, cognitive and perceptual deficits influence driving behavior (Anstey et al., 2005; Betz & Lowenstein, 2010; Molnar et al., 2013a,b; Molnar et al., 2015). Consistent with past research, we found that greater concerns about cognitive function and poorer perceived physical function helped explain observed differences in driving exposure and space between participants with high and low FOF. FOF may be a marker for perceived sensory, cognitive or physical deficits that lead participants to limit their driving exposure. FOF has previously been shown to be associated with restrictions in amount and type of physical and social activity (Scheffer et al., 2008). Our findings demonstrate that FOF is similarly associated with reduced driving mobility, with potentially serious adverse effects on access to goods and services and on social and civic engagement (Dickerson et al., 2019; Oxley & Whelan, 2008; Satariano et al., 2019; Webber et al., 2010).

We found little evidence that either FOF or a history of falls was associated with avoidance of difficult driving situations (e.g., driving during rush hour or at night). Evidence from prior studies is limited and inconsistent. Crizzle et al. (2013) similarly found no evidence that those with a history of falls had greater objectively-measured driving avoidance. In contrast, Vance et al. (2006) found that the number of self-reported falls was positively correlated with a composite driving avoidance score. Regardless of fall history or FOF, older drivers enrolled in our study generally avoided challenging driving situations, making relatively small percentages of trips on highways, during rush hour or at night.

Both a history of falls and a high FOF were associated with a modestly higher rate of rapid deceleration events (RDEs), whch indicate hard braking. Crizzle et al. (2013) also found significantly more hard braking among those with a fall history in a small sample with Parkinson's disease during two weeks of objectively-measured driving. Rapid deceleration events (RDEs) may relate to near crashes or crashes (Chevalier et al., 2017; Dingus et al., 2006; Yan et al., 2008), and have also been associated with driving violations (Zhao et al., 2012) and declining functional abilities (Eby et al., 2019). Cognitive impairment, which is associated with both fall risk (Deandrea et al., 2010) and FOF (Scheffer et al., 2008), may contribute to rapid deceleration events (Eby et al., 2019) and poorer driving performance (Hird et al., 2016; Jekel et al., 2015). In our sample, participants who had fallen assessed their cognitive function to be poorer than did those who had not fallen. We also found that participants with a fall history were more likely to currently take CNS medications, which may adversely impact driving performance. Use of drugs affecting the CNS are independent risk factors for both falls (Hartikainen et al., 2007; Park et al., 2015) and impaired driving (Hill et al., 2020; Hetland & Carr, 2014). Accounting for CNS medication use amd participants' concerns about their cognition attenuated the estimated association of past-year falls with rapid deceleration events, supporting the concept of underlying risk leading to both increased falls and poorer driving performance. It must be noted that the median difference in rapid deceleration events between those who fell and those who did not was less than one event per 1000 miles driven. Whether such a small difference translates into meaningful differences in crash risk is uncertain.

Our results showed that participants with high concern about falling had significantly lower education level and household income than those with low FOF. Numerous studies of community-dwelling older adults in diverse countries have similarly documented an association between lower education and increased fear of falling, after accounting for demographic, social and physical risk factors (Braga Lde et al., 2016; Choi et al., 2015;

Curcio et al., 2020; Dierking et al., 2016; Kumar et al., 2014; Lee et al., 2018; Mane et al., 2014; Oh et al., 2017), although a few studies found no association (Malini et al., 2016; Moreira et al., 2017; Pirrie et al., 2020). Several studies have also noted a relationship of FOF with lower socioeconomic status (SES) (Kumar et al., 2014; Vellas et al., 1997), although most studies examining SES have reported no association in adjusted models (Choi et al., 2015; Lee et al., 2018; Lee et al., 2018; Malini et al., 2016; Mane et al., 2014). Substantial evidence exists that lower education and income are correlated with poor health (Glymour et al., 2014). Thus, the greater odds of fear of falling reported by older drivers with lower education and income in our study may primarily reflect poorer health, although we did adjust for diverse measures of health in models of driving outcomes, such as cognitive and physical function, selected medications and hospitalization. Low education and socioeconomic status may also have a more direct effect on fear of falling, as Paiva et al. (2020) suggests: "Individuals who live in situations of social vulnerability experiencing material deprivation, a higher level of stress, fewer options, ..., and limited access to healthcare services suffer more intense consequences of falls." Thus, greater FOF may reflect awareness on the part of socioeconomically disadvantaged older adults of the potentially more serious consequences for them in the event of a fall. Regardless of the mechanism, this suggests that social disadvantage may lead to greater reductions in access to goods and services and to social and civic engagement as a consequence of higher FOF and associated reduced driving mobility, thus further exacerbating social inequalities in older adults.

This study had several limitations. Past-year fall history was self-reported. Research suggests under-reporting of falls by older adults (Ganz et al., 2005; Peel, 2000), which may have biased results toward the null. However, in the systematic review by Scott et al. (2017) all studies that specified the method of fall assessment used self-report; hence, our measure is consistent with other studies examining falls in relation to driving and crash risk. There were few speeding and rapid deceleration events, which reduced the study's power to identify differences between those who did and did not fall in the past year. We were unable to determine from our data whether participants' driving habits had changed subsequent to their fall or to examine temporal relationships between falls and use of CNS medication. We examined self-reported health characteristics as covariates. While perceived health is likely to be an important influence on both FOF and driving habits, we acknowledge that inclusion of objective sensory and physical function may have yielded differing results in adjusted models. Further, we lacked data on fall injuries, precluding evaluations of fall injuries' influence on driving habits. Although study participants were mostly relatively affluent, well-educated older drivers, approximately-one-third lacked a bachelor's degree and more than half had household incomes below the US median. Nevertheless, the adverse effects of FOF on driving may have been underestimated due to the sample characteristics. These same characteristics may also reduce the generalizability of our findings to other more socioeconomically diverse populations. Participants were recruited at sites that were selected for geographic diversity and may not represent the general US population. Among the strengths of this study are its inclusion of a large sample of older drivers recruited at geographically diverse sites, the use of objective driving data over a 12-month period after baseline, and the ability to account for differences in demographics, health and functional ability, and health care utilization between those with and without a fall history, and between those with high versus low FOF.

In conclusion, we found little evidence that the previously observed motor vehicle crash risk associated with a history of falls in older adults could be explained by differences in driving habits subsequent to the fall(s), when driving behaviors are measured objectively. The few differences in driving habits observed between those who fell and those who did not were largely explained by differences in perceived cognitive function and use of CNS medications. Rather than directly causing crash risk, falls may serve as a marker for older drivers who are at higher risk for motor vehicle crashes due to underlying age-related changes. We also found that older drivers with high FOF drove fewer days and miles, made fewer trips, and drove closer to home. Whether FOF itself reduces driving mobility or is a marker for actual or perceived physical, visual, and/or mental declines associated with aging or disease, remains unclear. Further studies with prospective collection of data on falls, fall injuries, and FOF, and examination of changes in driving habits in relation to each of these, can help to clarify the underlying relationships among them. Clinical trials may determine whether addressing the underlying factors that may have led to the fall, for example, careful assessment of cognitive function and consideration of dosage, frequency or class of medications taken that are known to have CNS effects, is more effective for reducing crashes than on-road driver training or similar measures to improve driving practices among those who have fallen. Similarly, clinical trials in persons with high fear of falling should evaluate the effects on mobility, quality of life, and independence of interventions to counteract physical, visual, and/or mental declines associated with aging, and of interventions to identify and access alternative transportation resources.

Acknowledgement

The LongROAD Research Team also includes Howard Andrews, David LeBlanc, Lindsay Ryan, and Robert Santos. Scott Bogard led development of the GPS-derived driving behaviors.

Institutional review board approval

This study was approved under the following Institutional Review Board (IRB) protocol/human subjects approval numbers: Colorado Multiple IRB (COMIRB No. 14-0528, Principal Investigator [PI]: C. DiGuiseppi), Columbia University IRB (Study No. IRB-AAAN9950, PI: G. Li), Johns Hopkins Bloomberg School of Public Health IRB (Study No. 00006200, PI: V. Jones), Mary Imogene Bassett Hospital IRB (No. IRB0000200, PI: D. Strogatz), University of California, San Diego IRB (IRB No. 141800, PI: L. Hill), and University of Michigan IRB (Study No. HUM00094031, PI: D. Eby).

Funding sources

This work was funded by the AAA Foundation for Traffic Safety (Washington, DC), with additional support for REDCap from NIH/NCATS Colorado CTSA Grant Number UL1 TR002535. The AAA Foundation for Traffic Safety contributed to the design and methods of the LongROAD study and reviewed a draft manuscript. The lead author (CGD) made the decision to submit the manuscript for publication. The report's findings and conclusions are solely those of the authors and do not necessarily reflect official positions of the AAA Foundation for Traffic Safety or the NIH.

Author contributions

CGD, MEB, DWE, LLH, VCJ, TJM, LJM, DS and GL conceptualized and designed the study and its methods; CGD, MEB, KAS, DWE, LLH, VCJ, TJM, LJM, DS and GL oversaw acquisition of subjects; CGD and HAH analyzed the data and drafted the manuscript; CGD, HAH, MEB and KAS interpreted the data; all authors revised the manuscript critically for important intellectual content and approved the version to be published.

Declaration of interest

The authors have no conflicts of interest to declare.

References

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychological Review, 25* (1), 45–65. https://doi.org/10.1016/j.cpr.2004.07.008.

Bergen, G., Stevens, M. R., & Burns, E. R. (2016). Falls and fall injuries among adults aged ≥65 years — United States, 2014. *Morbidity and Mortality Weekly Report, 65* (37), 993–998 https://doi.org/10.15585/mmwr.mm6537a2.

Betz, M. E., & Lowenstein, S. R. (2010). Driving patterns of older adults: Results from the Second Injury Control and Risk Survey. *Journal of the American Geriatrics Society, 58*(10), 1931–1935. https://doi.org/10.1111/j.1532-5415.2010.03010.x.

Braga Lde, S., Lima-Costa, M. F., César, C. C., & Macinko, J. (2016). Social inequalities on selected determinants of active aging and health status indicators in a large Brazilian City (2003–2010). *J Aging Health, 28*(1), 180–196. https://doi.org/10.1177/0898264315589575.

Chevalier, A., Coxon, K., Chevalier, A. J., Clarke, E., Rogers, K., Brown, J., Boufous, S., Ivers, R., & Keay, L. (2017). Predictors of older drivers' involvement in rapid deceleration events. *Accident; Analysis and Prevention, 98*, 312–319. https://doi.org/10.1016/j.aap.2016.10.010.

Choi, K., & Ko, Y. (2015). Characteristics associated with fear of falling and activity restriction in south Korean older adults. *Journal of Aging and Health, 27*(6), 1066–1083. https://doi.org/10.1177/0898264315573519.

Cox, A. E., & Cicchino, J. B. (2021). Continued trends in older driver crash involvement rates in the United States: Data through 2017–2018. *Journal of Safety Research, 77*, 288–295. https://doi.org/10.1016/j.jsr.2021.03.013.

Crizzle, A. M., Myers, A. M., Roy, E. A., & Almeida, Q. J. (2013). Associations between falls, balance confidence, driving speed, braking, and other driving practices in Parkinson's Disease. *Physical & Occupational Therapy in Geriatrics, 33*(1), 72–86. https://doi.org/10.3109/02703181.2014.991057.

Curcio, C. L., Wu, Y. Y., Vafaei, A., Barbosa, J. F. S., Guerra, R., Guralnik, J., & Gomez, F. (2020). A regression tree for identifying risk factors for fear of falling: The International Mobility in Aging Study (IMIAS). *The Journals of Gerontology: Series A, 75*(1), 181–188. https://doi.org/10.1093/gerona/glz002.

Deandrea, S., Lucenteforte, E., Foschi, R., La Vecchia, C., & Negri, E. (2010). Risk factors for falls in community-dwelling older people: A systematic review and meta-analysis. *Epidemiology, 21*(5), 658–668. https://doi.org/10.1097/ede.0b013e3181e89905.

Delbaere, K., Close, J. C. T., Mikolaizak, A. S., Sachdev, P. S., Brodaty, H., & Lord, S. R. (2010). The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. *Age and Ageing, 39*(2), 210–216. https://doi.org/10.1093/ageing/afp225.

Dickerson, A. E., Molnar, L. J., Bédard, M., Eby, D. W., Berg-Weger, M., Choi, M., Grigg, J., Horowitz, A., Meuser, T., Myers, A., O'Connor, M., & Silverstein, N. M. (2019). Transportation and aging: An updated research agenda to advance safe mobility among older adults transitioning from driving to non-driving. *The Gerontologist, 59*(2), 215–221. https://doi.org/10.1093/geront/gnx120.

Dierking, L., Markides, K., Al Snih, S., & Kristen, P. M. (2016). Fear of falling in older Mexican Americans: A longitudinal study of incidence and predictive factors. *Journal of the American Geriatrics Society, 64*(12), 2560–2565. https://doi.org/10.1111/jgs.14496.

Dingus, T. A., Klauer, S.G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., & Knipling, R.R. (2006, March). The 100-car naturalistic driving study, Phase II – Results of the 100-car field experiment. US Department of Transportation. https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/100carmain.pdf.

Dugan, E., & Lee, C. M. (2013). Biopsychosocial risk factors for driving cessation: Findings from the health and retirement study. *Journal of Aging & Health, 25*, 1313–1328.

Forrest, K. Y., Bunker, C. H., Songer, T. J., Coben, J. H., & Cauley, J. A. (1997). Driving patterns and medical conditions in older women. *Journal of the American Geriatrics Society, 45*, 1214–1218.

Eby, D. W., Molnar, L. J., Zakrajsek, J. S., Ryan, L. H., Zanier, N., St. Louis, R. M., Stanciu, S. C., LeBlanc, D., Bogard, S., Kostyniuk, L. P., Smith, J., Yung, R., Nyquist, L., Betz, M. E., DiGuiseppi, C., Jones, V., Li, G., Mielenz, T. J., Strogatz, D., & on behalf of the AAA LongROAD Research Team (2019). Factors related to rapid deceleration events among a large cohort of older drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 66*, 76–86. https://doi.org/10.1016/j.trf.2019.08.021.

Friedman, S. M., Munoz, B., West, S. K., Rubin, G. S., & Fried, L. P. (2002). Falls and fear of falling: Which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. *Journal of the American Geriatrics Society, 50*(8), 1329–1335. https://doi.org/10.1046/j.1532-5415.2002.50352.x.

Ganz, D. A., Higashi, T., & Rubenstein, L. Z. (2005). Monitoring falls in cohort studies of community-dwelling older people: Effect of the recall interval. *Journal of the American Geriatrics Society, 53*(12), 2190–2194. https://doi.org/10.1111/j.1532-5415.2005.00509.x.

Glymour, M. M., Avendano, M., & Kawachi, I. (2014). Socioeconomic status and health. *Social Epidemiology, 2*, 17–63.

Hartikainen, S., Lönnroos, E., & Louhivuori, K. (2007). The journals of gerontology. *Series A, Biological Sciences and Medical Sciences, 62*(10), 1172–1181. https://doi.org/10.1093/gerona/62.10.1172.

HealthMeasures (2020, December). PROMIS® Score and Interpret. http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis.

Hetland, A., & Carr, D. B. (2014). Medications and impaired driving. *The Annals of Pharmacotherapy, 48*(4), 494–506. https://doi.org/10.1177/1060028014520882.

Hill, L. L., Andrews, H., Li, G., DiGuiseppi, C. G., Betz, M. E., Strogatz, D., Pepa, P., Eby, D. W., Merle, D., Kelley-Baker, T., Jones, V., & Pitts, S. (2020). Medication use and driving patterns in older drivers: Preliminary findings from the LongROAD study. *Injury Epidemiology, 7*(1), 38 https://dx.doi.org/10.1186/s40621-020-00265-y.

Hird, M. A., Egeto, P., Fischer, C. E., Naglie, G., & Schweizer, T. A. (2016). A systematic review and meta-analysis of on-road simulator and cognitive driving assessment in Alzheimer's Disease and mild cognitive impairment. *Journal of Alzheimer's Disease: JAD, 53*(2), 713–729. https://doi.org/10.3233/jad-160276.

Jekel, K., Damian, M., Wattmo, C., Hausner, L., Bullock, R., Connelly, P. J., Dubois, B., Eriksdotter, M., Ewers, M., Graessel, E., Kramberger, M. G., Law, E., Mecocci, P., Molinuevo, J. L., Nygård, L., Olde-Rikkert, M. G., Orgogozo, J., Pasquier, F., Peres, K., ... Frölich, L. (2015). Mild cognitive impairment and deficits in instrumental activities of daily living: A systematic review. *Alzheimers Research & Therapy, 7* (1), 17. https://doi.org/10.1186/s13195-015-0099-0.

Karthaus, M., & Falkenstein, M. (2016). Functional changes and driving performance in older drivers: Assessment and interventions. *Geriatrics (Basel), 1*(2), 12 https://dx.doi.org/10.3390/geriatrics1020012.

Kumar, A., Carpenter, H., Morris, R., Iliffe, S., & Kendrick, D. (2014). Which factors are associated with fear of falling in community-dwelling older people? *Age and Ageing, 43*(1), 76–84. https://doi.org/10.1093/ageing/aft154.

Lee, S., Lee, C., Ory, M. G., Won, J., Towne, S. D., Jr, Wang, S., & Forjuoh, S. N. (2018). Fear of outdoor falling among community-dwelling middle-aged and older adults: The role of neighborhood environments. *Gerontologist, 58*(6), 1065–1074. https://doi.org/10.1093/geront/gnx123.

Lee, S., Oh, E., & Hong, G. S. (2018). Comparison of factors associated with fear of falling between older adults with and without a fall history. *International Journal of Environmental Research and Public Health, 15*(5), 982. https://doi.org/10.3390/ijerph15050982.

Li, G., Eby, D. W., Santos, R., Mielenz, T. J., Molnar, L. J., Strogatz, D., Betz, M. E., DiGuiseppi, C., Ryan, L. H., Jones, V., Pitts, S. I., Hill, L. L., DiMaggio, C. J., LeBlanc, D., Andrews, H. F., & for the LongROAD Research Team (2017). Longitudinal Research On Aging Drivers (LongROAD): Study design and methods. *Injury Epidemiology, 4*(1), 22. https://doi.org/10.1186/s40621-017-0121-z.

Liu, M., Hou, T., Li, Y., Sun, X., Szanton, S. L., Clemson, L., & Davidson, P. M. (2021). Fear of falling is as important as multiple previous falls in terms of limiting daily activities: A longitudinal study. *BMC Geriatrics, 21*, 350. https://doi.org/10.1186/s12877-021-02305-8.

Lyman, J. M., McGwin, G., & Sims, R. V. (2001). Factors related to driving difficulty and habits in older drivers. *Accident Analysis & Prevention, 33*, 413–421.

MacLeod, K. E., Satariano, W. A., & Ragland, D. R. (2014). The impact of health problems on driving status among older adults. *Journal of Transportation & Health, 1*, 86–94.

Malini, F. M., Lourenço, R. A., & Lopes, C. S. (2016). Prevalence of fear of falling in older adults, and its associations with clinical, functional and psychosocial factors: The Frailty in Brazilian Older People-Rio de Janeiro study. *Geriatrics & Gerontology International, 16*(3), 336–344. https://doi.org/10.1111/ggi.12477.

Mane, A. B., Sanjana, T., Patil, P. R., & Sriniwas, T. (2014). Prevalence and correlates of fear of falling among elderly population in urban area of Karnataka, India. *Journal of Mid-life Health, 5*(3), 150–155. https://doi.org/10.4103/0976-7800.141224.

Marie Dit Asse, L., Fabrigoule, C., Helmer, C., Laumon, B., & Lafont, S. (2014). Automobile driving in older adults: Factors affecting driving restriction in men and women. *Journal of the American Geriatrics Society, 62*(11), 2071–2078. https://doi.org/10.1111/jgs.13077.

Molnar, L. J., Charlton, J. L., Eby, D. W., Bogard, S. E., Langford, J., Koppel, S., Kolenic, G. E., Marshall, S., & Ma-Son-Hing, M. (2013). Self-regulation of driving by older adults: Comparison of self-report and objective driving data. *Transportation Research Part F: Traffic Psychology and Behaviour, 20*, 29–38. https://doi.org/10.1016/j.trf.2013.05.001.

Molnar, L. J., Charlton, J. L., Eby, D. W., Langford, J., Koppel, S., Kolenic, G. E., & Marshall, S. (2013). Factors affecting self-regulatory driving practices among older adults. *Traffic Injury Prevention, 15*(3), 262–272. https://doi.org/10.1080/15389588.2013.808742.

Molnar, L. J., Eby, D. W., Zhang, L., Zanier, N., St. Louis, R. M., & Kostyniuk, L. P. (2015, December). Self-regulation of driving by older adults: A synthesis of the literature and framework for future research. AAA Foundation for Traffic Safety. https://aaafoundation.org/wp-content/uploads/2017/12/SelfRegulationOfDrivingByOlderAdultsReport.pdf.

Moreira, B. S., Sampaio, R. F., Diz, J. B., Bastone, A. C., Ferriolli, E., Neri, A. L., Lourenço, R. A., Dias, R. C., & Kirkwood, R. N. (2017). Factors associated with fear of falling in community-dwelling older adults with and without diabetes mellitus: Findings from the Frailty in Brazilian Older People Study (FIBRA-BR). *Experimental Gerontology, 89*, 103–111. https://doi.org/10.1016/j.exger.2017.01.004.

Oh, E., Hong, G. S., Lee, S., & Han, S. (2017). Fear of falling and its predictors among community-living older adults in Korea. *Aging & Mental Health, 21*(4), 369–378. https://doi.org/10.1080/13607863.2015.1099034.

Owsley, C., Stalvey, B., Wells, J., & Sloane, M. E. (1999). Older drivers and cataract: Driving habits and crash risk. *The Journals of Gerontology Series A, Biological Sciences and Medical Sciences, 54*(4), M203–M211. https://doi.org/10.1093/gerona/54.4.m203.

Oxley, J., & Whelan, M. (2008). It cannot be all about safety: The benefits of prolonged mobility. *Traffic Injury Prevention, 9*(4), 367–378. https://doi.org/10.1080/15389580801895285.

Paiva, M. M., Lima, M. G., & Barros, M. B. A. (2020). Social inequalities in the impact of falls on health-related quality of life among older adults. *Cien Saude Colet, 25*(5), 1887–1896. https://doi.org/10.1590/1413-81232020255.34102019. English, Portuguese.

Park, H., Satoh, H., Miki, A., Urushihara, H., & Sawada, Y. (2015). Medications associated with falls in older people: Systematic review of publications from a recent 5-year period. *European Journal of Clinical Pharmacology, 71*(12), 1429–1440. https://doi.org/10.1007/s00228-015-1955-3.

Peel, N. (2000). Validating recall of falls by older people. *Accident; Analysis and Prevention, 32*(3), 371–372. https://doi.org/10.1016/s0001-4575(99)00066-4.

Pirrie, M., Saini, G., Angeles, R., Marzanek, F., Parascandalo, J., & Agarwal, G. (2020). Risk of falls and fear of falling in older adults residing in public housing in Ontario, Canada: Findings from a multisite observational study. *BMC Geriatrics, 20*(1), 11. https://doi.org/10.1186/s12877-019-1399-1.

Pomidor, A. (Ed.). (2019). *Clinician's guide to assessing and counseling older drivers* (4th ed.. New York: The American Geriatrics Society.

Satariano, W. A., Guralnik, J., Jackson, R. J., Marottoli, R. A., Phelan, E. A., & Prohaska, T. R. (2019). Mobility and aging: New directions for public health action. *American Journal of Public Health, 102*(8), 1508–1515 https://dx.doi.org/10.2105/AJPH.2011.300631.

Scheffer, A. C., Schuurmans, M. J., van Dijk, N., van der Hooft, T., & de Rooij, S. E. (2008). Fear of falling: Measurement strategy, prevalence, risk factors and consequences among older persons. *Age and Ageing, 37*(1), 19–24. https://doi.org/10.1093/ageing/afm169.

Scott, K. A., Rogers, E., Betz, M. E., Hoffecker, L., Li, G., & DiGuiseppi, C. (2017). Associations between falls and driving outcomes in older adults: Systematic review and meta-analysis. *Journal of the American Geriatrics Society, 65*(12), 2596–2602. https://doi.org/10.1111/jgs.15047.

Takemoto, M., Carlson, J. A., Moran, K., Godbole, S., Crist, K., & Kerr, J. (2015). Relationship between objectively measured transportation behaviors and health characteristics in older adults. *International Journal of Environmental Research and Public Health, 12*(11), 13923–13937. https://doi.org/10.3390/ijerph121113923.

Tinetti, M. E., & Powell, L. (1993). Fear of falling and low self-efficacy: A case of dependence in elderly persons. *Journal of Gerontology, 48*, 35–38. https://doi.org/10.1093/geronj/48.special_issue.35.

Vance, D. E., Roenker, D. L., Cissell, G. M., Edwards, J. D., Wadley, V. G., & Ball, K. K. (2006). Predictors of driving exposure and avoidance in a field study of older drivers from the state of Maryland. *Accident; Analysis and Prevention, 38*(4), 823–831. https://doi.org/10.1016/j.aap.2006.02.008.

Vellas, B. J., Wayne, S. J., Romero, L. J., Baumgartner, R. N., & Garry, P. J. (1997). Fear of falling and restriction of mobility in elderly fallers. *Age and Ageing, 26*(3), 189–193. https://doi.org/10.1093/ageing/26.3.189.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for reporting observational studies. *BMJ, 335*(7624), 806–808. https://doi.org/10.1136/bmj.39335.541782.AD.

Webber, S. C., Porter, M. M., & Menec, V. H. (2010). Mobility in older adults: A comprehensive framework. *The Gerontologist, 50*(4), 443–450. https://doi.org/10.1093/geront/gnq013.

Williams, A. F., Kyrychenko, S. Y., & Retting, R. A. (2006). Characteristics of speeders. *Journal of Safety Research, 37*(3), 227–232. https://doi.org/10.1016/j.jsr.2006.04.001.

Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. *Accident; Analysis and Prevention, 40*(1), 274–288. https://doi.org/10.1016/j.aap.2007.06.007.

Zhao, N., Mehler, B., Reimer, B., D'Amdrosio, L. A., Mehler, A., & Coughlin, J. F. (2012). An investigation of the relationship between the driving behavior questionnaire and objective measures of highway driving behaviors. *Transportation Research Part F: Traffic Psychology and Behaviour, 15*, 676–685. https://doi.org/10.1016/j.trf.2012.08.001.

Dr. Carolyn DiGuiseppi, a board-certified preventive medicine physician, is Professor of Epidemiology in the Colorado School of Public Health and Professor of Pediatrics and Director of Evidence-Based Medicine in the School of Medicine at the University of Colorado Anschutz Medical Campus. She has published more than 200 scientific journal articles, book chapters and scholarly reviews. A primary focus of her research has been the epidemiology and prevention of unintentional injuries, including motor vehicle injuries and falls. Her recent work focuses on the epidemiology of injuries among older adults and among persons with autism spectrum disorder. She has served on a variety of federal and state advisory committees and is currently on the editorial boards of *Injury Prevention* and *Injury Epidemiology*.

Journal of Safety Research 83 (2022) 210–222

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Biomechanical assessment of a passive back-support exoskeleton during repetitive lifting and carrying: Muscle activity, kinematics, and physical capacity

Billy Chun Lung So, Chunzhuo Hua, Tingting Chen, Qingwen Gao, Siu Shing Man *

Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Article history:
Received 27 August 2021
Received in revised form 24 April 2022
Accepted 25 August 2022
Available online 6 September 2022

Keywords:
Muscle activity
Kinematics
Physical capacity
Manual handling operation
Passive back-support exoskeleton

ABSTRACT

Introduction: Most people have experienced low back pain (LBP) more or less in their lifetime. Heavier load weight could increase the risk of LBP, especially in repetitive lifting and carrying tasks. The risk could also increase with the frequency of lifting. This study aims to investigate the effects of a passive back-support exoskeleton (PBSE) on trunk muscle activation, kinematics, and physical capacity in a repetitive lifting task and a carrying task in consideration of load weights in a laboratory setting. *Results*: Results showed that using the PBSE, the activities of the thoracic erector spinae and lumbar erector spinae muscles were reduced significantly by nearly 7% MVC and 3% MVC in the repetitive lifting task and the carrying task, respectively. There was no significant effect of the PBSE on the spine kinematics and physical capacity. *Practical Applications:* This study supports the use of the PBSE to reduce trunk muscle activity in repetitive lifting and carrying tasks.

1. Introduction

1.1. Low back pain and manual handling tasks

Low back pain (LBP) is a significant occupational health issue among workers (Kudo, Yamada, & Ito, 2019), a leading cause of disability (Ferguson et al., 2019), and falls within the realm of musculoskeletal disorders (Wang, Liu, Lu, & Koo, 2015). Most people experience LBP more or less in their lifetime. The most frequently affected area is between the inferior margin of the 12th rib and inferior gluteal folds (Koes, Van Tulder, & Thomas, 2006). The incidence of LBP increases with age, and this incidence reaches a peak in the 30-year-old population (Hoy, Brooks, Blyth, & Buchbinder, 2010). It was reported that LBP prevalence rates may range from 4% (Freburger et al., 2009) to 69% (Ganasegeran, Perianayagam, Nagaraj, & Al-Dubai, 2014), which depends on the length of time and pain intensity. The least commonly used remedies for LBP were taking time off from work (i.e., lost work time) and seeking medical care (Davis & Kotowski, 2015).

Manual handling tasks of heavy loads may enhance the risk of LPB in different industrial sectors, such as the healthcare industry (Nourollahi, Afshari, & Dianat, 2018) and construction industry (Wang, Dai, & Ning, 2015). The risk can also increase with the frequency of lifting (Hoogendoorn et al., 2000). According to a prospective cohort study of personal risk factors for LBP, bending for more than two hours every day is strongly and independently associated with the subsequent LBP, and the risk was even higher in those bending both forward and sideways (Ramond-Roquin et al., 2015). The degree of trunk flexion is also a risk factor of LBP. When a worker's trunk is a minimum of 60 degrees of flexion for more than 5% of the daily working time, working on the extreme flexion angle can increase the risk of LBP (Hoogendoorn et al., 2000). In a three-year follow-up study, the researchers found that prolonged standing, awkward lifting, and squatting/kneeling were highly related to mechanical LBP (Sterud & Tynes, 2013). Flexion of the trunk was classified as a moderate risk factor for LBP, especially at greater levels of exposure during repetitive lifting (Hoogendoorn et al., 2000). From the perspectives of kinematics, the risk of low back disorder due to working at limit angles can be reduced by limiting the peak flexion angle of lumbar spine (Adams & Hutton, 1982, 1986; Adams, McNally, Chinn, & Dolan, 1994). Passive back-support exoskeleton (PBSE) for preventing workers from LBP is attracting remarkable attention of safety researchers (Amandels, het Eyndt, Daenen, & Hermans, 2018; Bosch, van Eck, Knitel, & de Looze, 2016; Motmans, Debaets, & Chrispeels, 2018).

* Corresponding author.
E-mail address: siu-shing.man@polyu.edu.hk (S.S. Man).

https://doi.org/10.1016/j.jsr.2022.08.017
0022-4375/

1.2. Previous studies on passive back-support exoskeletons in manual handling tasks

Bosch et al. (2016) found that using a PBSE in a prolonged forward-bended work task can reduce 35–38% of lower muscle activity, lower physical discomfort in the low back, and increase endurance time, but increase physical discomfort in the chest region. For dynamic lifting and static holding tasks, using a PBSE can reduce approximate 10–40% of back muscle activity (De Looze, Bosch, Krause, Stadler, & O'Sullivan, 2016). Using a PBSE can also reduce Trapezius muscle activity during a reaching-bending task (Amandels et al., 2018). Motmans et al. (2018) found that during order picking the back muscle activity was reduced by 9–12% when wearing a PBSE. Huysamen et al. (2018) investigated a PBSE for lifting and lowering tasks and found that the PBSE can reduce Biceps Femoris and Erector Spinae muscle activity by 15% and 5%, respectively. PBSEs can significantly reduce perceived musculoskeletal effort in the trunk region (Huysamen et al., 2018). Koopman, Kingma, Faber, de Looze, and van Dieën (2019) investigated the effect of a PBSE on the activity of back and abdominal muscles, the flexion of hip and lumbar during a static bending task at five various hand heights. Subsequently, Alemi, Geissinger, Simon, Chang, and Asbeck (2019) and Alemi, Madinei, Kim, Srinivasan, and Nussbaum (2020) focused on repetitive lifting tasks and found that the use of PBSEs reduces the average peak and mean muscle activity of back and leg muscles in the tasks with different lifting technics (symmetric or asymmetric, kneeling, stoop, squat, and freestyle).

1.3. Research rationale and objective

Although the effects of PBSEs on muscle activity in manual handling tasks have been widely investigated in the literature, the knowledge about the effects of PBSEs on the activity of trunk muscles during repetitive lifting and carrying is insufficient, leaving a research gap. Spine kinematics is a risk factor of LBP (Hemming, Sheeran, van Deursen, & Sparkes, 2018; Papi, Bull, & McGregor, 2018) and physical capacity is one of interest in LBP related studies (Demoulin et al., 2013; Jakobsson, Gutke, Mokkink, Smeets, & Lundberg, 2019; Rasmussen et al., 2013). However, whether using PBSEs affects the spine kinematics and physical capacity of workers in repetitive lifting and carrying tasks has not been examined in the literature. Therefore, the objective of the current study was to examine the effects of a PBSE on the activity of trunk muscles, the spine kinematics and physical capacity of workers during repetitive lifting and carrying tasks. The hypotheses tested were: (1) using a PBSE can reduce trunk muscle activity (including thoracic erector spinae and lumbar erector spinae) during repetitive lifting and carrying tasks; (2) using a PBSE can reduce trunk muscle activity with the load weight; (3) using a PBSE can reduce more trunk muscle activity in lifting task than carrying task; (4) using a PBSE could benefit the users in both lifting and carrying tasks with different load weights by decreasing the trunk flexion; and (5) using a PBSE can increase the physical capacity during repetitive lifting task. The results of this study are expected to improve the existing literature on PBSEs in repetitive lifting and carrying tasks, enhance the understanding of how the PBSEs may affect muscle activity and body motion, elucidate the role of exoskeleton in preventing LBP in people, and provide practical recommendations for designing effective PBSEs.

2. Methodology

2.1. Study design

This was a cross-sectional experimental study conducted in a laboratory setting from February 2021 to May 2021. Each participant was required to perform two manual handling tasks (a repetitive load-lifting task and a load-carrying task). The muscle activities and kinematics of the trunk in each condition, and the physical capacity in the lifting task were recorded. All participants provided their informed consent as approved by the Human Subject Ethics Subcommittee.

2.2. Participants

Convenience sampling was used for this study. Twenty healthy males were recruited from the Hong Kong Polytechnic University. Participant inclusion criteria were: (a) Chinese male adults, and (b) Mandarin or English speakers. Participant exclusion criteria were: (a) medical history of cardiovascular disease, (b) recent (past 3 months) musculoskeletal injury or pain, and (c) exposure to COVID-19 cases in the last 28 days.

2.3. Exoskeleton

A PBSE device named "Muscle Suit Every" (Innophys, Japan) (Fig. 1) was used in this study. The actuator of this device is McKibben artificial muscles, which consists of an internal elastomeric bladder surrounded by a woven braided shell. The advantages of this device are lightweight, flexible, and simple construction. It could provide the spine with the maximum assistive force of 25.5 kg (56.2 lb) during lifting and carrying tasks. The exoskeleton has two sizes, which are small/medium size (Height × Width × Depth = 805 mm × 465 mm × 170 mm, weight = 4.3 kg) for body height of 150 cm to 165 cm, and medium/large size (Height × Width × Depth = 840 mm × 465 mm × 170 mm, weight = 4.4 kg) for body height of 160 cm to 185 cm. If the body height of the subjects was between 160 cm and 165 cm, they can select either the small/medium size or medium/large size of the exoskeleton, depending on which one can provide a better fit for them. Instruction sheets provided by the manufacturer of the exoskeleton for the proper use of the exoskeleton were given to the subjects. Table 1 shows the nine steps taken to properly use the exoskeleton. Additionally, a 30-min training session was provided to the subjects to ensure they selected and used the exoskeleton properly for the lifting and carrying tasks.

2.4. Experimental tasks

The tasks include a load-lifting task and a load-carrying task. Each task involved four conditions, with a PBSE or without a PBSE, 5 kg or 15 kg load weight. Those two tasks were selected because they were closely related to LBP (Coenen, Kingma, Boot, Bongers, & van Dieën, 2014; Hoogendoorn et al., 2000; Wai, Roffey, Bishop, Kwon, & Dagenais, 2010). According to a psychophysical

Fig. 1. Muscle Suit Every (Innophys, Japan).

Table 1
Steps taken to properly use the exoskeleton.

Steps	Content
1.	Using the shoulder belts and putting the exoskeleton on like a backpack.
2.	Pulling the shoulder belt adjustors and adjusting the exoskeleton so that the waist belt is at waist height.
3.	After fastening the waist belt, pulling the left and right adjustors to firmly tighten the belt.
4.	Adjusting the length of the hip belt.
5.	Bringing the thigh pads around to the front.
6.	Pumping the air pump 45 times and filling the exoskeleton with air.
7.	Allowing enough space for one fist to fit in the space between back and the suit.
8.	Fastening the left and right chest adjustors.
9.	The process is complete.

study, the maximum acceptable workloads for repetitive lifting during an 8-hour workday in industrial populations is 17.5 kg (Legg & Myles, 1981). In order to minimize the risk of injury of participants, 5 kg and 15 kg were selected as the load weights in this study for male adults (Heydari, Hoviattalab, Azghani, Ramezanzadehkoldeh, & Parnianpour, 2013; Huysamen, Power, & O'Sullivan, 2020).

2.4.1. Repetitive Load-Lifting task

In lifting task, participants were asked to repetitively lift a two-handled toolbox (Fig. 2) until subjective fatigue using a squat posture in four conditions of load weights (5 kg and 15 kg) and interventions (without the PBSE and with the PBSE). Subjective fatigue in this study was defined as overall fatigue. Lumbar Spine and thoracic spine are supposed to be the most affected joints during such tasks. The lifting task was a standardized task in which one full cycle included lifting the load from the ground level to the waist level with the angle of the elbow joint reaching 90 degrees, and then lowering the box down on the floor. Participants cannot move their feet during the full cycle. The lifting-and-lowering frequency was fixed at 10 full cycles per minute and controlled by a metronome. Participants performed repetitive lifting until subjective fatigue was reached (i.e., the participant could not complete a cycle of lifting even with a strong verbal encouragement). A 20-min recovery period was provided after each condition (Theurel, Desbrosses, Roux, & Savescu, 2018).

2.4.2. Load-carrying task

In the carrying task, participants were asked to walk a distance of eight meters at a free chosen (a usual walking pace) speed, carrying the two-handled toolbox (Fig. 3) in four conditions of load weights (5 kg and 15 kg) and interventions (without exoskeleton and with exoskeleton). During the carrying task, the participant kept the trunk upright and looked ahead, carrying the box at the waist level and with the elbow flexion at 90 degrees. For each condition, the carrying task was repeated three times, and a break of 10 seconds was given between each repetition. A 5-min recovery period was provided after each condition. This rest arrangement was adopted from a previous study of Theurel et al. (2018) to avoid any carryover fatigue across conditions.

2.5. Outcome measures

2.5.1. Surface electromyography (sEMG)

Muscle activity level of the thoracic erector spinae (TES) and lumbar erector spinae (LES) muscles were recorded on left and right sides using sEMG device (aktos, myon AG, Schwarzenburg, Switzerland) (Fig. 4). Before attaching the electrodes, the skin was shaved and cleaned by water, sandpaper and 75% alcohol to ensure the skin impedance was below 5 kΩ (Theurel et al., 2018). Standardized electrodes placement were conducted according to SENIAM Guidelines (Hermens & Freriks, 1997). Then, two pairs of wireless bipolar Ag/AgCl surface electrodes with 20 mm inter-electrode distance were attached bilaterally to the left and right sides of TES and LES. These muscles were selected because of their relevance to the tasks (Vleeming, Pool-Goudzwaard, Stoeckart, van Wingerden, & Snijders, 1995).

Prior to the lifting task, the participant was instructed to perform maximum voluntary isometric contraction (MVIC) of trunk muscles (TES and LES). The trunk MVIC was performed by stretching backwards against maximum resistance. The participant was in the prone position with the trunk suspended. One researcher pressed the ankle of the participant to provide stability, another researcher applied resistance near the scapular region. When performing the task, the provider was asked to increase the resisting force gradually until they reached maximum effort (Fig. 5). Non-threatening verbal encouragement was provided throughout. The participant maintained the MVIC for 5 seconds, and repeated MVIC trial twice with 2-min rest between trials. sEMG signals were recorded at 2000 Hz, and filtered at 20–500 Hz bandpass. Then the maximum root mean square (RMS) of each muscle was identified over successive periods of 100 ms sliding windows using MATLAB software (the MathWorks Inc., Natrick, MA, USA). The highest RMS sEMG signal of each muscle was chosen for normalization.

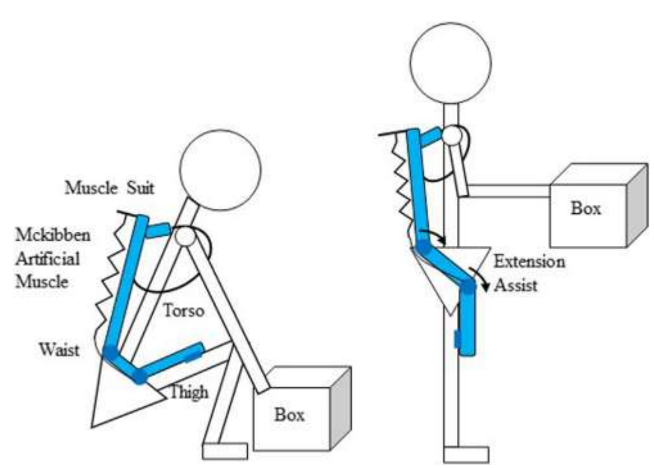

Fig. 2. Load-lifting task with the PBSE.

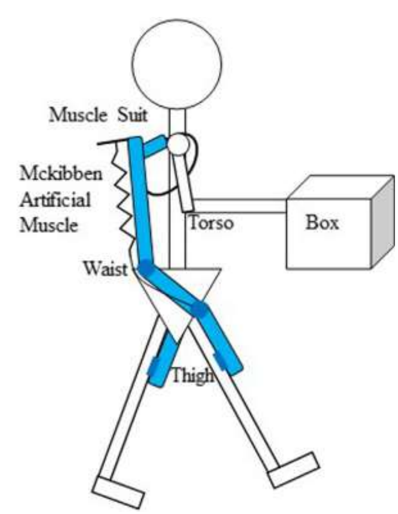

Fig. 3. Load-carrying task with exoskeleton.

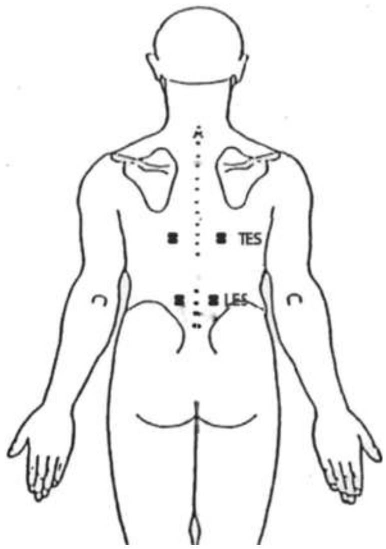

Fig. 4. Position of the sEMG electrodes (TES: thoracic erector spinae; LES: lumbar erector spinae).

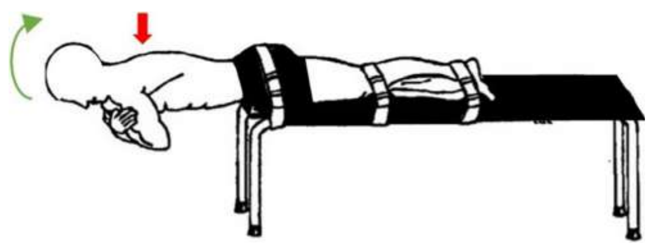

Fig. 5. MVIC test of trunk muscles.

Full-wave rectification and signal smoothing with a constant window of 100 ms RMS algorithm were used to process all sEMG signals. For the lifting task and carrying task, the maximum RMS values of EMG signals of various muscles from the entire motion were identified and subsequently divided by the corresponding maximum RMS of MVC to obtain the percentage of muscle activity (%MVC).

2.5.2. Spine kinematics

Kinematics data for flexion angle of thoracic and lumbar segments were measured using the MyoMotion system (Noraxon USA, Scoosdale, Arizona), and three motion sensors were placed over T4, L1, and S1 (Fig. 6). Previous studies have shown that trunk flexion is a risk factor for low back injuries, and workers have an increased risk of LBP with a minimum of 60 degrees of trunk flexion over 5% of the working time (Hoogendoorn et al., 2000). Also, body kinematics are closely associated with muscle activity as muscle performances changing at different joint angles (Leinonen, Kankaanpää, Airaksinen, & Hänninen, 2000). Participants were required to stand upright as the reference standing position for the calibration before each condition of tasks. Peak angles and flexion range of selected segments during the lifting and carrying were recorded for analysis.

2.5.3. Physical capacity

The number of the full cycles in the lifting task that participants can make with each condition until their subjective fatigue was recorded. The physical capacity of the participants in the lifting task can be reflected by the number of full cycles. The greater the number of the full cycles, the higher the physical capacity of the participants in the lifting task.

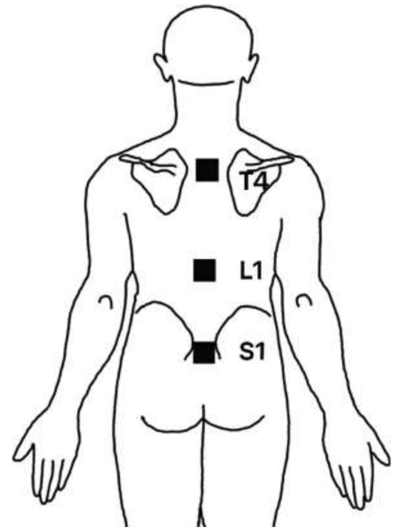

Fig. 6. Position of the motion sensors (T4: the 4th Thoracic column; L1: the first lumbar column; S1: sacrum).

2.6. Protocol

The experimental process was shown in Fig. 7. Before the experiment, the researchers explained the experimental process and showed the experimental tasks to the participants. Also, the participants were required to read the information sheet and sign the consent form, and then familiarize with the exoskeleton. The demographic data of the participant including age, height, weight, body mass index (BMI), and exercise frequency were recorded first.

After the skin preparation, electrodes were attached to the participants to test the MVIC of TES and LES. Then, the participant wore the motion sensors on selected regions to perform the experimental task. First, the participant was instructed to perform the lifting task until exhaustion in four conditions according to the order generated by Latin square in advance: (a) lifting 5 kg load weight with exoskeleton; (b) lifting 15 kg load weight with exoskeleton; (c) lifting 5 kg load weight without exoskeleton; and (d) lifting 15 kg load weight without exoskeleton. A 20-min break was given between two conditions. Before each condition, the participant was required to keep upright as the reference standing position to calibrate the motion sensor. The muscle activities and kinematics of the trunk in each condition as well as the number of the full cycles in each condition of the lifting task were recorded.

Subsequently, the carrying task was performed at the daily walking speed in four conditions based on the order generated by Latin square in advance: (a) carrying 5 kg load weight with exoskeleton; (b) carrying 15 kg load weight with exoskeleton; (c) carrying 5 kg load weight without exoskeleton; and (d) carrying 15 kg load weight without exoskeleton. Each condition was repeated three times, with 10 seconds rest between each repetition and 5 min rest between two conditions. Before each condition, the participant was required to keep upright as the reference standing position to calibrate the motion sensor. The muscle activities and kinematics of the trunk in each condition of the carrying task were recorded.

2.7. Statistical analysis

Demographic data including age, BMI, and exercise frequency were analyzed by descriptive statistics, including means and stan-

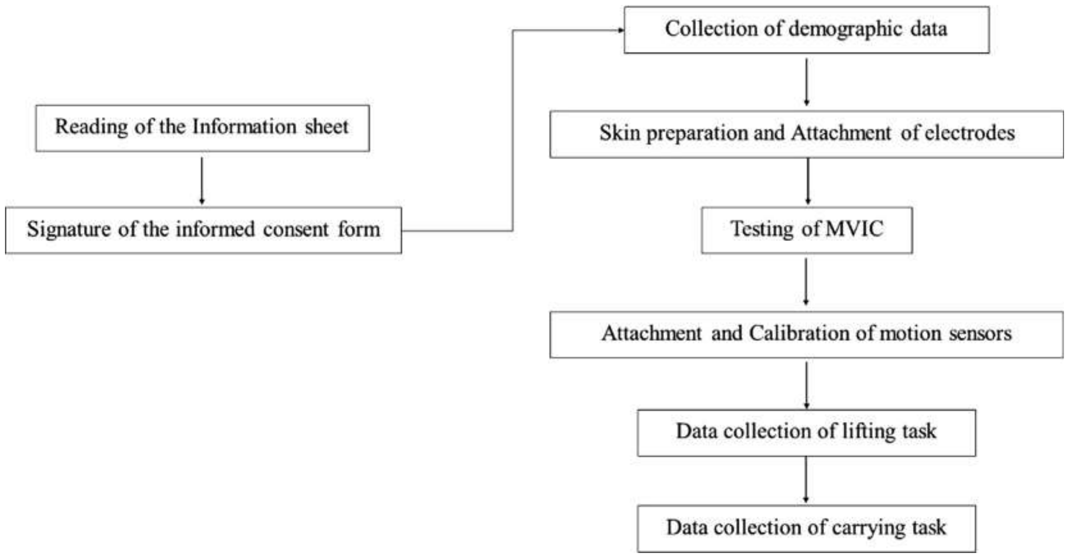

Fig. 7. Flow diagram of experimental process.

dard deviations. The raw sEMG and kinematics data were processed by MatLab software (The MathsWorks Inc., Natrick, MA, USA).

The independent variables in this study were two levels of load weights (5 kg and 15 kg), two styles of tasks (lifting task and carrying task), and two interventions (with exoskeleton and without exoskeleton). Three dependent variables were muscle activities for four muscles (lTES, rTES, lLES, and rLES), peak flexion angle and flexion range of thoracic and lumbar segments, and repetitions of lifting. SPSS Statistics Software Version 26 (IBM, USA) was used to do statistical analysis. The normality and the homogeneity of the variance of all data were tested by the Shapiro-Wilk test. The data violated the assumption of normality, therefore, all statistical tests performed were non-parametric. Wilcoxon Signed-Rank test was conducted for each outcome variable separately (Park, 2021). The α value at 0.05 and the level of statistical significance at $p < 0.05$ were set.

3. Results

3.1. Demographic characteristic of the participants

Table 2 showed the demographic characteristic of the 20 male participants [mean age = 24.15 ± 3.20]. The mean body height and weight are 1.75 ± 0.08 m and 69.17 ± 12.89 kg, respectively. The BMI of 20 male participants ranged from 17.07 to 29.80 kg/m^2, with an average of 22.56 ± 3.51 kg/m^2. For exercise frequency, the mean value was 2.40 ± 1.62 times per week.

3.2. Changes in trunk muscle activity

The mean and standard deviation of trunk muscle activities are shown in Table 3. The trunk muscle activities in mean %MVIC ranged from 14.39 ± 6.54 to 48.55 ± 14.98 when the exoskeleton was used, while between 16.52 ± 5.98 and 55.61 ± 14.57 when the exoskeleton was not used. The exoskeleton significantly reduced the muscle activities of all trunk muscles, including left thoracic erector spinae (lTES) ($p < 0.01$), right thoracic erector spinae (rTES) ($p < 0.05$), left lumbar erector spinae (lLES) ($p < 0.01$), and right lumbar erector spinae (rLES) ($p < 0.01$), under two levels of load weights and two levels of tasks (Fig. 8). The trunk muscle activities showed significant reduction in mean % MVIC ranged from 2.14 ± 2.93 to 8.17 ± 7.02 with the exoskeleton ($p < 0.05$). This finding supported the first hypothesis that the PBSE could benefit the users in both lifting and carrying tasks with different load weights by reducing trunk muscle activity.

3.2.1. Changes in trunk muscle activity in consideration of load weights

Based on the beneficial effect of the exoskeleton on trunk muscle activity, we further evaluated the effectiveness of the exoskeleton on trunk muscle activity in consideration of load weights by comparing the reduction of muscle activity in two levels of load weights. With 5 kg load weight, the exoskeleton reduced trunk muscle activities in mean %MVIC of 2.14 ± 2.93 to 7.33 ± 4.12, while 3.01 ± 3.59 to 8.17 ± 7.02 reduction in mean %MVIC was observed with 15 kg load weight. Overall, with the increase in load weight, the effectiveness of the exoskeleton on the trunk muscles increased slightly, but no significant difference was observed ($p > 0.05$) except for rTES with lifting task ($p < 0.05$) (Fig. 9). This finding rejected the second hypothesis that the benefits of the PBSE in reducing trunk muscle activity could increase with the load weight.

3.2.2. Changes in trunk muscle activity in consideration of tasks

The effectiveness of the exoskeleton on two tasks by comparing the reduction of muscle activity in lifting task and carrying task was further evaluated. In lifting tasks, the exoskeleton reduced

Table 2
Demographic data of participants.

	Mean	SD	Min	Max
Age	24.15	3.20	19	30
Height (m)	1.75	0.08	1.64	1.96
Weight (kg)	69.17	12.89	49.2	91.1
Body Mass Index (kg/m^2)	22.56	3.51	17.07	29.80
Exercise Frequency (times/week)	2.40	1.62	1	6

Table 3
Muscle activity of thoracic erector spinae and lumbar erector spinae muscles under the two conditions (with exoskeleton vs without exoskeleton).

Muscle	Task	Load weight	%MVIC with Exoskeleton Mean (SD)	%MVIC without Exoskeleton Mean (SD)	% Changed Mean (SD)	p value
lTES	Lifting	5 kg	25.32(4.95)	32.65 (5.43)	−7.33 (4.12)	**<0.001
		15 kg	41.91(8.83)	50.08 (9.19)	−8.17 (7.02)	**<0.001
	Carrying	5 kg	14.66 (7.26)	17.20 (8.17)	−2.54 (2.10)	**<0.001
		15 kg	31.33 (15.75)	34.50 (15.89)	−3.18 (5.24)	**0.009
rTES	Lifting	5 kg	26.15(6.07)	31.89 (6.54)	−5.74 (3.57)	**<0.001
		15 kg	42.04(11.31)	50.03 (11.79)	−7.98 (4.61)	**<0.001
	Carrying	5 kg	14.39 (6.54)	16.52 (5.98)	−2.14 (2.93)	**0.003
		15 kg	31.43 (16.63)	34.56 (16.62)	−3.12 (6.05)	*0.015
lLES	Lifting	5 kg	36.52(13.30)	41.72 (12.12)	−5.20 (4.91)	**0.001
		15 kg	48.55(14.98)	55.61 (14.57)	−7.05 (3.84)	**<0.001
	Carrying	5 kg	17.89 (8.15)	21.28 (9.58)	−3.40 (4.02)	**0.002
		15 kg	33.22 (14.38)	36.22 (14.34)	−3.01 (3.59)	**0.004
rLES	Lifting	5 kg	33.00(10.37)	39.56 (12.16)	−6.56 (5.90)	**<0.001
		15 kg	44.36(11.60)	50.19 (11.30)	−5.83 (3.57)	**<0.001
	Carrying	5 kg	17.68 (8.87)	21.02 (11.04)	−3.34 (4.49)	**0.002
		15 kg	31.76 (15.75)	35.30 (15.25)	−3.53 (3.70)	**0.002

lTES - left thoracic erector spinae, rTES - right thoracic erector spinae, lLES - left lumbar erector spinae, rLES - right lumbar erector spinae.
*significant at $p < 0.05$, **significant at $p < 0.01$.

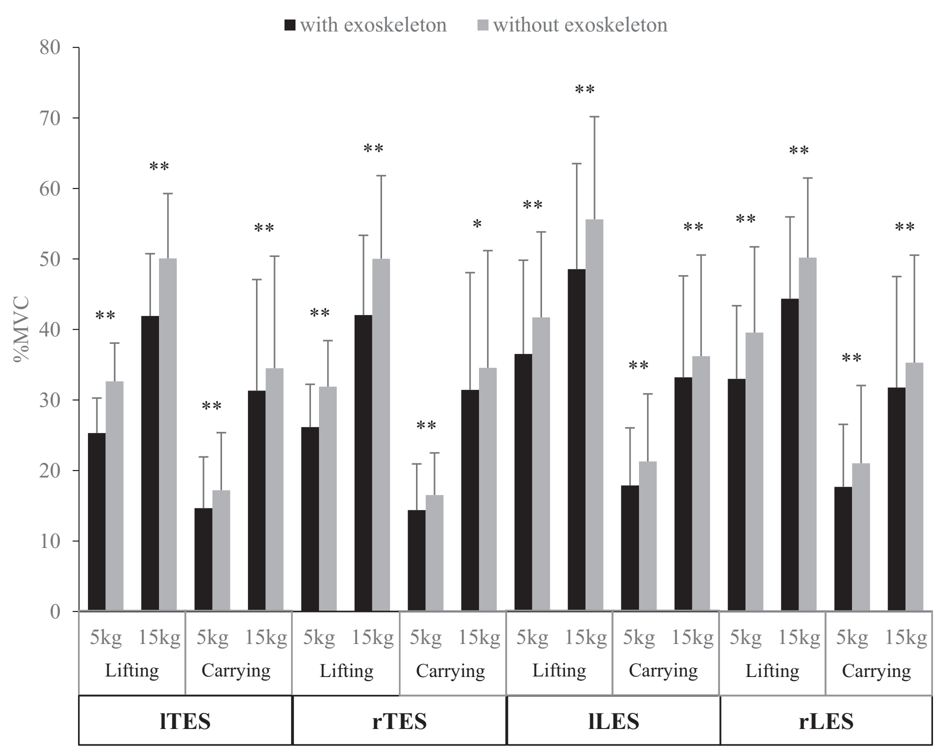

Fig. 8. Muscle activity of thoracic erector spinae and lumbar erector spinae muscles under the two conditions (with exoskeleton vs without exoskeleton). lTES - left thoracic erector spinae, rTES - right thoracic erector spinae, lLES - left lumbar erector spinae, rLES - right lumbar erector spinae. *significant at $p < 0.05$, **significant at $p < 0.01$.

the trunk muscle activities in mean %MVIC of 5.20 ± 4.91 to 8.17 ± 7.02, while decreasing 2.14 ± 2.93 to 3.53 ± 3.70 in carrying task. Significantly more reduction was observed on muscle activities of the TES in the lifting task ($p < 0.05$), while for LES, only a significant difference was observed in the case of lLES with 15 kg load weight ($p < 0.01$) (Fig. 10). This finding supported the third hypothesis that the PBSE could help more in lifting task than carrying task.

3.3. Changes in spine kinematics

For lifting task, the maximal flexion angle of thoracic increased 1.12 ± 10.24 to 4.61 ± 9.44 and minimal angle increased 0.95 ± 11.02 to 4.62 ± 10.58 when the exoskeleton was used (Fig. 11). Significant difference of peak flexion angle of thoracic was observed in condition with 15 kg load weight ($p < 0.05$). Peak flexion angle of lumbar decreased slightly, but no significant difference was observed ($p > 0.05$). As for the flexion range, it decreased in thoracic and lumbar segments when the exoskeleton was used, but no significant difference was observed ($p > 0.05$) (Fig. 12). For carrying task, the peak flexion angle of thoracic increased, but significant difference only observed on minimal flexion angle in condition with 5 kg load weight ($p < 0.05$) (Fig. 13). The peak flexion angle of lumbar decreased, and significant difference only observed on minimal flexion angle in condition with 5 kg load weight ($p < 0.05$). As for flexion range, it decreased in thoracic and lumbar segments when the exoskeleton was used, but no significant differ-

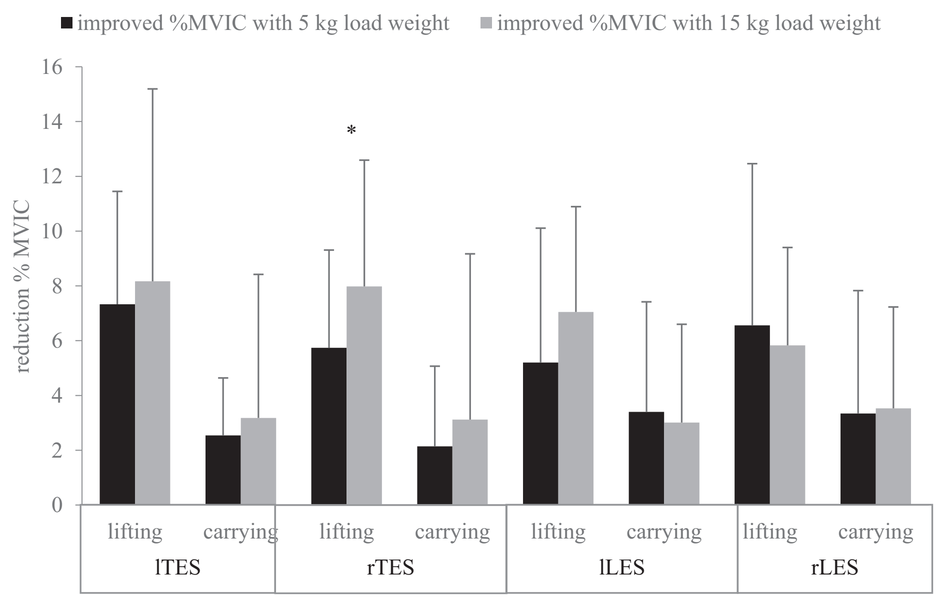

Fig. 9. Reduction of muscle activity on thoracic erector spinae and lumbar erector spinae muscles by exoskeleton in consideration of load weights lTES - left thoracic erector spinae, rTES - right thoracic erector spinae, lLES - left lumbar erector spinae, rLES - right lumbar erector spinae. *significant at $p < 0.05$, **significant at $p < 0.01$.

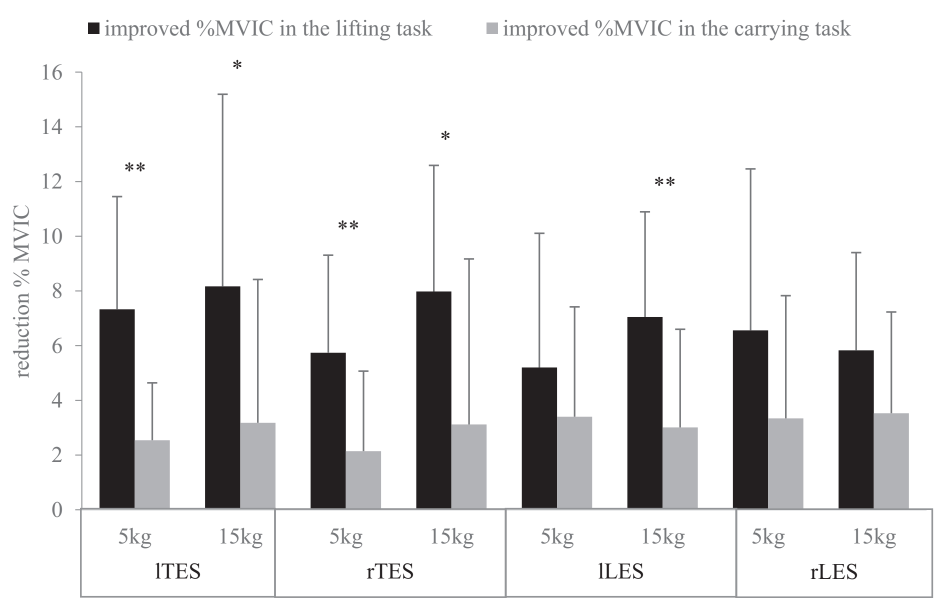

Fig. 10. Reduction of muscle activity on thoracic erector spinae and lumbar erector spinae muscles by exoskeleton in consideration of tasks. lTES - left thoracic erector spinae, rTES - right thoracic erector spinae, lLES - left lumbar erector spinae, rLES - right lumbar erector spinae. *significant at $p < 0.05$, **significant at $p < 0.01$.

ence was observed ($p > 0.05$) (Fig. 14). These findings partially supported the fourth hypothesis that the PBSE could benefit the users in both lifting and carrying tasks with different load weights by decreasing the trunk flexion.

3.4. Changes in physical capacity

For physical capacity, participants can make 2–3 more full cycles in the lifting task when the exoskeleton was used (Table 4). However, there was no significant difference between two conditions (with exoskeleton vs without exoskeleton) ($p > 0.05$) (Fig. 15). This finding rejected the fifth hypothesis that using a PBSE can increase the physical capacity during repetitive lifting task.

4. Discussion

This study aimed to investigate the effects of the PBSE on trunk muscle activation, kinematics, and physical capacity in repetitive lifting task and carrying tasks in consideration of load weights. The results showed that the exoskeleton can effectively reduce trunk muscle activity in repeated lifting and carrying tasks, and the peak flexion angle showed significant improvement in some conditions. These results were basically consistent with the hypotheses of this study. As for physical capacity in the repetitive lifting task, the exoskeleton showed no significant increase by analyzing the number of repetitive lifting. Moreover, the effectiveness of the exoskeleton did not change with the increase in load weight, and the exoskeleton was more effective for lifting tasks compared

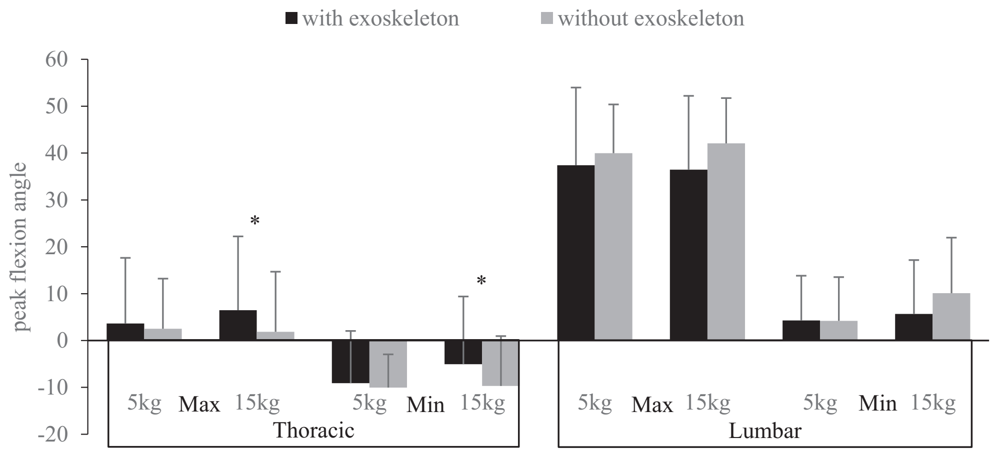

Fig. 11. Kinematics data of peak flexion angle of thoracic and lumbar segments during lifting task under the two conditions (with exoskeleton vs without exoskeleton) *significant at $p < 0.05$, **significant at $p < 0.01$.

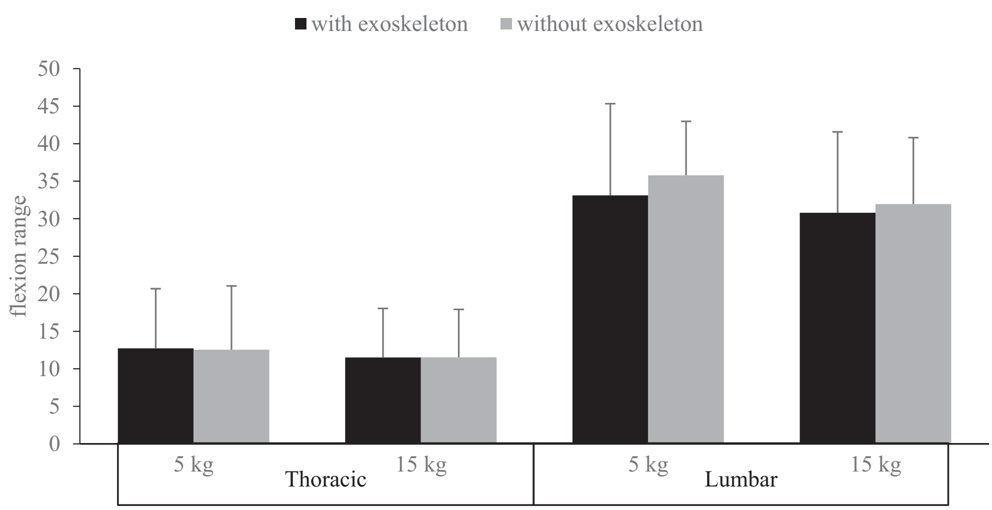

Fig. 12. Kinematics data of flexion range of thoracic and lumbar segments during lifting task under the two conditions (with exoskeleton vs without exoskeleton) *significant at $p < 0.05$, **significant at $p < 0.01$.

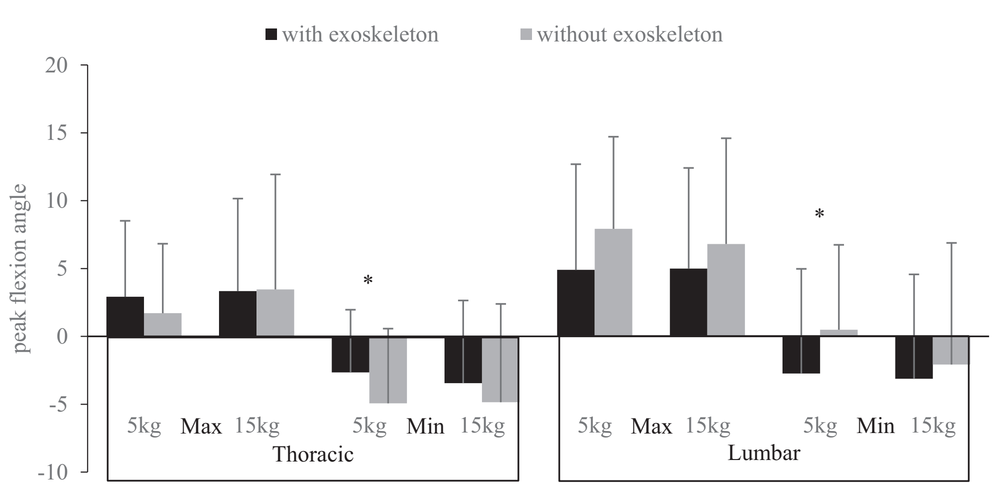

Fig. 13. Kinematics data of peak flexion angle of thoracic and lumbar segments during carrying task under the two conditions (with exoskeleton vs without exoskeleton) *significant at $p < 0.05$, **significant at $p < 0.01$.

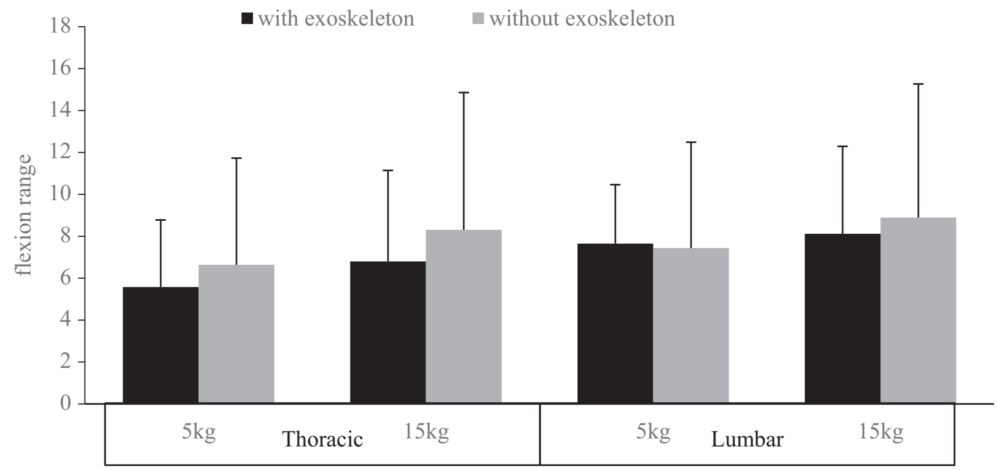

Fig. 14. Kinematics data of flexion range of thoracic and lumbar segments during carrying task under the two conditions (with exoskeleton vs without exoskeleton) *significant at $p < 0.05$, **significant at $p < 0.01$.

Table 4
Physical capacity during lifting task under the two conditions (with exoskeleton vs without exoskeleton).

Load Weight	Intervention	Physical capacity	changed	P value
5 kg	with exoskeleton	39.35 (18.54)	3.40 (14.91)	0.723
	without exoskeleton	35.95 (14.41)		
15 kg	with exoskeleton	22.10 (9.10)	2.45 (8.07)	0.239
	without exoskeleton	19.65 (5.00)		

*significant at $p < 0.05$, **significant at $p < 0.01$.

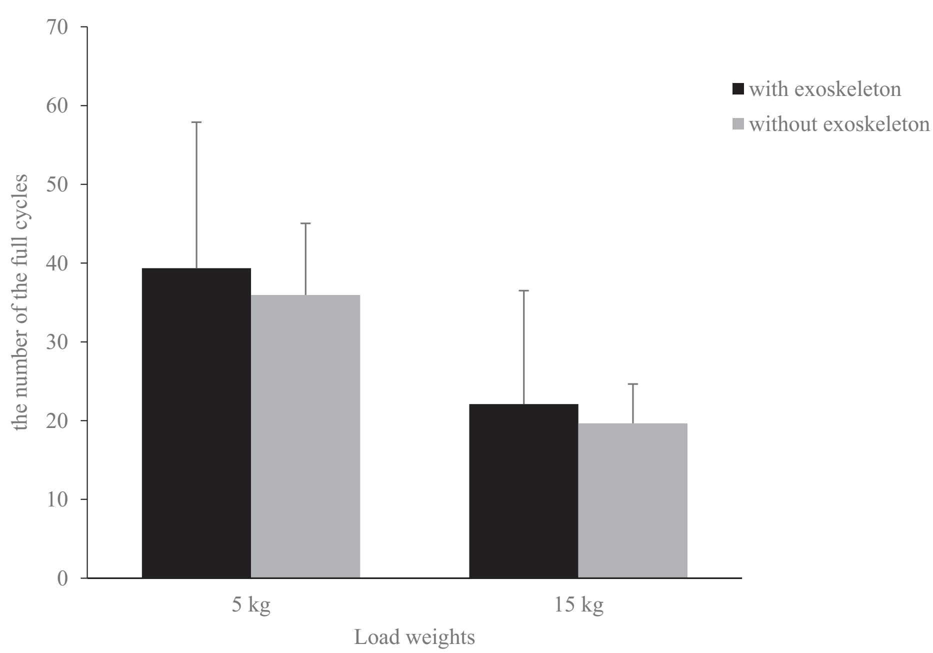

Fig. 15. Physical capacity during lifting task under the two conditions (with exoskeleton vs without exoskeleton) *significant at $p < 0.05$, **significant at $p < 0.01$.

to carrying task. The PBSE not only can effectively reduce trunk muscle activity (presented by %MVIC) and peak flexion angle of the trunk, but also did not affect the physical capacity.

4.1. The effect of exoskeleton on trunk muscle activity

The results showed that the exoskeleton decreased the muscle activity of TES and LES by 5–9% in the repetitive lifting task and 2–4% in the carrying task. Our findings advocated the previous PBSE studies (Alemi et al., 2020; Poliero et al., 2020; von Glinski et al., 2019; Yin, Yang, Wang, & Qu, 2019). This may be caused by similar devices design and principles of the PBSE. The exoskeleton takes advantage of "Artificial Muscles," which are parallel to the erector spinae muscles and could store energy during the bending process and release energy when standing, thereby reducing the demand of the relevant muscles (Abdoli-E, Agnew, & Stevenson, 2006).

The extensor strength related to the amplitude of the electromyogram is the main factor leading to the compression force of the lumbar intervertebral joints (Potvin, Norman, & McGill, 1991). A decrease in paraspinal muscle activity (muscle amplitude) is corresponding to a decrease in the load on the intervertebral disc (Healey, Fowler, Burden, & McEwan, 2005). Since severe and chronic LBP was most often induced by the excessive compression force on lumbar intervertebral disc (Adams, 2004), PBSEs can reduce the load of the intervertebral disc by reducing the activity of the paravertebral muscles. It can help workers reduce the muscle demand of tasks in the first few weeks after returning to work and also effectively reduce the recurrence of LBP and restore health (Abdoli-E et al., 2006). In addition, LBP caused by repetitive lifting or load carrying tasks can be effectively avoided by the PBSE to reduce the damage on low back that would otherwise occur easily. Therefore, the PBSE also helps to protect the low back that has been injured in the past, and thus reduce the occurrence of LBP.

However, it was also noticed that the decrease in trunk muscle activity (about 10%) during repetitive lifting task was lower than the 10–40% decrease in sEMG reported in previous studies (De Looze et al., 2016). This result can be explained in the following ways. First, although previous studies reported a large decline in trunk muscle activity, the working postures utilized were all stoop postures. The stoop posture puts more pressure on the non-contracting connective tissue, resulting in excessive bending of the lumbar spine, and therefore has a higher risk of injury (Kingma, Faber, & Van Dieën, 2010; McGill, Hughson, & Parks, 2000; Straker, 2003). However, this study utilized the squat position that is relatively more protective to the lower back (Bazrgari, Shirazi-Adl, & Arjmand, 2007; van Dieën, Hoozemans, & Toussaint, 1999). This study has proved that the exoskeleton can still effectively reduce the back muscle activity in the protective squatting posture. The results of previous studies on the effectiveness of PBSEs considering posture are inconsistent. Koopman et al. (2020) and Abdoli-E et al. (2006) reported that the effectiveness on reduction of comprehensive electromyography of thoracic and lumbar spine of PBSEs in squatting and bending posture in lifting tasks was similar (Abdoli-E et al., 2006; Koopman et al., 2020), while Alemi et al. (2019) reported that PBSE in the squatting posture can reduce back muscle activity more than the stoop posture (Alemi et al., 2019). This inconsistency should be investigated in the future studies. Secondly, the exoskeleton cannot be flexibly adjusted to fit each participant very well. The exoskeleton utilized in this study has a limited adjustable range, but the basic parameters of the participants are quite different (Table 2). In some cases, the exoskeleton cannot achieve the most desired effect. Similar design drawbacks have also been reported in previous study (Koopman, Toxiri, et al., 2019). Thirdly, the participants in this study moved at the same speed regardless of interventions. Previous studies reported that the effectiveness of PBSEs on reducing muscle activity in lifting tasks may be affected by slower movement speed (Alemi et al., 2019; Koopman, Toxiri, et al., 2019). Compared with the normal lifting movements, a significantly higher peak torque for the fast lifting movements was reported (Kingma et al., 2001), and as the trunk flexion angular acceleration increased, the muscle activity decreased, among which the erector spinae muscles decrease the most (Marras & Mirka, 1993). However, in this study, participants were asked to perform weightlifting tasks at 10 full cycles per minute in each condition.

4.1.1. Effectiveness of exoskeleton on sEMG in consideration of load weights

In order to evaluate the effectiveness of the exoskeleton on sEMG in consideration of load weights, we compared the reduction of muscle activity with the load of 5 kg and with 15 kg respectively. Only significant difference on rTES in repetitive lifting task was observed. The heavier the load, the more improvement of the exoskeleton on rTES ($p = 0.030$). As for other muscles and carrying tasks, although the exoskeleton reduced the muscle activity more as weight increases, the difference was not significant. Based on the literature review, no research explored this point, leaving a research gap.

4.1.2. Effectiveness of exoskeleton on sEMG in consideration of tasks

In order to evaluate the improvement of the exoskeleton on task, we compared the reduction of muscle activity in lifting and carrying task. This study found that the exoskeleton was more effective in lifting task in comparison with carrying task. In fact, few studies directly compared the effect of exoskeleton in different types of tasks. Based on the literature review, one study on active PBSE reported similar results (Poliero et al., 2020). This may be because the support of the PBSE relies on the trunk flexion to generate power, and the trunk flexion in the carrying task was less than that of the lifting task (Baltrusch, Van Dieën, Van Bennekom, & Houdijk, 2018).

4.2. The effect of exoskeleton on kinematics

According to the kinematics results, the flexion angles with the exoskeleton at the thoracic spine were greater than the angle without the exoskeleton, with an increase of 1 to 5 degrees, although most of the differences were not statistically significant. These findings were basically similar with the conclusions of some previous experiments that PBSEs could benefit in reducing the situation of people working with an improper flexion angle (Kim, Madinei, Alemi, Srinivasan, & Nussbaum, 2020; Koopman, Toxiri, et al., 2019; Sadler, Graham, & Stevenson, 2011; Simon, Alemi, & Asbeck, 2021; Ulrey & Fathallah, 2013). Given a greater flexion angle at the thoracic spine, the exoskeleton can provide a greater assistive force using the actuator (i.e., McKibben artificial muscles). As a result, thoracic muscle sEMGs are much smaller with the device than those without the device in both tasks.

Previous studies have pointed out that it may take more than two days for the participants to adapt to PBSE (Alemi et al., 2019; Gordon & Ferris, 2007). During our experiment, we did not give the participants enough time to adapt to the exoskeleton, which may affect the results of motion. The effect of the exoskeleton also depends on the body height and body weight of the user as well as the degree of relaxation of the participants (Simon et al., 2021). However, participants in our study have a large variation in body shape, in some cases, it was observed that the exoskeleton was prone to slip relative to the limb during the task, which was also reported in previous studies (Accoto et al., 2014; Neckel, Wisman, & Hidler, 2006). Also, according to the feedbacks of participants, the design of the PBSEs cannot avoid this slippage, and this is consistent with the findings in the previous study (Alemi et al., 2020), which may reduce the supportive effect of the exoskeleton. Some PBSE designs use a flexible beam to allow the trunk to move freely (Näf et al., 2018), which may affect the results. However, for 'Muscle Suit Every,' only limited misalignment compensating mechanisms could be met due to the existence of flexible beams and the lack of flex slider (which is a flexible back structure that allows the flexible beam on the back to be further elongated). These also cause discomfort during the use. According to the feedbacks of participants, walking seems more difficult because of the restriction from thigh structure by the frame and pads of the exoskeleton. This is not surprising, because in a passive device, the users have to work against the device while pushing the leg forward. In addition, unlike bending over to lift loads, the peak lumbar flexion angle reached by squat lifting is smaller and therefore avoids the risk of working at limit angle.

By limiting the range of lumbar flexion, the PBSEs could reduce the risk of low back disorder due to working at limit angles. Because the smaller the flexion angles of the lumbar, the compression on the anterior portion of the lumbar vertebral discs is less (Adams & Hutton, 1982, 1986; Adams et al., 1994). However, the prolonged flexion of the trunk could increase the laxity of the non-contractile connective tissues, thereby inducing the low back injury (Ulrey & Fathallah, 2013). By limiting the trunk flexion, the compression on the lumbar intervertebral discs and the amount of stretching in the posterior ligaments of the spine could be decreased (Ulrey & Fathallah, 2013). Furthermore, the laxity of the passive tissues could also increase with the cumulative mechanical low back load throughout the work, thereby increasing the range of motion of the trunk (Adams & Dolan, 1996; Adams, Dolan, & Hutton, 1987; Coenen et al., 2014). From this point of view, the PBSEs could prevent people from working at the limit flexion angle by limiting the motion of trunk during the lifting (Bonato et al., 2003; Ulrey & Fathallah, 2013).

4.3. The effect of passive exoskeleton on physical capacity

In repetitive lifting task, this study found that the exoskeleton did not significantly affect the number of the full cycles regardless of the load weight, which showed that the exoskeleton may not increase the physical activity of the participants. This finding was consistent with a previous study that used metabolic demand as the indicator of physical capacity (Whitfield, Costigan, Stevenson, & Smallman, 2014). But considering the weight of the exoskeleton itself (4.3 kg and 4.4 kg), it may suggest that the weight of the exoskeleton itself may not bring additional burdens to the users. Although one study that also utilized the number of repetitions as the outcome measure reported significant increase with another model of a PBSE (Miura et al., 2018), the task in that study was carried out with a stoop posture. The squat technique requires more metabolic energy than the stoop technique (Baltrusch et al., 2019; Whitfield et al., 2014), which may be the potential cause of inconsistent results.

4.4. Practical implications

This study focused on the repetitive lifting and load carrying tasks that cause the risk of LBP, and found some favorable effects of the exoskeleton on trunk muscle activity and trunk flexion angle. Moreover, unlike the studies based on the stoop posture (Ulrey & Fathallah, 2013; Yin et al., 2019), this experiment proved that the PBSE can further reduce the load on the back with the protective squat posture. In addition, we included the outcome measure of the number of the full cycles in repetitive lifting task, which further proved that the exoskeleton, as the device with certain weight and volume, may not affect the flexibility and physical ability of the users. This study provides new supportive evidence for the application of PBSE on reducing the risk of injury in the low back for people executing highly demanding tasks, such as repetitive lifting and load carrying. Since the PBSE helped more in lifting task than carrying task, this study especially recommended the application of this exoskeleton to repetitive lifting tasks. However, during the experiment, it was observed that in some cases the exoskeleton could not match the participants properly. Therefore, the PBSE may need to be designed more adjustable. Also, it was observed that the exoskeleton was prone to slip relative to the limb during the task. It suggested that a flex slider may be added to the exoskeleton to avoid this problem. Some participants reported that the exoskeleton itself was so heavy that wearing the exoskeleton increased subjective fatigue. Thus, the exoskeleton can be optimized by using light-weight and strong materials. Besides, sufficient time should be given to the users to become familiar with the exoskeleton before practical implication.

4.5. Limitations of present study

There were some limitations in this study. First, only the trunk muscles were measured, and we did not check for additional loading in other joints. Previous studies found that the reduced load on the low back can be transferred to the gluteal muscles (Vleeming et al., 1995). Also, it was found that PBSEs reduce erector spinae muscle activity while increasing leg muscle activity (Barrett, 2001). Therefore, it is unknown whether the exoskeleton may transfer the reduced load of the low back to the lower limbs and increase the risk of injury of lower limbs. Secondly, only biomechanical measures were involved in this study, but the practical application of the exoskeleton should take into account the impact on functional tasks and subjective acceptance of the user. Some studies with functional measures have reported that PBSEs were not conducive to walking, ladder climbing, and other functional tasks (Baltrusch et al., 2018; Kozinc, Baltrusch, Houdijk, & Šarabon, 2021). Taking into account the satisfaction and wearing comfort of user on the exoskeleton, subjective indicators may need to be included in the outcome measures to explore the acceptance of users on the exoskeleton. Lastly, this is a cross-sectional study, ideally, randomized controlled trials can provide strong evidence.

5. Conclusion

This study successfully explored the biomechanical effects of the PBSE in repetitive lifting task and carrying task in consideration of load weights. The biomechanical measures such as muscle activity demonstrated that the exoskeleton decreased trunk muscle activity significantly during repetitive lifting task and carrying task, implying potential effects of reducing back load in both the thoracic and lumbar regions. The kinematics parameters also showed some improvement in peak flexion angle while the flexion range did not change, indicating that the exoskeleton contributed to reducing the situation of people working on the limited flexion angle, thereby reducing the risk of injury without affecting flexibility. Meanwhile, this study further explored the effectiveness of the exoskeleton, and results showed that the effectiveness did not change with the increase in load weight, and the exoskeleton was more beneficial to lifting compared to carrying. Moreover, physical capacity did not show significant difference, implying that the weight of the exoskeleton did not affect the physical capacity of the user. These results support the positive benefits of the exoskeleton in high-risk tasks that cause LBP. Future research is recommended to involve more joints and subjective measures and to comprehensively determine the PBSE effects.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdoli-E, M., Agnew, M. J., & Stevenson, J. M. (2006). An on-body personal lift augmentation device (PLAD) reduces EMG amplitude of erector spinae during lifting tasks. *Clinical Biomechanics*, *21*(5), 456–465. https://doi.org/10.1016/j.clinbiomech.2005.12.021.

Accoto, D., Sergi, F., Tagliamonte, N. L., Carpino, G., Sudano, A., & Guglielmelli, E. (2014). Robomorphism: A nonanthropomorphic wearable robot. *IEEE Robotics & Automation Magazine*, *21*(4), 45–55. https://doi.org/10.1109/MRA.2014.2360276.

Adams, M., & Dolan, P. (1996). Time-dependent changes in the lumbar spine's resistancc to bending. *Clinical Biomechanics, 11*(4), 194–200. https://doi.org/10.1016/0268-0033(96)00002-2.

Adams, M., Dolan, P., & Hutton, W. (1987). Diurnal variations in the stresses on the lumbar spine. *Spine, 12*(2), 130–137. https://doi.org/10.1097/00007632-198703000-00008.

Adams, M., & Hutton, W. (1982). Prolapsed intervertebral disc: A hyperflexion injury. *Spine, 7*(3), 184–191. https://doi.org/10.1097/00007632-198205000-00002.

Adams, M., & Hutton, W. (1986). The effect of posture on diffusion into lumbar intervertebral discs. *Journal of Anatomy, 147*, 121.

Adams, M., McNally, D., Chinn, H., & Dolan, P. (1994). The clinical biomechanics award paper 1993 posture and the compressive strength of the lumbar spine. *Clinical Biomechanics, 9*(1), 5–14.

Adams, M. A. (2004). Biomechanics of back pain. *Acupuncture in Medicine, 22*(4), 178–188. https://doi.org/10.1136/aim.22.4.178.

Alemi, M. M., Geissinger, J., Simon, A. A., Chang, S. E., & Asbeck, A. T. (2019). A passive exoskeleton reduces peak and mean EMG during symmetric and asymmetric lifting. *Journal of Electromyography and Kinesiology, 47*, 25–34. https://doi.org/10.1016/j.jelekin.2019.05.003.

Alemi, M. M., Madinei, S., Kim, S., Srinivasan, D., & Nussbaum, M. A. (2020). Effects of two passive back-support exoskeletons on muscle activity, energy expenditure, and subjective assessments during repetitive lifting. *Human factors, 62*(3), 458–474. https://doi.org/10.1177/0018720819897669.

Amandels, S., Het Eyndt, H. O., Daenen, L., & Hermans, V. (2018). Introduction and testing of a passive exoskeleton in an industrial working environment. *Paper presented at the Congress of the International Ergonomics Association*.

Baltrusch, S., Van Dieën, J., Bruijn, S., Koopman, A., Van Bennekom, C., & Houdijk, H. (2019). The effect of a passive trunk exoskeleton on metabolic costs during lifting and walking. *Ergonomics*. https://doi.org/10.1080/00140139.2019.1602288.

Baltrusch, S., Van Dieën, J., Van Bennekom, C., & Houdijk, H. (2018). The effect of a passive trunk exoskeleton on functional performance in healthy individuals. *Applied ergonomics, 72*, 94–106. https://doi.org/10.1016/j.apergo.2018.04.007.

Barrett, A. L. (2001). Evaluation of four weight transfer devices for reducing loads on lower back during agricultural stoop labor. In *Paper presented at the ASAE Annual International Meeting*, 2001.

Bazrgari, B., Shirazi-Adl, A., & Arjmand, N. (2007). Analysis of squat and stoop dynamic liftings: Muscle forces and internal spinal loads. *European Spine Journal, 16*(5), 687–699. https://doi.org/10.1007/s00586-006-0240-7.

Bonato, P., Ebenbichler, G. R., Roy, S. H., Lehr, S., Posch, M., Kollmitzer, J., & Della Croce, U. (2003). Muscle fatigue and fatigue-related biomechanical changes during a cyclic lifting task. *Spine, 28*(16), 1810–1820. https://doi.org/10.1097/01.BRS.0000087500.70575.45.

Bosch, T., van Eck, J., Knitel, K., & de Looze, M. (2016). The effects of a passive exoskeleton on muscle activity, discomfort and endurance time in forward bending work. *Applied Ergonomics, 54*, 212–217. https://doi.org/10.1016/j.apergo.2015.12.003.

Coenen, P., Kingma, I., Boot, C. R., Bongers, P. M., & van Dieën, J. H. (2014). Cumulative mechanical low-back load at work is a determinant of low-back pain. *Occupational and Environmental Medicine, 71*(5), 332–337. https://doi.org/10.1136/oemed-2013-101862.

Davis, K. G., & Kotowski, S. E. (2015). Prevalence of musculoskeletal disorders for nurses in hospitals, long-term care facilities, and home health care: A comprehensive review. *Human Factors, 57*(5), 754–792.

De Looze, M. P., Bosch, T., Krause, F., Stadler, K. S., & O'Sullivan, L. W. (2016). Exoskeletons for industrial application and their potential effects on physical work load. *Ergonomics, 59*(5), 671–681. https://doi.org/10.1080/00140139.2015.1081988.

Demoulin, C., Huijnen, I. P., Somville, P.-R., Grosdent, S., Salamun, I., Crielaard, J.-M., ... Volders, S. (2013). Relationship between different measures of pain-related fear and physical capacity of the spine in patients with chronic low back pain. *The Spine Journal, 13*(9), 1039–1047.

Ferguson, S. A., Merryweather, A., Thiese, M. S., Hegmann, K. T., Lu, M.-L., Kapellusch, J. M., & Marras, W. S. (2019). Prevalence of low back pain, seeking medical care, and lost time due to low back pain among manual material handling workers in the United States. *BMC Musculoskeletal Disorders, 20*(1), 1–8.

Freburger, J. K., Holmes, G. M., Agans, R. P., Jackman, A. M., Darter, J. D., Wallace, A. S., ... Carey, T. S. (2009). The rising prevalence of chronic low back pain. *Archives of Internal Medicine, 169*(3), 251–258.

Ganasegeran, K., Perianayagam, W., Nagaraj, P., & Al-Dubai, S. (2014). Psycho-behavioural risks of low back pain in railway workers. *Occupational Medicine, 64*(5), 372–375.

Gordon, K. E., & Ferris, D. P. (2007). Learning to walk with a robotic ankle exoskeleton. *Journal of Biomechanics, 40*(12), 2636–2644. https://doi.org/10.1016/j.jbiomech.2006.12.006.

Healey, E. L., Fowler, N. E., Burden, A. M., & McEwan, I. M. (2005). Raised paraspinal muscle activity reduces rate of stature recovery after loaded exercise in individuals with chronic low back pain. *Archives of Physical Medicine and Rehabilitation, 86*(4), 710–715. https://doi.org/10.1016/j.apmr.2004.10.026.

Hemming, R., Sheeran, L., van Deursen, R., & Sparkes, V. (2018). Non-specific chronic low back pain: Differences in spinal kinematics in subgroups during functional tasks. *European Spine Journal, 27*(1), 163–170.

Hermens, H., & Freriks, B. (1997). *SENIAM: The state of the art on sensors and sensor placement procedures for surface electromyography, a proposal for sensor placement procedures*. Deliverable of the SENIAM Project.

Heydari, H., Hoviattalab, M., Azghani, M. R., Ramezanzadehkoldeh, M., & Parnianpour, M. (2013). Investigation on a developed wearable assistive device (WAD) in reduction lumbar muscles activity. *Biomedical Engineering: Applications, Basis and Communications, 25*(03), 1350035.

Hoogendoorn, W. E., Bongers, P. M., De Vet, H. C., Douwes, M., Koes, B. W., Miedema, M. C., ... Bouter, L. M. (2000). Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: Results of a prospective cohort study. *Spine, 25*(23), 3087–3092. https://doi.org/10.1097/00007632-200012010-00018.

Hoy, D., Brooks, P., Blyth, F., & Buchbinder, R. (2010). The epidemiology of low back pain. *Best Practice & Research Clinical Rheumatology, 24*(6), 769–781. https://doi.org/10.1016/j.berh.2010.10.002.

Huysamen, K., de Looze, M., Bosch, T., Ortiz, J., Toxiri, S., & O'Sullivan, L. W. (2018). Assessment of an active industrial exoskeleton to aid dynamic lifting and lowering manual handling tasks. *Applied Ergonomics, 68*, 125–131.

Huysamen, K., Power, V., & O'Sullivan, L. (2020). Kinematic and kinetic functional requirements for industrial exoskeletons for lifting tasks and overhead lifting. *Ergonomics, 63*(7), 818–830. https://doi.org/10.1080/00140139.2020.1759698.

Jakobsson, M., Gutke, A., Mokkink, L. B., Smeets, R., & Lundberg, M. (2019). Level of evidence for reliability, validity, and responsiveness of physical capacity tasks designed to assess functioning in patients with low back pain: A systematic review using the COSMIN standards. *Physical Therapy, 99*(4), 457–477.

Kim, S., Madinei, S., Alemi, M. M., Srinivasan, D., & Nussbaum, M. A. (2020). Assessing the potential for "undesired" effects of passive back-support exoskeleton use during a simulated manual assembly task: Muscle activity, posture, balance, discomfort, and usability. *Applied Ergonomics, 89* 103194.

Kingma, I., Baten, C. T., Dolan, P., Toussaint, H. M., van Dieën, J. H., de Looze, M. P., & Adams, M. A. (2001). Lumbar loading during lifting: A comparative study of three measurement techniques. *Journal of Electromyography and Kinesiology, 11*(5), 337–345. https://doi.org/10.1016/S1050-6411(01)00011-6.

Kingma, I., Faber, G. S., & Van Dieën, J. H. (2010). How to lift a box that is too large to fit between the knees. *Ergonomics, 53*(10), 1228–1238. https://doi.org/10.1080/00140139.2010.512983.

Koes, B., Van Tulder, M., & Thomas, S. (2006). Diagnosis and treatment of low back pain. *Bmj, 332*(7555), 1430–1434. https://doi.org/10.1136/bmj.332.7555.1430.

Koopman, A. S., Kingma, I., Faber, G. S., de Looze, M. P., & van Dieën, J. H. (2019). Effects of a passive exoskeleton on the mechanical loading of the low back in static holding tasks. *Journal of Biomechanics, 83*, 97–103. https://doi.org/10.1016/j.jbiomech.2018.11.033.

Koopman, A. S., Näf, M., Baltrusch, S. J., Kingma, I., Rodriguez-Guerrero, C., Babič, J., ... van Dieën, J. H. (2020). Biomechanical evaluation of a new passive back support exoskeleton. *Journal of Biomechanics, 105* 109795.

Koopman, A. S., Toxiri, S., Power, V., Kingma, I., van Dieën, J. H., Ortiz, J., & de Looze, M. P. (2019). The effect of control strategies for an active back-support exoskeleton on spine loading and kinematics during lifting. *Journal of Biomechanics, 91*, 14–22. https://doi.org/10.1016/j.jbiomech.2019.04.044.

Kozinc, Ž., Baltrusch, S., Houdijk, H., & Šarabon, N. (2021). Short-term effects of a passive spinal exoskeleton on functional performance, discomfort and user satisfaction in patients with low back pain. *Journal of Occupational Rehabilitation, 31*(1), 142–152. https://doi.org/10.1007/s10926-020-09899-7.

Kudo, N., Yamada, Y., & Ito, D. (2019). Age-related injury risk curves for the lumbar spine for use in low-back-pain prevention in manual handling tasks. *Robomech Journal, 6*(1), 1–10.

Legg, S. J., & Myles, W. (1981). Maximum acceptable repetitive lifting workloads for an 8-hour work-day using psychophysical and subjective rating methods. *Ergonomics, 24*(12), 907–916. https://doi.org/10.1080/00140138108924913.

Leinonen, V., Kankaanpää, M., Airaksinen, O., & Hänninen, O. (2000). Back and hip extensor activities during trunk flexion/extension: Effects of low back pain and rehabilitation. *Archives of Physical Medicine and Rehabilitation, 81*(1), 32–37. https://doi.org/10.1016/S0003-9993(00)90218-1.

Marras, W., & Mirka, G. (1993). Electromyographic studies of the lumbar trunk musculature during the generation of low-level trunk acceleration. *Journal of Orthopaedic Research, 11*(6), 811–817. https://doi.org/10.1002/jor.1100110606.

McGill, S. M., Hughson, R. L., & Parks, K. (2000). Changes in lumbar lordosis modify the role of the extensor muscles. *Clinical Biomechanics, 15*(10), 777–780. https://doi.org/10.1016/S0268-0033(00)00037-1.

Miura, K., Kadone, H., Koda, M., Abe, T., Kumagai, H., Nagashima, K., ... Funayama, T. (2018). The hybrid assistive limb (HAL) for Care Support successfully reduced lumbar load in repetitive lifting movements. *Journal of Clinical Neuroscience, 53*, 276–279. https://doi.org/10.1016/j.jocn.2018.04.057.

Motmans, R., Debaets, T., & Chrispeels, S. (2018). Effect of a passive exoskeleton on muscle activity and posture during order picking. *Paper presented at the congress of the international ergonomics association*.

Näf, M. B., Koopman, A. S., Baltrusch, S., Rodriguez-Guerrero, C., Vanderborght, B., & Lefeber, D. (2018). Passive back support exoskeleton improves range of motion using flexible beams. *Frontiers in Robotics and AI, 5*, 72. https://doi.org/10.3389/frobt.2018.00072.

Neckel, N., Wisman, W., & Hidler, J. (2006). Limb alignment and kinematics inside a Lokomat robotic orthosis. *Paper presented at the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society*.

Nourollahi, M., Afshari, D., & Dianat, I. (2018). Awkward trunk postures and their relationship with low back pain in hospital nurses. *Work, 59*(3), 317–323.

Papi, E., Bull, A. M., & McGregor, A. H. (2018). Is there evidence to use kinematic/kinetic measures clinically in low back pain patients? A systematic review. *Clinical Biomechanics, 55*, 53–64.

Park, C.-H. (2021). Changes in serratus anterior muscle activity according to sling angle in participants with shoulder instability. *The Journal of Korean Academy of Orthopedic Manual Physical Therapy, 27*(1), 13–20.

Poliero, T., Lazzaroni, M., Toxiri, S., Di Natali, C., Caldwell, D. G., & Ortiz, J. (2020). Applicability of an active back-support exoskeleton to carrying activities. *Frontiers in Robotics and AI, 7*, 157. https://doi.org/10.3389/frobt.2020.579963.

Potvin, J., Norman, R., & McGill, S. (1991). Reduction in anterior shear forces on the L4L5 disc by the lumbar musculature. *Clinical Biomechanics, 6*(2), 88–96.

Ramond-Roquin, A., Bodin, J., Serazin, C., Parot-Schinkel, E., Ha, C., Richard, I., ... Roquelaure, Y. (2015). Biomechanical constraints remain major risk factors for low back pain. Results from a prospective cohort study in French male employees. *The Spine Journal, 15*(4), 559–569.

Rasmussen, C. D., Jørgensen, M. B., Clausen, T., Andersen, L. L., Strøyer, J., & Holtermann, A. (2013). Does self-assessed physical capacity predict development of low back pain among health care workers? A 2-year follow-up study. *Spine, 38*(3), 272–276.

Sadler, E. M., Graham, R. B., & Stevenson, J. M. (2011). The personal lift-assist device and lifting technique: A principal component analysis. *Ergonomics, 54*(4), 392–402. https://doi.org/10.1080/00140139.2011.556259.

Simon, A. A., Alemi, M. M., & Asbeck, A. T. (2021). Kinematic effects of a passive lift assistive exoskeleton. *Journal of Biomechanics, 120*. https://doi.org/10.1016/j.jbiomech.2021.110317 110317.

Sterud, T., & Tynes, T. (2013). Work-related psychosocial and mechanical risk factors for low back pain: A 3-year follow-up study of the general working population in Norway. *Occupational and Environmental Medicine, 70*(5), 296–302.

Straker, L. (2003). Evidence to support using squat, semi-squat and stoop techniques to lift low-lying objects. *International Journal of Industrial Ergonomics, 31*(3), 149–160. https://doi.org/10.1016/S0169-8141(02)00191-9.

Theurel, J., Desbrosses, K., Roux, T., & Savescu, A. (2018). Physiological consequences of using an upper limb exoskeleton during manual handling tasks. *Applied Ergonomics, 67*, 211–217. https://doi.org/10.1016/j.apergo.2017.10.008.

Ulrey, B. L., & Fathallah, F. A. (2013). Effect of a personal weight transfer device on muscle activities and joint flexions in the stooped posture. *Journal of Electromyography and Kinesiology, 23*(1), 195–205. https://doi.org/10.1016/j.jelekin.2012.08.014.

van Dieën, J. H., Hoozemans, M. J., & Toussaint, H. M. (1999). Stoop or squat: A review of biomechanical studies on lifting technique. *Clinical Biomechanics, 14*(10), 685–696. https://doi.org/10.1016/S0268-0033(99)00031-5.

Vleeming, A., Pool-Goudzwaard, A., Stoeckart, R., van Wingerden, J.-P., & Snijders, C. (1995). The posterior layer of the thoracolumbar fascia. *Spine, 20*(7), 753–758. https://doi.org/10.1097/00007632-199504000-00001.

von Glinski, A., Yilmaz, E., Mrotzek, S., Marek, E., Jettkant, B., Brinkemper, A., ... Geßmann, J. (2019). Effectiveness of an on-body lifting aid (HAL® for care support) to reduce lower back muscle activity during repetitive lifting tasks. *Journal of Clinical Neuroscience, 63*, 249–255. https://doi.org/10.1016/j.jocn.2019.01.038.

Wai, E. K., Roffey, D. M., Bishop, P., Kwon, B. K., & Dagenais, S. (2010). Causal assessment of occupational carrying and low back pain: Results of a systematic review. *The Spine Journal, 10*(7), 628–638. https://doi.org/10.1016/j.spinee.2010.03.027.

Wang, D., Dai, F., & Ning, X. (2015). Risk assessment of work-related musculoskeletal disorders in construction: State-of-the-art review. *Journal of Construction Engineering and Management, 141*(6), 04015008.

Wang, S. Y., Liu, L. C., Lu, M. C., & Koo, M. (2015). Comparisons of musculoskeletal disorders among ten different medical professions in Taiwan: A nationwide, population-based study. *PloS One, 10*(4) e0123750.

Whitfield, B. H., Costigan, P. A., Stevenson, J. M., & Smallman, C. L. (2014). Effect of an on-body ergonomic aid on oxygen consumption during a repetitive lifting task. *International Journal of Industrial Ergonomics, 44*(1), 39–44. https://doi.org/10.1016/j.ergon.2013.10.002.

Yin, P., Yang, L., Wang, C., & Qu, S. (2019). Effects of wearable power assist device on low back fatigue during repetitive lifting tasks. *Clinical Biomechanics, 70*, 59–65. https://doi.org/10.1016/j.clinbiomech.2019.07.023.

Billy Chun Lung So is the Assistant Professor (Physiotherapy) in the Department of Rehabilitation (RS) at the Hong Kong Polytechnic University. Prior to joining the Hong Kong Polytechnic University, he was appointed as a Lecturer (2011–2012) and Adjunct Assistant Professor (2012–2015) at Department of Orthopaedics and Traumatology of CUHK. Billy has worked as a physiotherapist for over 18 years and gained solid clinical experience in different clinical settings. He is an active educator at different health-related areas and he has extensive teaching experience at the community and tertiary levels. He is the recipient of the RS Outstanding Teaching Award 2016/17 and FHSS Teaching Prize 2016/17 of PolyU. He has been appointed as a Guest Lecturer of Occupational Safety and Health Council (OSHC), Lecturer of Physical Fitness Association of Hong Kong, China (HKPFA) and Hong Kong Sport Institute (HKSI). He is serving in different professional organizations. He is an Executive Committee Member of the Hong Kong Physiotherapy Association, Associate Editor of Hong Kong Physiotherapy Journal, the Vice President of Hong Kong Ergonomics Society, a Specialist of Hong Kong Council for Accreditation of Academic and Vocational Qualifications. His research interest primarily focuses on applying ergonomic principle in preventing work-related musculoskeletal disorders. His PhD thesis was awarded the Best Occupational Safety and Health Project Award 2011 by Occupational Safety and Health Council. He was also awarded the Canadian Institutes of Health Research Scholarship for Post-Doctoral training on the Work Disability Prevention CIHR Strategic Training Program at University of Toronto (2013-2015). He has published papers in international journals and he is serving as a member of Work journal editorial board.

Chunzhuo Hua obtained her M.Sc. in Rehabilitation Sciences. Her research interest is rehabilitation intervention.

Tingting Chen obtained her M.Sc. in Rehabilitation Sciences. Her research interest is rehabilitation intervention.

Qingwen Gao obtained her M.Sc. in Rehabilitation Sciences. Her research interest is rehabilitation intervention.

Siu Shing Man obtained his B.Sc. in Industrial Engineering and Engineering Management from the City University of Hong Kong. He was then obtained his Ph.D. from the same university. His research focus is occupational safety and health, industrial design and ergonomics.

Journal of Safety Research 83 (2022) 139–151

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Can non-crash naturalistic driving data be an alternative to crash data for use in virtual assessment of the safety performance of automated emergency braking systems?

Pierluigi Olleja [a,*], Jonas Bärgman [a], Nils Lubbe [b]

[a] Division of Vehicle Safety at the Department of Mechanics and Maritime Sciences, Chalmers University of Technology, 412 96 Göteborg, Sweden
[b] Autoliv Research, Wallentinsvägen 22, 447 83 Vårgårda, Sweden

ARTICLE INFO

Article history:
Received 28 July 2021
Received in revised form 1 April 2022
Accepted 17 August 2022
Available online 26 August 2022

Keywords:
GIDAS
highD
SHRP2
AEB
Crash avoidance

ABSTRACT

Introduction: Developers of in-vehicle safety systems need to have data allowing them to identify traffic safety issues and to estimate the benefit of the systems in the region where it is to be used, before they are deployed on-road. Developers typically want in-depth crash data. However, such data are often not available. There is a need to identify and validate complementary data sources that can complement in-depth crash data, such as Naturalistic Driving Data (NDD). However, few crashes are found in such data. This paper investigates how rear-end crashes that are artificially generated from two different sources of non-crash NDD (highD and SHRP2) compare to rear-end in-depth crash data (GIDAS). *Method:* Crash characteristics and the performance of two conceptual automated emergency braking (AEB) systems were obtained through virtual simulations – simulating the time-series crash data from each data source. *Results:* Results show substantial differences in the estimated impact speeds between the artificially generated crashes based on both sources of NDD, and the in-depth crash data; both with and without AEB systems. Scenario types also differed substantially, where the NDD have many fewer scenarios where the following-vehicle is not following the lead vehicle, but instead catches-up at high speed. However, crashes based on NDD near-crashes show similar pre-crash criticality (time-to-collision) to in-depth crash data. *Conclusions:* If crashes based on near-crashes are to be used in the design and assessment of preventive safety systems, it has to be done with great care, and crashes created purely from small amounts of everyday driving NDD are not of much use in such assessment. *Practical applications:* Researchers and developers of in-vehicle safety systems can use the results from this study: (a) when deciding which data to use for virtual safety assessment of such systems, and (b) to understand the limitations of NDD.

1. Introduction

Traffic crashes are the eighth leading cause of death worldwide; every year 1.35 million people lose their lives in traffic crashes (WHO, 2018). Fatal rear-end crashes accounted for 7.2% of all fatal crashes in the United States (U.S.) in 2017 (NHTSA, 2019).

Safety measures have for many years been developed to address traffic safety. Available safety measures include preventive safety systems, aimed at avoiding or mitigating the consequences of a possible crash before impact, and protective safety systems, aimed at protecting the occupants from the consequences of a crash during impact. One example of an effective protective safety system for rear-end crashes is whiplash protection with energy-controlling structures and optimized headrest designs (Kullgren, Krafft, Lie, & Tingvall, 2007; Kullgren, Stigson, & Krafft, 2013). For preventive safety, Automated Emergency Braking (AEB) has been shown to be an effective preventive safety system, reducing the number and severity of rear-end crashes substantially (Cicchino, 2017; Fildes et al., 2015).

There are different types of AEB algorithms. Early AEB systems only included time-to-collision (TTC; based on the vehicles' relative distance, speeds and accelerations) and the braking response by the driver of the following vehicle (FV) (Brännström, Sjöberg, & Coelingh, 2008), while more mature systems may also consider the FV driver's ability to steer away comfortably (in addition or as an alternative to braking; see Brännström, Coelingh, & Sjöberg, 2010, 2014; Sander, 2018). This consideration reduces false positives, which are activations when either there is no real need for

* Corresponding author.
E-mail addresses: ollejap@chalmers.se (P. Olleja), jonas.bargman@chalmers.se (J. Bärgman), nils.lubbe@autoliv.com (N. Lubbe).

https://doi.org/10.1016/j.jsr.2022.08.011
0022-4375/

it or the driver does not feel it is warranted (Bliss & Acton, 2003; Coelingh, Jakobsson, Lind, & Lindman, 2007; Sander & Lubbe, 2016). In the latter case, drivers may still avoid the crash while remaining inside their comfort zone. We have chosen to quantify the comfort zone boundary by selecting an acceptable lateral acceleration for the steering-away maneuvers. The term "comfortable steering" will be used to describe maneuvers that do not take drivers out of their comfort zone, even when performed when braking is no longer an option to avoid a crash. An AEB considering comfort zone boundaries delays the intervention until the driver cannot avoid the crash by steering in a comfortable way (e.g., when the lateral acceleration crosses the driver's comfort zone boundary; see Bärgman, Smith, and Werneke, 2015; Summala, 2007).

To quantify the benefit of safety systems, such as AEB, developers need both assessment methods and data. They need to assess to what extent a specific concept (or even the specific application of a system) will affect safety all through the systems' life cycle (Alvarez et al., 2017). This type of assessment is prospective, predicting the potential safety benefit of a system before data are available from real-world crashes. There are different methods available for the prospective assessment of AEB systems. One increasingly popular method is virtual simulations in computers. This allows for early assessment of a system's potential, and enables fast and iterative improvement of its safety effectiveness. Virtual traffic simulation is one such approach. The movement of traffic participants is modeled with respect to vehicle dynamics, and driver behavior and control, or alternatively, with added automated control (Fahrenkrog et al., 2019; Helmer, 2014). When traffic characteristics, driver, and automation behavior are modeled accurately, one can expect to create an accurate representation of crashes as the outcome of interest; however, such detailed modeling is highly ambitious (Dobberstein et al., 2021). The benefit of automation can be determined by comparing a simulation with only human drivers to simulations with automation, either by re-simulating selected critical events (Fahrenkrog et al., 2019; Hallerbach, 2020) or only the crash events generated in the human-driver-only simulations (Tanaka, 2015). Traffic simulation for safety benefit assessment holds the promise of creating an essentially unlimited amount of parameter variations and crash events. However, as driver behavior is complex and modeling them accurately (enough) is difficult (Markkula, 2015), such models "only represent the behavior of real drivers to a certain extent" (Bjorvatn et al., 2021, p. 123) and results are sensitive to the variations of such models (ISO, 2021). Note that the traffic simulation approach is often targeting assessments of higher levels of automation.

Counterfactual or "what-if" simulations is another simulation-based assessment method that has been used extensively to assess advanced driver assistance systems (ADAS; Bärgman, Lisovskaja, Victor, Flannagan, & Dozza, 2015; Davis, Hourdos, Xiong, & Chatterjee, 2011; McLaughlin, Hankey, & Dingus, 2008; Scanlon et al., 2021). These simulations typically assess safety by using pre-crash kinematics from real-world data (Kusano & Gabler, 2012; Lindman & Tivesten, 2006; Sander, 2018; Scanlon et al., 2021), simulating each event with and without an algorithm modeling the preventive safety system under assessment (Kusano & Gabler, 2012; Sander, 2018). The results are typically provided in the form of the proportion of crashes that were avoided with the system, and the impact speed (or injury risk) distribution of the crashes that still occurred after the system was applied. In this way it is possible to virtually compare the original, baseline event with the modified ("what-if") events that include the AEB system.

Data on pre-crash kinematics are needed to perform the counterfactual AEB simulations. Different sources of pre-crash kinematics data include in-depth crash reconstruction, event data recorders (EDRs), and naturalistic driving data (NDD), which are collected either in-vehicle or on-site (i.e., monitoring a specific piece of road; see Krajewski et al., 2018).

In-depth reconstructed crash databases include information not only about the crash, but also about the pre-crash phase (Bakker et al., 2017). Experts can reconstruct the pre-crash kinematics and document many other aspects of the crash, such as the road geometry and other environmental factors—as well as the injuries sustained by the humans involved in the crash (Otte, Krettek, Brunner, & Zwipp, 2003). Typically, however, very little information is available about the pre-crash phase (Schubert, Erbsmehl, & Hannawald, 2013). Also, in-depth crash data with reconstructed pre-crash kinematics are not available in all countries or regions for which safety systems (such as AEB) should be evaluated prospectively. One example of an in-depth crash database is the German In-Depth Accident Study (GIDAS), which started collecting crash data in 1999. Approximately 2,000 crashes from the cities of Hannover and Dresden and their surroundings are added every year (Otte et al., 2003; Liers, 2018). GIDAS crashes are all reconstructed with estimates of crash kinematics and impact speed.

For a subset of crashes in the GIDAS crash database a Pre-Crash Matrix (PCM) is created, which includes the pre-crash kinematics of the vehicles involved up to five seconds before the collision. The crashes are reconstructed using a structured approach (Schubert et al., 2013). As of February 2018, the GIDAS PCM database contained 9,729 crashes (VUFO, 2020). Reconstruction of pre-crash kinematics has also been performed for other in-depth databases, such as the Initiative for the Global Harmonization of Accident Data (IGLAD) (Spitzhüttl, Petzold, & Liers, 2015) and the Road Accident Sampling System India (RASSI) (Shaikh & Sander, 2018).

The pre-crash kinematics data from reconstructed crashes can be used directly in counterfactual simulations (Rosén, 2013; Sander, 2018; Scanlon et al., 2021). Typically, the system under assessment is applied to the pre-crash kinematics and, for each timestep, a threat assessment analysis is performed. The simulation framework always includes a vehicle model, and often a model of the driver. The outcomes of the simulations consist of avoided or (hopefully) crashes with reduced impact speed and, thus, mitigated injury risk. That is, counterfactual simulations can also include collision models, so that in case of a crash the occupants' injury risks can be studied (Sander & Lubbe, 2016, 2018).

As an alternative to using data from in-depth crash investigations, real-world pre-crash kinematics for counterfactual simulations can be extracted from event data recorders. These recorders are already mandatory in new vehicles in several countries (NHTSA, 2006; UNECE, 2019), and more countries are following suit (Šajn, 2019). The event data recorders of today typically record, among other things, the vehicle speed in the few seconds leading up to the crash and the acceleration during the crash. However, the pre-crash data are often recorded at a low frequency (1–5 Hz), so it is often not known exactly when (within the 200 ms to 1 s that the sample frequency provides) the impact occurred, which reduces reconstruction quality (Thomson et al., 2013) and, naturally, impacts simulation validity. Nevertheless, event data recorders are a useful information source when reconstructing crashes: similar to in-depth reconstructed crash databases, they represent real-world crashes and have been used extensively as a basis for counterfactual simulations (Bareiss, Scanlon, Sherony, & Gabler, 2019; Kusano & Gabler, 2012; Scanlon, Kusano, & Gabler, 2015; Scanlon, Page, Sherony, & Gabler, 2016).

Lastly, NDD can also be a source of pre-crash kinematics data for counterfactual simulations. NDD are recorded unobtrusively in real-traffic, and two main types of such data exist: site-based NDD and in-vehicle NDD.

Site-based NDD are collected at one or more specific sites, where, typically, cameras, radars or LIDARs collect data about road-user movements over a time duration from minutes to months (Bock et al., 2020; Krajewski, Moers, Bock, Vater, & Eckstein, 2020; Krajewski et al., 2018; Laureshyn, 2010; Smith, Thome, Blåberg, & Bärgman, 2009). The data are post-processed to produce trajectories and other information, such as speed and acceleration, about the road users captured in the recordings. Most of the data from site-based NDD collections capture normal everyday driving without any critical events; they contain very few crashes (Van Nes, Christoph, Hoedemaeker, & Van Der Horst, 2013). The highD dataset is one example of recent site-based NDD which only includes normal everyday driving data (Krajewski et al., 2018). The data were collected by drones recording video of six stretches of highway in North Rhine-Westphalia, Germany in 2017 and 2018. There were 60 recordings, and the drones recorded 17 minutes per recording, on average. A total of 110,000 vehicle trajectories were recorded, with a typical length (longitudinal road segment) of 420 m. The highD data are freely available for research purposes (Krajewski et al., 2018).

In contrast to site-based NDD, in-vehicle NDD are collected from vehicles instrumented with a data acquisition system that collects vehicle information such as speed and acceleration, driver information such as glance behavior, and data about surrounding traffic (typically using radar and cameras). The largest in-vehicle NDD study to date is SHRP2: data were collected on 3,247 drivers who drove a total of almost 80 million km over a period of three years in the United States. Since SHRP2 collected so much data, more than 1,000 crashes of different severities were recorded (SHRP2 crash severity levels 1–3), as were other critical events (e.g., near-crashes: see Blatt et al., 2015; VTTI, 2020).

NDD have been used to study the safety benefit of, for example, forward collision warning (FCW) and AEB (Bärgman, Boda, & Dozza, 2017; Woodrooffe et al., 2012). When NDD are used in counterfactual simulations of rear-end crashes, the evasive maneuver of the FV in each event can be replaced by an evasive maneuver created by a quantitative driver response model (Bärgman et al., 2017). The main reason for this replacement is that each crash or near-crash is just one instance of the behavior of that driver, which just happened to produce that particular crash or near-crash. If the driver had acted differently, a crash may have been a near-crash or a more severe crash, or a near-crash could have become a crash. Here the underlying mathematical models of driver behavior (glance and response models) are fundamental for exploring the various possibilities (Bärgman et al., 2017). The simplest possible replacement behavior is to assume the driver sleeping. That is, that the driver does not act at all during the crash. This can be considered a worst-case behavior in any particular situation. Another way of using NDD that includes normal driving (and, possibly near-crashes) is to get distributions for stochastic variations, which can be used to both define the exposure to driving scenarios and to vary scenario characteristics. These distributions can then be used in, or together or compared with, virtual traffic simulations. There is research quantifying the relationship between near-crash increase and crash increase (Guo, Klauer, McGill, & Dingus, 2010; Victor et al., 2015), but the same cause-effect behavior is less noticeable when using normal driving data. For these reasons, working with exposure to scenarios from normal driving in relation to crash occurrence needs to be done with caution (Woodrooffe et al., 2012).

The choice of data source (and whether to remove evasive maneuvers from the original event) in a counterfactual simulation is driven by several factors, including what systems are to be assessed (e.g., whether driver behavior is to be evaluated), and whether the data are available. The availability of in-depth crash data with reconstructed pre-crash kinematics and even data recorder data is limited. When a preventive (or protective) safety system is to be developed for a specific market where in-depth reconstructed crashes or event data recorder data are not available, alternatives are needed. One option is to collect NDD and create synthetic crashes based on the structured application of models of driver behavior to non-crashes, as described above (Bärgman et al., 2017).

In this study we investigate the feasibility of using site-based NDD (highD) and in-vehicle NDD (SHRP2) non-crashes to create counterfactually simulated crashes, by comparing the resulting crash characteristics with those from reconstructed in-depth crashes (GIDAS). Comparing highD to GIDAS is comparing two samples of German highway rear-end crashes, hence we believe the comparison gives direct insights in how well the generated NDD crashes represent the reconstructed actual crashes. To study the suitability of generated crashes from the U.S. SHRP2 data, they would ideally be compared to a crash sample of identical sampling criteria, which we did not do, as such time series pre-crash data were not readily available. However, as SHRP2 is by far the most comprehensive NDD in the world to date, and GIDAS-PCM is one of the most commonly used high-quality crash datasets for counterfactual benefit assessment (of, for example, AEB), our comparison aims to study general comparability of data and results from the application of two different AEB system, rather than focusing on regional comparability. We evaluate both the crash avoidance and false positive rates of AEB systems. If the simulated crash characteristics are comparable to those of the reconstructed crashes, in the comparison of German highway rear-end crashes, then it may be feasible to use the more readily available and affordable NDD for early prospective assessments of preventive safety systems. If the U.S. NDD data were comparable to German highway crashes, then results are not sensitive to data choice, suggesting a liberal interpretation on generalization is suitable, at least for the relatively similar driving cultures of the United States and Germany.

The aim of this study can be divided into three parts: first, to compare crash characteristics generated from NDD with real reconstructed crashes; second, to quantify the influence of the data source on a comparison of the practical safety benefits of two AEB algorithms (one basic and one more advanced based on driver comfort zone boundaries, which seeks to reduce false positive activations by accounting for FV steering maneuvers); third, to demonstrate the use of NDD for assessing AEB algorithm false positive rates (and to confirm the hypothesis that the more advanced AEB system has a lower false positive rate).

2. Method

This section describes the data used in the study, the crash generation process, the AEB system application, and the simulation framework. Finally, the analysis steps are outlined.

2.1. Data

Three different sources of data were used in this study: GIDAS (Otte et al., 2003), SHRP2 (Victor et al., 2015; VTTI, 2020) and highD (Krajewski et al., 2018).

2.1.1. GIDAS – PCM

In this study-two subsets of GIDAS data, released in July 2018, were used. The first subset consists of all the highway rear-end crashes for which PCM was available. This subset was used as a reference for the counterfactual simulations assessing AEB. The second subset contained all GIDAS highway rear-end crashes in which the following vehicle (FV) did not perform an evasive braking maneuver prior to impact with a braking lead vehicle (LV), representing crashes as similar as possible to the crashes generated from NDD (where we, in the crash generation, assumed sleeping drivers). This subset includes cases for which PCM data were unavailable. Because impact speed is usually estimated for all GIDAS crashes even when PCM is not available, this subset was used as a reference in the assessment of the differences in impact speed. Within this subset (crashes with no driver evasive maneuver beforehand), PCM was available only in 7 of the 46 crashes. AEB assessment was not performed separately on these seven crashes as the low number of cases would have been too few for a relevant comparison. Note that we filtered out all but highway rear-end crashes from the GIDAS data (both with and without PCM), to maximize the match with the highD data.

The vehicles' relative longitudinal distance and lateral overlap in the pre-crash phase are used to predict the collision path in the AEB implementation. However, PCM does not directly code those values, so they were derived using other metrics in the PCM data. The predicted future path of the FV was generated as an arc with an assumed constant yaw rate for the cases when the FV turned, or as a straight line for the cases when the FV went straight. See Appendix A for a detailed description of the overlap calculations.

2.1.2. highD

The first naturalistic driving dataset used in this study is the highD dataset (Krajewski et al., 2018). It consists of processed drone video recordings of vehicles on German highways. Relevant time-varying parameters such as position, speed, and acceleration of the vehicles were extracted. The criticality of the interaction between each pair of vehicles (FV and LV) was assessed by extracting each LV's lowest acceleration value (its harshest braking maneuver) along with the time it occurred. The time headway between the FV and the LV was noted at that instant in time. Fig. 1 shows the relationship between the minimum acceleration and the time headway for all vehicle pairs in the dataset. In this study, only FV/LV interactions with a minimum LV acceleration of -2 m/s^2 or less and a time headway of five seconds or less were considered potentially critical and used in the crash generation process.

The highD datasets also provided information about the amount of lateral overlap between the LV and the FV. Fig. 2 shows the distribution of the overlaps at the time of minimum acceleration (and time headway extraction; see Fig. 1). The mode of the distribution is at approximately 1.70 m, a reasonable vehicle width in Germany. The distribution in Fig. 2 was obtained by taking the lesser of the left and right overlaps, assuming that the FV can always choose to steer left or right of the LV and that the FV and LV trajectories are always parallel to the road (information about yaw angle of the vehicles were thus not included in the data).

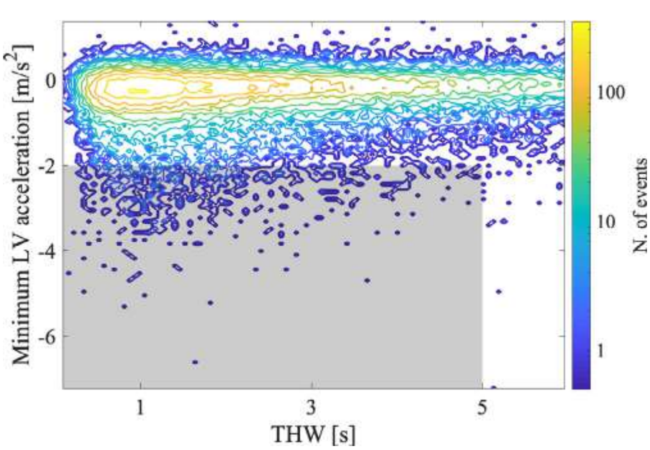

Fig. 1. Contour map of the acceleration and the time headway for all events in the highD dataset. The potential criticality of the scenario increases to the left and down. The grey area contains all the events considered for crash generation.

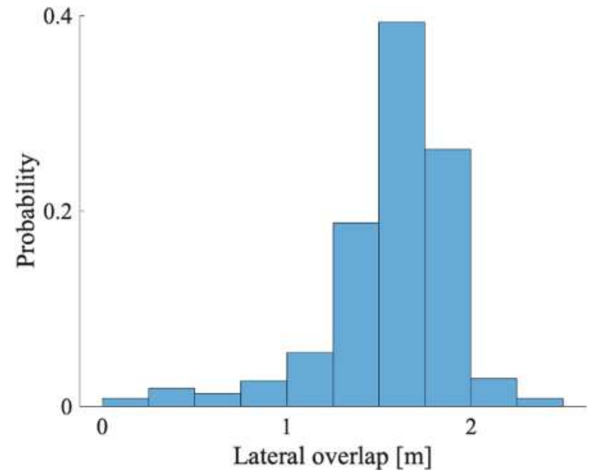

Fig. 2. Distribution of minimum lateral overlap within highD dataset.

2.1.3. SHRP2

The second naturalistic driving dataset used in this study comprises a subset of the SHRP2 naturalistic driving study. The subset originally contained 46 crashes and 211 near-crashes (Victor et al., 2015). In this study only the near-crashes were used. These near-crashes were manually reviewed by expert annotators, after an initial pre-filtering using kinematic or proximity triggers (e.g., longitudinal acceleration; see Hankey et al., 2016 p. 25–26 for details about the near-crash definition used). Of the 211 near-crashes, only 190 had the full kinematics data for both LV and FV vehicles. In 17 cases the crash generation procedure (described in Section 2.2) did not result in a crash, so those were discarded. An additional 42 of the generated crashes were not included, as the LV performed more than one braking maneuver, increasing and decreasing speed multiple times. The number of (near-crash-based) crashes used in the final analysis was thus 131. Note that, in contrast to the closely matched GIDAS PCM and highD data, the U.S. SHRP2 data were not restricted to highway crashes; it included crashes across several road types (such as rural, urban, suburban, and highway).

The comfort-based AEB algorithm (CAEB algorithm, see Section 2.3) requires lateral offset distances to make steering avoidance assessments feasible, but this information was not available in the SHRP2 dataset. Therefore, the offset distribution in the highD data was also used for the SHRP2-generated crashes, assuming parallel trajectories. Multiple simulations were run for each of the original crashes, applying each offset (bin) from the distribution in Fig. 2. The relative probabilities (weights) for each offset (bin) were considered in post-processing and the final results were calculated by weighting the simulation outcomes by their relative probabilities (per bin).

An illustration of the data usage and crash generation process in the study can be found in Fig. 3. For each dataset, only a subset was used for the crash generation and AEB application.

2.2. Crash generation

Crashes were generated from the two naturalistic driving datasets. The original kinematics of the events from highD and SHRP2 were used to define the moment the LV started braking as the moment when the LV reached an acceleration of -1 m/s^2. Decelerations closer to zero were probably caused by the driver's foot lift-

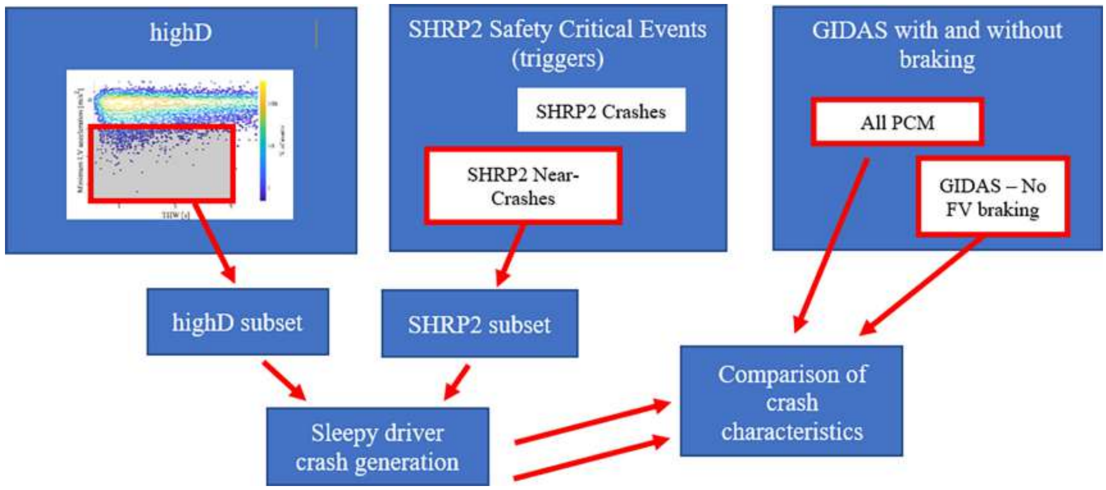

Fig. 3. Visual representation of the data selection process for crash generation and comparison, for highD, SHRP2 and GIDAS databases.

ing slightly from the accelerator pedal, and therefore were not considered further in this work. From the LV's start of braking, the speed of the FV was set to be constant until the crash happened. That is, the FV driver never performed any deceleration in response to the LV deceleration—basically simulating a sleeping driver. Fig. 4 shows this modification process, which was applied for all the highD scenarios and SHRP2 near-crashes used in the study. In many of the selected highD cases, although criticality was established by the LV kinematics, the event ended before the FV could reach the LV because it occurred near the end of the segment of road that each drone covered and recorded. (The segment was only 420 m long, and the LV decelerations happened at different points over this distance). A total of 361 events were excluded from highD dataset as a result. In total, 378 crashes were generated from highD and 131 (all) from SHRP2.

2.3. AEB algorithm descriptions

This work used a reference AEB algorithm (RAEB) and an AEB based on drivers' comfort zone boundaries with respect to lateral acceleration (comfort-based AEB; CAEB). Each algorithm was applied to the crashes from all three datasets in order to compare crash avoidance and mitigation results.

The RAEB, based on the work of Brännström et al. (2008), only considers possible longitudinal avoidance by the system. That is, it does not include the driver's capacity to avoid the crash by comfortable steering, instead identifying the moment when the deceleration required by the system to avoid the crash passes a threshold (the point in time after which the FV would be unable to avoid a crash). The RAEB considers the current speed, acceleration, and relative distance of the FV and LV as well as the maximum braking performance of the FV braking system. The values used to quantify braking performances were the maximum deceleration reachable by the vehicle ($-10\ m/s^2$), the jerk reachable by the braking system ($-50\ m/s^3$), and the time delay of the braking system (0.08 s, from Bärgman et al., 2017; Brännström et al., 2008). After the RAEB activation, the FV evasive maneuver (braking) played out according to the values used as input for the braking system limits. If the vehicle was traveling in a straight line, the braking maneuver was simulated in the same direction of travel. If the vehicle was turning (e.g., in a curve), the braking maneuver was simulated with the assumption that the FV traveled at a constant steering angle.

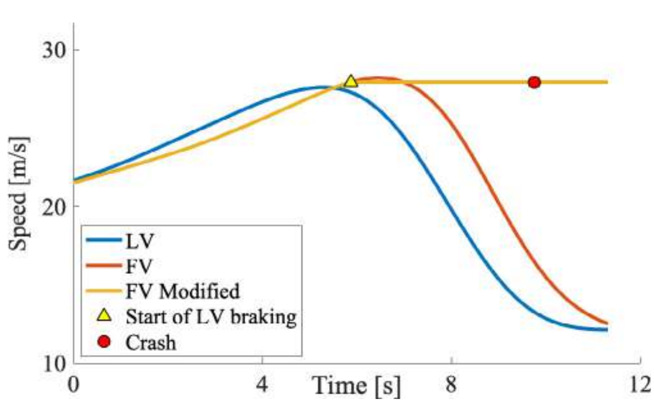

Fig. 4. Conceptual demonstration of the removal of the FV's braking.

The CAEB algorithm was also applied to all crashes in the study. Unlike the RAEB, this algorithm took into account the capability of drivers to avoid crashes by performing a comfortable steering maneuver. Depending on the relative speeds involved in the event, a driver might still comfortably perform an evasive steering maneuver to avoid a crash even when the BAEB may have already triggered (Brännström et al., 2014). Thus the CAEB may eliminate some early interventions (potential false positives). The algorithm also includes parameters of driver comfort limits in terms of lateral acceleration and a basic single-track bicycle model that defined the lateral dynamics of the vehicle. The bicycle model is only a first approximation of a vehicle but it was considered sufficiently accurate for this study.

The CAEB algorithm simulated an S-shaped maneuver by the driver: at the end of its trajectory, the FV is parallel to its position at the beginning and is next to the LV. That is, the FV and the LV are in the same longitudinal position but separated by a lateral safety distance (see Appendix A). The S shape was designed as follows: the angle of the steering wheel was gradually increased, considering the limit for steering wheel speed of 720°/s (Brännström et al., 2014) and the driver comfort limit for lateral acceleration of $5\ m/s^2$ (Sander, 2018). The latter is determined by the vehicle turning (yaw) rate and the vehicle speed. When the maximum tolerable lateral acceleration (the driver's comfort zone boundary) was reached, the steering angle was kept constant until the FV steered back, ending its trajectory parallel to its position at the start of the maneuver.

A key parameter for the generation of the FV's trajectory was the lateral distance the FV needed to traverse to avoid a crash with the LV, which depends on the lateral overlap (see Appendix A). The simulations of the trajectories included the lateral overlap, and the additional safety distance simulations included run-time estimations of the lateral distance traveled by the FV. To speed up simulations, only the first half of the FV trajectory was simulated: after the FV had traveled half the required lateral distance, the second half of the trajectory was assumed to mirror the first half. Additional vehicle dynamics during lane changing, such as tire slip, were ignored.

The CAEB algorithm was designed to reduce false positive activations and thus activates later than RAEB. Hence, RAEB intervention timing was used as the starting point for the activation decision of CAEB. From this starting point the vehicles were simulated as follows: (a) the LV continued along its original path and (b) the FV was projected along the newly created steering trajectory. If the FV was able to complete the steering trajectory and avoid a crash with the LV, then the AEB intervention was delayed (by one time-step); otherwise the AEB activated. This process was iterated, with the FV getting closer to the LV at each iteration, until the new trajectory resulted in a crash with the LV. Once the steering trajectory resulted in a crash, the AEB was activated and the evasive braking maneuver was applied (simulated).

2.3.1. False positive assessment

The false positive performances of the two AEB algorithms were assessed by applying them to the original events used to generate the crashes for the AEB assessments. Only the potentially critical highD events used to generate crashes (378 events, see Section 2.2) were considered of interest for the false positive assessment; all the other events recorded weak or absent LV braking maneuvers. For the false positive assessment of SHRP2, the original events included only previously selected near-crashes (131 events, see Section 2.1.3). In summary, critical (non-crash) events from the original highD and SHRP2 data (including the original braking behavior of the FV) were simulated with the RAEB and CAEB systems.

3. Results

In this section the results of the study are presented. First, for all datasets, generated crashes are compared to real-world crashes. Second, the results of the AEB algorithms' application to the data are shown, followed by the results from the analysis of false positives.

3.1. Crash comparison

Fig. 5a shows a comparison of the cumulative distributions of the maximum level of deceleration reached by the LV in the pre-crash phase across the three datasets. The SHRP2 and PCM events show harsher braking maneuver s (higher values of deceleration) than the highD events. However, there are more PCM events with relatively low maximum decelerations, in which the LV did not brake or only braked slightly. Note that this also includes LVs that stand still, for example in a traffic jam.

Fig. 5b shows a comparison of the time elapsed from the start of LV braking to the crash across the three datasets. The distributions for SHRP2 and highD consider all the generated crashes, while the PCM distribution only includes crashes in which the LV braked with a deceleration of at least -1 m/s^2 (60 crashes, from Fig. 5a), the same threshold used for the highD crash generation. PCM crashes where the LV applied more than -1 m/s^2 deceleration show a time-to-crash comparable to that of crashes generated from SHRP2 near-crashes. However, there is a substantial difference in the time-to-crash between SHRP2 and PCM on the one hand and highD on the other.

Fig. 5c compares the impact speeds across the three datasets (without AEB applied to the data). As the analyzed crashes are rear-end crashes only, the impact speed was computed as the relative speed between the vehicles, assuming that they were driving parallel to each other. The crashes generated from the naturalistic datasets show an overall substantially lower impact speed compared to the real crashes in PCM. The distribution of crashes in the GIDAS database where the LV was braking but the FV did not brake is also shown. Because not all crashes in GIDAS have been reconstructed into PCM (recall that PCM data were available only for seven of these crashes), simulations were performed on all rear-end PCMs (including those with FV braking), increasing the case count to N = 134. The distributions of the impact speed for all rear-end PCM crashes and the GIDAS no-FV-braking crashes are similar, especially for impact speeds between 10 m/s and 20 m/s, with larger differences in the tails.

Fig. 5d shows the relative FV-LV speed at the point when the LV deceleration reaches -1m/s^2, or, if the LV did not brake, when the FV starts braking or, if also the FV did not brake, when it crashes. The two NDD are very similar in the initial conditions, while the PCM data has much higher initial relative speeds.

Fig. 5e shows the comparison of the lateral overlaps (see Sec 2.1.1) at the time of the crash, indicating that the overlaps were lower for the PCM crashes than for the highD-based crashes. As noted, SHRP2-based crashes did not include information about the lateral overlap, so they were not included in the comparison.

3.2. The influence of data source choice on the comparison of AEB safety performance

The two AEB algorithms were applied to the crashes of all datasets. The RAEB only considers longitudinal kinematics, while the more advanced CAEB aims to decrease early (nuisance) interventions by accounting for the potential of the driver's evasive action (comfortable steering). Fig. 6a shows the cumulative frequencies of impact speed when the RAEB is applied and the crashes are mitigated, but not completely avoided (non-crashes are excluded; remaining crashes are N = 22 for PCM, N = 11 for highD and N = 12 for SHRP2). The remaining crashes all have lower impact speeds than the original crashes, and the crashes generated from highD and SHRP2 have lower impact speeds than the crashes from the PCM. Fig. 6b shows the cumulative frequencies of impact speeds of the mitigated crashes when the CAEB is applied to all three datasets. The impact speeds are higher than those obtained with the RAEB, and fewer original crashes were avoided (N = 34 for PCM, N = 41 for highD and N = 19.9 for SHRP2). (Recall that for SHRP2 crashes, the results include the weighting process described in Section 2.1.3, applying the offset distribution from highD crashes, which is why there are non-integer crash results for SHRP2). Fig. 6c shows the crash avoidance performances (as percentages of original crashes) of the two tested algorithms. As expected, the RAEB avoided more crashes than the CAEB for all the datasets tested.

3.3. False positive analysis

False positives were analysed in the original highD and SHRP2 no-crash events. The RAEB and CAEB algorithms were applied and the results were compared. The application of RAEB to highD resulted in four false positives, all occurring at very low-speed (<3 m/s) events in traffic jams, with FV and LV vehicles closely following each other. CAEB did not avoid any of these four false positive interventions, probably because of the short distances between vehicles (as they were low-speed events), together with the fact that steering is much less effective at low speeds. For SHRP2, the analysis of false positives was first performed on all

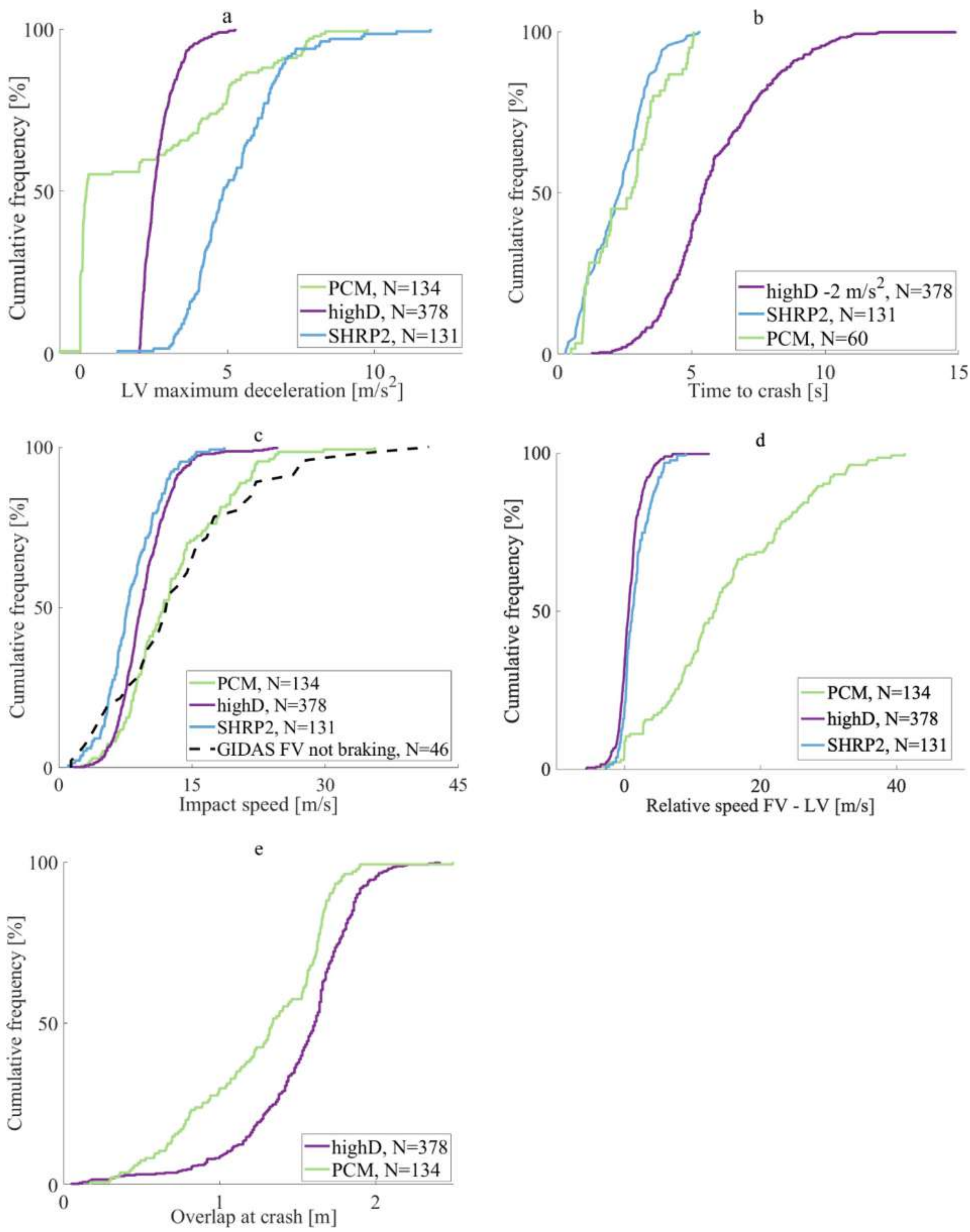

Fig. 5. (a) Cumulative frequency of LV maximum deceleration during the pre-crash phase. (b) Cumulative frequency of the time elapsed from LV braking initiation to the crash (at -2 m/s^2). (c) Cumulative frequency of original impact speed of all datasets. (d) Cumulative frequency of the relative speeds at the start of the event. (e) Cumulative frequency of lateral overlap between LV and FV at the time of crash.

131 events, and then on a subset of these events—those in which at least one of the vehicles reached a speed of 60 km/h (N = 42). This subset was considered more similar to the other datasets in this study. The RAEB application resulted in 28 false positives in 131 events (21.3%); seven occurred in the 42 high-speed events (16.7%). As noted previously, for the application of CAEB to SHRP2 data the simulations used the overlap distribution from the highD data (see Fig. 2). The simulation results were weighted according to the probability of each simulated overlap. The results of this procedure were not necessarily integers. To make this apparent to the

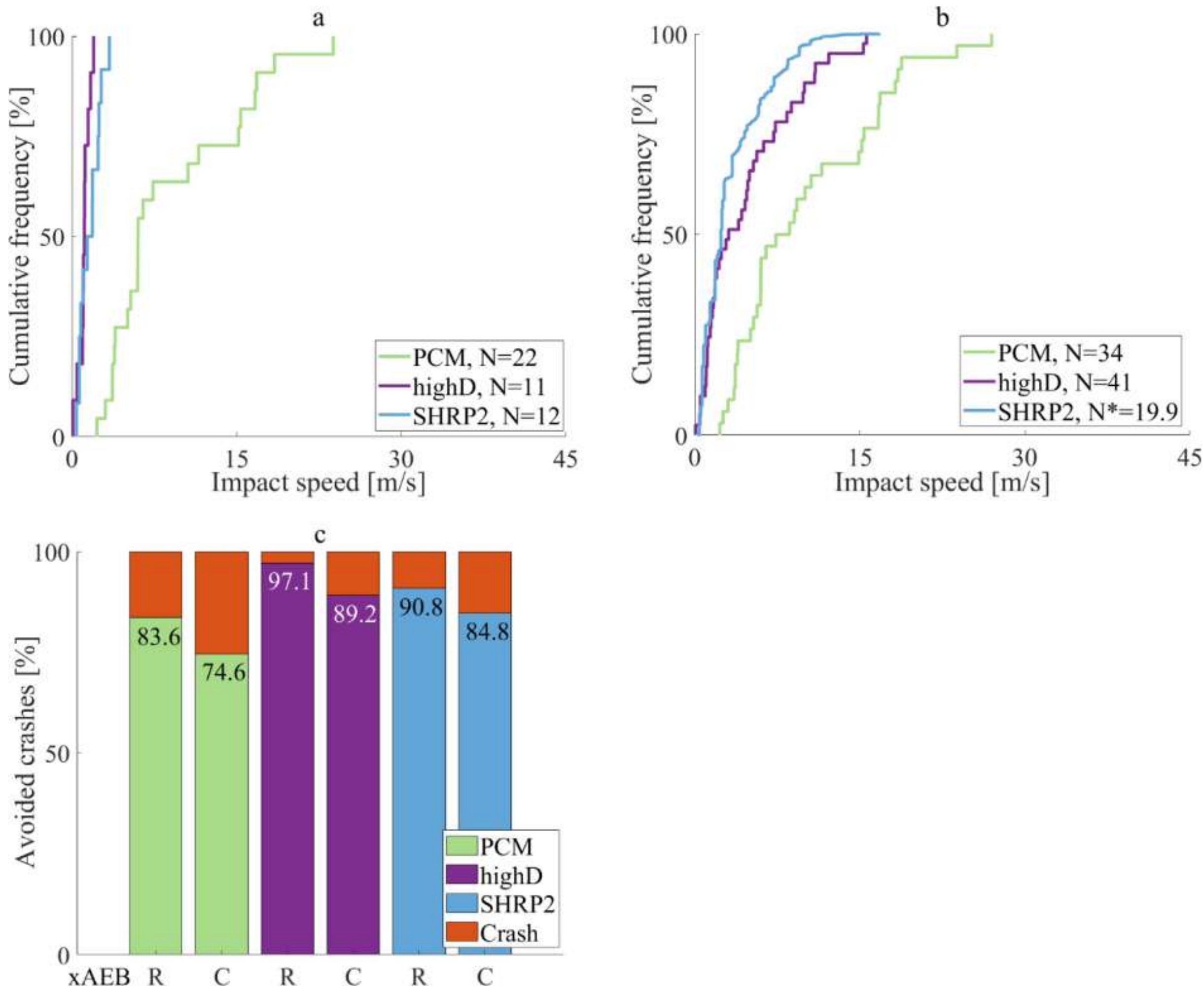

Fig. 6. (a) Cumulative frequency of the impact speed in crashes remaining after basic AEB application. (b) Cumulative frequency of the impact speeds in crashes remaining after advanced AEB application. N* is the theoretical number of avoided crashes resulting from the weighting process of the probabilities of the overlaps. (c) Bar plot showing the percentage of avoided (original) crashes for each dataset and both AEB systems.

reader the false positive counts were, in Table 1, intentionally left with one decimal place. The false positives decreased from 28 to 24.8 (from seven to 5.9 at high speed) when the CAEB was applied. The results are summed up in Table 1.

As expected, when SHRP2 crashes were simulated with small overlaps (see Fig. 2), the CAEB produced fewer false positives than the RAEB. As an example, the CAEB avoided six false positives (two of which were at high speed) at the (Fig. 2) bin with the smallest overlap.

4. Discussion

This study explored the possibility of using crashes generated from non-crash naturalistic driving data (NDD) to complement, or, in some instances even replace, real crashes with time-series kinematics (e.g., from reconstructions) for counterfactual assessment of AEB systems. This possibility would be useful for analysis of safety benefit (and, potentially, system optimization) for countries lacking crash data with time-series kinematics. We have shown that, in general, crashes generated from NDD have substantially lower impact speeds than real crashes (GIDAS PCM), but the pre-crash criticality in crashes generated from U.S. NDD (SHPR2) near-crashes is comparable to the criticality in real German crashes.

Table 1
False positive counts in highD and SHRP2 events for the application of RAEB and CAEB.

	highD		SHRP2	
	RAEB	CAEB	RAEB	CAEB*
Low speed	4	4	21	18.9
High speed	0	0	7	5.9
Total	4	4	28	24.8

* Values for CAEB applied to SHRP2 are weighted according to the overlap probability, possibly resulting in non-integers.

4.1. Comparing crash characteristics across datasets

4.1.1. Crash generation

The events from the two naturalistic datasets used in this study, the highD everyday highway driving and the SHRP2 near-crashes, by definition are not crashes. As an aim of this study was to investigate the suitability of using non-crash NDD to simulate counterfactual crashes, only near-crashes from SHRP2 were included in the study (i.e., not crashes). To generate a crash, the FV braking maneuver had to be removed, so the start of the FV braking had to be defined. In other studies this process was done manually (Bärgman et al., 2017) or computationally by fitting a piecewise linear model to the FV deceleration (Markkula, Engström, Lodin, Bärgman, & Victor, 2016; Svärd, Markkula, Engström, Granum, &

Bärgman, 2017). In this study, however, a different approach was used: an LV deceleration threshold of -1 m/s^2 defined the point in time when the FV speed was set to remain constant, eliminating any further deceleration. That is, the FV driver may or may not have initiated braking at that time. This process can be seen as simulating a substantially "distracted" FV driver or, maybe more accurately (as the duration of eyes-off-road is typically quite long) a sleeping driver, as the driver does not react to the unfolding of the critical event. This approach was used because of the different nature of the datasets analyzed compared to those in previous studies. Removing the evasive maneuver can be considered the worst possible outcome (unless the FV driver accelerates into the crash, which would be even worse — but would be unlikely).

The times needed for the FV to crash into the LV (Fig. 5b) were similar between GIDAS PCM and SHRP2, reflecting the fact that SHPR2 data actually capture critical events—but they were substantially different for highD and SHRP2, reflecting the different origins of the two datasets (everyday driving vs critical events). This difference indicates that highD data are likely not suitable as an artificial source of critical events, neither with respect to timing (criticality) nor impact speed. Some highD events proved to be more safety–critical than others, but a much more extensive data collection (capturing more critical events) is needed to make a highD-like dataset even marginally useful for, for example, AEB assessment. Note that SHRP2 collected over 3,958 driving man-years (the total number of years that data were collected of participants' everyday driving), and still only captured 125 rear-end crashes. Although highD captured many vehicles simultaneously, the amount of data collected (time per vehicle) was several orders of magnitude less than that of SHRP2; further, all data were recorded on straight highways.

In addition to differences in timing and impact speed, the differences in LV decelerations between the datasets (Fig. 5a) is also likely to affect the AEB assessment. In particular, there were many PCM crashes that had low, or no, LV deceleration, while the SHRP2 LV decelerations were much higher compared to highD, and the highD decelerations were quite uniform (and never lower than 1 m/s^2, per definition). The LV decelerations for highD are simply everyday driving decelerations. In order to produce the high impact speeds of the PCM, they are likely not simple car-following events with the LV braking creating the crash; instead, the LV is at a standstill, or driving at a substantially lower speed than the FV, with the FV catching up (or the LV is cutting in front of the FV, at high relative speed). The differences in scenarios can also be seen in Fig. 5d, where 13% of the PCM cases had a relative speed of up to 10 km/h, while 75% of the SHRP2 cases and 89% of the highD cases had low relative speeds (up to 10 km/h). These findings demonstrate that much care should be taken when (if) using crashes generated from near-crashes for virtual safety assessment. Selection criteria (incl. event categorization – specific subtype of the events used) must be at the forefront of the considerations. If the algorithms act differently in the two types of events (e.g., LV braking and FV catch-up), the results are likely to be misleading.

4.1.2. Crash comparison

The comparison of impact speed (Fig. 5c) makes it clear that overall, the GIDAS PCM crashes have much higher impact speeds compared to the generated crashes. Although the PCM crashes were only highway rear-end scenarios, there were sometimes specific circumstances resulting in more critical events, such as the very late appearance of the LV in the FV's path (e.g., due to a cut-in). Therefore, the higher impact speeds were expected, especially when compared to the SHRP2-based crashes, which were mostly lower-speed events (and only a relatively small proportion occurred on highways). That is, the low initial speeds (and possibly short time headway) in the SHRP2-based crashes—and the fact that most were car-following situations (and not FV catch-up situations, with large relative speed differences)—kept them from becoming high impact-speed crashes such as those found in GIDAS PCM data.

What makes the situations less critical in the highD than in the GIDAS PCM is that the LVs typically decelerate less in the former (see Fig. 5a). Further, in the latter, although the LV did not always brake, there were also cases with high LV deceleration (when the LV appeared suddenly in the path of the FV). These cases did not occur in highD: as noted, the deceleration in highD was more uniform.

In summary, the results of this study show that the impact speeds of crashes generated from site-based or in-vehicle NDD near-crashes during normal driving are substantially lower than the impact speeds of real crashes—at least, the real crashes from the GIDAS PCM used here. One reason for the difference is probably that, for an event to be included, the GIDAS PCM required that at least one person be suspected of being injured. If the PCM data also included crashes without personal injury, it may be that the crashes generated from NDD would be similar to the lower tail of the PCM data's impact speed distribution. However, as safety assessment typically prioritizes avoiding human injuries, this observation may not be relevant.

The comparison of lateral overlap at the time of the crash (Fig. 5e) showed similarities between the highD and GIDAS PCM data, but PCM on average had smaller overlaps (0.26 m smaller). The difference in distributions is noticeable for medium overlaps (0.8–1.6 m), while for small overlaps (<0.5 m) the distributions are relatively similar. This difference could be due to an evasive steering maneuver performed by the driver of the FV in the PCM events, resulting in a crash involving only one portion of the vehicle front bumper. In highD the cases with small overlaps are rarer, as the crashes were generated from normal highway driving scenarios, with the vehicles usually driving in the middle of the lane.

The highD overlaps represent normal driving behavior, so we believe that it is reasonable to virtually apply them to NDD near-crashes as part of the crash generation process. However, although the SHRP2 and GIDAS overlaps were similar, the limited event matching between highD and SHRP2 reduces the validity of applying the highD overlaps to SHRP2 data—yet another argument for aiming to use better matched datasets. In fact, the effect on the results if the assumption of similarity of overlaps is violated is not obvious. Actually, it may be counterintuitive: if the overlap in SHRP2 is smaller than that applied from highD (Fig. 2; requiring less steering to avoid a crash), the AEB would likely avoid fewer crashes (since the AEB response would be delayed)—and vice versa (a larger overlap would result in more avoided crashes).

4.2. The influence of data source choice on the comparison of AEB safety performance

4.2.1. Comparing crash characteristics across AEB algorithms

Two different AEB algorithms were applied to the crashes from the three datasets. The first algorithm, the RAEB, performs a run-time threat assessment based exclusively on the longitudinal kinematics of the two vehicles. When the algorithm detects that the FV will soon be unable to brake in time to avoid a crash, it initiates an automated braking maneuver. The second algorithm, the CAEB, goes a step further and accounts for the possibility that the driver will avoid the crash by comfortable steering. Numerous studies consider steering as an opportunity to avoid crashes (Brännström et al., 2010, 2014; Sander, 2018). The purpose of this addition to the AEB algorithm is to try to avoid as many unnecessary interventions as possible, as there could be cases in which the driver is aware of the LV and is planning to perform a late, but comfortable, steering maneuver: for example, when about to overtake.

On the one hand, the number of interventions that are too early to be accepted by the driver could potentially be reduced by considering comfortable steering, but on the other hand these delayed AEB interventions result in more, or more severe, crashes (i.e., increased impact speed; see Fig. 6a,b). That is, if the FV driver does not perform the expected steering maneuver, the vehicle's AEB system may no longer have the time to brake to avoid the collision. This outcome is also shown in Fig. 6c, where the CAEB is less effective at avoiding crashes than the RAEB. Both the fewer avoided crashes and the less reduced impact speed are direct consequences of the delayed AEB intervention. This finding was expected; safety system manufacturers are constantly balancing the potential improvement in user acceptance (Bliss & Acton, 2003; Coelingh et al., 2007) against the decreased effectiveness of the safety system at reducing the number and severity of rear-end crashes.

If crashes generated from near-crashes are to be used for AEB safety assessment, it is important to perform sensitivity analyses on the effects of LV decelerations, and differences in performance between the car-following versus FV-catch-up-to-LV (or cut-in) scenarios, on the preventive systems performance; our analyses show that the differences between the PCM and near-crash generated crashes are large with respect to LV decelerations and scenario type, also for near-crashes. This will likely have different impact on the safety performance assessment, depending on the AEB algorithm. Possibly probability weighting on, for example, scenario type, based on more representative crash databases, may mitigate potential effects of data source differences with respect to safety performance.

4.2.2. False positive analysis

Analysis of false positives is an important part of the assessment of preventive safety systems (Bliss & Acton, 2003; Coelingh et al., 2007; Sander & Lubbe, 2016). With respect to our false positive analysis, it is important to note that when an AEB system was triggered for a near-crash in the SHRP2 data, the situation may not actually have been a false positive, as it actually was quite critical. However, for consistency we decided to consider all AEB triggers as false positives in our analysis and discussion.

There were substantial differences between highD and SHRP2 datasets in the results from the simulation and analysis of false positives. The few highD false positives were all cases in which the FV was following the LV closely at a very low speed, representative of a traffic jam—perhaps not very relevant for human injury prevention. The SHRP2 events, on the other hand, had more false positives, and the dynamics were more safety–critical (at least, the speeds were higher). These differences mirror the near-crash nature of the SHRP2 events compared to the normal highway driving in highD (see Table 1).

The number of false positive activations was smaller with CAEB than RAEB, as logic suggests, although the reduction was substantially less than expected. However, we do show that crashes generated from NDD near-crashes are potentially a good source of data for false positive analysis. In contrast, highD, which contains approximately 17 hours of data (60 recordings of 17 minutes) with approximately 30 vehicles in the image at all times and records FV/LV interactions in normal driving data, appears not to be very useful for understanding even AEB false positive performance – if an AEB system triggered in such events, it would truly be a poorly designed system.

4.3. Limitations and future work

A main limitation of this study is the assumption that FV drivers are not reacting at all to LV braking. That is, we are basically assuming the drivers are sleeping. However, it is possible to virtually add glance behaviors and brake responses to generated events (Bärgman et al., 2017; Bärgman, Lisovskaja, et al., 2015; Lee, Lee, Bärgman, Lee, & Reimer, 2018), using distributions of documented glance behaviors (Morando, Victor, & Dozza, 2019) and driver response models (Markkula et al., 2016; Svärd et al., 2017). This aggregation of knowledge about driver behavior enables the generation of synthetic crashes and counterfactual simulations that take other factors, such as driver distraction, into account (Bärgman, Lisovskaja, et al., 2015; Bärgman & Victor, 2020). However, although driver glance behavior models have been included in previous studies, the approach to apply glances to non-crashes as part of crash generation has not been systematically validated (using, e.g., in-depth crash data). Applying glances to non-crashes in a study similar to this one would bring us one step closer to validating the approach.

In addition to improving the generation of synthetic crashes by adding driver behaviors, investigating different types of crash scenarios could be a next step—expanding from rear-end crashes to, for example, intersection crashes. There are site-based intersection NDD available from drone collections (Bock et al., 2020), similar to the highD data; however, given the results of our study, they would likely be of little use for safety assessment of, for example, AEB. Going forward, the focus must be on identifying and using data sources containing critical events, instead of using small samples of everyday driving. It may, however, still be relevant to use crashes generated also from drone based NDD, such as Bock et al. (2020), for methodological work related to safety assessment, where correct absolute safety benefit estimates are not the main focus.

Further, the simulations themselves are simplifications of reality. For example, the inclusion of more advanced vehicle models and mature AEB algorithms (e.g., in production) would likely improve the generalizability of the simulations. Future research would benefit from working closely with vehicle manufacturers, which have access to both of these. Also, for the RAEB algorithm, all vehicles were assumed to have the same values of jerk and (reachable) deceleration, regardless of the varying performances that different cars can have, and the weather conditions (e.g., rain and ice—braking performance was not tuned to reflect possible changes in road friction coefficient).

It is also important to consider what specific variables are important for specific benefit analysis. This study has shown that the lateral overlap between the LV and FV is important when using rear-end NDD to assess rear-end crashes. Consequently, its precise and accurate recording should be a priority in future data collection – also in critical event NDD recordings.

In this study the scope was to investigate how normal driving and near-crash NDD can (or cannot) be used to generate crashes for counterfactual safety assessment, with focus on AEB (which will be part of all levels of vehicle automation for the foreseeable future). We are not studying how NDD could be used in the process to generate crashes through, for example, traffic simulations. Research and development should continue for both the traffic simulations-based approach (typically targeting higher levels of automation) and for counterfactual safety assessment, with focus on ways of ensuring precise and accurate results, while making sure stakeholders easily can understand assumptions and limitations, and their implications on results and interpretation. Further, as needs now arise to assess the safety of higher levels of vehicle automation, estimating crash rates and exposure "of the future" is crucial. However, the method used in this study – counterfactual simulations (with or without crash generation) – do not attempt to predict the future rates. Instead the crash rates from, for example, crash databases would typically be used (i.e., simply using the original base rates per scenario type). Further development of methods to estimate future crash rates and scenario exposure is needed, where the use of traffic simulations is one path that should be pursued, with focus on validation.

Finally, another main limitation of our analysis is the limited matching of events across datasets. Although the comparability between the highD data and the GIDAS data is high, as both were collected on German highways, the types of scenarios differ substantially: the highD events are almost all car-following, while many of the GIDAS crashes consist of a car catching up to a lead vehicle at high relative speed. The SHRP2 and GIDAS data differ in a different way: the SHRP2 near-crashes rarely occurred on highways and the driving speeds were substantially lower than in GIDAS; there was also a large difference in the proportion of car-following versus catch-up scenarios. Further, the fact that highD is site-based NDD, while SHRP2 and GIDAS data are collected in-vehicle (SHRP2) and crash-based (GIDAS), is likely to affect comparability. Site-based NDD is unlikely to capture crashes that are related to driving context and infrastructure, while continuous in-vehicle and crash-based data collection capture crashes (and for SHRP2, critical events) in all contexts. In the SHRP2 database the events were all safety–critical near-crashes, but only a fraction of the events was at high speed, unlike highD and GIDAS PCM databases, which included mostly high-speed driving events. Also, our use of German crash data (GIDAS) and U.S. in-vehicle NDD (SHRP2) makes it difficult to determine if differences are primarily due to differences between regions, or if they are fundamental data source differences. However, the similarities in criticality indicate that some liberal generalizations (also between regions) on the relative effect of different AEB systems may be possible, in our case likely due to the driving cultures of Germany and the United States being relatively similar. The validity of such liberal generalization is, however, likely much dependent on the driving cultures of the involved regions (see, e.g., the comparison between China and the United States, Bianchi Piccinini, Engström, Bärgman, & Wang, 2017)). Consequently, future studies generating and analyzing more closely matched events, differing only in that some are generated from near-crashes and others from in-depth crash databases, would be most valuable.

5. Conclusion

In-depth crash data with reconstructed pre-crash kinematics can be used to develop both protective and preventive safety systems that are highly effective. Since such data are not available everywhere, alternative data sources are needed to make at least rough estimates of system performance in regions without them, as part of the system development process. This work studied the suitability of using easier-to-obtain non-crash naturalistic driving data (NDD) as a complementary data source for use in virtual assessment of preventive safety systems (specifically AEB).

Results show that virtual AEB assessments based on site-based NDD recordings of everyday driving on highways had neither the criticality nor the impact speed of assessments based on traditional pre-crash kinematics from in-depth reconstructions of crashes. We have consequently shown that site-based NDD that only capture a few tens of hours of normal driving are not suitable for assessing preventive safety performance, crash avoidance, or impact and injury risk reduction.

However, our results also show that the event criticality and the proportion of avoided crashes (but not impact speeds or impact speed reductions) were similar between crashes based on U.S. near-crashes and those based on a traditional German in-depth crash database. Therefore, critical-event near-crash data may be useful to complement in-depth crash data when comparing the safety benefit of different systems. The near crash data also allows an assessment of false-positive activations highlighting differences between systems. However, our results show that data sources that include original crashes, such as in-depth crash data, are still very important and preferred.

With respect to practical applications of our research, the results from our study can be used by system developers and researchers when deciding which data to use for virtual safety assessment (e.g., if NDD is an option). The paper further provides insights into the limitations of NDD for safety assessment, which is important to understand when NDD are considered for use in virtual safety assessment.

Funding

This work was supported by Autoliv Research and the MSc program Master Program of Automotive Engineering at Chalmers University of Technology.

Conflict of interest

Nils Lubbe works at Autoliv Research, located in Vårgårda, Sweden. Autoliv Research is part of Autoliv (www.autoliv.com), a company that develops, manufactures and sells, for example, protective safety systems to car manufacturers. Autoliv is a tier 1 supplier. Results from this study may impact how Autoliv choose to develop their products. Pierluigi Olleja and Jonas Bärgman have no interests to declare.

Acknowledgements

We want to thank Xiaomi Yang and Bo Sui for valuable discussions, and Hanna Jeppsson for her support on understanding the PCM data. Further, we want to thank the Institute for Automotive Engineering RWTH Aachen University, which owns the highD data, for providing access for this research project (the paper is partially based on work performed as part of the main authors' MSc thesis). We also want to thank VTTI, the Transportation Research Board, and the National Academies for providing access to the SHPR2 data. The data was made available under the data licensing agreement SHRP2-DUL-16-A-204. The findings and conclusions of this paper are those of the authors and do not necessarily represent the views of VTTI, the Transportation Research Board or the National Academies.

Appendix A

In the case of an intersection of the predicted FV path with the LV position, the relative longitudinal distance was measured following the FV's predicted path (along either the arc or the line). The lateral overlap is the lateral relative distance between the FV and the LV. That is, if all vehicles were of equal width and directly behind each other (complete overlap), there would be only one overlap value: the vehicle width. That is, the following driver needs to move a full vehicle width to the left (or right) to just barely avoid crashing, in a critical rear-end situation. Overlap values lower than the vehicle width occur when the LV and the FV are travelling at different lateral positions in the same lane. Overlap was computed by measuring the distance of the four vertices of the rectangle representing the LV to the centreline of the path for left (d_{lLV}) and right (d_{rLV}) sides of the LV and adding half the width of the FV (d_{lFV} or d_{rFV}, right and left half of the FV width, respectively, based on whether the steering maneuver is about to take place to the left or to the right of the LV) and an additional safety distance of one metre (e), so that once the FV has completed the lateral movement it has some lateral clearance to the LV, rather than almost touching

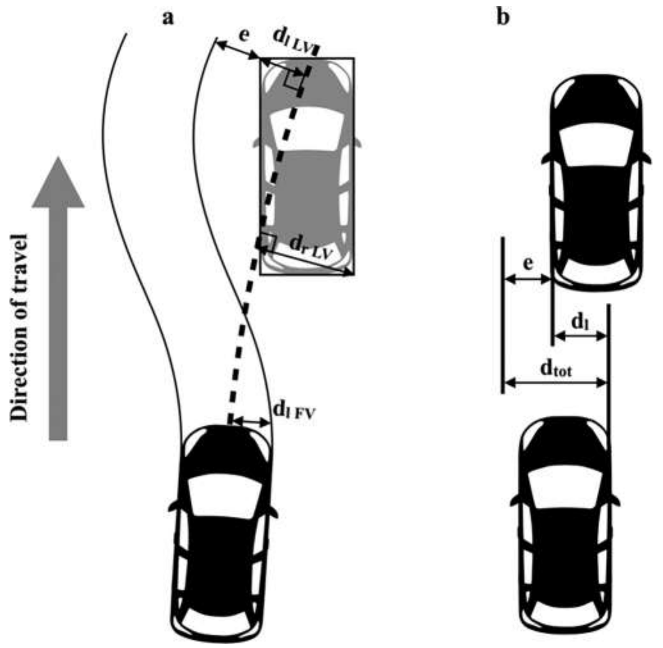

Fig. A1. Overlap measurements between LV and FV. The example shows two scenarios: a) FV steering right towards LV with a curved future path (dotted line); “$d_{l\ LV}$” and “$d_{r\ LV}$” are left and right distances from centreline of the future FV path to LV the furthest corners of the LV space to the left and the right, respectively; comfortable maneuver is chosen to be to the left of the LV (solid lines), moving laterally of “$d_l = d_{l\ LV} + d_{r\ LV} + e$”, where “e” is the additional safety distance; b) FV and LV parallel to each other; “d_{tot}” is the total clearance distance required to avoid a collision, comprising overlap “d_l” and “e”.

as it passes (see Fig. A1a,b). Fig. A1b shows how the total clearance distance (d_{tot}) was measured: it includes the overlap of the vehicles ($d_{lLV} + d_{lFV}$) and the additional safety distance (e). This procedure assured the availability of all the relevant metrics needed for the AEB application for the PCM crashes.

References

Alvarez, S., Page, Y., Sander, U., Fahrenkrog, F., Helmer, T., Jung, O., Hermitte, T., Düering, M., Döering, S., & Op den Camp, O. (2017). Prospective effectiveness assessment of ADAS and active safety systems via virtual simulation: A review of the current practices. *Proceedings of the 25th international conference on Enhanced Safety of Vehicles (ESV)* http://www-esv.nhtsa.dot.gov/Proceedings/25/25ESV-000346.pdf.

Bakker, J., Jeppsson, H., Hannawald, L., Spitzhüttel, F., Longton, A., & Tomasch, E. (2017). IGLAD - International Harmonized In-Depth Accident Data. In *Proceedings of the 25th international conference on Enhanced Safety of Vehicles (ESV)* (pp. 1–12).

Bareiss, M., Scanlon, J., Sherony, R., & Gabler, H. C. (2019). Crash and injury prevention estimates for intersection driver assistance systems in left turn across path/opposite direction crashes in the United States. *Traffic Injury Prevention, 20*(sup1), S133–S138. https://doi.org/10.1080/15389588.2019.1610945.

Bärgman, J., Boda, C. N., & Dozza, M. (2017). Counterfactual simulations applied to SHRP2 crashes: The effect of driver behavior models on safety benefit estimations of intelligent safety systems. *Accident Analysis and Prevention, 102*, 165–180. https://doi.org/10.1016/j.aap.2017.03.003.

Bärgman, J., Lisovskaja, V., Victor, T., Flannagan, C., & Dozza, M. (2015). How does glance behavior influence crash and injury risk? A “what-if” counterfactual simulation using crashes and near-crashes from SHRP2. *Transportation Research Part F: Traffic Psychology and Behavior, 35*, 152–169. https://doi.org/10.1016/j.trf.2015.10.011.

Bärgman, J., Smith, K., & Werneke, J. (2015). Quantifying drivers’ comfort-zone and dread-zone boundaries in left turn across path/opposite direction (LTAP/OD) scenarios. *Transportation Research Part F: Traffic Psychology and Behavior, 35*, 170–184. https://doi.org/10.1016/j.trf.2015.10.003.

Bärgman, J., & Victor, T. (2020). Holistic assessment of driver assistance systems: How can systems be assessed with respect to how they impact glance behavior and collision avoidance? *IET Intelligent Transport Systems*. https://doi.org/10.1049/iet-its.2018.5550.

Bianchi Piccinini, G., Engström, J., Bärgman, J., & Wang, X. (2017). Factors contributing to commercial vehicle rear-end conflicts in China: A study using on-board event data recorders. *Journal of Safety Research, 62*, 143–153. https://doi.org/10.1016/j.jsr.2017.06.004.

Bjorvatn, A., Page, Y., Fahrenkrog, F., Weber, H., Aittoniemi, E., Heum, P., Lehtonen, E., Silla, A., Bärgman, J., Borrack, M., Innamaa, S., Itkonen, T., Malin, F., Pedersen, K., Schuldes, M., Sintonen, H., Streubel, T., Hagleitner, W., Hermitte, T., ... Torrao, G. (2021). *L3Pilot deliverable D7.4 – impact evaluation results*. https://l3pilot.eu/fileadmin/user_upload/Downloads/Deliverables/Update_14102021/L3Pilot-SP7-D7.4-Impact_Evaluation_Results-v1.0-for_website.pdf.

Blatt, A., Pierowicz, J., Flanigan, M., Lin, P.-S., Kourtellis, A., Jovanis, P., Jenness, J., Wilaby, M., Campbell, J., Richard, C., Good, D., Czar, N., & Hoover, M. (2015). Naturalistic driving study: Field data collection. In *No. SHRP 2 Report S2-S07-RW-1*. https://doi.org/10.17226/22367.

Bliss, J. P., & Acton, S. A. (2003). Alarm mistrust in automobiles: How collision alarm reliability affects driving. *Applied Ergonomics, 34*(6), 499–509. https://doi.org/10.1016/j.apergo.2003.07.003.

Bock, J., Krajewski, R., Moers, T., Runde, S., Vater, L., & Eckstein, L. (2020). The inD dataset: A drone dataset of naturalistic road user trajectories at german intersections. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 1929-1934).

Brännström, M., Coelingh, E., & Sjöberg, J. (2010). Model-based threat assessment for avoiding arbitrary vehicle collisions. *IEEE Transactions on Intelligent Transportation Systems, 11*(3), 658–669. https://doi.org/10.1109/TITS.2010.2048314.

Brännström, M., Coelingh, E., & Sjöberg, J. (2014). Decision-making on when to brake and when to steer to avoid a collision. *International Journal of Vehicle Safety 1, 7*(1), 87–106. https://doi.org/10.1504/IJVS.2014.058243.

Brännström, M., Sjöberg, J., & Coelingh, E. (2008). A situation and threat assessment algorithm for a rear-end collision avoidance system. *IEEE Intelligent Vehicles Symposium, Proceedings, 102–107*. https://doi.org/10.1109/IVS.2008.4621250.

Cicchino, J. B. (2017). Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates. *Accident Analysis and Prevention, 99*, 142–152. https://doi.org/10.1016/j.aap.2016.11.009.

Coelingh, E., Jakobsson, L., Lind, H., & Lindman, M. (2007). Collision warning with auto brake: a real-life safety perspective. In *Innovations for safety: Opportunities and challenges.*

Davis, G. A., Hourdos, J., Xiong, H., & Chatterjee, I. (2011). Outline for a causal model of traffic conflicts and crashes. *Accident Analysis and Prevention, 43*(6), 1907–1919. https://doi.org/10.1016/j.aap.2011.05.001.

Dobberstein, J., Schmidt, D., Östling, M., Lindman, M., Bálint, A., & Höschele, P. (2021). Final results on detailed crash configurations from collisions expected to remain for automated vehicles, OSCCAR D1.3. www.osccarproject.eu.

Fahrenkrog, F., Wang, L., Platzer, T., Fries, A., Raisch, F., & Kompaß, K. (2019). Prospective safety effectiveness assessment of automated driving functions—From the methods to the results. *26th ESV conference*.

Fildes, B., Keall, M., Bos, N., Lie, A., Page, Y., Pastor, C., Pennisi, L., Rizzi, M., Thomas, P., & Tingvall, C. (2015). Effectiveness of low speed autonomous emergency braking in real-world rear-end crashes. *Accident Analysis and Prevention, 81*, 24–29. https://doi.org/10.1016/j.aap.2015.03.029.

Guo, F., Klauer, S. G., McGill, M. T., & Dingus, T. A. (2010). Evaluating the Relationship Between Near-Crashes and Crashes: Can Near-Crashes Serve as a Surrogate Safety Metric for Crashes? http://www.nhtsa.gov/DOT/NHTSA/NVS/Human%20Factors/Safety%20Problem%20Identification/DOT%20HS%20811%20382.pdf.

Hallerbach, S. (2020). Simulation-based testing of cooperative and automated vehicles. https://oops.uni-oldenburg.de/id/eprint/4649.

Hankey, J. M., Perez, M. A., & McClafferty, J. A. (2016). Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets. http://hdl.handle.net/10919/70850.

Helmer, T. (2014). *Development of a methodology for the evaluation of active safety using the example of preventive pedestrian protection*. Springer.

ISO (2021). ISO/TC 22/SC 36/WG 7, ISO/TR 21934-1 Road vehicles—Prospective safety performance assessment of precrash technology by virtual simulation—Part 1: State-of-the-art and general method overview.

Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems. In *IEEE conference on intelligent transportation systems*. https://doi.org/10.1109/ITSC.2018.8569552.

Krajewski, R., Moers, T., Bock, J., Vater, L., & Eckstein, L. (2020). The rounD dataset: A drone dataset of road user trajectories at roundabouts in Germany. In *2020 IEEE 23rd international conference on intelligent transportation systems, ITSC 2020*. https://doi.org/10.1109/ITSC45102.2020.9294728.

Kullgren, A., Krafft, M., Lie, A., & Tingvall, C. (2007). The effect of whiplash protection systems in real-life crashes and their correlation to consumer crash test programmes. In *Proceedings of the 20th international conference on Enhanced Safety of Vehicles (ESV)*, pp. 1–7. https://www-esv.nhtsa.dot.gov/Proceedings/20/07-0468-O.pdf.

Kullgren, A., Stigson, H., & Krafft, M. (2013). Development of whiplash associated disorders for male and female car occupants in cars launched since the 80s in different impact directions. In *Proceedings of IRCOBI conference* (pp. 51–62).

Kusano, K. D., & Gabler, H. C. (2012). Safety benefits of forward collision warning, brake assist, and autonomous braking systems in rear-end collisions. *IEEE Transactions on Intelligent Transportation Systems, 13*(4), 1546–1555. https://doi.org/10.1109/TITS.2012.2191542.

Laureshyn, A. (2010). *Application of automated video analysis to road user behavior*. Lund University (Doctoral dissertation) (Issue 253).
Lee, J. Y., Lee, J. D., Bärgman, J., Lee, J., & Reimer, B. (2018). How safe is tuning a radio? Using the radio tuning task as a benchmark for distracted driving. *Accident Analysis and Prevention, 110*, 29–37. https://doi.org/10.1016/j.aap.2017.10.009.
Lindman, M., & Tivesten, E. (2006). A method for estimating the benefit of autonomous braking systems using traffic accident data. *SAE Technical Papers, 2006*(724). https://doi.org/10.4271/2006-01-0473.
Markkula, G. (2015). Driver behavior models for evaluating automotive active safety: From neural dynamics to vehicle dynamics. https://research.chalmers.se/en/publication/212952.
Markkula, G., Engström, J., Lodin, J., Bärgman, J., & Victor, T. (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. *Accident Analysis and Prevention, 95*, 209–226. https://doi.org/10.1016/j.aap.2016.07.007.
McLaughlin, S. B., Hankey, J. M., & Dingus, T. A. (2008). A method for evaluating collision avoidance systems using naturalistic driving data. *Accident Analysis and Prevention, 40*(1), 8–16. https://doi.org/10.1016/j.aap.2007.03.016.
Morando, A., Victor, T., & Dozza, M. (2019). A reference model for driver attention in automation: Glance behavior changes during lateral and longitudinal assistance. *IEEE Transactions on Intelligent Transportation Systems, 20*(8), 2999–3009. https://doi.org/10.1109/TITS.2018.2870909.
NHTSA (2006). Event Data Recorders (EDRs).
NHTSA (2019). Traffic safety facts 2017: A compilation of motor vehicle crash data, pp. 24.
Otte, D., Krettek, C., Brunner, H., & Zwipp, H. (2003). Scientific approach and methodology of a new in-depth- investigation study in Germany so called GIDAS. *Proceedings of the 18th international conference on Enhanced Safety of Vehicles (ESV)* http://www-esv.nhtsa.dot.gov/Proceedings/18/18ESV-000161.pdf.
Rosén, E. (2013). Autonomous emergency braking for vulnerable road users. In *Proceedings of IRCOBI conference* (pp. 618–627).
Šajn, N. (2019). *General safety of vehicles and protection of vulnerable road user (Issue April 2019)*. European Parliament | Members' Research Service.
Sander, U. (2018). *Predicting Safety Benefits of Automated Emergency Braking at Intersections Virtual simulations based on real-world accident data*. Chalmers University of Technology https://research.chalmers.se/publication/504728.
Sander, U., & Lubbe, N. (2016). Prediction of accident evolution by diversification of influence factors in computer simulation: Opportunities for driver warnings in intersection accidents. *Aktive Sicherheit Und Automatisiertes Fahren - Methodenentwicklung Im Expertendialog, 29*.
Sander, U., & Lubbe, N. (2018). Market penetration of intersection AEB: Characterizing avoided and residual straight crossing path accidents. *Accident Analysis and Prevention, 115*, 178–188. https://doi.org/10.1016/j.aap.2018.03.025.
Scanlon, J. M., Kusano, K. D., Daniel, T., Alderson, C., Ogle, A., & Victor, T. (2021). Waymo Simulated Driving Behavior in Reconstructed Fatal Crashes within an Autonomous Vehicle Operating Domain. *Waymo LLC*. https://storage.googleapis.com/waymo-uploads/files/documents/Waymo-Simulated-Driving-Behavior-in-Reconstructed-Collisions.pdf.
Scanlon, J. M., Kusano, K. D., & Gabler, H. C. (2015). Analysis of driver evasive maneuvering prior to intersection crashes using event data recorders. *Traffic Injury Prevention, 16*, S182–S189. https://doi.org/10.1080/15389588.2015.1066500.
Scanlon, J. M., Page, K., Sherony, R., & Gabler, H. C. (2016). Using event data recorders from real-world crashes to evaluate the vehicle detection capability of an intersection advanced driver assistance system. *SAE technical paper, 2016-01–1457*. https://doi.org/10.4271/2016-01-1457.
Schubert, A., Erbsmehl, C., & Hannawald, L. (2013). Standardized pre-crash-scenarios in digital format on the basis of the vufo simulation. 5th International Conference on ESAR "Expert Symposium on Accident Research".
Shaikh, J., & Sander, U. (2018). Creation of pre-crash time-series database for evaluation of active safety systems using RASSI data. *8th international conference on ESAR "Expert Symposium on Accident Research"*.
Smith, K., Thome, A., Blåberg, C., & Bärgman, J. (2009). An invariant may drive the decision to encroach at unsignalized intersections. *Proceedings of the 5th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*.
Spitzhüttl, F., Petzold, M., & Liers, H. (2015). PCM from iGLAD database. FAT Forschungsvereinigung Automobiltechnik.
Summala, H. (2007). Towards understanding motivational and emotional factors in driver behavior: Comfort through satisficing. In C. P. Cacciabue (Ed.), *Modelling driver behavior in automotive environments* (pp. 189–207). Springer.
Svärd, M., Markkula, G., Engström, J., Granum, F., & Bärgman, J. (2017). A quantitative driver model of pre-crash brake onset and control. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*, 339–343. https://doi.org/10.1177/1541931213601565.
Tanaka, S. (2015). Reconstructing accidents by simulation for developing active safety systems. In *Proceedings of the 3rd international symposium on future active safety technology towards zero traffic accidents* (pp. 597–603).
Thomson, R., Sandin, J., Bagdadi, O., Hjort, M., Augusto, B., & Andersson, H. (2013). EDR pre-crash data: Potential for applications in active safety testing. *Proceedings of the 23th international conference on Enhanced Safety of Vehicles (ESV)* https://www-esv.nhtsa.dot.gov/Proceedings/23/files/23ESV-000414.PDF.
UNECE (2019). Event Data Recorder (EDR) GRSG-116 Agenda item 17. April, pp. 3.
Van Nes, N., Christoph, M., Hoedemaeker, M., & Van Der Horst, R. A. (2013). The value of site-based observations complementary to naturalistic driving observations: A pilot study on the right turn maneuver. *Accident Analysis and Prevention, 58*, 318–329. https://doi.org/10.1016/j.aap.2013.06.026.
Victor, T., Dozza, M., Bärgman, J., Boda, C. N., Engström, J., Flannagan, C., Lee, J. D., & Markkula, G. (2015). Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk. *SHRP 2 report S2–S08A-RW-1*.
VTTI (2020). InSight data access website SHRP2 naturalistic driving study.
Liers, H. (2018) Traffic Accident Research in Germany and the German In-Depth Accident Study (GIDAS). Presented at ACMA-SIAM-VDA Conference On Safer & Sustainable Road Transportation, New Delhi, India. Available at: https://www.vufo.de/wp-content/uploads/2021/02/181120_SIAM_ACMA_VDA_Conference_Liers_GIDAS-Presentation.pdf.
VUFO (2020). PRE-CRASH-MATRIX (PCM) V5.0 FORMAT. https://www.vufo.de/pcm/?lang=en.
WHO (2018). Global status report on road safety 2018. https://www.who.int/publications-detail/global-status-report-on-road-safety-2018.
Woodrooffe, J., Blower, D., Bao, S., Bogard, S., Flannagan, C., Green, P. E., & LeBlanc, D. (2012). Performance characterization and safety effectiveness estimates of forward collision avoidance and mitigation systems for medium/heavy commercial vehicles.

Pierluigi Olleja received the B.Sc. and M.Sc. degree in automotive engineering from Politecnico di Torino, Italy, in 2018 and 2020, respectively. During his M.Sc. he spent one year in Sweden at Chalmers University of Technology in a study exchange programme, where he did his master's thesis at the Division of Vehicle Safety. From 2019 to 2021, he worked at the Division of Vehicle Safety as a project assistant in the Crash Analysis and Prevention group. In July 2021 he started his studies as a PhD candidate. His research focus is counterfactual simulations and driver behaviour modelling.

Jonas Bärgman is an associate professor at Chalmers University of Technology in Gothenburg, Sweden. He started his career in industry at Autoliv Research in 1997, but he has worked as a researcher at Chalmers since 2009. His research focuses on methods for virtual safety assessment of human behavior and in-vehicle technologies in traffic. His work include research into a range of sub-domains needed for virtual assessment, including development computational driver models and crash causation research needed for such development, as well as methodological aspects of integration of driver models in virtual safety assessment.

Nils Lubbe received his PhD degree in Engineering from Chalmers University of Technology. He started his career in vehicle safety engineering at Toyota Motor Europe and is currently Director of Research at Autoliv in Vårgårda, Sweden. His research interests include crash analysis and prevention as well as injury biomechanics.

Journal of Safety Research 83 (2022) 357–363

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Changing vehicles to reduce older driver fatalities: An effective approach?

Aimee E. Cox *, Jessica B. Cicchino, Eric R. Teoh

Insurance Institute for Highway Safety, 4121 Wilson Boulevard, 6th Floor, Arlington, VA 22203, United States

ARTICLE INFO

Article history:
Received 6 December 2021
Received in revised form 1 June 2022
Accepted 14 September 2022
Available online 23 September 2022

Keywords:
Side airbags
Electronic stability control
Vehicle size
Vehicle age

ABSTRACT

Introduction: Age-related frailty leaves older drivers with the greatest fatality risk when involved in a crash compared with younger demographics. This study explored how vehicle features differed between crash-involved older and middle-aged drivers and estimated how those differences contribute to excess older driver fatalities. *Methods:* We merged Florida's crash data from 2014–2018 with Insurance Institute for Highway Safety and Highway Loss Data Institute databases. We compared the distribution of passenger vehicle age, type, size, and safety features among crash-involved older drivers (ages 70 and older) with crash-involved middle-aged drivers (ages 35–54). From logistic regression models, we estimated declines in older driver fatalities if they drove vehicles like those driven by middle-aged drivers under all and side-impact crash scenarios. *Results:* Older drivers in crashes were more likely to be in vehicles that were lighter, older, and without standard electronic stability control, standard head-protecting side airbags, and ratings of good in two IIHS crash tests than middle-aged drivers. In adjusted models, the fatality risk for older drivers in all crashes was significantly higher when ESC was not standard (odds ratio [OR], 1.37; 95% confidence interval [CI], 1.12–1.68) or when driving small passenger cars relative to large SUVs (OR, 2.02; 95% CI, 1.25–3.26); in driver-side crashes, the fatality risk doubled when vehicles did not have standard head-protecting side airbags (OR, 2.03; 95% CI, 1.58–2.62). If older drivers drove vehicles similar to middle-aged drivers, we estimated 3.3% and 4.7% fewer deaths in all and side-impact crashes, respectively. *Conclusions:* These results contribute to evidence suggesting that newer, more crashworthy vehicles with crash mitigation features benefit older drivers because of their heightened risk of crash-related fatality. *Practical Applications:* At a minimum, older drivers should aim to drive equipped vehicles with widely available features proven to reduce fatalities.

1. Introduction

Police-reported crashes per vehicle miles traveled for older drivers ages 70 and older have improved drastically since their mid-1990s peak, with rates declining more among older drivers than middle-aged drivers through 2016–2017 (43% vs 13%) (Cox & Cicchino, 2021). The decline in crash involvements contributed to a 50% decline in fatal crashes per vehicle miles traveled among older drivers, but older driver deaths per 1,000 crashes declined just 17% during the same period (Cox & Cicchino, 2021). Crash survivability remains low among this demographic due to age-related fragility (Cicchino, 2015). Consequently, fatal crash rate increases beginning at age 70 persist and the risk of dying in a crash continues to climb after middle-age (Cox & Cicchino, 2021).

Although drivers in their 70s do not experience excess crashes per vehicle miles traveled compared with middle-aged drivers, certain crash scenarios become more common with rising age due to age-related cognitive and visual declines (Anstey et al., 2005; Cox & Cicchino, 2021; Lombardi et al., 2017; Owsley et al., 1991). Intersection-related crashes in which older drivers are the struck (rather than the striking) vehicle due to failure to yield or looking but not seeing are particularly common (Alam & Spainhour, 2008; Budd et al., 2012; Clarke et al., 2010; Lombardi et al., 2017; Mayhew et al., 2006; Stamatiadis et al., 1991; Stutts et al., 2009) with older drivers 76% more likely to be involved and nearly-seven times more likely to be killed than middle-aged drivers in left-turn crashes across the path of another vehicle (Cox et al., 2022). Countermeasures that target survivability and the unique crash scenarios older drivers face may help reduce fatalities, and the recent safety enhancements to vehicles show promise in further contributing to declines in motor-vehicle-crash fatality rates among this vulnerable population (Cicchino, 2015).

* Corresponding author.
E-mail address: acox@iihs.org (A.E. Cox).

https://doi.org/10.1016/j.jsr.2022.09.010
0022-4375/

Despite this potential, older drivers tend to drive vehicles that are less safe because they are older models, have fewer safety features, and are likely lighter than those driven by some younger demographics. A prevalence study using quasi-induced exposure methods discovered that median vehicle age increased from five years among drivers ages 25–64 to eight years among drivers ages 85 and older and that older drivers were less likely to drive vehicles equipped with safety features like electronic stability control (ESC) and side airbags than their middle-aged counterparts (Metzger et al., 2020). A recent survey found that drivers 70 and older reported keeping their vehicles longer and being less likely to have requested safety features like side airbags and crash avoidance technologies when purchasing their primary vehicle than middle-aged participants (Cox & Cicchino, 2022). Both studies found that older drivers primarily drove passenger cars, which tend to be smaller and lighter than other vehicle types (Cox & Cicchino, 2022; Metzger et al., 2020).

An overwhelming body of evidence has revealed that newer model year vehicles protect against serious and fatal injury among drivers of all ages (National Highway Traffic Safety Administration [NHTSA], 2013; Ryb et al., 2009; Ryb et al., 2011). More recently, Fausto and Tefft (2018) found comparable results among drivers ages 65 and older. In their study using nationally representative data from police-reported crashes and data from a census of fatal crashes from 2010–2015, newer vehicle age alone, defined as model year 2010 and newer, significantly lowered the risk of fatal injury compared with older model-year vehicles (Fausto & Tefft, 2018). However, the effect of vehicle age on driver fatality is confounded by advancements in both safety features and crashworthiness associated with newer vehicle models, which are the true drivers of motor-vehicle fatality reductions. Crashworthiness has steadily improved through vehicle design changes and the equipment of features like side airbags, which allow for better injury outcomes (Highway Loss Data Institute [HLDI], 2020b; Teoh & Lund, 2011). Driver death risk in 2009 model year vehicles was about 50% lower than in 1984 models and 8% lower than in 2008 models, reflecting the benefit of such design changes (Farmer & Lund, 2015). Furthermore, technologies like ESC and other crash avoidance systems can enhance driver safety by preventing or mitigating crashes, thus reducing motor-vehicle injuries (Benson et al., 2018; Cicchino, 2017; Farmer, 2010).

The impact that individual aspects of improved crashworthiness and the equipment of safety technologies or features have had on fatality outcomes among the general driving population are indisputable. The Insurance Institute for Highway Safety (IIHS) evaluates vehicle crashworthiness through crash testing, and each year the proportion of vehicles earning the top rating of good improves (HLDI, 2020b). Studies evaluating the real-world effects of the IIHS side and moderate overlap frontal crash tests found clear correlations between ratings and fatality outcomes of side and head-on collisions, respectively, with the degree of protection highest among vehicles rated good (Farmer, 2005; Teoh & Lund, 2011). ESC is particularly effective at reducing the risk of single-vehicle fatal crash involvement, and side airbags with head protection greatly reduce a driver's fatality risk in driver-side crashes (Farmer, 2010; McCartt & Kyrychenko, 2007). Although the changes that reduce fatalities have been incorporated across the fleet, larger, heavier vehicles offer more protection to their occupants than smaller, lighter vehicles, especially when struck by a heavier partner vehicle (Monfort & Nolan, 2019; Ossiander et al., 2014). It is clear how these advancements reduce fatalities among the general population, but less is known about how many of these advancements may benefit older drivers specifically. Regardless, it is troubling that older drivers tend to drive older vehicles absent of many of the features and characteristics shown to reduce motor-vehicle fatalities.

One aim of this study was to expand on Metzger et al. (2020) by using the unique vehicle information databases maintained by IIHS and HLDI to explore how the vehicle characteristics of curb weight and crash test ratings, in addition to ESC, side airbag type, vehicle type, and vehicle age, might also vary by driver age. Using the foundation that features associated with reduced fatality outcomes are less prevalent in vehicles driven by older drivers than in vehicles driven by middle-aged drivers, we explored how those differences might expressly contribute to excess fatality among older drivers at a population level. Since older adults experience the highest crash fatality rates of any driver population, encouraging older drivers to drive the safest vehicles attainable is a promising approach to offsetting the other effects of age-related fragility and the resulting crash-related fatalities and injuries in play, even if these effects cannot be eliminated (Cox & Cicchino, 2021; Metzger et al., 2020). We believe that this study will improve awareness of what those vehicles are and how much difference they could make.

2. Methods

2.1. Data sources

We obtained the files for police-reported motor-vehicle crashes for the years 2014–2018 from the Florida Department of Highway Safety and Motor Vehicles. We selected Florida out of convenience due to the considerable size of the state's older adult population. We decoded valid Vehicle Identification Numbers (VINs) to obtain vehicle make, series, and model year using VINDICATOR, a proprietary VIN-decoding software maintained by HLDI, an affiliate of IIHS. Information on the availability of side airbags and ESC, passenger vehicle type (car, SUV, minivan, pickup trucks) and size, and curb weight were obtained from a database on vehicle features maintained by HLDI. Another data set housed IIHS vehicle ratings in the IIHS moderate overlap frontal and side crash tests (good, acceptable, marginal, and poor). These data sets were merged with the crash data by make, series, and model year.

Side airbag availability was categorized as follows: (1) standard head protection including side airbags that protect the head, available as standard equipment regardless of torso protection; (2) standard torso protection including side airbags that protect the torso, available as standard equipment, with head-protecting airbags either optional or not available; (3) optional protection of the head, the torso, or both the head and torso, available as optional equipment (not available as standard equipment); and (4) not available. Most side airbags that protect the head also protect the torso. We used HLDI's vehicle size classifications for each vehicle type, which are determined based on a combination of vehicle shadow (overall length times width) and curb weight (HLDI, 2020a).

2.2. Analyses

A total of 1,521,128 drivers ages 35–54 and 70 and older and their crash-involved vehicles were included in the study. We limited analyses to drivers and excluded other occupants as not every crash-involved vehicle will have a passenger and because inconsistencies in passenger placement within a vehicle creates complexity in analyzing injury outcomes given the crash configuration. Vehicles analyzed were passenger vehicles regardless of driver fault status. We excluded drivers with nontraffic-related injuries and vehicles older than model year 1981 because VINDICATOR's cannot decode VINs before that model year. Analyses of rated vehicles in the side crash test excluded series with multiple ratings that varied based on optional equipment, and both analyses of crash test rat-

Journal of Safety Research 83 (2022) 357–363

ings excluded vehicle series that were unrated in the respective crash test.

We examined the frequency of vehicle characteristics of vehicle age, type and size combination, curb weight, availability of side airbags and ESC, and ratings in the IIHS moderate overlap frontal and side crash tests of crash-involved vehicles by age group. Drivers 70 and older were stratified into subgroups of drivers ages 70–74, 75–79, and 80 and older. We calculated rate ratios for categorical variables (relative proportion, computed as the frequency of vehicle feature among older driver subgroup divided by the frequency of vehicle feature among middle-aged drivers) to make comparisons across age groups. Confidence intervals (CI) for the rate ratios (RR) were computed using a normal distribution approximation given by the following equation:

$$95\%CI = e^{\ln(RR)\pm 1.96\sqrt{\frac{1}{a}-\frac{1}{a+b}+\frac{1}{c}-\frac{1}{c+d}}}$$

where *a, b, c,* and *d* are the frequencies in a 2 by 2 contingency table (Morris & Gardner, 1988). Differences in means of continuous variables across age groups were tested with two-sample *t* tests.

To estimate how differences in the vehicles driven by middle-aged and older drivers might contribute to excess fatality among older drivers, we developed a total of four logistic regression models, two each for drivers ages 70 and older and 80 and older. Logistic regression takes the form $\ln(p/(1\text{-}p)) = b_0 + b_1x_1 + b_2x_2 + \ldots b_kx_k$, where p is the probability of driver death, and b_i is the parameter estimate for the x_i covariate. The first set of models estimated declines in older driver fatalities in all crashes. Older drivers are overinvolved in side-impact crashes in which their vehicle is struck and have an especially elevated risk of dying in these crashes relative to middle-aged drivers (Cicchino, 2015; Li et al., 2003); so, the second set of regression models estimated reductions in driver-side crashes to focus on this vulnerable crash type. Given the high degree of multicollinearity of many of the vehicle-characteristic variables (e.g., vehicle size classifications are derived from curb weight, safety features are more common in newer vehicles), and because of the large proportion of crash test ratings that were either missing or a rating of good, the final model estimating the odds of driver fatality in all crashes included (1) ESC availability, (2) vehicle type and size combination, and (3) side airbag availability. The second model estimating the odds of older driver fatality in driver side-impact crashes was limited to (1) vehicle type and size combination and (2) side airbag equipment, because these features are most relevant to survivability in this crash configuration.

Two estimates of death risk for drivers ages 70 and older were calculated for all crashes—one based on the vehicles they drove, and the other based on vehicles driven by 35- to 54-year-olds. Using the adjusted model equation for drivers 70 and older, two sets of estimated logits were computed by multiplying parameter estimates and the population proportions of vehicle characteristics for nonreference categories and summing (i.e., the model estimate using average values for covariates); this was done once using the proportions of vehicle characteristics among 35- to 54-year-olds in all crashes and once using proportions among drivers 70 and older. These estimated logit values were transformed into risk estimates using the inverse logit function exp(logit)/(1 + exp(logit)). Then, the estimated risk for drivers 70 and older (had they driven vehicles similar to those driven by 35- to 54-year-olds) was taken as the estimated risk based on the vehicles of drivers 35- to 54-year-old divided by the estimated risk based on the vehicles of drivers 70 and older. We repeated this process to estimate risk for drivers 80 and older compared with 35- to 54-year-olds in all crashes, using the corresponding model equation and proportions of vehicle characteristics for nonreference categories. We then applied this process twice more for each older driver group compared with 35- to 54-year-olds for driver-side crashes only, using the resulting model equations and population proportions of vehicle characteristics for nonreference categories among those involved in driver-side crashes. This resulted in four estimates of fatality reduction potential, two each for drivers ages 70 and older and 80 and older for all crashes and for driver-side crashes. SAS 9.4 was used for all analyses.

3. Results

Table 1 displays the distribution of characteristics of all crash-involved vehicles by driver age, and rate ratios comparing frequencies for categorical variables or *t* tests comparing means of continuous variables among each older driver subgroup with middle-aged drivers. Drivers 70 and older were significantly more likely to be driving the oldest vehicles (16 years old and older) than middle-aged drivers (11.5–13.4% vs 10.4%) and were significantly less likely to be in the newest (<3 years old) vehicles (21.6–25.3% vs 25.9%). As driver age increased, vehicles were significantly less likely to be equipped with ESC as a standard feature (older: 51.1–58.5%; middle-aged: 59.4%). There was minimal variation in the availability of side airbags with standard head protection among drivers ages 35–54 and 70–74 (64.0–64.3%), but significant declines began at age 75 (63.1%) and fell to 61.1% among drivers 80 and older.

The prevalence of passenger cars rose significantly with driver age, from 51.3% among middle-aged to 68.5% among drivers 80 and older (Table 1). This pattern held for small and large cars, and the proportion of older drivers in midsize passenger cars rose beginning at age 80 and older. Similarly, the proportion of drivers in SUVs declined with age, from 28.9% among middle-aged to 19.3% among age 80 and older. While older drivers were similarly or more likely to be in small SUVs than their middle-aged counterparts, the proportion of drivers in midsize SUVs began to decline at age 75 (75+: 9.7–13.2%; middle-aged and 70–74: 14.7%) and older drivers were much less likely than middle-aged drivers to be in large SUVs (1.8–3.9% vs 6.5%). Average curb weight steadily and significantly declined with driver age, from a mean of 3,730 pounds among middle-aged drivers to 3,470 pounds among drivers 80 and older.

Table 2 displays the frequencies and rate ratios of moderate overlap frontal and side crash ratings across driver age, among vehicle models tested by IIHS. Crash-involved drivers 75 and older were significantly less likely to be in vehicles rated good on either test than middle-aged drivers. Drivers 75 and older were also significantly more likely to be in vehicles with a poor rating in the side crash test (ranging from 9.0–11%) compared with drivers ages 35–54 and 70–74 (8.0–8.1%).

Table 3 presents results from the two adjusted logistic regression equations modeling odds of driver fatality for drivers 70 and older and 80 and older in all crashes. Relative to large SUVs, all other vehicle type and size combinations were associated with increased odds of older driver fatality, with small passenger cars and pickup trucks statistically significant for drivers 70 and older. Vehicles lacking standard ESC were significantly associated with 37% higher odds of older driver fatality for drivers 70 and older and 32% higher odds for drivers 80 and older, relative to when ESC was standard.

Table 4 displays the results from the two adjusted logistic regression equations modeling odds of driver fatality among drivers 70 and older and 80 and older in driver-side crashes. Drivers in vehicles without standard head-protecting side airbags had over double the odds of fatality than those with (70 and older: 103% higher odds, 80 and older: 107% higher odds), and these were statistically significant for both age groups.

Table 1
Distribution of characteristics of crash-involved vehicles in Florida from 2014–2018 by driver age, listed as percent or mean (SD) and comparison of features between older driver subgroups and drivers ages 35–54.

Characteristic	Ages 35–54 n = 1,238,443	Ages 70–74 n = 115,534		Ages 75–79 n = 78,421		Ages 80 and older n = 88,730	
	Percent	Percent	RR (95% CI)	Percent	RR (95% CI)	Percent	RR (95% CI)
Vehicle age (years)							
<3	25.9	25.3	0.98 (0.97, 0.99)*	24.1	0.93 (0.92, 0.94)*	21.6	0.83 (0.82, 0.84)*
3 to 5	18.4	18.1	0.98 (0.97, 1.00)	17.6	0.96 (0.94, 0.97)*	16.5	0.90 (0.89, 0.91)*
6 to 10	24.9	24.5	0.99 (0.95, 1.00)	24.5	0.99 (0.97, 1.00)	25.5	1.02 (1.01, 1.03)*
11 to 15	20.4	20.6	1.01 (1.00, 1.02)	21.4	1.05 (1.03, 1.06)*	23.1	1.13 (1.11, 1.14)*
16 and older	10.4	11.5	1.10 (1.08, 1.12)*	12.4	1.19 (1.17, 1.21)*	13.4	1.29 (1.26, 1.31)*
Vehicle type and size							
Passenger car	51.3	53.4	1.04 (1.04, 1.05)*	57.7	1.12 (1.12, 1.13)*	68.5	1.34 (1.33, 1.34)*
Small	18.5	19.8	1.07 (1.06, 1.09)*	20.4	1.10 (1.08, 1.12)*	21.4	1.16 (1.14, 1.17)*
Midsize	23.8	22.1	0.93 (0.92, 0.94)*	23.4	0.98 (0.97, 1.00)	27.5	1.15 (1.14, 1.17)*
Large	9.0	11.5	1.28 (1.26, 1.30)*	14.0	1.55 (1.52, 1.58)*	19.6	2.18 (2.15, 2.21)*
SUV	28.9	28.8	1.00 (0.98, 1.01)	26.4	0.91 (0.90, 0.92)*	19.3	0.67 (0.66, 0.68)*
Small	7.6	10.3	1.34 (1.32, 1.37)*	10.0	1.31 (1.28, 1.34)*	7.8	1.020 (1.00, 1.04)
Midsize	14.7	14.7	1.00 (0.98, 1.01)	13.2	0.90 (0.88, 0.91)*	9.7	0.66 (0.64, 0.67)*
Large	6.5	3.9	0.59 (0.57, 0.61)*	3.1	0.48 (0.46, 0.50)*	1.8	0.28 (0.27, 0.30)*
Pickup truck	14.6	11.90	0.81 (0.80, 0.83)*	10.0	0.68 (0.67, 0.69)*	6.6	0.45 (0.44, 0.46)*
Small	3.3	3.71	1.13 (1.10, 1.17)*	3.4	1.03 (0.99, 1.07)	2.6	0.79 (0.76, 0.82)*
Large	11.4	8.2	0.72 (0.71, 0.74)*	6.7	0.58 (0.56, 0.59)*	4.0	0.35 (0.34, 0.36)*
Minivan	5.1	5.8	1.14 (1.12, 1.17)*	5.9	1.16 (1.13, 1.19)*	5.6	1.09 (1.06, 1.12)*
Curb weight, Mean (SD)‡	3,730 (887.5)	3,616 (781.2)	*p* value < 0.0001	3,561 (735.6)	*p* value < 0.0001	3,470 (659.3)	*p* value < 0.0001
ESC							
Standard	59.4	58.5	0.98 (0.98, 0.99)*	56.1	0.94 (0.94, 0.95)*	51.1	0.86 (0.86, 0.87)*
Optional	10.1	10.8	1.07 (1.05, 1.08)*	11.8	1.17 (1.15, 1.19)*	14.2	1.41 (1.38, 1.43)*
Not available	30.5	30.8	1.01 (1.00, 1.02)	32.1	1.05 (1.04, 1.06)*	34.7	1.14 (1.13, 1.15)*
Side airbag†							
Standard head protection	64.0	64.3	1.00 (1.00, 1.01)	63.1	0.98 (0.98, 0.99)*	61.1	0.95 (0.95, 0.96)*
Standard torso protection	2.2	2.5	1.15 (1.11, 1.19)*	2.7	1.22 (1.17, 1.27)*	3.1	1.42 (1.36, 1.47)*
Any optional	18.5	18.0	0.97 (0.96, 0.99)*	18.8	1.02 (1.00, 1.03)	20.9	1.13 (1.11, 1.14)*
Not available	15.3	15.2	0.99 (0.98, 1.01)	15.4	1.01 (0.99, 1.03)	14.9	0.97 (0.96, 0.99)*
Missing	< 0.01	< 0.01		< 0.01		< 0.01	

Note: CI = confidence interval. RR = rate ratio. SD = standard deviation.
Percentages do not always sum to 100% due to rounding.
† Standard head protection includes airbags that protect the head and may offer torso protection. Torso protection offers no head protection.
* Denotes statistical significance at $\alpha = 0.05$.
‡ Results of mean comparisons presented as *p* value from *t* test.

Table 2
Distribution (percent) of moderate overlap frontal and side test crash ratings of crash-involved vehicles in Florida from 2014–2018 by driver age, comparison between older driver subgroup and drivers ages 35–54.

	Ages 35–54	Ages 70–74		Ages 75–79		Ages 80 and older	
Crash test rating	Percent	Percent	RR (95% CI)	Percent	RR (95% CI)	Percent	RR (95% CI)
Moderate overlap frontal	n = 904,475	n = 87,727		n = 59,898		n = 67,917	
Good	85.0	84.5	0.99 (0.99, 1.00)	84.0	0.99 (0.98, 0.99)*	83.0	0.98 (0.97, 0.98)*
Acceptable	9.0	9.7	1.08 (1.05, 1.10)*	10.2	1.14 (1.11, 1.17)*	11.6	1.29 (1.26, 1.32)*
Marginal	3.4	3.3	0.98 (0.94, 1.02)	3.2	0.95 (0.91, 0.99)*	2.8	0.82 (0.79, 0.86)*
Poor	2.7	2.6	0.97 (0.93, 1.02)	2.6	0.96 (0.92, 1.02)	2.6	0.99 (0.94, 1.03)
Side	n = 714,195	n = 68,709		n = 46,529		n = 52,121	
Good	82.0	82.1	1.00 (1.00, 1.00)	81.1	0.99 (0.98, 0.99)*	77.4	0.94 (0.94, 0.95)*
Acceptable	4.9	5.7	1.16 (1.12, 1.20)*	6.0	1.24 (1.19, 1.29)*	7.6	1.56 (1.51, 1.61)*
Marginal	5.1	4.2	0.82 (0.79, 0.85)*	3.9	0.76 (0.73, 0.80)*	4.0	0.78 (0.75, 0.82)*
Poor	8.0	8.1	1.01 (0.98, 1.03)	9.0	1.13 (1.09, 1.16)*	11.0	1.38 (1.34, 1.41)*

Note: CI = confidence interval. RR = rate ratio.
Percentages do not always sum to 100% due to rounding.
* Denotes statistical significance at $\alpha = 0.05$.

If the distribution of vehicles driven by crash-involved older drivers was the same as middle-aged drivers in terms of vehicle type and size and the availability of standard ESC and side airbags that protect the head, we estimated 3.3% and 4.8% fewer fatalities for drivers 70 and older and 80 and older, respectively, in all crashes (not shown in tables). If the distribution of vehicles driven by older drivers involved in driver-side crashes was the same as middle-aged drivers in terms of vehicle type and size and the availability of standard side airbags protecting the head, we estimated there would have been 4.7% fewer deaths for drivers 70 and older and 7.5% fewer deaths for drivers 80 and older (not shown in tables).

4. Discussion

This study strengthened Metzger et al. (2020) by supporting their findings that older drivers are more likely to drive older passenger vehicles that are cars and without ESC and side airbags than middle-aged drivers. We also expanded on their study by demon-

Table 3
Logistic regression analyses of driver fatality in all crashes in Florida from 2014–2018, ages 70 and older and 80 and older.

Parameter	Ages 70 and older		Ages 80 and older	
	Estimate (SE)	Odds ratio (95% CI)	Estimate (SE)	Odds ratio (95% CI)
Intercept	−6.3439 (0.2383)		−5.7514 (0.4117)	
Electronic stability control (ESC)				
Standard (ref)		1.00 (ref)		1.00 (ref)
Not available/optional	0.3170 (0.1026)	1.37 (1.12, 1.68)*	0.2807 (0.1419)	1.32 (1.01, 1.75)*
Vehicle type and size				
Sport utility vehicle				
Large (ref)		1.00 (ref)		1.00 (ref)
Midsize	0.2467 (0.2579)	1.28 (0.77, 2.12)	0.1373 (0.4428)	1.15 (0.48, 2.73)
Small	0.3722 (0.2628)	1.45 (0.87, 2.43)	0.3210 (0.4451)	1.38 (0.58, 3.30)
Passenger car				
Large	0.4507 (0.2510)	1.57 (0.96, 2.57)	0.3185 (0.4233)	1.38 (0.60, 3.15)
Midsize	0.3382 (0.2470)	1.40 (0.86, 2.28)	0.2432 (0.4204)	1.28 (0.56, 2.91)
Small	0.7027 (0.2451)	2.02 (1.25, 3.26)*	0.5682 (0.4196)	1.77 (0.78, 4.02)
Minivan	0.3778 (0.2724)	1.46 (0.86, 2.49)	0.3547 (0.4514)	1.43 (0.59, 3.45)
Pickup truck (small and large)	0.5206 (0.2551)	1.68 (1.02, 2.78)*	0.3807 (0.4435)	1.46 (0.61, 3.49)
Side airbag†				
Standard head (ref)		1.00 (ref)		1.00 (ref)
None/optional/chest	0.1593 (0.1014)	1.17 (0.96, 1.43)	0.0985 (0.1390)	1.10 (0.84, 1.45)

* Denotes statistical significance at $\alpha = 0.05$.
† Standard head protection includes airbags that protect the head and may offer torso protection. Torso protection offers no head protection.

Table 4
Logistic regression analyses of driver fatality in driver-side crashes in Florida from 2014–2018, ages 70 and older and 80 and older.

Parameter	Ages 70 and older		Ages 80 and older	
	Estimate (SE)	Odds ratio (95% CI)	Estimate (SE)	Odds ratio (95% CI)
Intercept	−5.7743 (0.5055)		−4.8817 (0.7176)	
Vehicle type and size				
Sport utility vehicle				
Large (ref)		1.00 (ref)		1.00 (ref)
Midsize	0.0714 (0.5515)	1.07 (0.36, 3.17)	−0.2687 (0.7954)	0.76 (0.16, 3.63)
Small	0.3472 (0.5577)	1.42 (0.47, 4.22)	0.3831 (0.7746)	1.47 (0.32, 6.69)
Passenger car				
Large	0.6322 (0.5226)	1.88 (0.68, 5.24)	0.1215 (0.7356)	1.13 (0.27, 4.78)
Midsize	0.5470 (0.5165)	1.73 (0.63, 4.78)	0.2164 (0.7287)	1.24 (0.30, 5.18)
Small	0.6584 (0.5160)	1.93 (0.70, 5.31)	0.2145 (0.7314)	1.24 (0.30, 5.20)
Minivan	0.5966 (0.5547)	1.82 (0.61, 5.39)	0.3234 (0.7741)	1.38 (0.30, 6.30)
Pickup truck (small and large)	0.3177 (0.5447)	1.37 (0.47, 4.00)	−0.2600 (0.8083)	0.77 (0.16, 3.63)
Side airbag†				
Standard head (ref)		1.00 (ref)		1.00 (ref)
None/optional/chest	0.7082 (0.1292)	2.03 (1.58, 2.62)*	0.7270 (0.1736)	2.07 (1.47, 2.91)*

* Denotes statistical significance at $\alpha = 0.05$.
† Standard head protection includes airbags that protect the head and may offer torso protection. Torso protection offers no head protection.

strating that older drivers are more likely to drive lighter vehicles and less likely to drive vehicles with good ratings in the side and moderate overlap frontal crash tests. The features that are least prevalent in vehicles driven by older drivers, the driving population most vulnerable to frailty, have been associated with lower fatality risk in studies of the general population (Farmer, 2005, 2010; McCartt & Kyrychenko, 2007; Ossiander et al., 2014; Teoh & Lund, 2011). We established that this risk reduction extends to drivers 70 and older for ESC and bolstered previous research demonstrating that head-protecting side airbags decrease older driver death risk in driver-side crashes (McCartt & Kyrychenko, 2007). This builds on the work of Fausto and Tefft (2018), who found an association between vehicle age and fatality outcomes among older drivers by identifying how the absence of vehicle safety characteristics and features associated with newer vehicle age contribute to this relationship.

Encouraging older drivers to buy newer and different vehicles can be a useful strategy to further encourage fatality declines among this demographic. We have seen considerable reductions over the years in older driver crash involvement, which contributed to most of the decline in fatal crash involvement rates in the past decades for drivers 75 and older (Cicchino, 2015). However, progress in increasing crash survivability has not been as large for the oldest drivers; from the study periods of 2008–2009 and 2016–2017, drivers 80 and older experienced a reduction in their police-reported crash rate per mile traveled but no change in their risk of death when involved in a crash (Cox & Cicchino, 2021). We estimate that if older adults were driving vehicles of similar type and size and with similar safety features as the vehicles driven by middle-aged drivers, driver deaths could be reduced by approximately 3% for drivers 70 and older and by 5% for drivers 80 and older, which would translate to 93 lives saved when applied to fatally injured U.S. passenger vehicle drivers ages 70 and older in 2019. These reductions could be larger still if older drivers were in the safest vehicles rather than the vehicles driven by older people today. While the safest vehicles available cannot eliminate an older driver's risk of dying in a crash, these results suggest the strategy can achieve further reductions.

Older drivers are overrepresented in side-impact crashes at intersections, and their risk of dying in a crash is especially elevated in this configuration. For example, Cicchino (2015) reported that drivers 75 and older were about eight times as likely to die when involved in a side-impact crash relative to middle-aged drivers during 2005–2008 and about 4.5 times as likely in a frontal impact. Features that are protective in side impacts, such as side airbags with head protection and good side crash test ratings, are thus especially important to this demographic. Kahane (2013) found that side airbags reduce fatal injuries of front-seat occupants ages 70 and older significantly more than among younger occupants. Yet, drivers 70 and older are less likely to report having requested side or curtain airbags when purchasing their vehicles than middle-aged drivers (Cox & Cicchino, 2022). The lighter vehicles driven by older adults offer less occupant protection than other characteristically heavier vehicle types, especially when struck by a heavier crash partner (Monfort & Nolan, 2019; Ossiander et al., 2014). Our findings indicate that older people would see the largest reduction in fatality risk in side impacts when driving safer cars, especially drivers 80 and older, who would experience an estimated 8% fewer driver-side crash fatalities if they were in vehicles similar to those driven by middle-aged drivers.

There are several explanations for why older adults drive older and therefore less safe vehicles, none of which are mutually exclusive. Lower income and driving frequency are associated with older vehicle age, and since many older adults are on a fixed income and drive far less than they did in their working years, long vehicle retention is more common (Budd et al., 2012; Cox & Cicchino, 2022; Metzger et al., 2020). Some older adults also take a "retirement vehicle" approach, in which they purchase a vehicle upon retiring with the intent to keep it until they stop driving (Budd et al., 2012). We can expect fatality outcomes to improve with time as today's larger and safer new vehicles reach the older driver population, and we are beginning to see that unfold as differences in vehicles driven by middle-aged drivers and those aged 70–74 were minimal in this study. A U.S. federal mandate required ESC on all passenger vehicles beginning in September 2011, and side airbags are now standard equipment on nearly all new passenger vehicles due to changes in the federal side-impact protection regulations that took effect in 2010; so when an older adult buys a new "retirement vehicle" it will be equipped with both of these safety features (HLDI, 2020c; NHTSA, 2010). SUVs may provide better occupant protection than passenger cars due to their larger size and curb weight, are more likely to be driven by older drivers with higher income, and now account for the greatest proportion of the newest (<3 years old) vehicle types purchased by older drivers as they become more prevalent in the new vehicle fleet overall (Cox & Cicchino, 2022; Monfort & Nolan, 2019; Ossiander et al., 2014). In Florida, 2019 household income was about $20,000 less among householders ages 65 and older than those ages 45–64 and about $14,000 less than householders ages 25–44 (United States Census Bureau, n.d.), possibly contributing to the type and age of vehicles driven by older drivers in this study. However, it is also worth noting that SUVs, particularly large SUVs, might be impractical vehicle options for some older drivers because of age-related physical changes and subsequent difficulty with features characteristic of large SUVs like high entry height (Shaw et al., 2010) and their perceptions that essential driving tasks like parking and maneuvering are easier in passenger cars than SUVs (Cox & Cicchino, 2022).

Encouraging a change in purchasing patterns so that older adults do not stick with a single "retirement vehicle" throughout their remaining driving years may also be advantageous. When today's younger cohort of older drivers purchase new vehicles that they keep into old age, their vehicles' safety capabilities will eventually become inferior to those available in the new vehicles of the time. This means that the gap in safe vehicle ownership between age groups will persist if drivers hold onto these vehicles, and that the oldest drivers who could benefit from the safest vehicles the most because of their heightened fragility will continually be at a disadvantage. This is the case currently, as we saw that drivers 80 and older were driving vehicles most dissimilar from those driven by middle-aged drivers.

Limitations are worth noting. Driver or crash characteristics not accounted for in our analyses could have contributed to some of the increased fatality risk associated with vehicle characteristics or prevalence of features among crash-involved drivers by age group. There may be changes in how, when, and where vehicles are driven as they age (Blows et al., 2003; Poindexter, 2003); for example, older drivers who drive the least frequently are the most likely to drive older vehicles (Cox & Cicchino, 2022), and low-mileage drivers have a higher crash risk per mile traveled (Antin et al., 2017; Hakamies-Blomqvist et al., 2002; Janke, 1991). Pickup trucks were associated with an increased risk of dying in a crash despite their large size. Factors contributing to this effect could be that pickups are more likely to be driven in rural areas, where speeds are higher and crashes tend to be more severe (Zwerling et al., 2005), and that pickup drivers are less likely than drivers of other vehicle types to be belted (Goetzke & Islam, 2015). While our analyses of the risk of dying in a crash were limited to examining vehicle type and size and the availability of ESC and side airbags, additional vehicle safety characteristics introduced in recent years contribute to reduced motor-vehicle fatalities, so it is probable that improvements beyond ESC and side airbags that became more prevalent contemporaneously with these features contributed to the effects associated with them. Even with these limitations, this study provides evidence to support recommending that older drivers choose vehicles with safety features that have been proven to reduce fatality risk.

5. Conclusions

We hope that our findings promote an enhanced comprehension of and appreciation for the value that the safety features associated with newer vehicle age affords older drivers, encouraging their selection of the safest vehicles available and attainable. Maximizing the mobility of older drivers while protecting their safety is a continuing challenge, and this study highlighted the importance of informed vehicle selection among the most vulnerable vehicle occupants. Although crash rates of older adults have been consistently declining since the mid-1990s, driver fatality rates remain the highest in this demographic (Cox & Cicchino, 2021). Our results add to the growing body of evidence that suggests the benefit of newer vehicles by showing how improvements in crash outcomes can be achieved when older drivers choose vehicles with modern safety features.

6. Practical Applications

Older adults should embrace the vehicle safety features and technologies that have been proven to reduce occupant fatalities, and if possible, strive to purchase new vehicles equipped with the most up-to-date safety features available. For those whom the newest vehicles are unattainable, we encourage the safest vehicle affordable.

Conflict of Interest

None.

Acknowledgements

This work was supported by the Insurance Institute for Highway Safety.

References

Alam, B. M., & Spainhour, L. K. (2008). Contribution of behavioral aspects of older drivers to fatal traffic crashes in Florida. *Transportation Research Record: Journal of the Transportation Research Board, 2078*(1). https://doi.org/10.3141/2078-07.

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory, and physical factors enabling driving safety in older adults. *Clinical Psychology Review, 25*(1), 45–65. https://doi.org/10.1016/j.cpr.2004.07.008.

Antin, J. F., Guo, F., Fang, Y., Dingus, T. A., Perez, M. A., & Hankey, J. M. (2017). A validation of the low mileage bias using naturalistic driving study data. *Journal of Safety Research, 63*, 115–120. https://doi.org/10.1016/j.jsr.2017.10.011.

Benson, A. J., Tefft, B. C., Svancara, A. M., & Horrey, W. J. (2018). *Potential reductions in crashes, injuries and deaths from large-scale deployment of advanced driver assistance systems (Research Brief)*. AAA Foundation for Traffic Safety.

Blows, S., Ivers, R. Q., Woodwarnd, M., Connor, J., Ameratunga, S., & Norton, R. (2003). Vehicle year and the risk of car crash injury. *Injury Prevention, 9*, 353–356. https://doi.org/10.1136/ip.9.4.353.

Budd, L., Scully, M., Newstead, S. V., & Watson, L. (2012). *The potential crash and injury reduction benefits of safer vehicle choices for older drivers in Australia and New Zealand* (Report No. 315). Monash University Accident Research Centre.

Cicchino, J. B. (2015). Why have fatality rates among older drivers declined? The relative contributions of changes in survivability and crash involvement. *Accident Analysis & Prevention, 83*, 67–73. https://doi.org/10.1016/j.aap.2015.06.012.

Cicchino, J. B. (2017). Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates. *Accident Analysis & Prevention, 99A*, 142–152. https://doi.org/10.1016/j.aap.2016.11.009.

Clarke, D. D., Ward, P., Bartle, C., & Truman, W. (2010). Older drivers' road traffic crashes in the UK. *Accident Analysis & Prevention, 42*, 1018–1024. https://doi.org/10.1016/j.aap.2009.12.005.

Cox, A. E., & Cicchino, J. B. (2021). Continued trends in older driver crash involvement rates in the United States: Data through 2017–2018. *Journal of Safety Research, 77*, 288–295. https://doi.org/10.1016/j.jsr.2021.03.013.

Cox, A. E. & Cicchino, J. B. (2022). Older driver vehicle preferences and perceptions of safety: A survey. *Journal of Safety Research.*. https://doi.org/10.1016/j.jsr.2022.08.018.

Cox, A. E., Mueller, A. S., & Cicchino, J. B. (2022). *Safety potential of crash avoidance features, improved headlights, and V2V-enhanced technologies for older drivers.* Arlington, VA: Insurance Institute for Highway Safety.

Farmer, C. M. (2005). Relationships of frontal offset crash test results to real-world driver fatality rates. *Traffic Injury Prevention, 8*, 31–37. https://doi.org/10.1080/15389580590928981.

Farmer, C. M. (2010). *Effects of electronic stability control on fatal crash risk*. Insurance Institute for Highway Safety.

Farmer, C. M., & Lund, A. K. (2015). The effects of vehicle redesign on the risk of driver death. *Traffic Injury Prevention, 16*, 684–690. https://doi.org/10.1080/15389588.2015.1012584.

Fausto, B. A., & Tefft, B. C. (2018). Newer model years are associated with reduced risk of motor vehicle crash fatalities among older adults. *Transportation Research Record, 2672*(33), 101–108. https://doi.org/10.1177/0361198118798240.

Goetzke, F., & Islam, S. (2015). Determinants of seat belt use: A regression analysis with FARS data corrected for self-selection. *Journal of Safety Research, 55*, 7–12. https://doi.org/10.1016/j.jsr.2015.07.004.

Hakamies-Blomqvist, L., Raitanen, T., & O'Neill, D. (2002). Driver ageing does not cause higher accident rates per km. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(4), 271–274. https://doi.org/10.1016/S1369-8478(03)00005-6.

Highway Loss Data Institute (2020a). *HLDI automobile size and class definitions*.

Highway Loss Data Institute (2020b). IIHS crashworthiness evaluation programs and the U.S. vehicle fleet—A 2020 update. *Loss Bulletin, 37*(10).

Highway Loss Data Institute (2020c). Predicted availability and prevalence of safety features on registered vehicles—A 2020 update. *Loss Bulletin, 37*(11).

Janke, M. K. (1991). Accidents, mileage, and the exaggeration of risk. *Accident Analysis & Prevention, 23*(2–3), 183–188. https://doi.org/10.1016/0001-4575(91)90048-A.

Kahane, C. J. (2013). *Injury vulnerability and effectiveness of occupant protection technologies for older occupants and women* (Report No. DOT HS 811 766). National Highway Traffic Safety Administration.

Li, G., Braver, E. R., & Chen, L. (2003). Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. *Accident Analysis & Prevention, 35*, 227–235. https://doi.org/10.1016/S0001-4575(01)00107-5.

Lombardi, D. A., Horrey, W. J., & Courtney, T. K. (2017). Age-related differences in fatal intersection crashes in the United States. *Accident Analysis & Prevention, 99*(A), 20–29. https://doi.org/10.1016/j.aap.2016.10.030.

Mayhew, D. R., Simpson, J. M., & Ferguson, S. A. (2006). Collisions involving senior drivers: High risk conditions and locations. *Traffic Injury Prevention, 7*(2), 117–124. https://doi.org/10.1080/15389580600636724.

McCartt, A. T., & Kyrychenko, S. Y. (2007). Efficacy of side airbags in reducing driver deaths in driver side car and SUV collisions. *Traffic Injury Prevention, 8*, 162–170. https://doi.org/10.1080/15389580601173875.

Metzger, K. B., Sartin, E., Foss, R. D., Joyce, N., & Curry, A. E. (2020). Vehicle safety characteristics in vulnerable driver populations. *Traffic Injury Prevention, 21*(S1), S54–S59. https://doi.org/10.1080/15389588.2020.1805445.

Monfort, S. S., & Nolan, J. M. (2019). Trends in aggressivity and driver risk for cars, SUVs, and pickups: Vehicle incompatibility from 1989 to 2016. *Traffic Injury Prevention, 20*(S1), S92–S96. https://doi.org/10.1080/15389588.2019.1632442.

Morris, J. A., & Gardner, M. J. (1988). Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates. *British Medical Journal (Clinical Research Edition), 296*(6632), 1313–1316. https://doi.org/10.1136/bmj.296.6632.1313.

National Highway Traffic Safety Administration (2010). *Title 49 Code of Federal Regulations (CFR) Part 571 Section 214, side impact protection.* Office of the Federal Register, National Archives and Records Administration.

National Highway Traffic Safety Administration (2013). *How vehicle age and model year relate to driver injury severity in fatal crashes* (Report No. DOT HS 811 825). National Highway Traffic Safety Administration.

Ossiander, E. M., Koepsell, T. D., & McKnight, B. (2014). Crash fatality and vehicle incompatibility in collisions between cars and light trucks or vans. *Injury Prevention, 20*, 373–379. https://doi.org/10.1136/injuryprev-2013-041146.

Owsley, C., Ball, K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging, 6*(3), 403–415. https://doi.org/10.1037//0882-7974.6.3.403.

Poindexter, K. N. (2003). *Passenger vehicle crash involvement rates by vehicle model year* (Report No. DOT HS 809 691). National Highway Traffic Safety Administration.

Ryb, G. E., Dischinger, P. C., & Ho, S. (2009). Vehicle model year and crash outcomes: A CIREN study. *Traffic Injury Prevention, 10*(6), 560–566. https://doi.org/10.1080/15389580903271401.

Ryb, G. E., Dischinger, P. C., McGwin, G., & Griffin, R. L. (2011). Crash-related mortality and model year: Are newer vehicles safer? *Annals of Advances in Automotive Medicine, 55*, 113–121.

Shaw, L., Polgar, J. M., Vrkljan, B., & Jacobson, J. (2010). Seniors' perceptions of vehicle safety risks and needs. *American Journal of Occupational Therapy, 64*, 215–224.

Stamatiadis, N., Taylor, W. C., & McKelvey, F. X. (1991). Older drivers and intersection traffic control devices. *Journal of Transportation Engineering, 117*(3), 311–319. https://doi.org/10.1061/(ASCE)0733-947X(1991)117:3(311).

Stutts, J., Martell, C., & Staplin, L. (2009). *Identifying behaviors and situations associated with increased crash risk for older drivers* (Report No. DOT HS 811 093). National Highway Traffic Safety Administration.

Teoh, E. R., & Lund, A. K. (2011). IIHS side crash test ratings and occupant death risk in real-world crashes. *Traffic Injury Prevention, 12*, 500–507. https://doi.org/10.1080/15389588.2011.585671.

United States Census Bureau (n.d.). *2019 American Community Survey 1-year estimates.* https://data.census.gov/cedsci/table?q=median%20income&t=Age%20and%20Sex&g=0400000US12&tid=ACSST1Y2019.S1903.

Zwerling, C., Peek-Asa, C., Whitten, P. S., Choi, S.-W., Sprince, N. L., & Jones, M. P. (2005). Fatal motor vehicle crashes in rural and urban areas: Decomposing rates into contributing factors. *Injury Prevention, 11*, 24–28. https://doi.org/10.1136/ip.2004.005959.

Aimee Cox is a Research Associate at the Insurance Institute for Highway Safety. Since joining the Institute in 2020, she has authored papers on the safety of older drivers. She holds an M.P.H. in Epidemiology from the University of South Florida and a B.S. from Eastern Michigan University.

Jessica Cicchino is Vice President, Research at the Insurance Institute for Highway Safety. Since joining the Institute in 2012, she has authored papers on topics such as the effectiveness and use of advanced driver assistance systems, automated traffic enforcement, and the safety of older drivers, pedestrians, bicyclists, motorcyclists, and child passengers. Dr. Cicchino holds a B.A. from Vassar College and a Ph.D. from Carnegie Mellon University.

Eric Teoh is the Director of Statistical Services at the Insurance Institute for Highway Safety. Since joining the Institute in 2006, he has conducted numerous studies quantifying the state of highway safety and identifying ways to improve it. His research has focused on motorcycles, young drivers, large trucks, and passenger vehicle safety.

Journal of Safety Research 83 (2022) 419–426

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Special Report from the CDC

Disparities in traumatic brain injury-related deaths—United States, 2020 ☆

Alexis B. Peterson [a,*], Hong Zhou [b,#], Karen E. Thomas [b,#]

[a] Applied Sciences Branch, Division of Injury Prevention, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, United States
[b] Data Analytics Branch, Division of Injury Prevention, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, United States

ARTICLE INFO

Article history:
Received 6 October 2022
Accepted 6 October 2022
Available online 18 October 2022

Keywords:
Traumatic brain injury
Fatal injury
Heath disparities
Surveillance
ICD-10

ABSTRACT

Introduction: Traumatic brain injury (TBI) affects how the brain functions and remains a prominent cause of death in the United States. Although preventable, anyone can experience a TBI and epidemiological research suggests some groups have worse health outcomes following the injury. *Methods:* We analyzed 2020 multiple-cause-of-death data from the National Vital Statistics System to describe TBI mortality by geography, sociodemographic characteristics, mechanism of injury (MOI), and injury intent. Deaths were included if they listed an injury International Classification of Diseases, Tenth Revision (ICD-10) underlying cause of death code and a TBI-related ICD-10 code in one of the multiple-cause-of-death fields. *Results:* During 2020, 64,362 TBI-related deaths occurred and age-adjusted rates, per 100,000 population, were highest among persons residing in the South (20.2). Older adults ($\geq$75) displayed the highest number and rate of TBI-related deaths compared with other age groups and unintentional falls and suicide were the leading external causes among this older age group. The age-adjusted rate of TBI-related deaths in males was more than three times the rate of females (28.3 versus 8.4, respectively); further, males displayed higher numbers and age-adjusted rates compared with females for all the principal MOIs that contributed to a TBI-related death. American Indian or Alaska Native, Non-Hispanic (AI/AN) persons had the highest age-adjusted rate (29.0) of TBI-related deaths when compared with other racial and ethnic groups. Suicide was the leading external cause of injury contributing to a TBI-related death among AI/AN persons. *Practical application:* Prevention efforts targeting older adult falls and suicide are warranted to reduce disparities in TBI mortality among older adults and AI/AN persons. Effective strategies are described in CDC's Stopping Elderly Accidents, Deaths, & Injuries (STEADI) initiative to reduce older adult falls and CDC's Preventing Suicide: A Technical Package of Policy, Programs, and Practices for the best available evidence in suicide prevention.

1. Introduction

Although preventable, traumatic brain injury (TBI) has contributed to the deaths of more than one million Americans during the last 20 years (Daugherty, Waltzman, Sarmiento, & Xu, 2019). TBI is caused by a bump, blow, or jolt to the head or penetrating head injury and results in the disruption of normal brain function (Centers for Disease Control and Prevention [CDC], 2022). A TBI can be unintentional, self-inflicted, or result from an assault. Anyone is at risk for sustaining a TBI; however, epidemiological research suggests that not all persons who sustain one are affected equally and that some groups experience worse outcomes, such as death, following the injury (Daugherty et al., 2019; CDC, 2022; CDC, 2021; CDC, 2019). Age and sex differences in TBI-related deaths are well documented in the United States, with older adults aged $\geq$ 75 -years having the highest rates when compared with all other age groups. In addition, males have higher age-adjusted TBI death rates than females (CDC, 2022; CDC, 2019; Daugherty et al., 2019). Differences in fatal outcomes of TBI by U.S. geographic location are also well described, with people residing in rural areas displaying higher rates compared to those living in urban areas (Daugherty, Sarmiento, Waltzman, & Xu, 2022; Daugherty, Zhou, Sarmiento, & Waltzman, 2021). An examination of TBI-related deaths by race

The Journal of Safety Research has partnered with the Office of the Associate Director for Science, Division of Injury Prevention, National Center for Injury Prevention and Control at the CDC in Atlanta, Georgia, USA, to briefly report on some of the latest findings in the research community. This report is the 72nt in a series of "Special Report from the CDC" articles on injury prevention.

* Corresponding author at: 4770 Buford Highway, Mail Stop S106-9, Atlanta, GA 30341, United States.
E-mail address: APeterson4@cdc.gov (A.B. Peterson).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.jsr.2022.10.001
0022-4375/

and ethnicity revealed higher annual age-adjusted rates among American Indian/Alaska Native (AI/AN), non-Hispanic persons when compared to other racial and ethnic groups over an 18-year study period (2000–2017; Daugherty et al., 2019). In order to reduce disparities in TBI-related deaths among population subgroups (e.g., age, sex, race and ethnicity, geographical location) it is important to examine epidemiological data describing differences in injury intent and mechanisms of injury (MOI) and to incorporate this information into the development of culturally-responsive prevention strategies.

The leading causes of TBI-related deaths vary by population subgroup. In the United States, TBI-related deaths are most frequently caused by suicide (primarily driven by the underlying MOI of firearms), unintentional falls, and motor vehicle crashes (CDC, 2022; CDC, 2021; Daugherty et al., 2019). However, the leading MOIs of TBI-related deaths differ when epidemiological studies are stratified by age or race and ethnicity. During 2018 and 2019, the leading MOI of TBI-related deaths was unintentional falls among older U.S. adults aged ≥ 75 while motor vehicle crashes and homicide contributed to most TBI-related deaths among children aged birth to 17 years (CDC, 2022). A recent CDC analysis of death data from the National Vital Statistics System (NVSS) revealed that intentional injuries (i.e., suicide and homicide) contributed to higher rates of TBI-related deaths among White, non-Hispanic persons and Black, non-Hispanic persons, while unintentional injuries (e.g., motor vehicle crashes, falls) contributed to higher rates of TBI-related deaths among AI/AN, non-Hispanic persons and Hispanic persons during 2015 to 2017 (Daugherty et al., 2019). Understanding differences in the leading MOIs contributing to TBI-related deaths is paramount for developing prevention activities that target those at greatest risk of sustaining this injury. This epidemiological report describes health disparities in TBI-related deaths by geography, sociodemographic characteristics, MOI, and injury intent utilizing the most recent NVSS data available. We analyzed 2020 multiple-cause-of-death data from the NVSS. Public health officials may use the findings from this analysis to identify priority areas, such as suicide and unintentional falls, for TBI prevention programs aimed at populations disproportionately affected.

2. Methods

We examined TBI-related deaths using the 2020 NVSS multiple-cause-of-death files. NVSS is a partnership between the National Center for Health Statistics (NCHS) and state and local jurisdictions that results in the compilation of records of all deaths in the United States (NCHS, 2016a). Deaths were included if they had an underlying cause of death code of an injury (ICD-10 codes: V01–Y36, Y85–Y87, Y89, U01–U03, World Health Organization, 2022), and a TBI-related ICD-10 code in one of the 20 multiple-cause-of-death fields. The following TBI-related ICD–10 codes correspond to the established TBI death surveillance definition (CDC, 2019): S01, S02.0, S02.1, S02.3, S02.7-S02.9, S04.0, S06, S07.0, S07.1, S07.8, S07.9, S09.7-S09.9, T90.1, T90.2, T90.4, T90.5, T90.8, T90.9. For this analysis, the injury mechanism/intent categories of interest were motor vehicle traffic crashes, unintentional falls, unintentionally struck by or against an object, other or unspecified unintentional injury, all mechanisms of suicide (e.g., firearm, drowning, poisoning), all mechanisms of homicide, and other, which includes TBIs of undetermined intent and those caused by legal intervention or war. To identify the cause of injury, codes listed in Appendix A were searched for in the underlying cause of death field. These codes are consistent with the ICD-10 external cause of injury matrix (NCHS, 2021a). The public use multiple-cause-of-death file was used for most of this study (NCHS, 2022a). For the analysis stratified by region of decedents' residence, deaths were obtained from the multiple-cause-of-death files available through CDC WONDER (NCHS, 2022b). Regions include the Northeast, Midwest, South, and West as defined by the U.S. Census (United States Census Bureau, 2021).

Suicides among children < 10 years were not presented because it is unclear whether children < 10 are able to form suicidal intent (Crepeau-Hobson, 2010). Rates for suicides were age-adjusted to the 10 years and older population. Any suicides in the 0 to 9 years age group were moved to the "other" cause category so that the sum of causes equaled the total number of TBI-related deaths.

Rates were calculated using bridged race population estimates obtained from NCHS as the denominator (NCHS, 2021b). Age-adjusted rates were calculated by the direct method of age adjusting using the 2000 standard U.S. population (Klein & Schoenborn, 2001). While deaths are a complete census of all occurrences, confidence intervals were presented to account for random variation (Kochanek, Murphy, Xu, & Arias, 2019).

3. Results

During 2020, TBI contributed to 64,362 deaths, equating to approximately 176 TBI-related deaths each day in the United States (Table 1).

Among TBI-related deaths with known age, children from birth to 17 years accounted for 4.3 % (data not shown) of decedents. Older U.S. adults aged ≥ 75 years (76.8 per 100,000 population), 65–74 years (23.7), and those aged 55–64 years (19.3) had the highest rates of TBI-related deaths per 100,000 population. Males displayed an age-adjusted rate of TBI-related deaths that was more than three times that of females (28.3 versus 8.4 per 100,000 population, respectively). When compared to other racial and ethnic groups, AI/AN, non-Hispanic persons had the highest age-adjusted rate (29.0 per 100,000 population) while Asian/Pacific Islander, non-Hispanic persons had the lowest rates (7.7) of TBI-related deaths. Age-adjusted rates of TBI-related deaths were highest among persons residing in the South (20.2 per 100,000 population), followed by persons residing in the Midwest (19.2), West (17.0), and Northeast (12.7).

During 2020, more than half of TBI-related deaths (54 %, N = 34,715) were categorized as unintentional injuries (i.e., motor vehicle crashes, falls, struck by or against an object, other unintentional injury with mechanism unspecified), while 45 % (N = 28,649) were categorized as intentional injuries (i.e., all mechanisms of homicide and suicide). Suicide, unintentional falls, and motor vehicle crashes contributed to the highest age-adjusted rates of TBI-related deaths (7.3 per 100,000 population, 4.8 and 3.2, respectively) (Table 2).

The data show variation by age group when stratifying TBI-related deaths by injury intent and MOI. Motor vehicle crashes and homicides contributed to the highest rates of TBI-related deaths among children (analyzed separately) aged birth to 17 years (1.1 and 1.1 per 100,000 population, respectively; data not shown). Among older adults aged ≥ 75 years, unintentional TBIs, combined across MOI, contributed to rates of TBI-related deaths that were more than four times higher than those due to intentional injuries (63.3 and 13.3, respectively). This difference is particularly driven by the rate of unintentional falls (55.1) among older adults. Rates of TBI-related deaths attributable to motor vehicle crashes were the greatest cause of TBI among those aged 15–24 years (4.8) and 25–34 years (4.7).

Intentional injuries contributed to a higher age-adjusted rate of TBI-related deaths among males than unintentional injuries (14.5 per 100,000 compared with 13.4) (Table 3). This difference is particularly driven by males' rate of suicide (13.1). In contrast, among females, unintentional TBIs contributed to a higher age-adjusted rate of TBI-related deaths than intentional injuries (5.6 per 100,000 compared with 2.7). This difference is particularly driven by the rate of unintentional falls (3.4) among females. Males had higher total numbers and age-adjusted rates of all the examined causes that contributed to TBI-related deaths (e.g., motor vehicle crashes, falls, being struck by or against an object, suicide, homicide) compared with females (Table 3).

The leading cause of TBI-related death, with respect to intent and MOI, varied by race and ethnicity (Table 4). Among Black, non-Hispanic persons, intentional injuries contributed to a higher age-adjusted rate of TBI-related deaths than unintentional injuries, (11.4 compared with 8.4 per 100,000 population); this difference was driven by the age-adjusted rate of homicide (7.6). In contrast, unintentional injuries contributed to a higher age-adjusted rate of TBI-related deaths than intentional injuries among White, non-Hispanic (9.8 compared with 9.1 per 100,000 population), Hispanic all races (7.4 compared with 4.5), and Asian/Pacific Islander, non-Hispanic (5.3 compared with 2.2) persons. However, suicide was the leading cause of TBI-related deaths among White, non-Hispanic persons. Unintentional falls were the leading MOI of TBI-related deaths among Hispanic (3.7 per 100,000 population, age-adjusted rate) and Asian/Pacific Islander, non-Hispanic (3.7) persons. AI/AN, non-Hispanic persons had age-adjusted rates with overlapping confidence intervals for unintentional TBIs and intentional TBIs (14.4 and 13.1, respectively). Suicide was the leading cause of TBI-related deaths among this group.

4. Discussion

More than 64,000 TBI-related deaths occurred in the U.S. population in 2020, with rates varying by age group, sex, and race and ethnicity. The age-adjusted rate of TBI-related deaths per 100,000 population in 2020 (18.0 per 100,000, age-adjusted) represents a 6.5% increase from 2019 (16.9 age-adjusted rate; CDC, 2022). The highest rates occurred among older adults aged ≥ 75 years, males, and among AI/AN, non-Hispanic persons which is consistent with previous CDC surveillance reports (CDC, 2022; CDC, 2021) and epidemiological research (Daugherty et al., 2019). Children aged birth to 17 years accounted for less than 5% of all TBI-related deaths with known age, a finding consistent with previous CDC surveillance reports (CDC, 2022; CDC, 2021). Suicide and unintentional falls were the most common causes of TBI-related death in 2020. Further examination of TBI-related deaths by injury intent and within

Table 1
Number and rate* of traumatic brain injury-related deaths† by selected sociodemographic characteristics — National Vital Statistics System, United States, 2020.

Socio-demographic characteristics		Number	Rate* (95% CI)
Age (years)			
Birth-17		2,774	3.8 (3.7–4.0)
Birth-4		612	3.2 (2.9–3.4)
5–9		282	1.4 (1.2–1.6)
10–14		621	3.0 (2.8–3.2)
15–24		7,389	17.4 (17.0–17.8)
25–34		8,301	18.0 (17.6–18.4)
35–44		6,761	16.0 (15.7–16.4)
45–54		6,716	16.6 (16.2–17.0)
55–64		8,199	19.3 (18.9–19.8)
65–74		7,725	23.7 (23.2–24.3)
75+		17,755	76.8 (75.7–78.0)
Sex			
Male	**Crude**§	47,668	29.4 (29.1–29.6)
	Adjusted¶	47,667	28.3 (28.0–28.6)
Female	**Crude**§	16,694	10.0 (9.8–10.1)
	Adjusted¶	16,694	8.4 (8.3–8.6)
Race/ethnicity**			
White, non-Hispanic persons	**Crude**§	46,281	23.1 (22.9–23.3)
	Adjusted¶	46,281	19.2 (19.0–19.4)
Black, non-Hispanic persons	**Crude**§	8,852	20.3 (19.9–20.7)
	Adjusted¶	8,852	20.2 (19.8–20.7)
Hispanic persons	**Crude**§	6,683	10.9 (10.6–11.2)
	Adjusted¶	6,683	12.2 (11.9–12.5)
American Indian/Alaskan Native non-Hispanic persons	**Crude**§	785	28.4 (26.4–30.4)
	Adjusted¶	785	29.0 (26.9–31.0)
Asian/Pacific Islander, non-Hispanic persons	**Crude**§	1,626	7.6 (7.3–8.0)
	Adjusted¶	1,626	7.7 (7.3–8.0)
U.S census region of decedent's residence			
Northeast	**Crude**§	8,264	14.8 (14.5–15.1)
	Adjusted¶	8,264	12.7 (12.5–13.0)
Midwest	**Crude**§	14,413	21.1 (20.8–21.4)
	Adjusted¶	14,413	19.2 (18.9–19.5)
South	**Crude**§	27,317	21.6 (21.3–21.8)
	Adjusted¶	27,316	20.2 (20.0–20.5)
West	**Crude**§	14,368	18.3 (18.0–18.6)
	Adjusted¶	14,368	17.0 (16.7–17.3)
Total	**Crude**§	64,362	19.5 (19.4–19.7)
	Adjusted¶	64,361	18.0 (17.8–18.1)

Abbreviations: CI = confidence interval.
* Per 100,000 population.
† Record-Axis Condition codes were used (usually includes conditions listed in both Part I and Part II of the death certificate).
§ Deaths with missing age were included.
¶ Deaths with missing age were excluded. Rates were age-adjusted to the 2000 U.S. Census population using 12 age groups: 0–4, 5–9, 10–14, 15–19, 20–24, 25–34, 35–44, 45–54, 55–64, 65–74, 74–84, and ≥ 85 years.
** Deaths with unknown Hispanic ethnicity were excluded (135 deaths).

Table 2
Number and rate* of traumatic brain injury-related deaths† by intent, mechanism of injury, and age group—National Vital Statistics System, United States, 2020.

	Age group (years)																				
	0–4		**5–9**		**10–14**		**15–24**		**25–34**		**35–44**		**45–54**		**55–64**		**65–74**		**75+**		**Total (all ages)**	
Intent and Mechanism of Injury	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Rate (95% CI)**	**Count**	**Age-Adjusted Rate (95% CI)** §§
Total unintentional TBI-related deaths	**292**	**1.5 (1.3–1.7)**	**181**	**0.9 (0.8–1.0)**	**250**	**1.2 (1.1–1.4)**	**2,449**	**5.8 (5.5–6.0)**	**2,746**	**6.0 (5.7–6.2)**	**2,473**	**5.9 (5.6–6.1)**	**2,801**	**6.9 (6.7–7.2)**	**4,176**	**9.8 (9.5–10.1)**	**4,720**	**14.5 (14.1–14.9)**	**14,627**	**63.3 (62.3–64.3)**	**34,715**	**9.2 (9.1–9.3)**
Unintentional motor vehicle crashes	162	0.8 (0.7–1.0)	120	0.6 (0.5–0.7)	170	0.8 (0.7–0.9)	2,053	4.8 (4.6–5.0)	2,158	4.7 (4.5–4.9)	1,614	3.8 (3.6–4.0)	1,408	3.5 (3.3–3.7)	1,415	3.3 (3.2–3.5)	874	2.7 (2.5–2.9)	713	3.1 (2.9–3.3)	10,687	3.2 (3.1–3.3)
Unintentional falls§	14	–	–	–	–	–	71	0.2 (0.1–0.2)	209	0.5 (0.4–0.5)	421	1.0 (0.9–1.1)	845	2.1 (2.0–2.2)	1,956	4.6 (4.4–4.8)	3,194	9.8 (9.5–10.2)	12,728	55.1 (54.1–56.0)	19,454	4.8 (4.7–4.8)
Unintentionally struck by or against an object	–	–	–	–	–	–	20	0.0 (0.0–0.1)	41	0.1 (0.1–0.1)	43	0.1 (0.1–0.1)	57	0.1 (0.1–0.2)	69	0.2 (0.1–0.2)	56	0.2 (0.1–0.2)	58	0.3 (0.2–0.3)	365	0.1 (0.1–0.1)
Other unintentional injury, mechanism unspecified¶	106	0.5 (0.4–0.7)	49	0.2 (0.2–0.3)	65	0.3 (0.2–0.4)	305	0.7 (0.6–0.8)	338	0.7 (0.7–0.8)	395	0.9 (0.8–1.0)	491	1.2 (1.1–1.3)	736	1.7 (1.6–1.9)	596	1.8 (1.7–2.0)	1,128	4.9 (4.6–5.2)	4,209	1.2 (1.1–1.2)
Total intentional TBI-related deaths	**277**	**1.4 (1.3–1.6)**	**88**	**0.4 (0.3–0.5)**	**351**	**1.7 (1.5–1.9)**	**4,814**	**11.3 (11.0–11.6)**	**5,397**	**11.7 (11.4–12.0)**	**4,105**	**9.7 (9.4–10.0)**	**3,765**	**9.3 (9.0–9.6)**	**3,871**	**9.1 (8.8–9.4)**	**2,918**	**9.0 (8.6–9.3)**	**3,063**	**13.3 (12.8–13.7)**	**28,649**	**8.4 (8.3–8.5)**
Suicide (includes all mechanisms)	**	**	**	**	212	1.0 (0.9–1.2)	3,046	7.2 (6.9–7.4)	3,682	8.0 (7.7–8.3)	2,990	7.1 (6.8–7.4)	3,033	7.5 (7.2–7.8)	3,307	7.8 (7.5–8.1)	2,614	8.0 (7.7–8.3)	2,855	12.4 (11.9–12.8)	21,739	7.3 (7.2–7.4)
Homicide (includes all mechanisms)	277	1.4 (1.3–1.6)	88	0.4 (0.3–0.5)	139	0.7 (0.6–0.8)	1,768	4.2 (4.0–4.3)	1,715	3.7 (3.5–3.9)	1,115	2.6 (2.5–2.8)	732	1.8 (1.7–1.9)	564	1.3 (1.2–1.4)	304	0.9 (0.8–1.0)	208	0.9 (0.8–1.0)	6,910	2.2 (2.1–2.2)
Other††	**43**	**0.2 (0.2–0.3)**	**13**	**0.1 (0.0–0.1)**	**20**	**0.1 (0.1–0.1)**	**126**	**0.3 (0.2–0.3)**	**158**	**0.3 (0.3–0.4)**	**183**	**0.4 (0.4–0.5)**	**150**	**0.4 (0.3–0.4)**	**152**	**0.4 (0.3–0.4)**	**87**	**0.3 (0.2–0.3)**	**65**	**0.3 (0.2–0.4)**	**997**	**0.3 (0.3–0.3)**

Abbreviations: CI = confidence interval.
* Per 100,000 population.
† Record-Axis Condition codes were used (usually includes conditions listed in both Part I and Part II of the death certificate).
§ Falls of undetermined intent were not included.
–Suppressed for deaths <=10 and rates based on < 20 deaths.
¶ External cause of injury codes specify that the injury was unintentional but do not specify the actual mechanism of injury.
** Age < 10 years were excluded because determining intent in younger children can be difficult. Rates for TBI-related deaths due to suicide were age-adjusted to the population 10 years and older.
†† Includes TBIs of undetermined intent and those caused by legal intervention or war.
§§ Deaths with missing age were excluded. Rates were age-adjusted to the 2000 U.S. Census population using 12 age groups: 0–4, 5–9, 10–14, 15–19, 20–24, 25–34, 35–44, 45–54, 55–64, 65–74, 74–84, and ≥ 85 years.

Table 3
Number and age-adjusted rate* of traumatic brain injury-related deaths† by intent, mechanism of injury, and sex—National Vital Statistics System, United States, 2020.

	Sex			
	Male		Female	
Intent and Mechanism of Injury	Number	Rate (95% CI)	Number	Rate (95% CI)
Total unintentional TBI-related deaths	**22,764**	**13.4 (13.2–13.6)**	**11,951**	**5.6 (5.5–5.7)**
Unintentional motor vehicle crashes	7,957	4.8 (4.7–4.9)	2,730	1.6 (1.6–1.7)
Unintentional falls§	11,463	6.6 (6.5–6.7)	7,991	3.4 (3.3–3.4)
Unintentionally struck by or against an object	311	0.2 (0.2–0.2)	54	0.0 (0.0–0.0)
Other unintentional injury, mechanism unspecified¶	3,033	1.8 (1.7–1.8)	1,176	0.6 (0.6–0.6)
Total intentional TBI-related deaths	**24,144**	**14.5 (14.3–14.7)**	**4,505**	**2.7 (2.6–2.8)**
Suicide** (includes all mechanisms)	19,058	13.1 (13.0–13.3)	2,681	1.8 (1.7–1.9)
Homicide (includes all mechanisms)	5,086	3.2 (3.1–3.3)	1,824	1.1 (1.1–1.2)
Other††	**759**	**0.5 (0.4–0.5)**	**238**	**0.1 (0.1–0.2)**
Total	**47,667**	**28.3 (28.0–28.6)**	**16,694**	**8.4 (8.3–8.6)**

Abbreviation: CI = confidence interval.
* Per 100,000 population, age-adjusted to the 2000 U.S. standard population using 12 age groups: 0–4, 5–9, 10–14, 15–19, 20–24, 25–34, 35–44, 45–54, 55–64, 65–74, 74–84, and ≥ 85 years.
† Record-axis condition codes were used (usually includes conditions listed in both Part I and Part II of the death certificate).
§ Falls of undetermined intent were not included.
¶ External cause of injury codes specify that the injury was unintentional but do not specify the actual mechanism of injury.
** Age < 10 years were excluded because determining intent in younger children can be difficult. Rates for TBI-related deaths due to suicide were age-adjusted to the population 10 years and older.
†† Includes TBIs of undetermined intent and those caused by legal intervention or war.

each MOI also revealed differences by age group, sex, and race and ethnicity.

Older age is a known risk factor for TBI (Thompson, McCormick, & Kagan, 2006) and U.S. trauma centers are seeing a greater proportion of elderly patients with more comorbid diseases as the population ages (Dutton et al., 2010). Pre-existing comorbidities (e.g., diabetes mellitus, hypertension, coronary heart disease) at the time of the TBI are associated with reduced functional independence 2 to 4 years post-injury (Lecours, Sirois, Ouellet, Boivin, & Simard, 2012) and increased 1-year mortality (Selassie, McCarthy, Ferguson, Tian, & Langlois, 2005) among this older age group (Gardner, Dams-O'Connor, Morrissey, & Manley, 2018). Further, anticoagulant therapies (e.g., non-vitamin K oral anticoagulants, warfarin [Coumadin]) and aspirin are often routinely used to manage chronic conditions among older adults. In this older population, anticoagulant use can result in an increased likelihood of intracranial hemorrhage (Maegele et al., 2017) and further complications from TBIs. Unintentional falls were the most common MOI for TBI-related deaths among those aged ≥ 75 years. This MOI is most common among older adults due to functional declines including vision, muscle strength, and balance. This is consistent with older age being a known major risk factor for falls (Moreland, Kakara, and Henry, 2020; Jin, 2018; Ambrose, Cruz, & Paul, 2015). CDC's Stopping Elderly Accidents, Deaths, & Injuries (STEADI) can help health care providers incorporate fall prevention for older patients into their routine clinical practice. The STEADI initiative has three core components of screening older patients for fall risk, assessing modifiable risk factors, and intervening to reduce risk using effective clinical and community strategies (National Center for Injury Prevention and Control (NCIPC, 2019). Effective clinical and community strategies include reviewing and managing patient medications, physical therapy, and exercises that improve gait, balance, and strength (e.g., tai chi) (Stevens & Lee, 2018). The public can actively prevent falls by screening themselves or their loved ones for fall risk using the Falls Free Checkup[1]; talking to their health care provider about their, or their loved one's, fall risk; having an annual eye exam; performing balance and strength exercises; and working with an occupational therapist to modify the home to increase safety (e.g., remove tripping hazards).

Our study is congruent with previous epidemiological data that has consistently found higher counts and age-adjusted rates of TBI-related deaths among males compared with females for overall incidence and across all intentional and unintentional MOI. TBI research suggests males are more likely than females in the general adult population to sustain (Faul & Coronado, 2015) and die from a TBI (Hong et al., 2022). Systematic reviews of epidemiological research report an increased propensity to sustain more severe TBIs among males when compared with females (Chang, Guerriero, & Colantonio, 2015; Toccalino, Colantonio, & Chan, 2021), which may contribute to the higher age-adjusted rates of TBI-related deaths across all intentional and unintentional MOI examined in this study. For example, the higher age-adjusted rate of TBI-related deaths due to unintentional falls among males might be related to circumstances of the fall, such as a larger proportion of males falling from heights (e.g., ladders) (Timsina et al., 2017), which increases the likelihood of moderate to severe injuries, including TBI.

In the United States, racial and ethnic differences persist with respect to injury intent and MOI of TBI-related deaths, and our findings corroborate a recent epidemiological analysis of TBI-related deaths over an 18-year (2000 to 2017) study period (Daugherty et al., 2019). Intentional TBIs attributed to homicide disproportionately affected Black, non-Hispanic persons compared with all other racial/ethnic groups. The underlying reasons for this disparity are complex and likely include the person's opportunities for education and their economic and household stability, as well as physical characteristics of their built environment (Schleimer et al., 2022). Further, structural racism and longstanding systemic inequities among various and racial and ethnic groups, have resulted in limited economic, educational, and housing opportunities associated with inequities in risk for violence (Bailey, Krieger, Agenor, Graves, Linos, & Basset, 2017). Implementing evidence-based strategies for preventing violence before it begins can help decrease rates of TBI-related homicides. CDC's NCIPC has created resources[2] that outline the best available evidence-based strategies to be used in combination with a multilevel, multisector effort to prevent multiple types of violence (i.e., adverse childhood experi-

[1] Available here: Falls Free CheckUp (https://www.ncoa.org).

[2] Technical packages for violence prevention are available from: Technical Packages for Violence Prevention |Violence Prevention|Injury Center|CDC.

Table 4
Number and age-adjusted rate* of traumatic brain injury-related deaths† by intent, mechanism of injury, and race/ethnicity§—National Vital Statistics System, United States, 2020.

Intent and Mechanism of Injury	Race/ethnicity: White, non-Hispanic persons: Number	Rate (95% CI)	Black, non-Hispanic persons: Number	Rate (95% CI)	Hispanic persons: Number	Rate (95% CI)	American Indian/Alaskan Native non-Hispanic persons: Number	Rate (95% CI)	Asian/Pacific Islander, non-Hispanic persons: Number	Rate (95% CI)
Total unintentional TBI-related deaths	**25,875**	**9.8 (9.7–9.9)**	**3,545**	**8.4 (8.1–8.6)**	**3,722**	**7.4 (7.2–7.7)**	**385**	**14.4 (12.9–15.9)**	**1,121**	**5.3 (5.0–5.7)**
Unintentional motor vehicle crashes	6,640	3.3 (3.2–3.3)	1,914	4.3 (4.1–4.5)	1,694	2.7 (2.6–2.9)	167	6.0 (5.1–7.0)	248	1.1 (1.0–1.3)
Unintentional falls¶	15,893	5.2 (5.1–5.3)	1,121	2.8 (2.6–3.0)	1,501	3.7 (3.5–3.9)	155	6.0 (5.1–7.0)	756	3.7 (3.4–3.9)
Unintentionally struck by or against an object	284	0.1 (0.1–0.1)	18	–	51	0.1 (0.1–0.1)	–	–	–	–
Other unintentional injury, mechanism unspecified**	3,058	1.2 (1.2–1.3)	492	1.2 (1.0–1.3)	476	0.9 (0.8–1.0)	61	2.3 (1.7–2.9)	107	0.5 (0.4–0.6)
Total intentional TBI-related deaths	**19,886**	**9.1 (9.0–9.3)**	**5,095**	**11.4 (11.1–11.7)**	**2,769**	**4.5 (4.3–4.6)**	**361**	**13.1 (11.7–14.4)**	**476**	**2.2 (2.0–2.4)**
Suicide†† (includes all mechanisms)	17,716	9.4 (9.2–9.5)	1,714	4.5 (4.3–4.7)	1,676	3.2 (3.0–3.4)	238	9.9 (8.6–11.2)	350	1.9 (1.7–2.1)
Homicide (includes all mechanisms)	2,170	1.1 (1.1–1.2)	3,381	7.6 (7.3–7.8)	1,093	1.7 (1.6–1.8)	123	4.5 (3.7–5.4)	126	0.6 (0.5–0.7)
Other§§	**520**	**0.3 (0.2–0.3)**	**212**	**0.5 (0.4–0.5)**	**192**	**0.3 (0.3–0.4)**	**39**	**1.5 (1.1–2.0)**	**29**	**0.1 (0.1–0.2)**
Total	**46,281**	**19.2 (19.0–19.4)**	**8,852**	**20.2 (19.8–20.7)**	**6,683**	**12.2 (11.9–12.5)**	**785**	**29.0 (26.9–31.0)**	**1,626**	**7.7 (7.3–8.0)**

Abbreviation: CI = confidence interval.
* Per 100,000 population, age-adjusted to the 2000 U.S. standard population using 12 age groups: 0–4, 5–9, 10–14, 15–19, 20–24, 25–34, 35–44, 45–54, 55–64, 65–74, 74–84, and ≥ 85 years.
† Record-axis condition codes were used (usually includes conditions listed in both Part I and Part II of the death certificate).
§ Deaths with unknown Hispanic ethnicity were excluded (135 deaths).
¶ Falls of undetermined intent were not included.
–Suppressed for deaths <=10 and rates based on < 20 deaths.
** External cause of injury codes specify that the injury was unintentional but do not specify the actual mechanism of injury.
†† Age < 10 years were excluded because determining intent in younger children can be difficult. Rates for TBI-related deaths due to suicide were age-adjusted to the population 10 years and older.
§§ Includes TBIs of undetermined intent and those caused by legal intervention or war.

ences, child abuse and neglect, youth violence, intimate partner violence and sexual violence).

Among White, non-Hispanic and AI/AN, non-Hispanic persons, suicide contributed to the highest age-adjusted rates of TBI-related death and is consistent with previous data documenting its increasing prevalence among these populations when compared with other racial/ethnic groups (Ivey-Stephenson, Crosby, Jack, Haileyesus, Kresnow-Sedacca, 2017). The underlying reasons for this disparity are complex and can include alcohol and/or substance use dependence, mental and physical health problems, poverty, and taboos around seeking mental health support (systematic review by Odafe, Talavera, Soumia, Hong, & Walker, 2016). CDC encourages using the best available evidence-based strategies, such as identifying and supporting persons at risk of suicide, creating protective environments (e.g., reducing access to lethal means), teaching coping and problem-solving skills, and strengthening access and delivery of suicide care (Stone et al., 2017). CDC is currently working with funded partners, Southern Plains Tribal Health Board and Wabanaki Public Health and Wellness, to increase capacity to adapt, implement, and evaluate ongoing suicide prevention programs with the best available evidence (Stone et al., 2017). Understanding racial and ethnic differences in injury intent and MOI in TBI mortality is a first step in developing targeted prevention strategies for groups at high risk for this injury. Further, health care workers should recognize that racial and ethnic disparities persist within the full spectrum of the TBI experience starting with social and environmental factors and conditions leading to the injury, acute diagnosis and care, and rehabilitation through long-term health outcomes (Saadi, Bannon, Watson, & Vranceanu, 2022) and across health care more generally (NCHS, 2016b).

This study is subject to several limitations. First, in cases of multiple injuries, non-TBI diagnoses might have contributed to the deaths included in this analysis. Second, incomplete reporting or misclassification of the cause of death field on the death certificate could lead to underestimation or overestimation of TBI-related deaths. Third, the specificity of conclusions drawn regarding the leading contributors of TBI-related deaths is limited due to the broad categorization of the principal MOI. Fourth, race and Hispanic origin on death certificates can be misclassified, particularly for AI/AN, Asian/PI, and Hispanic populations (Arias, Heron, & Hakes, 2016; Arias, Xu, Curtin, Bastian, & Tejada, 2021). This can lead to an underestimation of TBI-related deaths among these groups.

Understanding the leading intentional and unintentional MOI of TBI-related deaths and identifying groups at increased risk is important in targeted prevention of this injury. Health care providers can play a critical role by assessing patients at increased risk (e.g., persons at risk for suicide or unintentional falls) and by providing evidence-based and culturally-responsive interventions or referrals when warranted.

Disclaimer

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

IRB Statement

The Centers for Disease Control and Prevention (CDC) reviewed this activity which was deemed not to be research; it was conducted consistent with applicable federal law and CDC policy (i.e., 45 C.F.R. part 46.102(l)(2), 21 C.F.R. part 56; 42 U.S.C. Sect. 241(d); 5 U.S.C. Sect. 552a; 44 U.S.C. Sect. 3501 et seq.).

Conflict of Interest

None.

Appendix A

Category of ICD-10 codes for mechanism of injury and/or injury intent.

Mechanism and/or intent	ICD-10 codes
Unintentional motor vehicle crashes	[V02–V04](0.1,0.9), V09.2, [V12–V14](0.3–0.9), V19 (0.4–0.6), [V20–V28](0.3–0.9), [V29–V79](0.4–0.9), V80 (0.3–0.5), V81.1, V82.1, [V83–V86](0.0–0.3), V87 (0.0–0.8), V89.2
Unintentional falls	W00–W19
Unintentionally struck by/ against	W20–W22, W50–W52
Unintentional, other	All other codes in the V01–X59, Y85–Y86 ranges
Suicide	U03, X60–X84, Y87.0
Homicide	U01–U02, X85–Y09, Y87.1
Other (undetermined intent or due to legal intervention or war)	Y10–Y34, Y87.2, Y89.9, Y35–Y36, Y89(0.0,0.1)

References

Ambrose, A. F., Cruz, L., & Paul, G. (2015). Falls and Fractures: A systematic approach to screening and prevention. *Maturitas, 82*(1), 85–93. https://doi.org/10.1016/j.maturitas.2015.06.035.

Arias E., Heron M., Hakes J., National Center for Health Statistics, US Census Bureau. (2016). The validity of race and Hispanic-origin reporting on death certificates in the United States: an update. *Vital & Health Statistics* 2, (172), 1–21. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_172.pdf.

Arias E., Xu J. Q., Curtin S., Bastian B., & Tejada Vera B. (2021). Mortality profile of the non-Hispanic American Indian or Alaska Native population, 2019. National Vital Statistics Reports; vol. 70 no 12. Hyattsville, MD: National Center for Health Statistics. https://dx.doi.org/10.15620/cdc:110370.

Bailey, Z. D., Krieger, N., Agénor, M., Graves, J., Linos, N., & Bassett, M. T. (2017). Structural racism and health inequities in the USA: Evidence and interventions. *Lancet, 389*, 1453–1463.

Centers for Disease Control and Prevention. (2019). Surveillance Report of Traumatic Brain Injury-related Emergency Department Visits, Hospitalizations, and Deaths—United States, 2014. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.

Centers for Disease Control and Prevention. (2021). Surveillance Report of Traumatic Brain Injury-related Hospitalizations and Deaths by Age Group, Sex, and Mechanism of Injury—United States, 2016 and 2017. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.

Centers for Disease Control and Prevention. (2022). Surveillance Report of Traumatic Brain Injury-related Deaths by Age Group, Sex, and Mechanism of Injury—United States, 2018 and 2019. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.

Chang, V. C., Guerriero, E. N., & Colantonio, A. (2015). Epidemiology of work-related traumatic brain injury: A systematic review. *American Journal of Industrial Medicine, 58*(4), 353–377. https://doi.org/10.1002/ajim.22418.

Crepeau-Hobson, F. (2010). The psychological autopsy and determination of child suicides: A survey of medical examiners. *Archives of Suicide Research, 14*(1), 24–34. https://doi.org/10.1080/13811110903479011. PMID: 20112141.

Daugherty, J., Sarmiento, K., Waltzman, D., & Xu, L. (2022). Traumatic Brain Injury-Related Hospitalizations and Deaths in Urban and Rural Counties-2017. *Annals of Emergency Medicine, 79*(3), 288–296.e1. https://doi.org/10.1016/j.annemergmed.2021.09.433.

Daugherty, J., Waltzman, D., Sarmiento, K., & Xu, L. (2019). Traumatic Brain Injury-Related Deaths by Race/Ethnicity, Sex, Intent, and Mechanism of Injury — United States, 2000–2017. *MMWR Morbidity and Mortality Weekly Report, 68*, 1050–1056. https://doi.org/10.15585/mmwr.mm6846a2.

Daugherty, J., Zhou, H., Sarmiento, K., & Waltzman, D. (2021). Differences in State Traumatic Brain Injury-Related Deaths, by Principal Mechanism of Injury, Intent, and Percentage of Population Living in Rural Areas - United States, 2016–2018. *MMWR Morbidity and Mortality Weekly Report, 70*(41), 1447–1452. https://doi.org/10.15585/mmwr.mm7041a3.

Dutton, R. P., Stansbury, L. G., Leone, S., Kramer, E., Hess, J. R., & Scalea, T. M. (2010). Trauma mortality in mature trauma systems: Are we doing better? An analysis of trauma mortality patterns, 1997–2008. *Journal of Trauma, 69*(3), 620–626. https://doi.org/10.1097/TA.0b013e3181bbfe2a.

Faul, M., & Coronado, V. (2015). Epidemiology of traumatic brain injury. *Handbook of Clinical Neurology, 127*, 3–13. https://doi.org/10.1016/B978-0-444-52892-6.00001-5.

Gardner, R. C., Dams-O'Connor, K., Morrissey, M. R., & Manley, G. T. (2018). Geriatric traumatic brain injury: Epidemiology, outcomes, knowledge gaps, and future directions. *Journal of Neurotrauma, 35*(7), 889–906. https://doi.org/10.1089/neu.2017.5371.

Hong, Z. J., Firek, M., Zachary, B., Mörs, K., Schindler, C., Marzi, I., ... Coimbra, R. (2022). The effect of age and sex on outcomes following isolated moderate to severe traumatic brain injury. *European Journal of Trauma and Emergency Surgery, 48*(2), 871–880. https://doi.org/10.1007/s00068-020-01491-1 https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htm.

Ivey-Stephenson, A. Z., Crosby, A. E., Jack, S. P. D., Haileyesus, T., & Kresnow-Sedacca, M. J. (2017). Suicide trends among and within urbanization levels by sex, race/ethnicity, age group, and mechanism of death—United States, 2001–2015. *MMWR Surveillance Summary, 66*(SS-18).

Jin, J. (2018). Prevention of falls in older adults. *JAMA, 319*(16), 1734. https://doi.org/10.1001/jama.2018.4396.

Klein R. J., & Schoenborn C. A. (2001). Age adjustment using the 2000 projected U.S. population. Healthy People Statistical Notes, no. 20. Hyattsville, Maryland: National Center for Health Statistics. https://www.cdc.gov/nchs/data/statnt/statnt20.pdf.

Kochanek K. D., Murphy S. L., Xu J. Q., & Arias, E. (2019). Deaths: Final data for 2017. National Vital Statistics Reports; vol. 68 no 9. Hyattsville, MD: National Center for Health Statistics. https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_09-508.pdf.

Lecours, A., Sirois, M. J., Ouellet, M. C., Boivin, K., & Simard, J. F. (2012). Long-term functional outcome of older adults after a traumatic brain injury. *Journal of Head Trauma Rehabilitation, 27*, 379–390.

Maegele, M., Schöchl, H., Menovsky, T., Marechal, H., Marlund, N., Buki, A., & Stanworth, S. (2017). Coagulopathy and haemorrhagic progression in traumatic brain injury: Advances in mechanisms, diagnosis, and management. *Lancet Neurology, 16*, 630–647. https://doi.org/10.1016/S1474-4422(17)30197-7.

Moreland, B., Kakara, R., & Henry, A. (2020). Trends in Nonfatal Falls and Fall-Related Injuries Among Adults Aged ≥65 Years - United States, 2012–2018. *MMWR Morbidity and Mortality Weekly Report, 69*(27), 875–881. https://doi.org/10.15585/mmwr.mm6927a5. PMID: 32644982; PMCID: PMC7732363.

National Center for Health Statistics (US). (2016b). Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities. Hyattsville (MD): National Center for Health Statistics (US), Report No.: 2016-1232. PMID: 27308685.

National Center for Health Statistics. (2016a). About the National Vital Statistics System. Available from https://www.cdc.gov/nchs/nvss/about_nvss.htm.

National Center for Health Statistics. (2021a). ICD–10: External cause-of-injury mortality matrix. Available from: https://www.cdc.gov/nchs/injury/injury_tools.htm.

National Center for Health Statistics. (2021b). Bridged-race population estimates—data files and documentation. Available from:

National Center for Health Statistics. (2022a). Mortality Multiple Cause File Downloadable Data Files. Available from: https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm#Mortality_Multiple.

National Center for Health Statistics. (2022b). Mortality Data on CDC WONDER. Available from: https://wonder.cdc.gov/mcd.html.

National Center for Injury Prevention and Control. (2019). Stopping Elderly Accidents, Deaths, and Injuries (STEADI) Initiative. Available from: https://www.cdc.gov/steadi/. Accessed on May 25, 2022.

Odafe, M. O., Talavera, D. C., Cheref, S., Hong, J. H., & Walker, R. (2016). Suicide in Racial and Ethnic Minority Adults: A Review of the Last Decade. *Current Psychiatry Reviews, 12*(2), 181–198.

Saadi, A., Bannon, S., Watson, E., & Vranceanu, A. M. (2022). Racial and Ethnic Disparities Associated with Traumatic Brain Injury Across the Continuum of Care: A Narrative Review and Directions for Future Research. *Journal of Racial and Ethnic Health Disparities, 9*(3), 786–799.

Schleimer, J. P., Buggs, S. A., McCort, C. D., Pear, V. A., Biasi, A., Tomsich, E., ... Wintemute, G. J. (2022). Neighborhood Racial and Economic Segregation and Disparities in Violence During the COVID-19 Pandemic. *American Journal of Public Health, 112*(1), 144–153. https://doi.org/10.2105/AJPH.2021.306540.

Selassie, A. W., McCarthy, M. L., Ferguson, P. L., Tian, J., & Langlois, J. A. (2005). Risk of posthospitalization mortality among persons with traumatic brain injury, South Carolina 1999–2001. *Journal of Head Trauma Rehabilitation, 20*, 257–269.

Stevens, J. A., & Lee, R. (2018). The potential to reduce falls and avert costs by clinically managing fall risk. *American Journal of Preventive Medicine, 55*(3), 290–297. https://doi.org/10.1016/j.amepre.2018.04.035.

Stone, D. M., Holland, K. M., Bartholow, B. N., Crosby, A. E., Davis, S. P., & Wilkins, N. (2017). *Preventing suicide: A technical package of policies, programs, and practice*. Atlanta, GA: US Department of Health and Human Services, CDC https://www.cdc.gov/suicide/pdf/suicideTechnicalPackage.pdf.

Thompson, H. J., McCormick, W. C., & Kagan, S. H. (2006). Traumatic brain injury in older adults: Epidemiology, outcomes, and future implications. *Journal of the American Geriatrics Society, 54*, 1590–1595. https://doi.org/10.1111/j.1532-5415.2006.00894.

Timsina, L. R., Willetts, J. L., Brennan, M. J., Marucci-Wellman, H., Lombardi, D. A., Courtney, T. K., & Verma, S. K. (2017). Circumstances of fall-related injuries by age and gender among community-dwelling adults in the United States. *PLoS ONE, 12*, e0176561.
Toccalino, D., Colantonio, A., & Chan, V. (2021). Update on the epidemiology of work-related traumatic brain injury: A systematic review and meta-analysis. *Occupational and Environmental Medicine, 78*, 769–776.
United States Census Bureau. (2021). 2010 Census regions and divisions of the United States. Available from: https://www.census.gov/geographies/reference-maps/2010/geo/2010-census-regions-and-divisions-of-the-united-states.html.
World Health Organization. (2022). International Statistical Classification of Diseases and Related Health Problems (ICD). Available from: https://www.who.int/standards/classifications/classification-of-diseases.

Alexis Peterson, PhD, is a health scientist on the Traumatic Brain Injury (TBI) Team in the Division of Injury Prevention (DIP) at CDC's Injury Center. As a health scientist, her role on the TBI team is to promote the accurate reporting of TBI incidence and burden to inform the general public and stakeholders. She is also responsible for the development of projects and scientific manuscripts addressing multiple aspects of TBI, including sports- and recreation-related TBI.

Karen Thomas, MPH, is a data manager and statistician on the Economics and Statistics Team in DIP at CDC's Injury Center.

Hong Zhou, MS and MPH, is a statistician on the Economics and Statistics Team in DIP at CDC's Injury Center.

Journal of Safety Research 83 (2022) 1–7

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Do employees' work schedules put them at-risk? The role of shift scheduling and holidays in predicting near miss and incident likelihood

Matthew M. Laske [a], Philip E. Hinson [b], Yalcin Acikgoz [b,*], Timothy D. Ludwig [b], Anne M. Foreman [c], Shawn M. Bergman [b]

[a] University of Kansas, United States
[b] Appalachian State University, United States
[c] National Institute of Occupational Safety and Health, United States

ARTICLE INFO

Article history:
Received 11 November 2020
Received in revised form 16 November 2021
Accepted 28 July 2022
Available online 11 August 2022

Keywords:
Shift-schedule
Consecutive shifts
Holiday schedule
Injury prediction
Near miss prediction

ABSTRACT

Introduction: Using crew scheduling and injury incident data from a Fortune 500 manufacturing company, this study analyzed the effect of consecutive shifts and shifts near holidays on near misses and incidents. *Methods:* Logistic regressions were conducted with consecutive workdays, days near holidays, and time of shift as predictors of incident and near miss outcomes. *Results:* The logistic regression analysis indicated that working consecutive day shifts increases the probability of an incident occurring, with the fourth consecutive shift resulting in the most risk. The consecutive shift pattern did not replicate to employees working the night shift. However, the first and second shifts when transferring to a night schedule appear to have a greater chance of incident. Shifts near holidays did not have a significantly higher risk than other shifts. *Practical application:* The current research suggests that organizations can use similar analytic techniques to determine if shift scheduling might be related to increased risk and allocate resources to mitigate hazards during those peak probability shifts.

1. Introduction

In 2019, there were 2.8 million workplace injuries and illnesses, with 888,220 lost-time injuries (Bureau of Labor Statistics [BLS], 2020a). Additionally, there were 5,333 fatalities in the private sector (Bureau of Labor Statistics [BLS], 2020b). Each injury or fatality brings costs to the employee and their families (Schulte, 2005). Weil (2001) found that workplace injury costs are highly underestimated when the social costs (e.g., work disability, earning losses) are omitted from the calculations. Reviews of the social costs resulting from injuries indicate negative relationships with earnings and household family activities (Boden, 2005). Workers who experienced serious injuries were more susceptible to psychological issues, substance abuse, and marital problems (Texas Workers' Compensation Research Center, 2005). These findings illustrate the impact that workplace safety incidents have on employees, their families, and organizations. As such, organizations across many industries, particularly those whose employees are exposed to hazardous conditions, are turning to data analytics to help identify and mitigate risk factors associated with incidents.

One specific risk factor associated with workplace incidents is fatigue and labor scheduling practices, as previous literature has shown that scheduling can have profound effects on the safety and well-being of employees (Dembe, Erickson, Delbos, & Banks, 2005; Lombardi, Folkard, Willetts, & Smith, 2010; Nakata, 2011; Olds & Clarke, 2010). In the present study, we investigated the impact of consecutive days worked, day or night shifts, and holiday scheduling on injuries and near misses (i.e., an unplanned event that did not result in injury, illness, or damage – but had the potential to do so; National Safety Council, 2013). Data from worker crews at a large Fortune 500 chemical manufacturer in the southeastern United States were analyzed. Crews were selected as the level of analysis to protect animosity of employee identities.

1.1. Employee work schedules

1.1.1. Work hours and scheduling

Physical and psychological demands of a job and the workplace are related to employee fatigue (Li, Jiang, Yao, & Li, 2013). Employees with less than 11 hours of rest a day have been shown to have higher levels of psychological distress (Tsuchiya, Takahashi, Miki, Kubo, & Izawa, 2017) and increased fatigue (Vedaa et al., 2016). Work schedules can sometimes restrict the amount of rest

* Corresponding author.
E-mail address: acikgozy@appstate.edu (Y. Acikgoz).

https://doi.org/10.1016/j.jsr.2022.07.015
0022-4375/

between shifts to under 11 hours and can be considered both physically and psychologically demanding. This can result in acute negative health outcomes, such as lack of sleep and fatigue (Vedaa et al., 2016). Other research suggests that daily rest periods under 13 hours do not allow employees to adequately recover from fatigue (Ikeda et al., 2017).

Because of their associations with employee health and fatigue, work hours and scheduling have been shown to have direct relationships with employee safety outcomes. Dembe et al. (2005) found that jobs with overtime schedules, shifts of 12 or more hours, or workweeks of 60 or more hours have significantly higher self-reports of injuries than jobs without these characteristics. Across various industries, including both the commercial driving and medical industries, increased work hours are predictive of adverse safety outcomes (Dembe et al., 2005; Soccolich et al., 2013). These findings are supported by Yamauchi et al. (2019), who found that near miss likelihood was significantly higher for employees working over 41 hours a week compared to those working 35–40 hours a week. Further, incident likelihood was significantly higher for employees working over 51 hours a week compared to those working 35–40 hours a week.

There is some evidence that working multiple days in a row exacerbates inadequate rest between shifts, contributing to further fatigue and injuries. Thompson (2019) investigated how fatigue accumulates across three consecutive 12-hour shifts in healthcare. Reaction time, attention, and muscle function all depreciated over those consecutive shifts. Rotating night-shift schedules, such as the Dupont schedule, during which workers alternate weeks between working consecutive night shifts consecutive day shifts, have been shown to conflict with employees' circadian rhythms, resulting in sleepiness and worse job performance (Akerstedt, 1990). Industrial employees with rotating shift schedules tend to get less sleep in the initial days of a series of consecutive shifts (Budnick, Lerman, Baker, Jones, & Czeisler, 1994). Variance in employee well-being created by shift scheduling could result in an increased chance of injury.

A meta-analysis by Folkard and Lombardi (2006) modeled the compounding effects that longer shifts, non-daytime shifts, and consecutive shifts have on incident risk. First, they examined these factors in isolation, finding that night shifts were riskier than afternoon shifts, which were riskier than day shifts. Similarly, 12-hours shifts were riskier than 10-hour shifts, which were riskier than 8-hour shifts. Incident risk also increased for each consecutive shift worked. When examining these factors in tandem, they found that the safest way for an employee to work a 48-hour week was to work six consecutive 8-hour day shifts. This option was 20% safer than working four consecutive 12-hour day shifts, 40% safer than working six consecutive 8-hour night shifts, and 50% safer than working four consecutive 12-hour night shifts. A similar pattern was observed when modeling a 60-hour workweek. Folkard and Lombardi (2006) summarized two general findings from their results. First, working more but shorter shifts is safer than fewer but longer shifts. Second, day shifts tend to be safer than night shifts. Based on the previous studies, we expect to find that the more consecutive shifts a crew has worked, the higher the chance of an incident or near miss.

Hypothesis 1. The more consecutive days a crew works, the higher their risk of experiencing (a) an incident or (b) a near miss.

1.1.2. Employee work around holidays

In the United States, the average manufacturing employee receives nine paid holiday days per year (Bureau of Labor Statistics, 2019). Employees also average eight paid vacation days within their first year on the job and 12 days after five years of job tenure. Some employees may wish to use their vacation days around the holiday to extend time off. For manufacturing organizations operating 365 days a year, days around holidays can result in personnel changes that deviate from normal scheduling. When employees work around the holidays, their tasks and workload may change due to these personnel changes. Those changes may result in increased stress. Nawijn, de Bloom, and Geurts (2013) reported that prior to holidays, increased workload demands predicted decreases in self-reported health and well-being leading up to a vacation. Syrek, Weigelt, Kühnel, and de Bloom (2018) found that employees with a larger amount of unfinished work prior to the holiday were more likely to return to work with lower levels of positive affect. Therefore, work schedules prior to and post-holiday can be stressful for employees.

Workplace stressors are related to occupational injuries. For example, Haruyama et al. (2014) provided evidence for associations between job demands, physical and psychological stress, and reported cuts and burns in kitchen staff. A study of dam construction workers found a positive relationship between job stress and occupational injuries (Hussen, Dagne, & Yenealem, 2020). Self-reported time pressure, increased workload, excessive work, and working multiple job roles were related to occupational injuries among firefighters (Kim, Ahn, Kim, Yoon, & Roh, 2016). Similar relationships have also been found in the manufacturing industry (Kim, Min, Min, & Park, 2009; Nakata et al., 2006).

Based on this research, it is hypothesized that crews will be more likely to experience an incident or near miss on shifts near a holiday.

Hypothesis 2. Crews will be more likely to experience an incident or near miss while working shifts two days prior and following a company holiday.

1.2. Exploratory variable

1.2.1. Day or night shift

Across many jobs and industries, non-standard shifts (i.e., those that deviate from the conventional nine-to-five workday) are associated with a higher risk of injuries and illnesses (Dembe et al., 2005). For example, laborers who work past midnight have been shown to have poorer mental health (Sato, Kuroda, & Owan, 2020). There is some evidence that employees who switch from a non-night shift to a night shift may have an increased chance of developing depressive or anxiety disorders (Beltagy, Pentti, Vahtera, & Kivimäki, 2018). The same study found that when employees switched from a night to day shift, there was an increased recovery rate from these disorders. Night shifts have also been associated with a greater risk of injury than day shifts (Smith, Folkard, & Poole, 1994). Other studies have found no association between working night shifts and injuries (Nielsen et al., 2019). For example, a recent review of over 13,000 occupational injuries found no differences between day and night shifts and occupational injuries. Due to mixed findings in the literature, no hypotheses were made regarding day or night shift risk differences prior to conducting the analyses.

1.3. Overview

To evaluate the hypotheses described above, we conducted analyses of incident data collected at a chemical manufacturing plant over a three-year period in conjunction with human resources data regarding specific days and shifts those employees worked and holiday schedules. Analyzing organizational data has several advantages over analyzing self-reported survey data. Survey measures have limitations in the accuracy of self-reported information. Kessler et al. (2003) found that employees can overestimate their hours worked and underreport their work absenteeism compared to payroll records. Underreported and undercounted injuries often result in injury estimates much greater than reported data (Leigh et al., 2004, 2014). Across industries, unreported accidents occur at a greater rate than reported accidents (Probst & Estrada, 2010). Using self-report measures can potentially lead to misinformed analytic applications because of errors and inaccuracies in the data.

The objective of the present research is to examine three years of scheduling data to create prediction models for incident and near-miss outcomes. Logistic regression models were created to assess the impact of consecutive days worked and holiday scheduling on injuries and near misses. The hypotheses evaluated were: H1: *The more consecutive days a crew works, the higher their risk of experiencing (a) an incident or (b) a near miss.* H2: *Crews will be more likely to experience an incident or near miss while working shifts two days prior and following a company holiday.* We also explored the relationship between the type of shift (day or night) and injuries and near misses.

2. Materials and methods

2.1. Participants and setting

Data for the study were made available by a chemical manufacturer in the Southeastern United States that specializes in the production of various advanced materials, chemicals, and fibers. The scope of this study was limited to the division that manufactures advanced fiber materials. At the time of the study, the division contained approximately 350 operations employees. Within the division, there were five departments, with each department containing four crews. Employees did not alternate across crews. These crews follow a 12-hour shift DuPont schedule where they work a series of three or four days or nights in a row, followed by one to seven days off. Work tasks within the departments included (a) collecting samples of chemical materials, (b) switching out equipment configuration, (c) emptying excess chemical material from the system, and (d) transporting raw material with forklifts.

2.2. Measurement

Using R software, work scheduling variables were created based on the chemical company's crew schedule calendars for 2016–2018, totaling 2,144 observations. Safety outcome data were retrieved from the company's safety data tracking system and filtered to only include the participating departments' incident and near miss data. These outcomes were then linked to the specific crew that was impacted. All data were aggregated to the crew level across all five departments because multiple crews worked the same schedule across the different departments. Additionally, information on individual employees involved in incidents or near misses was excluded to protect the identity of those individuals. Therefore, each observation included a crew number, the crew's current shift in their work schedule, and the number of incidents and near misses for that crew.

2.2.1. Consecutive work days

A variable, ranging from one to four, was calculated based on the crews' shift calendar to indicate how many consecutive days a crew had worked prior to and including the current shift.

2.2.2. Near holiday

A binary variable was coded based on the crews' shift calendar to indicate whether a shift occurred two days before or after a company holiday. Company holidays include the following U.S. holidays: New Year's Day, Good Friday, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, and Christmas Day. Additional holiday leave was also granted the day following Independence Day, Thanksgiving Day, and proceeding Christmas Day.

2.2.3. Time of shift

A binary variable was created based on the crews' shift calendar to indicate whether a shift occurred during the day or the night.

2.2.4. Incidents and near misses

Two binary safety outcome variables were created to indicate whether an incident or near miss occurred during a given shift. The chemical manufacturer classifies an event as an incident when it results in employee injuries (both recordable incidents and first aid events that were reported), fatality, property damage, or unintended chemical release. Events are classified as near misses when hazardous energy (e.g., electrical, gravitational, hydraulic, pneumatic, mechanical) is released or modified, and an incident resulting in personal harm nearly occurs. Binary variables for each of these were deemed more appropriate than continuous counts due to the low base rate of these events (i.e., there were only four days in which two near misses occurred and three days in which two incidents occurred).

2.3. Analytic approach

To examine the effect of the number of consecutive days on incidents and near misses, we used a binary logistic regression model. The odds ratios associated with the consecutive days worked were examined in predicting the probability of an incident or near miss. The specific probabilities that an incident or near miss would occur for each level of consecutive days worked variable were also obtained. A chi-square analysis was used to examine the possible effect of holidays on incident and near miss occurrence by tabulating the type of shift with the binary incident and near miss variables separately. These binary variables were tabulated against whether a day was within two days of a holiday. The same methodology was applied to test associations between day versus night shift and the probability of an incident or near miss occurring. Lastly, to examine the combined effect of consecutive days worked and proximity to holiday, a binary logistic regression model was created in which these three variables were entered as covariates, and incident and near miss occurrence were entered as the dependent variable, separately.

3. Results

The means, standard deviations, and correlations among the variables used in the study are presented in Table 1. The mean value for the incident binary variable, which indicates whether or not there was an incident on a day, was 0.02 (SD = 0.16), indicating that 2% of the days in our dataset had an incident occurring. Similarly, 5% of the days had a near miss occurrence, and 8% of the days were considered to be near a holiday. The mean values for consecutive day and night shift variables indicate that, on average, the crews in the dataset worked 1.12 consecutive days and 1.13 consecutive nights, indicating that it was not common to work multiple back-to-back shifts.

The first hypothesis suggested that as the number of consecutive days worked increased, there would be a higher probability of (a) an incident and (b) a near miss occurring. To test this, a binary logistic regression analysis was performed. The number of consecutive days was entered as the predictor, and the binary incident and near miss variables were entered as the dependent variables in

Table 1
Descriptive statistics and correlations between the variables examined in the study.

	Incident Binary	Near Miss Binary	Near Holiday	Consecutive Days	Consecutive Nights
Incident Binary	–				
Near miss Binary	−0.02	–			
Near Holiday	0.03	0.04	–		
Consecutive Days	0.08***	0.03	−0.04	–	
Consecutive Nights	−0.05*	−0.05*	−0.04	−0.71***	–
Mean	0.02	0.05	0.08	1.12	1.13
SD	0.16	0.21	0.27	1.22	1.34

* p <.05.
*** p <.001.

separate analyses. Table 2 summarizes the results of the logistic regression analyses. Only one of the analyses resulted in a significant finding. Specifically, for employees working consecutive day shifts, the probability of an incident occurring significantly increased with a higher number of consecutive days worked, such that an increase of one day resulted in a 41.4% increase in the odds of an incident (95%CI = 2.7−94.7%). As further evidence, the correlation between consecutive days and incident variables was significant and positive (see Table 1), indicating that the probability of an incident significantly increased with each consecutive day shift. However, the relationship between consecutive days worked and near misses was not significant. Overall, Hypothesis 1 was partially supported. Table 3 summarizes the probabilities of incident or near miss occurring associated with each additional consecutive day employees worked. For example, if a crew was working their fourth consecutive day shift, they faced a 5.84% probability of an incident and a 4.93% probability of a near miss occurring during their shift.

The second hypothesis, which suggested that there would be a higher occurrence of incidents and near misses in days close to holidays, was tested using a Chi-square test of independence in which we tabulated the binary holiday variable with the binary incident variable indicating whether an incident or near miss had occurred during that shift. The relationship between these variables and incidents was not significant, $X^2(1, N = 2144) = 2.17$, p =.141. Similarly, we observed non-significant results in predicting near misses, $X^2(1, N = 2144) = 2.64$, p =.104.

In an exploratory analysis, we examined whether night shifts would have a higher likelihood of an incident occurring compared to day shifts. This was tested using a Chi-square test of independence in which we tabulated the shift variable (day vs night) with the binary incident variable indicating whether an incident had occurred during that shift. We found that the relationship between these variables was significant, $X^2(1, N = 2144) = 8.53$, $p < .01$. However, the direction of the effect was the opposite of previous findings from the literature, whereby day shifts had a higher likelihood of an incident than night shifts. The same pattern of results was observed for near misses, suggesting that day shifts may have a higher likelihood of those events, $X^2(1, N = 2144) = 4.67$, $p < .05$.

Table 4 shows the raw numbers of day and night shifts with incidents and near misses, grouped by consecutive day/night shifts. The first night shift session has the second highest probability of an incident occurring and the highest near miss occurring among consecutive night shifts. To further explore this pattern of findings and complement the analysis in Hypothesis 1, we also tested whether the probability of an incident or near miss occurring significantly increased or decreased with each additional night shift by running logistic regression analyses, which yielded non-significant results. However, as seen in Table 1, the correlations between consecutive nights and both incident and near miss variables were significantly negative, indicating that the risk for an incident or near miss decreased with each consecutive night shift. More research is needed to explore this nuanced relationship between consecutive day and night shifts and the probability of an incident or near miss occurring.

Table 3
Predicted probabilities of an incident or near miss occurring in each consecutive day/night.

	Incident	Near Miss
Day 1	2.14%	6.09%
Day 2	3.01%	5.68%
Day 3	4.20%	5.29%
Day 4	5.84%	4.93%
Night 1	1.39%	4.69%
Night 2	1.47%	3.76%
Night 3	1.56%	3.00%
Night 4	1.65%	2.39%

4. Discussion

The current study examined the effect of scheduling on safety outcomes, specifically on the likelihood of an incident or near miss occurring in each shift. The predictors examined were the number of consecutive days worked by employees, whether the shift was shortly before or after a holiday, and in an exploratory manner, whether it was a day shift or a night shift. The results indicate that the probability of an incident significantly increased as employees worked consecutive day shifts. With night shifts, a reverse pattern was observed such that the first and second nights seemed to have a higher risk of incident or near miss compared to shifts occurring after the second consecutive night. There were more incidents and near misses occurring on day shifts compared to night shifts. Finally, the proximity of shift work to an upcoming holiday break was not associated with increases in incidents or near misses. The null finding for incidents and near misses around holidays

Table 2
Logistic regressions predicting incident and near miss occurrence.

IV/DV	B	Wald X^2	p	Odds Ratio (OR)	OR 95% CI
Consecutive Day Shifts/Incidents	0.35	4.51	0.034	1.41	1.03–1.95
Consecutive Day Shifts/Near Misses	−0.08	0.33	0.569	0.93	0.72–1.20
Consecutive Night Shifts/Incidents	0.06	0.06	0.811	1.06	0.66–1.71
Consecutive Night Shifts/Near Misses	−0.23	1.96	0.162	0.79	0.57–1.10

Note. Each line in the table represents a separate logistic regression analysis.

Table 4
Raw Numbers of 1st, 2nd, 3rd, and 4th Day & Night Shifts with an Incident or Near Miss.

	Day Shift		Night Shift	
	Incident	Near Miss	Incident	Near Miss
1st	5/322 (1.6%)	22/322 (6.8%)	6/317 (1.9%)	15/317 (4.7%)
2nd	14/311 (4.5%)	16/311 (5.1%)	4/311 (1.3%)	13/311 (4.2%)
3rd	9/299 (3%)	12/299 (4%)	1/301 (0%)	6/301 (2%)
4th	9/140 (6.4%)	10/140 (7.1%)	5/143(3.5%)	5/143 (3.5%)

Note. The numerator indicates the number of shifts with a reported incident/near miss. The denominator indicates the total shifts in the dataset.

may be due to other variables that are affected by the holiday, such as planned decreases in production requirements, fewer employees working, and some employees taking extended vacation days.

One potential explanation for the difference between consecutive days worked versus consecutive nights worked concerns the possible effect of circadian rhythms. The crews involved in this study worked alternating weeks of day and night shifts. After working their consecutive days on the day shift for one week, the crew was shifted to the night shift. Therefore, the entire crew experienced the shift from day to night. It is possible that the dramatic shift in sleep schedule creates the most fatigue the first night shift after having worked the day shift for a week, as it may take some time for the employees' circadian rhythms to adjust (Violanti et al., 2012). Employees may be less attentive at work on their first night shift (Budnick et al., 1994), leading to an increase in incidents. Note that for night shifts, this would contradict our first hypothesis, which suggested that each consecutive shift may increase the probability of an incident occurring. The crews in our dataset were rotating shifts every week, meaning that the first night shift was always their first shift after working day shifts.

There was not a significant increase in near miss probability with each consecutive day worked during either day or night shifts. Since near misses are considered leading indicators of incidents (Occupational Safety and Health Administration, 2019), we would expect to see a similar increase in the probability of near misses with more consecutive day shifts. One possible explanation is that while reporting an incident is mandatory, reporting a near miss is optional. It may be possible for the near miss metrics to be influenced by idiosyncratic reporting behaviors, which likely vary across employees and situations, introducing a larger amount of error variance in near miss measures.

4.1. Practical Applications

When the potential for an incident is identified, an organization should direct resources (e.g., additional staffing) and safety initiatives to address that risk (e.g., increased observations/audits). Based on the results of the present study, adjustments in staffing, increases in break time, and additional safety initiatives (e.g., observations, audits, coaching) on the fourth consecutive day shifts and first night shifts after transitioning from day to night may be considered to mitigate the risk. If logistically possible, risk might be reduced by modifying shift schedules such that consecutive 12-hour shifts are limited to three consecutive days. In addition, given the finding that the first night shift appeared to have increased risk, crews may stay on day/night shifts for a longer amount of time instead of alternating every week.

4.2. Recommendations for future research

The current study examined the effect of consecutive workdays across day and night shifts on incidents and near misses with limited data containing relatively low base rates of both incidents and near misses. Regardless, we were able to tentatively identify increased incident and near miss probability after a number of consecutive days worked. Future research should use larger datasets to replicate and extend our findings to provide organizations with more specific guidelines regarding work scheduling.

Another area needing more extensive data analysis is the difference in the probability of an injury between night shifts and day shifts, as the findings in this area are mixed (Folkard & Lombardi, 2006; Fransen et al., 2006; Nielsen et al., 2019). While our analyses found that, on average, night shifts had a smaller number of incidents compared to day shifts, we were not able to control for the number of employees at work during those times. This may explain the difference in frequency of incidents (i.e., fewer workers equals fewer opportunities for injury). In addition, while the current study identified an increased risk on the first night shift worked after consecutive day shifts, this finding may only be a trend because of the low base rate of incidents in our data. However, this finding supports previous research showing that rotating shift schedules are associated with a greater risk of injury (Bagheri Hosseinabadi et al., 2019; Dembe et al., 2005; Wong, McLeod, & Demers, 2011). Finally, the organization providing the data for the current study utilized a DuPont work schedule, thus making it impossible to compare this schedule to other work schedule arrangements (Folkard & Lombardi, 2006). Future research could compare different work schedules in terms of safety outcomes such as incidents and near misses.

4.3. Limitations

This study has several limitations. First, data only included crews working in an advanced fibers manufacturing division within a chemical manufacturing company. Findings must be replicated in other settings and industries to support their generalizability. Second, data on day/night and consecutive shift scheduling were created using crew scheduling. This was later combined with a separate dataset indicating dates of incidents for employees within the crews. Consequently, when an incident occurred on a given date and time, we associated it with the crew assigned to work that day based on their day/night shift schedule. This meant we could not control for certain employee-level variables in our analyses, such as employee experience or tenure. In addition, while unlikely, there is a possibility that the employee might have worked a modified schedule for that week different from their crew. The third limitation involves the number of employees working in each shift. The number of labor hours directly influences the number of hours employees are exposed to hazards and risks. It is also plausible that when there is more activity on the shop floor (i.e., more employees working in the area), there will be a higher likelihood of an incident or near miss. Because the dataset did not include this information, we could not control the number of employees in each shift for our analyses.

5. Summary

Results from the current study provide evidence that an employee's work schedule is predictive of the probability of a safety incident. Logistic regression analysis indicated that working consecutive day shifts increases the probability of an incident occurring, with the fourth consecutive shift resulting in the highest probability. The consecutive shift pattern did not replicate to the night shift. However, after transferring to a night schedule, the first and second shifts appear to have a greater chance of incident than later night shifts. The current research suggests that industrial organizations can use similar analytic techniques to determine if shift scheduling might be related to increased risk and allocate resources to mitigate hazards during peak probability shifts.

References

Akerstedt, T. (1990). Psychological and psychophysiological effects of shift work. *Scandinavian Journal of Work, Environment and Health, 16*(1), 67–73. https://doi.org/10.5271/sjweh.1819.

Bagheri Hosseinabadi, M., Khanjani, N., Etemadinezhad, S., Samaei, S. E., Raadabadi, M., & Mostafaee, M. (2019). The associations of workload, individual and organisational factors on nurses' occupational injuries. *Journal of Clinical Nursing, 28*(5–6), 902–911. https://doi.org/10.1111/jocn.14699.

Beltagy, M. S., Pentti, J., Vahtera, J., & Kivimäki, M. (2018). Night work and risk of common mental disorders: Analyzing observational data as a non-randomized pseudo trial. *Scandinavian Journal of Work, Environment and Health, 44*(5), 512–520. https://doi.org/10.5271/sjweh.3733.

Boden, L. I. (2005). Running on empty: Families, time, and workplace injuries. *American Journal of Public Health, 95*(11), 1894–1897. https://doi.org/10.2105/AJPH.2005.062232.

Budnick, L. D., Lerman, S. E., Baker, T. L., Jones, H., & Czeisler, C. A. (1994). Sleep and alertness in a 12-hour rotating shift work environment. *Journal of Occupational Medicine, 36*(12), 295–300. https://doi.org/10.1097/00043764-199412000-00010.

Bureau of Labor Statistics [BLS] (2020a). Employer-reported workplace injuries and illnesses – 2019 (USDL-20-2030). Washington, DC: U.S. Department of Labor. https://www.bls.gov/news.release/pdf/osh.pdf.

Bureau of Labor Statistics [BLS] (2020b). National census of fatal occupational injuries in 2019 (USDL-20-2265). Washington, DC: U.S. Department of Labor. https://www.bls.gov/news.release/pdf/cfoi.pdf.

Bureau of Labor Statistics (2019). National compensation survey: Employee benefits in the United States, March 2019 (Bulletin 2791). Washington, DC: U.S. Department of Labor. https://www.bls.gov/ncs/ebs/benefits/2019/employee-benefits-in-the-united-states-march-2019.pdf.

Dembe, A. E., Erickson, J. B., Delbos, R. G., & Banks, S. M. (2005). The impact of overtime and long work hours on occupational injuries and illnesses: New evidence from the United States. *Occupational and Environmental Medicine, 62*(9), 588–597. https://doi.org/10.1136/oem.2004.016667.

Folkard, S., & Lombardi, D. A. (2006). Modeling the impact of the components of long work hours on injuries and "accidents". *American Journal of Industrial Medicine, 49*(11), 953–963. https://doi.org/10.1002/ajim.20307.

Fransen, M., Wilsmore, B., Winstanley, J., Woodward, M., Grunstein, R., Ameratunga, S., & Norton, R. (2006). Shift work and work injury in the New Zealand Blood Donors' Health Study. *Occupational and Environmental Medicine, 63*(5), 352–358. https://doi.org/10.1136/oem.2005.024398.

Haruyama, Y., Matsuzuki, H., Tomita, S., Muto, T., Haratani, T., Muto, S., & Ito, A. (2014). Burn and cut injuries related to job stress among kitchen workers in Japan. *Industrial Health, 52*(2), 113–120. https://doi.org/10.2486/indhealth.2013-0143.

Hussen, J., Dagne, H., & Yenealem, D. G. (2020). Factors associated with occupational injury among hydropower dam construction workers, south east Ethiopia, 2018. *BioMed Research International, 2020*, 12–15. https://doi.org/10.1155/2020/6152612.

Ikeda, H., Kubo, T., Izawa, S., Takahashi, M., Tsuchiya, M., Hayashi, N., & Kitagawa, Y. (2017). Impact of daily rest period on resting blood pressure and fatigue: A one-month observational study of daytime employees. *Journal of Occupational and Environmental Medicine, 59*(4), 397–401. https://doi.org/10.1097/JOM.0000000000000968.

Kessler, R. C., Barber, C., Beck, A., Berglund, P., Cleary, P. D., McKenas, D., et al. (2003). The World Health Organization Health and Work Performance Questionnaire (HPQ). *Journal of Occupational and Environmental Medicine, 45*(2), 156–174. https://doi.org/10.1097/01.jom.0000052967.43131.51.

Kim, H. C., Min, J. Y., Min, K. B., & Park, S. G. (2009). Job strain and the risk for occupational injury in small- to medium-sized manufacturing enterprises: A prospective study of 1,209 Korean employees. *American Journal of Industrial Medicine, 52*(4), 322–330. https://doi.org/10.1002/ajim.20673.

Kim, Y. K., Ahn, Y. S., Kim, K., Yoon, J. H., & Roh, J. (2016). Association between job stress and occupational injuries among Korean firefighters: A nationwide cross-sectional study. *BMJ Open, 6*(11). https://doi.org/10.1136/bmjopen-2016-012002.

Leigh, J. P., Du, J., & McCurdy, S. A. (2014). An estimate of the U.S. government's undercount of nonfatal occupational injuries and illnesses in agriculture. *Annals of Epidemiology, 24*(4), 254–259. https://doi.org/10.1016/j.annepidem.2014.01.006.

Leigh, J. P., Marcin, J. P., & Miller, T. R. (2004). An estimate of the U.S. government's undercount of nonfatal occupational injuries. *Journal of Occupational and Environmental Medicine, 46*(1), 10–18. https://doi.org/10.1097/01.jom.0000105909.66435.53.

Li, F., Jiang, L., Yao, X., & Li, Y. (2013). Job demands, job resources and safety outcomes: The roles of emotional exhaustion and safety compliance. *Accident Analysis and Prevention, 51*, 243–251. https://doi.org/10.1016/j.aap.2012.11.029.

Lombardi, D. A., Folkard, S., Willetts, J. L., & Smith, G. S. (2010). Daily sleep, weekly working hours, and risk of work-related injury: Us national health interview survey (2004–2008). *Chronobiology International, 27*(5), 1013–1030. https://doi.org/10.3109/07420528.2010.489466.

Nakata, A. (2011). Effects of long work hours and poor sleep characteristics on workplace injury among full-time male employees of small- and medium-scale businesses. *Journal of Sleep Research, 20*(4), 576–584. https://doi.org/10.1111/j.1365-2869.2011.00910.x.

Nakata, A., Ikeda, T., Takahashi, M., Haratani, T., Hojou, M., Fujioka, Y., et al. (2006). Impact of psychosocial job stress on non-fatal occupational injuries in small and medium sized manufacturing enterprises. *American Journal of Industrial Medicine, 49*(8), 658–669. https://doi.org/10.1002/ajim.20338.

Nawijn, J., de Bloom, J., & Geurts, S. (2013). Pre-vacation time: Blessing or burden? *Leisure Sciences, 35*(1), 33–44. https://doi.org/10.1080/01490400.2013.739875.

National Safety Council (2013). Near miss reporting systems. National Safety Council. https://www.nsc.org/Portals/0/Documents/WorkplaceTrainingDocuments/Near miss-Reporting-Systems.pdf.

Nielsen, H. B., Dyreborg, J., Hansen, Å. M., Hansen, J., Kolstad, H. A., Larsen, A. D., et al. (2019). Shift work and risk of occupational, transport and leisure-time injury. A register-based case-crossover study of Danish hospital workers. *Safety Science, 120*, 728–734. https://doi.org/10.1016/j.ssci.2019.07.006.

Olds, D. M., & Clarke, S. P. (2010). The effect of work hours on adverse events and errors in health care. *Journal of Safety Research, 41*(2), 153–162. https://doi.org/10.1016/j.jsr.2010.02.002.

Occupational Safety and Health Administration (2019). Using leading indicators to improve safety and health outcomes. https://www.osha.gov/leadingindicators/docs/OSHA_Leading_Indicators.pdf.

Probst, T. M., & Estrada, A. X. (2010). Accident under-reporting among employees: Testing the moderating influence of psychological safety climate and supervisor enforcement of safety practices. *Accident Analysis and Prevention, 42*(5), 1438–1444. https://doi.org/10.1016/j.aap.2009.06.027.

Sato, K., Kuroda, S., & Owan, H. (2020). Mental health effect of long work hours, night and weekend work, and short rest periods. *Social Science and Medicine, 246*, 112774. https://doi.org/10.1016/j.socscimed.2019.112774.

Schulte, P. A. (2005). Characterizing the burden of occupational injury and disease. *Journal of Occupational and Environmental Medicine, 47*(6), 607–622. https://doi.org/10.1097/01.jom.0000165086.25595.9d.

Smith, L., Folkard, S., & Poole, C. (1994). Increased injuries on night shift. *The Lancet, 344*, 1137–1139. https://doi.org/10.1016/S0140-6736(94)90636-X.

Soccolich, S. A., Blanco, M., Hanowski, R. J., Olsen, R. L., Mrgan, J. F., Guo, F., & Wu, S. (2013). An analysis of driving and working hour on commercial monitor vehicle driver safety using naturalistic data collection. *Accident Analysis and Prevention, 58*, 249–259. https://doi.org/10.1016/j.aap.2012.06.024.

Syrek, C. J., Weigelt, O., Kühnel, J., & de Bloom, J. (2018). All I want for Christmas is recovery–changes in employee affective well-being before and after vacation. *Work and Stress, 32*(4), 313–333. https://doi.org/10.1080/02678373.2018.1427816.

Texas Workers' Compensation Research Center (2005). Economic outcomes of injured workers with permanent impaimrents. *The Research Review, 27*, 1–4.

Thompson, B. J. (2019). Does work-induced fatigue accumulate across three compressed 12 hour shifts in hospital nurses and aides? Article e021175 *PLoS ONE, 14*(2). https://doi.org/10.1371/journal.pone.0211715.

Tsuchiya, M., Takahashi, M., Miki, K., Kubo, T., & Izawa, S. (2017). Cross-sectional associations between daily rest periods during weekdays and psychological distress, non-restorative sleep, fatigue, and work performance among information technology workers. *Industrial Health, 55*(2), 173–179. https://doi.org/10.2486/indhealth.2016-0140.

Vedaa, Ø., Harris, A., Bjorvatn, B., Waage, S., Sivertsen, B., Tucker, P., & Pallesen, S. (2016). Systematic review of the relationship between quick returns in rotating shift work and health-related outcomes. *Ergonomics, 59*(1), 1–14. https://doi.org/10.1080/00140139.2015.1052020.

Violanti, J. M., Fekedulegn, D., Andrew, M. E., Charles, L. E., Hartley, T. A., Vila, B., & Burchfiel, C. M. (2012). Shift work and the incidence of injury among police officers. *American Journal of Industrial Medicine, 55*(3), 217–227. https://doi.org/10.1002/ajim.22007.

Weil, D. (2001). Valuing the economic consequences of work injury and illness: A comparison of methods and findings. *American Journal of Industrial Medicine, 40*(4), 418–437. https://doi.org/10.10002/ajim.1114.

Wong, I. S., McLeod, C. B., & Demers, P. A. (2011). Shift work trends and risk of work injury among Canadian workers. *Scandinavian Journal of Work, Environment and Health, 37*(1), 54–61. https://doi.org/10.5271/sjweh.3124.

Yamauchi, T., Sasaki, T., Takahashi, K., Umezakim, S., Takahashi, M., Toshikawa, T., Suka, M., & Yanagisawa, H. (2019). Long working hours, sleep-related problems, and near-misses/injuries in industrial settings using a nationally representative sample of workers in Japan Article e0219657. *PLoS ONE, 14*(7). https://doi.org/10.1371/journal.pone.0219657.

Matthew M. Laske is a doctoral student in the Department of Applied Behavioral Science at the University of Kansas. He received his MA from Appalachian State University in Industrial Organizational Psychology and Human Resources Management.

Philip E. Hinson is a research associate at HumRRO. He received his MA from Appalachian State University in Industrial Organizational Psychology and Human Resources Management.

Yalcin Acikgoz is an assistant professor in the Department of Psychology at Appalachian State University. He received his Ph.D. from Middle East Technical University in I-O Psychology. His research interests focus on employee recruitment and job search, applicant reactions in recruitment and selection, applicant decision-making, applications of social media in human resources, and impression management in the workplace.

Timothy D. Ludwig is a professor in the Department of Psychology at Appalachian State University. He earned his Ph.D. at Virginia Tech researching the benefits of employee-driven behavioral safety programs. He has over 30 years of experience in research and practice in behavioral safety where he integrates empirical findings into his safety consulting.

Anne M. Foreman is an associate service fellow at the National Institute of Occupational Safety and Health (NIOSH). She received her Ph.D. from West Virginia University in Psychology with an emphasis on Behavior Analysis. She has a background in basic learning processes, decision making, and applied animal behavior.

Shawn M. Bergman is a professor in the Department of Psychology at Appalachian State University. He received his Ph.D. from The University of Tennessee-Knoxville in I-O Psychology. His teaching and research focus on the application of quantitative methods and analytics to solve practical problems.

Journal of Safety Research 83 (2022) 26–34

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Driver injuries in heavy vs. light and medium truck local crashes, 2010–2019

Terry Lee Bunn [a,b,*], Madison Liford [b], Michael Turner [b], Ashley Bush [b]

[a] Department of Preventive Medicine and Environmental Health, College of Public Health, University of Kentucky, 111 Washington Ave, Lexington, KY, USA
[b] Kentucky Injury Prevention and Research Center, 333 Waller Ave., Suite 242, Lexington, KY, USA

ARTICLE INFO

Article history:
Received 11 May 2021
Received in revised form 13 January 2022
Accepted 1 August 2022
Available online 10 August 2022

Keywords:
Light truck crash injuries
Medium truck crash injuries
Injury scenarios
Wholesale and retail trade industries
Rear-end crash injuries

ABSTRACT

Objective: Multiple heavy truck driver injury studies exist, but there is a paucity of research on light and medium truck driver injuries. The objective of this study was to use first report of injury (FROI) data to: (a) compare demographic and injury characteristics; (b) assess workers' compensation (WC) claim disposition and lost work time status; and (c) describe injury scenarios by vehicle type for heavy truck and light/medium truck driver local crashes. *Method:* Kentucky Department of Workers' Claims FROI quantitative and free text data were analyzed for years 2010–2019. Of 800 total FROIs, 451 involved heavy trucks and 349 involved light or medium trucks. *Results:* There was a higher light/medium truck driver crash FROI rate compared to the heavy truck driver crash FROI rate. There was a higher proportion of younger light/medium truck driver crash FROIs compared to younger heavy truck driver crash FROIs. The retail trade industry made up the largest percentage of light/medium truck local crash FROIs (47%); the transportation and warehousing industry was most frequently cited in heavy truck FROIs (46%). The heavy truck types most frequently identified in FROIs were semi-trucks (13%) and dump trucks (11%). The most common light/medium truck type identified was delivery trucks (30%). Most commonly, heavy truck crash FROIs involved rollovers, driving off/overcorrecting on narrow roadways, and driving downhill/unable to downshift. Light/medium truck crash FROIs most frequently involved being rear-ended, running red lights, and turning in front of other vehicles. *Conclusions:* The utilization of WC FROI data highlighted top injury scenarios and specific vehicle types for targeting driver safety training among truck drivers, particularly light/medium truck drivers. Road safety policies regarding driver training, crash reviews, and in-vehicle monitoring systems are needed for truck drivers with previous crash injuries, especially for light and medium truck drivers. *Practical applications:* Enhanced safety training on speeding on narrow roadways, on nearing intersections, and on downshifting on hills is needed for semi-truck, dump truck, and coal truck drivers with previous crash injuries. Rear-end crash prevention training (e.g., gradual stopping and checking mirrors) is needed for drivers of furniture, automotive parts and accessories, and groceries and soft drink delivery trucks with previous crash injuries.

1. Introduction

In 2019, the Bureau of Labor Statistics (BLS) Census of Fatal Occupational Injuries (CFOI) counted 2,122 commercial transportation incident deaths, of which 471 were in truck transportation, 83 were in wholesale trade, and 59 were in retail trade (Bureau of Labor Statistics, 2021). Within the truck transportation industry, the fatality rate was 27.2 worker deaths per 100,000 full-time equivalent workers, compared to the U.S. overall worker fatality rate of 3.5/100,000. The nonfatal injury incidence rate of 3.6 injuries and illnesses per 100 full-time workers in the truck transportation industry is 20% higher than the U.S. overall rate of 3.0/100 (Bureau of Labor Statistics, 2020). These high numbers and rates have served as the justification for multiple research studies on injuries in heavy vehicles such as semi-trucks in truck transportation (Bunn, Slavova, & Robertson, 2013; Bunn, Slavova, & Rock, 2019; Combs, Heaton, Raju, Vance, & Sieber, 2018; McKnight & Bahouth, 2009; Zheng, Lu, & Lantz, 2018).

Relatively few studies have been published on injuries to light and medium vehicle drivers. A study by Karaca-Mandic and Lee (2014) on car and light truck crashes using linked crash and hospital discharge data found that light truck drivers had reduced odds for hospitalizations and fatalities compared to passenger car dri-

* Corresponding author at: Kentucky Injury Prevention and Research Center University of Kentucky, College of Public Health, 333 Waller Ave., Suite 242, Lexington, KY 40504, USA.

E-mail address: tlbunn2@uky.edu (T.L. Bunn).

https://doi.org/10.1016/j.jsr.2022.08.001
0022-4375/© 2022 The Author(s). Published by the National Safety Council and Elsevier Ltd.

vers. Pratt and Bell (2019) showed that age, gender, and job tenure were significant driver collision risk factors in a light vehicle healthcare industry fleet. Lack of seat belt use and high speed have also been shown to be associated with light vehicle serious crash injuries (Doecke, Baldock, Kloeden, & Dutschke, 2020; Pipkorn, Iraeus, Lindkvist, Puthan, & Bunketorp, 2020; Stuckey, Glass, LaMontagne, Wolfe, & Sim, 2010). In a follow-up study by Stuckey, LaMontagne, Glass, and Sim (2010) using linked vehicle registration and crash data, the authors estimated light vehicle fatality rates to be much higher than those reported using workers' compensation (WC) data alone.

Using an in situ driving data set, Hanowski, Hickman, Wierwille, and Keisler (2007) found that most vehicle crashes, near crashes, and crash-related conflicts involving both light and heavy vehicles were initiated by the light vehicle driver, with aggressive driving being the primary contributing factor, whereas the heavy vehicle driver-initiated events involved poor driving habits. Chen, Amandus, and Wu (2014) analyzed U.S. CFOI data and determined that of the total driver/sales worker and truck driver occupational fatalities, 85% were among heavy truck and tractor-trailer drivers and 9.5% were among light truck drivers. The authors acknowledged that their study was limited by a lack of appropriate employment data on driver occupation by industry and of worker characteristics for the occupational subcategories. They recommended that "data on circumstances and scenarios leading to a fatal truck crash are needed to better understand risk factors associated with highway fatalities in the group of truck drivers and driver/sales workers.".

WC data may not be the optimum data source for calculation of injury rates since its lack of information on self-employed workers and calculated injury rates would lead to an underestimation (Stuckey, LaMontagne, & et al., 2010). The primary advantage of analyzing WC data is that it contains free-text injury narratives that can provide additional information on the circumstances of a crash and specific driving actions performed by the drivers prior to the crash that other data sources (such as electronic crash data with no free-text narratives) cannot provide (Chandler, Bunn, & Slavova, 2017). The objectives of this study were to use WC first report of injury (FROI) data to: (a) compare demographic and injury characteristics between heavy truck driver, and light and medium truck driver local crash FROIs; (b) assess WC disposition and lost work time status associated with local heavy, medium, and light truck crash FROIs; and (c) describe injury scenarios by specific vehicle type to obtain additional information on the crash injury circumstances. Results from this study can be used to target new and enhance current light, medium, and heavy truck driver injury prevention strategies.

2. Methods

2.1. Data source

De-identified Kentucky WC FROIs for years 2010–2019 were obtained from the Kentucky Department of Workers' Claims (KDWC); information regarding reimbursement for medical expenses related to injuries was not available. The WC data set does not include FROIs on self-employed worker injuries. According to the KDWC, the following are acceptance criteria for FROIs:

1. All worker injuries that require at least one day off from work or result in a disability that extends beyond 60 days are required to be reported;
2. When a worker has lost at least seven days of work due to an injury or has a permanent partial disability with no missed work days due to an injury, the worker is eligible for indemnity and/or lump sum payments. Indemnity payments associated with FROIs or claims (litigated FROIs) are defined as paid income benefits to compensate for lost wages, functional impairment, or death; and.
3. When a worker has lost at least two weeks of work due to an injury, the worker is eligible for lost wage compensation retroactive to the first day of work lost.

The FROIs reflect the number of crash injuries (in reports submitted to KDWC) and do not reflect individual drivers involved in crashes.

2.2. Study selection and inclusion criteria

Inclusion criteria for heavy, medium, and light truck driver collision FROIs were based on a WC standard cause of injury code of (45) collision or sideswipe, (46) collision with a fixed object, (48) vehicle upset or rollover, or (50) motor vehicle "not otherwise classified." FROI standard cause of injury code 50 includes injuries due to sudden start or stop, being thrown against the interior of the vehicle, and vehicle contents being thrown against the occupant. Additional study inclusion criteria for FROIs included: (a) all accepted FROIs, including open and closed FROIs; (b) FROIs including all ages and those with unknown ages; and (c) FROIs regardless of the injury location (in-state and out-of-state FROIs). Based on these criteria, 11,790 FROIs were selected for inclusion.

2.3. Identification of industry, truck type, and local transportation

North American Industry Classification System (NAICS) super sectors were identified based on BLS classification (https://www.bls.gov/sae/additional-resources/naics-supersectors-for-ces-program.htm). NAICS codes were poorly populated in the final data set (3,017 of the 11,390 total had no or incomplete codes), and Standard Occupational Classification (SOC) codes were not available in the dataset, though an occupation description was included. To identify industry and occupation codes (SOC codes), the data set was processed using the National Institute for Occupational Safety and Health's (NIOSH's) auto-coding program (the NIOSH Industry and Occupation Computerized Coding System, or NIOCCS, https://csams.cdc.gov/nioccs/), a reasonably effective data set auto-coder for industry and occupation (Buckner-Petty, Dale, & Evanoff, 2019; Schmitz & Forst, 2016). Seventy-four percent (n = 8,373) of the total FROIs (n = 11,390) were coded with NIOCCS for industry and occupation. The selected FROIs were then narrowed to those coded with a heavy truck driver occupation SOC code of 53-3032 or a light (and medium) truck driver occupation 53-3033 SOC code and manually reviewed to exclude long-distance driver FROIs (n = 146 [145 heavy truck and one light/medium truck driver]) and passenger FROIs (n = 54); 800 final FROIs remained, and 32 of the 800 (4%) required additional manual review by three reviewers to assign codes. The final data set contained 800 total local crash FROIs: 451 heavy truck driver and 349 light or medium truck driver crash FROIs.

Analysis of the heavy, medium, and light truck FROI free-text narratives and industry and occupation codes was performed to identify the specific truck type involved in the crash and to better understand the precipitating factors that led to the driver injury. Truck classifications conformed to the Federal Highway Administration's (FHWA's) vehicle and weight class definitions (https://afdc.energy.gov/data/10380):

1. Heavy vehicle (26,001 lbs. and over): furniture truck, high-profile semi, home fuel truck, medium semi-tractor, refuse (solid waste) vehicle, tow truck, cement mixer, dump truck, fire truck, fuel truck, heavy semi-tractor, refrigerated van, and semi sleeper;
2. Medium vehicle (10,001 lbs. to 26,000 lbs.): beverage truck, rack truck, single axle van, stake body truck, bucket truck, city delivery truck, large walk-in vehicle, conventional van, landscape utility vehicle, and medium walk-in vehicle; and.
3. Light vehicle (0 to 10,000 lbs.): Utility van-type trucks.

Light and medium truck FROIs were combined for the analysis, as SOC codes are not specific for light and medium truck drivers in the FROIs and the NIOCCS needed to be used to code truck driver occupation. Only two SOC codes delineate truck type: heavy truck driver and light truck driver. SOC defines light truck drivers as those who "Drive a light vehicle, such as a truck or van, with a capacity of less than 26,001 pounds Gross Vehicle Weight (GVW), primarily to pick up merchandise or packages from a distribution center and deliver. May load and unload vehicle." Both medium trucks (10,001 to 26,000 pounds) and light trucks (0 to 10,000 pounds) fall within the SOC light truck driver code, so they could not be separated.

FROIs do not have a "vehicle type" data variable as crash data does, so the NIOCCS code was the only code available to us to identify truck type. To supplement the NIOCCS code, the injury narratives were searched for specific truck type; however, only 42% of the FROI narratives mentioned the general vehicle type. Mentions of solid waste trucks, concrete mixers, tow trucks, coal trucks, semi-trucks, and dump trucks in the FROI narrative were included in the heavy truck category, and mentions of delivery trucks, large vans, and other vehicles such as box trucks in the FROI narrative were included in the light and medium truck category.

Local transportation was identified through National Council on Compensation Insurance (NCCI) class codes (used to categorize and classify businesses to underwrite workers' compensation insurance); NAICS codes; and manual review of the industry field in the FROIs. Codes used to identify local transportation were: (a) NCCI class code 7228, Trucking: Local Hauling Only—All Employees and Drivers; (b) NAICS 48422, Specialized Freight (except Used Goods) Trucking, Local ("Local trucking establishments provide trucking within a metropolitan area that may cross state lines. Generally, the trips are same-day return"); (c) NAICS 484110, General Freight Trucking, Local ("Local general freight trucking establishments usually provide trucking within a metropolitan area which may cross state lines"); and (d) driver residence and crash location county. Specific distances were not available in the FROIs, but crashes were identified as local when the residence county was in Kentucky and the crash location county occurred in Kentucky or in a border state. Free-text industry names in the industry data field were also determined by the study authors to be local transportation (see definition above for NAICS codes) if the listed industries were florists, furniture stores, automotive parts and accessories stores, and other industries highly likely to only serve a local area. Long-distance transportation FROIs identified through NAICS code 48412, General freight trucking, long-distance, and key words in the free text (long-haul, over-the-road, etc.) were excluded from the study. Two reviewers reviewed the truck type and local transportation coding by the first reviewer and inter-reliability checks were performed.

Driver residence and crash location counties were determined through analysis of county names. Appalachian county designation was used as the proxy to identify rural versus urban area truck crashes. The counties identified as Appalachian in this study were derived from the Appalachian Regional Commission (https://www.arc.gov/appalachian-counties-served-by-arc/).

The institutional review board approved the study; because the study involved the analysis of secondary data with no personal identifiers, informed consent was neither required nor obtained.

2.4. Groupings of injured body parts, nature of Injury, and lost time

Injured body part and nature of injury codes are defined by the KDWC using the International Association of Industrial Accident Boards and Commissions' coding framework (https://www.wcio.org/Document%20Library/InjuryDescriptionTablePage.aspx). Injured body parts coded in FROIs were collapsed into five categories: (1) head, face, and neck; (2) back, torso, chest, abdomen, and groin; (3) upper extremities including shoulder; (4) lower extremities including pelvis; and (5) multiple body systems, whole body, or other. Nature of injury codes were collapsed into five categories: (1) concussion; (2) contusion/laceration; (3) fracture/dislocation; (4) strain/sprain; and (5) other. The "other" categories for both Nature of Injury and Body Part included FROIs where the individuals required medical attention but the record did not list a traditional injury (e.g., pregnancy concerns, elevated blood pressure).

Extent of lost time was grouped into three distinct categories (no lost time, lost time [defined as greater than one day of lost time], and fatality [first report of injury was a fatal injury]) and grouped by days of lost time (0–1 day, 2–6 days, 7–29 days, and 30+ days of lost time). Job tenure and days of lost time due to injury were calculated using dates in the dataset. Job tenure was defined as the number of days between the reported hire date and date of injury. Number of days of lost time due to injury was defined as the number of days between the date of injury and the return to work date. Completion of the return to work date field is optional in the FROI, so the days of lost time due to injury was poorly populated (59% missing data).

2.5. Injury scenarios

Free-text narratives were analyzed using keyword searches and manual review of each FROI to identify the specific truck types by the leading industries. Two reviewers reviewed the narrative coding of the truck type by the first reviewer and inter-reliability checks were performed. The top injury activity scenarios were described for each major truck group type that was identified.

2.6. Statistical analysis

This study incorporated a cross-sectional design. Frequencies were determined for demographics, industry, injury outcomes (e.g., injured body part, lost time, and cause and nature of injury), and award disposition variables. Chi-square tests were performed to assess the significant differences between the heavy truck, and medium and light truck local crash FROI groups on the above variables. All statistical analysis was performed using SAS Enterprise 8.2.

Denominator numbers by age and gender were not available to calculate age-adjusted local crash FROI rates; instead, Kentucky crude FROI rates for the heavy truck driver (SOC code 53-3032) and for the light/medium truck driver (53-3033 SOC code) occupational categories were calculated for years 2010–2019. Long distance could not be separated from local distance for the denominator, so two numerator types were utilized to develop FROI rates: (1) the final 451 heavy truck driver and 349 light or medium truck driver *local crash FROIs* based on all exclusion and inclusion study criteria and (2) the 596 heavy truck driver and 350 light/medium truck driver *local and long distance crash FROIs* based on all exclusion and inclusion criteria with the exception

Table 1
Demographic characteristics of heavy vs light and medium truck driver local crash first reports of injuries, 2010–2019.[1]

Demographic Characteristic	All First Reports of Injuries n = 800	Light/Medium Truck First Reports of Injuries n = 349 (%)	Heavy Truck First Reports of Injuries n = 451 (%)
Sex	**n = 798**	**n = 348**	**n = 450**
Male	732	300 (86%)	432 (96%)
Female	66	48 (14%)	18 (4%)
Age (Years)	**n = 800**	**n = 349**	**n = 451**
<25	60	42 (12%)	18 (4%)
25–44	321	140 (40%)	181 (40%)
45–65	360	134 (38%)	226 (50%)
>65	59	33 (9%)	26 (6%)
Job Tenure	**n = 735**	**n = 323**	**n = 412**
Less than 1 year	366	152 (47%)	214 (52%)
1–4 years	200	96 (30%)	104 (25)
5–10 years	91	40 (12%)	51 (12%)
Over 10 years	78	35 (11%)	43 (10%)
NAICS Industry[1]	**n = 800**	**n = 349**	**n = 451**
Natural Resources & Mining	18	<5	17 (4%)
Construction	37	10 (3%)	27 (6%)
Manufacturing	60	23 (7%)	37 (8%)
Trade, Transportation, & Utilities	518	245 (70%)	273 (61%)
Wholesale Trade	*105*	72 (21%)	33 (7%)
Retail Trade	*187*	164 (47%)	23 (5%)
Transportation & Warehousing	*215*	9 (3%)	206 (46%)
Utilities	*11*	0 (0%)	11 (2%)
Information	<5	<5	0 (0%)
Financial Activities	6	<5	<5
Professional & Business Services	80	18 (5%)	62 (14%)
Education & Health Services	<5	0 (0%)	<5
Leisure & Hospitality	30	19 (5%)	11 (2%)
Other Services	42	25 (7%)	17 (4%)
Government	6	<5	<5
County of Injury Region	**n = 800**	**n = 349**	**n = 451**
Appalachia	246	74 (21%)	172 (38%)
Non-Appalachia	504	254 (73%)	250 (55%)
Out of State	50	21 (6%)	29 (6%)

[1] Numbers less than five are suppressed in accordance with state data management policy.

of local versus long distance. The denominator data used for both numerator types were Kentucky heavy and light/medium truck driver occupational employment data (local plus long distance) obtained from BLS Occupational Employment and Wage Statistics (https://www.bls.gov/oes/tables.htm).

3. Results

3.1. Demographic characteristics

While the majority of all local light, medium, and heavy truck driver crash FROIs involved males, there was a higher percentage of female light and medium truck driver local crash FROIs (14%) compared to female heavy truck crash FROIs (4%) (Table 1).

The highest proportion of local heavy truck crash FROIs was for the 45–65 year age driver group (50%), while the 24–44 year age driver group represented the highest proportion of local light and medium truck crash FROIs (40%). The youngest (less than 25 years) age group was represented at a higher proportion in local light and medium truck crash FROIs compared to the same age group in the heavy truck crash FROIs (12% of young local light and medium truck driver FROIs vs 4% of younger heavy truck driver FROIs). There was no difference between the light and medium truck driver and heavy truck driver FROIs in regards to job tenure. About one-half of the heavy truck and light and medium truck FROIs were for workers with less than one year of job tenure, and between 25% and 30% were for workers with one to four years of job tenure.

There was a higher percentage of heavy truck driver crash FROIs in the transportation and warehousing industry (46%), and professional and business services industries (14%) compared to the light/medium truck driver FROIs that had the highest percentage of FROIs in the retail (47%) and wholesale trade (21%) industries. A higher percentage of the heavy truck driver local crash FROIs occurred in the rural Appalachian region compared to local light and medium vehicle driver crash FROIs (38% among heavy vehicle drivers vs 21% among light and medium vehicle drivers). In contrast, a higher percentage of local light and medium truck crash FROIs occurred in urban non-Appalachian regions (73%) compared to heavy truck crash FROIs (55%).

3.2. Driver injury characteristics

Collision or sideswipe with another vehicle was the primary cause of injury in the light and medium truck local crash FROIs (46%) and the second highest cause of injury in the heavy truck local crash FROIs (32%). A higher percentage of heavy truck driver crash injuries was due to vehicle upset, rollover, or jackknife (27%) compared to the light and medium truck local crash FROIs (9%). Crashes not otherwise classified, including sudden start or stop, represented 37% of both heavy and light/medium vehicle FROIs (Table 2).

Sprains and strains accounted for the highest percentage of light and medium truck driver crash FROIs (39%), whereas "Other" was the primary injury type listed for local heavy truck driver crash FROIs (43%), followed by sprains and strains (24%). When body part injured was examined, "multiple injuries" accounted for similar percentages of the light and medium truck driver crash FROIs (41%) compared to local heavy truck driver crash FROIs (45%). Multiple parts was the most frequent body part injured for both truck size FROIs. Head and neck injuries were more common

Table 2
Driver injury characteristics in heavy vs light and medium truck local crash first reports of injuries, 2010–2019.

Injury Characteristic	Light and Medium Truck First Reports of Injury n = 349 (%)	Heavy Truck First Reports of Injury n = 451 (%)	Chi-Square *p*-value
Cause of Injury	**n = 349**	**n = 451**	**<0.0001**
Collision or sideswipe with another vehicle	162 (46%)	145 (32%)	
Collision with fixed object	24 (7%)	19 (4%)	
Vehicle upset, rollover, or jackknife	33 (9%)	122 (27%)	
Not Otherwise Classified, including sudden start or stop	130 (37%)	165 (37%)	
Nature of Injury	**n = 349**	**n = 451**	**<0.001**
Concussion	8 (2%)	7 (2%)	
Contusion/laceration	52 (15%)	92 (20%)	
Fracture/dislocation	29 (8%)	50 (11%)	
Sprain/strain	135 (39%)	108 (24%)	
Other	125 (36%)	194 (43%)	
Body Part Injured	**n = 324**	**n = 437**	**0.410**
Head, face, and neck	51 (16%)	50 (11%)	
Back, torso, chest, abdomen, and groin	71 (22%)	88 (20%)	
Upper extremities, including shoulder	45 (14%)	64 (15%)	
Lower extremities, including pelvis	25 (8%)	37 (8%)	
Multiple parts, whole body, or other	132 (41%)	198 (45%)	

Table 3
Disposition status and lost work time in heavy vs light and medium truck local crash first reports of injury, 2010–2019.[1]

Disposition and Lost Time Status	Light/Medium Truck First Reports of Injury n = 349 (%)	Heavy Truck First Reports of Injury n = 451 (%)	Chi-Square *p*-value
First Report of Injury Resulted in Workers' Compensation Award[2]	**n = 349**	**n = 451**	**0.644**
No	263 (77%)	324 (76%)	
Yes	78 (23%)	104 (24%)	
Extent of Lost Time due to Injury	**n = 349**	**n = 451**	**<0.01**
No lost time	23 (7%)	46 (10%)	
Lost time	323 (93%)	389 (86%)	
Fatality	<5	16 (4%)	
Days of Lost Time due to Injury[3]	**n = 349**	**n = 451**	**0.064**
0–1 day	16 (13%)	38 (19%)	
2–6 days	45 (37%)	56 (27%)	
7–29 days	28 (23%)	36 (18%)	
30+ days	32 (26%)	75 (37%)	
Missing Return- to-Work Date[4]	228	246	

[1] Numbers less than five are suppressed in accordance with state data management policy.
[2] FROI resulted in Workers' Compensation award if disposition was in agreement approved—Administrative Law Judge (ALJ), award—ALJ, lump sum agreement on first report, agreement approved on first report. FROI resulted in no award if disposition was none, case dismissed—ALJ, consolidated ALJ dismissal, consolidated ALJ no money, medical dispute dismissed/denied. FROIs under review included assigned to ALJ, held in abeyance, medical dispute closed, ready to set for pre-hearing conference, scheduled for pre-hearing, set for hearing, submitted for ALJ decision, medical dispute set for proof time, proof time, ALJ opinion, medical dispute program.
[3] The days off work after injury field, being not required, is poorly populated, with 59% of 800 observations missing.
[4] Excluded from statistical analysis.

among light and medium truck driver crash FROIs (16%) compared to heavy truck driver FROIs (11%).

3.3. Disposition status and lost work time

Approximately-three-quarters of the FROIs for both groups did not result in a workers' compensation award, a proxy for injury severity since there was no settlement benefit for disability, impairment, or death; medical benefits could have been paid out on the FROI, but medical benefit data were not available, as KDWC does not maintain the medical data set. There was a significantly higher percentage of light and medium truck driver local crash FROIs with lost time due to injuries compared to heavy truck driver FROIs (93% of light and medium truck driver FROIs vs 86% of heavy truck driver FROIs) (Table 3).

While there was no significant difference between the two groups for the number of days of lost time, there was an indication that the drivers in heavy vehicle local crash FROIs had a higher percentage of 30 days or more of lost work time (37%) compared to drivers in light and medium vehicle FROIs (26%).

3.4. Driver injuries by truck type

Through free-text narrative analysis, the primary truck types identified in the heavy truck crash FROIs were semi-trucks (13%), dump trucks (11%), and solid waste trucks, tow trucks, and coal trucks (6% each of the total) (Table 4).

This is an undercount of the identification of the specific truck types since 54% of all heavy truck crash FROIs did not contain enough information in the free-text to identify the specific vehicle involved in the crash injury report. Not surprisingly, general freight trucking (local) and local trucking industries were listed most frequently for the semi-truck crashes and dump truck crash FROIs. Solid waste collection and refuse systems were the industries most frequently listed for the solid waste truck crash FROIs. Surprisingly, general freight trucking (local) was the industry listed for almost

Table 4
Driver injuries in heavy vs light and medium truck local crash first reports of injury by vehicle type and leading primary industry, 2010–2019.

Heavy Truck Type	n = 451 (%)	Leading Primary Industries	n (%)
Semi-Truck	59 (13%)		
		General Freight Trucking (Local), Local Trucking	17 (29%)
		Recyclable Material Merchant Wholesalers	7 (12%)
Dump Truck	50 (11%)		
		General Freight Trucking (Local)	18 (37%)
		Site Preparation Contractors	6 (12%)
Solid Waste Truck	27 (6%)		
		Solid Waste Collection	14 (52%)
		Refuse Systems	11 (41%)
Tow Truck	26 (6%)		
		General Freight Trucking (Local)	11 (42%)
		Automotive Services, General Automotive Repair	8 (31%)
Coal Truck	25 (6%)		
		General Freight Trucking (Local)	17 (68%)
		Bituminous Coal Underground Mining, Support Activities for Coal Mining	8 (32%)
Concrete Mixer	20 (4%)		
		Ready-Mix Concrete Manufacturing, Ready-Mixed Concrete, Concrete Work	16 (80%)
Unidentified Truck	244 (54%)		
Light and Medium Truck Type	**n = 349 (%)**	**Primary Industries**	**n (%)**
"Delivery Truck"	104 (30%)	Furniture	20 (19%)
		Automotive Parts and Accessories	13 (13%)
		Groceries and Soft Drinks	9 (9%)
		Florist	7 (7%)
		Motor Vehicle Supplies and New Parts Merchant Wholesalers	8 (8%)
		Tire Dealers and Manufacturing	6 (6%)
		Pharmacies and Drug Stores	6 (6%)
Van	9 (3%)		
Other Truck Types	14 (4%)		
Unidentified Trucks	**222 (64%)**		

half of the tow truck crash FROIs (42%); only 23% listed automotive services and general automotive repair as the associated industry.

For light and medium truck FROIs, the specific vehicle types identified were delivery trucks (30%), other vehicles such as box trucks (4%), and vans (3%). These percentages are again an undercount of the identification of the specific truck types since 64% of all light and medium truck FROIs were unidentified in the free-text narrative. The industries most frequently listed for delivery truck FROIs were the furniture industry (19%), automotive parts and accessories (13%), and groceries and soft drinks (9%).

3.5. Injury scenarios

The top injury scenarios associated with each identified truck type were extracted from the free-text narratives (Table 5).

For semi-truck FROIs, the top injury scenarios involved rollovers (41%) in situations such as the inability to stop in time when a vehicle was stopped or slowed in front of them. For dump truck FROIs, top injury scenarios involved rollovers (50%) and incidents where the dump truck ran off the road (20%) in incidents such as driving off the lanes of narrow roadways and overcorrecting while trying to bring the vehicle back onto the pavement, as well as the inability to stop in time for red lights, indicating that the truck may have been going too fast while approaching the intersection. Coal trucks running off the roadway (48%) in situations such as driving downhill and not being able to downshift was the most common scenario, as was jumping from the vehicle while it was in motion.

In the light and medium truck local crash FROI group, top injury scenarios for the furniture industry FROIs involved being rear-ended by other vehicles (35%) while the furniture truck was in motion or while stopped at red lights. The most common injury scenarios in the automotive parts and accessories industry FROIs involved trucks being rear-ended (46%). For the groceries and soft drinks industry FROIs, being rear-ended (67%) while in motion or while stopped were the most common injury scenarios.

Fig. 1 shows the heavy and medium/light trucker driver FROI rates by occupational category (SOC codes). The heavy truck driver occupation (local and local distance) comprised 1.5% of all Kentucky employment occupations, whereas the medium/light truck driver occupation comprised only 0.7% of all Kentucky employment occupations (data not shown). Using our study exclusion and inclusion criteria for the numerator (local distance only), the light/medium truck driver FROI rate was 64% higher than the heavy truck driver occupation FROI rate. When we examined the rate before our final long distance exclusion criterion was applied, the light/medium truck driver FROI rate (number of local + long distance FROIs/number of local + long distance truck drivers employed) was still 24% higher than the heavy truck driver FROI rate, indicating that, regardless of distance driven, light/medium truck drivers had higher FROI rates than heavy truck driver crash FROI rates.

4. Discussion

Our study results show that there were higher percentages of collision or sideswipe with another vehicle crash, lost time, and rear end crash FROIs involving light and medium truck drivers compared to FROIs involving heavy truck driver. Almost one-half of the light and medium truck driver local crash FROIs occurred in the retail trade industry, and approximately-three-quarters occurred in urban areas (non-Appalachia). The 2018 National Occupational Research Agenda (NORA) for the Wholesale and Retail Trade industries recommends additional motor- vehicle crash research on wholesale and retail trade drivers by vehicle type, driver abilities, and the need for refresher training (NORA, 2018). NORA also recommends identification of risk factors for the observed elevated transportation incidence rates in the following industry sectors: automotive parts and accessories, grocery and related product wholesalers, motor vehicles and parts wholesalers, and druggist goods and merchant wholesalers. The industries des-

Table 5
Top driver injury scenarios in heavy vs light and medium truck local crash first reports of injury, 2010–2019.[1]

Heavy Truck Type	Primary Industry	Incident type	Top Injury Scenarios
Semi-Truck (n = 59)	General Freight Trucking (Local), Local Trucking	Rollover (n = 24; 41%) Rear-end (n = 10; 17%) Ran off roadway (n = 6; 10%) Other (Sideswipe/backing up/Head-on/Pothole/Struck object) (n = 10; 17%) Not enough information (n = 9; 15%)	"Driver was driving the tractor/trailer on a haul road when he went into a curve and turned the tractor/trailer over." "He was driving a semi-truck. A car stopped in front of him to avoid an accident. He could not get stopped in time to avoid impact." "While employee was driving tractor/trailer unit he encountered a motorcycle in his lane. He swerved to avoid hitting motorcycle and the unit overturned."
Dump Truck (n = 50)	General Freight Trucking (Local)	Rollover (n = 25; 50%) Ran off roadway (n = 10; 20%) Rear-end (n = 8; 16%) Sideswipe/Not enough information (n = 7; 14%)	"Operating a dump truck dropped off roadway and overcorrected turning the truck on its side." "Driving dump truck on narrow road, overcorrected and turned truck over." "While driving a dump truck the worker applied the brakes and crossed an intersection and turned over." "Hauling dirt in dump truck. Worker approached a yellow light and couldn't stop the truck. Turned left to avoid oncoming car and tipped the truck." "Driver was loaded delivering rock when shoulder broke off the road and truck overturned."
Coal Truck (n = 25)	General Freight Trucking (Local)	Ran off roadway (n = 12; 48%) Rollover (n = 7; 28%) Head-on/Rear-end/Sideswipe/Not enough information (n = 6; 24%)	"Transporting driver was shifting gears going downhill and could not get truck in gear." "On way to pick up coal traveling on a haul road. Going downhill lost control and turned over." "Driving truck transporting coal; couldn't change gears and truck wouldn't stop so had to jump out of moving truck." "Worker was transporting coal when jumped from a moving vehicle."
Light and Medium Truck Type	**Primary Industry**		**Injury Scenarios**
"Delivery Truck"	Furniture (n = 20)	Rear-end (n = 7; 35%) Ran off roadway/Rollover/Sideswipe/Struck object (n = 8; 40%) Not enough information (n = 5; 25%)	"Employee was in delivery truck; motor vehicle struck truck in rear." "Employee was heading to customer's house for delivery in company truck when struck from behind by another vehicle." "Stopped in company vehicle at stop light when he was hit in the rear." "Rear-ended by another vehicle (passenger)."
	Automotive Parts and Accessories (n = 13)	Rear-end (n = 6; 46%) Struck by/struck object/Not enough information (n = 7; 54%)	"Driver was rear ended & has head injury." "Lost control of vehicle, ran into median and hit a tree. This was the first day of work for employee." "Employee was stopped at red light and was rear-ended." "Driver alleges he was blinded by the sun and did not see the red light. Allegedly ran the red light into the path of other vehicle." "Vehicle allegedly turned left in front of other vehicle and was struck."
	Groceries and Soft Drinks (n = 9)	Rear-end (n = 6; 67%) Not enough information (n = 3; 33%)	"Associate was rear-ended by a passenger vehicle." "Associate was involved in a motor vehicle accident. Struck from behind while stopped." "Employee was making a delivery and was rear-ended." "Employee was involved in a rear-end collision and sustained injuries."

[1] Numbers less than five are suppressed in accordance with state data management policy so incident types are collapsed into broader categories with the exception of the "Not enough information" category in Groceries and Soft Drinks.

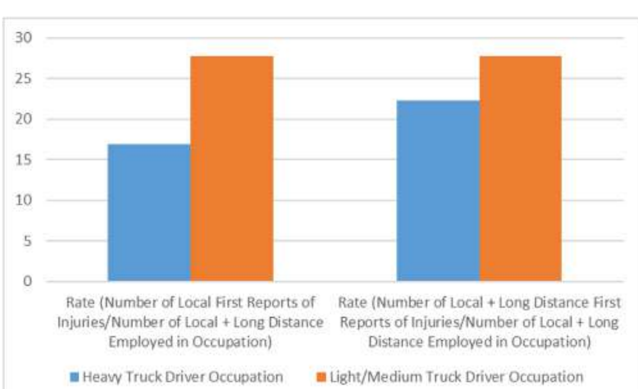

Fig. 1. Heavy and light/medium truck driver occupational first report of injury rates, 2010–2019.

ignated by NORA to need motor-vehicle crash research correspond to the industries and nature of the trucks identified in this study in the free-text narratives.

The NIOSH Center for Motor Vehicle Safety Strategic Plan identified the transportation and retail trade industries as priority industries for moto- vehicle safety research (NIOSH, 2020), particularly regarding the need for motor-vehicle safety programs in the wholesale and retail trade industries. There was a high percentage of FROIs for younger drivers of light/medium trucks who were employed in the retail trade industry (48%). Enhanced refresher driver safety training for younger drivers who have been in previous crashes with injuries is needed in the retail trade industry and especially in the furniture, automotive parts and accessories, and grocery subsectors identified in this study. Vivoda, Pratt, & Gillies, 2019, identified safety practices and policies such as road safety program duration and timely updating, company safety commitment, driver training, crash review and scorecard, and fatigue risk management as 'necessary ingredients' for improving driver safety on the roads.

Overall, light and medium truck driver FROIs involved more lost time due to injuries (higher percentages of crashes with another vehicle [rear-end crashes], sprains and strains, as well as head and neck-related and back and torso-related injuries) compared to heavy truck driver FROId. Due to the high frequency of light and medium truck rear-end collision FROIs described in the free-text injury scenarios, light and medium truck employers should consider the inclusion of targeted curricula for light and medium truck drivers with previous crash FROIs that addresses distracted driving and emphasizing the prevention of rear-end crashes.

Our top injury scenarios for semi-truck rollover FROIs correlate with previous findings by McKnight and Bahouth (2009), where the authors found that one-half of the rollovers were due to high speed, unsafe brakes, and intersections. This study shows that local dump truck driver and coal truck driver FROIs also had a high frequency of rollovers due to tires dropping off the shoulder of the road and to losing control from being unable to downshift while driving downhill. States should consider implementing refresher heavy truck driver safety training, particularly for drivers with previous crash FROIs, that includes driving on narrow roads and roadway departure prevention and shifting gears on hills, in addition to speeding, unsafe brakes, and intersections, as highlighted by McKnight and Bahouth (2009).

4.1. Limitations

This study was limited in the ability to identify the specific truck types in the FROIs. Vehicle type is not a mandatory data field within KDWC; therefore, we needed to rely on the mention of the specific truck type in the free-text narrative. Approximately 60% of all local truck FROIs did not contain enough information in the free text to identify the specific vehicle types involved in the crash; therefore, the results may not be generalizable to all local-distance light, medium, and heavy truck crash FROIs.

Also, a study limitation was that this is a database of first reports of injuries and not a database of available truck drivers, therefore, conclusions are limited to those who were involved in a crash with a first report of injury. This is a first study to comprehensively describe injuries of light and medium truck drivers involved in local crashes compared to heavy truck drivers.

Last, another limitation was the inability to calculate age-adjusted FROI rates to accurately measure younger truck driver exposures. In the absence of Kentucky light, medium, and heavy truck driver occupation employment data by age and/or distance driven, we were only able to calculate crude FROI rates by heavy truck and light/medium truck occupation employment. Using this denominator, we show that, overall, FROI rates for light and medium truck drivers had higher crash rates compared to FROIs for drivers of heavy trucks.

5. Conclusions and practical applications

The utilization of WC data highlighted top injury scenarios and specific vehicle types for targeting driver safety training among truck drivers, particularly light/medium truck drivers. Road safety policies regarding driver training, crash reviews, and in-vehicle monitoring systems are needed for truck drivers, especially light and medium truck drivers. Implementation of entry-level (less than one year of employment) and refresher driver training for light/medium truck drivers should be considered to reduce truck driver crashes, similar to Federal Motor Carrier Safety Administration-mandated heavy truck entry-level driver training. In addition, in-vehicle monitoring systems have promising effectiveness in increasing driver safety (Furlan et al., 2020), particularly when monitoring includes both supervisory review and discussion as well as in-cab warning lights (Bell, Taylor, Chen, Kirk, & Leatherman, 2017).

Enhanced driver safety training on speeding on narrow roadways, nearing intersections, and downshifting on hills is needed for drivers of heavy trucks, particularly drivers of semi-trucks, dump trucks, and coal trucks. Driver safety training on the prevention of rear-end crashes (e.g., gradual stopping and checking mirrors) is needed for light and medium truck drivers of furniture trucks, automotive parts and accessories trucks, and groceries and soft drink trucks.

Future studies using WC data linked with crash data are needed to comprehensively identify the specific industries, vehicle types, and circumstances surrounding light and medium truck crashes.

Disclosure statement

The authors have no conflict of interest or financial interest to disclose.

Funding Statement

This work was supported by the Centers for Disease Control and Prevention under Cooperative Agreement Number 6 U60 OH 008483-16. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services. The funding source had no involvement in the study design; data collection, analysis and interpretation of data; writing of the manuscript; or in the decision to submit the article for publication.

References

Bell, J. L., Taylor, M. A., Chen, G., Kirk, R. D., & Leatherman, E. R. (2017). Evaluation of an in-vehicle monitoring system (IVMS) to reduce risky driving behaviors in commercial drivers: Comparison of in-cab warning lights and supervisory coaching with videos of driving behavior. *Journal of Safety Research, 60*, 125–136. https://doi.org/10.1016/j.jsr.2016.12.008. Epub 2016 Dec 21. PMID: 28160807 PMCID: PMC5427714.

Buckner-Petty, S., Dale, A. M., & Evanoff, B. A. (2019). Efficiency of autocoding programs for converting job descriptors into standard occupational classification (SOC) codes. *American Journal of Industrial Medicine, 62*(1), 59–68. https://doi.org/10.1002/ajim.22928. Epub 2018 Dec 5 PMID: 30520070.

Bunn, T., Slavova, S., & Robertson, M. (2013). Motor vehicle injuries among semi truck drivers and sleeper berth passengers. *Journal of Safety Research, 44*, 51–55. https://doi.org/10.1016/j.jsr.2012.09.003. Epub 2012 Nov 20.

Bunn, T. L., Slavova, S., & Rock, P. J. (2019). Association between commercial vehicle driver at-fault crashes involving sleepiness/fatigue and proximity to rest areas and truck stops. *Accident Analysis and Prevention, 126*, 3–9. https://doi.org/10.1016/j.aap.2017.11.022. Epub 2017 Nov 22.

Bureau of Labor Statistics (2020). U.S. Department of Labor, in cooperation with state, New York City, District of Columbia, and federal agencies. December 2020. Census of fatal occupational injuries. URL accessed on June 24, 2021, at https://www.bls.gov/iif/oshwc/cfoi/cftb0331.htm.

Bureau of Labor Statistics (2021). U.S. Department of Labor, in cooperation with participating state agencies. Survey of occupational injuries and illnesses. URL accessed on June 25, 2021, at https://www.bls.gov/iif/oshwc/osh/os/summ1_00_2019.htm.

Chandler, M. D., Bunn, T. L., & Slavova, S. (2017). Narrative and quantitative analyses of workers' compensation-covered injuries in short-haul vs. long-haul trucking. *International Journal of Injury Control and Safety Promotion, 24*(1), 120–130. https://doi.org/10.1080/17457300.2016.1170041. Epub 2016 May 4. PMID: 27142559.

Chen, G. X., Amandus, H. E., & Wu, N. (2014). Occupational fatalities among driver/sales workers and truck drivers in the United States, 2003–2008. *American Journal of Industrial Medicine, 57*(7), 800–809. https://doi.org/10.1002/ajim.22320. Epub 2014 May 8.

Combs, B., Heaton, K., Raju, D., Vance, D. E., & Sieber, W. K. (2018). A descriptive study of musculoskeletal injuries in long-haul truck drivers: A NIOSH national survey. *Workplace Health & Safety, 66*(10), 475–481.

Doecke, S. D., Baldock, M. R. J., Kloeden, C. N., & Dutschke, J. K. (2020). Impact speed and the risk of serious injury in vehicle crashes. *Accident Analysis and Prevention, 144*, 105629. https://doi.org/10.1016/j.aap.2020.105629. Epub 2020 Jun 19.

Furlan, A. D., Kajaks, T., Tiong, M., Lavallière, M., Campos, J. L., Babineau, J., et al. (2020). Advanced vehicle technologies and road safety: A scoping review of the evidence. *Accident Analysis and Prevention, 147*, 105741. https://doi.org/10.1016/j.aap.2020.105741. Epub 2020 Sep 24 PMID: 32979820.

Hanowski, R. J., Hickman, J. S., Wierwille, W. W., & Keisler, A. (2007). A descriptive analysis of light vehicle–heavy vehicle interactions using in situ driving data. *Accident Analysis and Prevention, 39*(1), 169–179. https://doi.org/10.1016/j.aap.2006.06.016. Epub 2006 Aug 24 PMID: 16934736.

Karaca-Mandic, P., & Lee, J. (2014). Hospitalizations and fatalities in crashes with light trucks. *Traffic Injury Prevention, 15*(2), 165–171. https://doi.org/10.1080/15389588.2013.803279.

McKnight, A. J., & Bahouth, G. T. (2009). Analysis of large truck rollover crashes. *Traffic Injury Prevention, 10*(5), 421–426. https://doi.org/10.1080/15389580903135291. PMID: 19746305.

NIOSH (2020). NIOSH Center for motor vehicle safety strategic plan, 2020–2029. By Pratt S, Retzer K, Rodríguez-Acosta R, Olsavsky R, Fosbroke D. Morgantown. WV: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health. DHHS (NIOSH) Publication No. 2020-126.

NORA (2018). National occupational research agenda for wholesale and retail trade. NORA Wholesale and Retail Trade Sector Council, August 2018.

Pipkorn, B., Iraeus, J., Lindkvist, M., Puthan, P., & Bunketorp, O. (2020). Occupant injuries in light passenger vehicles—A NASS study to enable priorities for development of injury prediction capabilities of human body models. *Accident Analysis and Prevention, 138*, 105443. https://doi.org/10.1016/j.aap.2020.105443. Epub 2020 Feb 12.

Pratt, S. G., & Bell, J. L. (2019). Analytical observational study of nonfatal motor vehicle collisions and incidents in a light-vehicle sales and service fleet. *Accident Analysis and Prevention, 129*, 126–135. https://doi.org/10.1016/j.aap.2019.05.009. Epub 2019 May 28 PMID: 31150919.

Schmitz, M., & Forst, L. (2016). Industry and occupation in the electronic health record: An investigation of the national institute for occupational safety and health industry and occupation computerized coding system. *JMIR Medical Informatics, 4*(1), e5. Published online 2016 Feb 15. PMCID: PMC4771928.

Stuckey, R., Glass, D. C., LaMontagne, A. D., Wolfe, R., & Sim, M. R. (2010a). Risk factors for worker injury and death from occupational light vehicle crashes in New South Wales (Australia). *American Journal of Industrial Medicine, 53*(9), 931–939. https://doi.org/10.1002/ajim.20854. PMID: 20698024.

Stuckey, R., LaMontagne, A. D., Glass, D. C., & Sim, M. R. (2010b). Estimating fatality rates in occupational light vehicle users using vehicle registration and crash data. *Australian and New Zealand Journal of Public Health, 34*(2), 142–145. https://doi.org/10.1111/j.1753-6405.2010.00498.x. PMID: 23331357.

Vivoda, J. M., Pratt, S. G., & Gillies, S. J. (2019). The relationships among roadway safety management practices, collision rates, and injury rates within company fleets. *Safety Science, 120*, 589–602.

Zheng, Z., Lu, P., & Lantz, B. (2018). Commercial truck crash injury severity analysis using gradient boosting data mining model. *Journal of Safety Research, 65*, 115–124. https://doi.org/10.1016/j.jsr.2018.03.002. Epub 2018 Mar 17.

Terry Lee Bunn, PhD, is a Professor of Preventive Medicine and Environmental Health in the University of Kentucky's College of Public Health and directs the Kentucky Injury Prevention and Research Center. Research interests include the prevention of motor vehicle injuries in the working and general populations, and the prevention of drug overdoses.

Madison Liford, MPH, is an Epidemiologist at the Kentucky Injury Prevention and Research Center focusing on occupational injury. Madison works primarily with the Kentucky Fatality Assessment and Control Evaluation (FACE) and Kentucky Occupational Safety and Health Surveillance (KOSHS) programs.

Michael David Turner, MS, is the Project Manager of the Kentucky Occupational Safety and Health Surveillance Program within the Kentucky Injury Prevention and Research Center. Research interests include the prevention of work-related injuries to older workers and foreign-born workers.

Ashley M. Bush, DrPH is a Research Program Administrator in the Kentucky Injury Prevention and Research Center, where she serves as the Practice Core Director. Research interests include injury surveillance and prevention, specifically transportation safety, child maltreatment, occupational safety, and rural health disparities. She values the present and future partnerships essential for effectively and efficiently minimizing the risk of injury, violence, and associated risk factors.

JOURNAL OF SAFETY RESEARCH

Editor-in-Chief: **Thomas W. Planek,** National Safety Council, USA
Editors: **Ken Kolosh,** National Safety Council, USA
Sergey Sinelnikov, National Safety Council, USA
Jonathan Thomas, National Safety Council, USA
Managing Editor: **Kathleen Porretta,** National Safety Council, USA

Journal of Safety Research 83 (2022) 339–348

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Effectiveness of implementing a Graduated Driver Licensing (GDL) law among young Serbian drivers

Predrag Stanojević [a], Timo Lajunen [b], Dragana Jakšić [a], Dragan Jovanović [c], Boško Matović [d,*]

[a] Academy of Applied Studies of Kosovo and Metohija, Serbia
[b] Department of Psychology on the Norwegian University of Science and Technology, Trondheim, Norway
[c] Department of Transport and on the Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia
[d] Faculty of Mechanical Engineering, University of Montenegro, Podgorica, Montenegro

ARTICLE INFO

Article history:
Received 24 July 2021
Received in revised form 1 June 2022
Accepted 13 September 2022
Available online 22 September 2022

Keywords:
Young drivers
Graduated driver licensing
Attitudes
Risk behavior

ABSTRACT

Introduction: Young novice drivers have higher rates of engagement in road crashes worldwide, often owing to unfavorable attitudes toward road safety or lack of knowledge, experience, and risk consciousness. However, the implementation of graduated driver licensing (GDL) systems has proven effective in reducing the high incidence of young drivers involved in crashes. *Method:* The purpose of this study was to compare the change in driving outcomes (e.g., attitudes toward traffic safety, behavior patterns while driving, risk assessment in traffic, assessment of personal driving skills, and involvement in traffic crashes) of young drivers prior to and following the implementation of a GDL law. Respondents (n = 642) completed a battery of questions, including a driver attitudes questionnaire (*Behaviour of Young Novice Drivers Scale),* a self-assessed driving ability questionnaire, and a risk perception questionnaire. Of the total sample size, 324 drivers passed the old system of training driver's license candidates, and 318 drivers passed the new GDL system. *Results:* The results showed that drivers licensed with GDL reported safer attitudes toward traffic rule violations and speed, and higher levels of safety orientation with regard to their driving abilities. They also reported much higher levels of risk perception and lower exposure to risky situations (risky driving exposure). There were no differences between GDL drivers and non-GDL drivers in terms of self-reported crashes or transient or fixed violations. In addition, GDL was not related to the number of traffic crashes, the number of fatalities, or serious and slight injuries in crashes involving young drivers in crashes obtained from official records. *Conclusions:* The results suggest that GDL contributed to the improvement of drivers' attitudes and understanding of risk but did not contribute to significant changes in the behavior of young drivers and traffic crashes. In addition, the GDL program in Serbia only ranks fair on the Insurance Institute for Highway Safety (IIHS) scale. Strengthening the GDL program in Serbia with additional components in line with GDL programs rated as good by the IIHS scale could improve the safety of young and novice drivers in traffic.

1. Introduction

Road traffic crashes have been the leading cause of death and serious injury among young adults in the last decade. Based on data provided by the World Health Organization (WHO, 2018), road traffic crashes are a leading cause of death globally, and young drivers aged 16–25 are significantly overrepresented among those killed and seriously injured.

Research has shown that young drivers commonly engage in risky driving behaviors, including exceeding the speed limit (ECMT, 2006; Scott-Parker et al., 2013), driving under the influence of alcohol (ECMT, 2006; NHTSA, 2018), dangerous overtaking (Fernandes et al., 2010), and distracted driving (Korpinen & Pääkönen, 2012; Lipovac et al., 2017). Using Serbia state year data on fatal crashes during 2016–2019, it has been found that among young drivers aged 18–24, 22 % and 8 % of all fatal traffic accidents involved inappropriate speed in relation to traffic and road conditions and alcohol-impaired driving (RTSA, 2021), respectively.

The main factors related to young driver traffic crashes are inexperience, sex, and age (ECMT, 2006). Studies that included the factors of inexperience and age have shown that these two factors are the critical predictors of crashes among young drivers (Chapman et al., 2014; McCartt et al., 2003; Twisk & Stacey, 2007). McCartt et al. (2009) reviewed 11 studies that attempted to separate the

* Corresponding author at: Faculty of Mechanical Engineering, University of Montenegro, Podgorica, Bulevar Džordža Vašingtona bb, Montenegro.

E-mail address: boskom@ucg.ac.me (B. Matović).

https://doi.org/10.1016/j.jsr.2022.09.008
0022-4375/

crash effects of age and experience, represented by the length of licensure. The results showed that age and experience have important, independent effects on crash risk, even after differences in driving mileage are accounted for. Meanwhile, studies on sex have demonstrated that male drivers are prevalent in traffic crashes involving young drivers. In Australia, for example, the death rate in 2017 was three times higher for males than for females in the 17–25 age category (BITRE, 2018). Further, studies have reported that drivers' attitudes, sensation-seeking, and personality traits could predict poor driving outcomes and high-risk driving behavior in young individuals (e.g., Hassan & Abdel-Aty; 2013; Khan et al., 2020; Preece et al., 2018; Ulleberg & Rundmo, 2003).

To reduce risks to young drivers, a range of countermeasures needs to be applied. Such countermeasures should include improvements in the areas of training, education, testing, communication, enforcement, and technology (ECMT, 2006; Molloy et al., 2019; Scott-Parker et al., 2015; Twisk & Stacey, 2007). A reasonable assumption is that education and training can correct attitudes toward traffic safety and establish safe behavioral patterns in young drivers. Driver training refers to teaching people driving skills to control and operate a vehicle to obtain a license, whereas driver education is a broader term that implies not only driver training but also general road safety concepts, road laws, behaviors, and awareness as well (Barua et al., 2014). Research findings on the effects of driver education in terms of fewer traffic violations and crashes have been mixed: Some studies show no significant effects of education on crashes or violations (Bates et al., 2020; Lonero & Mayhew, 2010); others show reductions in crashes or violations for those receiving driver education (Shell et al., 2015; Vanlaar et al., 2009).

Various licensing systems are used worldwide. Generally, the following categories of licensing systems can be distinguished: (1) traditional licensing systems and two-phase systems and (2) graduated driver licensing (GDL) systems. Two-phase systems represent a variation in probationary licensing systems, in which candidates must go through two phases of driving theory and training before becoming a fully licensed driver. After completing the first phase of the licensing process, candidates obtain a provisional or probationary license that allows them to drive solo under given conditions; a full license is issued after completion of the second phase of theory and training, without further testing (ECMT, 2006). Similar to other related systems, such as two-phase systems or systems that include a probationary license, GDL aims to address certain specifics of the "young driver problems." GDL systems include an extension of the learning period, supervised driving, and after obtaining the initial license, restrictions on driving with peer passengers and nighttime driving. In essence, GDL programs focus on inexperienced drivers and are tasked with ensuring that they are exposed only to driving conditions appropriate to their skills while they are in the process of developing additional skills and experience (Foss, 2007). GDL systems have proven highly effective in reducing the risk of crashes among young drivers (Bates et al., 2018; Shope, 2007). In October 2012, a GDL system was introduced in Serbia, comprising educational training and the GDL component proper. The educational training component consists of several phases. The first phase includes 40 hours of education (acquisition of theoretical knowledge for safe participation in traffic). The second phase includes mastering driving skills, and the training includes practical driving (on the street) after obtaining the certificate for the completion of the theoretical test. The third phase includes a first-aid examination. First-aid training and examination are organized by the Red Cross of Serbia. When all of these phases are completed, the driving test proper is conducted. The new GDL system component in Serbia implies first obtaining a probationary driver's license with restrictions on speed, night driving, cell phone use, and alcohol use, and with required supervised driving. A full license is obtained after one year of driving with a probationary license. Supervised driving is required for holders of a probationary driver's license obtained at the age of 17 and lasts one year, that is, until the expiration of the probationary license. Supervised driving is not required for drivers who obtain a probationary license after the age of 18. All other restrictions affect all holders of a probationary driver's license. The police are responsible for enforcing restrictions on young drivers. They can readily identify drivers subject to restrictions based on a "P" sticker placed on drivers' vehicle and check whether young drivers adhere to the legally defined restrictions. Further, for the enforcement of restrictions on holders of a probationary driver's license obtained at the age of 17, the person who supervises the driver (who should hold a license valid for at least five years) is responsible.

In the traditional training system that was in place in Serbia before the GDL system, the novice driver was fully licensed after passing the driving test, and no special conditions were placed. Novice drivers had 40 hours of driving training (polygon and street driving), and after completing the training, a driving exam was conducted, which was composed of two parts: theoretical and practical. After passing the driving test, candidates instantly obtained a full license. A more detailed presentation of the new GDL system and its comparison with the traditional training system are presented in Table 1.

Evaluation of the effects of driver education on intermediate outcome criteria should include examinations of changes in behavior, attitudes, knowledge, and exposure to risk (Lonero & Mayhew, 2010). In their literature review, Thomas et al. (2012) found little solid evidence that driver education affects teen crashes or other outcomes. However, recent studies have revealed that education and training within GDL programs are linked to a smaller number of crashes and traffic violations (Mayhew et al. 2014; Shell et al., 2015). Based on the reviewed literature, we hypothesized that: drivers licensed with GDL have more favorable attitudes toward traffic safety (H1); drivers licensed with GDL have safer behavior patterns while driving (H2); drivers licensed with GDL have better risk assessment in traffic (H3); drivers licensed with GDL have better assessment of personal driving skills (H4); and drivers licensed with GDL have fewer crash involvements (H5).

2. Methods

2.1. Participants and procedure

The data in this study were collected by mail. The names and addresses of 1,000 individuals (young drivers aged 17–26 years with driving experience of up to three years) with valid driver's licenses were obtained from a registry of motor-vehicle owners. Participants were instructed to thoroughly read all the questions in the study questionnaire and honestly provide their answers, which would be treated as anonymous. Participants were also assured that their data would be treated as confidential and would only be used for the purpose of the study. To reduce the possibility of participants giving socially desirable answers, participants were not asked to provide their names in the questionnaire. Questionnaires were collected on two occasions. At the end of 2012, 500 questionnaires were distributed to young drivers who had passed the old system of training driver's license candidates (pre-GDL). At the end of 2016, another 500 questionnaires were distributed to young drivers who passed the new system of training driver's license candidates (GDL). In both cases, a reminder was sent to non-respondents. A total of 642 participants completed and returned the distributed questionnaires: 265 were female and 377 were male. Based on data obtained from the questionnaires, a comparison of young drivers licensed pre-GDL and those licensed

Table 1
Comparison of the content of the traditional training system with the GDL system in Serbia.

	Pre-GDL	GDL
Learner's entry age	18	16
Minimum age for obtaining a license	18	17
Full license	Immediately after passing the driving exam	After one-year of driving with a probationary license
Learner's holding period	about 1 month	About 3 months
Theoretical education	No theoretical education	40 hours of theoretical education (acquisition of theoretical knowledge for safe participation in traffic)
Driving training	40 hours of driving (driving training is not conditioned by taking the theoretical test)	40 hours of driving (prior to driving training, passing the theoretical test is obligatory)
First-aid training	Was not applicable	8 hours of training followed by the examination
Probationary driver's license	Did not exist	Probationary driver's license with one-year validity. For the holders of a probationary driver's license who obtain a license at the age of 17, the restrictions are as follow: (a) Supervised driving with a person who holds a license with a validity of at least 5 years (b) Not allowed to drive at a speed of over 110 km/h on a highway, 90 km/h on a motorway, and 90 % of the speed allowed on the part of the road on which driving takes place (other roads) (c) Night driving restriction from 11:00p.m. to 5:00 a.m. (d) The driver may not use a cell phone while driving (e) Allowed BAC 0.0 g/l (f) Vehicle must have a "P" sticker on a visible place on the front and the rear side of the vehicleFor the holders of a probationary driver's license who obtained the license at the age of 18, the same restrictions apply except under paragraph a).
Price	From 250 to 300 Euros (280 to 330 USD)	From 500 to 550 Euros (550 to 600 USD)
Video surveillance	Did not exist	Obligatory when taking tests and during the lectures

with GDL was conducted using various measures, including attitudes, risky behavior, self-assessed driving ability, risk perception, and the number of self-reported crashes.

The demographic characteristics of the overall sample are presented in Table 2.

To obtain a more complete picture of the effects of GDL, a comparison of the rates of crashes with causalities, fatalities, serious injuries, and slight injuries per 100,000 licensed drivers (aged 18–26) before and after the implementation of GDL was performed. Data on traffic crashes and the number of young drivers were obtained from official records (Road Traffic Safety Agency and Ministry of Interior). The time coverage was the period from 2002 to 2020, and the rates of traffic crashes with causalities, fatalities, serious injuries, and slight injuries in the period from 2002 to 2012 (before the implementation of GDL) were compared with the period from 2013 to 2020 (after the implementation of GDL).

2.2. Measures

2.2.1. Drivers' attitudes

Drivers' attitudes were measured using a questionnaire that comprised 23 items. The responses were marked on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). We used this questionnaire to test the following attitudes: the attitude toward the violation of traffic regulations (8 items; e.g., 'There are many traffic rules which cannot be obeyed to keep up the traffic flow,' or 'Sometimes it is necessary to bend the rules to keep traffic going'); the attitude toward speed (7 items; e.g. 'If you have good skills, speeding is OK,' or 'Driving 5 or 10 miles above the speed limit is OK because every-one does it'); the attitude toward drinking and driving (5 items; e.g., 'People can drive safely even if they are under the influence of alcohol,' or 'People should be allowed to decide themselves how much they can drink and drive after that'); and the attitude toward joyriding (6 items; e.g., 'Adolescents have a need for fun and excitement in traffic,' or 'Speeding and excitement belong together when you are driving'). The attitude scales were based on questionnaires developed by Ulleberg and Rundmo (2003) and Iversen and Rundmo (2004).

2.2.2. Behaviour of young novice drivers scale

The *Behaviour of Young Novice Drivers Scale* (BYNDS, Scott-Parker et al., 2010) was used to measure the risky behavior of young novice drivers. The BYNDS contains five subscales that are added to obtain a composite risky driving score. Transient violations (13 items) measure risky driving behaviors that can change throughout the journey, such as driving speed and speaking on a mobile phone. Fixed violations (10 items) explore risky driving behaviors that are unlikely to change throughout the journey, such as alcohol and/or drug intoxication, and the wearing of seatbelts. Misjudgment (9 items) captures the novice driving errors that

Table 2
Participants' demographic characteristics.

	pre-GDL	GDL	P value
Response rate	64.8 %	63.6 %	
Sex			
Male (%)	195 (60.2)	182 (57.2)	0.447
Female (%)	129 (39.8)	136 (42.8)	
Age			
Mean (S.D.)	20.21 (2.17)	21.12 (2.43)	<0.001
Range	18–26	18–26	
Distribution of ages			
17–18 (%)	91 (28.1)	64 (20.1)	<0.001
19–20 (%)	106 (32.7)	80 (25.2)	
21–22 (%)	76 (23.5)	76 (23.9)	
23–24 (%)	33 (10.2)	58 (18.2)	
25–26 (%)	18 (5.6)	40 (12.6)	
Duration of driving experience in moths			
Mean (S.D.)	21.07	15.14	<0.001
Range	0–84	0–42	
Month mileage in km			
Mean (S.D.) - Total sample	516 (1059)	272 (604)	<0.001
Mean (S.D.) - Males	705 (1261)	336 (643)	
Mean (S.D.) - Females	230 (527)	185 (537)	
Education			
College degree or more (%)	63 (19.4)	102 (32.1)	<0.001
Less than college degree (%)	261 (80.6)	216 (67.9)	
Number of self-reported crashes from the obtaining the license			
Range	0–6	0–6	0.427
Mean (S.D.)	0.40 (0.89)	0.35 (0.88)	

place the young driver at an increased risk of crash, such as underestimating the distance required to stop and misjudging the gap when overtaking or turning across traffic. Risky exposures (9 items) gauge the risky circumstances of the young novices' driving, including driving at night and with their friends as their passengers. Driver moods (3 items) demonstrate the extent to which the young novices are driving in response to their emotions, including anger and frustration (Scott-Parker & Proffitt, 2015).

Respondents were asked to indicate how often they have engaged in the following behaviors. Responses were recorded on a 5-point Likert scales that ranged from 1 to 5 (1 = *never* to 5 = *almost all the time*).

2.2.3. Self-assessed driving ability questionnaire

The measurement instrument applied to examine self-assessed driving ability comprised 22 indicators (Tronsmoen, 2008). The indicators fell into four dimensions. The first dimension related to self-assessment of general driving ability (8 items) included skills such as speeding, anticipating, driving in slippery conditions, and driving in the dark. The second dimension, safety orientation (6 items), referred to the driver's perception of his/her own ability to identify risk, danger and his/her perception of their ability to drive with satisfactory safety margins. The third dimension was the body dimension (5 items), which measured the feeling of unity with and control of the car. The fourth dimension was specific task skills (8 items) and items under this dimension consisted of judgment of the ability for precise and effective parallel parking, reversing into a garage, as well as reversing using the rear-view mirrors. The subjects were asked how well the statements fit the way they manage and perceive driving a car. The responses were marked on a 5-point scale (1 = *does not fit me at all* to 5 = *fits me perfectly*).

2.2.4. Risk perception questionnaire

The risk perception questionnaire (Harbeck & Glendon, 2013) comprised 10 items. The respondents were asked to estimate how risky the given actions were in traffic (e.g., driving under influence of alcohol, driving over the 20 km/h speed limit). The responses were recorded on a 5-point Likert-type scale (from *1 = not risky* to *5 = extremely risky)*.

2.2.5. Demographic variables

Respondents answered questions about their age, sex, education, and crash involvement during the previous three years (number of self-reported crashes [crashes with fatalities + crashes with material damage]), and were also asked to estimate their monthly mileage.

2.2.6. Data from official records

Data on traffic crashes of young drivers age 18–26 (data on the total number of crashes with casualties, killed and injured young drivers can be found in Appendix A Table A1) are extracted from the integrated database of characteristics of traffic safety RTSA (RTSA, 2021).

The number of licensed young drivers aged 18–26 was obtained from the Ministry of Interior of the Republic of Serbia. This was used for the calculation of rates pertaining to crashes, fatalities, serious injuries, and slight injuries per 100,000 licensed drivers (aged 18–26) for each year of investigation.

2.3. Statistical analyses

Instruments used in this study had not yet been validated on a Serbian sample. Because some cultural, social, and economic factors could result in a different factor structure, we decided to analyze all scales using factor analysis. Considering the recommendations that confirmatory factor analysis (CFA) should be used when the underlying structure has been established based on prior empirical and theoretical grounds, the factor structure of all the instruments used in the study was initially examined using CFA. However, if the model on some scale produced a poor fit to the data, principal component analysis was run to examine the factor structure of the given instrument. Reliability for finalized measures and established scales was assessed using internal consistency (Cronbach's alpha). The values for asymmetry and kurtosis between −2 and + 2 are considered acceptable in order to prove normal univariate distribution (George & Mallery, 2010). A one-way between groups analysis of covariance (ANCOVA) was used to identify differences between young drivers licensed pre-GDL with those licensed with a GDL, after controlling for sex and mileage. As the self-reported yearly accident rate did not follow normal distribution, Poisson or negative binomial regression analyses (see Lord, Washington, & Ivan, 2005) were performed to test the relationship between GDL program and traffic crashes. In this analysis, the dependent variable was traffic crashes and independent variables were mileage, GDL program and their interactions (GDL × mileage). Interaction between GDL program and mileage was included in the model because there is a statistically significant difference between young drivers licensed pre-GDL and those licensed with a GDL (see Table 2). Therefore, there is a possibility that GDL is associated with the reduction of mileage at young drivers that further leads to the reduction of the number of traffic crashes. Finally, a comparison of traffic crashes (obtained from the official records) of young drivers before and after the implementation of GDL was performed. The time coverage of the research is the period 2002–2020.

3. Results

3.1. Factor analysis of the scales for examining attitudes, behaviors, self-assessed driving ability, and risk perception of young drivers

Confirmatory factor analysis (CFA) was used to test the internal structure of the driver attitudes questionnaire, *Behaviour of Young Novice Drivers Scale* (BYNDS), self-assessed driving ability questionnaire, and risk perception questionnaire used in this study. Model fit was evaluated using the χ^2/degree of freedom ratio, root mean square error of approximation (RMSEA), goodness-of-fit index (GFI), comparative fit index (CFI), and standardized root mean square residual (SRMR). In general, a good model fit should have a 2:1 or 5:1 χ^2/degree of freedom ratio and GFI > 0.90, CFI > 0.90 (preferably > 0.95), RMSEA < 0.08 or 0.10 (preferably < 0.06), and SRMR < 0.08 (preferably < 0.05) indices (Hu & Bentler, 1999; Russell, 2002). All scales except the BYNDS model showed a good fit for the data (the results of the CFA can be found in Appendix A, Table A2).

Methodologically, CFA and exploratory factor analysis (EFA) models cannot be applied to the same data. Therefore, the initial group of 642 participants was randomly divided into two subgroups. The first subgroup (Sample 1) was used to perform the CFA and consisted of 317 drivers (179 males and 138 females). Given that the application of CFA provided a poor fit for the BYNDS (the results are given in Appendix A Table A2), the data were reexamined within an EFA framework. An EFA was conducted with the second subgroup (Sample 2, 325 drivers, out of which 198 were male and 127 were female) with principal components by using an oblique promax rotation with Kaiser normalization. Initially, 44 items of the BYNDS translated into the Serbian language were used. However, the final solution included only 35 items because four items (TR3, TR11, TR12, FI10, and MS9) had a factor loading below 0.4, and four items (TR7, TR8, FI3, and EX7) had a high loading (>0.4) on two factors. The Kaiser–Meyer–Olkin measure of sampling adequacy was acceptable at 0.893, and Bartlett's test of sphericity was significant at $p < 0.001$. The EFA of the remaining 36 items revealed a structure of five factors, explaining 55.93 % of the variance in risky young driving behavior (the results of the EFA can be found in Appendix A Table A3).

3.2. Scale statistics

The items within each factor were summed and comprised four subscales of "attitudes toward traffic," five subscales of "Serbian BYNDS," four subscales of "self-assessed driving ability," and one scale of "driving risk perception" (the range [min–max], means, standard deviations, skewness, kurtosis, and Cronbach's alpha for all subscales can be found in Appendix A Table A4). The skewness and kurtosis values indicated that the distributions of the subscales did not deviate substantially from normality (skewness < 2; kurtosis < 2 in all cases). All subscales had good internal consistency coefficients, with Cronbach's alpha ranging from 0.73 to 0.90.

3.3. Results of analysis of covariance (ANCOVA)

ANCOVA was conducted to test whether the mean differences in the dependent variables (attitudes toward traffic, Serbian BYNDS, self-assessed driving ability, and risk perception) between groups (pre-GDL versus GDL) would be significant after removing the effects of covariates (sex and mileage). For this purpose, 14 ANCOVAs were performed. In each ANCOVA, driver licensing systems (pre-GDL vs GDL) were the independent variables, and the sex and mileage variables were treated as covariates. Table 3 shows that after adjusting for sex and mileage, there were significant differences between young drivers licensed pre-GDL and those licensed with GDL with regard to attitude toward violation of traffic regulations, attitude toward speed, risky driving exposure, safety orientation, and driving risk perception. Drivers licensed with GDL reported safer attitudes toward traffic rule violations and speed and higher levels of safety orientation, which refers to a driver's perception of his or her ability to identify risks and dangers and the driver's perception of his or her ability to drive with satisfactory safety margins. They also reported higher levels of risk perception and less exposure to risky situations (i.e., risky driving exposure).

3.4. Results of the regression analyses

The distribution of crashes did not follow a normal distribution. Thus, Poisson or negative binomial regression analyses were performed. In this analysis, the dependent variable was self-reported traffic crashes, and the independent variables were mileage, the GDL program, and their interactions (GDL $\times$ mileage). We used goodness-of-fit statistics to test the appropriateness of the regression models based on Poisson distribution. These statistics indicated a misfit of the Poisson distribution for self-reported yearly accident involvement, χ^2 (638) = 1187.21, $p < 0.001$, therefore, models based on negative binomial distribution were constructed. As shown in Table 4, mileage ($Z = -1.07$, $p = 0.283$), the GDL program ($Z = -0.33$, $p = 0.740$), and GDL $\times$ mileage ($Z = 1.63$, $p = 0.103$) were not significantly related to the number of traffic crashes.

The stepwise regression analysis was used to explain the variance in the direct road safety outcome measures reported by the police (see Table 1 in Appendix A). The results showed that the change in calendar years explained significant amount of variability in the number of crashes with casualties ($R^2 = 0.293$; $p = 0.017$), the number of fatalities ($R^2 = 0.703$; $p = 0.000$), and the number of serious injuries ($R^2 = 0.688$; $p = 0.000$) in crashes involving young drivers (18–26). However, calendar years did not predict a significant amount of variance in the number of slight injuries ($R^2 = 0.077$; $p = 0.249$). The inclusion of the pre/post GDL variable did not predict a significant level of variance in the road safety outcomes.

4. Discussion

In Serbia, young drivers are more frequently involved in car crashes than any other age group. The aim of this study was to determine the effects of adopting GDL on various important factors directly related to the safety of young drivers. A comparison between young drivers licensed pre-GDL and those licensed with GDL was conducted using various measures, including attitudes, risky behavior, self-assessed driving ability, risk perception, and the number of self-reported crashes and crashes obtained from official records.

Responses for two of the four dimensions measuring attitudes showed a tendency to report attitudes indicative of greater safety among drivers licensed with GDL compared with drivers licensed

Table 3
Differences between young drivers licensed pre-GDL and licensed with a GDL.

	pre-GDL		GDL		*F*(1,737)	Eta^2
	Mean	SE	Mean	SE		
Attitudes towards traffic						
Attitude towards violation of traffic regulations	18.20	0.34	16.80	0.34	**8.49****	0.013
Attitude towards speed	18.22	0.36	16.05	0.36	**17.99****	0.027
Attitude towards drinking and driving	8.51	0.22	7.98	0.22	2.91	0.005
Attitude towards joyriding	13.89	0.27	13.29	0.27	2.55	0.004
Serbian BYNDS						.
Transient rule violations	17.66	0.33	16.99	0.33	2.07	0.003
Fixed rule violations	12.19	0.24	11.66	0.25	2.27	0.004
Misjudgement	14.29	0.27	14.30	0.28	0.01	0.000
Risky driving exposure	28.39	0.39	27.20	0.39	**4.55***	0.007
Driver mood and aggressive driving	12.38	0.27	11.89	0.27	1.57	0.002
Self-assessed driving ability						.
General driving ability	24.93	0.36	25.57	0.36	1.55	0.002
Safety orientation	21.70	0.26	20.84	0.26	**5.49***	0.010
The body dimension	17.36	0.23	17.31	0.23	0.02	0.000
Specific task skills	10.36	0.17	10.39	0.17	0.02	0.000
Driving risk perception	34.34	0.35	36.32	0.35	**15.86****	0.024

Note: Attitudes towards traffic: lower totals scores express safer attitudes; Serbian BYNDS: lower total scores indicate the safer behaviour in traffic; Self-assessed driving ability: lower total scores indicate critical view of their own driving skills; Driving risk perception: lower total scores indicate high perceived risk. The means are adjusted for sex and mileage. SE = standard error. $^{*}p < 0.05$; $^{**}p < 0.001$.

Table 4
Negative binomial regression analyses on crash involvement during the previous 3 years.

Variables	Incidence rate ratios (IRR)	Std. Err.	Z-value	95 % conf. interval	Pseudo R^2
GDL program (1 = pre-GDL, 2 = GDL)	0.793	0.171	−1.07	0.519–1.211	0.023
Mileage	0.999	0.000	−0.33	0.999–1.001	
GDL × mileage	1.000	0.000	1.63	0.999–1.001	

pre-GDL. Drivers licensed with GDL reported more favorable attitudes toward traffic regulations and speed. This result was expected because traffic regulations and speed are commonly represented in education programs for drivers licensed with GDL. These results are consistent with those of other studies that have shown that educational programs can change drivers' attitudes (Floreskul et al., 2016; Mann & Lansdown, 2009). However, there were no significant differences between young drivers licensed pre-GDL and those licensed with GDL in terms of attitudes toward drinking and driving and toward joyriding. These results are not consistent with our hypothesis. There are several potential reasons for these results. The total score on the attitude toward drinking and driving scale in both groups was low, which indicates that young drivers are largely opposed to drinking and driving; thus, there is not much room to upgrade this attitude.

Other studies using the questionnaire on the attitude toward drinking and driving that we used in our study also reported low total scores, that is, in range of our results (Endriulaitienė et al., 2018; Ulleberg & Rundmo, 2002). Moreover, studies in Serbia conducted in the period relevant to our research (2010 and 2017) showed that there have been no significant changes in drinking and driving. For example, in 2010 and 2017, 90.4 % and 89 % of car drivers respectively reported that they had not driven above the legal limit the past month (Antov et al., 2012; RTSA, 2019). Attitude toward joyriding represents a specific attitude, and topics of the training program do not cover this area.

Further, the differences in participants' risk perception were expected. Risk perception is directly related to informing the driver, and it seems that the content of the theoretical part of the training has had an effect, that is, the young driver's better understanding of the risks in traffic. This result is consistent with studies that have shown that training and campaigns can increase the risk perception of traffic among young drivers (Rosenbloom et al., 2008; Rundmo & Iversen, 2004). Interestingly, a statistically significant difference was observed in self-assessed driving ability, in the sense that drivers licensed pre-GDL had higher scores on the safety orientation scale than those licensed with GDL. One possible explanation for this result is that GDL education and training contribute to addressing the problem of overconfidence and may lead young drivers to adopt a more self-critical view of their driving abilities. No difference was observed in transient and fixed violations, which was unexpected, considering the existing differences in risk perception and attitudes toward regulations and speed. Notably, a difference in the total scores of transient and fixed violations was observed, but the difference was not statistically significant. Other studies have provided evidence of small changes in target behavior when attempting to change attitudes (Elvik et al., 1997; OECD, 1994). Studies have shown that applying a combination of different measures is critical for determining whether there are significant effects on behavior (Aarø & Rise, 1996; Delhomme, 1999).

The results show that the application of the GDL is not related to the number of self-reported crashes. In addition, the GDL is not related to the total number of crashes with casualties, fatalities, serious injuries, and slight injuries among young drivers. The data were obtained from official records. It seems that the application of GDL did not reduce the number of traffic crashes and their consequences among young drivers. The new Road Traffic Safety Law in Serbia was implemented in 2010, while the application of GDL was delayed to 2013. Based on data on the number of total crashes and killed and injured young drivers, it can be seen that the decline in the number of killed and injured young drivers started in 2010 (see Appendix A Table A1) under the influence of the new law, and it seems that the application of GDL *per se* did not have expected effects. These results are unexpected, considering the results of studies that have shown that the application of GDL significantly contributes to the reduction of traffic crashes (e.g., Dee et al., 2005; Males, 2007).

It seems that GDL increased the awareness of young drivers with regard to traffic risks, but not sufficiently to be evident in changing behavior and lessening involvement in traffic crashes. Several studies have reported that the implementation of the new Road Traffic Safety Law in Serbia has had mild effects on traffic safety (Antić et al., 2011; Nazif-Munoz & Nikolic, 2018). It seems that strengthening the GDL program in Serbia with additional components is necessary to change drivers' behavior and reduce young drivers' traffic crashes. Strengthening should take into account two directions. First, education and training of novice drivers should be adjusted to the GDL framework. It seems that education and training should include content that would provide risk-prevention and self-evaluation skills related to the strategic and personal levels of driving behavior. Although several studies have shown little or no convincing evidence that driver training reduces crashes or other outcomes (e.g., Lonero & Mayhew, 2010; Thomas et al., 2012), it is important to not abandon but improve driver education. Second, the GDL system in Serbia does not include components such as driving under the supervision of an adult with driving experience after the age of 18 and driving without the presence of peers. In addition, the learner's license-holding period is quite short (three months). Studies have reported that supervised driving has multiple advantages and significantly reduces crash risk in young drivers (ECMT, 2006; Masten et al., 2015; Siskind et al., 2019). Further, passenger restrictions are associated with a substantial reduction in young drivers' fatal crash rates (Hirschberg & Lye, 2020; McCartt et al., 2010; Senserrick & Williams, 2015). In addition, multiple studies have shown that license-holders who spent a longer time in the learner stage had fewer crashes (Ehsani et al. 2013; Senserrick & Williams, 2015).

5. Limitations and future research

This study has some methodological limitations that should be considered. First, because this study is based on self-reported data, it suffers from the usual perceived weaknesses such as social desirability bias. However, because no names were collected, and data collection was undertaken remotely, the impact of social desirability bias is unlikely to have significantly affected the results. Second, drivers' attitudes, risk perception, self-assessed driving ability, and driver behavior could have been influenced by other factors (e.g., traffic safety campaigns and traffic enforcement), not only by education and training. The Road Traffic Safety Law in Serbia (which came into force in 2010) has brought many traffic-related novelties, with the most important of these being strategic actions in traffic safety, a traffic safety financing system, a penalty points system, improvement of road infrastructure safety, and traffic safety campaigns. However, most of these measures have not yet been implemented in practice. For example, there have been no differences between the periods before and after the enactment of the new law in terms of campaign implementation, intensity of traffic enforcement, and strategic actions in traffic safety. Considering that there were no significant changes in our other factors, it can be assumed that driver training was the factor that contributed the most to the differences related to drivers' attitudes, risk perception, self-assessed driving ability, and behavior of young drivers.

The results of the present study show that the education of novice drivers leads to a change in the attitudes and risk perception of young drivers, but there is no change in behavior. Future studies should take different approaches in driver education that would contribute to understanding the importance of appropriate safety attitudes and behavior to avoid accident involvement in traffic.

6. Conclusion

The effects of GDL in Serbia have been limited. GDL has contributed to the improvement of drivers' attitudes and understanding of risk, but it has not contributed to significant changes in the behavior of young drivers and traffic crashes. The results of this study suggest that it is necessary to further improve the education and training of new drivers. The Insurance Institute for Highway Safety developed a rating system that can be applied to various driver education programs. The rating system includes components such as minimum age limit, permit holding periods, required practice hours, nighttime and passenger restrictions, and durations of the restrictions. Laws are rated as good, fair, marginal, or poor (for additional information, see McCartt et al., 2010). Based on its components, the Serbian GDL program may be rated as fair. McCartt et al. (2010) provide evidence that a strong graduated licensing law can serve as an effective countermeasure for reducing fatal crash involvements of teenage drivers aged 15–17 years. Laws rated good were associated with a 30 % and 19 % reduction in the fatal crash rate compared with laws rated poor and fair, respectively. The GDL program in Serbia, in relation to programs rated as good, has the following deficiencies: short learner's holding period (about three months), restriction on night driving in the probationary phase starting from 11:00p.m., and passenger restrictions not being included in the probationary phase. It seems that strengthening the GDL program in Serbia with additional components would be necessary to change drivers' behavior and reduce traffic crashes of young drivers.

Acknowledgement

This work was supported by the Provincial Secretariat for HigherEducation and Scientific Research (Autonomous Province of Vojvodina) through project "Development of a spatially based intelligent traffic safety management system" [grant number 142-451-2690/2021-01/02].

Appendix A

See Tables A1–A4.

Table A1
The total number of crashes with casualties, the number of fatalities, serious and slight injuries in crashes involving young drivers (18–26) and the number of licensed young drivers (18–26), 2002–2020.

Year	Total crashes with casualties[a] (rates[b])	Fatalities (rates[c])	Serious injuries (rates[d])	Slight injuries (rates[e])	Number of licensed young drivers	Licensing systems
2002	4369 (102.54)	72 (1.69)	410 (9.62)	1317 (30.91)	426,065	Pre-GDL
2003	4706 (104.24)	64 (1.42)	453 (10.03)	1455 (32.23)	451,476	Pre-GDL
2004	5090 (111.87)	73 (1.60)	472 (10.37)	1648 (36.22)	455,007	Pre-GDL
2005	5042 (107.87)	73 (1.56)	466 (9.97)	1668 (35.69)	467,412	Pre-GDL
2006	5461 (110.68)	84 (1.70)	547 (11.09)	1887 (38.25)	493,394	Pre-GDL
2007	6531 (131.49)	100 (2.01)	595 (11.98)	2479 (49.91)	496,692	Pre-GDL
2008	6452 (129.68)	89 (1.79)	649 (13.04)	2564 (51.54)	497,518	Pre-GDL
2009	5972 (117.77)	81 (1.60)	576 (11.36)	2401 (47.35)	507,095	Pre-GDL
2010	5234 (101.58)	58 (1.13)	454 (8.81)	2042 (39.63)	515,283	Pre-GDL
2011	4825 (94.51)	69 (1.35)	409 (8.01)	1877 (36.77)	510,515	Pre-GDL
2012	4390 (88.16)	60 (1.20)	376 (7.55)	1691 (33.96)	497,977	Pre-GDL
2013	4375 (92.74)	54 (1.14)	306 (6.49)	1711 (36.27)	471,753	GDL
2014	4135 (96.02)	39 (0.91)	298 (6.92)	1529 (35.51)	430,639	GDL
2015	4313 (109.35)	45 (1.14)	289 (7.33)	1569 (39.78)	394,412	GDL
2016	4622 (124.20)	41 (1.10)	277 (7.44)	1774 (47.67)	372,145	GDL
2017	4755 (132.25)	39 (1.08)	299 (8.32)	1804 (50.18)	359,537	GDL
2018	4498 (128.40)	38 (1.08)	264 (7.54)	1657 (47.30)	350,298	GDL
2019	4615 (129.09)	39 (1.09)	252 (7.05)	1757 (49.15)	357,492	GDL
2020	4238 (117.30)	44 (1.22)	280 (7.75)	1685 (46.64)	361,282	GDL

Note:
[a] Sum of fatal and injury crashes.
[b] Rates of crashes with causalities per 10,000 licensed young drivers.
[c] Rates of fatalities per 10,000 licensed young drivers.
[d] Rates of serious injuries per 10,000 licensed young drivers.
[e] Rates of slight injuries per 10,000 licensed young drivers.

Table A2
Fit indexes from confirmatory factor analysis.

Model	χ^2/df	RMSEA (95 % CI)	GFI	CFI	SRMR
Attitudes towards traffic	859.8/283 = 3.0	0.06 (0.05 to 0.06)	0.90	0.91	0.05
Self-assessed driving ability	631.3/194 = 3.3	0.06 (0.05 to 0.06)	0.92	0.94	0.04
Behaviour of Young Novice Drivers Scale	2937.4/892 = 3.3	0.09 (0.08 to 0.09)	0.68	0.74	0.09
Driving risk perception (one factor)	102.3/24 = 4.2	0.07 (0.06 to 0.07)	0.97	0.96	0.05

Note: A good fit of model should, in general, have 2:1 or 5:1 χ^2/df, GFI > 0.90, CFI > 0.90 (preferably > 0.95), RMSEA < 0.08 or 0.10 (preferably < 0.06), and SRMR < 0.08 (preferably < 0.05) indices.

Table A3
Serbian version of the Behaviour of Young Novice Drivers Scale.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Risky driving exposure					
EX2. You drove in the rain	,868				
EX4. You drove at night	,867				
EX3. You drove at peak times in the morning	,851				
EX5. You drove at dusk or dawn	,812				
EX1. You drove on the weekend	,770				
EX6. You carried your friends as passengers at night	,627				
EX8. Your car was full of your friends as passengers	,521				
EX9. You went for a drive with your mates giving you directions to where they wanted to go	,458				
Misjudgement					
MS3. You misjudged the gap when you were turning left		,882			
MS6. You misjudged the gap when you were overtaking another vehicle		,880			
MS2. You misjudged the speed of an oncoming vehicle		,756			
MS4. You misjudged the stopping distance you needed		,707			
MS1. You misjudged the speed when you were exiting a main road		,686			
MS7. You missed your exit or turn		,471			
MS5. You turned left into the path of another vehicle		,461			
MS8. You entered the road in front of another vehicle		,400			
Transient violations					
TR2. You went 10–20 km/h over the speed limit (e.g. 72 km/h in a 60 km/h zone, 112 km/h in a 100 km/h zone)			,925		
TR1. You drove over the speed limit in areas where it was unlikely there was a radar or speed camera			,827		
TR5. You went up to 10 km/h over the speed limit (e.g. 65 km/h in a 60 km/h zone, 105 km/h in a 100 km/h zone)			,824		
TR6. You drove more than 20 km/h over the speed limit (e.g. 60 km/h in a 40 km/h zone, 120 km/h in a 100 km/h zone)			,773		
TR9. You sped up when the lights went yellow			,678		
TR10. You didn't always indicate when you were changing lanes			,544		
TR4. You sped at night on roads that were not well lit			,444		
Driver mood and risky driving					
FI5. You drove without a valid licence as because you had not applied for one yet or it had been suspended				,810	
FI7. If there was no red light camera, you drove through intersections on a red light				,703	
DM3. You drove faster if you were in a bad mood				,692	
FI2. You drove after taking an illicit drug such as marijuana or ecstasy				,678	
DM2. You allowed your driving style to be influenced by your mood				,648	
DM1. Your driving was affected by negative emotions such as anger or frustration				,585	
FI9. You drove when you thought you may have been over the legal alcohol limit				,508	
Fixed violations					
FI6. You did not wear a seatbelt if it was only for a short trip					,803
FI4. You did not always wear your seatbelt					,746
FI1. Your passengers did not wear seatbelts					,628
TR13. You spoke on a mobile that you held in your hands					,487
FI8. You carried more passengers than there were seatbelts for in your car					,413
Eigenvalues	9.95	4.57	2.06	1.54	1.46
Variance (%)	28.43	13.07	5.87	4.40	4.16

Extraction method: principal components, rotation method: promax with Kaiser Normalisation.

Table A4
Range, means, standard deviations, skewness, kurtosis, and Cronbach's alpha for all subscales for total sample.

Scale	Min-Max	M	SD	Skewness	Kurtosis	Cronbach's alpha
Attitudes towards traffic						
Attitude towards violation of traffic regulations	8–40	17.50	6.17	0.31	−0.32	0.79
Attitude towards speed	7–35	17.14	6.58	0.28	−0.61	0.85
Attitude towards drinking and driving	5–23	8.24	3.92	1.51	1.97	0.80
Attitude towards joyriding	6–30	13.59	4.93	0.48	−0.16	0.73
Serbian BYNDS						
Transient rule violations	7–35	17.33	6.13	0.26	−0.64	0.88
Fixed rule violations	5–25	11.93	4.52	0.41	−0.41	0.79
Misjudgement	8–36	14.29	4.91	1.11	1.60	0.86
Risky driving exposure	8–40	27.80	7.22	−0.49	−0.16	0.88
Driver mood and risky driving	7–33	12.14	4.99	1.30	1.59	0.83
Self-assessed driving ability						
General driving ability	8–40	25.25	6.71	−0.17	−0.15	0.89
Safety orientation	6–30	21.27	4.65	−0.76	0.81	0.80
Body dimension	5–25	17.33	4.30	−0.37	−0.09	0.85
Specific task skills	3–15	10.37	3.07	−0.41	−0.39	0.82
Driving risk perception	11–50	35.32	6.40	−0.17	0.20	0.84

References

Aarø, L. E., & Rise, J. (1996). *Den menneskelige faktor: Kan ulykker forebygges gjennom holdningspåvirkning*. In Oslo: Skadeforebyggende forum/Esso Norge.

Antić, B., Vujanić, M., Jovanović, D., & Pesić, D. (2011). Impact of the new road traffic safety law on the number of traffic casualties in Serbia. *Scientific Research and Essays, 6*(29), 6176–6184. https://doi.org/10.5897/sre11.1419.

Antov, D., Banet, A., Barbier, C., Bellet, T., Bimpeh, Y., Boulanger, A., ... & Zavrides, N. (2012). *European road users' risk perception and mobility: the SARTRE 4 survey*. IFSTTAR.

Barua, S., Sidawi, B., & Hoque, S. (2014). Assessment of the role of training and licensing systems in changing the young driver's behavior. *International Journal of Transportation Science and Technology, 3*(1), 63–78. https://doi.org/10.1260/2046-0430.3.1.63.

Bates, L. J., Filtness, A., & Watson, B. (2018). Driver education and licensing programs. In *Safe mobility: Challenges, methodology and solutions*. Emerald Publishing Limited. https://doi.org/10.1108/S2044-994120180000011002.

Bates, L., Evenhuis, A., & Lennon, A. (2020). Effectiveness of a pre-licensing driver education program on five psycho-social factors over twelve months. *Accident Analysis & Prevention, 148*. https://doi.org/10.1016/j.aap.2020.105806 105806.

Bureau of Infrastructure, Transport and Regional Economics (BITRE) (2018). *Road Trauma Australia, 2017 Statistical Summary*. BITRE, Canberra ACT.

Chapman, E. A., Masten, S. V., & Browning, K. K. (2014). Crash and traffic violation rates before and after licensure for novice California drivers subject to different driver licensing requirements. *Journal of safety research, 50*, 125–138. https://doi.org/10.1016/j.jsr.2014.05.005.

Dee, T. S., Grabowski, D. C., & Morrisey, M. A. (2005). Graduated driver licensing and teen traffic fatalities. *Journal of Health Economics, 24*(3), 571–589. https://doi.org/10.1016/j.jhealeco.2004.09.013.

Delhomme, P. (Ed.) (1999). *GADGET-Project, deliverable 4: evaluated road safety media campaigns: an overview of 265 evaluated campaigns and some meta analysis on crashes*. RR-00-006-FR, INRETS, Arcueill.

Ehsani, J. P., Bingham, C. R., & Shope, J. T. (2013). The effect of the learner license Graduated Driver Licensing components on teen drivers' crashes. *Accident Analysis & Prevention, 59*, 327–336. https://doi.org/10.1016/j.aap.2013.06.001.

Elvik, R., Borger-Mysen, A., & Vaa, T. (1997). *Trafikksikkerhekshandbok (Traffic Safety Handbook)*. Oslo, Norway: Institute of Transport Economics.

Endriulaitienė, A., Šeibokaitė, L., Žardeckaitė-Matulaitienė, K., Markšaitytė, R., & Slavinskienė, J. (2018). Attitudes towards risky driving and Dark Triad personality traits in a group of learner drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 56*, 362–370. https://doi.org/10.1016/j.trf.2018.05.017.

European Conference of Ministers of Transport (ECMT) (2006). *Young drivers: The Road to Safety*. European Conference of Ministers of Transport, OECD Publication. https://doi.org/10.1787/9789282113356-en.

Fernandes, R., Hatfield, J., & Job, R. S. (2010). A systematic investigation of the differential predictors for speeding, drink-driving, driving while fatigued, and not wearing a seat belt, among young drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 13*(3), 179–196. https://doi.org/10.1016/j.trf.2010.04.007.

Floreskul, V., Žardeckaite-Matulaitien, K., Endriulaitiene, A., & Šeibokaite, L. (2016). Effectiveness of pre-driver education programme for high school students: Application of Theory of Planned Behaviour on road risk taking behavior. *Journal of Behavior, Health & Social Issues, 8*(1), 8–16. https://doi.org/10.1016/j.jbhsi.2017.08.003.

Foss, R. D. (2007). Improving graduated driver licensing systems: A conceptual approach and its implications. *Journal of Safety Research, 38*(2), 185–192. https://doi.org/10.1016/j.jsr.2007.02.006.

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update* (10a ed.). Boston: Pearson.

Harbeck, E. L., & Glendon, A. I. (2013). How reinforcement sensitivity and perceived risk influence young drivers' reported engagement in risky driving behaviors. *Accident Analysis and Prevention, 54*, 73–80. https://doi.org/10.1016/j.aap.2013.02.011.

Hassan, H. M., & Abdel-Aty, M. A. (2013). Exploring the safety implications of young drivers' behavior, attitudes and perceptions. *Accident Analysis & Prevention, 50*, 361–370. https://doi.org/10.1016/j.aap.2012.05.003.

Hirschberg, J., & Lye, J. (2020). Impacts of graduated driver licensing regulations. *Accident Analysis & Prevention, 139*. https://doi.org/10.1016/j.aap.2020.105485 105485.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Convential criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118.

Iversen, H., & Rundmo, T. (2004). Attitudes towards traffic safety, driving behaviour and accident involvement among the Norwegian public. *Ergonomics, 47*(5), 555–572. https://doi.org/10.1080/00140130410001658709.

Khan, K., Zaidi, S. B., & Ali, A. (2020). Evaluating the Nature of Distractive Driving Factors towards Road Traffic Accident. *Civil Engineering Journal, 6*(8), 1555–1580. https://doi.org/10.28991/cej-2020-03091567.

Korpinen, L., & Pääkönen, R. (2012). Accidents and close call situations connected to the use of mobile phones. *Accident Analysis and Prevention, 45*, 75–82. https://doi.org/10.1016/j.aap.2011.11.016.

Lipovac, K., Đerić, M., Tešić, M., Andrić, Z., & Marić, B. (2017). Mobile phone use while driving-literary review. *Transportation Research Part F: Traffic Psychology and Behaviour, 47*, 132–142. https://doi.org/10.1016/j.trf.2017.04.015.

Lonero, L., & Mayhew, D. (2010). Large-scale evaluation of driver education: Review of the literature on driver education evaluation 2010 update. *Washington, DC: AAA Foundation for Traffic Safety*. https://doi.org/10.1037/e612252011-001.

Lord, D., Washington, S. P., & Ivan, J. N. (2005). Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory. *Accident Analysis & Prevention, 37*(1), 35–46. https://doi.org/10.1016/j.aap.2004.02.004.

Males, M. (2007). California's graduated driver license law: Effect on teenage drivers' deaths through 2005. *Journal of Safety Research, 38*(6), 651–659. https://doi.org/10.1016/j.jsr.2007.09.002.

Mann, H. N., & Lansdown, T. (2009). Pre-driving adolescent attitudes: Can they change? *Transportation Research Part F: Traffic Psychology and Behaviour, 12*(5), 395–403. https://doi.org/10.1016/j.trf.2009.05.003.

Masten, S. V., Thomas, F. D., Korbelak, K. T., Peck, R. C., & Blomberg, R. D. (2015). *Meta-Analysis of Graduated Driver Licensing Laws* (No. DOT HS 812 211). United States. National Highway Traffic Safety Administration.

Mayhew, D., Marcoux, K., Wood, K., Simpson, H., Vanlaar, W., Lonero, L., & Clinton, K. (2014). Evaluation of Beginner Driver Education Programs: Studies in Manitoba and Oregon.AAA Foundation for Traffic Safety, Washington, D.C.

McCartt, A. T., Mayhew, D. R., Braitman, K. A., Ferguson, S. A., & Simpson, H. M. (2009). Effects of age and experience on young driver crashes: Review of recent literature. *Traffic injury prevention, 10*(3), 209–219. https://doi.org/10.1080/15389580802677807.

McCartt, A. T., Shabanova, V. I., & Leaf, W. A. (2003). Driving experience, crashes and traffic citations of teenage beginning drivers. *Accident Analysis and Prevention, 35*(3), 311–320. https://doi.org/10.1016/S0001-4575(02)00006-4.

McCartt, A. T., Teoh, E. R., Fields, M., Braitman, K. A., & Hellinga, L. A. (2010). Graduated licensing laws and fatal crashes of teenage drivers: A national study. *Traffic Injury Prevention, 11*(3), 240–248. https://doi.org/10.1080/15389580903578854.

Molloy, O., Molesworth, B. R., & Williamson, A. (2019). Which cognitive training intervention can improve young drivers' speed management on the road? *Transportation Research Part F: Traffic Psychology and Behaviour, 60*, 68–80. https://doi.org/10.1016/j.trf.2018.09.025.

Nazif-Munoz, J. I., & Nikolic, N. (2018). The effectiveness of child restraint and seat belt legislation in reducing child injuries: The case of Serbia. *Traffic Injury Prevention, 19*(sup1), S7–S14. https://doi.org/10.1080/15389588.2017.1387254.

NHTSA (National Highway Traffic Safety Administration) (2018). Traffic Safety Facts 2016 Data - Young Drivers. *U.S. Department of Transportation, Washington, DC.*

OECD (1994). *Improving Road Safety by Attitude Modifications*. Report prepared by OECD Scientific Expert Group. Paris: OECD. ISBN 9264141472; Dewey Number 363.125; Libraries Australia ID 11055150.

Preece, C., Watson, A., Kaye, S. A., & Fleiter, J. (2018). Understanding the psychological precursors of young drivers' willingness to speed and text while driving. *Accident Analysis and Prevention, 117*, 196–204. https://doi.org/10.1016/j.aap.2018.04.015.

Rosenbloom, T., Shahar, A., Elharar, A., & Danino, O. (2008). Risk perception of driving as a function of advanced training aimed at recognizing and handling risks in demanding driving situations. *Accident Analysis & Prevention, 40*(2), 697–703. https://doi.org/10.1016/j.aap.2007.09.007.

RTSA (Road Traffic Safety Agency), Republic of Serbia (2019). Review report, Attitudes of traffic participants on alcohol, speed, protection system and mobile phone use in the Republic of Serbia. Retrieved April 20, 2021 from https://www.abs.gov.rs/%D1%81%D1%80/analize-i-istrazivanja/realizovani-projekti.

RTSA (Road Traffic Safety Agency), Republic of Serbia (2021). *The integrated database of characteristics of traffic safety*. RTSA. Retrieved November 21, 2021 from http://195.222.99.60/ibbsPublic/.

Rundmo, T., & Iversen, H. (2004). Risk perception and driving behaviour among adolescents in two Norwegian counties before and after a traffic safety campaign. *Safety science, 42*(1), 1–21. https://doi.org/10.1016/S0925-7535(02)00047-4.

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and social psychology bulletin, 28*(12), 1629–1646. https://doi.org/10.1177/014616702237645.

Scott-Parker, B., & Proffitt, C. (2015). Validation of the Behaviour of Young Novice Drivers Scale (BYNDS) in a New Zealand young driver population. *Accident Analysis & Prevention, 77*, 62–71. https://doi.org/10.1016/j.aap.2015.01.019.

Scott-Parker, B., Goode, N., & Salmon, P. (2015). The driver, the road, the rules… and the rest? A systems-based approach to young driver road safety. *Accident Analysis & Prevention, 74*, 297–305. https://doi.org/10.1016/j.aap.2014.01.027.

Scott-Parker, B., Hyde, M. K., Watson, B., & King, M. J. (2013). Speeding by young novice drivers: What can personal characteristics and psychosocial theory add to our understanding? *Accident Analysis and Prevention, 50*, 242–250. https://doi.org/10.1016/j.aap.2012.04.010.

Scott-Parker, B., Watson, B., & King, M. J. (2010). The risky behaviour of young drivers: developing a measurement tool. In *Proceedings of the 20th Canadian multidisciplinary road safety conference* (pp. 1–15). Canadian Association of Road Safety Professionals.

Senserrick, T. M., & Williams, A. F. (2015). *Summary of literature on the effective components of a graduated licensing scheme for car drivers. Austroads research Report APR476-15*. Sydney, New South Wales, Australia: Austroads Ltd. https://trid.trb.org/view/1346164.

Shell, D. F., Newman, I. M., Córdova-Cazar, A. L., & Heese, J. M. (2015). Driver education and teen crashes and traffic violations in the first two years of driving in a graduated licensing system. *Accident Analysis & Prevention, 82*, 45–52. https://doi.org/10.1016/j.aap.2015.05.011.

Shope, J. T. (2007). Graduated driver licensing: Review of evaluation results since 2002. *Journal of Safety Research, 38*(2), 165–175. https://doi.org/10.1016/j.jsr.2007.02.004.

Siskind, V., Faulks, I. J., & Sheehan, M. C. (2019). The impact of changes to the NSW graduated driver licensing system on subsequent crash and offense experience. *Journal of Safety Research, 69*, 109–114. https://doi.org/10.1016/j.jsr.2019.02.006.

Thomas, F. D., Blomberg, R. D., & Fisher, D. L. (2012). *A fresh look at driver education in America* (No. DOT HS 811 543). United States. National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/1919.

Tronsmoen, T. (2008). Associations between self-assessment of driving ability, driver training and crash involvement among young drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 11*(5), 334–346. https://doi.org/10.1016/j.trf.2008.02.002.

Twisk, D. A., & Stacey, C. (2007). Trends in young driver risk and countermeasures in European countries. *Journal of Safety Research, 38*(2), 245–257. https://doi.org/10.1016/j.jsr.2007.03.006.

Ulleberg, P., & Rundmo, T. (2002). Risk–taking attitudes among young drivers: The psychometric qualities and dimensionality of an instrument to measure young drivers' risk–taking attitudes. *Scandinavian Journal of Psychology, 43*(3), 227–237. https://doi.org/10.1111/1467-9450.00291.

Ulleberg, P., & Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. *Safety Science, 41*(5), 427–443. https://doi.org/10.1016/s0925-7535(01)00077-7.

Vanlaar, W., Mayhew, D., Marcoux, K., Wets, G., Brijs, T., & Shope, J. (2009). An evaluation of graduated driver licensing programs in North America using a meta-analytic approach. *Accident Analysis & Prevention, 41*(5), 1104–1111. https://doi.org/10.1016/j.aap.2009.06.024.

WHO (2018). Global Status Report on road safety 2018. Geneva: World Health Organization. Retrieved from: https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/.

Predrag Stanojević, Ph.D. is a Professor at The Academy of Applied Studies of Kosovo and Metohija, Serbia. His research interests are in Traffic Safety Psychology and Behaviour. He has participated in more than 15 research projects and studies, as well as in more than 10 scientific committees in Serbia and abroad and he has published more than 50 scientific papers. Author and coauthor of number of scholarly or professional papers published in many national and international Journals and Proceedings of International and National Conferences. His recent publications have appeared in Accident Analysis and Prevention and Transportation Research Part F-Traffic Psychology and Behaviour.

Timo Lajunen, Ph.D. is a Full Professor at Norwegian University of Science and Technology, Department of Psychology, Norway. His main field of research is Traffic Safety Psychology and Behaviour. He has participated in more than 30 research projects and studies, as well as in more than 15 scientific committees in Turkey and abroad and he has published more than 100 scientific papers. His recent publications have appeared in Accident Analysis and Prevention and Transportation Research Part F-Traffic Psychology and Behaviour.

Dragana Jakšić, Ph.D. is Teaching Assistant at at The Academy of Applied Studies of Kosovo and Metohija, Serbia. Her research interests are in Traffic Safety Psychology and Behaviour. She has participated in more than 10 research projects and studies in Serbia. He is the author and coauthor of a number of publications in journals and conference proceedings. Her recent publication has appeared in Traffic Injury Prevention.

Dragan Jovanović, Ph.D. is a Full Professor and Head of Department at University of Novi Sad, Faculty of Technical Sciences, Department of Transport Engineering, Novi Sad, Serbia. His main field of research is Road Safety. He has participated in more than 20 research projects and studies, as well as in more than 15 scientific committees in Serbia and abroad and he has published more than 80 scientific papers. His recent publications have appeared in Accident Analysis and Prevention, Safety Sciences, Transportation Research Part F-Traffic Psychology and Behaviour.

Bosko Matović, Ph.D. is an Assistant Professor at University of Montenegro, Faculty of Mechanical Engineering, Podgorica, Montenegro. His research interests are in Traffic Safety Psychology and Behaviour. He has participated in more than 10 research projects and studies in Serbia and Montenegro. He is the author and coauthor of a number of publications in journals and conference proceedings. His recent publications has appeared in Accident Analysis and Prevention and Traffic Injury Prevention.

Journal of Safety Research 83 (2022) 152–162

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Effect of transit-oriented design on pedestrian and cyclist safety using bivariate spatial models

Mankirat Singh [a], Yongping Zhang [a], Wen Cheng [a,*], Yihua Li [b], Edward Clay [a]

[a] Department of Civil Engineering, California State Polytechnic University, Pomona, Pomona, CA 91768, United States
[b] Department of Logistics Engineering, Logistics and Traffic College, Central South University of Forestry and Technology, Hunan 410004 30, China

ARTICLE INFO

Article history:
Received 6 May 2021
Received in revised form 15 November 2021
Accepted 18 August 2022
Available online 28 August 2022

Keywords:
Pedestrian and bicyclist involved crashes
Crash frequency models
Bivariate models
Transit stops

ABSTRACT

Introduction: Walking and cycling for transportation provide immense benefits (e.g., health, environmental, social). However, pedestrians and bicyclists are the most vulnerable segment of the traveling public due to the lack of protective structure and difference in body mass compared with motorized vehicles. Numerous studies are dedicated to enhancing active transportation modes, but very few studies are devoted to the safety analysis of the transit stops, which serve as the important modal interface for pedestrians and bicyclists. Method: This study bridges the gap by developing joint models based on the multivariate conditional autoregressive (MCAR) priors with distance-oriented neighboring weight matrix. For this purpose, transit-oriented design (TOD) related data in Los Angeles County were used for model development. Feature selection relying on both random forest (RF) and correlation analysis was employed, which leads to different covariates inputs to each of the two joint models, resulting in increased model flexibility. An integrated nested Laplace approximation (INLA) algorithm was adopted due to its fast, yet robust, analysis. For a comprehensive comparison of the predictive accuracy of models, different evaluation criteria were utilized. Results: The results demonstrate that models with correlation effect perform much better than the models without a correlation of pedestrians and bicyclists. The joint models also aid in the identification of the significant covariates contributing to the safety of each of the two active transportation modes. The findings show that population density, employment density, and bus stop density positively influence bicyclist-involved crashes, suggesting that an increase in population, employment, or the number of bus stops leads to more active modes involved collisions. Practical Applications: The findings of this study may prove helpful in the development and implementation of the safety management process to improve the roadway environment for the active modes in the long run.

1. Introduction

The challenge of meeting the mobility requirements of the 21st century requires a shift in mindset from designing an automobile-focused highway system to operating a transportation network that accommodates all users and modes safely and conveniently. The goal may be achieved by implementing complete streets design, which allows the flexibility to enhance traffic safety and strategic urban mobility planning (Smith et al., 2010). Typical elements of the complete streets include sidewalks, bicycle lanes (or wide, paved shoulders), shared-use paths, designated bus lanes, safe and accessible transit stops, and frequent and safe crossings for pedestrians and bicyclists. Among the different roadway facilities, the transit stop plays a crucial role in successfully implementing complete streets programs due to its unique position as an intermodal interface. It is worth mentioning that here "transit" refers to both bus transit and rail transit (road-based and segregated rail transit). Compared with other modes, non-motorized transportation modes provide enormous health, environmental, social benefits, and many others. However, the non-motorists are a vulnerable segment of the traveling public due to the lack of a protective structure and difference in body mass as compared to motor vehicles, which renders them prone to heightened injury susceptibility in case of a collision (Mader & Zick, 2014; Cai et al., 2017). Therefore, incorporating transportation network attributes into traffic safety would facilitate the development of safety

* Corresponding author.
E-mail addresses: mankirats@cpp.edu (M. Singh), yongpingz@cpp.edu (Y. Zhang), wcheng@cpp.edu (W. Cheng), yhli@csuft.edu.cn (Y. Li), erclay@cpp.edu (E. Clay).

https://doi.org/10.1016/j.jsr.2022.08.012
0022-4375/

programs and better strategies to engender a safe environment for all roadway users.

On the other hand, the Sustainable Communities and Climate Protection Act of 2008, Senate Bill (SB) 375 in California, requires that Metropolitan Planning Organizations (MPOs) develop a Sustainable Communities Strategy (SCS) to reduce per capita greenhouse gas emissions through integrated transportation, land use, housing, and environmental planning. SB 375 creates incentives for residential or mixed–use residential projects that may be exempt from, or subject to a limited review of, the California Environmental Quality Act (CEQA), provided they are consistent with the MPO's adopted SCS. These "transit priority projects" must, among other criteria, be located within one–half mile of a major transit stop or high–quality transit corridor (HQTC).

SB 743, signed into law in 2013, provides further opportunities for CEQA exemption and streamlining to facilitate transit-oriented development (TOD). Specifically, certain types of projects within "transit priority areas" (TPAs) can benefit from a CEQA exemption if they are consistent with an adopted specific plan and the SCS. A TPA is an area within a half-mile of a major transit stop that is existing or planned, if the planned stop is scheduled to be completed within the planning horizon included in a Federal Transportation Improvement Program (FTIP).

The 2016–2040 SCAG Regional Transportation Plan/Sustainable Communities Strategy (RTP/SCS) plans for focusing new growth around transit, particularly in the High Quality Transit Areas (HQTAs). While HQTAs account for only 3% of total land area in SCAG region, they are planned and projected to accommodate 46% of the region's future household growth and 55% of the future employment growth. With more development and growth focused around transit stops, it is anticipated that the traffic accidents around transit stops will increase too, without proper proactive planning and design. Meanwhile, the Fixing America's Surface Transportation (FAST) Act calls for establishing performance measures and standards on traffic safety. The Federal Highway Administration (FHWA) is now requiring State Department of Transportation (DOTs) to work with MPOs to assess fatalities and serious injuries on all public roads and to set annual performance measures.

To address this urgent need, previous studies have strived to obtain valuable insights by considering various empirical and methodological aspects of non-motorist safety modeling. Crash frequency models are often used to identify the factors influencing the propensity of active-modes-related crashes. As the crash frequency data are non-negative integers, the most widely used crash frequency models assume the Poisson distribution of crash counts. The initial Poisson regression models are subject to the limitation of equality between mean and variance of crash counts (Kim et al., 2002; Miranda-Moreno, 2006), which means Poisson models cannot handle over-dispersion and can be adversely influenced by low-sample means. The presence of such issues in data could result in biased and inconsistent parameters, leading to erroneous inferences and predictions relating to the factors that estimate crash-frequencies (Oh et al., 2006; Lord & Mannering, 2010). In this regard, subsequent enhancement contains various model alternatives including, but not limited to, Poisson gamma or negative binomial (Hauer, 2001; El-Basyouny & Sayed, 2006; Lord & Bonneson, 2007), Poisson lognormal (Park & Lord, 2007; Lord & Miranda-Moreno, 2008; Aguero-Valverde & Jovanis, 2008; Daniels et al., 2010), and zero-inflated models (Malyshkina & Mannering, 2010; Washington et al., 2020; Aguero-Valverde, 2013a,b), which can address crash over-dispersion using different model formulations.

Among the models mentioned above, the most extensively used format is the univariate model, which contains only one dependent variable for the data interpretation (Anarkooli et al., 2019). However, the univariate setting cannot address the unobserved heterogeneities that might be common to various crash types or severities (Mannering & Bhat, 2014). In response, multivariate models have been employed extensively to explicitly account for the possible correlations among the distinct response variables. Some papers relied on the bivariate framework for various applications such as angled injury security (Russo et al., 2014) and investigation of bicycle conflict location (Conway et al., 2013), while others took advantage of multivariate models to address response variable of multiple discrete outcomes like different crash types (Serhiyenko et al., 2016) and crash involving distinct modes (Huang et al., 2017). Another benefit of using bivariate/multivariate is the explicit consideration of correlation among different crash outcomes (Bijleveld, 2005; Song et al., 2006; Park & Lord, 2007; Aguero-Valverde & Jovanis, 2009). Even with well-documented benefits over the univariate alternatives (Mothafer et al., 2016), studies are still dedicated to further enhancing the bivariate models.

In addition to the above-mentioned different models, one popular strategy is the explicit consideration of spatial effect in the safety analysis of active transportation. For example, some research incorporated pedestrian safety into urban road space allocation (Chen et al., 2020), some investigated the effect of road network configuration urban design on incidences of pedestrian and cyclist crashes (Dumbaugh et al., 2013), yet others explored the use of new data types such as household travel survey data and bike sharing usage data to estimate the pedestrian and cyclist crash exposure (Branion-Calles et al., 2021; Ding et al., 2020) Incorporating geographic information collected by sophisticated software such as geographical information system (GIS) allows to include the influential factors relating to the spatial perspective. Nevertheless, the data from the same geographic area may share unobserved effects and arise spatial correlation problems. Given this context, various studies develop Bayesian spatial models such as conditional autoregressive (CAR) (Song et al., 2006; Soroori et al., 2019; Zeng et al., 2020), simultaneous autoregressive (SAR) (Quddus, 2008; Chiou et al., 2014; Hosseinpour et al., 2018), multivariate conditional autoregressive (MCAR) (Aguero-Valverde, 2013a,b; Lee et al., 2015; Cai et al., 2018) to address the spatial correlation issue. Another prevalent method to account for spatial correlation is the inclusion of random effects models in which common unobserved effects are assumed to be uncorrelated with independent variables and distributed over spatial units. In the context of crash frequencies, random effects models have been used by a large number of studies.

All in all, active transportation has gained ever-increasing popularity due to its multiple benefits over the typical motorized modes. However, how to improve the safety of non-motorists plays a pivotal role in promoting such healthy, economical, and environmentally friendly modes, especially at various transit stops where different transportation modes interact with one another. Compared with other transportation facilities such as intersections, sidewalks, and bike lanes, there is considerably less research dedicated to safety analysis along with the transit stops. Until now, it remains unclear which factors constitute the main contributing ones to the walking and biking safety conditions in the areas adjacent to transit stops, given the complexity of the influential factors and their interactions. To bridge this gap, the present study aims to rank the importance and quantify the impact of pedestrian and bicyclists traffic safety-pertinent variables near the transit (bus transit and rail transit) stops, which are imperative for the effective design of complete streets and transportation planning policies. For this purpose, bivariate spatial models were utilized owing to the frequent advantages reported in previous research associated with multivariate settings and spatial heterogeneity consideration. Specifically, joint models based on the multivariate conditional autoregressive (MCAR) priors with distance-oriented neighboring weight matrix were used. In order to take advantage of a substantial reduction in computational time for estimation under a complex model scenario (Serhiyenko et al., 2016; Blangiardo et al., 2013), the current study employs an integrated nested Laplace approximation (INLA)-oriented package, INLAMSM (Palmi-Perales et al., 2019), to carry out approximate Bayesian inference within a bivariate spatial framework. In addition, feature selection using both random forest (RF) and correlation analysis is employed, which yields different covariates to each of the two active transportation modes and leads to an increase in the model flexibility. Moreover, high-quality transit-oriented development (TOD) related data, including the built environment, socioeconomic, demographic, and crash data aggregated at the 0.5-mile-radius circular zone surrounding transit stops, were used for statistical analysis. Not only the number of bus and rail transit stops, but also the characteristics of transit stops (such as the configuration of transit stops, accessible routes, and crossing) are considered in this study. Finally, for a comprehensive comparison of the predictive accuracy of models, different evaluation criteria were utilized, which include deviance information criterion (DIC), widely applicable information criterion (WAIC), posterior mean deviance ($\bar{D}$) and the effective number of parameters (P_D).

The proposed project expects to promote active transportation and enhance the multimodal traffic safety conditions adjacent to the HQTAs, which is imperative for successive implementation of complete street design and SCAG's RTP/SCS. The research results are anticipated to furnish transportation professionals with additional insights to create safer access to transit and thus promote active transportation in the United States.

The paper is structured as follows: Section 2 shows the existing literature relevant to this study by highlighting previously published studies. Section 3 describes the methodology used in this paper. Section 4 presents the data collected for the study. Then, the empirical estimation results and discussion are included in Section 5. Finally, Section 6 concludes the paper with recommendations.

2. Literature review

A plethora of studies in the safety field has employed bivariate/-multivariate models for crash count data. For example, Wang et al. (2013) developed a Poisson-lognormal conditional autoregressive model for their bivariate spatial analysis of pedestrian crash counts across census tracts in Austin, Texas. The results indicate that a bivariate cross-correlation of serious (fatal and major injury) and non-serious crash rates shows covariates' impacts across severity levels are more local in nature (e.g., lighting conditions or local sight obstructions along with spatially lagged cross-correlation). Ma et al. (2008) used a multivariate Poisson-lognormal specification to investigate different crash counts at different severity levels. Their findings show that the bivariate/multivariate Poisson-lognormal model aids in showing the statistically significant correlations between crash counts at a different level of injury severity.

Explicit consideration of spatial autocorrelation in the bivariate/multivariate settings is one popular strategy and includes both random effect (Aguero-Valverde, 2013a,b) and random parameter models (Barua et al., 2016; Imprialou et al., 2016). However, bivariate/multivariate models are complex to estimate as the correlation matrix formulation is required (Lord & Mannering, 2010). Employing another method to account for spatial correlation, many studies consider the random effects models where the common unobserved effects are assumed to be uncorrelated with independent variables and distributed over spatial units. A study conducted by Hausman et al. (1984) examined random effects and fixed effects in negative binomial models for panel data in their research. The findings suggest that random effects help to account for unobserved factors shared by distinct, discrete outcomes. In the context of crash frequencies, random effects have been used by a large number of previous studies. For instance, Ma et al. (2017) proposed a series of multivariate models under the framework of Full Bayesian with different random effects to predict the crash frequencies of different injury severity levels over a one-year period in Colorado. Cheng et al. (2017) developed the multivariate Poisson lognormal models with random effects to predict the motorcycle injury severity crashes using weather data during the years 2008–2013 in the city of San Francisco. Hou et al. (2018) developed the random effect negative binomial (RENB) model by investigating the effects of traffic characteristics and freeway design elements on crash frequency in China. Besharati et al. (2020) utilized the bivariate spatial negative binomial Bayesian model with random effects to examine the association between the fuel consumption in the transportation sector and the annual fatal and nonfatal injury counts from 2005 to 2015 in Iran. Nonetheless, the random effects models only influence the intercept of the model. The extension of random effects models is the random parameter models that provide the flexibility to accommodate the site-specific unobserved heterogeneity by allowing each estimated parameter to vary across each individual observation in the data (Anastasopoulos & Mannering, 2009; Milton et al., 2008).

To carry out the Bayesian inference, the Markov Chain Monte Carlo (MCMC) simulation method is the most popular approach used in the safety field. Park and Lord (2007) adopted the MCMC simulation method in multivariate Poisson-lognormal models to evaluate covariates' impact on crash counts. Similarly, El-Basyouny and Sayed (2006) used multivariate Poissonlognormal models with the MCMC simulation approach through the WinBUGS platform to jointly analyze a dataset of crash counts categorized by two injury severity levels. However, the Bayesian framework using the MCMC simulation method may be computationally challenging under a complex model scenario and time-consuming, especially for large datasets (Narayanamoorthy et al., 2013). To address this issue, the present research adopts an alternative Bayesian approach, Integrated Nested Laplace Approximation or INLA (Rue et al., 2009), to carry out approximate Bayesian inference. The INLA method aids in reducing the computational time and efforts involved in the estimation of complex and large datasets (Serhiyenko et al., 2016; Blangiardo et al., 2013) sevenfold compared to the MCMC simulation method (Serhiyenko, 2015).

3. Data description

The data used in this research were obtained from multiple sources: Southern California Association of Governments (SCAG), Transportation Injury Mapping System (TIMS), Los Angeles Metropolitan Transportation Authority (LACMTA), InfoUSA, and U.S. Census Bureau (Census). All data were compiled into GIS databases. The dependent variables include the total number of pedestrian-involved crashes, the total number of bike-involved crashes, and the total number of vehicle-only crashes (i.e., crashes with no pedestrian or bike involved). The independent variables are categorized into various groups, including the following: socioeconomic characteristics, employment characteristics, diversity/mixed use of land, built environment/access to active transportation and transit, land development characteristics such as Transit-Oriented Development (TOD), and biking-/walking-related built environment variables.

SCAG's High-Quality Transit Area (HQTA) is within one half-mile from major transit stops and high-quality transit corridors

(HQTC), and it was developed based on the language in Senate Bill 375 (Barbour, 2016). According to SCAG, the definitions of major transit stops and HQTC are as follows:

Major Transit Stop: a site containing an existing rail transit station, a ferry terminal served by either a bus or rail transit service, or the intersection of two or more major bus routes with a frequency of service interval of 15 minutes or less during the morning and afternoon peak commute periods (C.A. Public Resource Code Section 21064.3). It also includes major transit stops that are included in the applicable regional transportation.

HQTC: a corridor with fixed-route bus service, and the service intervals are no longer than 15 minutes during peak commute hours. (SCAG, 2020).

It is worth mentioning that, for the research purpose, the analysis is focused on the one-half mile buffer zones around high-quality transit stops in the High-Quality Transit Area (HQTAs) in the SCAG region for the year 2016 (see Fig. 1). As mentioned above, SCAG adopted the one-half mile radius from the major transit stop to define the HQTA and HQTC. To be consistent with SCAG's definition, the observation unit in this study is the one-half mile buffer zone surrounding the major transit stop. Most variables used in this study were obtained or derived from SCAG's travel demand and land-use models. There are 948 major transit stops in the SCAG region, including 155 rail stations. Fig. 1 shows the location of these stops for the year 2016. Out of 948 stops, 870 stops are in Los Angeles County (LAC). Without losing the representativeness, this research focuses on the major transit stops in LAC. Furthermore, to avoid the duplication among stops that are too close to each other because of the half-mile buffer, 293 stops that are within one mile to each other are removed in the analysis. This results in 655 stops in LAC.

Transit-oriented development (TOD) is mixed-use development designed/planned to be near transit stops. TOD is characterized by "a mix of residential, commercial, and civic uses within walking distance (a half-mile radius) from a transit stop; pedestrian-friendly streets with sidewalks and walkable destinations; reduced parking; high-density development; preservation of open space; and a variety of housing types and prices" (Policy & Tools: Transit-Oriented Development (TOD), https://www.forworking-families.org/resources/policy-tools-transit-oriented-development-tod).

To better characterize TOD, a variety of built environment-related variables were derived at SCAG's Tier 2 TAZ (traffic analysis zones) level. There are 11,267 Tier 2 TAZs in the SCAG travel demand model. Table 1 shows the descriptive statistics of the variables used in the research. The bottom row indicates the statistics of the stop-to-stop distance (on network) matrix used for spatial autocorrelation analysis in the model development.

As shown in Table 1, variables directly associated with transit stops include:

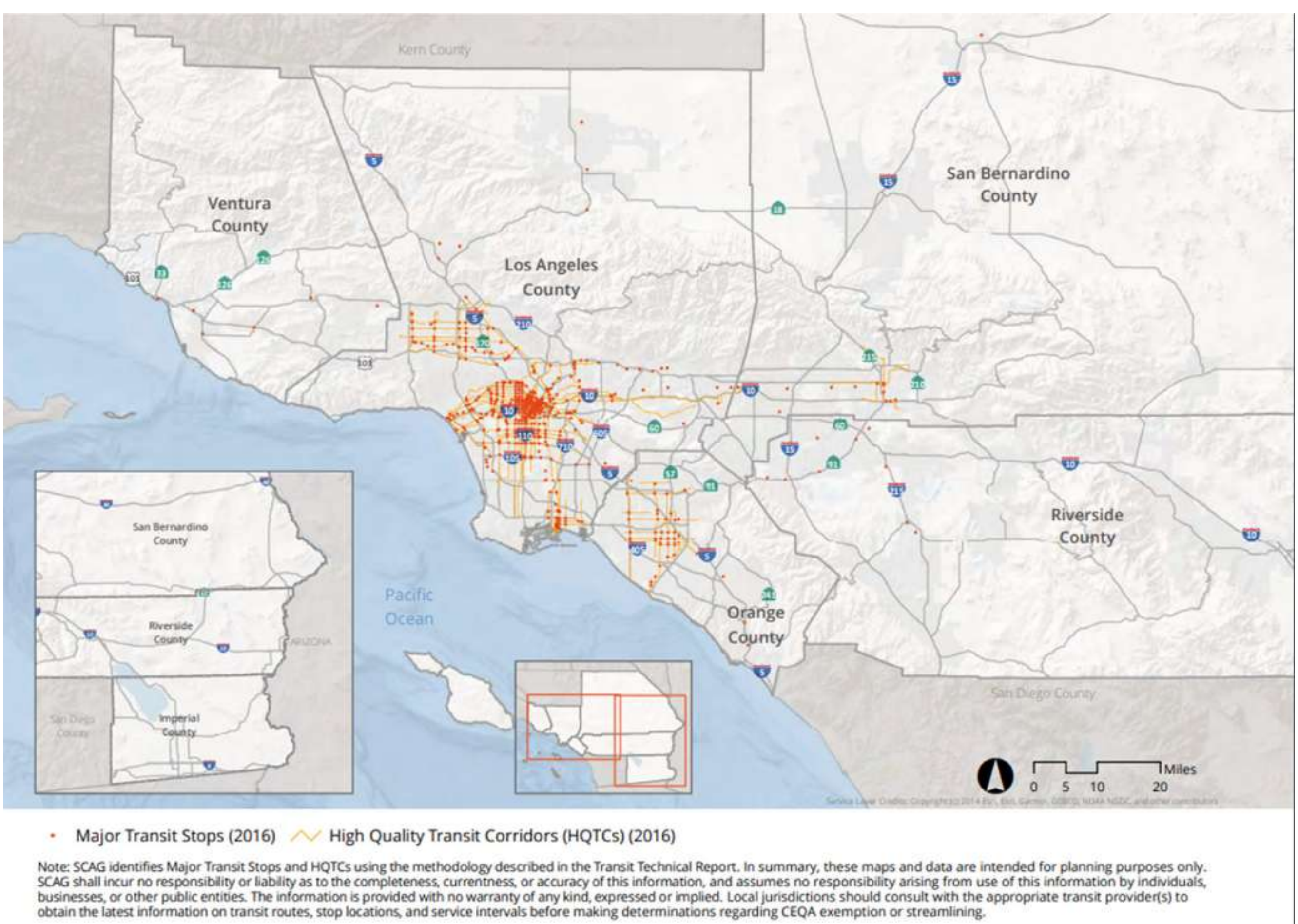

Fig. 1. 2016 Existing High-Quality Transit Corridors in the SCAG Region (Source: SCAG 2045 RTP/SCS).

ExBus_D1	Stop density for Express Bus and BRT (number of stops per acre).
HFLbus_D1	High-Frequency Bus Stop Density (local bus headway $\leq$ 20 mins).
TTbus_D1	Total Bus Stop Density.
HQTA_pct1	Percent of TAZ area are in non-freeway HQTA (high-quality transit area).
TPA_pct1	Percent of TAZ area are in TPA.
Distance	The distance between each pair of the 655 transit stops (miles).

A total of 250,817 non-freeway collisions in LA County from 2012 through 2017 were collected from the Transportation Injury Mapping System (TIMS) website, which provides quick, easy, and free access to California crash data maintained by the Statewide Integrated Traffic Records System (SWITRS). Each collision was geocoded to the actual location. First, a GIS layer with the selected major transit stops was used to create the one-half-mile buffer zones. Then, the GIS layer was intersected with SCAG's Tier 2 TAZ, where the built-environment information is available. Lastly, the above two layers were overlaid with another GIS layer that stores collision information in LA County from 2012 through 2017 (see Fig. 2 for the screenshot of the collision layer overlaying transit stop buffer zones). The integrated dataset was cleaned and developed, containing transit-oriented crash count, built environment, and other related information.

4. Methodology

This section presents the methodological details, including modeling specification, variable importance by random forest, and evaluation criteria.

4.1. Modeling specification

In the current study, under the framework of INLA (Lindgren & Rue, 2015), INLAMSM package (Palmi-Perales et al., 2019) is used to draw the inferences due to its capability to address bivariate latent effects. The Bayesian hierarchical approach is used to estimate the Poisson process:

$$y \sim Poisson(\boldsymbol{\lambda}) \tag{1}$$

where y is the observed crash count, and λ is the Bayesian mean expected crashes, which can be modeled as a function of the covariates following a lognormal distribution as shown below:

$$\log(\lambda) = \beta_0 + \boldsymbol{\beta X} + \phi + \boldsymbol{\varepsilon} \tag{2}$$

Table 1
Variable description of collected data.

Variables	Description
Density (TAZ level)	
Pop_den1	Population density (persons/acre)
HH_den1	Household density (households/acre)
Emp_den1	Employment density (jobs/acre)
Ret_den1	Retail job density (jobs/acre)
RetSer_den1	Retail + Service (retail + FIRE + Food + Other Serv.) job density
Diversity/Mixed Use of Land (TAZ Level)	
Jobmix131	Employment mix (13sectors); 1 = highest mix (jobs are equal for all sectors)
Jobmix91	Employment mix (9 sectors); 1 = highest mix (jobs are equal for all sectors)
Emix131	Employment mix (13sectors); 1 = highest mix (jobs are equal for all sectors)
Emix91	Employment mix (9 sectors); 1 = highest mix (jobs are equal for all sectors)
EH_ratio1	Job/Household ratio
EP_ratio1	Job/Population ratio
Built Environment/Access to Active Transportation and Transit (TAZ Level)	
Int34_Den1	Intersection density (number of intersections/acre)
BKlnAcc1	Bike lane access (1 = if a TAZ has a bike lane)
Rail1	1 = at least one rail stop in a TAZ
ExBus_D1	Stop density for Express Bus and BRT (number of stops per acre)
HFLbus_D1	High-Frequency Bus Stop Density (local bus headway $\leq$ 20 mins)
TTbus_D1	Total Bus Stop Density
Land Development Characteristics: TOD (HQTA/TPA) (TAZ Level)	
Mlt_pct1	Percent of households living in multiple units
HQTA_pct1	Percent of TAZ area are in non-freeway HQTA (high-quality transit area)
TPA_pct1	Percent of TAZ area are in TPA (transit prority area)
Additional Biking or Walking Related Built Environment Variables (TAZ Level)	
BLdenIND1	Bike Lane Density Indicator = Sum (Bike Lane Density/Distance to Home TAZ within 3 miles) Bike Lane Density for Each TAZ = ((Street15-25 mph) * 1 + (Street 35 mpg) * 2 + Bike Lane Class1 * 3 + Bike Lane Class2 * 4 Bike Lane Class 3 * 5)/Total TAZ area (excluding speed > 60mph)
Blck_len1	Estimated block length = Local St/Int34new (Total street length/number intersection), but freeways and state highways are excluded
WalkAcc1	Walk Accessibility (RS_den2/block_len) = (weighted retail + service density)/estimated block length
Pct_Art1	Percent of main arterial (45–55 mph) of TAZ – higher % means more difficult to across street (also larger block to across street); can be used with WalkAcc
Natural Log Transformation	
L_Pden1	L_Pden = Ln (Pop_den + 0.001)
L_Hden1	L_Hden = Ln (HH_den + 0.001)
L_Eden1	L_Eden = Ln (Emp_den + 0.001)
L_REden1	L_REden = Ln (Ret_den + 0.001)
L_RSEden1	L_RSEden = Ln (RetSer_den + 0.001)
Crash Count (Geo-coded to the actual location)	
Ped	Pedestrian-involved accidents count.
Bike	Bike-involved accidents count.
Distance	
Distance	The distance between each pair of the 655 transit stops (miles)

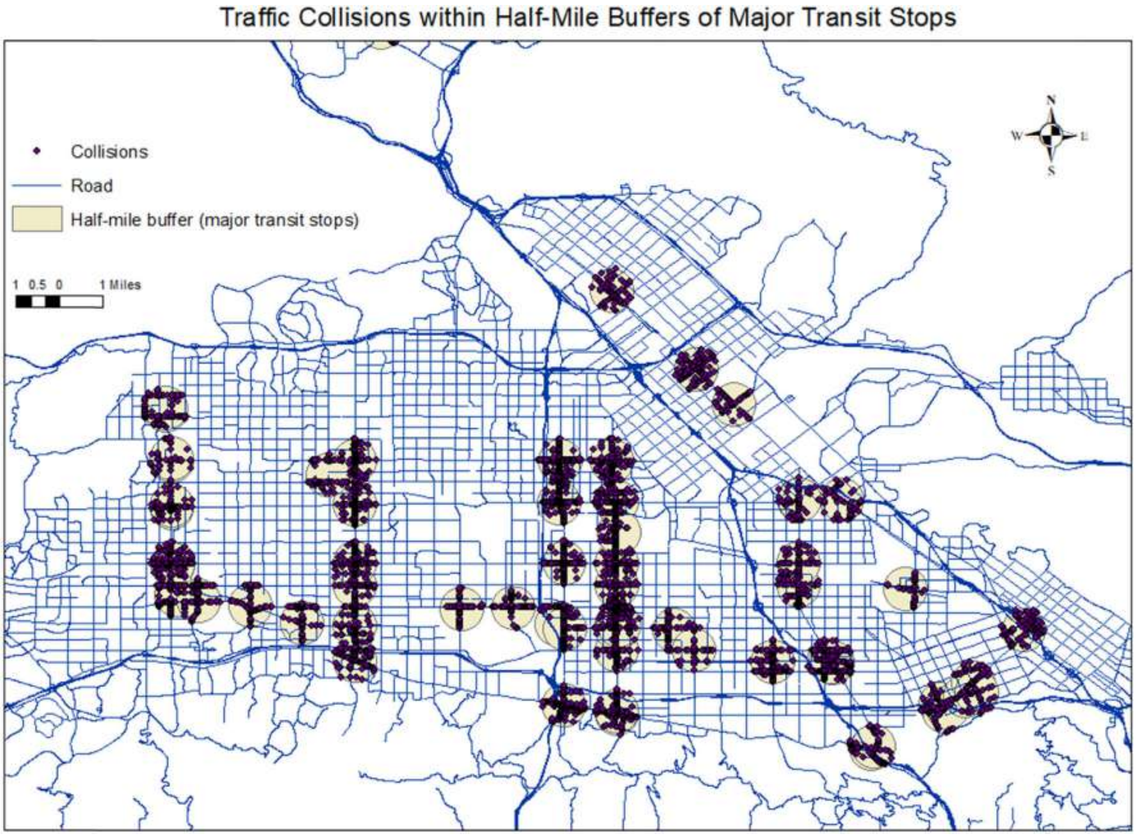

Fig. 2. Traffic Collisions within Half-Mile Buffers of Major Transit Stops in the San Fernando Valley (Part of LA County).

where β_0 is the global intercept, $\boldsymbol{\beta}$ is a fixed coefficient vector, $\boldsymbol{X}$ is the covariate matrix, ϕ is the spatially structured error term, and fit by the multivariate conditional autoregressive (MCAR), $\boldsymbol{\varepsilon}$ represents the white noise matrix. For the bivariate model, correlated priors in the random effects vector are estimated using normal priors (Ma & Kockelman, 2006; Park & Lord, 2007):

$$\boldsymbol{\varepsilon} \sim Normal\left(\mu, \sum\right) \tag{3}$$

$$\text{where } \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \sum = \begin{pmatrix} \tau_1 & \sqrt{\tau_1\tau_2}/\rho_{12} \\ \sqrt{\tau_1\tau_2}/\rho_{12} & \tau_2 \end{pmatrix} \tag{4}$$

In the above equations, *Normal* represents the bivariate normal distribution, $\boldsymbol{\varepsilon}$ is the independent random effect matrix which captures the extra-Poisson heterogeneity among locations, μ is the vector consisting of the mean value for each transportation mode, and $\sum$ is called the precision matrix where the diagonal elements τ's represent the marginal precision of each of the transportation modes, while the off-diagonal elements represent the inverse of covariance, calculated as the ratio of $\sqrt{\tau_1\tau_2}$ and ρ_{12} (or the correlation coefficient between the two response variables). If no correlation between the transportation modes is assumed, the off-diagonal elements can be specified to zero. In this research, both correlated and non-correlated modes are considering for model performance purposes. This inverse of the precision matrix is defined by:

$$\sum\nolimits^{-1} \sim Wishart(I, n) \tag{5}$$

where the $\sum^{-1}$ is a symmetric positive definite matrix, I is the scale matrix (Congdon, 2007), and J (J = 2) is the degree of freedom, resulting in a non-informative specification (Heydari et al., 2017). The covariates coefficient was specified with a normally distributed vague priors N (0,100). Such diffused normal distribution with mean values of zero and a large variance is commonly employed as a vague prior to posterior estimates *n* the absence of sufficient knowledge of priori distribution (Osama & Sayed, 2017; Cheng et al., 2018).

To tackle the spatial dependency, the authors employed the MCAR algorithm initially derived by Mardia (1998) from the results in Besag (1974). Let $\phi^T = (\phi_1^T \cdots \ldots \phi_m^T)$, where Φ is nm $\times 1$, with each ϕ being an n-dimensional vector. In the present study, n = 2 representing the two transportation modes, and m = 655 representing the 655 transit stops in Los Angeles County. Considering a bivariate Gaussian distribution for Φ:

$$(\Phi) = (2\pi)^{-\frac{nm}{2}}|S|^{\frac{1}{2}} \exp\left(-\frac{1}{2}\Phi^T S \Phi\right) \tag{6}$$

where S is a nm $\times$ nm precision matrix. The matrix can also be considered as a m $\times$ m block matrix n $\times$ n blocks S. Following the Mardia (1998), the zero-entered MCAR, which has a conditional normal density, is shown as follows:

$$\phi_i|\phi_{j,} \sum_i \sim N_k\left(\sum_{j\sim 1} C_{ij}, \phi_{j,} \sum\nolimits_i\right) \tag{7}$$

where subscripts *i* and *j* refer to a transit stop and its neighbors *j*, each $\sum_i$ is an n $\times$ n positive definite matrix representing the conditional precision matrix, C_{ij} is a distance matrix of the same dimensions $\sum_i$ (Jonathan et al., 2016). The precision matrix $\sum^{-1}$ follows the Wishart distribution as shown in Eq. (5).

4.2. Variable importance by random forest

Random forest (RF) is an ensemble classifier that consists of many decision trees and outputs the model estimates by individual trees. The method combines bagging and the random selection of features to construct a collection of decision trees with a controlled variation. Using ensembles of predictors for classification or regression has proved to yield more accurate results than using a single predictor. This technique has an advantage over the traditional decision trees of obtaining unbiased error estimates without separating the cross-validation test dataset. When a particular tree is grown from a bootstrap sample, usually-one-third of the training cases, called out-of-bag (OOB) data, are left out and not used to grow the tree.

The efficient implementation of the RF algorithm relies on two important components: the number of trees to grow, and the number of predictors that would be selected to split each node to produce stable results and a minimum OOB error rate. Once the proper values of the tree number and predictor size were determined, the variable importance ranking was reported based on the mean decrease of accuracy in predictions on the OOB samples when a given variable is excluded from the model. A similar practice is shown in some previous research in the traffic safety field (Abdel-Aty & Haleem, 2011; Siddiqui et al., 2012; Jiang et al., 2016). As illustrated in these studies, compared with the typical feature selection techniques such as forward or backward selection based on statistical model metrics, the RF is free of specific data distribution and features the enhanced capability to handle data complexity, especially those with a high order of interactions. The interested readers can refer to the pertinent document (James et al., 2013) for details of RF and variable importance ranking.

4.3. Evaluation criteria

For Bayesian hierarchical model evaluation, deviance information criterion (DIC) is a popular criterion used to assess the models' complexity and goodness of fit (Spiegelhalter et al., 2003). As a hierarchical modeling generalization of the Akaike information criterion (AIC), DIC can be expressed as using the following formulation:

$$\mathrm{DIC} = \mathrm{D}\left(\bar{\theta}\right) + 2P_D = \bar{D} + P_D \tag{8}$$

where D $\left(\bar{\theta}\right)$ is the deviance evaluated at the posterior means of estimated unknowns $\left(\bar{\theta}\right)$, and posterior mean deviance $\bar{D}$ can be taken as a Bayesian measure of data-fitting. P_D denotes the effective number of parameters in a model, as the difference between D $\left(\bar{\theta}\right)$ and $\bar{D}$. In general, the difference between observed and model-predicted data decreases as the number of parameters in a model increases. Therefore, the P_D term is mainly used to compensate for this effect by favoring models with a smaller effective number of parameters. The larger the DIC value, the worse the model tends to perform. As a general rule of thumb suggested by Lunn et al. (2012): the models with a DIC value less than five are considered to have the same performance. The models with a greater DIC value by 5 and 10 points are slightly worse, and the models with a larger DIC by more than 10 points are significantly worse. Overall, DIC may be regarded as the measure of an indirect assessment of the out-of-sample errors as it is based on in-sample errors ($\bar{D}$), while also accounting for the bias due to overfitting usually resulting from more model parameters (James et al., 2013).

Like DIC, the widely applicable information criterion (WAIC) is another generalized version of AIC. For Bayesian models, WAIC (Watanabe, 2010) can be viewed as an improvement on the DIC, which is not fully Bayesian since it is based on a point estimate (van der Linde, 2005; Plummer, 2008). By contrast, WAIC is fully Bayesian, which is invariant to parametrization and closely approximates Bayesian cross-validation using leave-one-out techniques. Like DIC, the model with smaller WAIC is more preferred (Gelman et al., 2013). WAIC was also used in the present research as an additional criterion to assess model performance from a different perspective.

5. Results

For the bivariate spatial crash prediction models, the different covariates were first selected for each of the two transportation modes. The R-package "INLAMSM" was utilized to develop the models and generate the posterior mean of the model parameters. Distinct evaluation criteria were finally employed to assess the model performance.

5.1. Feature selection

Based on the parsimony rule, it is often desirable to reduce the model data load to the fewest number of inputs with maximal predictive accuracy. The typical feature selection techniques integrated with statistical model development like backward-forward feature selection are usually subject to the strong assumption of the particular distribution function and lack the capability to handle the possible complex variable interactions. Therefore, the present study performs feature selection using the correlation analysis and one of the ensemble techniques (or RF), which has recently gained popularity due to its benefits over the typical techniques.

The RF model was developed via the R package "randomforest" (Cutler et al., 2012). During the tree-growing, four predictor variables were randomly sampled as candidates at each split, where the OOB error rate was found to be at a minimum of 0.264, with 63.24% of data variability being explained by the model. Once the RF model was generated, the variable importance ranking was determined based on the contribution of variables to reduce the mean squared errors (MSE) in the OOB samples. The larger the contribution, the more important the particular variable tends to be for model development. The plots for the variable importance are shown in Fig. 3 with a decreasing order (vertically) for both pedestrian and bike counts.

This study also performed a correlation analysis to avoid supplying redundant information fed into the models. To determine whether the variables are highly correlated or not, the popular cut line 0.6 for the correlation coefficient with the significance level of 0.05 was used. The correlated variables were removed in multiple steps using the engineering judgment to avoid excluding any significant variables. This procedure acts as a tradeoff between omitted variable bias and multi-collinearity. At last, out of 32 variables, 14 were retained to perform modeling development, as shown in Table 2 as variables marked in bold. It is noteworthy that different covariates are used for different transportation modes, which enhances the model flexibility and accuracy, with more related important variables being used for the respective modes.

5.2. Model parameter estimates and performance evaluation

Once the proper covariates were selected for pedestrian and bicyclist-involved collisions, two bivariate spatial models were developed with and without explicitly considering the correlation of the two modes.

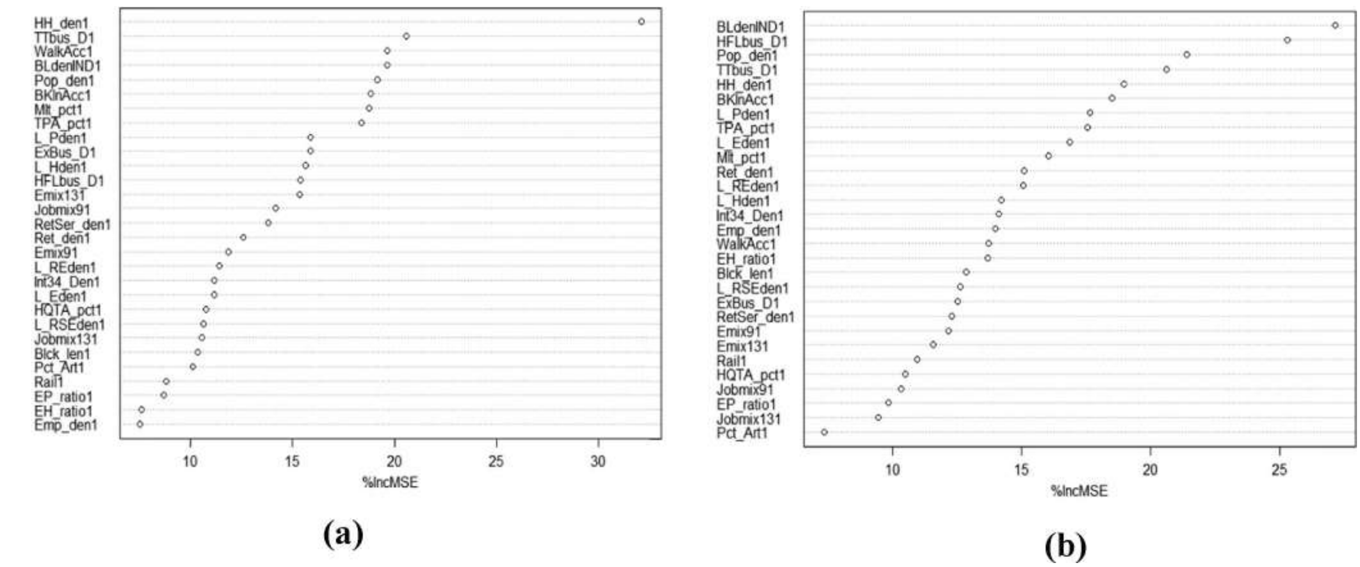

Fig. 3. A Variable Importance Plot for (a) Pedestrian-Involved Crash Counts (a) Bike-Involved Crash Counts. Note: "%IncMSE" represents the percentage of the drop of mean squared errors with certain variables being excluded from the model development.

Table 2
Descriptive Statistics of Collected Data.

Variables	Minimum	Maximum	Mean	S.D.
Pop_den1	0.00	76.86	22.13	12.74
HH_den1	0.00	30.59	7.56	5.03
Emp_den1	0.05	127.48	11.81	12.17
Ret_den1	0.00	7.11	1.12	1.03
RetSer_den1	0.02	26.62	3.84	4.13
Jobmix131	0.23	0.86	0.65	0.08
Jobmix91	0.25	0.79	0.61	0.08
Emix131	0.30	0.77	0.66	0.05
Emix91	0.14	0.55	0.41	0.06
EH_ratio1	0.00	10495.77	94.07	757.99
EP_ratio1	0.00	10495.55	85.17	817.95
Int34_Den1	0.01	0.58	0.20	0.06
BKlnAcc1	0.00	1.00	0.56	0.30
Rail1	0.00	0.72	0.08	0.14
ExBus_D1	0.00	0.71	0.03	0.05
HFLbus_D1	0.00	0.47	0.05	0.05
TTbus_D1	0.00	2.71	0.42	0.29
Mlt_pct1	0.00	0.49	0.28	0.12
HQTA_pct1	0.00	1.00	0.89	0.23
TPA_pct1	0.00	1.00	0.71	0.33
BLdenIND1	0.03	11.82	6.32	2.83
Blck_len1	0.11	1.12	0.24	0.08
WalkAcc1	0.01	53.41	6.94	6.75
Pct_Art1	0.00	0.27	0.04	0.05
L_Pden1	0.01	6.91	2.78	0.90
L_Hden1	0.01	6.91	1.84	0.97
L_Eden1	0.06	4.13	1.81	0.94
L_REden1	0.00	5.62	1.04	0.76
L_RSEden1	0.00	4.55	0.80	0.66
Ped	0.00	222.00	48.12	38.50
Bike	0.00	177.00	35.93	29.65
Distance of pairs of stops	0.00	156.97	20.90	17.61

Notes: 1. S.D. represents standard deviation. 2. Variables fed into final spatial model development are marked in a bold font.

Table 3 reveals posterior estimates of model parameters with and without correlation of pedestrian and bike. It can be observed that models with explicit consideration of the correlation of pedestrian and bike highlight more statistically significant covariates than the models without correlation being considered. For instance, the variables 'ExBus_D' (Stop density for Express Bus and BRT), 'WalkAcc1'(Walk Accessibility), and 'L_Pden1′ (Population density) for pedestrian and EH_ratio1 (Job/Household ratio) for the bike were found to be significant in models with correlation. Such finding indicates that consideration of correlation between the modes greatly impacts not only the coefficient magnitude but also the precision and the associated confidence level. Interestingly, only one variable, 'L_REden1′ (Retail job density), was found to have a statistically significant adverse impact across both modes, either with or without correlation. The possible explanation may be that the local area traffic management at retail areas reduces vehicle operating speeds and thereby decreases the likelihood of pedestrian and bicyclist collisions.

At the individual model level, three covariates which include 'HH_den1l' (Household density), 'TTbus_D1′ (Total bus stop density), and 'Emix131′ (Job mix 13 sectors), appear to have a statistically positive impact on the pedestrian crash counts in both cases (with and without correlation), indicating the propensity of pedestrians to be involved in collisions with the increase of the number of households (Noland & Quddus, 2004; Ponnaluri & Nagar, 2010; Sze et al., 2019), bus stops (Truong & Somenahalli, 2011; Quistberg et al., 2015; Craig et al., 2019), and jobs (Hess et al., 2004; Schneider et al., 2010; Miranda-Moreno et al., 2011; Lee et al., 2015; Nesoff et al., 2018). Likewise, for bikes, there are four statistically significant variables with consistent signs in both situations, which contain Blck_len1 (Estimated block length), 'ExBus_D1′ (Stop density for Express Bus and BRT), 'L_Pden1′ (Population density), and 'L_Eden1′ (Employment density). As expected, the latter three variables demonstrate positive coefficient values, suggesting that the increase in population, employment, and the number of express bus stops lead to more bicyclist-involved collisions. Interestingly, 'Blck_len1′ (Estimated block length) seems to have a statistically negative influence on bike-involved crashes. The phenomenon may be due to the longer reaction time and more environmental adaptability rendered to the bicyclists by the greater street block lengths.

As previously mentioned, this study employed both DIC and WAIC to assess the performance of different models from different perspectives. DIC, a penalized criterion, acts as a trade-off between model fit and model complexity, which are represented by its two components: posterior deviance ($\bar{D}$) and effectiveness number of parameters (P_D). WAIC, a fully Bayesian approach, was adopted as a cross-validation measure to assess the model performance from another angle. The models with comparatively smaller values

Table 3
Description of Model Parameter Estimates.

Variables	With Correlation of Ped and Bike				Without Correlation of Ped and Bike			
	Pedestrian		Bike		Pedestrian		Bike	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Model Parameters								
(Intercept)	**1.572**	**0.153**	**3.788**	**0.735**	**1.604**	**0.153**	**3.792**	**0.734**
HH_den1	**0.188**	**0.028**	NA	NA	**0.453**	**0.052**	NA	NA
Jobmix91	0.015	0.044	NA	NA	0.002	0.044	NA	NA
Emix131	**0.135**	**0.024**	NA	NA	**0.286**	**0.044**	NA	NA
Emix91	NA	NA	0.021	0.040	NA	NA	0.048	0.036
EH_ratio1	NA	NA	**−0.072**	**0.019**	NA	NA	0.029	0.032
BKlnAcc1	−0.031	0.044	0.034	0.040	−0.030	0.036	0.018	0.031
ExBus_D1	**0.234**	**0.045**	**0.270**	**0.040**	−0.007	0.042	**0.130**	**0.031**
TTbus_D1	**0.095**	**0.025**	NA	NA	**0.524**	**0.050**	NA	NA
BLdenIND1	NA	NA	**0.251**	**0.022**	NA	NA	**−0.684**	**0.038**
Blck_len1	NA	NA	**−0.086**	**0.026**	NA	NA	**−0.322**	**0.041**
WalkAcc1	**0.108**	**0.027**	NA	NA	−0.034	0.046	NA	NA
L_Pden1	**0.295**	**0.049**	**0.289**	**0.044**	0.073	0.045	**0.085**	**0.037**
L_Eden1	NA	NA	**0.100**	**0.033**	NA	NA	**0.089**	**0.043**
L_REden1	**−0.216**	**0.050**	**−0.286**	**0.050**	**−0.098**	**0.046**	**−0.139**	**0.047**

Notes: 1. The bold fonts represent the variables that are statistically significant at the significance level of 0.05.
2. S.D. represents standard deviation.
3. Refer to Table 1 for a detailed variable description.

Table 4
Good-of-fit criteria results.

Good-of-fit Criteria	With Correlation of Ped and Bike	Without Correlation of Ped and Bike
DIC	**8752.37**	9285.00
WAIC	**8706.51**	9186.79
$\bar{D}$	**7801.69**	8502.66
P_D	950.68	**782.34**

Note: The bold fonts indicate the best performance under specific criteria.

of DIC and WAIC indicate better performance (Gelman et al., 2013). The results for DIC and WAIC are illustrated in Table 4. It is obvious that the models with consideration of correlation of pedestrian and bike counts demonstrate the superior performance under DIC, WAIC, and $\bar{D}$. This superiority may be attributed to incorporating the correlation structure that explicitly allows the flexibility to capture spatial heterogeneity. However, the value of P_D (950.68) in the models with a correlation structure is substantially higher (with a difference of 168.34 points) than in the models without correlation. It suggests that the advantage associated with models with correlation is accompanied by the dramatic increase in the model complexity due to the inclusion of the correlation coefficient. Overall, the significantly better performance accompanied by models with the included correlation clearly justifies the benefits of addressing the correlation between the transportation modes.

Table 5 illustrates the marginal precision and correlation coefficient between walking and biking crash counts. As exhibited, the statistically positive correlation (see Eq. (4)) between the transportation modes ($\rho = 0.953$) was observed within the correlated effect models, which shows that the close spatial proximity (0.5-mile-radius zone) may be attributed to shared unobserved factors such as road surface type, lighting condition, day/night, and weather condition between pedestrian and bicyclist crashes. The results again corroborate the sensibility of using the bivariate spatial framework in this study.

6. Conclusions and recommendations

The main objective of this study was to quantify the impact of TOD influential variables on the pedestrian and bicyclist's traffic safety near those stops. For this purpose, joint models based on the multivariate conditional autoregressive (MCAR) priors with distance-oriented neighboring weight matrix were employed. First, R package INLAMSM was employed to take advantage of the approximate Bayesian inference within a bivariate spatial framework. Second, different covariates to each of the two active transportation modes were identified via the feature section using both random forest and correlation analyses. The mode-specific covariates used in the jointed models are anticipated to increase both model flexibility and accuracy. Third, the rarely collected transit stop-centered data allow the transportation practitioners to better understand the safety-pertinent factors at these important intermodal interfaces. Finally, evaluation criteria relying on both in-sample and out-sample errors clearly reveal the performance of models in both cases, or, with and without correlation between the modes being explicitly taken into consideration. Overall, the following conclusions are drawn:

1. The models with the correlation of pedestrians and bicyclists exhibited better results than those without the correlation being considered. Such finding highlights the importance of employing the bivariate spatial models, rather than two sepa-

Table 5
Marginal precision and correlation coefficient.

	With Correlation of Ped and Bike		Without Correlation of Ped and Bike	
	Pedestrian (S.D)	Bike (S.D)	Pedestrian (S.D)	Bike (S.D)
τ(Tau)	0.001 (4.2E−05)	0.002 (4.9E−05)	0.002 (4.2E−05)	0.003 (1.6-E04)
ρ(Rho)	0.953 (0.002)		–	

Note: 1. τ represents marginal precision of pedestrian and bike. 2. ρ illustrates the correlation coefficient between pedestrian and bike. 3. S.D shows the standard deviation.

rate univariate ones, to capture the unobserved heterogeneity shared by the crashes involving both modes. It is noteworthy the inclusion of the correlation parameter would substantially enhance the model complexity per the model results.

2. Regarding model parameters, retail job density appeared to have a statistically significant negative impact on both modes-related collisions in cases of both correlation treatments. The increase in households, jobs, and transit stops would lead to the rise of pedestrian-involved crashes may be due to the increase in pedestrian exposure. For bicyclists, the stop density for express bus and BRT, population density, and employment density seem to influence bicyclist-related crashes significantly. Interestingly, longer street block length was demonstrated to enhance bicyclist safety near the transit stops. The potential reason may be the longer reaction time and more surrounding adaptability rendered to the bicyclists by the greater street block lengths.

The findings mentioned above from this study reflect a better understanding of transit stop-related factors and their impacts on active transportation safety. However, it is important to be aware of some caveats. First, the current findings are based on the empirical results obtained from the bicycle and pedestrian crash data of Los Angeles County only. The superiority of specific models may not hold when employed at a different spatial level. Second, even though the present study focused on a set of influential factors to pedestrian and bicycle safety, some other factor attributes including road network configuration, accessible routes to public transit station, and traffic management and control on pedestrian and cyclist safety are also important to the active transportation safety, and shall be further explored in the future studies.. Third, the present study used the MCAR formulation to account for the spatial correlation problem. It is recommended to explore the other spatial formulations that might produce different findings from the present study. Fourth, other feature selection techniques may lead to different covariates used and hence the different coefficient values. Fifth, the zone with a fixed radius of 0.5 miles was utilized for model development and evaluation purposes. The zones with different radii or varying radii corresponding to zones' characteristics might generate different findings. Sixth, the present paper uses the typical variables such as population density and household density to represent the proxy exposure. It's preferred that household travel survey data and bike sharing data may be adopted to estimate the pedestrian and bicyclist exposure, as shown in more recent studies. Finally, the random effects were employed to address the unobserved heterogeneity. Future studies may adopt a more flexible approach of random parameters, which might lead to different parameter estimates from the current study.

Acknowledgements

This study was funded by the California State University Transportation Consortium research program under California's Senate Bill 1 (ZSB12017-SJAUX). We are grateful to the Southern California Association of Governments (SCAG) for providing various data surrounding the transit stops. The efforts and comments from the anonymous reviewers are significantly appreciated as well.

References

Abdel-Aty, M., & Haleem, K. (2011). Analyzing angle crashes at unsignalized intersections using machine learning techniques. *Accident Analysis & Prevention, 43*(1), 461–470.

Aguero-Valverde, J. (2013a). Full Bayes Poisson gamma, Poisson lognormal, and zero inflated random effects models: Comparing the precision of crash frequency estimates. *Accident Analysis & Prevention, 50*, 289–297.

Aguero-Valverde, J. (2013b). Multivariate spatial models of excess crash frequency at area level: Case of Costa Rica. *Accident Analysis & Prevention, 59*, 365–373.

Aguero-Valverde, J., & Jovanis, P. P. (2008). Analysis of road crash frequency with spatial models. *Transportation Research Record, 2061*(1), 55–63.

Aguero-Valverde, J., & Jovanis, P. P. (2009). Bayesian multivariate Poisson lognormal models for crash severity modeling and site ranking. *Transportation Research Record, 2136*(1), 82–91.

Anarkooli, A. J., Persaud, B., Hosseinpour, M., & Saleem, T. (2019). Comparison of univariate and two-stage approaches for estimating crash frequency by severity-case study for horizontal curves on two-lane rural roads. *Accident Analysis & Prevention, 129*, 382–389.

Anastasopoulos, P. C., & Mannering, F. L. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention, 41*(1), 153–159.

Barbour, E. (2016). Evaluating sustainability planning under California's senate bill 375. *Transportation Research Record: Journal of The Transportation Research Board, 2568*(1), 17–25.

Barua, S., El-Basyouny, K., & Islam, M. T. (2016). Multivariate random parameters collision count data models with spatial heterogeneity. *Analytic methods in accident research, 9*, 1–15.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological), 36*(2), 192–225.

Besharati, M. M., Kashani, A. T., Li, Z., Washington, S., & Prato, C. G. (2020). A bivariate random effects spatial model of traffic fatalities and injuries across provinces of Iran. *Accident Analysis & Prevention, 136*, 105394.

Bijleveld, F. D. (2005). The covariance between the number of accidents and the number of victims in multivariate analysis of accident related outcomes. *Accident Analysis & Prevention, 37*(4), 591–600.

Blangiardo, M., Cameletti, M., Baio, G., & Rue, H. (2013). Spatial and Spatio-Temporal Models with R-INLA. *Spatial and Spatio-Temporal Epidemiology, 4*, 33–49.

Branion-Calles, M., Teschke, K., Koehoorn, M., Espin-Garcia, O., & Harris, M. A. (2021). Estimating walking and bicycling in Canada and their road collision fatality risks: The need for a national household travel survey. *Preventive Medicine Reports, 22*, 101366.

Cai, Q., Abdel-Aty, M., & Lee, J. (2017). Macro-level vulnerable road users crash analysis: A Bayesian joint modeling approach of frequency and proportion. *Accident Analysis & Prevention, 107*, 11–19.

Cai, Q., Abdel-Aty, M., Lee, J., Wang, L., & Wang, X. (2018). Developing a grouped random parameters multivariate spatial model to explore zonal effects for segment and intersection crash modeling. *Analytic Methods in Accident Research, 19*, 1–15.

Chen, T., Sze, N. N., Chen, S., & Labi, S. (2020). Urban road space allocation incorporating the safety and construction cost impacts of lane and footpath widths. *Journal of Safety Research, 75*, 222–232.

Cheng, W., Gill, G. S., Dasu, M., & Jia, X. (2018). An empirical evaluation of multivariate spatial crash frequency models. *Accident Analysis & Prevention, 119*, 290–306.

Cheng, W., Gill, G. S., Sakrani, T., Dasu, M., & Zhou, J. (2017). Predicting motorcycle crash injury severity using weather data and alternative Bayesian multivariate crash frequency models. *Accident Analysis & Prevention, 108*, 172–180.

Chiou, Y. C., Fu, C., & Chih-Wei, H. (2014). Incorporating spatial dependence in simultaneously modeling crash frequency and severity. *Analytic Methods in Accident Research, 2*, 1–11.

Congdon, P. (2007). *Bayesian statistical modelling* (Vol. 704). John Wiley & Sons.

Conway, A., Cheng, J., Peters, D., & Lownes, N. (2013). Characteristics of multimodal conflicts in urban on-street bicycle lanes. *Transportation Research Record: Journal of The Transportation Research Board, 2387*(1), 93–101.

Craig, C. M., Morris, N. L., Van Houten, R., & Mayou, D. (2019). Pedestrian safety and driver yielding near public transit stops. *Transportation Research Record, 2673*(1), 514–523.

Cutler, A., Cutler, D., & Stevens, J. (2012). Random forests. *Ensemble Machine Learning*, 157–175.

Daniels, S., Brijs, T., Nuyts, E., & Wets, G. (2010). Explaining variation in safety performance of roundabouts. *Accident Analysis & Prevention, 42*(2), 393–402.

Ding, H., Sze, N. N., Li, H., & Guo, Y. (2020). Roles of infrastructure and land use in bicycle crash exposure and frequency: A case study using Greater London bike sharing data. *Accident Analysis & Prevention, 144*, 105652.

Dumbaugh, E., Li, W., & Joh, K. (2013). The built environment and the incidence of pedestrian and cyclist crashes. *Urban Design International, 18*(3), 217–228.

El-Basyouny, K., & Sayed, T. (2006). Comparison of two negative binomial regression techniques in developing accident prediction models. *Transportation Research Record, 1950*(1), 9–16.

Gelman, A., Hwang, J., & Vehtari, A. (2013). Understanding predictive information criteria for Bayesian models. *Statistics and Computing, 24*, 997–1016.

Hauer, E. (2001). Overdispersion in modeling accidents on road sections and in empirical Bayes estimation. *Accident Analysis and Prevention, 33*(6), 799–808.

Hausman, J.A., Hall, B.H., and Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship.

Hess, P. M., Moudon, A. V., & Matlick, J. M. (2004). Pedestrian safety and transit corridors. *Journal of Public Transportation, 7*(2), 5.

Heydari, S., Fu, L., Miranda-Moreno, L., & Jopseph, L. (2017). Using a flexible multivariate latent class approach to model correlated outcomes: A joint analysis of pedestrian and cyclist injuries. *Analytic Methods in Accident Research, 13*, 16–27.

Hosseinpour, M., Sahebi, S., Zamzuri, Z. H., Yahaya, A. S., & Ismail, N. (2018). Predicting crash frequency for multi-vehicle collision types using multivariate

Poisson-lognormal spatial model: A comparative analysis. *Accident Analysis & Prevention, 118*, 277–288.
Hou, Q., Tarko, A. P., & Meng, X. (2018). Investigating factors of crash frequency with random effects and random parameters models: New insights from Chinese freeway study. *Accident Analysis & Prevention, 120*, 1–12.
Huang, H., Zhou, H., Wang, J., Chang, F., & Ma, M. (2017). A multivariate spatial model of crash frequency by transportation modes for urban intersections. *Analytic Methods in Accident Research, 14*, 10–21.
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning, Vol. 112*, 18.
Imprialou, M. I. M., Quddus, M., Pitfield, D. E., & Lord, D. (2016). Re-visiting crash–speed relationships: A new perspective in crash modelling. *Accident Analysis & Prevention, 86*, 173–185.
Jiang, X., Abdel-Aty, M., Hu, J., & Lee, J. (2016). Investigating macro-level hotzone identification and variable importance using big data: A random forest models approach. *Neurocomputing, 181*, 53–63.
Jonathan, A. V., Wu, K. F. K., & Donnell, E. T. (2016). A multivariate spatial crash frequency model for identifying sites with promise based on crash types. *Accident Analysis & Prevention, 87*, 8–16.
Kim, H., Sun, D., & Tsutakawa, R. K. (2002). Lognormal vs. Gamma: Extra variations. *Biometrical Journal: Journal of Mathematical Methods in Biosciences, 44*(3), 305–323.
Lee, J., Abdel-Aty, M., & Jiang, X. (2015). Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. *Accident Analysis & Prevention, 78*, 146–154.
Lindgren, F., & Rue, H. (2015). Bayesian spatial modelling with R-INLA. *Journal of Statistical Software, 69*(19).
Lord, D., & Bonneson, J. A. (2007). Development of accident modification factors for rural frontage road segments in Texas. *Transportation Research Record, 2023*(1), 20–27.
Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research part A: Policy and Practice, 44*(5), 291–305.
Lord, D., & Miranda-Moreno, L. F. (2008). Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective. *Safety Science, 46*(5), 751–770.
Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to bayesian analysis*. CRC Press.
Ma, J., & Kockelman, K. (2006). Bayesian multivariate Poisson regression for models of injury count, by Severity. *Transportation Research Record: Journal of The Transportation Research Board, 1950*(1), 24–34.
Ma, X., Chen, S., & Chen, F. (2017). Multivariate space-time modeling of crash frequencies by injury severity levels. *Analytic Methods in Accident Research, 15*, 29–40.
Ma, J., Kockelman, K. M., & Damien, P. (2008). A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis & Prevention, 40*(3), 964–975.
Mader, E. M., & Zick, C. D. (2014). Active transportation: Do current traffic safety policies protect non-motorists? *Accident Analysis & Prevention, 67*, 7–13.
Malyshkina, N. V., & Mannering, F. L. (2010). Zero-state Markov switching count-data models: An empirical assessment. *Accident Analysis & Prevention, 42*(1), 122–130.
Mannering, F., & Bhat, C. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1–22.
Mardia, K. (1998). Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *Journal of Multivariate Analysis, 24*, 265–284.
Milton, J. C., Shankar, V. N., & Mannering, F. L. (2008). Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis & Prevention, 40*(1), 260–266.
Miranda-Moreno, L. F. (2006). *Statistical models and methods for the identification of hazardous locations for safety improvements (Doctoral dissertation)*. Department of Civil Engineering, University of Waterloo. Ph. D. Thesis.
Miranda-Moreno, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. *Accident Analysis & Prevention, 43*(5), 1624–1634.
Mothafer, G. I., Yamamoto, T., & Shankar, V. N. (2016). Evaluating crash type covariances and roadway geometric marginal effects using the multivariate Poisson gamma mixture model. *Analytic Methods in Accident Research, 9*, 16–26.
Narayanamoorthy, S., Paleti, R., & Bhat, C. R. (2013). On accommodating spatial dependence in bicycle and pedestrian injury counts by severity level. *Transportation research part B: methodological, 55*, 245–264.
Nesoff, E. D., Milam, A. J., Branas, C. C., Martins, S. S., Knowlton, A. R., & Furr-Holden, D. M. (2018). Alcohol outlets, neighborhood retail environments, and pedestrian injury risk. *Alcoholism: Clinical and Experimental Research, 42*(10), 1979–1987.
Noland, R. B., & Quddus, M. A. (2004). Analysis of pedestrian and bicycle casualties with regional panel data. *Transportation Research Record, 1897*(1), 28–33.
Oh, J., Washington, S. P., & Nam, D. (2006). Accident prediction model for railway-highway interfaces. *Accident Analysis & Prevention, 38*(2), 346–356.
Osama, A., & Sayed, T. (2017). Investigating the effect of spatial and mode correlations on active transportation safety modeling. *Analytic Methods in Accident Research, 16*, 60–74.
Palmi-Perales, F., Gomez-Rubio, V., & Martinez-Beneito, M. A. (2019). Bayesian multivariate spatial models for lattice data with INLA. arXiv preprint arXiv:1909.10804.
Park, E., & Lord, D. (2007). Multivariate poisson-lognormal models for jointly modeling crash frequency by Severity. *Transportation Research Record: Journal of The Transportation Research Board, 2019*(1), 1–6.
Plummer, M. (2008). Penalized loss functions for Bayesian model comparison. *Biostatistics, 9*(3), 523–539.
Ponnaluri, R. V., & Nagar, F. (2010). Road crash risk among vulnerable population groups in Andhra Pradesh, India (No. 10-0037).
Quddus, M. A. (2008). Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. *Accident Analysis & Prevention, 40*(4), 1486–1497.
Quistberg, D. A., Koepsell, T. D., Johnston, B. D., Boyle, L. N., Miranda, J. J., & Ebel, B. E. (2015). Bus stops and pedestrian–motor vehicle collisions in Lima, Peru: A matched case–control study. *Injury Prevention, 21*(e1), e15–e22.
Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology), 71*(2), 319–392.
Russo, B. J., Savolainen, P. T., Schneider, W. H., & Anastasopoulos, P. C. (2014). Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered Probit model. *Analytic Methods in Accident Research, 2*, 21–29.
SCAG, 2020, https://gisdata-scag.opendata.arcgis.com/datasets/SCAG::high-quality-transit-areas-hqta-2045-scag-region/about.
Schneider, R. J., Diogenes, M. C., Arnold, L. S., Attaset, V., Griswold, J., & Ragland, D. R. (2010). Association between roadway intersection characteristics and pedestrian crash risk in Alameda County, California. *Transportation Research Record, 2198*(1), 41–51.
Serhiyenko, V. (2015). Dynamic modeling of multivariate counts-Fitting, diagnostics, and applications. *Doctoral Dissertations, 858*.
Serhiyenko, V., Mamun, S. A., Ivan, J. N., & Ravishanker, N. (2016). Fast Bayesian inference for modeling multivariate crash counts. *Analytic Methods in Accident Research, 9*, 44–53.
Siddiqui, C., Abdel-Aty, M., & Huang, H. (2012). Aggregate nonparametric safety analysis of traffic zones. *Accident Analysis & Prevention, 45*, 317–325.
Smith, R., Reed, S., & Baker, S. (2010). Street Design: Part 1. Complete Streets. *Public Roads, 74*(1), 12–17.
Song, J. J., Ghosh, M., Miaou, S., & Mallick, B. (2006). Bayesian multivariate spatial models for roadway traffic crash mapping. *Journal of Multivariate Analysis, 97*(1), 246–273.
Soroori, E., Mohammadzadeh Moghaddam, A., & Salehi, M. (2019). Application of local conditional autoregressive models for development of zonal crash prediction models and identification of crash risk boundaries. *Transportmetrica A: Transport Science, 15*(2), 1102–1123.
Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2003). WinBUGS user manual.
Sze, N. N., Su, J., & Bai, L. (2019). Exposure to pedestrian crash based on household survey data: Effect of trip purpose. *Accident Analysis & Prevention, 128*, 17–24.
Truong, L. T., & Somenahalli, S. V. (2011). Using GIS to identify pedestrian-vehicle crash hot spots and unsafe bus stops. *Journal of Public Transportation, 14*(1), 6.
Van Der Linde, A. (2005). DIC in variable selection. *Statistica Neerlandica, 59*(1), 45–56.
Washington, S., Karlaftis, M. G., Mannering, F., & Anastasopoulos, P. (2020). *Statistical and econometric methods for transportation data analysis*. CRC Press.
Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research, 11*(Dec), 3571–3594.
Zeng, Q., Wen, H., Wong, S. C., Huang, H., Guo, Q., & Pei, X. (2020). Spatial joint analysis for zonal daytime and nighttime crash frequencies using a Bayesian bivariate conditional autoregressive model. *Journal of Transportation Safety & Security, 12*(4), 566–585.

Mankirat Singh, is a research associate in the Department of Civil Engineering at California State Polytechnic University, Pomona. His research focuses on highway safety and statistical modeling.

Yongping Zhang, Ph.D., P.E., is an assistant professor in the Department of Civil Engineering at California State Polytechnic University, Pomona. His research focuses on transportation planning and highway safety.

Wen Cheng, Ph.D., P.E., PTOE, is a professor in the Department of Civil Engineering at California State Polytechnic University, Pomona. He received a M.S. from University of Arizona, and a Ph.D. from Arizona State University. His research focuses on highway safety, statistical modeling, traffic signal design and traffic simulation.

Yihua Li, Ph.D., is an associate professor in the Department of Logistics Engineering, Logistics and Traffic College at Central South University of Forestry and Technology. He is a visiting scholar in Civil Engineering Department at California State Polytechnic University, Pomona. His research focuses on statistical modeling and highway safety.

Edward Clay, is a research assistant in the Department of Civil Engineering at California State Polytechnic University, Pomona. Her research focuses on highway safety and data mining.

Journal of Safety Research 83 (2022) 238–247

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Employer safety obligations, safety climate, and safety behaviors in the ready-made garment context in Bangladesh

Md. Shamsul Arefin [a,*], Ishita Roy [a], Swapna Chowdhury [b], Md. Shariful Alam [c]

[a] Department of Management Studies, Faculty of Business Studies, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj 8100, Bangladesh
[b] Department of Business Administration, University of Development Alternative, Dhaka, Bangladesh
[c] Faculty of Business and Economics, United International University, Dhaka, Bangladesh

ARTICLE INFO

Article history:
Received 20 May 2020
Received in revised form 26 April 2022
Accepted 30 August 2022
Available online 10 September 2022

Keywords:
Psychological contract
Safety attitudes
Prosocial safety behaviors
Proactive safety behaviors
Safety compliance behaviors

ABSTRACT

Purpose: The impact of employer safety obligations on safety climate and safety outcomes has become an important area of research in organizational and safety sciences. Evidence shows that employer safety obligations positively impact safety outcomes, including safety climate and safety behaviors. However, these relationships have not been thoroughly explored within the garment settings. This study is one of the first known studies to examine the relationships between employer safety obligations, safety climate, and safety behavior outcomes in a sample of garment employees. *Methods:* Two-wave time-lagged data were collected from 347 garment employees and their supervisors in Bangladesh. Hierarchical regression analysis was applied to examine hypothesized relationships using AMOS a SPSS. *Results:* Employer safety obligations positively influenced safety climate perceptions among garment employees. Safety climate perceptions are positively and significantly associated with safety behaviors, including safety compliance behaviors, prosocial safety behaviors, and proactive safety behaviors. Moreover, the safety climate mediates the influence of employer safety obligations on safety behaviors. *Conclusions:* These findings provide important evidence of the relationships between employer safety obligations, safety climate, and safety behaviors in the garment industry of Bangladesh. *Practical Applications:* Ultimately, these findings guide the government, garment manufacturers, and managers to bolster garment employees' safety outcomes.

1. Introduction

Workplace safety is an integral part of the national economy, organizations, and society. Organizations strive to ensure safety issues to maintain a safe workplace and achieve sustained competitive advantage in an industry. Although previous research examined safety in different industries, there has been a lack of apparel-related research on safety, which has led to more severe accidents and injuries in the garment industry in recent years. The collapse of the Rana Plaza building, the Tazreen Fashion factory fire in Bangladesh, and the Ali Enterprises factory fire in Pakistan accounted for three out of the seven world's deadliest industrial disasters in the last 20 years (McClure, 2018). The Rana Plaza building collapse in Bangladesh killed at least 1,132 people and injured more than 2500 people (ILO, 2018). A few months before the Rana plaza incident, at least 112 workers had lost their lives in another tragic accident in Bangladesh, trapped inside the burning Tazreen Fashions factory (ILO, 2018). Since the Rana Plaza tragedy, at least 109 accidents have occurred in Bangladesh. Among these, at least 35 were garment and textile industry incidents in which 491 workers were injured and 27 were killed (ILO, 2018). Therefore, the garment industry is considered the deadliest industry in Bangladesh; fire and other health and safety incidents caused 1,303 deaths and 3,875 injuries from 2012 to 2018 (Solidarity Center, 2018). These statistics accentuate the extreme importance and urgency of reducing garment accidents and improving the safety performance of garment employees.

Three common causes of accidents are attributed such as technical issues, organizational issues, and human factors. Despite the development and implementation of modern technology in improving safety management systems, human casualties and injuries are still mounting in the garment setting due to organizational and human factors. Previous studies on safety-related behaviors mainly focused on the influence of organizational factors and leadership (Smith et al., 2016; Chen & Chen, 2014; Clarke, 2013); however, the psychological contract from employers' per-

* Corresponding author.
E-mail addresses: shamsul.arefin@bsmrstu.edu.bd (M.S. Arefin), sharifdu@bus.uiu.ac.bd (M.S. Alam).

https://doi.org/10.1016/j.jsr.2022.08.020
0022-4375/

spective (i.e., employer safety obligations) on individuals is still rare. Considering that employees hold beliefs about safety obligations in their workplace (Walker, 2010), exploring further how these beliefs transfer into employees' safety-related attitudes and behaviors within the garment settings is necessary. When observing the fulfillment of safety obligations by the employer, individual employees may be more likely to comply with safety rules and show safety citizenship behaviors.

Although a plethora of research identified the outcomes of a safety climate, little attention has been paid to the reciprocity between employer and employees, predicting an unthreatened environment in the organization (Mullen et al., 2017). In the social exchange theory, Blau (1964) suggests that mutual exchange happens between the two parties when one individual does something in exchange for the action of other individuals or entities. More specifically, employers are obliged to formulate and implement safety policies and procedures to shape a safety climate – and in response, employees are more likely to exhibit safety behaviors. This reciprocity leads to psychological contracts (Rousseau, 1990) that are stemmed from social exchange theory (Blau, 1964) and the norm of reciprocity (Gouldner, 1960). In the realm of safety, psychological contracts are individual beliefs in reciprocal safety obligations between employees and employers viewed from the employee's perspective (Rousseau, 1990). An employer's obligations to safety represent fulfilling implicit or explicit promises (Walker, 2010). Perceived safety obligations are defined as employees' beliefs regarding organizational safety activities stemming from the influence of society and organizations, which have become a significant predictor of organizational safety behavior in recent years (Mullen et al., 2017). Prior research shows that employees bear the beliefs of safety obligations in their working environment (Walker, 2010; Walker, 2013). Walker and Hutton (2006) established that individuals show safety behaviors in response to the organizational applications of safety responsibilities. This study focused on how employer safety obligations impact employees' in-role and extra-role safety behaviors—namely, safety compliance behaviors, prosocial safety behaviors, and proactive safety behaviors. Safety compliance behaviors are in-role task behaviors that are somewhat mandatory to perform, whereas prosocial and proactive safety behaviors are extra-role behaviors that are volitional. Therefore, the first goal of this study is to identify the influence of employer safety obligations on employee safety behaviors.

Although few researchers have shown a keen interest in revealing the influence of psychological contracts on safety climate in recent years (e.g., Walker, 2013; Newaz et al., 2019b), not much is known about the ready-made garment industries. Safety climate is the shared perception of an organization's safety-related policies, procedures, and practices (Griffin & Neal, 2000). This study assumes that employer safety obligations influence building safety climate in the garment organization. Employees expect employers to provide necessary training to operate equipment, avoid accidents, and respond to an emergency (Mullen et al., 2017). In addition, employers are expected to offer proper training for coworkers and monitor the safety practices of coworkers to prevent the organization from safety violations (Walker, 2010). Employees want to ensure that the equipment is well-maintained and operating correctly as a safety precaution. Accordingly, employees' safety expectations induce employers to shape a favorable safety climate. The previous study found the positive influence of the psychological contract of safety on safety climate (e.g., Walker, 2013; Newaz et al., 2019b). Thus, the second goal of this study is to examine the impact of perceived employer safety obligations on garment employees' safety climate attitudes.

The safety climate development process will be more accurately understood when its connection to employer safety obligations is investigated. Therefore, this study posits that employer safety obligations in the ready-made garment industry will be predicted by their employees' safety perceptions (i.e., safety climate), which ultimately affects employee compliance and their prosocial and proactive safety behaviors. Evidence shows that safety behaviors are related to safety climate (Clarke, 2006; Liu et al., 2019). Moreover, the safety climate developmental process is checked by investigating the mutual relationship between employers and employees. Therefore, the employer safety obligations model is implemented to clarify the effect of perceived safety obligations on a safety climate in the ready-made garment industry. Hence, the third goal of this paper is to investigate whether employer safety obligations impact garment employees' perception of safety climate, which, in turn, influences their safety compliance and safety citizenship behaviors.

The study moves forward previous research in several ways. First, we empirically investigated the consequences of employer safety obligations on two important employee safety behaviors: safety compliance and citizenship behaviors. Little is known about the relationship between employer safety obligations and employee prosocial and proactive safety behaviors. Second, this study investigates how employer safety obligations impact employee safety behaviors by examining the mediating role of employees' perceived safety climate. Previous studies largely ignored the intervening mechanism through with employer safety obligations affect employee behaviors. Finally, no empirical research was found reflecting the employee safety attitudes and behaviors in the garment industry. Moreover, the role of employer safety obligations in developing a safety climate is articulated in a few studies—mostly in Western settings. Therefore, further research is warranted to validate the association in the garment industry in a South Asian country, Bangladesh.

2. Literature review

2.1. Perceived employer safety obligations and safety behaviors

The safety literature reveals that employees show favorable attitudes and behaviors in response to the social responses in part from their organizations. Based on the social exchange theory (Blau, 1964), psychological contract theory (Rousseau, 1990) postulates that in reaction to organizational transactional (e.g., safety resource) and relational (e.g., safety commitment) influences, favorable safety attitudes and behaviors are also demonstrated (Walker, 2010). Over their employment period, employees build their perceptions about the safety obligations of both employer and employee (Walker & Hutton, 2006). Based on their stay in the organization, employees make some expectations regarding safety in their workplace that help them create perceptions of safety-related obligations that demand reciprocal actions.

When it comes to expectations, garment employees focus on having the essential skills to safely operate machinery and equipment and prevent accidents from occurring (Mullen et al., 2017). Employees frequently complain about a lack of training in identifying potential hazards and responding to an emergency. Furthermore, if they are doing an unsafe job, they want their company to stop them and safeguard them from harming their health and safety (Walker, 2013). Although ensuring safety is a shared responsibility, even a single employee's disregard for safety might endanger the well-being of their coworkers and the entire organization (Walker, 2010). Employees want to be assured that no one is allowed to engage in dangerous tasks without the proper training. Employees try to protect themselves from the potential unsafe behaviors of their coworkers. In order to prevent the company from safety violations, employees expect that their employer will provide adequate training to coworkers and keep an eye on their coworkers' safety practices (Mullen et al., 2017). A well-maintained and working piece of machinery can help protect employees from the risk of equipment failures. In general, garment organizations are concerned with ensuring that safety rules and initiatives are appropriately implemented.

These safety expectations constitute psychological contracts of safety. Employees think that the organization will look after their safety and establish appropriate safety policies regarding their welfare. When employees observe that their organization fulfills safety-related obligations and transactional responsibilities, such as providing regular safety training and the necessary resources to cope with the safety measures, it signals to employees that their safety and well-being are valued within the organization (Mullen et al., 2017; Walker & Hutton, 2006). Conversely, there may be a feeling of less fulfillment of obligations if the obligations are not periodically met. If it continues over a long period, employees may perceive it as a breach of contract (Walker, 2013), resulting in dissatisfaction among the employees and provoking them to demonstrate negligence in their safety behaviors (Walker, 2013). If employers comply with safety obligations, it indicates to the employees that their safety concerns are highly emphasized in the organization. In response, employees feel obligated to demonstrate safe behaviors (Hofmann & Morgeson, 1999; Kath et al., 2010). Several meta-analyses of safety ensure the positive association between organizational safety practices and employee safety behaviors, thus supporting the concept of employer-employee reciprocity of safety (Clarke, 2006; Liu et al., 2019).

Employer safety obligations are based on the social exchange; therefore, there are perceived obligations on the part of employees to reciprocate their responsibility by doing something that benefits employers (Blau, 1964; Gouldner, 1960). This study proposes a direct relationship between employer safety obligations and two types of employee safety behaviors – namely, safety compliance and safety-related organizational citizenship behaviors (OCBs). Safety compliance represents safety performance related to mandatory safety-related tasks, such as organizational safety processes and relevant safety procedures. On the contrary, safety citizenship behavior represents the voluntary safety-related behaviors directed toward the organization and its employees, such as helping employees cope with safety issues and providing safety-related suggestions. Safety citizenship behaviors are the discretionary behaviors that benefit the organization, but the formal reward system does not recognize these behaviors (Organ et al., 2006).

OCBs are a comparatively less developed research domain within the field of safety and are primarily tested as a single construct for safety citizenship research (Curcuruto & Griffin, 2018; Curcuruto et al., 2015; Curcuruto et al., 2019). The most common categories of OCBs are prosocial and proactive (van Dyne & LePine, 1998; Curcuruto & Griffin, 2018). Prosocial behaviors focus on affiliation with other individuals and are exhibited through cooperating with colleagues and seeking their welfare (Curcuruto & Griffin, 2018). They concentrate mainly on the safety concern related to group members and the establishment of the interpersonal relationship among its members. On the other hand, proactive behaviors are inherently challenging, and they try to make beneficial changes in work policies (Curcuruto et al., 2019; Mei et al., 2018). These behaviors concentrate less on interpersonal relations and more on changing the existing practices (Curcuruto et al., 2016). Although these behaviors are discretionary, they are separate from each other. Prosocial behaviors concentrate on collaboration and proactive behaviors, focusing on challenges. Following the study by Curcuruto et al. (2015), this research considers initiating change as proactive behavior and helping others as prosocial behavior. These categories are reflected in the overall OCB literature and indicate that safety behaviors can also be divided in order to examine their differing impacts on outcomes (Reader et al., 2017). This notion is supported by recent research that demonstrates that these behaviors work with safety procedures differently. Employees may reciprocate their responsibilities by expanding their obligatory roles consistent with the type of behavior valued in their work environment. Specifically, in response to fulfilling safety obligations, employees payback their employers by engaging in safety citizenship behaviors that benefit the employer and others in the work setting. Given this evidence, we assume that perceived employer safety obligations are related to compliance, prosocial, and proactive safety behaviors. Hence, the following hypotheses are drawn.

Hypothesis 1a. There is a positive relationship between perceived employer safety obligations and employee safety compliance behaviors.

Hypothesis 1b. There is a positive relationship between perceived employer safety obligations and employee prosocial safety behaviors.

Hypothesis 1c. There is a positive relationship between perceived employer safety obligations and employee proactive safety behaviors.

2.2. Safety climate and safety behaviors

Safety climate is the shared perception of safety affairs related to risky operations (Zohar, 2000). Like the organizational climate, safety climate perception instills behaviors that the organization rewards and supports. Zohar (1980) defined an organizational safety climate as 'a summary of molar perceptions that employees share about their work environments ... a frame of reference for guiding appropriate and adaptive task behaviors' (p. 96). In a broader sense, safety climate covers the manifestation of ongoing safety culture and management's attitude towards safety (Saedi et al., 2020; Mullen et al., 2017). The focus on merely counting the number of mishaps has been transformed into predicting and controlling the safety scenario to reduce employee accidents and injuries (Mearns et al., 2003). A strong safety climate in an organization leads to fewer occurrences of individual unsafe behaviors that cause accidents and injuries (Lee et al., 2019). When operationalizing the safety climate, the shared perception of employees on the relative priority of safety over other priorities is represented. In measuring their safety climate, employees express their perception of their supervisor's priority in safety matters, while other competing demands related to achieving the mission must be completed (Zohar & Luria, 2004).

Researchers used different constructs to assess safety climate, such as supervisor support, safety communication, safety motivation, safety knowledge, safety policy, safety training, safety attitude and commitment, safety priority, safety budget, coworker support, and safety awareness (e.g., Zohar & Luria, 2005; Liu et al., 2015; Olsen, 2010; O'Connor et al., 2011; Brondino et al., 2012; Huang et al., 2016; Lu & Yang, 2011). These constructs vary from industry to industry (Glendon & Litherland, 2001), thus requiring specifying the relevant constructs that represent the safety climate of the garment industry. Extant research has been conducted to explore the safety climate in different sectors such as nuclear power operations (e.g., Morrow et al., 2014), manufacturing (e.g., Zohar, 2000, 2002; Griffin & Neal, 2000; Liu et al., 2015), construction (e.g., Huang et al., 2006), passenger ferry (e.g., Lu & Yang, 2011), transport (e.g., Fugas et al., 2012), and hospitals (e.g., Neal et al., 2000). However, there remains a gap in the relevant safety climate scale for the garments industry. After reviewing the literature on safety climate constructs in the different industrial sectors and the garment industry's safety practices, this study includes management values, communication, training, and systems. Unlike breaking down the safety climate into several components, this research considers safety climate as a single construct to examine its relationship with the study variables (Neal et al., 2000; Lu & Tsai, 2010; Lu & Yang, 2011). The employees' positive perception of the safety climate induces them to engage in safety compliance, proactive safety, and prosocial behaviors (Curcuruto & Griffin, 2018; Jiang et al., 2019). Several meta-analyses established the positive influence of safety climate on employee safety compliance and safety citizenship behaviors (e.g., Clarke, 2006; Liu et al., 2019). Taken together, we posit the following hypotheses.

Hypothesis 2a. Perceived safety climate is positively related to employee safety compliance behaviors.

Hypothesis 2b. Perceived safety climate is positively related to employee prosocial safety behaviors.

Hypothesis 2c. Perceived safety climate is positively related to employee proactive safety behaviors.

2.3. Perceived employer safety obligations and safety climate

As discussed earlier, when employees experience any safety-related social influence within the organization, they are more likely to reciprocate the benefits through positive safety attitudes and behaviors (Hofmann et al., 2003; Hofmann & Morgeson, 1999). In response to fulfilling an employer's safety-related transactional and relational influences, employees develop a positive attitude toward the organizational safety climate. The felt obligation induces favorable safety culture in the garment organizations. Employees expect their company to provide adequate training to handle the hazardous elements of their jobs and cope with emergencies. Employees place a high value on the ability to safely operate machinery and equipment and the ability to prevent accidents. Employees are constantly on the lookout for ways to keep themselves and their coworkers safe. Employees want the employer to provide appropriate training to their coworkers and monitor their safety behaviors to protect the organization from violation of safety procedures. Employees want to make certain safety policies and procedures are implemented correctly. On a sample of healthcare workers in Australia, Walker (2013) found that employer breach of safety obligations negatively influences employees' safety climate attitudes and employee fulfillment of safety obligations positively influences safety climate attitudes. Similarly, in a study on construction employees in Australia, Newaz et al. (2019b) found that the psychological contract of safety positively influences different components of safety climate. In line with the findings and arguments, we assume that employee safety obligations positively impact employee safety climate attitudes. Thus, the following hypothesis is predicted.

Hypothesis 3. Perceived employer safety obligations are positively related to employee safety climate attitudes.

2.4. Safety climate, perceived employer safety obligations, and safety behaviors

In line with the previous discussion, a comprehensive investigation would explain how fulfilling employer safety obligations impact employees' safety climate attitudes and, ultimately, their safety behaviors (Walker, 2013; Newaz et al., 2019a; Mullen et al., 2017). Based on psychological contract theory (Rousseau, 1990) and social exchange theory (Blau, 1964), this study further assumes that employer safety obligations positively influence safety climate, affecting employee safety behaviors. More specifically, when employer safety obligations are met, employees reciprocate positive attitudes toward safety policies and practices in their organization, leading them to engage in certain safety-related behaviors. Therefore, the following hypothesis is drawn. The proposed model is depicted in Fig. 1.

Hypothesis 4a. Employee perceived safety climate mediates the positive relationship between perceived employer safety obligations and employee safety compliance behaviors.

Hypothesis 4b. Employee perceived safety climate mediates the positive relationship between perceived employer safety obligations and employee prosocial safety behaviors.

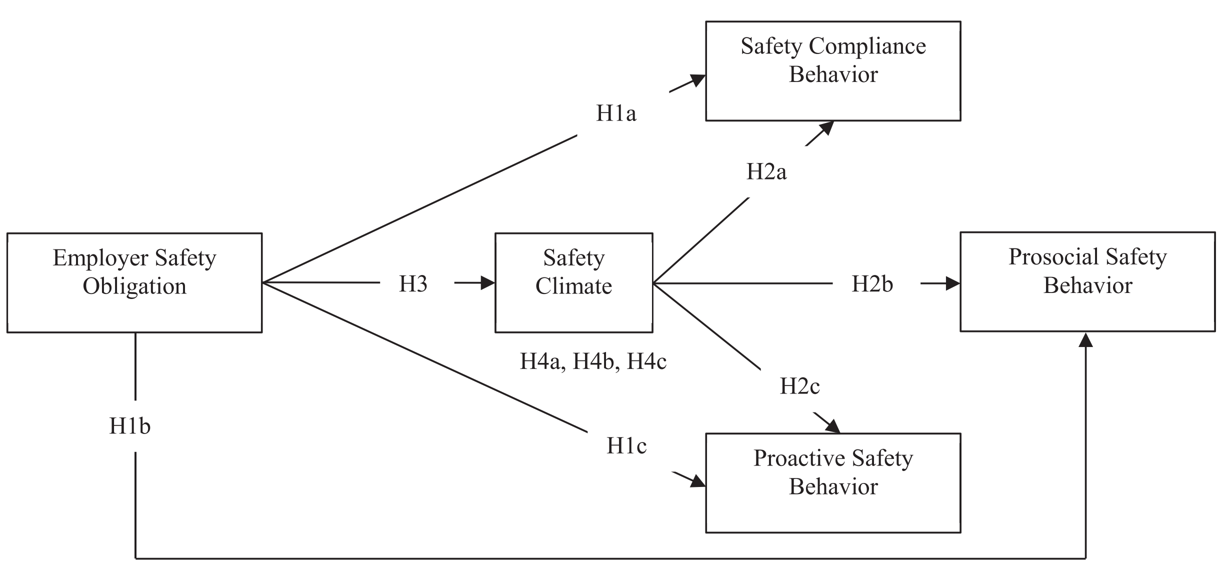

Fig. 1. The proposed model.

Hypothesis 4c. Employee perceived safety climate mediates the positive relationship between perceived employer safety obligations and employee proactive safety behaviors.

3. Methodology

3.1. Participants and procedures

The respondents were full-time employees from a large ready-made garment organization located in the Gazipur district in Bangladesh. We contacted the human resource manager for consent to the survey. After getting approval, the human resource manager is requested to provide a list of full-time employees. We randomly chose the employees from the list and distributed the questionnaire. We employed a back-translation method to convert English to Bengali because all of the items used in this study were initially written in English (Brislin, 1980). Two sets of questionnaires were developed for employees and their immediate supervisors. To ensure the face and content validity, we approached two professors who taught management and four senior safety officers to adjust the items with the work practices and culture in the study context. Further, a pilot study was conducted to ensure the content validity of the items, and a group of garment employees (n = 32) and their supervisors (n = 8) at a garment organization in the Gazipur district were surveyed. The alpha reliability of all the constructs was tested, and found the reliability ranged between 0.73 and 0.87. The final questionnaire was delivered face-to-face during their working hours.

We conducted our surveys in two phases with a lag of four weeks. Collecting data from different periods might prevent possible common method bias problems (Podsakoff et al., 2003), and this method supports the proposition that the perceptions of employer safety obligations might affect perceived safety climate and employee safety behavior. Data with different time lags were designed to decrease the participants' fatigue and promote the temporal separation among the study variables, such as independent and mediating variables. Participation in the survey was voluntary. At Phase 1, the structured questionnaire was distributed to 550 respondents. A cover letter addressing the study's purpose and assuring the confidentiality of their response was mentioned in the questionnaire. The respondents were requested to submit their completed questionnaires the following day. The questionnaire contained the independent variable that includes the perception of employer safety obligations. A total of 378 employees (68.73%) completed the survey. Four weeks later, all initial participants were surveyed. The second phase questionnaire contained the mediating variable that included employee-perceived safety climate. In Phase 2, their immediate supervisors were also surveyed. Codes were allocated to ensure the identification of the dyadic relationship and anonymity. The questionnaire completed by the supervisors included safety compliance behaviors, prosocial safety behaviors, and proactive safety behaviors. Overall, dyadic responses were received from 358 employees (65.10%). A sample of 347 employees (63.10%) was obtained after eliminating 11 incomplete questionnaires. Due to the fact that the respondents were surveyed during their working time, a high response rate was achieved. Most of the employees were female (75.8%), and the organization's average employee tenure was approximately-three years. The average age of the employees was about 24; around 68% of the participants received a high school education (HSC) or below. A total of 78 supervisors responded. Most of the supervisors were male (62.4%), and the organization's average supervisor tenure was approximately-seven years. The supervisors' average age was about 39; around 57% of the participants received a Bachelor's degree or above.

3.2. Measures

All the measures used in this study were collected from the established literature.

3.2.1. Perceived employer safety obligations

An 11-item measure, originated by Mullen, Kelloway, and Teed (2017), of perceived employer safety obligations was used in this study. These items were developed from the work of Walker (2010) and focused on the transactional (as opposed to relational) employer obligations (Walker, 2010). Table 1 shows the items of perceived employer safety obligations. A five-point Likert scale was used from 1 (strongly disagree) to 5 (strongly agree). The reliability alpha was 0.94.

3.2.2. Safety climate

A 10-item measure of safety climate was used in this study, developed by previous studies (i.e., Neal et al., 2000; Lu & Tsai, 2010; Lu & Yang, 2011). This scale measured the perception of safety climate, focusing on management values, communication, training, and safety systems. Exploratory principal axis factor analysis resulted in the extraction of one factor accounting for 64% of item variance. All items loaded substantially (>0.70) on the factor. A sample item reads, 'My company has assigned safety issues as a top priority.' A five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) was used to measure the safety climate. Cronbach's alpha was 0.96.

3.2.3. Safety compliance behavior

The supervisors completed a three-item scale developed by Neal and Griffin (2006) was used to measure safety compliance. The sample items include 'This employee use all the necessary safety equipment to do his/her job.' A five-point Likert scale ranging from 1 (never) to 5 (frequently) was used to measure this behavior, and the reliability coefficient was 0.86.

3.2.4. Proactive safety behavior

The supervisors completed a four-item scale developed by Hofmann et al. (2003) to measure proactive safety behavior. This study established the initiation of safety-related change behavior as one of the proactive safety behaviors. One sample item was 'This employee tries to improve safety procedures.' A five-point Likert scale ranging from 1 (never) to 5 (frequently) was used to measure this behavior. Cronbach's alpha was 0.84.

3.2.5. Prosocial safety behavior

The supervisors completed a six-item scale developed by Hofmann et al. (2003), which was used to measure voice behavior as a prosocial safety behavior. One sample item was 'This employee assists others to make sure they perform their work safely.' A five-

Table 1
Items of perceived employer safety obligations.

1. Provided me with safety training
2. Showed me how to prevent accidents
3. Pointed out aspects of the job that could potentially harm me
4. Taught me how to respond to emergency situations
5. Prevented me from carrying out potentially dangerous work
6. Prevented me from performing a task that I have not been properly trained to do
7. Taught me how to properly use equipment and machinery
8. Ensured that my coworkers were properly trained before performing a job
9. Monitored the safety behavior of my coworkers to ensure they do not injure someone
10. Implemented safety policies and practices
11. Ensured the equipment is maintained and properly functioning

point Likert scale ranging from 1 (never) to 5 (frequently) was used to measure this behavior. The reliability coefficient of this measure was 0.94.

3.2.6. Control variables

This study controlled age, experience, gender, and education (Wang et al., 2019). Gender was measured using a dummy variable (0 for male and 1 for female), and the educational level was measured using a scale of 1 through to 5 (1 for 'below secondary school certificate;' 2 for 'secondary school certificate;' 3 for 'higher secondary school certificate;' 4 for 'undergraduate;' 5 for 'postgraduate'). Age and experience were both measured in the number of years.

4. Results

4.1. Preliminary analysis

The descriptive statistics are reported in Table 2. The study variables were significantly related to each other. All the correlation values are below the threshold value of 0.9, which indicates that the multicollinearity problem is not a concern in this study (Hair et al., 2006). We conducted the confirmatory factor analysis (CFA) using AMOS v23 statistical software. The maximum likelihood estimation was taken. Several model fit indices were evaluated, such as comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). The result of the CFA showed a five-factor model having a sound model fit ($\chi2 = 2835.961$, $df = 488$, $\chi2/df = 1.713$, CFI = 0.970, TLI = 0.966, RMSEA = 0.045, SRMR = 0.046).

Table 3 showed the factor loadings of all five latent factors to be significant and above the threshold value of 0.7. Items six and eight of employer safety obligations were not above them and were thus dropped from the analysis (Tabachnick et al., 2007). We further tested the convergent and discriminant validity of the latent variables. Convergent validity is about how much of a given construct's indicators converge or have a high share of variance in common (Hair et al., 2014). We estimated factor loadings and average variance extracted (AVE) to identify the convergent validity. Discriminant validity refers to how different a latent factor is from other latent factors (Hair et al., 2014). The square root of AVE was used to check the discriminant validity. Table 2 shows that all the AVE values are above the threshold value of 0.5 (Fornell & Larcker, 1981). Table 2 also revealed that the square root of AVE that represents in the diagonals was greater than the respective correlation values displayed in the rows and columns – which determines discriminant validity. Thus, the data of this study confirmed the reliability, convergent, and discriminant validity issues.

Table 3
Factor loadings of the items.

	ESO	SC	SPA	SCB	SPS
ESO1	0.706				
ESO2	0.831				
ESO3	0.879				
ESO4	0.778				
ESO5	0.837				
ESO6	0.651				
ESO7	0.717				
ESO8	0.549				
ESO9	0.884				
ESO10	0.855				
ESO11	0.743				
SC1		0.774			
SC2		0.885			
SC3		0.858			
SC4		0.851			
SC5		0.743			
SC6		0.759			
SC7		0.768			
SC8		0.772			
SC9		0.904			
SC10		0.881			
SPA1			0.750		
SPA2			0.753		
SPA3			0.767		
SPA4			0.773		
SCB1				0.837	
SCB2				0.832	
SCB3				0.790	
SPS1					0.824
SPS2					0.886
SPS3					0.879
SPS4					0.905
SPS5					0.828
SPS6					0.731

Notes: SC = Safety Climate, ESO = Employer Safety Obligation, SPA = Safety Proactive Behavior, SCB = Safety Compliance Behavior, SPS = Safety Prosocial Behavior.

4.2. Hypothesis testing

This study conducted a hierarchical regression analysis to test the hypotheses. We predicted that employer safety obligations influenced safety compliance behaviors, prosocial behaviors, and proactive behaviors in hypotheses 1a, 1b, and 1c, respectively. As per our assumption, Table 4, Model 2 showed that employer safety obligations positively influenced safety compliance behaviors ($\beta = 0.41$, SE = 0.05, $p < 0.01$). Thus Hypothesis 1a was supported. As shown in Table 5, Model 5, proactive safety behaviors predicted perceived employer safety obligations ($\beta = 0.57$, SE = 0.04, $p < 0.01$). Therefore, Hypothesis 1b was also supported. Table 5, Model 8 showed that perceived employer safety obligations influenced prosocial safety behaviors ($\beta = 0.55$, SE = 0.05, $p < 0.01$). Hence, Hypothesis 1c was accepted.

Table 2
Descriptive statistics, CR, AVE, and correlation.

Variables	Mean	SD	Alpha	CR	AVE	SC	ESO	SPA	SCB	SPS
SC	3.57	0.75	0.96	0.95	0.68	**0.82**				
ESO	3.77	0.61	0.94	0.94	0.61	0.42	**0.78**			
SPA	3.99	0.54	0.84	0.85	0.58	0.49	0.59	**0.76**		
SCB	3.91	0.58	0.86	0.86	0.67	0.45	0.46	0.50	**0.82**	
SPS	4.00	0.62	0.94	0.94	0.71	0.47	0.59	0.58	0.61	**0.84**
Age	23.84	3.94								
Gender[a]	1.24	0.43								
Education level	2.23	0.81								
Organization tenure	5.16	3.88								

Note: SC = Safety Climate, ESO = Employer Safety Obligation, SPA = Safety Proactive Behavior, SCB = Safety Compliance Behavior, SPS = Safety Prosocial Behavior, CR = Composite Reliability, AVE = Average Variance Extracted, Values in diagonal represents the square root of AVE.
[a] Male = 0, Female = 1.

Table 4
Hierarchical regression effect of safety climate and safety compliance behavior.

Variables	SC	Safety Compliance Behavior		
	Model 1	Model 2	Model 3	Model 4
Control variables				
Age	−0.03(0.04)	−0.05(0.03)	−0.045(0.03)	−0.04(0.03)
Gender[a]	0.11(0.09)	0.03(0.07)	0.00(0.07)	0.00(0.06)
Education level	0.06(0.05)	0.00(0.04)	−0.03(0.04)	−0.01(0.03)
Organization tenure	0.05(0.04)	0.06(0.03)	0.05(0.03)	0.05(0.03)
Independent Variables				
Employer Safety Obligation	0.51**(0.06)	0.41**(0.05)		0.30**(0.05)
Safety Climate (SC)			0.32**(0.04)	0.22**(0.04)
F	15.58**	16.71**	15.10**	20.08**
R^2	0.20	0.20	0.20	0.26
Adjusted R^2	0.17	0.19	0.17	0.25
ΔR^2				0.06

Notes: Values in parentheses are standard errors; entries are unstandardized coefficients.
** $p < 0.01$.
[a] Male = 0, Female = 1.

Table 5
Hierarchical regression effect of Proactive safety behavior and prosocial safety behavior.

Variables	Proactive Safety Behavior			Prosocial Safety Behavior		
	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Control variables						
Age	0.00(0.02)	0.00(0.03)	0.01(0.02)	0.01(0.03)	0.02(0.03)	0.02(0.03)
Gender[a]	0.07(0.05)	0.05(0.06)	0.05(0.05)	0.13(0.07)	0.10(0.07)	0.10(0.06)
Education level	0.02(0.03)	−0.02(0.03)	0.01(0.03)	0.00(0.04)	−0.04(0.04)	−0.01(0.03)
Organization tenure	0.01(0.03)	0.00(0.03)	0.00(0.02)	−0.00(0.03)	−0.01(0.03)	−0.01(0.03)
Independent Variables						
Employer Safety Obligation	0.57**(0.04)		0.50**(0.04)	0.55**(0.05)		0.44**(0.05)
Safety Climate (SC)		0.31**(0.04)	0.14**(0.03)		0.36**(0.04)	0.22**(0.04)
F	50.24**	16.66**	47.70**	30.62**	14.16**	32.70**
R^2	0.42	0.20	0.46	0.31	0.21	0.37
Adjusted R^2	0.42	0.19	0.45	0.30	0.20	0.36
ΔR^2			0.26			0.16

Notes: Values in parentheses are standard errors; entries are unstandardized coefficients.
** $p < 0.01$.
[a] Male = 0, Female = 1.

We further predicted that employee perceived safety climate was positively associated with supervisor-rated safety compliance behaviors (2a), proactive safety behaviors (2b), and prosocial safety behaviors (2c). As shown in Table 4, Model 3, perceived safety climate was positively associated with safety compliance behavior (β = 0.32, SE = 0.04, p < 0.01). Table 5, Model 6 showed that perceived safety climate significantly influenced proactive safety behavior (β = 0.31, SE = 0.04, p < 0.01). As shown in Table 5, Model 9, prosocial safety behaviors predicted perceived safety climate (β = 0.36, SE = 0.04, p < 0.01). Thus hypotheses 2a, 2b, and 2c received support. Additionally, we predicted that perceived employer safety obligations were related to perceived safety climate. As shown in Table 5, Model 9, perceived safety climate predicted perceived employer safety obligations (β = 0.51, SE = 0.06, p < 0.01). Hence, Hypothesis 3 was not rejected.

In Hypotheses 4a, 4b, and 4c, this study projected that safety climate would mediate the relationship between employer safety obligations and safety compliance, proactive safety, and prosocial safety behaviors. To test these hypotheses, we followed MacKinnon's (2008) four-step procedure to establish the mediation effect, which demands: (a) a significant relation between employer safety obligations and three dependent variables such as safety compliance, proactive safety, and prosocial safety behaviors; (b) a significant relationship between employer safety obligations and safety climate; (c) a significant relationship between safety climate and three dependent variables such as safety compliance, proactive safety, and prosocial safety behaviors while controlling for employer safety obligations; and (d) a significant coefficient for the indirect path between the employer safety obligations and three dependent variables, such as safety compliance, proactive safety, and prosocial safety behaviors via safety climate. The bias-corrected percentile bootstrap method determines whether the last condition is satisfied or not. This study estimated parameters for the mediation effect with PROCESS macro (Model 4) by Hayes (2013). This study has included respondents' age, gender, education, and organizational tenure as covariates throughout the analyses.

The first and second steps of MacKinnon's procedure were demonstrated earlier in analyzing hypotheses 1, 2, and 3, and the relationships among the study variables were significant. In the third step, when this model was controlled for employer safety obligations, safety climate was significantly related to safety compliance behaviors (β = 0.22, p < 0.01; Table 4, Model 4); proactive safety behaviors (β = 0.14, p < 0.01; Table 5, Model 7); and prosocial safety behaviors (β = 0.22, p < 0.01; Table 5, Model 10). We utilized the Sobel test (Sobel, 1982) to calculate the critical ratio as a test of whether the indirect effect of the employer safety obligations on the safety behaviors via safety climate is significantly different from zero. The Sobel test confirms the significance of the indirect relationships of employer safety obligations with safety compliance behaviors (z = 4.586, SE = 0.024, p < 0.01), proactive safety behaviors (z = 3.965, SE = 0.018, p < 0.01), and prosocial safety behaviors (z = 4.607, SE = 0.023, p < 0.01). Finally, the bias-corrected percentile bootstrap method showed that the indirect effect of employer safety obligations on safety compliance behaviors via safety climate was significant, ab = 0.23, SE = 0.06, 95% CI = [0.1304, 0.4017]. This mediation effect accounted for 26% of the total effect. Similarly, the indirect effect of employer safety obligations on proactive safety behaviors via safety climate was significant, ab = 0.31, SE = 0.07, 95% CI = [0.1503, 0.4254]. The mediation effect accounted for 46% of the total effect. The indirect effect of employer safety obligations on prosocial safety behaviors via safety climate was significant, ab = 0.26, SE = 0.05, 95% CI = [0.1035, 0.3654]. This mediation effect accounted for 37% of the total effect. Overall, the study results satisfied the four criteria for establishing a mediation effect. Therefore, hypotheses 4a, 4b, and 4c were supported.

5. Discussion

The primary purpose of this study was to identify the relationship between employer safety obligations and employee safety behaviors such as safety compliance, prosocial, and proactive safety behaviors. This study also examined the mediating role of safety climate in the association between employer safety obligations and employee safety behaviors. Time-lagged analyses identify the positive effect of employer safety obligations on employee safety behaviors. The findings of this study showed that employer safety obligations are significantly associated with employee perceived safety climate. This study also revealed that employee perceived safety climate positively related to supervisor-rated employee safety compliance, prosocial safety behaviors, and proactive safety behaviors. The results of this study confirmed the mediating role of employee perceived safety climate in the relationship between employer safety obligations and employee safety compliance behaviors and prosocial and proactive behaviors. These results endorse earlier studies in the safety literature.

5.1. Theoretical contribution

This study examined and found the positive influence of employer safety obligations on employees' compliance, proactive, and prosocial safety behaviors separately based on social exchange theory and psychological contract theory. Mullen et al. (2017) found that employer safety obligations were positively related to employee safety behaviors in line with the findings. This study focused on the three distinct safety behaviors to examine the relative impacts largely ignored in past studies. This study found that employees were more likely to engage in safety citizenship behaviors than safety compliance behaviors. The most surprising finding of this study was that proactive safety behaviors were more likely to be displayed relative to prosocial behaviors if employees perceived that employers fulfilled their safety obligations. This finding indicated that employees were more likely to take charge and initiate safety-related change when they perceived that employers performed safety obligations.

Our results contributed to the extant safety research that revealed that proactive employee safety compliance predicted safety climate attitudes and prosocial safety behaviors. The association between safety climate and prosocial safety behaviors remained unexplored, and this study found a positive linkage. Past studies supported the impact of safety climate attitudes on safety compliance behaviors (Griffin & Neal, 2000; Clarke, 2006; Liu et al., 2019). Moreover, Fugas et al. (2012) found an indirect relationship between safety climate and employee self-reported proactive safety behavior. The results also showed that employer safety obligations impacted the construction of garment employees' safety climate attitudes. Specifically, employees reciprocated favorable safety climate attitudes in response to employer safety obligations. The results were consistent with previous studies that indicated that safety at work could be influenced by organizational safety factors such as management's safety concerns (e.g., Griffin & Neal, 2000; Walker, 2013).

The psychological contract of safety literature primarily investigated the direct relationships between employee safety attitudes (e.g., Newaz et al., 2019b) and safety behaviors (e.g., Mullen et al., 2017). However, no study investigated how employer safety obligations influenced employees' behaviors. Based on social exchange theory and psychological contract theory, this study examined employees' safety climate attitudes as a mediating mechanism through which employer safety obligations translated into employee safety behaviors. This study's findings suggested that employees' safety climate attitudes significantly mediated the positive relationship between employer safety obligations and employee compliance, proactive and prosocial safety behaviors. The fulfillment of employer safety obligations sends a message to employees that the organization values them. In turn, employees reciprocate positive attitudes toward safety policies and practices in their organization, which ultimately leads them to engage in certain safety-related behaviors.

Previous safety literature primarily concentrated on employee self-rated measures of safety behaviors. However, research suggested articulating relevant others to rate employees' safety behaviors (Postlethwaite et al., 2009). This study focused on supervisor-rated employee safety behaviors. Moreover, the data were collected at two points with a four-week lag, which might reduce the potential common method biases. Although the garment industry is full of risks and hazards (Alamgir & Banerjee, 2019), the industry's safety issues are grossly overlooked in the current literature. To the best of our knowledge, no study focused on organizational and human factors of safety in the garment industry. Thus, this study sheds light on how organizational factors (i.e., employer safety obligations) impact garment employees' safety attitudes and behaviors in Bangladesh.

5.2. Practical implication

Employers should provide a safe working environment to employees and arrange compensation systems to safeguard them from accidents. The fulfillment of employer safety obligations ensures protection against injuries and fatalities. Managers must acknowledge that legislative duties related to health and safety (e.g., ensuring safety training and providing safe equipment) can be essential to enhance safety at work and employees' safety behaviors. Employers in the garment industry are expected to provide the necessary knowledge and skills relevant to handling equipment, performing the job safely, and preventing accidents. In addition, garment organizations should constantly monitor and assess the employees' tasks so that no one is forced to participate in potentially hazardous activities. Workplace safety might be jeopardized if coworkers engage in unsafe behavior. It is crucial for organizations to make sure that every employee has received proper safety training and demonstrates their safety behaviors by adhering to the company's safety rules and regulations. Employers have to make sure that the equipment is well-maintained and running effectively in order to protect them from prospective failure. In essence, managers should emphasize their obligations to formulate and implement garment organizations' health and safety systems.

Additionally, when managers fulfill the safety-related obligations, the employee will perceive the safety climate positively. In the garment industry, the employees search for and select extremely concerned employers about safety. Therefore, managers need to realize that the effectiveness of safety policies and programs depends on fulfilling the employers' safety obligations. Proactive and prosocial safety behaviors should be promoted to enhance the safety issues in garment organizations. Managers in garment organizations should gradually emphasize change-oriented safety behaviors to improve organizational safety. In the same vein, helping-related safety behaviors should be encouraged so that employees can assist their coworkers in understanding and maintaining safety concerns. However, if employers do not meet the desired safety standards, employees may not engage in safety citizenship behaviors. Therefore, the manager should test the candidates as to whether they possess proactive and prosocial behaviors in recruiting employees. Furthermore, managers should provide the required training, devise appropriate appraisal systems, and arrange proper incentives to promote these behaviors in garment employees.

5.3. Limitations and future research directions

This study has some limitations. First, in order to avoid potential common method variance (Podsakoff et al., 2003), the data of this study were gathered at two-time phases and from both employees and their immediate supervisors. Employees were asked to rate both independent (i.e., employer safety obligations) and mediating variables (i.e., safety climate), so the data might be inflated. Second, the cross-sectional nature of this study cannot confirm causality. A comprehensive longitudinal study might explore the causal relationships among the study variables in future studies. Third, this study had a significant gender imbalance in the participant sample, with 75.8% of the participants being female because the data were sought from the female-dominated ready-made garment industry. Further studies may concentrate on a balanced selection of male and female participants to generalize the findings. Finally, this research focused on the ready-made garment industry in Bangladesh, and therefore it is not clear to what extent the results generalize to other organizational settings and countries. Further studies may examine other safety-related organizational factors (e.g., servant leadership) and safety performance to enhance understanding of why and how employer safety obligations affect safety outcomes in different occupations and organizational settings.

6. Conclusion

The purpose of the study was to identify why and how employer safety obligations affect employee safety behaviors. This study examined the effect of employer safety obligations on safety attitudes and behaviors. Our time-lagged findings indicated a positive emphasis on social exchange relations in organizations. The fulfillment of employer safety obligations influences employees' safety climate attitudes and exhibits positive safety behaviors. We also found strong associations between employer safety obligations and safety behaviors through safety climate. Thus, we suggest that employer safety obligations play a crucial part in establishing safe working conditions. The results of this study may help organizations facilitate the necessary safety policies and practices.

References

Alamgir, F., & Banerjee, S. B. (2019). Contested compliance regimes in global production networks: Insights from the Bangladesh garment industry. *Human Relations, 72*(2), 272–297.

Blau, P. M. (1964). *Exchange and power in social life*. New York: Wiley.

Brislin, R. W. (1980). Translation and content analysis of oral and written material. *Handbook of Cross-Cultural Psychology, 2*(2), 349–444.

Brondino, M., Silva, S. A., & Pasini, M. (2012). Multilevel approach to organizational and group safety climate and safety performance: Coworkers as the missing link. *Safety Science, 50*(9), 1847–1856.

Chen, C. F., & Chen, S. C. (2014). Measuring the effects of Safety Management System practices, morality leadership and self-efficacy on pilots' safety behaviors: Safety motivation as a mediator. *Safety science, 62*, 376–385.

Clarke, S. (2013). Safety leadership: A meta-analytic review of transformational and transactional leadership styles as antecedents of safety behaviours. *Journal of Occupational and Organizational Psychology, 86*(1), 22–49.

Clarke, S. (2006). The relationship between safety climate and safety performance: A meta-analytic review. *Journal of Occupational Health Psychology, 11*(4), 315.

Curcuruto, M., & Griffin, M. A. (2018). Prosocial and proactive "safety citizenship behaviour"(SCB): The mediating role of affective commitment and psychological ownership. *Safety Science, 104*, 29–38.

Curcuruto, M., Conchie, S. M., Mariani, M. G., & Violante, F. S. (2015). The role of prosocial and proactive safety behaviors in predicting safety performance. *Safety Science, 80*, 317–323.

Curcuruto, M., Mearns, K. J., & Mariani, M. G. (2016). Proactive role-orientation toward workplace safety: Psychological dimensions, nomological network and external validity. *Safety Science, 87*, 144–155.

Curcuruto, M., Parker, S. K., & Griffin, M. A. (2019). Proactivity towards workplace safety improvement: An investigation of its motivational drivers and organizational outcomes. *European Journal of Work and Organizational Psychology, 28*(2), 221–238.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research, 18*(3), 328–388.

Fugas, C. S., Silva, S. A., & Meliá, J. L. (2012). Another look at safety climate and safety behavior: Deepening the cognitive and social mediator mechanisms. *Accident Analysis & Prevention, 45*, 468–477.

Glendon, A. I., & Litherland, D. K. (2001). Safety climate factors, group differences and safety behaviour in road construction. *Safety Science, 39*(3), 157–188.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, 61–178.

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology, 5*(3), 347.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (Vol. 6) River, NJ: Pearson Prentice Hall.

Hair, J. F., Gabriel, M., & Patel, V. (2014). AMOS covariance-based structural equation modeling (CB-SEM): Guidelines on its application as a marketing research tool. *Brazilian Journal of Marketing, 13*(2).

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Hofmann, D. A., & Morgeson, F. P. (1999). Safety-related behavior as a social exchange: The role of perceived organizational support and leader–member exchange. *Journal of Applied Psychology, 84*(2), 286.

Hofmann, D. A., Morgeson, F. P., & Gerras, S. J. (2003). Climate as a moderator of the relationship between leader-member exchange and content specific citizenship: Safety climate as an exemplar. *Journal of Applied Psychology, 88*(1), 170–178.

Huang, Y. H., Ho, M., Smith, G. S., & Chen, P. Y. (2006). Safety climate and self-reported injury: Assessing the mediating role of employee safety control. *Accident Analysis and Prevention, 38*(3), 425–433.

Huang, Y. H., Lee, J., McFadden, A. C., Murphy, L. A., Robertson, M. M., Cheung, J. H., & Zohar, D. (2016). Beyond safety outcomes: An investigation of the impact of safety climate on job satisfaction, employee engagement and turnover using social exchange theory as the theoretical framework. *Applied ergonomics, 55*, 248–257.

ILO (2018). The Rana Plaza Accident and its aftermath, International Labor Organization. https://www.ilo.org/global/topics/geip/WCMS_614394/lang–en/index.htm (accessed on 5 June 2019).

Jiang, L., Lavaysse, L. M., & Probst, T. M. (2019). Safety climate and safety outcomes: A meta-analytic comparison of universal vs. industry-specific safety climate predictive validity. *Work & Stress, 33*(1), 41–57.

Kath, L. M., Marks, K. M., & Ranney, J. (2010). Safety climate dimensions, leader–member exchange, and organizational support as predictors of upward safety communication in a sample of rail industry workers. *Safety Science, 48*(5), 643–650.

Lee, J., Huang, Y. H., Cheung, J. H., Chen, Z., & Shaw, W. S. (2019). A systematic review of the safety climate intervention literature: Past trends and future directions. *Journal of Occupational Health Psychology, 24*(1), 66.

Liu, S., Yang, X., & Mei, Q. (2019). The effect of perceived organizational support for safety and organizational commitment on employee safety behavior: A meta-analysis. *International Journal of Occupational Safety and Ergonomics*, 1–37.

Liu, X., Huang, G., Huang, H., Wang, S., Xiao, Y., & Chen, W. (2015). Safety climate, safety behavior, and worker injuries in the Chinese manufacturing industry. *Safety Science, 78*, 173–178.

Lu, C. S., & Tsai, C. L. (2010). The effect of safety climate on seafarers' safety behaviors in container shipping. *Accident Analysis & Prevention, 42*(6), 1999–2006.

Lu, C. S., & Yang, C. S. (2011). Safety climate and safety behavior in the passenger ferry context. *Accident Analysis & Prevention, 43*(1), 329–341.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York, NY: Taylor & Francis Group.

McClure, K. (2018). 7 of the world's deadliest manufacturing disasters in the last 20 years, sourcify.com https://www.sourcify.com/7-of-the-worlds-deadliest-manufacturing-disasters-in-the-last-20-years/ (accessed 5 June 2019).

Mearns, K., Whitaker, S. M., & Flin, R. (2003). Safety climate, safety management practice and safety performance in offshore environments. *Safety Science, 41*(8), 641–680.
Mei, Q., Wang, Q., Liu, S., Zhou, Q., & Zhang, J. (2018). Effects of organizational safety on employees' proactivity safety behaviors and occupational health and safety management systems in Chinese high-risk small-scale enterprises. *International Journal of Occupational Safety and Ergonomics*, 1–11.
Morrow, S. L., Koves, G. K., & Barnes, V. E. (2014). Exploring the relationship between safety culture and safety performance in US nuclear power operations. *Safety Science, 69*, 37–47.
Mullen, J., Kelloway, E. K., & Teed, M. (2017). Employer safety obligations, transformational leadership and their interactive effects on employee safety performance. *Safety Science, 91*, 405–412.
Neal, A., & Griffin, M. A. (2006). A study of the lagged relationships among safety climate, safety motivation, safety behavior, and accidents at the individual and group levels. *Journal of Applied Psychology, 91*(4), 946.
Neal, A., Griffin, M. A., & Hart, P. M. (2000). The impact of organizational climate on safety climate and individual behavior. *Safety Science, 34*(1–3), 99–109.
Newaz, M. T., Davis, P., Jefferies, M., & Pillay, M. (2019a). Using a psychological contract of safety to predict safety climate on construction sites. *Journal of Safety Research, 68*, 9–19.
Newaz, M. T., Davis, P., Jefferies, M., & Pillay, M. (2019b). The psychological contract: A missing link between safety climate and safety behaviour on construction sites. *Safety Science, 112*, 9–17.
O'Connor, P., O'Dea, A., Kennedy, Q., & Buttrey, S. E. (2011). Measuring safety climate in aviation: A review and recommendations for the future. *Safety Science, 49*(2), 128–138.
Olsen, E. (2010). Exploring the possibility of a common structural model measuring associations between safety climate factors and safety behaviour in health care and the petroleum sectors. *Accident Analysis & Prevention, 42*(5), 1507–1516.
Organ, P., Podsakoff, P. M., & MacKenzie (2006). Organizational citizenship behavior: Its nature, antecedents, and consequences. *Sage, Thousand Oaks, 16*, 43–44.
Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879.
Postlethwaite, B., Robbins, S., Rickerson, J., & McKinniss, T. (2009). The moderation of conscientiousness by cognitive ability when predicting workplace safety behavior. *Personality and Individual Differences, 47*(7), 711–716.
Reader, T. W., Mearns, K., Lopes, C., & Kuha, J. (2017). Organizational support for the workforce and employee safety citizenship behaviors: A social exchange relationship. *Human Relations, 70*(3), 362–385.
Rousseau, D. M. (1990). New hire perceptions of their own and their employer's obligations: A study of psychological contracts. *Journal of Organizational Behavior, 11*(5), 389–400.
Saedi, A. M., Majid, A. A., & Isa, Z. (2020). Relationships between safety climate and safety participation in the petroleum industry: A structural equation modeling approach. *Safety Science, 121*, 240–248.
Smith, T. D., Eldridge, F., & DeJoy, D. M. (2016). Safety-specific transformational and passive leadership influences on firefighter safety climate perceptions and safety behavior outcomes. *Safety Science, 86*, 92–97.
Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology, 13*, 290–312.
Solidarity Center (2018). Fire and Other Health and Safety Incidents in the Bangladesh Garment Sector, https://www.solidaritycenter.org/wp-content/uploads/2018/04/Bangladesh-Safety-Incident-Chart.4.9.18.pdf (accessed on 5 June 2019).
Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5) Boston, MA: Pearson.
Van Dyne, L., & LePine, J. A. (1998). Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal, 41*(1), 108–119.
Walker, A., & Hutton, D. M. (2006). The application of the psychological contract to workplace safety. *Journal of Safety Research, 37*(5), 433–441.
Walker, A. (2010). The development and validation of a psychological contract of safety scale. *Journal of Safety Research, 41*(4), 315–321.
Walker, A. (2013). Outcomes associated with breach and fulfillment of the psychological contract of safety. *Journal of Safety Research, 47*, 31–37.
Wang, Q., Mei, Q., Liu, S., Zhou, Q., & Zhang, J. (2019). Demographic differences in safety proactivity behaviors and safety management in Chinese small–scale enterprises. *Safety Science, 120*, 179–184.
Zohar, D., & Luria, G. (2004). Climate as a social-cognitive construction of supervisory safety practices: Scripts as proxy of behavior patterns. *Journal of Applied Psychology, 89*(2), 322.
Zohar, D., & Luria, G. (2005). A multilevel model of safety climate: cross-level relationships between organization and group-level climates. *The Journal of Applied Psychology, 90*(4), 616–628.
Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology, 65*(1), 96.
Zohar, D. (2000). A group-level model of safety climate : Testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology, 85*(4), 587.
Zohar, D. (2002). The effects of leadership dimensions, safety climate, and assigned priorities on minor injuries in work groups. *Journal of Organizational Behavior, 23*(1), 75–92.

Dr. Md. Shamsul Arefin is an Associate Professor in the Faculty of Business Studies at the Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh. He pursued his Ph.D. from Huazhong University of Science and Technology, China. His research interest is in workplace safety, psychological contract, work-family interface, leadership, and knowledge management. He has written numerous articles that have been published in leading academic journals, such as Personnel Review, Journal of Knowledge Management, and International Journal of Emerging Markets.

Mrs. Ishita Roy is an Associate Professor in the Faculty of Business Studies at the Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh. Currently, she is pursuing her Ph.D. at the Faculty of Business Studies, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Bangladesh. She published several papers in peer-reviewed international journals. Her research interest is in human resources management, work-life balance, safety at work, and organizational behaviors.

Mrs. Swapna Chowdhury is an Assistant Professor in the Department of BBA at the University of Development Alternative, Bangladesh. She published several papers in peer-reviewed international journals. Her research interest is in corporate social responsibility, human resources management, workplace safety, wellbeing, organizational behavior, and leadership.

Dr. Md. Shariful Alam is an Associate Professor in the Faculty of Business and Economics at the United International University, Bangladesh. He pursued his Ph.D. from Wuhan University of Technology, China. He published over 40 articles in peer-reviewed international journals. His research interest is in workplace safety, human resource management, consumer psychology, organizational behavior, and leadership.

Journal of Safety Research 83 (2022) 323–328

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Ensuring data quality and maximizing efficiency in coding agricultural and forestry injuries: Lessons to improve occupational injury surveillance

Erika Scott [a,*], Liane Hirabayashi [a], Kevin Luschen [a], Nicole Krupa [b], Paul Jenkins [b]

[a] Northeast Center for Occupational Health and Safety in Agriculture, Forestry, and Fishing, Bassett Medical Center, Cooperstown, NY, United States
[b] Bassett Research Institute, Bassett Medical Center, Cooperstown, NY, United States

ARTICLE INFO

Article history:
Received 11 April 2022
Received in revised form 23 May 2022
Accepted 9 September 2022
Available online 20 September 2022

Keywords:
Kappa score
OIICS
FAIC
Free-text
Pre-hospital care reports
Percent agreement

ABSTRACT

Introduction: Specialized occupational injury surveillance systems are filling the gap in the undercount of work-related injuries in industries such as agriculture and forestry. To ensure data quality and maximize efficiency in the operation of a regional occupational injury surveillance system, the need for continued dual coding of occupational injury records was assessed. *Methods:* Kappa scores and percent agreement were used to compare interrater reliability for assigned variables in 1,259 agricultural and forestry injuries identified in pre-hospital care reports. The variables used for the comparison included type of event, source of injury, nature of injury, part of body, injury location, intentionality, and farm and agriculture injury classification (FAIC). *Results:* Kappa (κ) ranged from 0.2605 for secondary source to 0.8494 for event and exposure. Individual coder accuracy ranged from medium to high levels of agreement. Agreement beyond the first digit of OIICS coding was measured in percent agreement, and type of event or exposure, body part, and primary source of injury continued to meet levels of accord reaching 70% or greater agreement between all coders and the final choice, even to the most detailed 4th digit of OIICS. *Conclusions:* This research supports evidence-based decision making in customizing an occupational injury surveillance system, ultimately making it less costly while maintaining data quality. We foresee these methods being applicable to any surveillance system where visual inspection and human decisions are levied. *Practical Applications:* Assessing the rigor of occupational injury record coding provides critical information to tailor surveillance protocols, especially those targeted to make the system less costly. System administrators should consider evaluating the quality of coding, especially when dealing with free-text narratives before deciding on single coder protocols. Further, quality checks should remain a part of the system going forward.

1. Introduction

The agricultural, forestry, and fishing industries (AgFF) have consistently held the highest fatality rates of any other sector in the nation. The average worker fatality rate in these fields is seven times that of the national average, at 25.3 FTE versus 3.4 FTE (per 100,000 Full Time Employees) (Bureau of Labor Statistics, 2020). Further, public health surveillance systems commonly fail to attain numbers sufficiently representative of the impact of occupational morbidity among these industries, underrepresenting injury and acute events.

The Survey of Occupational Injuries and Illnesses (SOII) is a common source of data for nonfatal injuries and illnesses in the United States. While these data provide a snapshot into occupational morbidity, it does so through the collection of data for a random sampling of U.S. businesses while excluding the military, self-employed individuals, farms with less than 11 employees, and federal agencies (Bureau of Labor Statistics, 2015). Consequently, estimates are not truly representative of nonfatal injury and illness rates in smaller AgFF operations. As small-scale operations of this sort make up the majority of businesses in the Northeastern United States, large swathes of the AgFF industries ultimately go unaccounted for (Leigh, Marcin, & Miller, 2004; National Occupational Research Agenda, 2008; Ruser, 2008). Worsening this lack of representation, some ancillary reporting systems such as the Occupational Injury Surveillance of Production Agriculture (OISPA) have

* Corresponding author at: Bassett Medical Center, One Atwell Road, Cooperstown, NY 13326, United States

E-mail address: Erika.scott@bassett.org (E. Scott).

https://doi.org/10.1016/j.jsr.2022.09.006
0022-4375/

been discontinued due to unsustainable costs (Centers for Disease Control and Prevention, 2015).

There are numerous ongoing efforts to help fill this gap with accurate reporting. The National Children's Center for Agricultural Health and Safety monitor agricultural injuries through media reports, posted on the AgInjuryNews.org website. Data are gathered from news media, social media, and obituaries for intentional, unintentional, occupational, and non-occupational agricultural cases (Weichelt, Salzwedel, Heiberger, & Lee, 2018). The University of Nebraska Medical Center and Central States Center for Agricultural Safety and Health (CS-CASH) have been surveying self-employed farmers and ranchers since 2011. Their efforts allow for a more accurate understanding of injury rates, consequences, risk factors, and trends in their serviced region (Jadhav, Achutan, Haynatzki, Rajaram, & Rautiainen, 2017). The NIOSH Western States Division monitors the United States Coast Guard (USCG) investigative reports to develop a Commercial Fishing Incident Database (CFID) in order to identify hazards leading to death and/or injury within the fishing community. Their database is applicable at the national level, applying to all fishing industry workers (Case, Lincoln, & Lucas, 2018). Larson et al. developed a surveillance system to focus on migrants and seasonal workers. These workers are notoriously difficult to track accurately due to population movement, differences in definitions, duplicate counts, and a plethora of other hazards. To combat this, the Migrant and Seasonal Farmworker Enumeration Profiles Study (EPS) was created to gather accurate numbers at the county level, as well as state-level estimates for children and youths (Risto, 2021).

While not an exhaustive list of such systems, these efforts illustrate how more precise injury and illness rates are calculated, at a variety of costs. However, there remains a need to provide greater geographic coverage for these specific events and for systems to be comparable. This process is typically accomplished by human coders breaking apart the free-text data to assign consistent coding, often from the Occupational Injury and Illness classification (OIICS) and Farm and Agricultural Injury Classification (FAIC) systems, for example.

Established in 1992 by the Bureau of Labor Statistics, OIICS classifies occupational injuries, illnesses, and fatalities with four main categories: (1) nature of injury, (2) type of event or exposure, (3) source of injury or illness, and (4) body part affected (Bureau of Labor Statistics, 2012). Each category has up to four levels of increasing detail. The OIICS codes are available in a tree format online and are downloadable in desktop version of the tree structure and as Excel files for OIICS version 2.01, which has more than 3,000 individual codes.

Departing from this general occupational approach, the FAIC system focuses on agricultural injuries alone, providing a trove of detailed information from which researchers can pull. The FAIC coding system was developed by the American Society of Agricultural and Biological Engineers (ASABE) to identify if fatalities or injuries related to farming/ranching are occupational (American Society of Agricultural and Biological Engineers, 2020). There are 10 FAIC codes, with four of them linked directly to North American Industrial Classification System (NAICS) codes. In addition to the above, the NORA location specifies the locale of the incident, and intentionality informs as to whether the incident was unintentional or intentional, such as in the case of workplace violence or a suicide.

The accuracy of the final data requires following the established coding rules closely, and performing quality checks. This permits for the aggregation of reports with other sources of data - essentially converting them to a common tongue and allowing different systems to "speak the same language.".

Whenever dual coding is part of the research process, it is important to mitigate variation between coders and to identify potential sources for error. Northeast Center (NEC) researchers developed an algorithm to identify AgFF injury in pre-hospital care reports (PCR) which, until this analysis, employed a dual coding protocol. The mechanics of the machine learning algorithm can be found in previous publications (Hirabayashi, Scott, Jenkins, & Krupa, 2020; Scott, Hirabayashi, Levenstein, Krupa, & Jenkins, 2021). To test the accuracy of classifications by different coders, two types of analysis were conducted: the use of Kappa scores on the first tier of coding, followed by percent level of agreement for the following tiers; the accepted approach when coding categorical data (Landis & Koch, 1977). Gorucu et al. set out to determine levels of agreement regarding codification of the OIICS and FAIC classification systems (Gorucu, Weichelt, Redmond, & Murphy, 2020). They identified five themes in regards to accurate coding of these systems: (a) inclusions/exclusion based on classification system, (b) inconsistent and/or discrepant reports, (c) incomplete and/or nonspecific reports, (d) the effects of supplemental information on coding, and (e) differences in coder interpretation of code selection criteria. With the above in mind, the intent of this paper is to focus on coder interpretation, and describe methods for maintaining the quality of injury record coding for PCRs, while minimizing both time and costs expended in the codification of these surveillance systems.

2. Materials and methods

2.1. Description of dataset

The records used were from the 2011–2016 Maine and New Hampshire pre-hospital care reports (PCRs) that had been determined to definitely or possibly contain an injury or exposure related to the agriculture, forestry, or fishing (AgFF) industries. The validity of PCRs has been established in an international study through the University of Leeds, with a pooled specificity of 0.94 and sensitivity of 0.74 within multiple national databases. In plain language, this means that paramedics correctly excluded a diagnosis in 94% of patients, and correctly diagnosed a patient in 74% of all cases (Wilson, Harley, & Steels, 2018). The process of narrowing down the original dataset of 2,714,766 records to 29,099 records for visual inspection and assignment of AgFF case determination is described in a previous publication: *The Development of a Machine Learning Algorithm to Identify Occupational Injuries in Agriculture Using Pre-Hospital Care Reports* (Scott et al., 2021).

2.2. Classification systems

It is advantageous to use existing coding methods when developing a new system for injury analysis, as this allows others to access those systems for confirmation and further fine-tuning. The coding team applied four classification systems to the dataset: OIICS, FAIC, (both described previously) intentionality, and location. The coding team established a separate set of four codes to determine whether injury intentionality was: (1) unintentional injury; (2) intentional injury, self-inflicted; (3) intentional injury, inflicted by other; and (4) unknown. This simple classification allows researchers to easily identify and categorize violent injuries and deaths and self-harm. The National Occupational Research Agenda (NORA) Agriculture, Forestry, and Fishing Dictionary of Terms identified Location of Incident as one of the preferred categories to define specific characteristics of injuries occurring in production agriculture and support services (Agricultural, 2008:). There are 14 location codes, including but not limited to field/pasture, barn, milkhouse, and farm shop.

Table 1
Agreement thresholds.

Coding Variable		Options Per Variable	Statistical Test	Agreement Required for Future Single Coding
OIICS	Type of Event of Exposure	8*	Level 1: Cohen κ score	Level 1: 0.61
	Primary Source of Injury	10*	Level 2–4: Percent agreement	Level 2: 75%
	Secondary Source of Injury	10*		Level 3: 50%
	Nature of Injury	9*		Level 4: 25%
	Part of Body Affected	9*		
FAIC Code		10	Cohen κ score	0.61
Intentionality		4	Cohen κ score	0.61
Location		14	Cohen κ score	0.61

* For OIICS, the number of options listed per variable is for Level 1.

2.3. Coding interface and protocols

All four coding systems were imported into a Microsoft Access 2016 database, along with the 1,258 PCR records. A form was designed to allow a coder to review the information from the PCR record (narrative, date of birth, admit date, gender, dispatch reason, location, primary impression, and mechanism of injury) and easily assign OIICS, FAIC, intentionality, and NORA location of incident. Links to the online OIICS coding tree, the research team's surveillance manual, EMS abbreviations, and the FAIC code were embedded in the form.

The coding team consisted of nine Northeast Center staff members. The coders were trained by the surveillance team's principal investigator and research coordinator with an overview of the surveillance manual, the coding systems, the coding interface, and hands-on practice. The initial coders were paired into five teams, based on experience (veteran coder with new coder) and availability (number of records to code). Within each coding pair, one was designated as coder A and the other as coder B, which corresponded to the database form they were instructed to use. As a result, each coder within an assigned pair was coding independently of their counterpart. The coder pairs were assigned between 150 and 350 records, depending on their reported availability.

After completing 25 records, the coder pairs viewed and resolved discrepancies in their coding. If unable to agree on a resolution, the records with discrepancies were reviewed by the surveillance team as a whole. Space was provided on all forms for coders to enter comments or questions. The PI and research coordinator updated the surveillance user manual to reflect feedback from coders and reviewed the updates with the coders.

2.4. Guidelines for agreement levels

Before beginning analysis of interrater reliability, the research team established thresholds for coder agreement for these systems, which, if met or exceeded, would eliminate the need for dual coding in the future. The agreement thresholds are listed in Table 1. The Cohen κ score is considered the key statistics for measuring interrater reliability, as it controls for the possibility of chance agreement (Gorucu et al., 2020; Landis & Koch, 1977). Choosing 0.61 as the Cohen κ score threshold was based on Viera and Garrett's interpretation of Kappa (see Table 2) for categorical variables, with 0.61 as the lowest point of substantial agreement (Viera & Garrett, 2005).

The research team did not use the Cohen κ score for OIICS Levels 2–4 due to the high number of categorical options, as shown in Table 3; in situations such as these, where many response options may not be selected, the use of Cohen score loses its power (McHugh, 2012). In explanation, while OIICS Source Level 1 consistently has nine options, allowing for the use of a Kappa rating, Levels 2–4 have multiple options. For example, in Level 3 there are six possible branches with the Construction Machinery source, but only four choices with Agricultural machinery. As a result, percentage agreement was employed, which is the number of records where coders agreed divided by the total number of records reviewed. This protocol was approved by the Institutional Review Board of the primary institution.

Table 2
Viera & Garrett's Interpretation of Cohen κ Score.

Kappa Range	Agreement
< 0	Less than chance agreement
0.01–0.20	Slight agreement
0.21– 0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

3. Results

The results of the coding pairs can be found in Table 5, where kappa (κ) ranged from 0.2605 for secondary source, to 0.8494 for event and exposure. Comparisons between the coder choice and the final choice can also be found in Table 4. Agreement was almost perfect between the individual coders and the final coding choice for type of event or exposure, body part, and primary source of injury, and there was substantial agreement between the nature of the injury, intentionality, and NORA location.

Agreement beyond the first digit of OIICS coding was measured in percent agreement, and type of event or exposure, body part, and primary source of injury continued to meet high levels of accord when tested against the final coding, while the secondary source of injury and the nature of the injury were also high, as seen in Table 5. These scores were lower when comparing coders against each other, as shown in Table 6, with the highest levels of agreement occurring in source of injury, secondary source, and event exposure.

Individual coder accuracy ranged from medium to high levels of agreement. As shown in Table 7, Kappa scores reached perfect levels of agreement between some coders in nature of injury, source and secondary source, and event exposure; though one coder only scored fair agreement regarding secondary sources of injury. These levels of agreement continued into the 2nd, 3rd, and 4th digits, with primary and secondary sources of injury and event exposure maintaining high percentages of agreement throughout. Confidence intervals were all in the positive direction with the exception of FAIC codes.

4. Discussion

It is reasonable to assume that the accuracy of some variables is more critical to successful public health interventions than others. For example, knowing the primary source of injury and the type of event and exposure are important in our ability to improve safety.

Table 3
Number of options available for OIICS Levels 2–4.

	Event/Exposure Type	Source of Injury	Nature of Injury	Part of Body Affected
Level 2	48	78	40	45
Level 3	178	439	192	91
Level 4	304	1139	375	74

Table 4
Comparing Coder versus Coder / Coder vs Final Coding.

	Comparing Coder A vs Coder B			Comparing Coder vs Final Coding		
OIICS (1st digit)	N	Kappa	95% CI	N	Kappa	95% CI
Nature of Injury	1258	0.5727	0.5241–0.6212	2517	0.7653	0.7366–0.7939
Body Party	1250	0.7479	0.7211–0.7748	2509	0.8536	0.8382–0.8689
Source of Injury 1	1258	0.7330	0.7042–0.7618	2517	0.8515	0.8353–0.8677
Source of Injury 2	1258	0.2605	0.2020–0.3190	2517	0.5198	0.4829–0.5566
Event Exposure	1258	0.8494	0.8271–0.8717	2517	0.9195	0.9075–0.9315
FAIC	1258	0.5346	0.4983–0.5708	2517	0.5939	0.5697–0.6182
Intentionality	1259	0.6052	0.4842–0.7261	2518	0.7395	0.6608–0.8181
NORA location	1257	0.5635	0.5321–0.5949	2516	0.7532	0.7344–0.7719

Table 5
Coder (A and B) versus Final Choice: % Agreement for OIICS codes beyond the first digit.

	Nature of Injury (n = 2517)	Body Part (n = 2509)	Source of Injury 1 (n = 2517)	Source of Injury 2 (N = 2517	Event Exposure (n = 2517)
2 digits	78.94	78.68	85.66	82.48	87.41
3 digits	74.29	71.26	82.80	82.08	81.17
4 digits	70.56	70.67	76.52	81.33	75.29

Table 6
Coder A versus Coder B: % Agreement for OIICS codes beyond the first digit.

	Nature of Injury (n = 1258)	Body Part (n = 1250)	Source of Injury 1 (n = 1258)	Source of Injury 2 (N = 1258)	Event Exposure (n = 1258)
2 digits	62.80	63.12	74.48	79.65	76.79
3 digits	54.93	50.16	70.35	78.78	65.58
4 digits	49.28	48.96	59.86	76.79	54.29

Table 7
Coder range of agreement.

	KAPPA	95% CI	2 digit agreement (%)	3 digit agreement (%)	4 digit agreement (%)
Nature of Injury	0.5818–1	0.5971–1	54.09–91.17	44.03–86.93	38.99–84.10
Body Part	0.7986–0.9411	0.4563–1	66.03–88.69	56.41–84.45	55.13–83.75
Source of Injury 1	0.6904–1	0.6040–1	71.52–89.40	62.66–88.34	51.90–83.75
Source of Injury 2	0.2987–1	0.1343–1	80.44–87.28	79.34–87.28	77.13–87.28
Event Exposure	0.8517–1	0.7897–1	79.25–100.00	70.44–100.00	55.97–100.00
FAIC	0.3149–0.8811	−0.0522–1			
Intentionality	0.5687–0.8203	0.1303–1.0000			
NORA location	0.6475–0.8160	0.3237–1.0000			

Public health interventions often target these two factors through a variety of ways, be it elimination, substitution, engineering controls, administrative controls, or personal protective equipment. An example would be data indicating tractor (source of injury) rollovers (event or exposure) being a significant cause of injury, and an intervention targeting the installation of rollover protective structures. Though possible, interventions do not typically target a specific nature of injury (e.g., preventing leg fractures, but not leg crushing), therefore those variables, while helpful in fully understanding the burden of injury, are not always critical for intervention development. Near perfect agreement for event and exposure, primary source of injury, and body part, along with the substantial agreement for nature, location, and intent, gives the research team confidence in changing the surveillance system protocols to eliminate time-consuming dual coding. When disagreement did occur, it was typically not wildly disparate, but more to do with the ordering of an injury event or noting the many rules within the OIICS system. It is worth noting that, due to its infrequent use, there was greater variation in the kappa score for secondary source of injury. Given that secondary source is not frequently assigned, this is not a critical value in the decision to change from dual to single coding. FAIC Coding showed moderate agreement between coders and the final choice. This is an area where additional guidance and training has been warranted. Our ability to draw distinctions between production agricultural injury events, which are likely captured in systems such CFOI or SOII (for larger events), from bystander injures, such as children hurt on the farm, is important. The blurring of the farm as often a workplace but often a home necessitates a means to code beyond traditional definitions of "work" to capture true risk. These findings mirror the kappa scores and general suggestions of Gorucu et al. (Gorucu et al., 2020).

These analyses allowed us to assess and customize additional coder training. Without the rigor of dual coding, there will be continued need for quality assurance checks. These should involve the senior members of the research team visually inspecting a random sample of coded cases on a routine schedule, to ensure that data quality is maintained. In addition, we recommend that a coder training environment is established, where they can practice and get feedback on their kappa scores and percent agreement, before coding new data. These proposed protocols could be applied to a variety of surveillance systems, especially when there is a concern to reduce the staff-time involved in running the system, without sacrificing data quality.

Various limitations presented themselves over the course of this project. It was necessary to train reviewers in assigning codes, and some developed a more firm understanding of the principles behind the coding than did others. Further, it is natural for code selection to drift as one becomes more familiar with the subject matter, albeit ideally in the direction of increasing accuracy. At other times, the narratives themselves represented a limitation, as not all information necessary for OIICS coding was always present. In these cases, there is a tendency to make assumptions that may not be borne out by the available information; for example, assuming an injury was unintentional when the narrative does not contain the details necessary to determine that status. A further limitation is that the diagnosis offered by an EMS PCR does not always match the final diagnosis as settled upon by the attending hospital physician, particularly in regards to diagnosis sensitivity (Wilson et al., 2018). While still effective in tracking AFF occupational injuries, the OIICS coding might differ somewhat if hospital records were incorporated alongside EMS PCRs.

5. Conclusions

This research provides for evidence-based decision making for customizing an occupational injury surveillance system, ultimately making it less costly. The quality of the coded data was acceptable for variables important for injury epidemiology and intervention development. Good stewardship of public health resources is critical for the long-term success of such programs, and continued refinements and cost-savings should be considered an important part of the system.

5.1. Practical applications

Assessing the rigor of occupational injury record coding provides critical information to tailor surveillance protocols, especially those targeted to make the system less costly. System administrators should consider evaluating the quality of coding, especially when dealing with free-text narratives before deciding on single coder protocols. Further, quality checks should remain a part of the system going forward.

Funding

This work was supported by the National Institute of Occupational Safety and Health [Grant No 2U54OH007542].

Acknowledgements

We are grateful to the staff who spent considerable time learning the coding systems and reviewing cases, including but not limited to: Deb Dalton, Katherine Franck, Meghan Goodspeed, Judy Graham, Rebecca Meininger, Amanda Roberts, Kathy Smith, and Ryan Todd.

References

American Society of Agricultural and Biological Engineers. Farm and agricultural injury classification (FAIC) code. May, 2020. St. Joseph, MI: ASABE.

Bureau of Labor Statistics. *Occupational Injury and Classification Manual Version 2.01*. January 2012. Retrieved from https://www.bls.gov/iif/oshoiics.htm.

Bureau of Labor Statistics. Survey of Occupational Injuries and Illnesses (SOII) respondents home page. Accessed June 25, 2015. http://www.bls.gov/respondents/iif.

Bureau of Labor Statistics. (2020). Number and rate of fatal work injuries, civilian workers, by major occupational group. Bureau of Labor Statistics. Accessed May 20, 2022. https://www.bls.gov/charts/census-of-fatal-occupational-injuries/number-and-rate-of-fatal-work-injuries-by-occupation.htm.

Case, S. L., Lincoln, J. M., & Lucas, D. L. (2018). Fatal falls overboard in commercial fishing – United States, 2000–2016. *MMWR. Morbidity and Mortality Weekly Report, 67*(16), 465–469. https://doi.org/10.15585/mmwr.mm6716a2.

Centers for Disease Control and Prevention. (2015). Looking to the future for agriculture injury surveillance at NIOSH - agriculture, forestry and fishing resources. Accessed October 18, 2019. http://www.cdc/gov/niosh/agforfish/aginjurysurv.html.

Gorucu, S., Weichelt, B., Redmond, E., & Murphy, D. (2020). Coding agricultural injury: Factors affecting coder agreement. *Journal of Safety Research, 75*, 111–118. https://doi.org/10.1016/j.jsr.2020.08.006.

Hirabayashi, L., Scott, E., Jenkins, P., & Krupa, N. (2020). Occupational injury surveillance methods using free text data and machine learning: Creating a gold standard data set. *SAGE Research Methods Cases*. https://doi.org/10.4135/9781529720488.

Jadhav, R., Achutan, C., Haynatzki, G., Rajaram, S., & Rautiainen, R. (2017). Injury risk factors to farm and ranch operators in the Central United States. *American Journal of Industrial Medicine, 60*, 889–899. https://doi.org/10.1002/ajim.22757.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics., 33*(1), 159–174. https://doi.org/10.2307/2529310.

Leigh, J. P., Marcin, J. P., & Miller, T. R. (2004). An estimate of the U.S. government's undercount of nonfatal occupational injuries. *Ann Epidemiology., 46*(1), 10–18.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem Med., 22*(3), 276–282. PMID: 23092060; PMCID: PMC3900052.

National Occupational Research Agenda. National agriculture, forestry and fishing research agenda. 2008. Accessed December 7, 2021. http://www.cdc.gov/niosh/nora/comment/agendas/AgForFish/pdfs/AgForFishDec2008.pdf.

NORA Agricultural, F., and Fishing Sector Council. (2008). NORA National Agriculture, Forestry, and Fishing Agenda for Occupational Safety and Health in Research and Practice in the U.S. Agriculture, Forestry, and Fishing Sector. 77.

Risto, R. (2021). Surveillance of agriculture, forestry, and fishing injury, illness, and economic impacts. *Journal of Agromedicine, 26*(1), 59–61. https://doi.org/10.1080/1059924X.2021.1849508.

Ruser, J. (2008). Examining evidence on whether BLS undercounts workplace injuries and illnesses. Accessed February 19, 2015. http://www.bls.gov/opub/mlr/2008/08/art2full.pdf.

Scott, E., Hirabayashi, L., Levenstein, A., Krupa, N., & Jenkins, P. (2021). The development of a machine learning algorithm to identify occupational injuries in agriculture using pre-hospital care reports. *Health Information Science and Systems, 9*, 31. https://doi.org/10.1007/s13755-021-00161-9.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*(5), 360–363.

Weichelt, B., Salzwedel, M., Heiberger, S., & Lee, B. C. (2018). Establishing a publically available national database of U.S. news article reporting agriculture-related injuries and fatalities. *American Journal of Industrial Medicine, 61*, 59–61. https://doi.org/10.1002/ajim.22860.

Wilson, C., Harley, C., & Steels, S. (2018). Systematic review and meta-analysis of pre-hospital diagnostic accuracy studies. *Emergency Medicine, 35*(12), 757–764. https://doi.org/10.1136/emermed-2018-207588.

Erika Scott is the Deputy Director of the Northeast Center for Occupational Health and Safety. She works collaboratively to reduce the occupational morbidity and mortality in agriculture, forestry, and fishing. Her research in logging health and safety began at the New York State Department of Health analyzing tree-related fatalities, leading to a focus on forestry and logging when she started as a research coordinator at the Northeast Center. She later became a Principal Investigator, developing the Maine Logger Health and Safety Study, involving longitudinal surveys and in-person health assessments across Maine. She has gathered detailed data on nearly 400 Maine loggers.

Kevin Luschen is a research coordinator with a background in project management, data administration, and occupational health and safety. He has work experience including security, occupational health and safety in the oil and gas industry, teaching, and military service.

Liane Hirabayashi served as the research coordinator for two projects at the Northeast Center for Occupational Health and Safety in Agriculture, Forestry, and Fishing: "Assessing Overall Health and Improving Injury Surveillance of Maine Logging Workers"; and "Improving Methods for Traumatic Injury Surveillance in Agriculture, Forestry, and Fishing." She also assisted with coordination on "Giving Safety a Competitive Advantage: Increasing PFD Use Among Lobster Fishermen." Liane used her expertise in Microsoft Access to redesign existing databases to improve data entry and reporting, and her background in organization development

to streamline processes, including the database used for the coding in this paper.

Nicole Krupa is an Informatics Analyst, with 20 years of experience in data management. She is well adept at working with data in a variety of formats, and has used SAS extensively for the management and analysis of data. She has also worked with many external data sources.

Paul Jenkins is the senior statistician at the Bassett Healthcare Network Research Institute and has considerable experience in a wide range of statistical methods. As a doctoral level statistician, he has written the methods and analysis sections of 58 grant applications ranging from large multi-year/R^{-01}(−|-) proposals to smaller state and local applications and has served as the senior statistician on those proposals that were funded.

Journal of Safety Research 83 (2022) 388–399

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Evaluation of advanced curve speed warning system to prevent fire truck rollover crashes

Peter Simeonov *, Ashish Nimbarte [1], Hongwei Hsiao [2], Richard Current, Douglas Ammons, Hee-Sun Choi [3], Md Mahmudur Rahman [4], Darlene Weaver

Division of Safety Research, National Institute for Occupational Safety and Health, 1000 Frederick Lane Morgantown, WV 26505, USA

ARTICLE INFO

Article history:
Received 26 January 2022
Received in revised form 31 May 2022
Accepted 21 September 2022
Available online 5 October 2022

Keywords:
Fire tanker
Horizontal curves
Emergency response
Speeding
Overturn

ABSTRACT

Introduction: A disproportionately high number of deadly crash-incidents involve fire-tanker rollovers during emergency response driving. Most of these rollover incidents occur at dangerous horizontal curves (“curves”) due to unsafe speed. This study examined the effects of a curve speed warning system (CSWS) on fire tanker drivers’ emergency response behavior to develop system improvement suggestions. *Method:* Twenty-four firefighters participated in driving tests using a simulator. A fire tanker model, carrying a full tank of water, was used in emergency driving tests performed with and without CSWS. The CSWS was designed using the algorithm for passenger vehicles with a few initial modifications considering the unique requirements of heavy fire tanker and emergency driving. *Results:* The results indicated that the CSWS was effective in issuing preemptive warnings when the drivers were approaching curves with unsafe speed during emergency response. Warnings occurred more frequently at curves with smaller radius. Although the CSWS improved driving performance, it did not significantly reduce the number of rollover events. A detailed analysis of the rollover events provided suggestions for improvement of CSWS algorithms. *Conclusions:* To further improve the CSWS algorithm, the following may be considered: including increased safety speed margin below the rollover critical speed, moving the speed warning trigger from the curve apex to the curve entry point, extending the safe speed-control zone to cover the entire curve, and employing artificial intelligence to accommodate individual driving styles. *Practical Applications:* Fire tankers continue to be at increased risk of rollover during emergency response due to unsafe negotiation of dangerous curves. Development and use of advanced driver assist systems such as CSWS evaluated in this study may be an effective strategy to prevent deadly rollover crash-incidents. The knowledge generated by this study will be useful for system designers to improve the CSWS specifically designed for heavy emergency vehicles.

1. Introduction

Transportation-related injuries remain a significant problem for firefighters in the United States. In 2019, an estimated 15,350 collisions resulting in 575 firefighter injuries were directly related to emergency vehicles responding to or returning from incidents (Campbell & Evarts, 2020). A disproportionately high number of these incidents involve fire tankers. Tankers represent only 3 % of all fire apparatus in the United States but were involved in 21.9% of all fire vehicle fatal crashes that took place in the period 1990 to 2001 (FEMA, 2003). Rollover crashes are the most common and most deadly incidents for fire tankers – of the 63 crashes with 73 deaths involving tankers in the period 1977–1999, 77.8% of the crashes and 74.0% of the deaths involved a rollover (NIOSH, 2002).

Fire tankers, defined as “mobile water supply apparatus” (FEMA, 2003), are some of the heaviest apparatus operated by a fire department. The water carry capacity of typical fire tankers is in the range 5,678 to 11,356 liters (1,500 to 3,000 gallons) and weigh more than 25 tons (55,000 lb.) (NFPA, 2008). Furthermore, fire tankers have high center of gravity and are often referred as “top-heavy” vehicles. Vehicles that have a high center of gravity are more challenging to maneuver and control. Specifically, top-

* Corresponding author.
E-mail address: psimeonov@cdc.gov (P. Simeonov).
[1] Permanent Address: West Virginia University, 349 Engineering Science Building, Morgantown, WV 26506-6107, USA.
[2] Current Address: Texas A&M University – Corpus Christi, 6300 Ocean Drive, Corpus Christi, TX 78412, USA.
[3] Current Address: Texas Tech University, 2500 Broadway, Lubbock, TX 79409-2051, USA.
[4] Current Address: Purdue University, 315 N. Grant Street, West Lafayette, IN 47907-2023, USA.

https://doi.org/10.1016/j.jsr.2022.09.014
0022-4375/

heavy fire trucks driven at an unsafe speed through a horizontal curve ("curve") have a tendency to rollover. The tendency of a vehicle tipping/rolling over while moving through a curve is a matter of simple physics involving inertia and momentum - as the vehicle negotiates the turn, its weight leans in the direction opposite to which the vehicle is turning (FEMA, 2003).

Road-related factors, such as the curve radius and road banking or super-elevation, provide adequate drainage and facilitate safe and comfortable negotiation of the curve at a reasonable speed (FEMA, 2003). The smaller the radius of a curve the larger the tipping forces for the same speed of the vehicle. Flat curves, or curves without super-elevation are more challenging and require additional speed reductions to avoid the vehicle sliding or rolling over. The risk further increases if the roadway is banked toward the outside of the turn (i.e., the curve with negative super-elevation). The effects of these road geometry-related factors can be controlled if the vehicle is operated within an adequate speed. The real danger for a rollover is when the driver attempts to negotiate a dangerous curve at an unsafe speed, which may happen during emergency driving.

A critical review on rollover of heavy commercial vehicles (Winkler, 2000) revealed that "the rollover threshold of loaded heavy trucks extends well into the 'emergency' maneuvering capability of the vehicle and sometimes into the 'normal' maneuvering range, and it is relatively hard for truck drivers to perceive their proximity to rollover while driving. Rollover is very much like walking up to a cliff with your eyes closed: as you approach the edge, you are still walking on solid ground but once you've stepped over, it's too late. Further, the rollover threshold of a commercial truck changes regularly as the load changes, so drivers may not have the chance to get used to the stability of their vehicle." Driving a tanker carrying fluid (which may shift and slosh in the tank) under emergency conditions poses several additional challenges regarding rollover threshold and vehicle stability, due to the associated time pressure and excessive speed.

Excessive speed for specific driving conditions has been identified as a major contributing factor for vehicle-related firefighter fatalities (FEMA, 2003). Emergency responders are commonly allowed to travel at speeds above the posted limits (FEMA, 2003). Speeding associated with emergency response driving increases the amount of risk imposed upon firefighters and the apparatus in which they ride. Therefore, firefighters should be trained in safe driving practices, including getting to know the dangerous routes and curves in their response area. Furthermore, some fire apparatuses are equipped with safety technology such as speed limiters and data collection black boxes, and most of the modern fire apparatuses have stability control systems (NFPA, 2008). Despite these measures, the risk of speed-related crashes and injuries for firefighters riding in fire apparatus remains high.

A promising technological approach to assist fire apparatus drivers in controlling speed in curves may be the use of a curve speed warning system (CSWS) (Pomerleau et al., 1999); however, its effectiveness in emergency driving of heavy vehicles such as fire tankers has not been well studied. The CSWS is an advanced driver assistance system (CSWS-ADAS) that uses information from digital maps and the vehicle's current location and speed from a global positioning system (GPS) to issue a warning when the vehicle approaches a curve at an unsafe speed. The system calculates the required deceleration and estimates the distance ahead of the curve at which to issue a warning such that the driver can safely reduce the vehicle speed before entering the curve. A great advantage of the CSWS-ADAS is that it can use variable or dynamic speed limits with the ability to adapt to different road and weather conditions (Jimenez, Liang, & Aparicio, 2012). Furthermore, it can be based on standard GPS (Chowdhury, Faizan, & Hayee, 2020) or connected vehicle technologies (Wang, Wang, Zheng, & Zhixia, 2020), and be adaptive to individual driver behavior (Ahmadi & Ghanipoor Machiani, 2019).

The CSWS-ADAS has also been occasionally referred to as an intelligent speed assistance/adaptation (ISA) system. ISA is a generic term for a class of ADAS in which the driver is warned and/or vehicle speed is automatically limited when the driver is, intentionally or inadvertently, traveling over the posted speed limit, or some other pre-defined (fixed) speed threshold (Young, Regan, Triggs, Jontof-Hutter, & Newstead, 2010). The benefits of ISA technology are well documented to reduce speed, speed variability, speed violations, and injury and fatal crashes (Young et al., 2010). Devices that exercise a greater control over the driver are seen to be most beneficial, as opposed to simple advisory systems. However, these controlling systems are not necessarily appreciated by drivers (Young et al., 2010). Several negative effects have been observed with ISA. Two key issues are (1) acceptability of the system warnings and (2) driver adaptation or system over-reliance. System over-reliance is a particular concern as faster speeds in curves by some drivers have been observed. Research on ISA use in heavy trucks remains limited (Fitzharris et al., 2011), and it is not clear if the described ISA effects during the operation of general vehicles are applicable to heavy fire trucks.

Recently, Simeonov et al. (2021) evaluated the effectiveness, safety outcomes, and driver acceptance of a CSWS-ADAS during emergency response of a fire tanker using a driving simulator. The research findings suggested that the drivers reduced their driving speed at curve approaching and entering phases for most challenging curves, without affecting the overall time in completing the test route. Furthermore, drivers had a reduced number of severe braking and decreased average in-curve distance traveled over the safety speed limits, when the CSWS was in use. Drivers also rated the CSWS as assisting, effective, and useful. Overall, the study demonstrated that the CSWS can enhance fire truck safety during emergency driving without sacrificing drivers' precious response time.

In the study of Simeonov et al. (2021), the algorithm for the CSWS was adapted from the guidelines of Pomerleau et al. (1999) and was further modified to meet the demands (for reducing the risk of rollover) of a heavy fire tanker in emergency driving conditions. The working principle in the CSWS algorithm is based on estimating the maximum curve safety speed and using it to calculate the warning distance (Pomerleau et al., 1999). In the existing studies, which are predominantly related to passenger cars (with a low center of gravity), the safety speed profiles were determined based on the risk of sideslip considering the factors such as curve radius, road super-elevation, and friction factor (Chowdhury et al., 2020; Jimenez et al., 2012; Wang et al., 2020). The warning distance estimation in most of these studies was based on the curve apex serving as the curve safety speed target location, which is adequate for passenger cars since drivers usually continue decelerating well into the curve (Bella, 2014).

While formulating the CSWS algorithm for the fire tankers (with an elevated center of gravity) driven under emergency conditions, the factors such as mass distribution and rollover susceptibility were considered to estimate the safety speeds. Based on the recommendation that heavy trucks require earlier speed reduction to safely negotiate a curve (IAFF, 2010), the in-curve target for reducing the vehicle speed (at or below the curve safety speed) was moved from the curve apex to the entry-apex mid-point, thus establishing a safety speed zone (from the entry-apex mid-point to the curve apex). Despite these enhancements, during the simulated emergency responses, no significant difference in the rollover crash events was observed with and without CSWS conditions, signifying a need for further improvement in the CSWS algorithm. Therefore, in this study an in-depth analysis of the fire truck rollover crash events was conducted with a goal of developing meaningful guidance for further improving the CSWS algorithms for emergency driving of heavy fire trucks.

2. Methods

2.1. Participants

The study participants were active members (career or volunteer) of the fire departments in Morgantown, WV, and the surrounding area. The inclusion and exclusion criteria consisted of age >=18 years, a valid driver's license, more than 6 months of experience driving a fire truck, ability to follow the study protocol and give informed consent, and no symptoms of motion sickness. Mean age, height, and weight of the participants were 36 years (SD = 10.1 years), 182.9 cm (SD = 4.8 cm), and 103.5 kg (SD = 16.8 kg), respectively.

2.2. Equipment

2.2.1. Driving simulator

A motion-base simulator (Mechanical Simulation, Ann Arbor, MI) with three degrees of freedom motion (roll, pitch, and heave) was used in this study. The simulator consists of three 178 cm (70 in) high-definition display screens, a high-fidelity sound system for realistic sound effects, a precision steering system, commercial grade foot controls, and a reconfigurable instrument cluster. A TruckSim (Mechanical Simulation, Ann Arbor, MI) based tanker model with a 3-axle fire truck in a "laden" condition (i.e., carrying full tank of water; 11,356 liters/3000 gallons) with a total weight of 26,822 kg (59,008 lb.) was used in this study (Fig. 1a). The model featured accurate dynamic performance of a heavy fire truck tanker, including truck dimensions, geometry, mass distribution, engine power, acceleration, steering and braking performance, suspension, and tire-road interaction (friction). Unity (Unity Technologies, San Francisco, CA) software was used to develop road geometry and create interactive driving scenarios with advanced graphic design and performance.

2.2.2. Warning system interface

The CSWS graphic user interface (GUI) provided data on the current speed of the vehicle, the posted speed limit for the current route section, and the safety speed for the upcoming curve. The system status was shown using color-codes: blue – system inactive; green – normal or OK, yellow – caution with sound warnings with frequency 2.6–3.1beeps/s accompanied by a blinking arrow in the direction of the upcoming turn, and red with a steady arrow in the direction of the upcoming turn – danger with sound warnings with frequency 3.2–4.0 beeps/s. A touch screen tablet (Windows Surface Pro 4, 312 mm, Microsoft, Redmond, WA) was used to display the CSWS GUI (Fig. 1b, c).

The audible warning was with a fundamental frequency of 300 Hz and 15 harmonic components (Gonzalez, Lewis, Roberts, Pratt, & Baldwin, 2012). The audio warning had a pulse duration of 200 ms and variable inter-pulse interval (185–50 ms). The inter-pulse duration was regulated to increase the frequency of audio pulses to indicate increasing danger when speed reduction was insufficient or absent. The warning signal was issued with an increasing frequency, which was a function of increasing values of the calculated deceleration required to reach the safety speed to avoid a rollover event. In the situations with increasing danger, synchronized vibration signals at the steering wheel (actuated by

Fig. 1. Selected simulations: (a) fire tanker model; (b) view from the cab, with active CSWS while driving on a straight section within the safety speed limits (green screen); (c) approaching a curve with inappropriate speed – the active CSWS is issuing a warning (red screen).

the power steering system) were provided along with the audible warning signals.

2.2.3. Warning algorithm

The original curve speed warning system (CSWS) algorithm, mostly applicable to light passenger vehicles, is based on the following equation (Pomerleau et al., 1999):

$$a = \frac{V^2 - V_s^2}{2(d - t_r V)} \quad (1)$$

where:

a = deceleration required to reach V_s at curve apex.
V = vehicle speed.
V_s = curve maximum safe speed.
d = distance between vehicle position and curve apex.
t_r = driver reaction time (assumed to be 1.5 s) (Pomerleau et al., 1999).

Based on this algorithm, a warning is issued when the calculated deceleration value becomes higher than a preset average deceleration value (1.5 m/s^2). Considering the specific requirements of a fire truck driving under emergency conditions, the following modifications were made to the algorithm (Fig. 2) (Simeonov et al., 2021):

1. Recognizing that the most critical speed-related crash event in a curve for a heavy truck on a dry road is a rollover, the safety speed profile (V_s) was set as 90% of the rollover critical speed. The rollover critical speed (Vroll.cr) was determined according to (Eq. (2)). The maximum lateral acceleration ($a_{lat.max}$) was measured with a swept steer test at 40 mph (64 km/h) in TruckSim using Proving Grounds (Unity Scene) (by Mechanical Simulation, Ann Arbor, MI) (Simeonov et al., 2021).

$$V_{roll.cr} = \sqrt{Ra_{lat.max}}; \quad V_{roll} = 0.9V_{roll.cr} \quad (2)$$

The rollover critical speed is higher than the posted speed and the slip-related safety speed (based on the side friction factor assuming wet and icy conditions) and thus aids with reducing unnecessary warnings during an emergency response driving on a dry road.

2. The V_s target location was shifted from the curve apex to the middle (50%) of the entry-apex curve section. This alteration was made knowing that a heavy truck driver must reduce vehicle speed (and reach V_s) much earlier in the curve (IAFF, 2010) compared to light passenger vehicles.
3. Between the V_s target location and the curve apex, a “speed control zone” was established. Within this zone a simple control logic was implemented to issue a warning if the vehicle speed exceeded the curve safety speed.

2.3. Procedure

The data collection begun with the investigators describing the study and the experimental tasks to the participants, answering their questions, and obtaining their signatures on the informed consent forms approved by the NIOSH IRB. Participants were screened for susceptibility to motion sickness and to obtain baseline scores for the subsequent motion sickness monitoring (Hoffman, Molino, & Inman, 2003).

Each of the 24 participants completed a pre-test driving task followed by the main test – an emergency response driving task completed in two trials – one with CSWS “on” and one with CSWS “off” (for a total of 48 trials). The participants completed the CSWS “on” and “off” trials in a balanced order – half of the participants started with CSWS “on” followed by a trial with CSWS “off;” for the other half of the participants this order was reversed. Before each trial the participants were informed/instructed about the status (“on” or “off”) of the CSWS. The participants wore their firefighter protective pants and boots during the driving tasks. The pre-test task lasted for 10–15 minutes and was primarily designed to help the participants get familiar with the simulator environment. The pre-test route was 11.126 km (6.9 miles) long with 14 curves having radius (R) in the range of 51–612 m. Eight of the 14 curves had an R < 200 m and 5 of the 8 had an R < 100 m.

For the emergency response driving task, the route was 12.640 km (7.9 miles) long with 18 curves having R in the range of 45–1229 m. Eleven of the 18 curves had an R < 200 m and 7 of the 11 had an R < 100 m (Fig. 3). The driving environment consisted of a rural two-lane (lane width = 3.36 m/11 ft) road on a hilly terrain with varied vegetation that partially occluded some of the upcoming curves. The simulated task began with a radio emergency dispatch message. The participants then turned on the emergency lights and sirens and started to drive as fast as possible, but safely. The participants were instructed to address the warnings by

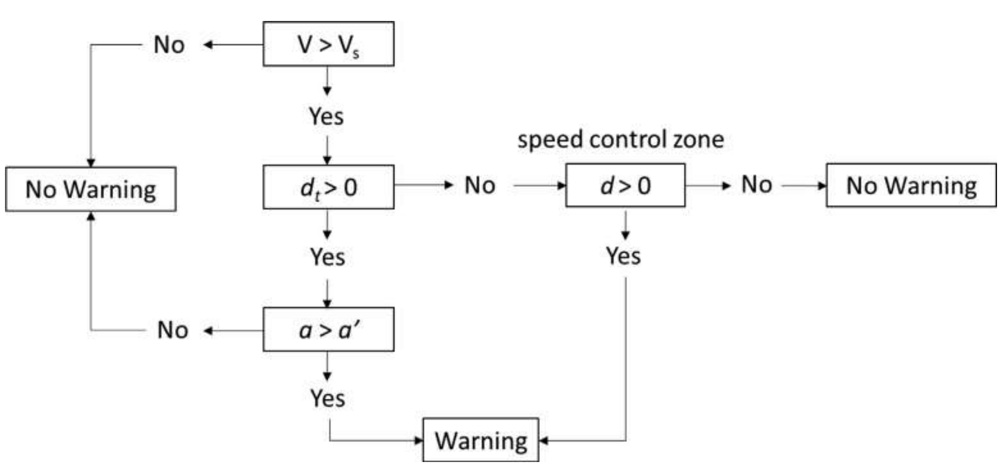

Fig. 2. Flow chart on working logic of the CSWS system. V = vehicle speed. V_s = curve maximum safe speed. d_t = distance between vehicle position and curve trigger point (trigger point = mid-point between curve entry and apex). d = distance between vehicle position and curve apex. a = deceleration required to reach V_s at curve trigger point (from equation (1) using trigger point instead of apex). a' = average deceleration (a'= 1.5 m/s^2). Speed control zone - between trigger point and curve apex.

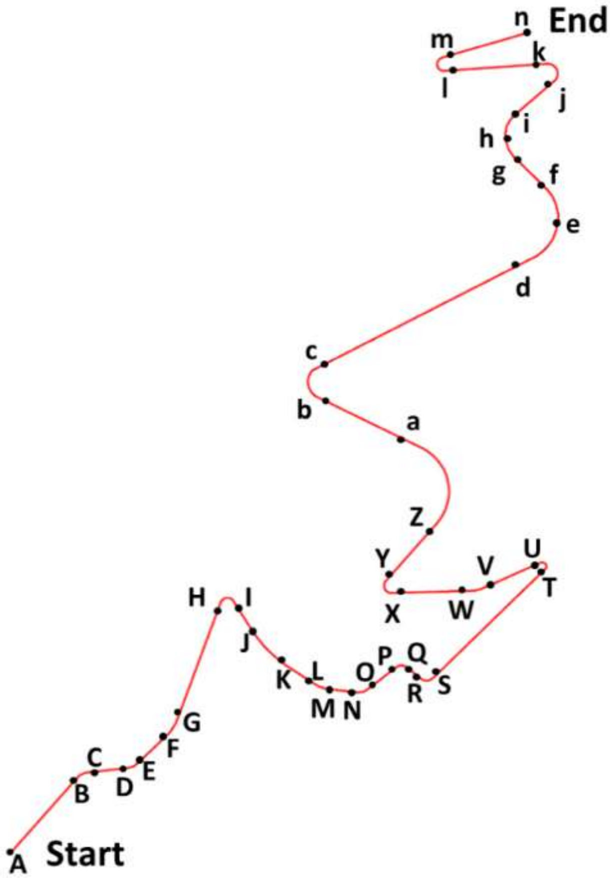

Fig. 3. Map of the test route with alphabetically indicated road segments.

gradually reducing the vehicle speed until the warnings stop. During the driving task, two additional incident-specific messages were provided to convey the emergency of the task. In case of a rollover crash event, the trial was restarted from the location of the crash.

Throughout the tests, participants were monitored closely for any symptoms of motion sickness (Hoffman et al., 2003). Rest breaks (10 min) were provided after each trial and additionally as needed at the request of study participants. After completion of the test session, the participants completed a standing balance test (semi-tandem Romberg test – Cobb, 1999; Simeonov et al., 2011) as an additional precautionary measure for no symptoms of carry-over motion sickness before they were compensated for their time and released.

2.4. Variables

2.4.1. Independent variables

The independent variables included the CSWS status and curve radius. The status of the CSWS had two levels: system "off" and system "on." The system "off" setting was used as a control or baseline and system "on" was used to evaluate the effect of the CSWS on the dependent variables. The curve radius (R) was considered in the range 46 m–196 m, including the 11 critical curves for which the safety speed was smaller than the maximum speed allowed at straight sections (V_s < 96 km/h).

2.4.2. Dependent variables

Two groups of dependent variables were used to evaluate and describe the CSWS performance and its effect on the safety outcomes (the occurrence of simulated rollover crashes) in fire tanker emergency driving.

2.4.2.1. CSWS performance variables. Warning occurrence: The warning occurrence (W_{occ}) variable indicates if any warning was issued when approaching and entering a curve. The warning occurrence is described using a binary value with 1 – indicating that a warning was activated, independent of the number or the duration of warnings issued for a specific curve, and 0 – indicating that no warnings were issued. The distance range analyzed for warning occurrences was from 200 m before the curve entry (Bella, 2014; Polus, Fitzpatrick, & Fambro, 2000) to the curve apex. The cumulative W_{occ} from all participants for each curve was expressed as a percentage (cumulative relative warning occurrence) of the total warning occurrences possible (24 for all study participants).

Intensity of Warning occurrence: The Intensity of Warning occurrence (W_{int}) was defined as the cumulative relative warning occurrence (%) per meter of the total distance with active warnings (%).

2.4.2.2. Safety outcome variable. Rollover crashes at curves: The count, location, and circumstances of rollover crashes with CSWS "on" and "off" were analyzed to obtain clues/guidance for system improvement.

2.5. Statistical analysis

Regression analyses were performed to determine the relationships of the variables warning occurrence (W_{occ}) and warning intensity (W_{int}) with curve radius (R). For the safety outcome variable "rollover crashes in curves," descriptive statistics was used to compare CSWS "on" and "off" conditions.

3. Results

3.1. CSWS performance

The cumulative warnings issued by the CSWS are displayed along with the average vehicle speed and the reference curve safety speed (V_s) for the test route in Fig. 4. The figure demonstrates that warnings were triggered by participants' driving behavior at the approaches of most curves. The cumulative warnings had a pattern culminating before the curve with different intensities for the individual curves.

3.1.1. Warning occurrence and intensity

To further characterize the performance of CSWS, the cumulative relative warning occurrence (W_{occ}) and intensity of waning occurrence (W_{int}) was regressed as a function of curve radius (R) for the 11 curves with R < 200 m. The cumulative relative warning occurrence (W_{occ}) increased with a decrease in curve radius (R), representing nearly perfect linear relation (R^2 = 0.95) (Fig. 5a). On the other hand, the intensity of warning occurrence (W_{int}), increased exponentially (R^2 = 0.86) with a decrease of curve radius (Fig. 5b).

3.2. Rollover events in curves

The emergency response tasks in the simulated driving environment resulted in a total of 19 rollover events (Table 1). Of the 19 rollover events, 10 occurred with the CSWS "off" and 9 with the system "on." The rollover events were experienced by 13 out of the 24 participants: 6 drivers had rollovers only with system "off," 4 only with system "on," and 3 drivers experienced rollovers both with system "off" and "on.".

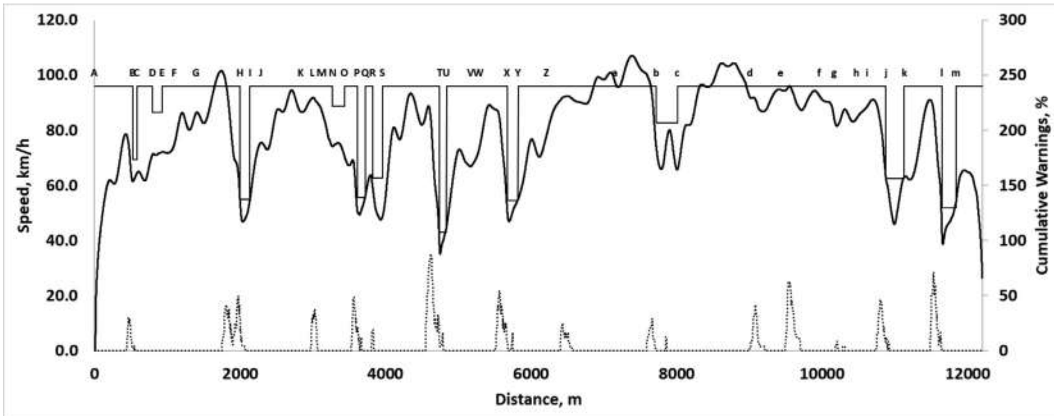

Fig. 4. Cumulative warnings (dotted line at the bottom) along the test route (plotted using secondary axis on the right) together with safety speed (V_s) (thin line at the top) and average speed (V) (thick line) (plotted using primary axis on the left).

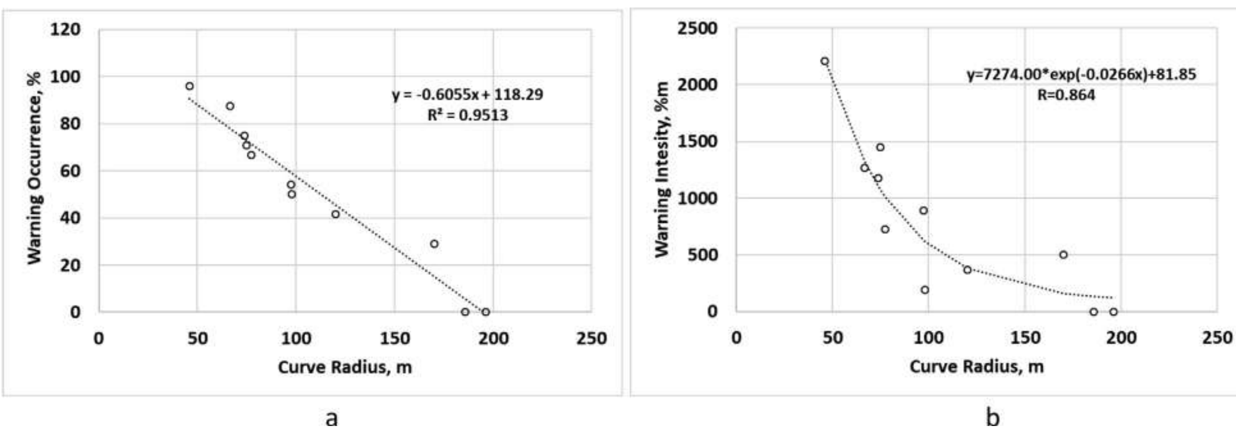

Fig. 5. Warning occurrence and warning intensity: (a) warning occurrence (W_{occ}) as a negative linear function of curve radius; (b) intensity of warning occurrence (W_{int}) as a negative exponential function of curve radius.

Table 1
Study participants who experienced rollover events during the emergency response driving trials.

Study	CSWS Off		CSWS On		Rollovers
Participant	Rollovers	Event #*	Rollovers	Event #*	Total
S1	1	#10	1	#1	2
S5	1	#13			1
S6	1	#17	2	#6, #7	3
S7	1	#14	1	#9	2
S10			2	#4, #8	2
S11	1	#11			1
S14	1	#18			1
S15	1	#16			1
S17	2	#15, #19			2
S18	1	#12			1
S20			1	#3	1
S22			1	#5	1
S24			1	#2	1
Total	10		9		19

*Event # - provided for cross-referencing with Table 2 and Fig. 7.

Rollovers occurred on 5 out of 18 curves. There was a trend for rollovers to occur at the ending section of the route (i.e., at the last two curves: "j-k" and "l-m"). Six out of 10 rollovers occurred at the last two curves when the CSWS was "off" and 8 out of 9 when the CSWS was "on" (Fig. 6). All rollovers occurred on curves with R < 100 m (safety speed drop > 30 km/h). There was also a tendency for rollovers to occur at longer curves without superelevation (the last two curves: "j-k" and "l-m"). The lengths of the curves alone, however, were not correlated with the rollover outcomes.

All rollover events (both with CSWS "on" and "off") were associated with in-curve max speed (V_{max}) at or above the curve safety speed limits (V_s) (Table 2, Fig. 7). For rollover events with CSWS system "off," there was a trend for higher curve entry speed (V_{ent}) and V_{max} as compared to rollover events with CSWS "on." The V_{max} associated with rollovers with CSWS "off" was >10% over V_s for most events, while V_{max} associated with rollovers with CSWS "on" was <10% over V_s (Table 2, Fig. 7). There was an outlier for V_{max} with CSWS "on," where the curve entry speed was 26% above the safety speed limits, in which case there was a proper warning, but the driver ignored the warning. In all other rollover events with CSWS "on," no warnings were issued by the CSWS.

In analyzing the rollovers at which the CSWS was "on" and warnings were not issued, we identified two distinct types of rollover events that were not prevented – at curve "entry" and at curve "exit" sections. In the five "entry" cases (Cases # 1–5, Table 2, and Fig. 7), the V_{max} was slightly above the safety speed limit V_s (0%-6%) within the first 20% of the entry-apex section. In all of these cases, warnings were not issued because over-speeding ($V_{max} > V_s$) was relatively small and sufficiently far from the target speed control zone for the algorithm. In three of the five cases (Cases # 3–5, Table 2, and Fig. 7), the vehicle was accelerating after the curve entry. In the three "exit" cases (Cases # 7–9, Table 2, and Fig. 7),

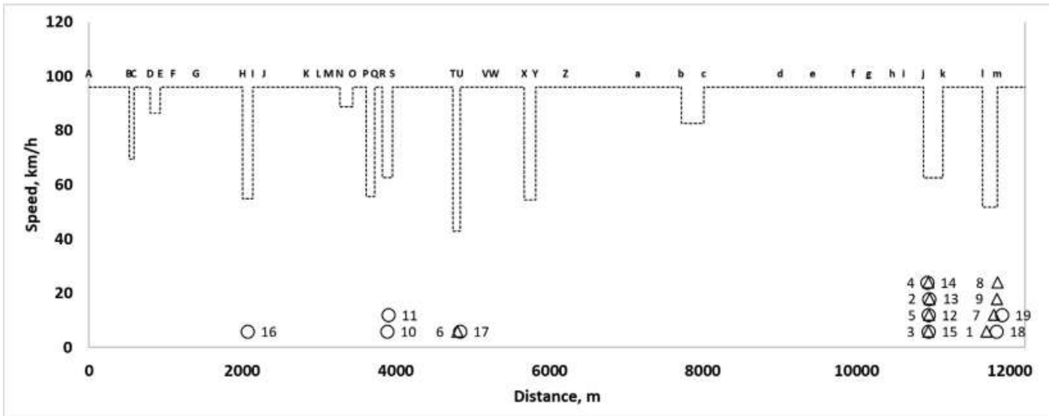

Fig. 6. Rollover events by on-route location (each symbol represents one rollover-crash occurrence: circle – system "off", triangle – system "on"). Labels next to each symbol (to the right for circles and to the left for triangles) indicate the rollover event case number as described in Table 2 and Fig. 7.

Table 2
Circumstances of rollover events and suggested measures for CSWS improvement.

Event	Curve	V_s	V_{max}	V_{max}%	Dist,%	Circumstances and Behaviors	Suggest CSWS improvements
Rollover Events with CSWS "On" in the Curve Entry-Apex Section							
1	l-m	51.7	54.8	6.0	−100	$V_{ent} > V_s$ breaking, no in curve warning	Trigger at Entry /Reduce V_s
2	j-k	62.5	64.9	3.8	−100	$V_{ent} > V_s$ coasting, no in curve warning	Trigger at Entry /Reduce V_s
3	j-k	62.5	62.3	−0.3	−82.5	$V_{ent} \sim V_s$ accelerating, no warning	Trigger at Entry /Reduce V_s
4	j-k	62.5	65.0	4.0	−82.5	$V_{ent} \sim V_s$ accelerating, no warning	Trigger at Entry /Reduce V_s
5	j-k	62.5	64.9	3.8	−80	$V_{ent} > V_s$ coast/accel, late in curve warning	Trigger at Entry /Reduce V_s
6	T-U	42.9	54.2	26.3	−100	$V_{ent} \gg V_s$ decel, adequate warning - no response	No suggested change
Rollover Events with CSWS "On" in the Curve Apex-Exit Section							
7	l-m	51.7	51.7	0.0	30.8	$V \sim V_s$ coasting, no in-curve warning	Test after Apex/Reduce V_s
8	l-m	51.7	54.0	4.4	83.3	$V > V_s$ accelerating, no in curve warning	Test after Apex/Reduce V_s
9	l-m	51.7	56.5	9.3	95.7	$V > V_s$ accelerating, no in curve warning	Test after Apex/Reduce V_s
Rollover Events with CSWS "Off" in the Curve Entry-Apex Section							
10	R-S	62.7	69.9	11.5	−100	$V_{ent} > V_s$ coasting; CSWS could help	No suggested change
11	R-S	62.7	69.1	10.2	−100	$V_{ent} > V_s$ decelerating; CSWS could help	No suggested change
12	j-k	62.5	66.0	5.6	−87.5	$V_{ent} > V_s$ coast/accel; CSWS would be too late	Trigger at Entry /Reduce V_s
13	j-k	62.5	66.8	6.9	−86.6	$V_{ent} > V_s$ accel/coast; CSWS would be too late	Trigger at Entry /Reduce V_s
14	j-k	62.5	69.7	11.5	−88.2	$V_{ent} > V_s$ decelerating; CSWS could help	No suggested change
15	j-k	62.5	72.4	15.8	−85	$V_{ent} > V_s$ coast/accel; CSWS could help	No suggested change
16	H-I	54.9	75.1	36.8	−100	$V_{ent} \gg V_s$ decelerating; CSWS could help	No suggested change
Rollover Events with CSWS "Off" in the Curve Apex-Exit Section							
17	T-U	42.9	47.0	9.6	12.5	$V > V_s$ accelerating to V_{max}; CSWS not effective	Test after Apex/Reduce V_s
18	l-m	51.7	56.1	8.5	77.3	$V > V_s$ accelerating to V_{max}; CSWS not effective	Test after Apex/Reduce V_s
19	l-m	51.7	66.9	29.4	100	$V > V_s$ accelerating to V_{max}; CSWS not effective	Test after Apex/Reduce V_s

"V_{max}" = in-curve maximum speed associated with a rollover event.
"V_{max}%" = $(V_{max}-V_s)/V_s*100$.
"Dist,%" = Distance from Apex; Curve Entry = -100%, Curve Exit = 100%.
"Reduce V_s" = Reduce Safety Speed to 0.85 of V_{roll} (currently = $0.9V_{roll}$).
"Trigger at Entry" = Move the algorithm trigger point to the curve entry point.
"Test after Apex" = Provide a warning for $V > V_s$ after Apex; and early warning for acceleration to $V > V_s$.

the rollover was associated with acceleration and over-speeding ($V_{max} > V_s$) after the curve apex, which was not a target zone for speed control and therefore no warnings were issued.

More detailed descriptive analyses for four representative cases of rollover events are presented in the next section that could be considered to improve CSWS algorithms. The case examples include rollover events for curve "Entry-Apex" section with CSWS "on" and "off" and for curve "Apex-Exit" section with CSWS "on" and "off.".

3.3. Analysis of circumstances (case study) of rollover events

Case #1: An example of a curve "Entry – Apex" rollover event with CSWS "on" is presented in Fig. 8. It represents 5 similar cases of "no adequate warning" after curve entry (cases # 1–5, Table 2, and Fig. 7). The rollover event in this case is characterized with insufficient speed reduction at curve entry and no adequate warning. The vehicle was approaching the curve at a high speed (~94.0 km/h); the driver (Participant S1) received an advanced warning (at about 150 m before the curve) by the CSWS; in response (41 m later) the driver applied the brakes (for about 60 m) and drastically reduced the vehicle speed (with ~40 km/h) to a level at which the warning stopped. The vehicle continued coasting and entered the curve at speed 54.8 km/h, which was above the curve safety speed (V_s = 51.7 km/h). Soon after entering the curve the vehicle lost control, and despite the last-minute braking, the vehicle crashed in a rollover (at 55 m into the curve). The CSWS did not issue an additional warning at the curve entry since the algorithm targeted control zone was from the mid-point of the curve-entry to apex (48.5 m) section to the curve apex (97 m), for which the required speed reduction (of 3.1 km/h) could be achieved at the nominal braking rate (1.5 m/s^2). The analysis of this

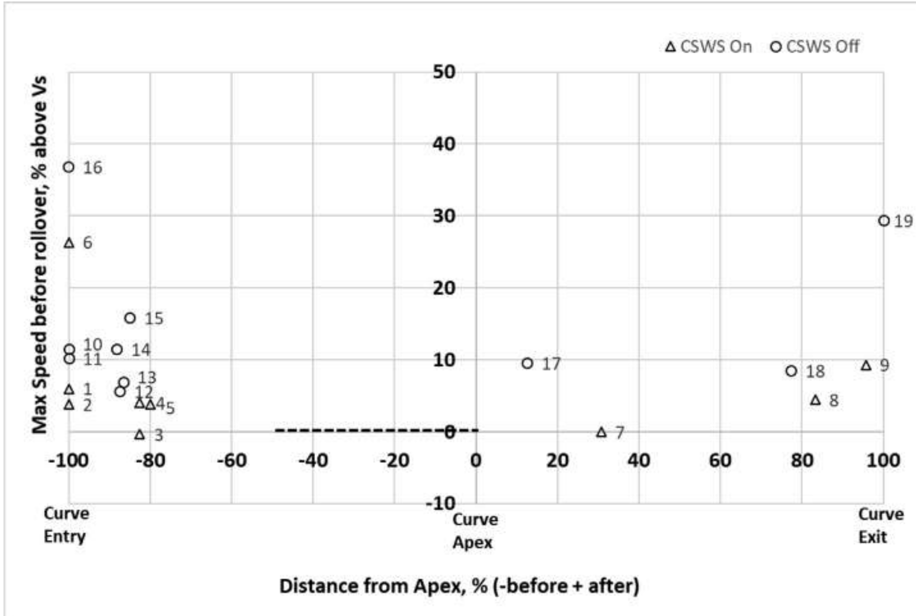

Fig. 7. Rollover event-related maximum speed (V_{max}) as % over safety speed limit (V_s) and its relative location within the curve; dashed line indicates the CSWS algorithm-targeted "speed control zone"; labels next to (the right of) each symbol (circle – CSWS "off", triangle – CSWS "on") indicate the rollover event case number as described in Table 2.

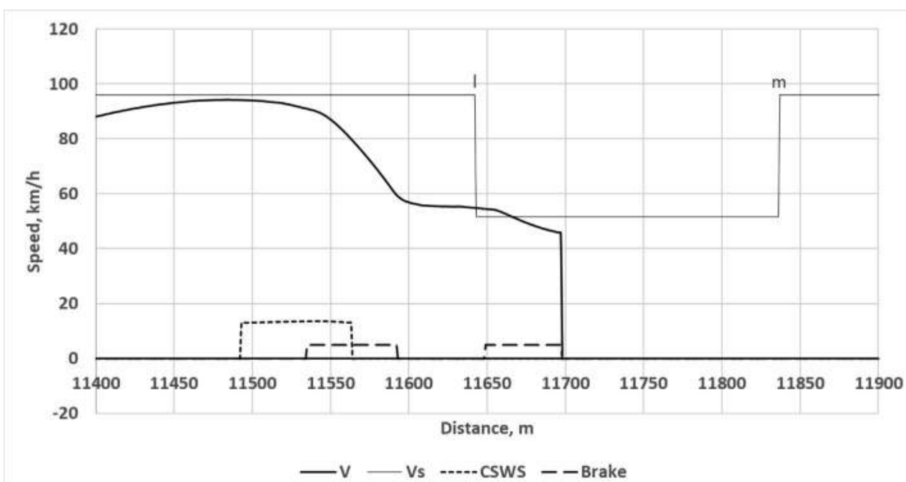

Fig. 8. Case #1: An example of an "Entry-Apex" overturn event with CSWS "on" for participant S1; the overturn occurred in the "Entry-Apex" zone of curve "l-m". Legend: thin line at the top indicates the safety speed (V_s) at the different road sections; thick line reflects the vehicle speed (V) for this trial; dashed line indicates where the brake was applied (using an arbitrary value of 5); dotted line reflects where a warning was issued by the CSWS, as a function of beeping frequency (values x5).

rollover event suggests the need for CSWS algorithm improvements by extending the targeted speed control zone (potentially to the curve entry), reducing the curve safety speed (using a safety margin > 10%, i.e., 15%), or using a combination of the two measures.

Case #7: An example of an "Apex - Exit" rollover event with CSWS "on" is presented in Fig. 9. It represents three similar cases of "no warning" after curve apex (cases # 7–9, Table 2, and Fig. 7). The rollover event in this case is characterized with accelerating after curve entry and coasting at the curve safety speed before and after the curve apex and no adequate warning. The vehicle was approaching the curve at the maximum for the straight section safety speed (~96.0 km/h). The driver (Participant S6) received an advanced warning by the CSWS (at about 163 m before the curve); in response (22 m later), the driver applied the brakes (for about 42 m) and substantially reduced the vehicle speed (with ~25 km/h) to a level at which the warning stopped. The vehicle continued coasting at ~69 km/h and upon getting closer to the curve received another short warning (at about 35 m before the curve); the driver applied the brakes (at about 33 m and all the way to curve entry) and entered the curve at speed 48.4 km/h, which is below the curve safety speed (51.7 km/h). After curve entry, the driver accelerated back to approximately the curve safety speed (51.7 km/h) and continued with this speed all the way through and beyond the curve apex. The vehicle lost control and crashed in a rollover at 148 m from curve entry, which is 51 m after the curve apex. The CSWS issued no warnings after the curve entry, since the vehicle never exceeded the curve safety speed in the speed control zone, and the algorithm did not include any control measures after the apex. This case suggests the need to

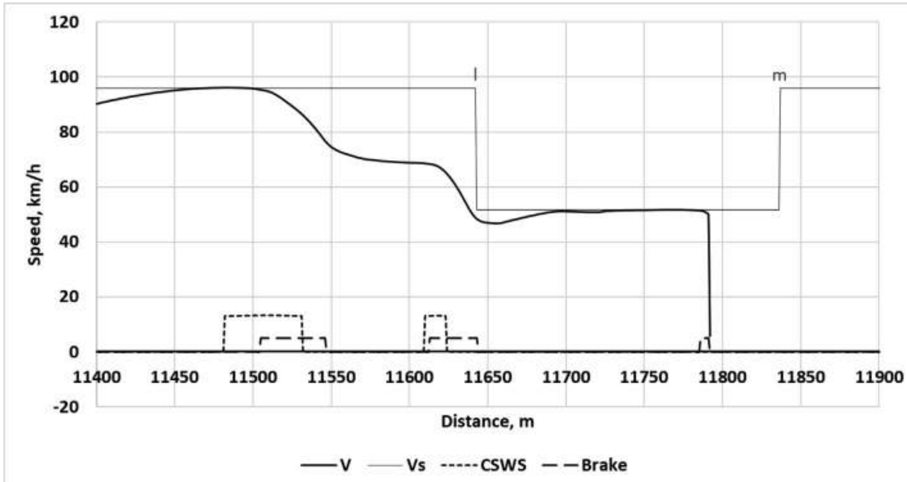

Fig. 9. Case #7: An example of an "Apex-Exit" overturn event with CSWS "on" for participant S6; the overturn occurred in the "Apex-Exit" zone of curve "l-m". Legend: thin line at the top indicates the safety speed at the different road sections (V_s); thick line reflects the vehicle speed for this trial; dashed line indicates where the brake was applied (using an arbitrary value of 5); dotted line reflects where a warning was issued by the CSWS, as a function of beeping frequency (values x5).

increase the safety speed margin to > 10% (i.e., 15%), extend the safety speed control zone beyond the curve apex and possibly all the way to the curve exit, or use a combination of these two measures. The algorithm may also be enhanced to include in-curve acceleration detection logic to predictively issue warnings if the vehicle is approaching the curve safety speed.

Case #16: An example of a curve "Entry – Apex" rollover event with CSWS "off" is presented in Fig. 10. It represents seven similar cases of entering a curve with speed higher than the curve safety speed and coasting/decelerating in which a warning could have potentially helped (cases # 10–16, Table 2, and Fig. 7). The rollover event in this case is characterized with insufficient speed reduction during curve approach and at curve entry. The vehicle was approaching the curve with a speed above the safety speed for the straight section (~100.0 km/h); in approaching the curve, the driver (Participant S15) started to reduce the speed by applying the brakes three consecutive times (at 209 m before the curve for about 49 m, at 151 m for about 23 m, and at 117 m for about 55 m) thus reducing the speed by~ 25 km/h. The vehicle continued coasting and entered the curve at ~75 km/h, which is substantially higher than the curve safety speed (54.9 km/h); shortly (8–10 m) after entering the curve, the driver applied the brakes, but lost control and the vehicle crashed in a rollover just 3.2 m before the curve apex (at 63.8 m in the curve). The existing CSWS algorithm model indicated that an active CSWS would have issued appropriate warnings and possibly prevented this rollover event.

Case #18: An example of an "Apex - Exit" rollover event with CSWS "off" is presented in Fig. 11. It represents three similar cases of acceleration above V_s within the curve after the apex (cases # 17–19, Table 2, and Fig. 7). The rollover event in this case is characterized by hard braking at curve entry followed by acceleration in the curve to above the curve safety speed after the apex. The vehicle was approaching the curve at the maximum (for the straight section) safety speed (~96.0 km/h) and coasting; in curve vicinity (47 m before curve entry), the driver (Participant S14), started applying the brakes (for about 62 m including 15 m in the curve) thus drastically reducing the speed by ~47 km/h. The vehicle entered the curve at 55.9 km/h and within 15 m in the

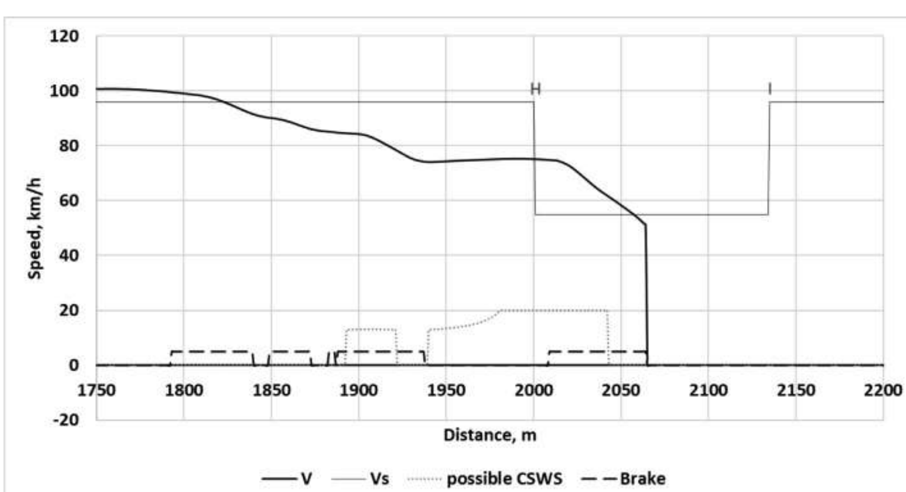

Fig. 10. Case #16: An example of an "Entry-Apex" rollover event with CSWS "off" for participant S15; the rollover occurred in the "Entry-Apex" zone of curve "H-I". Legend: thin line at the top indicates the safety speed at the different road sections (V_s); thick line reflects the vehicle speed for this trial; dashed line indicates where the brake was applied (using an arbitrary value of 5); dotted line reflects where a warning could be issued by the CSWS if it was "on", as a function of beeping frequency (values x5).

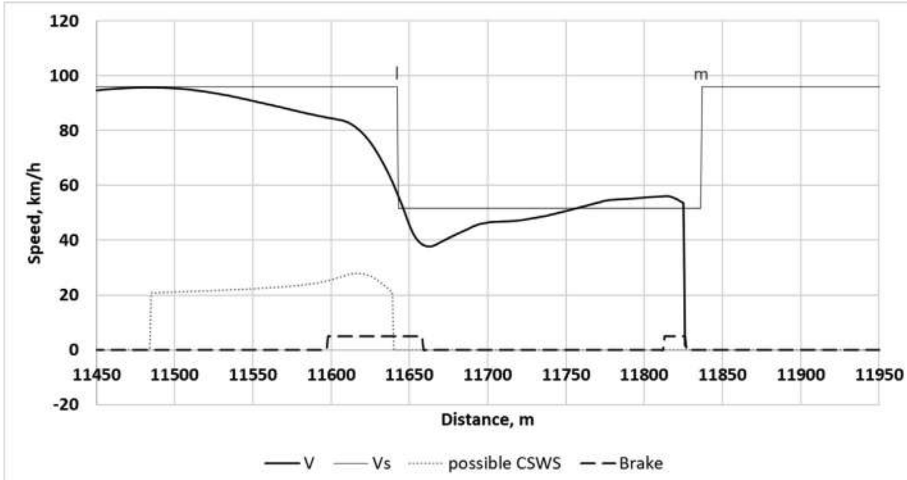

Fig. 11. Case #18: An example of an "Apex-Exit" rollover event with CSWS "off" for participant S14; the rollover occurred in the "Apex-Exit" zone of curve "l-m". Legend: thin line at the top indicates the safety speed at the different road sections (V_s); thick line reflects the vehicle speed for this trial; dashed line indicates where the brake was applied (using an arbitrary value of 5); dotted line reflects where a warning could be issued by the CSWS if it was "on", as a function of beeping frequency (values x5).

curve the speed was reduced to 37.8 km/h, which is substantially below the curve safety speed of 51.7 km/h. Immediately after that, the driver accelerated and continued accelerating after passing the curve Apex to reach speed of 56.1 km/h, well above the curve safety speed limit. At that time the driver lost control and despite the last-minute braking, the vehicle crashed in a rollover at 184 m in the curve (87 m after the curve Apex). The existing CSWS algorithm model suggested that a warning with increasing urgency (beeping frequency) could have been properly issued if the CSWS was active, and thus preventing the aggressive braking during the curve entry. The existing CSWS algorithm model, however, could not have detected the acceleration to above the safety speed after the curve Apex, indicating the need for further improvements as suggested at the end of the descriptive analysis of case #7 above.

4. Discussion

4.1. Warning system performance

When approaching curves in an emergency response mode without a CSWS, fire truck drivers are likely to speed through the curve and may underestimate the need for a timely speed reduction. Furthermore, the study results of the baseline condition (CSWS "off") confirmed that this driving behavior is independent of the curve radius (Simeonov et al., 2021). The evaluation of CSWS performance by the warning occurrence measure reflects the binary probability for a driver to receive or not receive any warning (independent of the number of warnings) when approaching a curve. The results indicate that this probability increases in a linear fashion with a decrease of the curve radius. The probability for a driver to receive a warning can be regarded as a result of the interaction between driver behavior and the CSWS algorithm. In the algorithm, the safety speed limit at a curve is a near-linear function of the curve radius.

The exponential increase of warning intensity (cumulative warning probability for each curve) associated with the decrease of curve radius further highlights the extent of driver behavior deviation from the algorithm predictions. The CSWS algorithm assumed a linear behavior in speed reduction using a constant deceleration rate to calculate the warning timing and duration. As compared to the linear CSWS prediction, drivers adjust the vehicle speed in a non-linear fashion – usually in the last moments before a sharp turn, which can trigger an increased number of warnings and with a longer total warning duration. Previous research on speed-reducing measures for curves (for passenger vehicles and non-emergency driving) has reported that clear reduction in speed in the baseline condition is not observed until approximately 100 m before a curve entry, and after this point, the deceleration becomes heavy into the curve (Comte & Jamson, 2000).

4.2. Effects of warning system on safety outcomes

The test route for this study was designed to be challenging for the fire truck drivers with many sharp turns preceded by long relatively straight sections. There was a total of 19 rollover crashes during the driving trials, of which 10 with CSWS "off" and 9 with CSWS "on." While the CSWS has been helpful in reducing the average distance traveled at speeds over the safety speed limits at the entry-to-apex section of a curve (Simeonov et al., 2021), the overall rollover counts suggested a need for further improvements in the warning system algorithms. An in-depth analysis showed that all rollovers occurred at or above the safety speed, which reflects an accurate vehicle model and safety speed calculations. Most of the rollovers in the simulation could be preventable if the vehicle speed is maintained below the safety speed limits. The tendency for more rollovers to occur later in route (8 of 9) when the system was "on" as compared to the CSWS "off" condition (6 of 10) may indicate an increased driver risk-taking behavior with over-reliance on the CSWS later in route, which also suggests the need for improvements of the CSWS.

The safety speed limit for the CSWS algorithm in this study was derived from the vehicle rollover speed ($V_s = 0.9^*V_{roll}$), in contrast to previous research using slide-out speed (Pomerleau et al., 1999). Considering the emergency response driving and the dry-road conditions where the leading crash risk for the heavy fire tanker is rollover, this setting was selected to minimize the unneeded false warnings, reduce drivers' annoyance, and improve system acceptance. However, this safety speed setting also leaves little room for errors and may be one of the causes for the observed rollovers with system "on." Furthermore, at some of the sharp curves, the driver's comfort speed limit (based on lateral acceleration) for the tested fire tanker was higher than the rollover speed, which leads to "undetectable risk" conditions for rollovers. Therefore, one possible direction for improvement of the CSWS algorithm is to lower the V_s to below 0.9*Vroll.cr levels (e.g., increase the safety margin to more than 10%, for example, 15%).

Analysis of the rollover circumstances indicated that in six (out of nine) of the rollovers with system "on," drivers lost control before the curve apex. In one case a warning was issued but ignored, and in five cases, warning was not issued since the drivers entered the curve at speeds close to the safety speed (V_s) and accelerated within the curve at a timing that did not trigger the algorithm. The existing guidelines (Pomerleau et al., 1999) set the trigger point at the curve apex. In contrast, in this study, the speed control zone for the algorithm was set to start (trigger) at mid-distance of the curve entry-apex to accommodate heavy emergency vehicles that need to reduce speed earlier in the curve (IAFF, 2010). This setting in the algorithm may need to be adjusted, since it allowed entering the curve at a speed above the safety speed and missed detecting some in-curve accelerations at speeds above V_s before the trigger point. To address this issue, the trigger point can be moved from the mid-distance of the curve entry-apex section to the curve entry in the CSWS algorithm.

In three out of nine rollovers with the CSWS system "on," the drivers lost control at locations after the curve apex, usually due to acceleration within the curve at speeds above V_s Obviously, to reduce the risk of these rollovers, the truck speed must be maintained below the safety speed limits even after the apex, especially for heavy emergency response trucks and long curves. Possible measures to reduce the risk of rollovers after the apex may include reducing the curve safety speed (as discussed above) and extending the safety speed control zone beyond the apex (or for the whole curve).

4.3. Suggested CSWS modifications, challenges, and directions for future work

The analysis of the rollover crash circumstances suggested the following modifications for improvement of the CSWS algorithm: (1) reduce the safety speed to below 0.9Vroll.cr, (2) move the trigger point to curve entry, (3) extend the safety speed-control zone beyond the apex (to possibly cover the whole curve), and (4) check for acceleration behavior to calculate and predictively issue a warning if speed will reach the safety speed limit while in curve. Implementing just one or a couple of these measures may be sufficient to reduce rollover crashes. It must be noted that some of these measures may introduce additional challenges, such as sub-optimal (too early or too late) warnings that may increase driver annoyance, reduce system acceptability, and result in compensatory (rebound) speeding behavior.

Preventing nuisance alarms is difficult because there is no commonly accepted benchmark for "correctly" negotiating a curve (Pomerleau et al., 1999). An effective way to reduce nuisance warnings is the development of an adaptive CSWS, which can account for the considerable variation in driver behavior, and specifically in speed profiles when negotiating a curve. A sophisticated adaptive CSWS may model an individual driver's curve negotiation behavior, including measures such as driver's reaction time, brake onset time, deceleration rate, and tolerance for lateral acceleration (Pomerleau et al., 1999; Ahmadi & Machiani, 2019). An innovative CSWS may also use artificial intelligence (AI) to fine-tune the optimal speed profiles by implementing learning algorithms using data from previous driving runs through specific routes of vehicles with similar dynamic characteristics and emergency response tasks.

4.4. Limitations

The use of a driving simulator as a modeling research tool in this study is associated with several limitations, related to the fidelity of the simulation and the previous gaming experience of the participants. The benefits of using driving simulations in vehicle safety research are well known. Simulators allow for a better control over the experimental conditions, lower expense and better efficiency, improved safety for participants and researchers, and convenience in data collection. However, driving simulations also have some major limitations, such as lack of realism associated with low-risk perception, limited physical laws (i.e., lack of appropriate vestibular and motion cues), moderate behavioral validity, and potential motion sickness (Godley, Triggs, & Fildes, 2002; Nilsson, 1993). Nevertheless, driving simulations have been validated for generating and generalizing relative speed in testing road-based speeding countermeasures (Godley et al., 2002) and in studies on curve negotiation in two-lane rural roads (Bella, 2008). However, previous research has shown that in a driving simulator, participants initiate braking later and brake much harder as compared to real roads (Boer, Girshick, Yamamura, & Kuge, 2000); and, in a driving simulator as compared to real roads, the curve-entry speed was faster in less challenging curves ($R > 582$ m) and slower in the most difficult curves ($R < 146$ m) (Bittner, Simsek, Levison, & Campbell, 2002).

Participants with extensive car-racing gaming experience may have been more used to simulated driving environments and drive more aggressively as compared to the average person. This study did not assess the gaming car-racing experience of the participants. However, during the tests some of the participants provided comments on their gaming car-racing experience (and real-life car-racing experience) and their perceptions of how it may have affected their performance. In video-racing games, drivers are reinforced for driving recklessly and systematically breaking traffic rules, and as a result, video-racing gaming experience may increase risk-taking driving behaviors (Fischer et al., 2009).

This study used a balanced experimental design in which all participants performed all the experimental conditions in a balanced order, which should help in cancelling out most of the effects of the abovementioned limitations.

5. Conclusions

A curve speed warning system (CSWS) was tested in this study for its effectiveness in preventing fire tanker rollover crashes at risky curves during simulated emergency response driving. The results demonstrated that the CSWS was effective in issuing preemptive warnings when drivers were approaching curves at an unsafe speed. Warnings were more likely to occur at curves with radius smaller than 200 m and occurred more often at sharper curves. There is limited information on developing a warning system for heavy emergency vehicles in preventing curve speed-related rollover crashes. While the CSWS algorithms tested in this simulation study did not show a significant reduction in the number of rollover crashes, the study results provided valuable suggestions for improving the CSWS algorithm, including programming a safety speed with a more than 10% safety margin below the rollover critical speed, establishing a speed warning trigger at the curve entry point instead of the mid-point between the curve entry and apex, extending the safe speed-control zone to cover the entire curve beyond the apex point, and employing artificial intelligence technologies to accommodate individuals with different driving styles.

6. Practical Applications

Fire tankers are top-heavy vehicles that continue to be at increased risk of rollover during emergency response due to unsafe negotiation of dangerous curves. Development and use of advanced driver assist systems such as CSWS evaluated in this study may be an effective strategy to prevent deadly rollover crash-incidents. The knowledge generated by this study will be useful for system designers to improve the CSWS specifically designed for heavy emergency vehicles.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Institute for Occupational Safety and Health (NIOSH), Centers for Disease Control and Prevention (CDC). Mention of any company or product does not constitute endorsement by NIOSH or CDC.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahmadi, A., & Ghanipoor Machiani, S. G. (2019). Drivers' performance examination using a personalized adaptive curve speed warning: Driving simulator study. *International Journal of Human-Computer Interaction, 35*(11), 996–1007.
Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. *Accident Analysis and Prevention, 40*, 1078–1087.
Bella, F. (2014). Driver Performance Approaching and Departing Curves: Driving Simulator Study. *Traffic Injury Prevention, 15*(3), 310–318.
Bittner, A. C., Jr., Simsek, O., Levison, W. H., & Campbell, J. L. (2002). On-road versus simulator data in driver model development driver performance model experience. *Transportation Research Record Journal. Transportation Research Board, 1803*, 38–44.
Boer, E. R., Girshick, A. R., Yamamura, T., & Kuge, N. (2000). Experiencing the Same Road Twice: A Driver Centered Comparison between Simulation and Reality. *DSC 2000: Driving Simulation Conference. Conference Proceedings. Paris, France.*.
Chowdhury, S., Faizan, M., & Hayee, M. (2020). Advanced Curve Speed Warning System using Standard GPS Technology and Road-level Mapping Information. In *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2020)* (pp. 464–472).
Cobb, S. V. G. (1999). Measurement of postural stability before and after immersion in a virtual environment. *Applied Ergonomics, 30*, 47–57.
Comte, S. L., & Jamson, A. H. (2000). Traditional and innovative speed reducing measures for curves: An investigation of driver behavior using a driving simulator. *Safety Science, 36*(3), 137–150.
Campbell, R., & Evarts, B. (2020). *United States Firefighter Injuries in 2019*. Quincy, MA: National Fire Protection Association (NFPA) Research.
Federal Emergency Management Agency (FEMA), 2003. Safe operation of fire tankers. Federal Emergency Management Agency, US Fire Administration.
Fischer, P., Greitemeyer, T., Morton, T., Kastenmüller, A., Postmes, T., Frey, D., ... Odenwälder, J. (2009). The Racing-Game Effect: Why Do Video Racing Games Increase Risk-Taking Inclinations? *Personality and Social Psychology Bulletin, 35*, 1395–1409.
Fitzharris, M., Stephan, K., Newstead, S., Truong, J., Healy, D., Collins, S., & Rowe, G. (2011). Final analysis of the TAC ISA heavy vehicle trial: effects of ISA and fuel efficiency training on speed choice. *Australasian Road Safety Research, Policing and Education Conference, 2011, Perth, Western Australia, Australia*.
Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis and Prevention, 34*, 589–600.
Gonzalez, C., Lewis, B. A., Roberts, D. M., Pratt, S. M., & Baldwin, C. L. (2012). Perceived urgency and annoyance of auditory alerts in a driving context. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (vol. 56 (1), pp. 1684–1687). Sage CA: Los Angeles, CA: Sage Publications.
Hoffman, R. B., Molino, J. A., & Inman, V. W. (2003). Driving simulator sickness management at Turner-Fairbank Highway Research Center. *Proceedings of the Driving Simulation Conference North America 2003, Dearborn, Michigan, October 8–10th*.
International Association of Fire Fighters (IAFF) (2010). *Best Practices for Emergency Vehicle and Roadway Operations Safety in the Emergency Services*. Washington, DC: International Association of Fire Fighters.
Jimenez, F., Liang, Y., & Aparicio, F. (2012). Adapting ISA system warnings to enhance user acceptance. *Accident Analysis and Prevention, 48*, 37–48.
National Fire Protection Association (NFPA) (2008). *NFPA1901: Standard for Automotive Fire Apparatus* (2009 Edition). Quincy, MA: National Fire Protection Association.
Nilsson, L. (1993). Behavioural research in an advanced driving simulator – experiences of the VTI system. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting.* Vol. 37, no. 1 pp. 612–616.
National Institute for Occupational Safety and Health (NIOSH) (2002). Fire Fighter Deaths from Tanker Truck Rollovers. *Applied Occupational and Environmental Hygiene, 17*(9), 598–599.
Polus, A., Fitzpatrick, K., & Fambro, D. (2000). Predicting Operating Speeds on Tangent Sections of Two-Lane Rural Highways. *Transportation Research Record, 1737*, 50–57.
Pomerleau, D., Jochem, T., Thorpe, C., Batavia, P., Pape, D., Hadden, J., ... Everson, J. (1999). Run-off-road collision avoidance using IVHS countermeasures. *DOT HS, 809*(170), 88.
Simeonov, P., Hsiao, H., Nimbarte, A., Current, R., Ammons, D., Choi, H.-S., ... Weaver, D. (2021). Evaluation of advanced curve speed warning system for fire trucks. *Applied Ergonomics, 97*, 103527.
Simeonov, P., Hsiao, H., Powers, J., Ammons, D., Kau, T. Y., & Amendola, A. (2011). Postural stability effects of random vibration at the feet of construction workers in simulated elevation. *Applied Ergonomics, 42*(5), 672–681.
Wang, S., Wang, Y., Zheng, Q., & Zhixia, L. (2020). Guidance-oriented advanced curve speed warning system in a connected vehicle environment. *Accident Analysis and Prevention, 148*, 105801.
Winkler, C. (2000). Rollover of Heavy Commercial Vehicles. UMTRI Research Review, Vol. 31, No. 4. Research Information and Publications Center of the University of Michigan Transportation Research Institute (UMTRI), Ann Arbor, Michigan.
Young, K. L., Regan, M. A., Triggs, T. J., Jontof-Hutter, K., & Newstead, S. (2010). Intelligent speed adaptation – effects and acceptance by young inexperienced drivers. *Accident Analysis and Prevention, 42*, 935–943.

Peter Simeonov is a Research Safety Engineer in the Protective Technology Branch, Division of Safety Research, National Institute for Occupational Safety and Health (NIOSH). He received his Ph.D. in Technical Sciences from the Bulgarian Academy of Sciences in 1987.

Ashish Nimbarte is a Professor in the Department of Industrial and Management Systems Engineering, West Virginia University. He received his Ph.D. in Engineering Science (Industrial Engineering) from Louisiana State University in 2009.

Hongwei Hsiao is a Professor in the School of Engineering and Computing Sciences, Texas A&M University-Corpus Christi and former chief of the Protective Technology Branch in the Division of Safety Research at the National Institute for Occupational Safety and Health (NIOSH). He received his Ph.D. in Industrial Engineering from the University of Michigan in 1990.

Richard Current is a Research General Engineer in the Protective Technology Branch in the Division of Safety Research at the National Institute for Occupational Safety and Health. He received his B.S. in Aerospace Engineering from West Virginia University in 1990.

Douglas Ammons is a Computer Engineer in the Protective Technology Branch in the Division of Safety Research at the National Institute for Occupational Safety and Health. He received his B.S. in Computer Engineering from West Virginia University in 1999.

Darlene Weaver is a Technical Information Specialist in the Protective Technology Branch in the Division of Safety Research at the National Institute for Occupational Safety and Health. She received his M.S. in Occupational Safety and Health Engineering from West Virginia University in 1996.

HeeSun Choi is an Assistant Professor in the Department of Psychological Sciences at Texas Tech University and former Associate Service Fellow in the Division of Safety Research at the National Institute for Occupational Safety and Health. She received her Ph.D. in Human Factors and Applied Cognition from North Carolina State University in 2016.

Md Mahmudur Rahman is a Visiting Assistant Professor in the School of Industrial Engineering at Purdue University and former Associate Service Fellow in the Division of Safety Research at the National Institute for Occupational Safety and Health. He received his Ph.D. in Industrial and Systems Engineering from Mississippi State University in 2016.

Journal of Safety Research 83 (2022) 371–378

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Examining associations between work-related injuries and all-cause healthcare use among middle-aged and older workers in Canada using CLSA data

Shahin Shooshtari [a,*], Verena Menec [a], Brenda M. Stoesz [b], Dimple Bhajwani [c], Nick Turner [d], Caroline Piotrowski [a]

[a] Department of Community Health Sciences, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada
[b] The Centre for the Advancment of Teaching and Learning, University of Manitoba, Winnipeg, Manitoba, Canada
[c] College of Medical Rehabilitation, University of Manitoba, Winnipeg, Manitoba, Canada
[d] Haskayne School of Business, University of Calgary, Calgary, Alberta, Canada

A R T I C L E I N F O

Article history:
Received 21 January 2022
Received in revised form 1 June 2022
Accepted 20 September 2022
Available online 29 September 2022

Keywords:
Aging population
Canada
CLSA
Healthcare utilization
Work-related injuries

A B S T R A C T

Introduction: Prior studies examining the relationship between work- related injuries and healthcare use among middle-aged and older workers were mainly cross-sectional and reported inconsistent results. *Objective:* The objective of this study was to examine the associations between work-related injuries and 10 types of healthcare service use for any cause among middle-aged and older Canadian workers using longitudinal data. *Methods:* Our study involved longitudinal analysis of baseline and 18-month follow-up Maintaining Contact Questionnaire data from the Canadian Longitudinal Survey on Aging (CLSA) for a national sample of Canadian males and females aged 45–85 years who worked or were recently retired ($N = 24{,}748$). *Results:* Among CLSA participants who worked or were recently retired, 361 per 10,000 reported a work-related injury within the year prior to the survey. Work-related injuries decreased with increasing age. Work-related injury was associated with emergency department visits, overnight hospitalization, visits to dentists, and visits to physiotherapists, occupational therapists, or chiropractors at follow-up in bivariate analyses. Compared to those with no work-related injuries, Canadians with work-related injuries had used, on average, a significantly higher number of health services within the last 12 months prior their survey. When controlling for the contribution of various socio-demographic, work-related, and health-related characteristics, work-related injuries remained a significant predictor of emergency department visits and visits to physiotherapists, occupational therapists, or chiropractors. *Conclusions:* The relationship between work-related injuries, emergency department visits, and visits to physiotherapists, occupational therapists, or chiropractors in middle-aged and older workers in Canada suggests that workplace injuries can be associated with ongoing health problems. *Practical Applications:* Healthcare services used by injured employees must be considered priorities for employment insurance coverage, if not already covered. Future research should more fully examine whether pre-existing health conditions predict both work-related injury and subsequent health problems. Injury-specific healthcare use following work-related injuries in middle-aged and older workers, as well as economic costs, should also be examined.

1. Introduction

Injuries are a serious public health concern in Canada and worldwide. The economic burden of all types of injuries, including costs associated with healthcare expenditures, and reduced productivity due to hospitalization, disability, and premature death, is approximately $26.8 billion per year (or 1.7% of the Gross Domestic Product [GDP]) in Canada (Parachute, 2015) and 4% of the average global GDP (International Labour Organization (ILO), 2005). Although information on work-related incidence and fatalities for Canadian workers by region is available (Morassaei & Breslin, 2013; Tucker & Keefe, 2019), less is known about post-injury treatment seeking among middle-aged and older workers.

* Corresponding author at: 219 Human Ecology Building, Department of Community Health Sciences, Rady Faculty of Health Sciences, Uiversity of Manitoba, Winnipeg, MB R3T 2N2, Canada.

E-mail address: Shahin.shooshatri@umanitoba.ca (S. Shooshtari).

https://doi.org/10.1016/j.jsr.2022.09.012
0022-4375/

The association between work-related injuries and healthcare use (e.g., general practitioner [GP] or emergency department visits, overnight hospitalizations) is an indicator of the presence of ongoing problems resulting from injuries, and an indirect measure of the economic costs of workplace injuries (Brown, McDonough, & Mustard, 2006). Examining the relationship between work-related injuries and healthcare use is important as a significant proportion of the working population in Canada is aged 55+ years (36%), which is predicted to increase to 40% by 2026 (Statistics Canada, 2017a), with similar increases expected worldwide (White, Burns, & Conlon, 2018). Increases in the age of the working population are due to a number of factors, including increased life expectancy and financial necessity (Statistics Canada, 2017b).

A recent systematic review (Stoesz, Chimney, & Deng, 2020)revealed that work-related injuries were less common among older (aged 40+ years) than younger workers, but that older workers experienced more severe work-related injuries (Chen, Chakrabarty, & Levine, 2013; Frickmann, Wurm, & Jeger, 2012; Konstantinidis, Talving, & Kobayashi, 2011), more adverse health and economic outcomes (Scott, Liao, & Fisher, 2018), and higher healthcare costs (Algarni, Gross, & Senthilselvan, 2015). The review also highlighted evidence of unmet healthcare needs. For example, older workers (aged 65+ years) in Alberta, Canada, were less likely than younger workers (aged 25–54 years) to be offered rehabilitation services and the delay between their work injuries and the initial rehabilitation assessments was significantly longer (Algarni et al., 2015). Health seeking behaviors following workplace injuries have also been reported in other jurisdictions (Berdahl & Zodet, 2020; Brown et al., 2006; Chen et al., 2013; Ruseckaite, Collie, & Prang, 2016). For example, using administrative and compensation data in British Columbia, Canada, Brown et al. (Brown et al., 2006)-found more GP visits, mental healthcare services, and hospitalizations among lost-time injured workers five years post-injury compared to one year pre-injury, and these rates were higher than those observed for non-injured workers. Administrative and compensation data in Australia showed that one third of injured workers of all ages visited GPs, which was followed by hospitalization services, physiotherapy, and psychologists 12 months post work-related injury, and that rates were higher with increasing age for some healthcare services (i.e., GP visits, physiotherapy) (Ruseckaite et al., 2016). However, some studies failed to find evidence of increases in healthcare use following work-related injuries. McCaig, Burt, and Stussman (1998), for example, did not find a significant difference in the proportion of emergency department visits between those with and without work-related injuries using the 1995–1996 National Hospital Ambulatory Medical Care Survey data. Other studies have reported the proportion of the study population who were working and experienced work-related injuries that led to utilization of specific health services (e.g., visits to emergency departments), but did not compare these with workers without work-related injuries (Tonozzi & Layne, 2016).

Research also documents other risk factors for increased healthcare use following work-related injuries. Females are more likely than males to use healthcare services following work-related injuries (Ruseckaite et al., 2016). U.S. emergency services data showed that the rate of occupational knee injuries increased from 10 to 15 per 10,000 for females ages 25–34 and 50–54 years, respectively, but rates decreased with age for males (Chen et al., 2013). In addition, individuals living in rural areas use healthcare services less often following workplace injury; this may be due to decreased availability of services compared to urban areas and increased traveling time (Ruseckaite et al., 2016; Young, Cifuentes, & Wasiak, 2009).

Working hours and type and severity of injuries have also been linked to healthcare service use among workers. For example, in Japan, significantly higher rates of physician visits were found for male workers aged 18–65 years who worked 100–200 hours per month compared to those who worked 201–250 hours per month (Sato, Yamazaki, & Hayashino, 2011). However, they were less likely to use over the counter medication. Individuals who worked more than 250 hours per month were also less likely to use over the counter medication, as well as dietary supplements for symptoms experienced than those working 201–250 hours per month. The authors concluded that full-time work might act as a barrier to access required health services. In terms of specific relationships between work-related injuries and post-injury treatment seeking, studies in the United States and Switzerland have shown that machine- and fall-related injuries and increased severity of injury are significantly associated with increased hospitalization rates (Pfortmueller, Kradolfer, & Kunz, 2013; Tadros, Sharon, & Chill, 2018).

Results of many studies examining healthcare use following workplace injuries cannot be generalized to the broader working population as these studies have focused on particular occupational categories (Nilsson, Pinzke, & Lundqvist, 2010; Pfortmueller et al., 2013; Tonozzi & Layne, 2016; Weigel, Armijos, & Beltran, 2014), or were based on administrative data (medical or workers' compensation claims), in which the incidence of underreporting of workplace injuries is high (Tucker & Keefe, 2019). In addition, previous studies were mostly cross-sectional in design (Berdahl & Zodet, 2020; McCaig et al., 1998; Ruseckaite et al., 2016; Seo, Chao, & Yeung, 2019; Weigel et al., 2014). To the best of our knowledge, only one Canadian study has used a retrospective longitudinal design based on the linked Medical Service Plan, hospital discharge and Workers Ccompensation Board (WCB) data for workers in the Canadian province of British Columbia to examine the link between work-related injuries and healthcare utilization (Brown et al., 2006).

To address methodological limitations and observed inconsistencies reported in the literature, we used national-level longitudinal data from the Canadian Longitudinal Study on Aging (CLSA) (Canadian Longitudinal Survey on Aging (CLSA), 2020; Raina et al., 2018, 2019) to examine the associations between work-related injuries and the utilization of 10 types of healthcare services including: (1) emergency department visits; (2) overnight hospitalizations; (3) nursing home or convalescent home use; (4) family doctor visits; (5) medical specialist visits; (6) psychologist visits; (7) optometrist visits; (8) visits to a physiotherapist, occupational therapist, or chiropractor; (9) visits to a social worker; and (10) dentist visits among middle-aged and older Canadian workers. We selected these 10 types of healthcare services as they were studied in relation to work-related injuries in previous research with inconsistent results, and were available in the CLSA data. We controlled for the contribution of various socio-demographic and health-related characteristics, which may relate to patterns of utilization.

2. Methods

2.1. Study design and data sources

We conducted longitudinal analyses using data from the CLSA (Canadian Longitudinal Survey on Aging (CLSA), 2020; Raina et al., 2018, 2019), which consists of two cohorts. The *Tracking Cohort* (N = 21,241) consists of a randomly (within age/sex strata) selected sample of Canadian males and females in 10 provinces, aged 45–85 years, who completed a *Baseline Questionnaire* administered via computer-assisted telephone interviews (CATI). The *Comprehensive Cohort* (N = 30,097) were randomly (within age/sex strata) selected Canadian males and females living near one of 11 data collection sites located across Canada. The Comprehensive Cohort completed the *Baseline Questionnaire* during in-home interviews; they also provided physical, biological, and clinical data at data collection sites. Approximately 18 months after the baseline, both cohorts (N = 51,338) completed the *Maintaining Contact Questionnaire* via CATI, which included questions on healthcare use. Individuals living in long-term care institutions, those with cognitive impairment, residents in the Canadian territories, individuals living on federal First Nations reserves, full-time members of the Canadian Armed Forces, unable to respond in English or French were excluded from participation in CLSA. For this study, *Baseline* and *Maintaining Contact Questionnaire* data for a subsample of the CLSA were examined.

2.2. Study sample

Given the present focus on work-related injuries, this study was based on a subsample of CLSA participants who were either still employed (full-time or part-time) at the time of the Baseline interview, or had recently retired. Recently retired participants were defined as those who reported having retired within a 12-month period prior to the date of the interview. This study was, therefore, based on data for 24,748 participants (48.2% of the full sample).

2.3. Study measures

2.3.1. Socio-demographic characteristics

Socio-demographic measures including age (45–54 years, 55–85 years), sex (male/female), marital status (married/living with a partner in a common law relationship, single/never married/never living with a partner/widowed/divorced/separated), education (<secondary graduation/secondary graduation/no post-secondary, some post-secondary/post-secondary degree/diploma), and total annual household income (less than $20,000/$20,000-$49,999/$50,000-$90,000/$100,000–149,999/$150,000 or more) were collected from CLSA participants.

2.3.2. Work-related injuries

In the Baseline Questionnaire, participants were asked the yes/no question, "In the last 12 months, have you had any injuries that were serious enough to limit some of your normal activities? For example, a broken bone, a bad cut or burn, a sprain or a poisoning." Those who responded "yes" were asked to indicate the causes of their injuries: (1) Fall, (2) Motor Vehicle Collision (including injuries sustained as a pedestrian), (3) An incident in your workplace, (4) None of the above, and (5) Don't know/No answer/Refused. Participants were then asked what type of activity they were doing when they were injured: (1) Sports or physical exercise (include school activities); (2) Leisure or hobby; (3) Working at a job or business (including travel to or from work); (4) Household chores, other unpaid work or education; (5) Sleeping, eating, personal care; or (6) Other. Using these two variables, those who reported an activity-limiting injury caused by "an accident in their workplace" or "working at a job or business including travel to or from work" were classified as having a work-related injury.

2.3.3. Work-related variables

2.3.3.1. Current working status. Respondents who indicated that they were currently working were asked, "What is your current working status? If you are self-employed, choose full-time or part time, as appropriate." Responses were classified into *full-time* or *part-time employee*.

2.3.3.2. Current work schedule. Respondents were asked, "Which of the following best describes your working schedule?: (1) Daytime schedule or shift; (2) Evening shift; (3) Night shift; (4) Rotating shift, changing periodically from days to evenings or nights; and (5) Seasonal, on-call or casual, no pre-arranged schedules, but called as need arises." Responses were classified into *Working daytime* or *Other types of work* schedules or shifts.

2.3.4. Vision

Respondents were asked, "Is your eyesight, using glasses or corrective lens if you use them excellent, very good, good, fair, or poor?" Responses were coded as Excellent/Very good/Good, or Fair/Poor/non-existent (blind).

2.3.5. Mood disorder

Respondents were asked, "Has a doctor ever told you that you have a mood disorder, such as depression (including manic depression), bipolar disorder, mania, or dysthymia?" Responses were coded as having or not having a mood disorder.

2.3.6. Memory problem

Respondents were asked if they were diagnosed with memory problems by a health professional. Responses were coded as *yes* or *no*.

2.3.7. Smoking

Participants were asked about their smoking behaviors during the past 12 months. Responses to smoking behavior questions were classified into the following categories: (1) Daily smoker; (2) Occasional smoker, but former daily smoker; (3) Occasional smoker; (4) Former daily smoker, but non-smoker now; and (5) Never smoked. In this study, we used a binary variable to differentiate: (1) Occasional smoker/Former daily smoker, but non-smoker now/Never smoked from those who were, from (2) Daily smoker/Occasional smoker, but former daily smokers. This classification was used to differentiate those with current heavier smoking behavior from others.

2.3.8. Health care utilization

In the *Maintaining Contact Questionnaire*, participants were asked: "During the past 12 months, have you had contact with any of the following about your physical or mental health? (a) Family doctor? (b) Medical specialist (such as a cardiologist, gynecologist, psychiatrist, or ophthalmologist)? (c) Psychologist? (d) Optometrist? (e) Physiotherapist, occupational therapist, or chiropractor? (f) Social worker." From this list, we identified those participants who had seen a family doctor (vs those who had not) and those who had seen a psychologist (vs not). Respondents were also asked: "Have you been seen in an Emergency Department during the past 12 months?", "Were you a patient in a hospital overnight during the past 12 months?", and "Were you a patient in a nursing home or convalescent home during the past 12 months?" We created 10 binary variables (yes/no) to indicate the use of each type of healthcare. In addition, we defined a new index variable indicating the sum of the 10 types of healthcare use to be able to examine and compare utilization of multiple types of healthcare within the same timeframe between those with and without work-related injuries.

2.4. Data analysis

We conducted descriptive analyses to describe the socio-demographic, work-related, and health-related characteristics of our study population. Sampling weights were applied to obtain population estimates. We used cross-tabulations to test associations between each study factor and work-related injuries. We also tested the bivariate association between work-related injuries and the 10 types of health services use. The average number of various types of healthcare utilization were estimated and compared

Table 1
Socio-demographic Characteristics of the Study Population.

Variable	Canada			
	Weighted *n*	%	99% CI	
Age (years)				
45–54	4,203,803	58.1	56.8	59.3
55–85	3,036,649	41.9	40.7	43.2
Total	7,240,452	100.0		
Sex				
Female	3,438,555	47.5	46.2	48.8
Male	3,801,897	52.5	51.2	53.8
Total	7,240,452	100.0		
Marital status				
Single, never married or never lived with a partner/Widowed/Divorced/Separated	1,493,185	20.6	19.7	21.6
Married/ Living with a partner in a common-law relationship	5,745,164	79.4	78.4	80.3
Total	7,238,350	100.0		
Education				
Less than secondary school graduation/Secondary school graduation, no post-secondary	1,087,650	15.1	14.1	16.0
Some post-secondary education/Post-secondary degree/diploma	6,136,157	84.9	84.0	85.9
Total	7,223,808	100.0		
Total household income				
Less than $50,000	1,097,548	15.8	83.3	85.1
$50,000 or more	5,861,597	84.2	14.9	16.7
Total	6,959,145	100.0		

between those with and without work-related injuries among our study population. Finally, we ran several multivariate regression models to examine the independent contribution of work-related injuries on subsequent health care use. Adjusted odd ratios (AOR) and their 99% confidence intervals (CIs) were used to identify significant predictors of subsequent health care use among middle-aged and older workers in Canada. Analytical weights that adjust for the CLSA complex survey design were applied for any statistical

Table 2
Comparing Profiles of the CLSA Participants With and Without Work-Related Injuries.

Variable	**Work-related injuries**		x^2
	Yes % (99% CI)	**No % (99% CI)**	
Age (years)			
45–54	63.5 (56.9, 70.2)	57.9 (56.6, 59.1)	13.4*
55–85	36.5 (29.8, 43.1)	42.1 (40.9, 43.4)	
Sex			
Male	57.9 (50.9, 64.9)	52.3 (51.0, 53.6)	3.6
Female	42.1 (35.1, 49.1)	47.7 (46.4, 49.0)	
Marital status			
Single, never married or never lived with a partner/Widowed/ Divorced/Separated	21.7 (16.5, 26.9)	20.6 (19.6, 21.6)	11.0*
Married/ Living with a partner in a common-law relationship	78.3 (73.1, 83.5)	79.4 (78.4, 80.4)	
Education			
Less than secondary school graduation/Secondary school graduation, no post-secondary	16.9 (11.8, 22.1)	15.0 (14.0, 16.0)	6.8*
Some post-secondary education/Post-secondary degree/ diploma	83.1 (77.9, 88.2)	85.0 (84.0, 86.0)	
Annual household income			
Less than $50,000	20.7 (15.0, 26.4)	15.6 (14.7, 16.5)	19.7*
$50,000 or more	79.3 (73.6, 85.0)	84.4 (83.5, 85.3)	
Current working status			
Full time employee (30 + hours/week)	86.1 (81.2, 91.0)	81.6 (8.5, 82.7)	7.1*
Part time employee	13.9 (9.0, 18.9)	18.4 (17.3, 19.5)	
Current work schedule			
Daytime schedule or shift	67.9 (60.5, 75.4)	82.7 (81.6, 83.8)	43.5*
Evening/night/rotating shift, seasonal, on call, or casual but called as need arises	32.1 (24.6, 39.5)	17.3 (16.2, 18.4)	
Vision			
Excellent/ Very good/ Good	89.2 (84.8, 93.6)	93.6 (92.9, 94.3)	20.8*
Fair/Poor or non-existent (blind)	10.8 (6.4, 15.2)	6.4 (5.7, 7.1)	
Smoking			
Occasional smoker/Former daily smoker, but non-smoker now/Never smoked	84.3 (79.3, 89.4)	90.0 (89.2, 90.8)	26.6*
Daily smoker/Occasional smoker (former daily smoker)	15.7 (10.6, 20.7)	10.0 (9.2, 10.8)	
Mood disorder			
Yes	11.6 (6.7, 16.5)	7.1 (6.5, 7.8)	19.2*
No	88.4 (83.5, 93.3)	92.9 (92.2, 93.5)	
Memory problem			
Yes	1.7 (0.0, 3.3)	1.0 (0.7, 1.2)	4.3*
No	98.3 (96.7, 100.0)	99.0 (98.8, 99.3)	

Note. CI = confidence interval; $^{*}p$ <. 01.

testing. SAS version 9.4 was used to conduct the analyses. Respondents' province of residence and whether they were in the Tracking vs Comprehensive Cohorts were also controlled in the multivariate analyses, as recommended by CLSA. We used p <.01 as our significance level for all analyses.

2.5. Ethics

Data access for this study was approved by the CLSA Data Access Committee. The Health Research Ethics Board of the University of Manitoba approved the study protocol.

3. Results

About 51.7% of the Canadian population aged 45–85 years were working at the time of their CLSA *Baseline Questionnaire* with 361 per 10,000 experiencing a work-related injury within the year prior to the survey (see Table 1). Significant age effects for work-related injuries were observed, such that a higher proportion of workers aged 45–54 years reported work-related injuries than those aged 55–85 years (see Table 2). Marital status, education level, annual household income, current working status, work schedule, and several health and health-related conditions and behaviors (i.e., vision, smoking behavior, mood disorder, and memory problem) were also significantly associated with a higher risk of work-related injuries.

Results from bivariate analyses revealed that work-related injuries were significantly associated with increased risk of overnight hospitalization, emergency department visits, visits to dentists, and visits to physiotherapists, or occupational therapists, or chiropractors, but not with the number of visits to family doctors, medical specialists, psychologists, optometrists, social workers, or nursing home/convalescent home use (Table 3).

The average number of healthcare services used within the last 12 months by those who reported work-related injuries (M = 3.54, SE = 0.09) was significantly higher than for those without work-related injuries (M = 3.31, SE = 0.02). However, the observed mean difference was not statistically significant [t(47,726) = -0.06, p = 0.95].

When controlling for the contribution of other factors, work-related injury remained a significant predictor of emergency department visits, and visits to physiotherapists, or occupational therapists, or chiropractors but not overnight hospitalizations, or dentist visits (see Table 4).

4. Discussion

Our results are based on analyses of a large national sample of Canadian males and females, aged 45–85 years. At the time of the CLSA, the incidence of work-related injuries among middle-aged and older workers in Canada was 361 per 10,000 workers, and the rate was higher among 45–54-year-olds than among 55–85-year-olds. A key finding of our study is the significant independent contribution of work-related injuries on subsequent healthcare use of specific type.

Although we found a statistically significant bivariate association between work-related injuries and overnight hospitalizations and visits to dentists, these associations were not significant in multivariate analyses in which we controlled for the contribution of various factors. In contrast, we found that work-related injuries were significantly associated with increased odds of emergency department visits, and visits to physiotherapists, or occupational therapists, or chiropractors even after controlling for the effects of age, sex, annual household income, education, and all the other factors associated with work-related injuries in our study. Although work-related injuries did not predict hospitalizations,

Table 3
Associations between Work-related Injuries and Healthcare Use.

Variable	Work-related injuries		X^2
	Yes% (99% CI)	No% (99% CI)	
Overnight hospitalization			
Yes	10.5 (5.5, 15.4)	5.8 (5.1, 6.4)	10.9*
No	89.5 (84.6, 94.5)	94.2 (93.6, 94.9)	
Emergency department visits			
Yes	29.3 (22.3, 36.3)	18.7 (17.7, 19.8)	24.1*
No	70.7 (63.7, 77.7)	81.3 (80.2, 82.3)	
Family physician visits			
Yes	87.7 (82.5, 92.9)	85.6 (84.6, 86.6)	2.6
No	12.3 (7.1, 17.5)	14.4 (13.4, 15.4)	
Psychologist visits			
Yes	5.3 (2.4, 8.2)	5.4 (4.8, 6.1)	2.1
No	94.7 (91.8, 97.6)	94.6 (93.9, 95.2)	
Medical Specialist visits			
Yes	44.0 (36.6, 51,5)	40.9 (39.5, 42.2)	2.1
No	56.0 (48.5, 63.4)	59.1 (57.8, 60.5)	
Dentist visits			
Yes	78.7 (73.1, 84.2)	84.0 (83.0, 85.0)	19.0*
No	21.3 (15.8, 26.9)	16.0 (15.1, 17.0)	
Optometrist visits			
Yes	52.3 (44.9, 59.7)	53.3 (51.9,54.7)	0.5
No	47.7 (40.3, 55.1)	46.7 (45.3, 48.1)	
Physiotherapist, occupational therapist, or chiropractor visits			
Yes	43.2 (35.8, 50.7)	35.5 (34.2, 36.9)	18.8*
No	56.8 (49.3, 64.2)	64.5 (63.1, 65.8)	
Social Worker visit			
Yes	2.9 (0.3, 5.6)	2.4 (2.0, 2.9)	0.0
No	97.1 (94.4, 99.7)	97.6 (97.1, 98.0)	
Nursing Home resident			
Yes	0.7 (0.0, 2.1)	0.2 (0.1, 0.3)	0.0
No	99.3 (97.9, 100.0)	99.8 (99.7, 99.9)	

Note. CI = confidence interval; *p <.01.

Table 4
Predictors of Healthcare Use in Middle-Aged and Older Canadian Workers aged 45 – 85 years of age.

Variables	Emergency Department Visits	Overnight Hospitalization	Dental Visits	Visits to physiotherapist, occupational therapist, or chiropractor
	OR (99% CI)	OR (99% CI)	OR (99% CI)	OR (99% CI)
Age: 45–54 years	0.943 (0.845, 1.051)	0.597* (0.497, 0.716)	1.080 (0.958, 1.218)	1.140* (1.045, 1.244)
Sex: Female	0.998 (0.890, 1.119)	0.948 (0.783, 1.149)	1.637* (1.441, 1.859)	1.322* (1.207, 1.447)
Marital status: Married/Living with a partner in a common-law relationship	1.024 (0.889, 1.178)	1.086 (0.863, 1.365)	0.980 (0.839, 1.144)	1.025 (0.916, 1.147)
Education: Less than secondary school graduation/Secondary school graduation, no post-secondary	1.168 (0.992, 1.375)	1.365* (1.062, 1.754)	0.586* (0.499, 0.687)	0.845 (0.734, 0.974)
Total household income: $50,000 or more	0.803* (0.680, 0.949)	0.704* (0.539, 0.921)	3.128* (2.660, 3.678)	1.270* (1.103, 1.463)
Work schedule: Daytime	0.872 (0.754, 1.009)	0.770* (0.614, 0.965)	1.286* (1.105, 1.496)	1.020 (0.903, 1.152)
Work status: Full time employee (30+ hours/week)	1.021 (0.880, 1.184)	0.893 (0.708, 1.126)	1.248* (1.067, 1.459)	0.986 (0.877, 1.108)
Work-related injuries	1.490* (1.145, 1.939)	1.409 (0.945, 2.100)	0.748 (0.552, 1.014)	1.471* (1.158, 1.870)
Memory problem	1.483 (0.859, 2.560)	1.684 (0.801, 3.542)	1.090 (0.573, 2.076)	1.278 (0.799, 2.044)
Mood disorder	1.390* (1.202, 1.606)	1.485* (1.176, 1.876)	1.094 (0.919, 1.301)	1.232* (1.092, 1.390)
Vision: Excellent/very good/good	0.949 (0.761, 1.185)	0.751 (0.547, 1.031)	1.376* (1.098, 1.725)	1.204* (1.000,1.448)
Smoking: Daily/Occasional (former daily)	1.320* (1.104, 1.579)	1.026 (0.761, 1.384)	0.506* (0.424, 0.602)	0.683* (0.581, 0.803)

Note. Province of residence and tracking versus comprehensive cohort are also controlled for because of the analytic weights used in the analyses.
CI = confidence interval; *Significant at p <.01.
Reference categories: Participants aged 55–85 years; males; Single, never married or never lived with a partner/Widowed/ Divorced/Separated; Some post-secondary education/Post-secondary degree/ diploma; No memory problem; less than $50,000 total household income; current working schedule Evening/night/rotating shift, seasonal, on call, or casual but called as need arises; Part time employee; No mood disorder; Fair/Poor or non-existent (blind) vision; Occasional smoker/Former daily smoker, but no-smoker now/Never smoked; not injured while working at a job.

lower education level and household income, and non-daytime work schedule significantly increased the odds of having been hospitalized overnight. This suggests that hospitalization was mainly driven by socio-economic factors. Given that socio-economic status is also a predictor of work-related injuries, a mediational analysis might have been informative, as it may have teased out possible indirect effects of work-related injuries and socio-economic status on hospitalization. The differential effect for emergency department visits and hospitalization could also be the result of model power issues as the number of middle-aged and older workers with histories of hospitalization was low (i.e., 10.5%). The lack of significant associations between work-related injuries and visits to family physicians and clinical psychologists may be due to ceiling and floor effects, respectively. Specifically, more than 95% of our study population reported family physician visits but only 5% reported visits to psychologists in the 12 months prior to their follow-up *Maintaining Contact Questionnaire*. It should be noted that there are more barriers to access mental health care than to access family physicians, including long waiting lists, out of pocket costs, and lower availability in rural contexts.

Although work-related injuries were associated with specific types of healthcare use, visits to emergency departments, and visits to physiotherapists, occupational therapists, or chiropractors, combining healthcare utilization into an index of overall healthcare use did not result in statistically significant differences between those with and without work-related injuries. These results suggest that the relationship between work-related injuries and subsequent healthcare use is quite complex.

In addition, the increased use of some of the healthcare services in this study is unlikely to be the direct result of the work injury, given that CLSA participants were asked about injuries in the year prior to the *Baseline Questionnaire*, whereas healthcare use was assessed following the *Baseline Questionnaire*. Therefore, a causal relationship between work-related injuries and subsequent healthcare use cannot be concluded. Nevertheless, the significant association between workplace injuries, emergency department visits, and visits to physiotherapists, occupational therapists, or chiropractors suggests that middle-aged and older workers experienced ongoing health problems that may be exacerbated following their injuries.

A strength of our study is its longitudinal design based on national sample of Canadian workers with all types of occupations are generalizable to the broader population of workers within the age groups that we studied. Our findings also build upon existing research demonstrating that socio-economic status and ongoing health problems are important considerations when examining the complex relationship between workplace injuries and healthcare use in middle-aged and older workers. To the best of our knowledge, prior studies on healthcare utilization outcomes of work-related injuries were mostly cross-sectional and/or focused on a particular occupational category (i.e., farmers) (Nilsson et al., 2010; Pfortmueller et al., 2013; Tonozzi & Layne, 2016; Weigel et al., 2014)with few studies of longitudinal design examined patterns of healthcare utilization following work-related injuries (Brown et al., 2006; López Gómez, Williams, & Boden, 2020; Tonozzi & Layne, 2016). Although studies such as these shed light on the association between work-related injuries and use of health services among injured workers, the models cannot be generalized to the entire working population in that jurisdiction as the existing evidence suggests that the incidence of underreporting is high, with less than half of workers filing work-related injuries claims (Tucker & Keefe, 2019).

Despite its strengths, we acknowledge limitations of our study. First, measures of work-related injuries and healthcare use are based on self-reported CLSA data, and therefore are susceptible to recall bias. Second, we did not examine the independent contribution of work-related injuries on subsequent healthcare utilization rates by age, sex, or type of injury or by occupation type. We looked at type of injuries in our study, but since the majority were classified as sprains or strains, and there was little variability in our study population, we did not include that variable in our analyses when examining healthcare utilization. Examining the contribution of type of injury or socio-economic factors other than the ones included in our analysis (e.g., ethno-cultural background) could be the focus of future studies. Further analysis to examine the link between work-related injuries and subsequent healthcare use by these characteristics is important as there is evidence from prior studies that the contribution might vary by age, sex, and type of injury. For example, younger workers (15–24 years) have been found to report twice the rate of emergency department visits due to knee injuries compared to older workers (55+ years), but older workers had more severe injuries than their younger counterparts (Chen et al., 2013). Other studies have found that older workers (60+ years) reported dramatically fewer but more severe work-related injuries of the lower extremities than younger workers (20–39 years) but here also the older workers reported greater injury severity scores (Frickmann et al., 2012; Konstantinidis et al., 2011). Male workers also have higher rates of emergency department visits for work-related injuries compared to female workers (McCaig et al., 1998). We also cannot rule out the possibility that the work-related injury and health care use was due to a common, underlying health problem. Although we controlled for some health-related variables in the analysis (e.g., mood disorder), future research should include additional potential health issues, such as arthritis which may affect mobility and dexterity. We also recommend conducting mediation analysis in future studies to further investigate the association between work-related injuries, socio-economic characteristics, and targeted health outcomes.

There are several practical implications emanating from these findings. First, organizations and WCBs could ensure that healthcare services used by injured employees are considered priorities for employment insurance coverage, if not already covered. For example, our findings show that injured employees are more likely than non-injured employees to use physiotherapists, occupational therapists, chiropractors, and dentists. This is particularly important because some provincial health care plans offer limited access to these types of services (e.g., a maximum of seven visits to chiropractic care per calendar year in Manitoba), whereas others do not. Although use of physical therapy practitioners was not surprising given the nature of common work injuries, we were surprised by the increased use of dentists. It is unclear whether increased use of dentists among injured employees co-occurs with increased use of other healthcare services; this latter finding deserves more research attention and consideration by organizations supporting injured employees. Second, consideration of these findings for employees once they retire and how their organizations can support injured retirees is worth discussion. For example, do these organizations ensure that these healthcare services are covered in retirees' benefits? Third, the stability of findings between our bivariate and multivariate analyses has implications for employers and provincial-level WCBs. When controlling for residency in Canadian provinces, there are differences among injured and non-injured employees in visits to dentists and physical therapy practitioners. For example, workers from the provinces of Ontario and Nova Scotia had increased odds of visits to dentists when compared to workers from Newfoundland and Labrador. When we examined use of physical therapy practitioners, workers from the provinces of Alberta, Saskatchewan, British Columbia, Manitoba, and Ontario all had increased odds of utilization compared to workers from Newfoundland and Labrador. This might indicate different levels of accessibility for these types of healthcare practitioners by employees in these provinces, and the extent to which workers compensation boards in those provinces provide coverage for healthcare services for injured employees.

5. Conclusions

Our study findings contribute to a greater understanding of the association between workplace injuries and subsequent healthcare use, specifically emergency department visits, and visits to physiotherapists, occupational therapists, and chiropractors, among middle-aged and older workers in Canada. Our findings of increased healthcare use following workplace injuries after controlling for a number of socio-demographic and health-related characteristics suggest a need for the focus on prevention of workplace injuries for all ages including middle-age and older workers. Examination of long-term health and health care utilization of work-related injuries of middle-aged and older workers is recommended.

Data Availability Statement

Data are available from the Canadian Longitudinal Study on Aging (www.clsa-elcv.ca) for researchers who meet the criteria for access to de-identified CLSA data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was supported by WorkplaceNL and the Canadian Centre for Advanced Leadership in Business. This research was made possible using the data/biospecimens collected by the Canadian Longitudinal Study on Aging (CLSA). Funding for the Canadian Longitudinal Study on Aging (CLSA) is provided by the Government of Canada through the Canadian Institutes of Health Research (CIHR) under grant reference: LSA 94473 and the Canada Foundation for Innovation, as well as the following provinces, Newfoundland, Nova Scotia, Quebec, Ontario, Manitoba, Alberta, and British Columbia. This research has been conducted using the CLSA Baseline Tracking Dataset Version 3.4; Comprehensive Dataset Version

4.0; Follow-up 1 Tracking Dataset Version 1.0 and Follow-up 1 Comprehensive Dataset Version 1.0, dataset under Application 190204. The CLSA is led by Drs. Parminder Raina, Christina Wolfson and Susan Kirkland. We thank Scott Nowicki for his assistance in data analysis.

Disclaimer

The opinions expressed in this manuscript are the authors' own and do not reflect the views of the Canadian Longitudinal Study on Aging (CLSA).

References

Algarni, F. S., Gross, D. P., Senthilselvan, A., et al. (2015). Ageing workers with work-related musculoskeletal injuries. *Occupational Medicine (Chic Ill), 65*, 229–237. https://doi.org/10.1093/occmed/kqu213.

Berdahl, T. A., & Zodet, M. (2020). Medical care utilization for work-related injuries in the United States 2002–2006.

Brown, J., McDonough, P., Mustard, C., et al. (2006). Healthcare use before and after a workplace injury in British Columbia, Canada. *Occupational and Environmental Medicine, 63*, 396–403. https://doi.org/10.1136/oem.2005.022707.

Canadian Longitudinal Survey on Aging (CLSA). (2020) Can. Longitud. Study Aging/Etude Longitud. Can. sur le Vieil. (CLSA elcv). https://www.clsa-elcv.ca/ (accessed 29 Mar 2020).

Chen, Z., Chakrabarty, S., Levine, R. S., et al. (2013). Work-related knee injuries treated in emergency departments in the United States. *Journal of Occupational and Environmental Medicine, 55*, 1091–1099. https://doi.org/10.1097/JOM.0b013e31829b27bf.

Frickmann, F., Wurm, B., Jeger, V., et al. (2012). 782 Consecutive construction work accidents: Who is at risk? A 10-year analysis from a Swiss university hospital trauma unit. *Swiss Medical Weekly, 142*. https://doi.org/10.4414/smw.2012.13674.

International Labour Organization (ILO). (2005). *Introductory report: Decent work – safe work*. https://www.ilo.org/wcmsp5/groups/public/—europe/—ro-geneva/—sro-moscow/documents/genericdocument/wcms_312093.pdf.

Konstantinidis, A., Talving, P., Kobayashi, L., et al. (2011). Work-related injuries: Injury characteristics, survival, and age effect. *American Surgeon, 77*, 702–707.

López Gómez, M. A., Williams, J. A. R., Boden, L., et al. (2020). The relationship of occupational injury and use of mental health care. *Journal of Safety Research, 74*, 227–232. https://doi.org/10.1016/j.jsr.2020.06.004.

McCaig, L. F., Burt, C. W., & Stussman, B. J. (1998). A comparison of work-related injury visits and other injury visits to emergency departments in the United States, 1995–1996. *Journal of Occupational and Environmental Medicine, 40*, 870–875.

Morassaei, S., Breslin, F. C., Ibrahim, S. A., et al. (2013). Geographic variation in work injuries: A multilevel analysis of individual-level data and area-level factors within Canada. *Annals of Epidemiology, 23*, 260–266. https://doi.org/10.1016/j.annepidem.2013.03.008.

Nilsson, K., Pinzke, S., & Lundqvist, P. (2010). Occupational injuries to senior farmers in Sweden. *J Journal of Agricultural Safety and Health, 16*, 19–29. https://doi.org/10.13031/2013.29246.

Parachute (2015). *The cost of injury in Canada*. Toronto, ON: Parachute.

Pfortmueller, C. A., Kradolfer, D., Kunz, M., et al. (2013). Injuries in agriculture - Injury severity and mortality. *Swiss Medical Weekly, 143*. https://doi.org/10.4414/smw.2013.13846.

Raina, P. S., Wolfson, C., Kirkland, S. A., et al. (2018). The Canadian Longitudinal Study on Aging (CLSA) report on healthy aging in Canada. Published Online First: https://www.clsa-elcv.ca/doc/2639.

Raina, P., Wolfson, C., Kirkland, S., et al. (2019). Cohort profile: The Canadian Longitudinal Study on Aging (CLSA). *International Journal of Epidemiology, 48*, 1752–1753j. https://doi.org/10.1093/ije/dyz173.

Ruseckaite, R., Collie, A., Prang, K. H., et al. (2016). Does medical certification of workers with injuries influence patterns of health service use? *Work, 54*, 669–678. https://doi.org/10.3233/WOR-162312.

Sato, K., Yamazaki, S., Hayashino, Y., et al. (2011). Associations between hours worked, symptoms and health resource utilization among full-time male Japanese workers. *Journal of Occupational Health, 53*, 197–204. https://doi.org/10.1539/joh.10-0039-OA.

Scott, K. A., Liao, Q., Fisher, G. G., et al. (2018). Early labor force exit subsequent to permanently impairing occupational injury or illness among workers 50–64 years of age. *American Journal of Industrial Medicine, 61*, 317–325. https://doi.org/10.1002/ajim.22817.

Seo, J. Y., Chao, Y. Y., Yeung, K. M., et al. (2019). Factors Influencing Health Service Utilization Among Asian Immigrant Nail Salon Workers in the Greater New York City Area. *Journal of Community Health, 44*. https://doi.org/10.1007/s10900-018-0544-7.

Statistics Canada. (2017a). Focus on Geography Series, 2016 Census. Statistics Canada Catalogue no. 98-404-X2016001. Ottawa, Ontario. Data products, 2016 Census. https://www12.statcan.gc.ca/census-recensement/2016/as-sa/fogs-spg/Facts-pr-eng.cfm?Lang=Eng&GK=PR&GC=10&TOPIC=7 (accessed 10 May 2020).

Statistics Canada. (2017b). Census in brief: Working seniors in Canada. http://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016027/98-200-x2016027-eng.pdf.

Stoesz, B., Chimney, K., Deng, C., et al. (2020). Incidence, risk factors, and outcomes of non-fatal work-related injuries among older workers: A review of research from 2010 to 2019. *Safety Science, 126*, 104668. https://doi.org/10.1016/j.ssci.2020.104668.

Tadros, A., Sharon, M., Chill, N., et al. (2018). Emergency department visits for work-related injuries. *American Journal of Emergency Medicine, 36*, 1455–1458. https://doi.org/10.1016/j.ajem.2018.04.058.

Tonozzi, T. R., & Layne, L. A. (2016). Hired crop worker injuries on farms in the United States: A comparison of two survey periods from the National Agricultural Workers Survey. *American Journal of Industrial Medicine, 59*, 408–423. https://doi.org/10.1002/ajim.22578.

Tucker, S., & Keefe, A. (2019) Report on work fatality and injury rates in Canada. Regina, Saskatchewan, Canada: https://www.uregina.ca/business/faculty-staff/faculty/file_download/2018-Report-on-Workplace-Fatalities-and-Injuries.pdf.

Weigel, M. M., Armijos, R. X., & Beltran, O. (2014). Musculoskeletal injury, functional disability, and health-related quality of life in aging Mexican immigrant farmworkers. *Journal of Immigrant and Minority Health, 16*, 904–913. https://doi.org/10.1007/s10903-013-9788-6.

White, M. S., Burns, C., & Conlon, H. A. (2018). The impact of an aging population in the workplace. *Workplace Health & Safety, 66*, 493–498. https://doi.org/10.1177/2165079917752191.

Young, A. E., Cifuentes, M., Wasiak, R., et al. (2009). Urban-rural differences in work disability following occupational injury: Are they related to differences in healthcare utilization? *Journal of Occupational and Environmental Medicine, 51*, 204–212. https://doi.org/10.1097/JOM.0b013e3181917d8e.

Dr. Shahin Shooshtari is a Professor in the Department of Community Health Sciences in the Max Rady College of Medicine at the University of Manitoba in Canada. Her program of research is focused on aging with disability.

Dr. Verena Menec is a Professor in the Department of Community Health Sciences in the Max Rady College of Medicine at the University of Manitoba in Canada. Her program of research is focused on older adults. She is the Manitoba lead researcher for the Canadian Longitudinal Study on Aging (CLSA).

Dr. Brenda M. Stoesz is an Adjunct Professor in the Department of Psychology at the University of Manitoba in Canada. She is also a Senior Faculty Specialist with the University of Manitoba Centre for the Advancement of Teaching and Learning.

Ms. Dimple Bhajwani is a graduate student in Medical Rehabilitation at the University of Manitoba in Canada.

Dr. Nick Turner is a Professor at Haskayne School of Business at the University of Calgary in Canada. His program of research is focused on occupational health and safety, leadership, and work design.

Dr. Caroline Piotrowski is an Associate Professor in the Department of Community Health Sciences in the Max Rady College of Medicine at the University of Manitoba in Canada. Her program of research is focused on violence and injury prevention.

Journal of Safety Research 83 (2022) 302–309

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Existence of the safety-in-numbers effect in the aspect of injury severity: A macroscopic analysis for bicyclists and pedestrians

Yanqi Lian [a], Enru Zhou [a], Jaeyoung Lee [a,b,*], Mohamed Abdel-Aty [b]

[a] School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan 410075, China
[b] Department of Civil, Environmental & Construction Engineering, University of Central Florida, United States

ARTICLE INFO

Article history:
Received 6 September 2021
Received in revised form 19 May 2022
Accepted 7 September 2022
Available online 16 September 2022

Keywords:
Safety-in-numbers
Injury severity
Bicyclists
Pedestrians
Fractional split multinomial logit model

ABSTRACT

Objective: Several studies have confirmed the existence of a safety-in-numbers effect in relation to vulnerable road users. The safety-in-numbers effect refers to a phenomenon wherein the number of bicyclists/pedestrians on a road is higher, and consequently, the risk of each bicyclist/pedestrian being involved in a crash is lower. Nevertheless, the existence of the safety-in-numbers effect in the aspect of injury severity in traffic crashes has not yet been investigated. Thus, this study aimed to explore whether traffic injuries are more (less) severe with fewer (more) pedestrians/bicyclists at the county level. *Method:* Using two fractional split multinomial logit models, the relationships between the number of bicyclists/pedestrians and the proportion of crashes involving bicyclists/pedestrians based on crash severity were investigated at the county level using crash data from Florida. In other words, we explored whether differing number of bicyclists/pedestrians could change the distribution of traffic injury severity levels. *Results:* The modeling results clearly revealed a lower proportion of severe injuries caused to bicyclists/pedestrians at a higher level of daily bicycle/pedestrian flows, indicating existence of the safety-in-numbers effect. Several variables (e.g., the percentage of people aged 65 years and older, the percentage of commuters using public transportation, and the proportion of recreational land use) were found to have a significant effect on the distribution of traffic injury severity among bicyclists/pedestrians. *Conclusion:* This study proves that a safety-in-numbers effect exists in the aspect of injury severity among bicyclists and pedestrians. *Practical applications:* These findings are expected to provide recommendations for promoting the use of active transportation, which will improve the safety of vulnerable road users in future.

1. Introduction

In recent decades, a number of transportation and public health researchers have agreed that walking and cycling should be promoted owing to their advantages of economic and environmental sustainability and benefits to public health (Mueller et al., 2018). However, more than half of all the road traffic deaths reported involve vulnerable road users, such as pedestrians, cyclists, and motorcyclists (World Health Organization, 2021), which presents a barrier to the shift from motorized transportation to sustainable modes of transportation. To encourage car users to choose walking or cycling as their primary mode of transportation, several efforts have been dedicated to explore and improve the safety of vulnerable road users. A phenomenon called safety-in-numbers was discovered while studying the safety of vulnerable road users (Brüde & Larsson, 1993).

The safety-in-numbers effect refers to the hypothesis that an individual in a larger group is more likely to be protected from mishaps or accidents. In other words, when the number of bicyclists/-pedestrians increases, the risk of crash involvement for each bicyclist/pedestrian decreases. Over the past two decades, the safety-in-numbers concept has attracted significant research interest. Some researchers have investigated the relationship between crashes involving bicyclists/pedestrians and the number of bicyclists/pedestrians, and the results have indicated that the safety-in-numbers effect exists on a microscopic scale, including intersections (Kröyer, 2016; Miranda-Moreno, Morency, & El-Geneidy, 2011; Murphy, Levinson, & Owen, 2017; Schneider et al., 2010; Xie, Dong, Wong, Huang, & Xu, 2018; Xu, Xie, Dong, Wong, & Huang, 2019), crosswalks (Elvik, 2016), and roundabouts (Daniels, Brijs, Nuyts, & Wets, 2010). In contrast, other studies have

* Corresponding author.
E-mail addresses: lyanqi@csu.edu.cn (Y. Lian), enru_zhou@csu.edu.cn (E. Zhou), lizaining@csu.edu.cn, jaeyoung@knights.ucf.edu (J. Lee), m.aty@ucf.edu (M. Abdel-Aty).

https://doi.org/10.1016/j.jsr.2022.09.004
0022-4375/

investigated the safety-in-numbers effect on a macroscopic scale, including states (Robinson, 2005), census tracts (Tasic, Elvik, & Brewer, 2017), and metropolitan statistical areas (Lee, Abdel-Aty, & Cai, 2020). Jacobsen (2015) examined the safety-in-numbers effect and proved its existence across different macro levels ranging from cities to countries.

In addition to exploring the existence of the safety-in-numbers effect on a microscopic or macroscopic scale, a few studies have analyzed more specific circumstances. Agent-based modeling was applied to replicate the safety-in-numbers effect in a simulated environment, and the results implied that the safety-in-numbers effect exists only when bicycle density increases over time (Thompson, Savino, & Stevenson, 2015). Lee, Abdel-Aty, Xu, and Gong (2019) explored the safety-in-numbers effect in areas with low pedestrian activities at intersections. The results indicate that the safety-in-numbers effect exists at intersections with larger pedestrian activities, whereas it does not exist at intersections with minimal pedestrian activities.

Although the above-mentioned studies have identified the existence of the safety-in-numbers effect, the reported strength of the effect varies significantly between studies. Elvik (2017) reviewed existing studies analyzing the safety-in-numbers effect using count regression models and revealed that the strength of the safety-in-numbers effect is inversely related to the number of cyclists and pedestrians. Another meta-analysis conducted by Elvik and Bjørnskau (2019) suggested that the strength of the safety-in-numbers effect is more likely to be larger for pedestrians than for cyclists and higher at the macro level than at the micro level.

Mechanisms underlying the safety-in-numbers effect have also been investigated in recent years. Some possible mechanisms that have been proposed include the behavioral adaptation of drivers (Jacobsen, 2015), improved interaction between road user groups (Phillips, Bjørnskau, Hagman, & Sagberg, 2011), and safer street regulations, design, and operation (Bhatia & Wier, 2011). These mechanisms have been tested in several studies. The improved interaction between road user groups, which refers to other road users creating more correct expectations when they become accustomed to the presence of bicyclists, was verified by analyzing data from Oslo, Norway, where bicycle use exhibits a substantial seasonal variation (Fyhri, Sundfør, Bjørnskau, & Laureshyn, 2017). Two studies that used simulated environments demonstrated that behavioral adaptation may be a sufficient but unnecessary input and that the safety-in-numbers effect still exists even in the absence of behavioral adaptation (Thompson et al., 2015, 2016). One mechanism involves the following hypothesis: more attention will be paid to cyclists if more drivers are also cyclists (cyclist drivers), which occurs when the number of cyclists increases. To test this mechanism, Johnson, Oxley, Newstead, and Charlton (2014) conducted an online survey among Australian drivers who did not cycle and cyclist drivers and discovered that cyclist drivers tended to report positive attitudes towards cyclists.

Based on the previous discussions, it can be established that considerable effort has been devoted to explore the safety-in-numbers effect. However, previous studies have analyzed the safety-in-numbers effect with regard to the crash frequency or crash rate. An analysis of the effect with regard to injury severity among bicyclists and pedestrians has not yet been investigated. The existence of this effect would indicate that with an increasing proportion of bicyclists/pedestrians, the proportion of more severe injuries would decrease, whereas that of less severe injuries would increase. The primary objective of the current study is, therefore, to evaluate the existence of the safety-in-numbers effect from the perspective of traffic injury severity among vulnerable road users. To accomplish this, the current study adopts two fractional split multinomial logit models to analyze the proportion of crashes based on their severity with data collected from Florida aggregated at the county level.

2. Data

Descriptive statistics of the data for bicyclists and pedestrians are summarized in Tables 1 and 2, respectively. These data were collected from 67 counties in Florida, USA. Four main datasets were integrated and used in the current study: bicycle and pedestrian crash data, Strava data, Florida socioeconomic and demographic data, and land-use data. The data used in this study are described in detail in this section.

2.1. Bicycle and pedestrian crash data

Bicycle and pedestrian crash data at different severity levels were obtained from the Florida Department of Transportation (FDOT) Crash Analysis Reporting System. The crash data contained five injury severity levels: fatal injury (K), incapacitating injury (A), non-incapacitating evident injury (B), possible injury (C), and property damage only (O). To obtain sufficient observations at each injury severity level, the crash data were divided into three levels: severe injury (KA), moderate injury (B), and minor/no injury (CO). Bicycle crash data from 2012–2016 and pedestrian crash data from 2012–2015 were used in this study. The crash data based on severity type were further aggregated according to the county, and the corresponding crash proportions were calculated. No bicycle crashes were reported for three counties during the study period. Therefore, these three counties were excluded from this study.

2.2. Exposure data

Strava tracks the activities of cyclists and pedestrians via smartphone applications based on the global positioning system. Thus, Strava data were obtained from the FDOT Unified Basemap Repository. Similar to the crash data, Strava data for bicyclists from 2012–2016 and for pedestrians from 2012–2015 were used in this study. The daily bicycle miles traveled (DBMT) and daily pedestrian miles traveled (DPMT) were calculated by multiplying the travel distance by the trips of bicyclists and pedestrians, respectively. The logarithms of the DBMT and DPMT were used as the traffic exposure variables in this study.

2.3. Socioeconomic and demographic data

Socioeconomic data such as the percentage of occupations in the primary industry sector (raw materials), secondary industry sector (manufacturing), and tertiary industry sector (services), percentage of families with no vehicles, unemployment rate, and median household income were obtained from the U.S. Census Bureau. Demographic data including the percentages of the population based on age group and the percentages of specific races (i.e., Asian, Black, and Hispanic) were also collected from the U.S. Census Bureau.

2.4. Commuting travel data and land-use data

Commuting travel data were acquired from the U.S. Census Bureau, and the percentages of commuters using specific travel modes (e.g., car, public transportation, and walking) were computed for each county. Land-use data containing details of land-use classification were acquired from the FDOT. This study includes the proportions of several major land-use types (i.e., agricultural, public, recreational, residential, and commercial land use).

Table 1
Descriptive statistics of data for bicycle analysis (N = 64).

Category	Variable	Mean	Stdev	Min	Max
Injury severity related variables	Proportion of severe injuries (KA)	0.212	0.148	0.000	1.000
	Proportion of moderate injuries (B)	0.357	0.162	0.000	0.857
	Proportion of minor/no injuries (CO)	0.643	0.162	0.143	1.000
Exposure variables	DBMT (daily bicycle miles traveled)	163,789	268,459	1326.1	1,382,029
Socioeconomic characteristics	Percentage of occupations in the primary industry sector	3.172	4.146	0.300	19.600
	Percentage of occupations in the secondary industry sector	14.200	3.084	5.800	23.400
	Percentage of occupations in the tertiary industry sector	82.628	6.107	61.800	93.800
	Percentage of families with no vehicles	5.666	1.872	1.900	10.300
	Percentage of unemployment	5.961	1.547	2.900	12.000
	Percentage of education level: college and higher	23.013	9.904	8.300	46.200
	Percentage of individuals without health insurance	12.811	3.590	5.600	25.600
	Median household income (in USD)	51,616	10,376	35,438	82,252
Demographic characteristics	Percentage of age group: 5–14 years	10.953	1.810	4.200	15.100
	Percentage of age group: 15–24 years	11.478	3.066	4.300	25.100
	Percentage of age group: 25–64 years	50.584	4.405	33.100	57.700
	Percentage of age group: 65 years and older	21.939	7.781	11.600	56.800
	Percentage of males	50.883	3.086	47.400	59.600
	Percentage of African Americans	14.091	9.295	3.000	56.100
	Percentage of Hispanics	14.572	13.327	2.600	68.500
	Percentage of Asian Americans	1.678	1.336	0.100	6.100
Commuting characteristics	Percentage of commuters using cars	89.795	3.764	76.000	96.900
	Percentage of commuters using public transportation	0.881	1.247	0.000	6.600
	Percentage of commuters who walk	1.427	0.873	0.200	4.700
	Percentage of workers at home	5.561	2.083	0.900	11.000
Land-use attributes	Proportion of agricultural land-use	0.486	0.251	0.000	0.938
	Proportion of public land-use	0.112	0.136	0.004	0.762
	Proportion of recreational land-use	0.080	0.141	0.000	0.700
	Proportion of residential land-use	0.094	0.076	0.006	0.421
	Proportion of commercial land-use	0.011	0.013	0.000	0.080

Table 2
Descriptive statistics of data for pedestrian analysis (N = 67).

Category	Variable	Mean	Stdev	Min	Max
Injury severity related variables	Proportion of severe injuries (KA)	0.341	0.154	0.000	1.000
	Proportion of moderate injuries (B)	0.347	0.133	0.000	1.000
	Proportion of minor/no injuries (CO)	0.311	0.122	0.000	0.600
Exposure variables	DPMT (daily pedestrian miles traveled)	10,261	18,368	42.629	99,147
Socioeconomic characteristics	Percentage of occupations in the primary industry sector	3.503	4.462	0.300	19.600
	Percentage of occupations in the secondary industry sector	14.181	3.041	5.800	23.400
	Percentage of occupations in the tertiary industry sector	82.316	6.297	61.800	93.800
	Percentage of families with no vehicles	5.724	1.909	1.900	10.300
	Percentage of unemployment	6.118	1.731	2.900	12.000
	Percentage of education level: college and higher	22.413	10.076	7.900	46.200
	Percentage of individuals without health insurance	12.851	3.548	5.600	25.600
	Median household income (in USD)	51,290	10,300	35,438	82,252
Demographic characteristics	Percentage of age group: 5–14 years	10.955	1.774	4.200	15.100
	Percentage of age group: 15–24 years	11.513	3.018	4.300	25.100
	Percentage of age group: 25–64 years	50.840	4.477	33.100	58.400
	Percentage of age group: 65 years and older	21.639	7.735	11.600	56.800
	Percentage of males	51.339	3.740	47.400	65.200
	Percentage of African Americans	14.542	9.464	3.000	56.100
	Percentage of Hispanics	14.378	13.082	2.600	68.500
	Percentage of Asian Americans	1.616	1.337	0.000	6.100
Commuting characteristics	Percentage of commuters using cars	89.954	3.823	76.000	98.000
	Percentage of commuters using public transportation	0.863	1.225	0.000	6.600
	Percentage of commuters who walk	1.434	0.873	0.200	4.700
	Percentage of workers at home	5.427	2.157	0.600	11.000
Land-use attributes	Proportion of agricultural land-use	0.500	0.254	0.000	0.938
	Proportion of public land-use	0.109	0.134	0.003	0.762
	Proportion of recreational land-use	0.079	0.139	0.000	0.700
	Proportion of residential land-use	0.091	0.076	0.006	0.421
	Proportion of commercial land-use	0.010	0.013	0.000	0.080

3. Preliminary analysis

Preliminary analyses were conducted to investigate the existence of the safety-in-numbers effect in the aspect of crash frequency. The relationship between bicycle/pedestrian volume and bicycle/pedestrian crash frequency was analyzed using the following general negative binomial regression model:

$$\text{Number of bicycle (pedestrian) crashes} = exp^{\beta_0} Veh^{\beta_1} Bic(Ped)^{\beta_2} \tag{1}$$

Table 3
Negative binomial model for bicycle and pedestrian crashes.

Variables	Bicycle crash frequency	Pedestrian crash frequency
Intercept	– 13.822*** (SE = 1.449)	– 11.666*** (SE = 1.093)
Daily pedestrian miles traveled (DPMT)		0.289*** (SE = 0.066)
Daily bicycle miles traveled (DBMT)	0.436*** (SE = 0.113)	
Daily vehicle miles traveled (DVMT)	0.930*** (SE = 0.163)	0.951*** (SE = 0.103)
Overdispersion (α)	0.422 (SE = 0.077)	0.164 (SE = 0.033)
Log-likelihood ratio (d.f. = 2)	146.460*** (d.f. = 2, p < 0.0001)	206.780*** (d.f. = 2, p < 0.0001)
Pseudo R^2	0.170	0.229

*** Significant at 99% confidence level.

Table 4
Parameter estimates for bicycle crash proportion based on severity.

Variables	Proportion of Severe Injuries	Proportion of Moderate Injuries	Proportion of Minor/No Injuries
Natural logarithm of DBMT	−0.207*** (0.074)	−0.065 (0.062)	Reference
Proportion of age group: 65 years and older	0.007 (0.007)	−0.012** (0.006)	
Proportion of commuters using public transportation	−0.104 (0.067)	−0.191*** (0.073)	
Intercept	1.611* (0.946)	1.330* (0.807)	
Pseudo log-likelihood at intercept	−67.810657		
Pseudo log-likelihood at convergence	−66.757660		

Numbers in parentheses represent standard errors of the estimated coefficients.
*** Significant at 99% confidence level.
** Significant at 95% confidence level.
* Significant at 90% confidence level.

Here, Veh, Bic, and Ped represent the annual average volumes of the motor vehicles, bicycles, and pedestrians, respectively. β_0 denotes the intercept, β_1 and β_2 represent the corresponding exponential coefficients of the vehicle volume and bicycle (pedestrian) volume, respectively.

As presented in Table 3, the exponential coefficients of vehicle volume and bicycle (pedestrian) volume are greater than zero and less than one, which indicates that the crash rate decreases as the vehicle volume and bicycle (pedestrian) volume increases; this confirms the existence of the safety-in-numbers effect in the aspect of crash frequency (Elvik & Bjørnskau, 2017; Elvik & Goel, 2019; Lee et al., 2019, 2020).

4. Method

The fractional split model was first proposed by Papke (1996) to analyze fractional bivariate dependent variables that range between zero and one. The multinomial version, the fractional multinomial split model, was further developed and has been applied in the field of transportation, including the evaluation of crash proportion based on severity levels (Yasmin, Eluru, Lee, & Abdel-Aty, 2016), crash proportion based on vehicle type (Lee, Yasmin, Eluru, Abdel-Aty, & Cai, 2018), and the proportion of speed limit violations across highway segments (Afghari, Haque, & Washington, 2018). The formulation of the fractional split multinomial logit model to analyze the proportion of crashes based on severity level is presented in this section.

In this study, the dependent variables are the proportions of crashes based on severity level, which consider continuous values between zero and one and add up to one for each county.

$$0 \leq y_{sc} \leq 1 \sum_{s=1}^{S} y_{sc} = 1 \tag{2}$$

The fraction of crashes based on injury severity s ($s = 1, \ldots, S$; $S = 3$) in county c ($c = 1,2, \ldots, C$), y_{sc}, is a function of a vector consisting of the relevant explanatory variables related to the characteristics of that county.

$$\mathrm{E}(y_s|x) = G_s(x, \beta) = \frac{\exp(x\beta_s)}{\sum_{s=1}^{S}\exp(x\beta_s)}, \quad s = 1, \ldots, S \tag{3}$$

The predetermined link function $G_s(\cdot)$ is assumed to follow a logistic distribution that responds to the fractional split multinomial logit model.

The fractional split multinomial logit model cannot be estimated using a conventional maximum likelihood function. Thus, the quasi-likelihood function $L_q(\beta)$ is employed in this study:

$$L_q(\beta) = \prod_{s=1}^{S} G_s(x_c, \beta)^{y_{sc}} \tag{4}$$

The quasi log-likelihood function, $\mathcal{L}(\beta)$, is defined as.

$$\mathcal{L}(\beta) = \sum_{c=1}^{C} \ln(L_q(\beta)) \tag{5}$$

5. Results

Tables 4 and 5 present the parameter estimation results of the fractional split multinomial logit models for bicyclists and pedestrians, respectively. In the fractional split multinomial logit model, each alternative has a propensity equation, and the model estimation requires one of the alternatives as a reference. In the current study, the model was estimated using the proportion of minor/no injuries as the reference. Thus, no coefficients are specific to the proportion of minor/no injuries in Tables 4 and 5.

Table 5
Parameter estimates for pedestrian crash proportion based on severity.

Variables	Proportion of Severe Injuries	Proportion of Moderate Injuries	Proportion of Minor/No Injuries
Natural logarithm of DPMT	−0.202*** (0.044)	−0.074* (0.040)	Reference
Proportion of recreational land-use	0.769* (0.444)	0.751** (0.364)	
Intercept	1.569*** (0.381)	0.634 (0.393)	
Pseudo log-likelihood at intercept	−73.530311		
Pseudo log-likelihood at convergence	−72.496468		

Numbers in parentheses represent standard errors of the estimated coefficients.
*** Significant at 99% confidence level.
** Significant at 95% confidence level.
* Significant at 90% confidence level.

Table 6
Marginal effects of explanatory variables for bicycle crash proportion based on severity.

Variables	Proportion of Severe Injuries	Proportion of Moderate Injuries	Proportion of Minor/No Injuries
Natural logarithm of DBMT	−0.028** (0.012)	0.003 (0.014)	0.025** (0.012)
Percentage of age group: 65 years and older	0.002* (0.001)	−0.004** (0.002)	0.001 (0.001)
Proportion of commuters using public transportation	0.0004 (0.008)	−0.037** (0.015)	0.036** (0.014)

Numbers in parentheses represent standard errors of the estimated coefficients.
***Significant at 99% confidence level.
** Significant at 95% confidence level.
* Significant at 90% confidence level.

Table 7
Marginal effects of explanatory variables for pedestrian crash proportion based on severity.

Variables	Proportion of Severe Injuries	Proportion of Moderate Injuries	Proportion of Minor/No Injuries
Natural logarithm of DPMT	−0.036*** (0.009)	−0.007 (0.009)	0.029*** (0.006)
Proportion of recreational land-use	0.080 (0.090)	0.080 (0.080)	−0.160** (0.071)

Numbers in parentheses represent standard errors of the estimated coefficients.
*** Significant at 99% confidence level.
** Significant at 95% confidence level; *Significant at 90% confidence level.

In the fractional split multinomial logit model, the positive (negative) coefficient of the alternative indicates an increased (decreased) proportion of this alternative compared with the proportion of minor/no injury. An insignificant coefficient suggests that the model does not find a significant difference between the effects of the variable on the alternative and proportion of minor/no injuries based on the data. It should be noted that the coefficients of severe (KA) and moderate (B) injuries in the models were obtained by comparison with the reference: minor/no injuries (CO).

5.1. Bicycle crash model

As presented in Table 4, positive intercepts indicate that the proportions of severe and moderate injuries are higher than the proportions of minor/no injuries in the absence of any other factors. An increase in the natural logarithm of DBMT is related to a decrease in the proportion of severe injuries. A higher proportion of people aged 65 years and older or a higher proportion of commuters using public transportation is inclined to decrease the proportion of moderate injuries. Meanwhile, it is possible that the proportion of commuters using public transportation demonstrates certain effects on the proportion of severe injuries, as the p-value is slightly greater than 0.1 ($p = 0.121$); however, this has not been proven with the data used in this study.

5.2. Pedestrian crash model

The positive intercept in Table 5 demonstrates that the proportion of severe injuries is higher than the proportion of minor/no injuries in the absence of any other factors. An increase in the natural logarithm of DPMT is associated with a decrease in the proportions of severe and moderate injuries, with the magnitude for severe injuries being higher. The proportion of recreational land use exerts a positive influence on the proportion of severe and moderate injuries.

Tables 6 and 7 present the marginal effects of the explanatory variables in the fractional split multinomial logit models for bicyclists and pedestrians, respectively. The marginal effect is usually used to measure the exact magnitude of the effect of the explanatory variables on dependent variables. In this study, marginal effects are defined as changes in the crash proportion of each injury severity category in response to a one-unit increase in the value of an explanatory variable while all other variables are maintained constant.

6. Discussion

It should be noted that Strava data cannot capture all bicyclist/pedestrian trips because not all bicyclists/pedestrians use the Strava application. A previous study discovered that Strava data

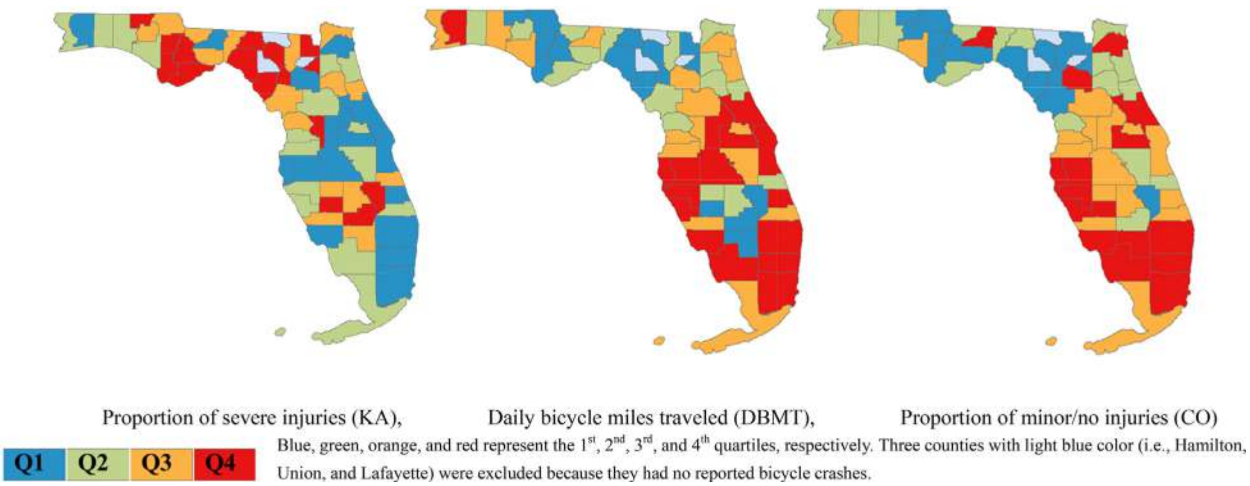

Fig. 1. Proportions of bicycle crashes based on injury severity and daily bicycle miles traveled.

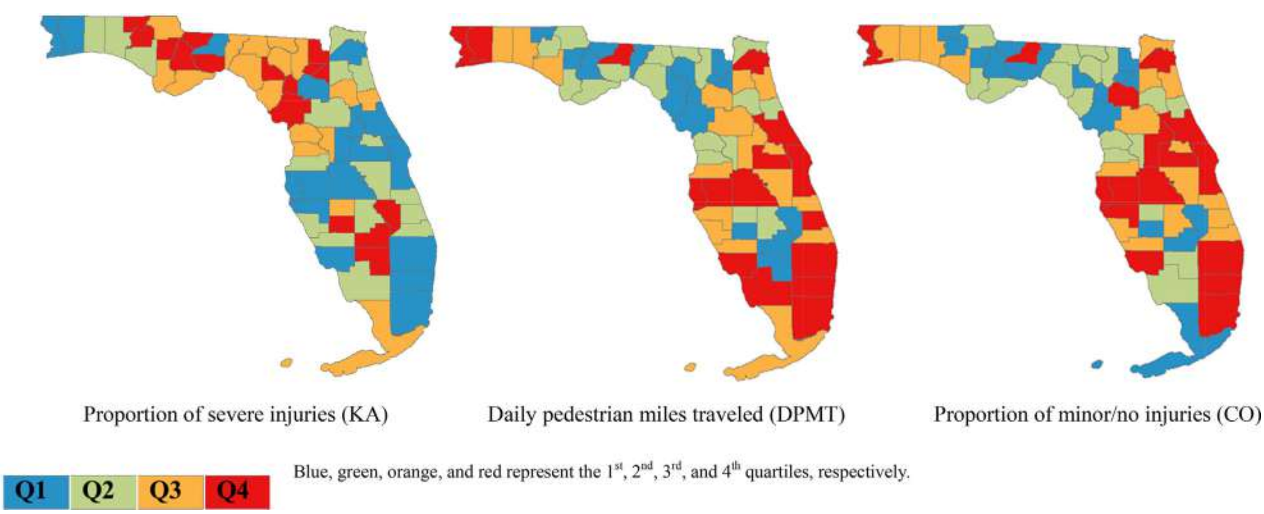

Fig. 2. Proportions of pedestrian crashes based on injury severity and daily pedestrian miles traveled.

are skewed towards young male cyclists (Hochmair, Bardin, & Ahmouda, 2019). Therefore, several other explanatory variables were introduced to complement the incomplete Strava bicyclist/-pedestrian trip data.

At first glance, some readers may assume that the results of the current study are counterintuitive or contradictory to those of previous studies. Nevertheless, it should be noted that the current study focused on the proportions of crashes based on severity level, whereas previous studies have focused on crash frequency; thus, the results of the two topics are, of course, different. It is worth reiterating that there is an intrinsic difference between the decrease in an alternative relative to a reference (i.e., minor/no injuries) and the decrease in the alternative itself. The decrease in an alternative relative to the reference does not imply that the alternative itself has decreased. Therefore, both estimated coefficients in the fractional split multinomial logit model and their marginal effect should be simultaneously analyzed to understand the impact of explanatory variables on the proportion of crashes based on injury severity level.

6.1. Bicycle crash model

According to the marginal effects of the explanatory variables, as DBMT increases, the proportion of severe injuries decreases, while the proportion of minor/no injuries increases. This result suggests that the safety-in-numbers effect still exists from the perspective of injury severity of bicyclists. In other words, bicyclists in counties with a larger proportion of bicyclists have a propensity to be less seriously injured when they are involved in a crash. A consistent conclusion can be drawn from Fig. 1. Counties with a larger DBMT are likely to have a smaller proportion of severe injuries and a higher proportion of minor/no injuries. This phenomenon can be explained based on two perspectives. First, drivers are accustomed to the presence of bicyclists and develop more accurate perceptions of bicyclists in counties with a larger population of bicyclists (Fyhri et al., 2017). Second, more drivers will also be bicyclists as the number of bicyclists increases. These drivers tend to adopt more positive attitudes towards bicyclists and safer driving behaviors (Fyhri et al., 2017). Drivers with greater caution towards bicyclists are more likely to react quickly when conflicting with bicyclists, which will lead to a reduction in injury severity. The importance of attitudes and safety behaviors in reducing fatal crashes is highlighted by comparing traffic safety performance in different countries (Page, 2001). Driving attitudes and behaviors are considered to be driver-related factors that affect injury severity (De Oña, De Oña, Eboli, Forciniti, & Mazzulla, 2014).

The results indicate that counties with a high proportion of people aged 65 and older are likely to have a higher proportion of severe injuries but a lower proportion of moderate injuries. This is because older people are less likely to ride a bicycle than younger people (Census Bureau, 2021; Survey, 2019). This results in a relatively small number of bicyclists; thus, such counties tend to provide relatively unsafe environments for bicyclists. Counties with a high proportion of elderly people should be targeted to improve the safety of bicyclists. A higher proportion of commuters using public transportation tends to decrease the proportion of moderate injuries but increase the proportion of minor/no injuries. This result suggests that a higher proportion of commuters using public transportation is related to lower injury severity among bicyclists. One possible explanation is that commuters using public transportation usually commute to bus stops by walking or cycling. Therefore, counties with a high proportion of commuters using public transportation usually have a high proportion of bicyclists and pedestrians and are more likely to provide a safer environment for bicyclists.

6.2. Pedestrian crash model

The results indicate that the DPMT is negatively related to the proportion of severe injuries among pedestrians, while it is positively related to the proportion of minor/no injuries. Thus, the existence of the safety-in-numbers effect is confirmed in the aspect of pedestrian injury severity. This viewpoint is also verified by the decrease in the proportions of severe injuries and moderate injuries relative to that of minor/no injuries, as shown in Table 5, and the spatial distributions of DPMT and the distribution of traffic injury severity among pedestrians in counties, as shown in Fig. 2. The reason for this phenomenon may be attributed to the fact that a larger number of pedestrians are more visible to motorists and are less likely to be overlooked by them (Jacobsen, Ragland, & Komanoff, 2015). Thus, drivers can undertake evasive actions earlier when encountered with a larger number of pedestrians, which can reduce injury severity in the event of a conflict.

Based on the proportion of minor/no injuries, the proportions of severe and moderate injuries increase as the proportion of recreational land use increases. It is speculated that this may be because the population in recreational areas is more likely to consume alcohol and is prone to excitement, which increases the proportion of severe injuries. The correlation between the alcohol-involved pedestrian crash rate and proportion of recreational land use was found to be positive ($r = 0.204$, $p = 0.098$), which verifies the above hypothesis. Accordingly, considerable attention should be paid to the environment (i.e., recreational areas) when planning pedestrian safety.

7. Conclusion

Walking and cycling are essential towards the development of sustainable transportation; however, crashes are obstacles to the promotion of this active mode of transportation. To promote walking and cycling among people, this study focused on exploring the existence of the safety-in-numbers effect in the aspect of injury severity among bicyclists and pedestrians. Two fractional split multinomial logit models were used at the macro level. The modeling results confirmed the presence of the safety-in-numbers effect from the perspective of injury severity among both bicyclists and pedestrians. This indicates that the probability of severe injuries resulting from a bicycle or pedestrian crash decreases as the number of bicyclists/pedestrians increases. This finding can encourage people to select active modes of transportation while traveling, thereby supporting long-term planning for sustainable transportation.

It is also revealed that a higher proportion of people aged 65 years and older is inclined to increase the proportion of severe injuries but decrease the proportion of moderate injuries among bicyclists. Counties with a high proportion of commuters using public transportation are likely to have a lower proportion of moderate injuries but a higher proportion of minor/no injuries among bicyclists. An increase in the proportion of recreational land use is related to an increase in the proportion of severe and moderate injuries relative to the reference (i.e., the proportion of minor/no injuries) among pedestrians. Therefore, more attention should be paid to counties with a high proportion of people aged 65 years and older or a high proportion of recreational areas to improve the safety of vulnerable road users.

We hope that the findings of this study will help policymakers and practitioners make informed decisions and take effective countermeasures to decrease injury severity among vulnerable road users. As Metropolitan Planning Organizations and other agencies encourage more people to walk and bike, it is important to have a target of their share in modal split to reach the safety-in-numbers effect and achieve less risk of severe injuries. Also, to convey these ideas to the public to encourage more of them to walk and bike, these concepts could be promoted at certain corridors, and more restrictions on vehicle travel could be suggested.

Although this study discovered key findings, it is not without limitations. First, the Strava data did not capture all bicyclists/pedestrians. This disadvantage of the Strava data may have led to biased results. Second, the study provides insights into the impacts of several explanatory variables on the proportion of injury severities among bicyclists and pedestrians; however, some variables (e.g., weather and air quality) were not considered. Finally, the current study was only conducted at the macro level. The heterogeneities between bicyclist/pedestrian individuals and the group size of bicycles/pedestrians cannot be considered in a macroscopic-level analysis. The existence of the safety-in-numbers effect with regard to injury severity at the micro level with a consideration of bicyclist/pedestrian individuals and the group size of bicycles/pedestrians is worth exploring using video data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the National Key R&D Program of China (No. 2020YFB1600400) and the Innovation-Driven Project of Central South University (No.2020CX013).

References

Afghari, A. P., Haque, M. M., & Washington, S. (2018). Applying fractional split model to examine the effects of roadway geometric and traffic characteristics on speeding behavior. *Traffic Injury Prevention, 19*(8). https://doi.org/10.1080/15389588.2018.1509208.

Bhatia, R., & Wier, M. (2011). "Safety in Numbers" re-examined: Can we make valid or practical inferences from available evidence? *Accident Analysis and Prevention, 43*(1). https://doi.org/10.1016/j.aap.2010.08.015.

Brüde, U., & Larsson, J. (1993). Models for predicting accidents at junctions where pedestrians and cyclists are involved. How well do they fit? *Accident Analysis & Prevention, 25*(5), 499–509.

Daniels, S., Brijs, T., Nuyts, E., & Wets, G. (2010). Explaining variation in safety performance of roundabouts. *Accident Analysis and Prevention, 42*(2). https://doi.org/10.1016/j.aap.2009.08.019.

De Oña, J., De Oña, R., Eboli, L., Forciniti, C., & Mazzulla, G. (2014). How to identify the key factors that affect driver perception of accident risk. A comparison between Italian and Spanish driver behavior. *Accident Analysis and Prevention, 73*. https://doi.org/10.1016/j.aap.2014.09.020.

Elvik, R. (2016). Safety-in-numbers: Estimates based on a sample of pedestrian crossings in Norway. *Accident Analysis and Prevention, 91*. https://doi.org/10.1016/j.aap.2016.03.005.

Elvik, R. (2017). Exploring factors influencing the strength of the safety-in-numbers effect. *Accident Analysis and Prevention, 100*. https://doi.org/10.1016/j.aap.2016.12.013.

Elvik, R., & Bjørnskau, T. (2017). Safety-in-numbers: A systematic review and meta-analysis of evidence. *Safety Science, 92*. https://doi.org/10.1016/j.ssci.2015.07.017.

Elvik, R., & Goel, R. (2019). Safety-in-numbers: An updated meta-analysis of estimates. *Accident Analysis and Prevention, 129*. https://doi.org/10.1016/j.aap.2019.05.019.

Fyhri, A., Sundfør, H. B., Bjørnskau, T., & Laureshyn, A. (2017). Safety in numbers for cyclists—conclusions from a multidisciplinary study of seasonal change in interplay and conflicts. *Accident Analysis and Prevention, 105*. https://doi.org/10.1016/j.aap.2016.04.039.

Hochmair, H. H., Bardin, E., & Ahmouda, A. (2019). Estimating bicycle trip volume for Miami-Dade county from Strava tracking data. *Journal of Transport Geography, 75*. https://doi.org/10.1016/j.jtrangeo.2019.01.013.

Jacobsen, P. L. (2015). Safety in numbers: More walkers and bicyclists, safer walking and bicycling. *Injury Prevention, 21*(4), 271–275. https://doi.org/10.1136/ip.9.3.205rep.

Jacobsen, P. L., Ragland, D. R., & Komanoff, C. (2015). Safety in numbers for walkers and bicyclists: Exploring the mechanisms. In. *Injury Prevention, Vol. 21, Issue 4*. https://doi.org/10.1136/injuryprev-2015-041635.

Johnson, M., Oxley, J., Newstead, S., & Charlton, J. (2014). Safety in numbers? Investigating Australian driver behaviour, knowledge and attitudes towards cyclists. *Accident Analysis and Prevention, 70*. https://doi.org/10.1016/j.aap.2014.02.010.

Kröyer, H. R. G. (2016). Pedestrian and bicyclist flows in accident modelling at intersections. Influence of the length of observational period. *Safety Science, 82*. https://doi.org/10.1016/j.ssci.2015.09.015.

Lee, J., Abdel-Aty, M., & Cai, Q. (2020). Investigation of safety-in-numbers for pedestrians and bicyclists at a macroscopic level with various exposure variables. In. *Transportation Research Record, 2674*(12). https://doi.org/10.1177/0361198120953795.

Lee, J., Abdel-Aty, M., Xu, P., & Gong, Y. (2019). Is the safety-in-numbers effect still observed in areas with low pedestrian activities? A case study of a suburban area in the United States. *Accident Analysis and Prevention, 125*. https://doi.org/10.1016/j.aap.2019.01.037.

Lee, J., Yasmin, S., Eluru, N., Abdel-Aty, M., & Cai, Q. (2018). Analysis of crash proportion by vehicle type at traffic analysis zone level: A mixed fractional split multinomial logit modeling approach with spatial effects. *Accident Analysis and Prevention, 111*. https://doi.org/10.1016/j.aap.2017.11.017.

Miranda-Moreno, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention, 43*(5). https://doi.org/10.1016/j.aap.2011.02.005.

Mueller, N., Rojas-Rueda, D., Salmon, M., Martinez, D., Ambros, A., Brand, C., ... Nieuwenhuijsen, M. (2018). Health impact assessment of cycling network expansions in European cities. *Preventive Medicine, 109*. https://doi.org/10.1016/j.ypmed.2017.12.011.

Murphy, B., Levinson, D. M., & Owen, A. (2017). Evaluating the Safety In Numbers effect for pedestrians at urban intersections. *Accident Analysis and Prevention, 106*. https://doi.org/10.1016/j.aap.2017.06.004.

Page, Y. (2001). A statistical model to compare road mortality in OECD countries. *Accident Analysis and Prevention, 33*(3). https://doi.org/10.1016/S0001-4575(00)00051-8.

Papke, L. E. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics, 11*(6). https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1.

Phillips, R. O., Bjørnskau, T., Hagman, R., & Sagberg, F. (2011). Reduction in car-bicycle conflict at a road-cycle path intersection: Evidence of road user adaptation? *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(2). https://doi.org/10.1016/j.trf.2010.11.003.

Robinson, D. L. (2005). Safety in numbers in Australia: More walkers and bicyclists, safer walking and bicycling. *Health Promotion Journal of Australia : Official Journal of Australian Association of Health Promotion Professionals, 16*(1). https://doi.org/10.1071/he05047.

Schneider, R. J., Diogenes, M. C., Arnold, L. S., Attaset, V., Griswold, J., & Ragland, D. R. (2010). Association between roadway intersection characteristics and pedestrian crash risk in Alameda County, California. *Transportation Research Record, 2198*. https://doi.org/10.3141/2198-06.

Tasic, I., Elvik, R., & Brewer, S. (2017). Exploring the safety in numbers effect for vulnerable road users on a macroscopic scale. *Accident Analysis and Prevention, 109*. https://doi.org/10.1016/j.aap.2017.07.029.

Thompson, J., Savino, G., & Stevenson, M. (2015). Reconsidering the Safety in Numbers Effect for Vulnerable Road Users: An Application of Agent-Based Modeling. *Traffic Injury Prevention, 16*(2). https://doi.org/10.1080/15389588.2014.914626.

Thompson, J., Savino, G., & Stevenson, M. (2016). A model of behavioural adaptation as a contributor to the safety-in-numbers effect for cyclists. *Transportation Research Part A: Policy and Practice, 85*. https://doi.org/10.1016/j.tra.2015.12.004.

U.S. Census Bureau (2021). 2017 National Population Projections Tables: Main Series. https://www.census.gov/data/tables/2017/demo/popproj/2017-summary-tables.html.

U.S. Department of Transportation, Federal Highway Administration, 2017 National Household Travel Survey. URL: https://nhts.ornl.gov.

World Health Organization. (2021). *Road traffic injuries*. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

Xie, S. Q., Dong, N., Wong, S. C., Huang, H., & Xu, P. (2018). Bayesian approach to model pedestrian crashes at signalized intersections with measurement errors in exposure. *Accident Analysis and Prevention, 121*. https://doi.org/10.1016/j.aap.2018.09.030.

Xu, P., Xie, S., Dong, N., Wong, S. C., & Huang, H. (2019). Rethinking safety in numbers: Are intersections with more crossing pedestrians really safer? *Injury Prevention, 25*(1). https://doi.org/10.1136/injuryprev-2017-042469.

Yasmin, S., Eluru, N., Lee, J., & Abdel-Aty, M. (2016). Ordered fractional split approach for aggregate injury severity modeling. *Transportation Research Record, 2583*. https://doi.org/10.3141/2583-15.

Ms. Yanqi Lian is a Ph.D. student at the School of Traffic & Transportation Engineering of the Central South University. She received bachelor's degree of Traffic & Transportation Engineering (Honors) from the Central South University in 2018. Her main research interests include big data analytics and vulnerable road users' safety.

Ms. Enru Zhou is a Master's student at the School of Traffic & Transportation Engineering of the Central South University. She received bachelor's degree of Transportation Engineering from the Shanghai University of Engineering Science in 2020. She has been working on research studies related to traffic safter analysis, public health and transportation.

Dr. Jaeyoung Lee is Professor at the School of Traffic & Transportation Engineering of the Central South University. Since the last decade, he has conducted research studies in traffic safety, traffic planning, operation and management of transportation, transportation big data, and intelligent transportation systems. He has published over 150 academic papers. He is a member of multiple TRB Standing Committees including Transportation Safety Management (ACS10) and Impairment in Transportation (ACS50). He is also a Courtesy Professor at the University of Central Florida.

Dr. Mohamed Abdel-Aty, PE is a Trustee Chair at the University of Central Florida (UCF). He is a Pegasus Professor and the Chair of the Civil, Environmental, and Construction Engineering Department at UCF. He is leading the Future City initiative at UCF. He is also the director of the Smart and Safe Transportation Lab1, the Winner of the USDOT Solving for Safety Visualization Challenge, Real-time crash risk visualization using integrated tools for traffic safety evaluation and management, November 2019. His main expertise and interests are in the areas of traffic safety, simulation, big data and data analytics, ITS, and CAV. He is the pioneer and well recognized nationally and internationally in work and research in real-time safety, Proactive traffic management, integrating road safety and transportation planning, Highway Safety Manual, and Connected Vehicles.

Journal of Safety Research 83 (2022) 128–138

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Factors impacting bike crash severity in urban areas

Ishita Dash *, Mark Abkowitz, Craig Philip

Department of Civil and Environmental Engineering, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA

A R T I C L E   I N F O

Article history:
Received 9 December 2021
Received in revised form 1 April 2022
Accepted 16 August 2022
Available online 30 August 2022

Keywords:
Bicycling
Sustainable transportation
Machine learning
Safety
Policy

A B S T R A C T

Introduction: Bicycling plays an important role as a major non-motorized travel mode in many urban areas. While increasingly serving as a key part of an integrated transportation demand management system and a sustainable mobility option, interest in biking as an active transportation mode has been unfortunately accompanied by an increase in the number of bike crashes, many with incapacitating injuries or fatal outcomes. Thus, to improve bicycling safety it is crucial to understand the critical factors that influence severe bicyclist crash outcomes, and to identify and prioritize policies and actions to mitigate these risks. *Method*: The study reported herein was conducted with this objective in mind. Our approach involves the use of classification models (logistic regression, decision tree and random forest), as well as techniques for treating unbalanced data by under sampling, oversampling, and weighted cost sensitivity (CS) learning, applied to bike crash data from the State of Tennessee's two largest urban areas, Nashville and Memphis. *Results*: The results indicate that random forest with weighted CS offers the potential for greater explanatory accuracy, an important observation given the paucity of efforts to date in applying random forest to bike safety studies. Inadequate lighting conditions, crashes on roadways, speed limits, average annual daily traffic, number of lanes, and weekends are the critical features identified. *Conclusion*: Based on these results, a series of specific, suggested policy changes are presented for implementation consideration. *Practical Applications*: There is existing guidance in FHWA Lighting Handbook and TDOT's Roadway Design Guidelines that spell out some engineering design solutions like lighting provisions, bicycle facility design, and traffic calming measures. These measures may alleviate the identified key features impacting fatal and incapacitating bicycle injuries. Further research should be conducted to gauge the efficacy of the solutions suggested.

1. Introduction

Bicycling represents a relatively small portion of the total commuting activity in the United States, but this non-motorized travel mode plays an important role in many of the nation's urban areas. Biking is a relevant part of many emerging integrated transportation demand management systems and offers a sustainable mobility option, with a lower carbon footprint should commuters choose to switch from modes that rely on traditional fuel sources. This has prompted several state and local agencies to take steps to promote biking by employing strategies such as sidewalk modifications and construction of dedicated bike lanes. In recent years, the number of cities with bicycle sharing programs has also increased dramatically. These developments have resulted, however, in an increase in bike crashes, many with incapacitating injuries or fatal outcomes. Therefore, it is important to improve our understanding of the critical factors impacting bike crashes in urban areas, aiming towards developing risk mitigation strategies to curb this crash trend. This paper discusses an analysis performed with this intent.

Our study objective is to determine bike road safety in select urban areas within the State of Tennessee using detailed crash data to investigate the factors affecting bike crashes with incapacitating and fatal outcomes, and to subsequently develop a classification model for fatal or incapacitating events. The paper concludes with a policy discussion directed at enhancements to transportation infrastructure and operations with bicycle safety in mind.

2. Background

Commuting on a bicycle is the third most utilized U.S. transportation mode and is quickly gaining popularity as a commuting option. The number of commuters biking to work has increased by 65% nationwide from 2000 to 2019 (US Census Bureau, 2014 & 2021). Unfortunately, with increased usage there are also alarming trends involving fatal and incapacitating bicycle crashes. Traffic

* Corresponding author.
E-mail address: ishita.dash@vanderbilt.edu (I. Dash).

https://doi.org/10.1016/j.jsr.2022.08.010
0022-4375/

hazards for bicyclists include poorly designed roads, high motor-vehicle speeds, and lack of responsibility exhibited by other road users (Furth et al., 2016; Jacobsen & Rutter, 2017). In 2019, bikers accounted for 0.5% of 156 million commuters; however, of all traffic crashes, bikers account for 1.78% of injury crashes and 2.3% fatal and serious (incapacitating) injuries for the entire nation. Fatal and serious bike injuries have seen a 36% increase since 2010 (NHTSA, DOT HS 813 197: Traffic Safety Facts, 2019), indicating bicyclists were among the most vulnerable users in terms of being disproportionately impacted (Jacobsen & Rutter, 2017; Smart Growth America, 2020).

The disturbingly high number of crashes involving bicycles resulting in fatal or incapacitating injury outcomes leads one to question whether the transportation infrastructure and operations lack accessible and safe facilities for bikers, which can be problematic when bikers must share roads with other users, particularly motor vehicles. Bicyclists are considered among the most vulnerable participants in mixed traffic because of the kinetic energy produced upon crashes between two differential masses where one is traveling at a higher velocity and mass (Jacobsen & Rutter, 2017). In case of an automobile colliding with a cyclist, speeds above 20 miles per hour increase the risk of severe road injury or fatality (Jacobsen & Rutter, 2017; Jurewicz et al., 2016). Heavily utilized urban corridors, therefore, impose a potentially significant danger to cyclists, if not provided with adequate safety measures (NTSB, 2019).

3. Literature review

Bicycle crashes have been studied by researchers worldwide. Many of these efforts have been directed at individual areas or regions for the purpose of identifying and rectifying safety issues within the bicycle infrastructure and operations. The most common modeling techniques have included use of the Poisson distribution, negative binomial models, linear regression models, logit models, ordered probit models, and multivariable logistic regression. Table 1 lists results of significant factors found in previous studies, organized according to field type and variable. Table 2 summarizes relevant study methodologies.

With regard to Table 2, note that the use of random forest modeling is not included. Studies modeling bicycle injury prediction using random forest are currently in their infancy, such as one examining bicyclist only crashes in Victoria, BC, Canada; however, the dataset consists of only 111 crashes and 234 near misses and was collected via surveys rather than from official crash records.

4. Data analysis

The bicycle crash data utilized in this analysis was obtained from the Tennessee Department of Transportation (TDOT) for the period of January 1, 2017 through December 31, 2020, covering the entire state. In Tennessee, a crash is reported when a driver of a vehicle is involved in a crash resulting in injury, death or property damage exceeding $50 (Tennessee Code Title 55. Motor and Other Vehicles § 55-10-106). A crash is also reported when a vehicle collides with an unattended vehicle (Tennessee Code Title 55. Motor and Other Vehicles § 55-10-104), such as one located in a parking lot. Crash data obtained from TDOT and used for this study consists of only bicyclist-motor vehicle crashes. Attributes associated with each crash record are listed in Appendix I.

During this period, 5,347 bike crashes were recorded for which there was complete information (see Table 3), distributed across the state as shown in Fig. 1. Of the 95 counties in Tennessee (TN), Shelby County and Davidson County recorded the highest bike crashes, collectively accounting for 2,942 incidents, more than one-half of the overall state total. This is to be expected since these two counties are densely populated and include the cities of Memphis and Nashville, respectively. As a result, these two locations subsequently became the focus of the modeling effort.

Bike crash severity results for the two counties are shown in Fig. 2. TDOT crash data include an injury severity attribute according to whether there was no injury, non-serious injury, serious (incapacitating) injury or fatality. An incapacitating injury is one that results in one or more of the following: (1) severe laceration resulting in exposure of underlying tissues/muscle/organs or resulting in significant loss of blood; (2) broken or distorted extremity (arm or leg); (3) crush injuries; (4) suspected skull, chest or abdominal injury other than bruises or minor lacerations; (5) significant burns (second and third degree burns over 10% or more of the body); (6)

Table 1
Significant bicycle crash factors from prior studies.

Variable Field	Variable Analyzed	Relevant Studies
Environmental	Lighting	Zangenehpour et al., 2016
	Weather	Yan et al., 2011
	Intersection Type	Klop and Khattak, 1999
	Speed Limit	Allen-Munley et al., 2004
	Traffic Control Device	Strauss et al., 2015
	Number of Lanes	Reynolds et al., 2009
	Road Curvature	Turner et al., 2011
	Traffic Volume (AADT)	Lee and Abdel-Aty (2005)
	Land Use (urban, rural, residential, industry, farmland, institutional, commercial)	Petritsch et al. (2006)
		Pai (2011)
		Schepers and den Brinker (2011)
		Dixon et al. (2012)
		Kim et al. (2007)
		Eluru et al. (2008)
		Oh et al. (2008)
		Vandenbulcke et al. (2014)
Crash Specific	Crash Type	Wang et al., 2015
	Severity	Klop and Khattak, 1999
		Allen-Munley et al., 2004
Time	Year	Wang et al., 2015
	Month	
	Day	
	Hour	

Table 2
Previous methodologies and modeling techniques.

Modeling Technique	Author	Study Focus	Variables Analyzed
Poisson Distribution	Oh et al. (2008)	Bicycle Crash at Urban Signalized Intersections	Average daily traffic volume, presence of bus stops, sidewalk widths, number of driveways, presence of speed restrict devices, and presence of crosswalks are all statistically significant risk factors.
Negative Binomial Model	Oh et al. (2008)	Bicycle Crash at Urban Signalized Intersections	Found different types of facility designs impact bicycle safety such as bike lanes, bike track, pavement markings or colors.
	Wang and Nihan (2004)	Bicycle - Motor Vehicle Crashes at a Signalized Intersection	Intersection design impacts on bicycle safety in multiple ways.
Linear Regression	Dixon et al. (2012)	State Highways	For intersection and network movement, hazardous crossings, right hook, left sneak and complicated interactions are potentially dangerous to bicyclists. Intersection safety influenced by vehicle volume, vehicle speed, percentage of heavy vehicles, among others.
Logit Model	Eluru et al. (2008)	Road Segments	Crashes on curved/non-flat roadways tend to result in more severe injuries.
	Kim et al. (2007)	Bicycle-Motor Vehicle Crashes	Curved rounds significantly increase the chance that a fatal or incapacitating injury will occur during a vehicle-bicycle crash.
	Pai (2011)	Road Segments	Horizontal and vertical curves can contribute to bicycle crashes.
	Schepers and den Brinker (2011)	Road Segments	Bicyclists colliding with a bollard, road narrowing or riding off a curve found to occur more than when bicyclists hit an obstacle. More crashes were observed where the bicycle had the right-of-way on a through movement at intersections with two-way bicycle tracks that are well marked and are reddish in color. Fewer crashes occurred when there are raised bicycle crossings (speed humps) or other speed reduction measures.
	Abdel-Aty and Keller (2005)	Signalized Intersections	The division of a minor road, as well as a higher speed limit on the minor road lowered the expected injury level, while a median on the minor road may prevent more head-on crashes, which were found to be more severe crashes.
	Haleem and Abdel-Aty (2010)	Unsignalized Intersections	Traffic volume on the major approach, number of through lanes on the minor approach, upstream and downstream distance to the nearest signalized intersection, left and right shoulder width, number of left-turn movements on the minor approach, and number of right- and left-turn lanes on the major approach are significant factors influencing bicycle risk.
Decision Tree	Rahman (2018)	Pedestrian & Bicycle Crashes	Highlighted the most significant predictor variables for pedestrian and bicycle crash count in terms of three broad categories: traffic, roadway, and socio demographic characteristics
Bayesian Model	Vandenbulcke et al. (2014)	Selected Controlled Sites or Bikeable Road Network	Right-of-way intersections equipped with bicycle lanes tend to have higher crash risk for cyclists, due to vehicles not respecting the right-of-way (i.e., right-hook crashes). Cyclists riding on marked bicycle lanes in roundabouts and signalized intersections with marked cycle lanes had higher crash risk, attributed to bicyclists being in drivers' blind spots. Additionally, complex intersections (high number of road legs, road users, high number of signs, dense traffic crossings, etc.), and therefore complex traffic situations, increase bicycle risk.
Safety Analyst and Clustering Algorithm	Dolatsara (2014)	Roadway Segments in Michigan	Exposure, the presence of bicycle lanes and bus stops, and the number of left-turn lanes at intersections are positively associated with bicycle crashes.

Table 3
TN bike crashes by year.

Year	Total Bike Crashes
2017	1,384
2018	1,299
2019	1,432
2020	1,232

unconsciousness when taken from the crash scene; and (7) paralysis. Particularly notable is that fatal and incapacitating injury collisions account for 27% of 2,942 recorded crashes.

Fig. 3 shows the distribution of bike crashes by time of day, where times have been grouped into the following categories: (1) Early Morning (midnight-5:00 am); (2) Morning (5:00–9:00 am); (3) Peak AM (9:00 am-1:00 pm); (4) Afternoon (1:00–5:00 pm); (5) Peak PM (5:00–9:00 pm); and (6) Late Evening (9:00 pm-midnight). Note that while the frequency of bicycle crashes tends to increase as the day goes along, the percentage of those that result in incapacitating and fatal injuries are highest during the earlier part of the day.

As displayed in Fig. 4, the largest number of bike crashes in general as well as those resulting in a fatality or incapacitating injury occurred on four lane roads, two lanes in each direction. It was also observed that four lane roads experience high bicycle injuries on medians and in turn lanes.

As seen in Fig. 5, roads with speed limits from 30 mph to 45 mph experience a significant number of bicycle crashes, with the proportion of those resulting in a fatality or incapacitating injury increasing at higher speeds. This observation is consistent with prior studies (Isalsson-Hellman & Toreki, 2019).

5. Modeling approach

An overview of the process used in developing a predictive model of bicycle crash severity is shown in Fig. 6. Model estimation was performed using Logistic Regression (LR), Decision Tree (DT), and Random Forest (RF). LR serves as a baseline for our binary classification problem and represents a widely used method to study risk factors impacting injury severity. DT is another frequently used classification algorithm for undestanding and interpreting data, where the top node is the root node, representing the best feature that divides the data. Each internal node is a feature and branches indicate the decision, with class label being represented by a leaf node. DT serves as a foundation for RF.

Although RF has not been extensively used as a classification algorithm for analyzing bicycle crashes, it was included because RF has been shown to improve modeling performance relative to a single tree classifier (e.g., DT) and LR. RF enables multiple uncor-

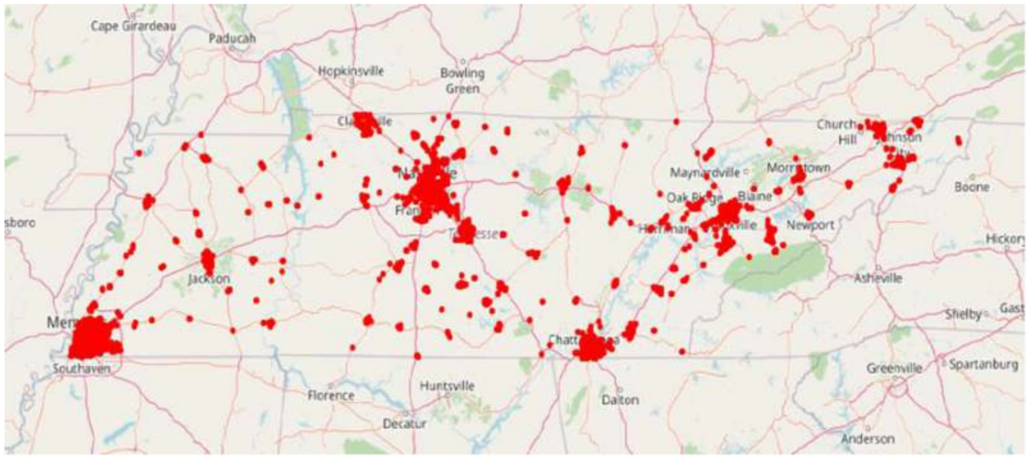

Fig. 1. Bicycle crashes cluster in TN (2017–2020).

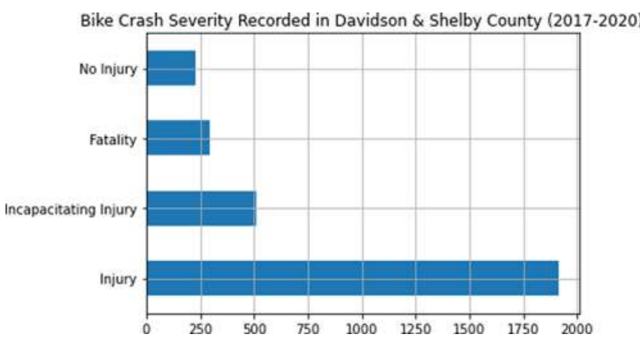

Fig. 2. Bike crash severity in Davidson and Shelby County.

related DTs to grow, thus creating a forest. RF uses a technique called feature bagging, where features are selected randomly for individual DTs, which is similar to bagging procedure. With feature bagging, the correlation between each DT is reduced but the overall accuracy of the model increases. RF performs better as it is more robust to noise, able to capture the non-linear tendencies by putting all the weak learners in an ensemble that is used to make the prediction. It also avoids overfitting because those individual learners are weak, so it is not one massive model that could lead to overfitting the data (Müller et al., 2017).

In this study, we elected to use LR followed by DT and RF to observe the model prediction outcome. It is not necessary to use models that build on the previous ones; however, this was done to tune the classifier and improve model performance.

The dependent variable was defined as a numerical Boolean variable, with a value of 1 indicating a fatal or incapacitating injury outcome, and 0 otherwise (i.e., minor injury or no injury). Prior to conducting model estimation, data pre-processing was performed to remove records with missing data, following that exploratory data analysis was performed. This resulted in the selection of the following candidate crash factors (attributes) to be considered as independent variables in model estimation: location, functional class, number of lanes, speed limit, average annual daily traffic (AADT), impaired driver, weather, lighting, and weekend. Categorical values for location (roadway, intersection, bridge, ramp), functional class (urban, rural), impaired driver (yes, no), weather (clear, cloudy, rain, fog, snow, severe cross wind, sleet, hail), lighting (dark, dawn, daylight, dusk), and weekend (yes, no) were converted to numerical Boolean variables (0 or 1). AADT, speed limit, and number of lanes were scaled to help decrease the magnitude as per a fixed ratio; this process assists with reducing fluctuations in model performance.

The data set was divided where 80% of the observations were used for training and the remaining 20% for testing. We attempted to balance the training data before model insertion. Note that, as

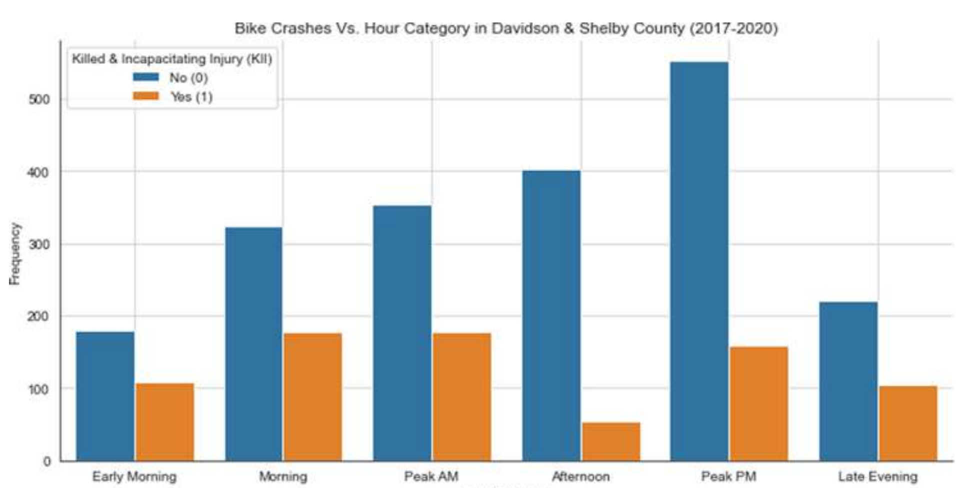

Fig. 3. Bike crashes by time of day in Davidson & Shelby County.

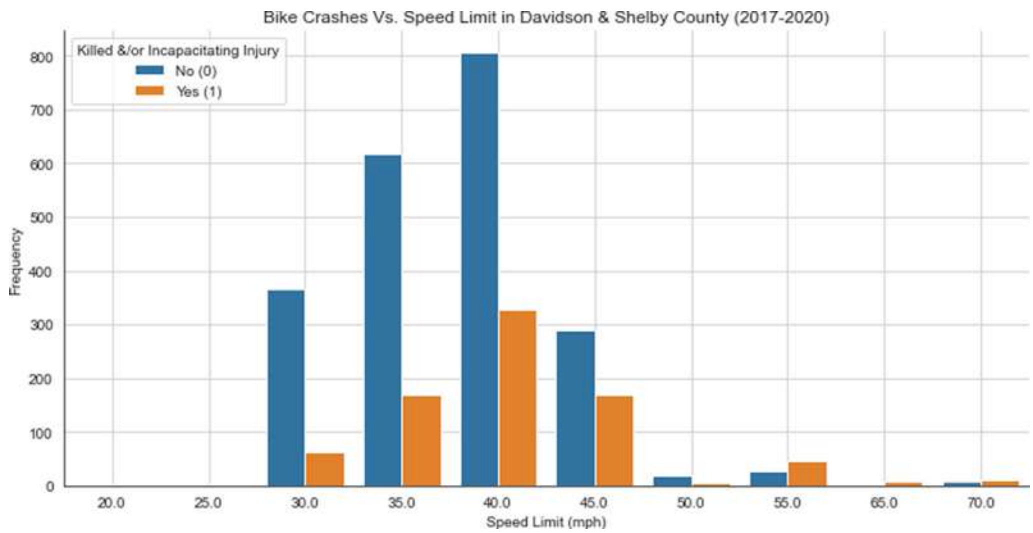

Fig. 4. Bike crashes by lane configuration.

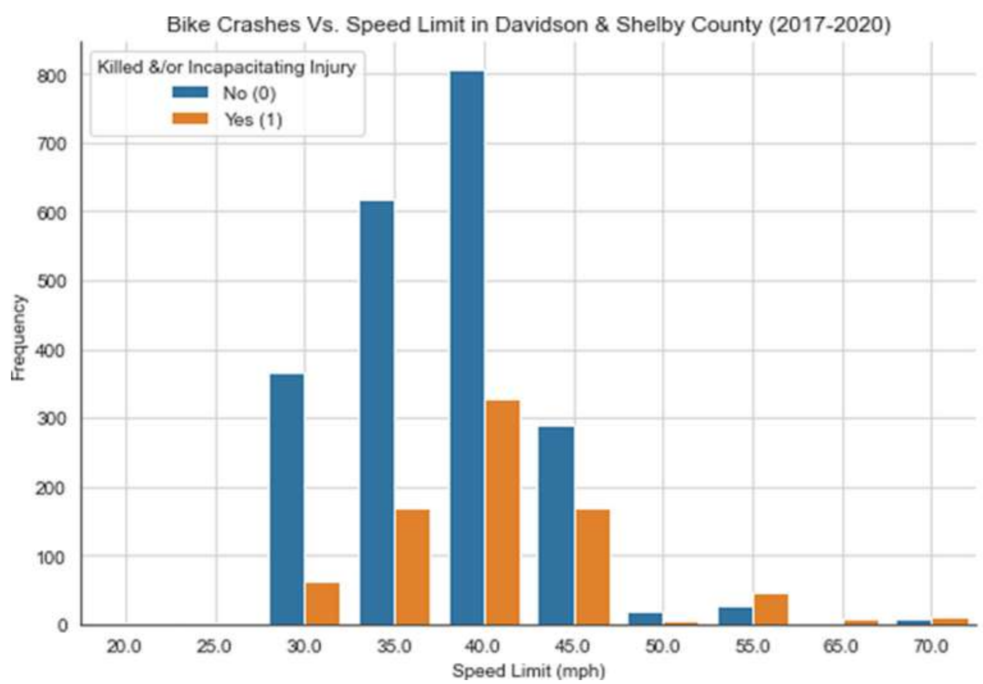

Fig. 5. Bike crashes by road speed limit.

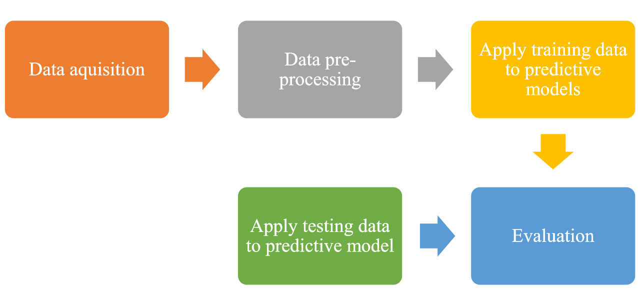

Fig. 6. Predictive model development framework.

shown in Fig. 7, the dependent variable is unevenly distributed in the training dataset, with 27% of bike crashes resulting in a fatality and/or incapacitating injury (i.e., minority class), and 73% of bike crashes resulting in no fatality and/or incapacitating injury (i.e., majority class). Attempts to balance the data for model estimation can be performed by decreasing the majority class sample size (under-sampling) or increasing the minority class sample size (oversampling). Another balancing approach involves the use of cost sensitive learning (CSL), whereby a larger weight is assigned to the minority class and a smaller weight is applied to the majority class.

All three of these sampling techniques were applied as part of the modeling effort. We used the NearMiss algorithm for under-sampling to prevent the problem of information loss in most traditional under-sampling techniques. Synthetic Minority Oversampling Technique (SMOTE) was applied for oversampling, a technique where synthetic samples are generated for the minority class that helps to overcome the overfitting problem posed by traditional random oversampling techniques.

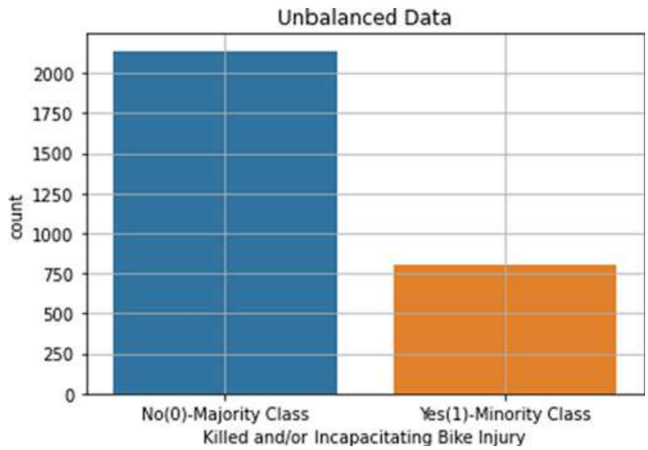

Fig. 7. Unbalanced data for the dependent variable.

As the dataset consists of both numerical and categorical inputs, three methods of feature selection were applied in sequential fashion: (1) Correlation coefficient, (2) DT feature importance, (3) Recursive Feature Elimination (RFE). A correlation coefficient threshold of ±0.7 was applied to eliminate highly correlated features. Non-zero features were selected using DT on the training data. The RFE starting point was the set of features filtered using DT in the training dataset.

Some limitations were identified in the crash data. The data used for this study consists of only bicyclist –motor-vehicle crashes. All the data are recorded at the scene of crash by law enforcement officers. Once this police report is filed, it is then entered into the data platform. Hence, these data can suffer from human error when reported, collected, and processed at the various stages. TDOT does not include near misses and unreported bike incidents. The data do not provide information on the cause of crash, which party involved in the crash was injured (although we assume that if any injuries are reported, it at a minimum involves a bicyclist), nor any details on how or in which direction the involved parties where moving (i.e., circumstances prior to crash).

6. Model results

The overall model performance measure is the extent to which the model can accurately predict whether a bike crash results in a fatality or incapacitating injury. Appendix II provides a list of relevant metrics and their corresponding definitions for indicators considered in evaluating the efficacy of model performance.

Table 4 summarizes the performance metrics for various estimated models using the features that emerged from the aforementioned elimination process: lighting (dark), number of lanes, speed limit, AADT, weekend, and location type (roadway).

As shown in Table 4, three models (oversampled LR, weighted CSL applied to both LR and RF) perform well. However, weighted CSL applied to RF performs slightly better due to its higher true negativity rate (0.63) and true positivity rate (0.77) and lower Type I and Type II errors. Moreover, RF with weighted CSL has the highest value of Geometric-mean (0.7) and weighted accuracy (0.7).

The Receiver Operating Characteristic (ROC) curve value for RF with weighted CSL is also high for the testing data (0.7) and varies the least (0.01) from the training data (0.71). This measure is derived from a curve plotted on a graph showing the performance of a classification model at different classification thresholds. This curve plots two parameters: true positive rate (TPR) and false positive rate (FPR). The ROC measures the area under the curve; when the ROC is closer to 1 but greater than 0.5, it indicates a strong model.

Shapley additive explanations (SHAP), as shown in Fig. 8, measure the contribution of a feature in model prediction (Apley & Zhu, 2020). Note that both classes use the same feature equally (i.e., all features have equal impact on model prediction). Among these features, dark lighting and roadway crash location are the most important factors affecting bike crash severity, while roads with higher motor-vehicle speed limits, heavy traffic, multilane roads, and weekend travel are also significant contributors.

Fig. 9 displays a bee swarm plot for the study data. This plot helps one understand how a variable may influence model prediction. In this plot, every record in the database is shown as a dot on each row. The color of the dot represents the value of that feature for the event, with red indicating a high value and blue a low value. Here, one can observe that for Class 1 (killed and/or incapacitating bike injury), when the lighting condition is inadequate and location type is roadway, it is more likely to result in a killed and/or incapacitating bike injury.

Understanding prediction for individual instances can provide meaningful information, as it explains how individual predictions are reached in terms of feature contribution. To illustrate, we selected this information for two bicycle crash records, one that resulted in a killed or incapacitating injury (Event 419), and another where the outcome was not a killed or incapacitating injury (Event 422). Using the feature inputs for Event 419, the model predicts a killed and/or incapacitating bike injury with 0.71 probability. This compares with when we do not know any features for a specific event, in which case the average model output over the training dataset is 0.4995 (base value). In the case of Event 422, the model predicts a no killed and/or incapacitating

Table 4
Performance metrics for various model estimation techniques.

Performance Metrics	True Negative Rate	True Positive Rate	False Negative Rate	False Positive Rate	Geometric-Mean	Weighted Accuracy	Receiver Operating Characteristics - Train	Receiver Operating Characteristics - Test
LR – Unbalanced	0.18	0.94	0.062	0.82	0.41	0.56	0.58	0.56
LR – Undersample	0.57	0.68	0.32	0.43	0.62	0.625	0.64	0.63
LR – Oversample	0.71	0.68	0.32	0.29	0.69	0.695	0.67	0.7
LR – Weighted CSL	0.68	0.69	0.31	0.32	0.68	0.685	0.68	0.68
DT– Unbalanced	0.31	0.89	0.11	0.69	0.53	0.6	0.68	0.6
DT– Undersample	0.62	0.66	0.34	0.38	0.64	0.64	0.69	0.64
DT – Oversample	0.5	0.78	0.22	0.5	0.62	0.64	0.75	0.64
DT – Weighted CSL	0.66	0.7	0.3	0.34	0.68	0.68	0.71	0.68
RF – Unbalanced	0.2	0.95	0.055	0.8	0.44	0.575	0.6	0.57
RF – Undersample	0.63	0.62	0.38	0.37	0.62	0.625	0.7	0.63
RF – Oversample	0.56	0.76	0.24	0.44	0.65	0.66	0.77	0.66
RF – Weighted CSL	0.63	0.77	0.23	0.37	0.70	0.7	0.71	0.70

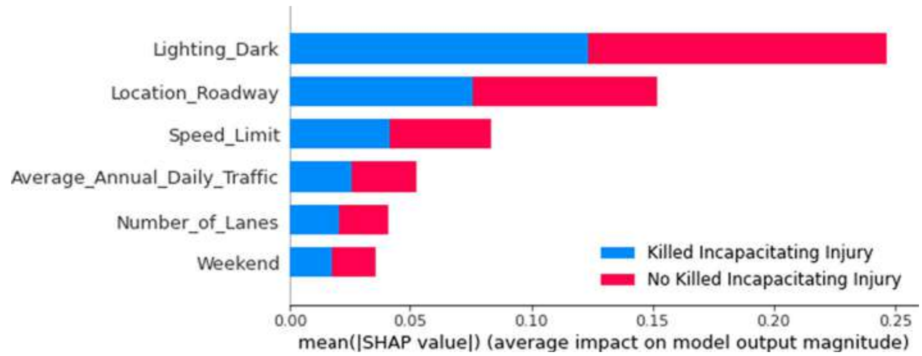

Fig. 8. Summary plot displaying SHAP values for model features.

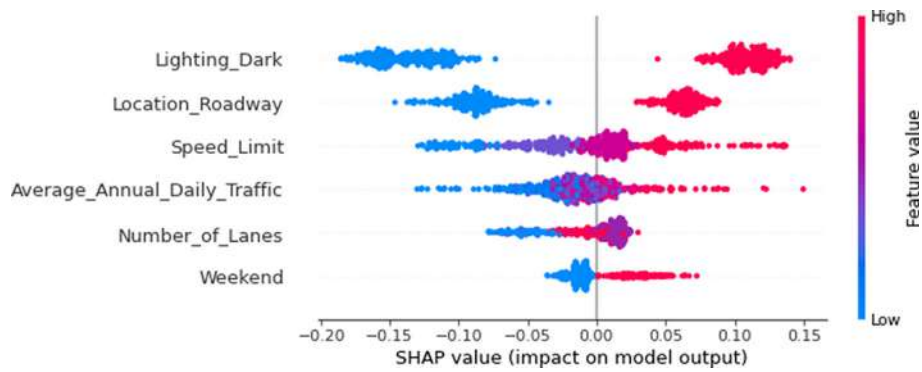

Fig. 9. Summary plot combining feature importance with feature effect for class 1 (killed and/or incapacitating bike injury).

bike injury with 0.74 probability, compared to a base value of 0.5005.

Fig. 10 helps identify groups of similar instances by using hierarchical agglomerative clustering to order the instances. Each position on the x-axis is an event in the database, where red plots increase the model prediction and blue decreases it. A cluster is observed towards the right of the curve with high prediction of killed and/or incapacitating bike injury.

The heavy influence of inadequate lighting conditions on bike crash severity is a finding consistent with prior studies and is the largest factor influencing bicycle injury severity (Asgarzadeh et al., 2018), with Kim et al. (2007) concluding that the probability of a fatal bike injury doubles in the absence of streetlighting. The magnitude of this factor in the model results suggests that risk mitigation strategies should seriously consider improvements to lighting infrastructure.

Many previous studies have focused on crashes along the intersections since they have the highest conflict points. However, within our study database, more than one-half of the bike crashes occurred on non-intersection segments, one reason why this feature emerged as a significant explanatory factor for serious injury outcomes. Asgarzadeh et al. (2017) similarly found these locations

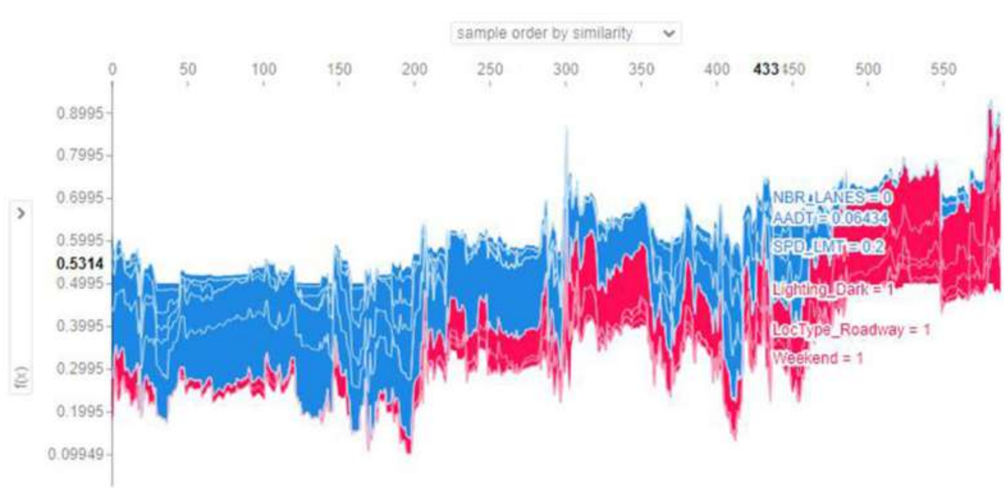

Fig. 10. Clustering based on features for class 1 (killed and/or incapacitating bike injury).

to be important, reporting that crashes on non-intersection segments are more likely to result in 1.31 times higher injury severity.

An increase in speed limit is also positively associated with a killed and/or incapacitating biker injury outcome. Chimba et al. (2012) noted a similar relationship when comparing crashes on roads with speed limits of 30 mph to those with a 35–45 mph speed limit. Fridman et al. (2020) describe several case studies that significantly reduce the likelihood of a killed and/or incapacitating bike (and pedestrian) injury by lowering road speed to 20 mph.

As a larger number of motor vehicles (AADT) travel across a road segment, it creates greater opportunity for crash exposure. Therefore, it is not surprising that biker injury frequency would increase; however, it is less clear based on model results that AADT alone accounts for more severe injury outcomes. This may be explained by the presence of other related factors such as vehicle speeds and number of lanes.

The same can be said for the significance of number of roadway lanes as an explanatory factor in predicting a serious biker injury outcome. In fact, the combination of multilane roads with higher speed limits being associated with higher risk of fatal or incapacitating injuries is one of the most consistent findings across the literature (Siddiqui et al., 2012; Huang et al., 2010; Lee et al., 2015; Noland & Quddus, 2004; Quddus, 2015; Wier et al., 2009; Yu & Zhu, 2015).

Finally, the weekend effect, albeit small, influences the likelihood of a bike crash causing a killed and/or incapacitating injury. Similar findings were observed in research performed by Shubo et al. (2021).

7. Policy implications

The feature importance associated with the selected model (see Fig. 8) provides insights into key factors that most influence serious biker safety outcomes as well as their relative contribution to those impacts. Foremost among these are issues related to lighting followed by roadway type; risk mitigation strategies aimed at these factors merit the most serious consideration.

Regarding lighting conditions, relatively simple risk mitigation strategies would include placement of street lighting along popular bike routes. In addition to improving illumination by providing more and improved street lights, it would also help if the bikers made their presence known on a roadway by wearing reflective materials and installing blinking lights on their bikes. The latter is required by the traffic law in Tennessee (TN), especially at nighttime. It goes without saying that personal protective wear, which includes helmets, should always be worn by the bicyclist.

While the relationship between roadways and serious bike crash outcomes is clear, the particular built environment and usage may influence exposure; hence, the reason why higher AADT's and number of lanes also contribute to the problem. Controlling for bicyclist exposure, Kaplan and Prato (2015) concluded that separated bicycle facilities reduce both bicyclist injury crashes and fatal crashes, whereas on-street bike lanes do not. This suggests that efforts to create dedicated bikeways that are physically separated from the roadway would be a more effective, albeit a more expensive, risk management strategy. In the absence of resources to provide these means, creating sufficient street width for an on-street bike lane is paramount, as most bicycle lanes today are placed between the vehicular route and the curb, often at widths of no more than four feet (including the 1–2 feet gutter pan as part of the bicycle lane). This problem is compounded by motorist expectations that bicyclists will remain in their dedicated lane, even when physically unable to do so. It is further exacerbated by the presence of "mixing zones," which are placed in advance of right-turn lanes to allow vehicles to cross the bicycle lane to enter the right-turn lane. When combined with adequate signage and other demarcations, these intervention strategies should help alleviate at least some crashes and reduce the impact of others when they occur.

Regarding speed limits, we recommend reviewing all urban streets with speed limits above 30 mph to assess whether the limit should be lowered. When this is not deemed a viable strategy, signage with dynamic message boards could be placed at vulnerable locations, reminding motorists to obey speed limits. Another strategy would be to deploy speed sensors coupled with speed cameras (either mobile or fixed) at vulnerable locations that display the actual speed of a passing vehicle, which flashes when the speed limit is being exceeded. Speed bumps and roundabouts are other options to slow vehicular traffic speed along the roadway and at intersections.

While recommendations for improving bike safety are encoded into bicycle design guidance (American Association of State Highway and Transportation Officials, 2014; National Association of City Transportation Officials, 2014), the widespread use of bike lanes generally, and mixing zones in particular, has been cited as an example of broader professional ignorance on matters of traffic safety (Hauer, 2016). There are recent and ongoing efforts to better understand bicyclist safety, including NCHRP 17-84: Pedestrian and Bicycle Safety Performance Functions for the Highway Safety Manual, NCHRP 15-73: Design Options to Reduce Turning Motor Vehicle – Bicycle Conflicts at Controlled Intersections and NCHRP 15-74: Safety Evaluation of On-Street Bicycle Facility Design Features. While a lack of crash and exposure data continues to be a hindrance to bike safety research, it has generally been accepted that as the biker population increases, the crash rate decreases (Elvik et al., 2009), perhaps an indication of greater awareness on the part of motorists of the need to share the road with this travel mode.

To that end, both Sweden and the Netherlands have developed approaches to address this challenge. Starting in the early 1990s, Sweden's Vision Zero and the Netherlands' Sustainable Safety Vision have integrated motorists and vulnerable road users with the concept of shared road responsibility to create homogeneous, multimodal transportation networks (Welle et al., 2018; Wegman et al., 2006). The same concept has been recently adopted in Davidson County, TN (Vision Zero, 2020).

8. Final thoughts

Bike safety has been a much-discussed topic, particularly of late, as interest in bicycling as a sustainable transportation alternative continues to gain popularity. Consequently, policy analysts and planners have been grappling with cost-effective methods to reduce bicycle crashes, particularly those with serious outcomes. We believe that the results of this study have shed additional light on the subject, in particular: (1) demonstrating the use of random forest modeling and select sampling techniques as having the potential to provide greater accuracy in predicting the likelihood of a fatal and serious bike injury, and (2) utilizing the feature weighting of the predictive model to prioritize the types of risk mitigation strategies that offer the greatest impact.

Funding Statement

This research was funded by a Congestion Mitigation and Air Quality (CMAQ) grant administered by the U.S. Federal Highway Administration and awarded by the Tennessee Department of Transportation [Grant No 40100-11319].

Disclosure Statement

There are no conflicts of interest to disclose.

Data availability

The data used in this publication are from the Tennessee Department of Transportation. All crashes are recorded and maintained in Tennessee Roadway Information Management System (TRIMS).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in the paper.

Appendix I. Data attributes

1. 'CASENO' – Case Number
2. 'Region' – Region where crash was recorded
3. 'MPO_RPO' - Metropolitan Planning Organization or Rural Planning Organization
4. 'County' - County where crash was recorded
5. 'City' – City where crash was recorded
6. 'DATEOFCRASH' – Crash date
7. 'TIMEOFCRASH' – Crash time
8. 'TOTALKILLED' – Total people killed in the crash
9. 'TOTALINJURED' – Total people injured in the crash
10. 'TOTAL_INCAP_INJ' – Total people suffering incapacitating injury
11. 'NBR_RTE' - Route number where crash occurred
12. 'RTE_NME' -Route name where crashed occurred
13. 'LocType' – Crash location type
14. 'FunctionalClass' – Functional classification of roadway where crash was recorded
15. 'NBR_LANES' – Number of lanes
16. 'SPD_LMT' – Speed limit
17. 'Median' – If crash occurred on median (yes/no)
18. 'TurnLane' – If crash occurred on turn lane (yes/no)
19. 'AADT' - Average annual daily traffic on the road segment where crash was recorded
20. 'CollisionType' – Type of collision
21. 'Hit_Run'- If it was a hit and run crash (yes/no)
22. 'Distraction' – If there was any indication of distracted driving (yes/no)
23. 'ImpairedDriver' – If the driver was impaired due to intoxication (yes/no)
24. 'Veh0_5yr' – Vehicle age 0-5 years
25. 'Veh5_10yr'- Vehicle age 5-10 years
26. 'Veh10_20yr' - Vehicle age 10-20 years
27. 'VehOver20yr' - Vehicle age over 20 years
28. 'TruckBus' – If truck/bus was involved in crash (yes/no)
29. 'Severity' – Severity of the crash (no injury, injury, incapacitating injury, and fatality)
30. 'TOTALVEHICLES' – Total vehicles involved in crash
31. 'PeopleInvolved'- Total people involved in crash
32. 'NonMotorist' – Non-motorist involved (bike and pedestrians)
33. 'Construction' – If there was construction (yes/no)
34. 'Weather' – Weather conditions when crash occurred
35. 'Lighting' – Lighting conditions when crash occurred
36. 'TrafficSegmentID' – Traffic segment identification number of the road
37. 'RdSegID'- Road segment identification number
38. 'POINT_X' - Longitude
39. 'POINT_Y' - Latitude

Appendix II. Definition of model outputs

Based on labeling the targets as equal to a value of 1 for killed or incapacitating bike injury, and 0 for no killed or no incapacitating bike injury.

Confusion matrix

Provides an overview of classification model performance and the types of errors produced by the model. Also generates a summary of correct and incorrect predictions broken down by each category. Four types of outcomes are possible:

True Positives (TP) – Occur when we predict that an observation belongs to a certain class and the observation actually belongs to that class.

True Negatives (TN) – Occur when we predict that an observation does not belong to a certain class and the observation actually does not belong to that class.

False Positives (FP) –Occur when we predict that an observation belongs to a certain class but the observation actually does not belong to that class. This type of error is called **Type I error.**

False Negatives (FN) –Occur when we predict that an observation does not belong to a certain class but the observation actually belongs to that class. This is a very serious error and it is called **Type II error.**

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Accuracy: Percentage of correctly classified observations. Calculated by dividing the number of correct predictions by the total number of predictions. i.e., Accuracy = (TP + TN)/(TP + TN + FP + FN).

Precision: Percentage of relevant observations that actually belong to a certain class among all the samples which were predicted to belong to the same class. i.e., Precision = TP/(TP + FP).

Recall: Percentage of observations that were predicted to belong to a certain class among all the samples that truly belong to that class. i.e., Recall = TP/(TP + FN). Also called Sensitivity or True Positive Rate (TPR).

False Negative Rate: It's the percentage of observations that were positive but identified as negative. It can be calculated by (1-Sensitivity) or FN /(TP + FN).

Specificity: Also known as True Negative Rate (TNR) is a measure that correctly identifies the proportions of negatives i.e., Specificity = TN/(TN + FP).

False Positive Rate: It's the percentage of observations that were negative but identified as positive. It can be calculated by (1-Specificity) or FP/(FP + TN).

F measure: or F-1 score, combines Precision and Recall as a measure of effectiveness of classification in terms of ratio of weighted importance on either Recall or Precision as determined by β coefficient. i.e., F measure = $((1 + \beta)^2 \times \text{Recall} \times \text{Precision})/(\beta^2 \times \text{Recall} + \text{Precision})$. β is usually taken as 1.

G-mean: Geometric mean measures the balance between classification performance on both true negative rate or specificity (TN/TN + FP) and true positive rate or sensitivity or recall (TP/TP

+ FN). A lower G-mean indicates poor classification performance. Whereas G-mean close to 1 has a good classification for both negative and positive class.

$$G - Mean = \sqrt{(Sensitivity \;\; X \;\; Specificity)}$$

Weighted Accuracy: To be more sensitive to the performance of individual classes, weights (w) can be assigned to each class. More weight means more influence of that particular class on the weighted accuracy. In our case since both classes are equally important, a weight of 0.5 is assigned to each. Weighted Accuracy = w* True Negative Rate+(1-w) *True Positive Rate.

ROC: Receiver Operator Characteristic or Area Under the Curve (AUC) is a curve that plots the true positive rate and false positive rate at various threshold. When AUC is 1 classifier is able to distinguish between majority and minority class perfectly. An AUC with 0.5 or lower is said to be a poor classifier.

References

Abdel-Aty, M., & Keller, J. (2005). Exploring the Overall and Specific Crash Severity Levels at Signalized Intersections. *Accident Analysis & Prevention, 37*(3), 417–425.

Allen-Munley, C., Daniel, J., & Dhar, S. (2004). Logistic model for rating urban bicycle route safety. *Transportation Research Record, 1878*, 107–115. TRB, National Research Council, Washington DC.

Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 82*(4), 1059–1086.

Asgarzadeh, M., Verma, S., Mekary, R. A., Courtney, T. K., & Christiani, D. C. (2017). The Role of Intersection and Street Design on Severity of Bicycle-Motor Vehicle Crashes. *Injury Prevention, 23*, 179–185.

Asgarzadeh, M., Fischer, D., Verma, S. K., Courtney, T. K., & Christiani, D. C. (2018). The impact of weather, road surface, time-of-day, and light conditions on severity of bicycle-motor vehicle crash injuries. *American Journal of Industrial Medicine, 61*(7), 556–565.

Chimba, D., Emaasit, D., & Kutela, B. (2012). Likelihood Parameterization of Bicycle Crash Injury Severities. *Journal of Transportation Technologies, 2*, 213–219.

Dixon, K., Monsere, C., Xie, F., & Gladhill, K. (2012). Calibrating the Future Highway Safety Manual Predictive Methods for Oregon State Highways Publication No. OTREC-RR-12-02.

Dolatsara, H. A. (2017). Development of Safety Performance Functions for Non-Motorized Traffic Safety.

Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention, 40*(3), 1033–1054.

Elvik, R., Hoye, A., Vaa, T., & Sorensen, M. (2009). The Handbook of Road Safety Measures, 2nd ed. Oslo, Norway.

Fridman, L., Ling, R., Rothman, L., et al. (2020). Effect of reducing the posted speed limit to 30 km per hour on pedestrian motor vehicle collisions in Toronto, Canada - a quasi experimental, pre-post study. *BMC Public Health, 20*, 56.

Furth, P. G., Mekuria, M. C., & Nixon, H. (2016). Network Connectivity for Low-Stress Bicycling. *Transportation Research Record: Journal of the Transportation Research Board, 2587*, 41–49. TRB, National Research Council, Washington DC.

Guide for the Development of Bicycle Facilities, 4th Edition (2014). American Association of State Highway and Transportation Officials (AASHTO). Washington, DC.

Haleem, K., & Abdel-Aty, M. (2010). Examining Traffic Crash Injury Severity at Unsignalized Intersections. *Journal of Safety Research, 41*(4), 347–357.

Huang, H., Abdel-Aty, M., & Darwiche, A. (2010). County-level crash risk analysis in Florida: Bayesian spatial modeling. *Transportation Research Record: Journal of the Transportation Research Board, 2148*, 27–37.

Hauer, E. (2016). An exemplum and its road safety morals. *Accident Analysis and Prevention, 94*, 168–179.

Isaksson-Hellman, I., & Töreki, J. (2019). The effect of speed limit reductions in urban areas on cyclist injuries in collision with cars. *Traffic Safety Prevention, 20* (sup3), 39–44. 2018 International Cycling Safety Conference (ICSC).

Jacobsen, P. L., & Rutter, H. (2012). Cycling Safety. In R. Buehler & J. Pucher (Eds.), *City Cycling* (pp. 141–156). Cambridge, Massachusetts: MIT Press.

Jurewicz, C., Sobhani, A., Woolley, J., Dutschke, J., & Corben, B. (2016). Exploration of vehicle impact speed – injury severity relationships for application in safer road design. *Transportation Research Procedia, 14*, 4247–4256.

Kaplan, S., & Prato, C. G. (2015). A spatial analysis of land use and network effects on frequency and severity of cylist – motorist crashes in Copenhagen region. *Traffic Injury Prevention, 16*, 724–731.

Kim, J. K., Kim, S., Ulfarsson, G. F., & Porrello, L. A. (2007). Bicyclist injury severities in bicycle– Motor vehicle accidents. *Accident Analysis & Prevention, 39*(2), 238–251.

Klop, J. R., & Khattak, A. J. (1999). Factors Influencing Bicycle Crash Severity on Two-Lane, Undivided Roadways in North Carolina. *Transportation Research Record., 1674*(1), 78–85.

Lee, C., & Abdel-Aty, M. (2005). Comprehensive Analysis of Vehicle-Pedestrian Crashes at Intersections in Florida. *Accident Analysis & Prevention, 37*(4), 775–786.

Lee, J., Abdel-Aty, M., & Jiang, X. (2015). Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. *Accident Analysis & Prevention, 78*, 146–154.

Müller, A. C., & Guido, S. (2017). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

National Transportation Safety Board (NTSB). (2019). Bicyclist Safety on US Roadways: Crash Risks and Countermeasures. Safety Research Report NTSB/ SS-19/01. Washington, DC.

Noland, R., & Quddus, M. (2004). Analysis of pedestrian and bicycle casualties with regional panel data. *Transportation Research Record: Journal of the Transportation Research Board, 1897*, 28–33. TRB, National Research Council, Washington DC.

Oh, J., Jun, J., Kim, E., & Kim, M. (2008). Assessing Critical Factors Associated with Bicycle Collisions at Urban Signalized Intersections. *Transportation Research Board*, 1–17. TRB, National Research Council, Washington DC.

Pai, C. W. (2011). Overtaking, Rear-End, and Door Crashes Involving Bicycles: An Empirical Investigation. *Accident Analysis & Prevention, 43*(3), 1228–1235.

Petritsch, T., Landis, B., Huang, H., & Challa, S. (2006). Sidepath safety model: Bicycle Sidepath design factors affecting crash rates. *Transportation Research Record: Journal of the Transportation Research Board, 1982*, 194–201. TRB, Washington DC.

Quddus, M. (2015). Effects of geodemographic profiles of drivers on their injury severity from traffic crashes using multilevel mixed-effects ordered logit model. *Transportation Research Record: Journal of the Transportation Research Board, 2514*, 149–157.

Rahman, M.S. (2018). Applying machine learning techniques to analyze the pedestrian and bicycle crashes at the macroscopic level.

Reynolds, C. C., Harris, M. A., Teschke, K., Cripton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: A review of the literature. *Environmental Health, 8*, 47.

Schepers, P., & den Brinker, B. (2011). What do cyclists need to see to avoid single-bicycle crashes? *Ergonomics, 54*(4), 315–327.

Shubo, W., Yuan, Q., Yan, Z., & Xu, Q. (2021). Analyzing Accident Injury Severity via an Extreme Gradient Boosting (XGBoost) Model. *Journal of Advanced Transportation, 2021*, 11. Article ID 3771640.

Siddiqui, C., Abdel-Aty, M., & Choi, K. (2012). Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accident Analysis & Prevention, 45*, 382–391.

Smart Growth America. (2020). Dangerous by Design 2020. Smart Growth America.

Strauss, J., Miranda-Moreno, L. F., & Morency, P. (2015). Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. *Accident Analysis & Prevention, 83*, 132–142.

Traffic Safety Facts: Bicyclists and other Cyclists 2019 Data. (2021). National Highway Traffic Safety Administration (NHTSA), US Department of Transportation, Washington, DC: USDOT.

Turner, S., Wood, G., Hughes, T., & Singh, R. (2011). Safety performance functions for bicycle crashes in New Zealand and Australia. *Transportation Research Record: Journal of the Transportation Research Board, 2236*, 66–73.

Urban Bike Design Guide, 2nd Edition (2014). National Association of City Transportation Officials (NACTO).

US Census Bureau (2014). Modes Less Traveled—Bicycling and Walking to Work in the United States: 2008–2012, American Community Survey Reports.

US Census Bureau (2021). Commuting by Public Transportation in the United States: 2019, American Community Survey Reports.

Vandenbulcke, G., Thomas, I., & Panis, L. I. (2014). Predicting Cycling Accident Risk in Brussels: A Spatial Case-Control Approach. *Accident Analysis & Prevention, 62*, 341–357.

Vision Zero. (2020). Mayor Cooper Commits to Vision Zero to Eliminate Traffic Fatalities and Severe Injuries in Nashville, Mayor's Office, Media Release, January 18, 2020. Accessed on 12 November 2021.

Wang, Y., & Nihan, N. L. (2004). Estimating the risk of collisions between bicycles and motor vehicles at signalized intersections. *Accident Analysis & Prevention, 36* (3), 313–321.

Wang, C., Lu, L., & Lu, J. (2015). Statistical analysis of bicyclists' injury severity at unsignalized intersections. *Traffic Injury Prevention, 16*(5), 507–512.

Wegman, F. C. M., Aarts, L. T., & Bax, C. (2006). *Advancing Sustainable Safety: National Road Safety Outlook for the Netherlands for 2005-2020*. Leidschendam: SWOV Institute for Road Safety Research.

Welle, B., Sharpin, A. B., Adriazola-Steil, C., Bhatt, A., Alveano, S., Obelheiro, M., Imamoglu, C. T., Job, S., Shotten, M., & Bose, D. (2018). *Sustainable and Safe: A Vision and Guidance for Zero Road Deaths. Ross Center, World Resources Institute* (p. 70). Washington DC: World Resources Institute.

Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., & Bhatia, R. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. *Accident Analysis & Prevention, 41*(1), 137–145.

Yan, X., Ma, M., Huang, H., Abdel-Aty, M., & Wu, C. (2011). Motor vehicle-bicycle crashes in Beijing: Irregular Maneuvers, Crash Patters, and Injury Severity. *Accident Analysis & Prevention, 43*(5), 1751–1758.

Yu, C.-Y., & Zhu, X. (2015). Planning for safe schools: Impacts of school siting and surrounding environments on traffic safety. *Journal of Planning Education and Research*, 1–11.

Zangenehpour, S., Strauss, J., Miranda-Moreno, L. F., & Saunier, N. (2016). Are signalized intersections with cycle tracks safer? A case-control study based on automated surrogate safety analysis using video data. *Accident Analysis & Prevention., 86*, 161–172.

Ishita Dash received her B.Sc. in Marine Engineering from Birla Institute of Technology and Science, India and M.Sc. in Reliability, Availability, Maintainability and Safety from Norwegian University of Science and Technology, Trondheim, Norway. She is a Ph.D. student and a graduate research assistant in the Civil and Environmental Engineering Department at Vanderbilt University, Nashville, Tennessee. Her research interests include transportation data analytics, transportation safety, and intelligent transportation systems. She brings a decade of work experience in design, construction, operations, and maintenance of offshore installations globally with a focus on risk, reliability, availability, maintainability and safety.

Mark Abkowitz is a Professor of Civil and Environmental Engineering at Vanderbilt University, Nashville, Tennessee. Dr. Abkowitz specializes in risk assessment, management and communication; community and infrastructure resilience; smart cities technologies and applications, and transportation safety and security. He has served as a researcher and consultant to a wide variety of businesses and government agencies, and is the founder and former chairman of Visual Risk Technologies (now known as Factor, Inc.). Dr. Abkowitz is also the author of *Operational Risk Management - A Case Study Approach to Effective Planning and Response*, published by John Wiley & Sons.

Craig Philip is a Research Professor of Civil and Environmental Engineering and the Director of Vanderbilt Center for Transportation and Operational Resiliency (VECTOR) at Vanderbilt University, Nashville, Tennessee. Dr. Philip's research focus includes infrastructure sustainability and the application of risk management tools to transportation systems, carrier safety management, and transport policy and regulation. Before joining Vanderbilt, he spent 35 years in the rail, intermodal and maritime industries, including serving as President and CEO of Ingram Barge Company. Dr. Philip has served on the Executive Committee of the Transportation Research Board and is a member of the National Academy of Engineering.

Journal of Safety Research 83 (2022) 260–268

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Have you met Angus? Development and evaluation of a social marketing intervention to improve personal flotation device use in commercial fishing

Theodore D. Teske [a,*], Samantha L. Case [b], Devin L. Lucas [b], Christy L. Forrester [c], Jennifer M. Lincoln [b,1]

[a] National Institute for Occupational Safety and Health, Western States Division, 315 E. Montgomery Avenue, Spokane, WA 99207, USA
[b] National Institute for Occupational Safety and Health, Western States Division, 4230 University Drive, Anchorage, AK 99508, USA
[c] National Institute for Occupational Safety and Health, Office of the Director, Patriots Plaza 1, 395 E Street, S.W., Suite 9200, Washington, DC 20201, USA

ARTICLE INFO

Article history:
Received 22 November 2021
Received in revised form 18 April 2022
Accepted 31 August 2022
Available online 13 September 2022

Keywords:
Occupational safety
Falls overboard
Behavior change
PFD
Fishermen

ABSTRACT

Introduction: Drowning is the leading cause of death among commercial fishermen in the United States. Approximately 30% of all commercial fishing fatalities are attributed to falls overboard. One of the simplest and most affordable ways to prevent these fatalities is for crewmembers to wear personal flotation devices (PFDs) while on deck. An examination of over 200 fatal falls overboard in the U.S. fishing industry revealed that none of the victims were wearing PFDs when they died. PFDs are not required to be worn by commercial fishermen in the United States, so this study was designed to encourage behavior change using targeted health communication and social marketing. *Methods*: This study developed, implemented, and evaluated a multi-media social marketing campaign featuring a fictitious, culturally-relevant spokesman designed to look, talk, and act like the target audience. The messages were crafted to address common barriers to PFD adoption and misconceptions about fleet-specific risks for fatalities from falls overboard. The campaign was evaluated over two seasons of fishing to look at message retention and intent toward action following exposure to the campaign materials. *Results*: Survey respondents indicated overall positive opinions about the spokesman and the messages. Results also show a reported change in behavior related to using PFDs while working on deck. *Discussion*: Targeted multi-media messaging can influence behavior of workers in high-risk occupations in remote locations. Safety message development should focus on occupational culture to create valid and authentic communication products for workers in high-risk industries.

1. Introduction

Drowning is the leading cause of death among commercial fishermen in the United States (National Institute for Occupational Safety and Health, 2021). During 2000–2019, 871 fishermen died while commercial fishing in the United States (National Institute for Occupational Safety and Health, 2021). This resulted in the highest fatality rate for any occupation in 2019, at 145 fatalities per 100,000 full-time equivalent (FTE) workers, 41 times higher than the average worker (Bureau of Labor Statistics, 2020). Most of these fatalities were due to drownings after a vessel disaster or from a crewmember falling overboard (Lincoln and Lucas, 2010; Lucas and Case, 2018; Case et al., 2018). Approximately 30% of all commercial fishing fatalities are attributed to falls overboard (Lincoln and Lucas, 2010; Lucas and Case, 2018).

While the number of deaths from falls overboard in the fishing industry has declined slightly since 2000, the lack of widespread adoption of prevention strategies persists (Case et al., 2018). For instance, one of the simplest and most affordable ways to prevent these fatalities is for crewmembers to wear personal flotation devices (PFDs) while on deck. An examination of over 200 fatal falls overboard in the U.S. fishing industry revealed that none of the victims were wearing PFDs when they died (Case et al., 2018). This finding is consistent with earlier studies focused in Alaska that highlighted the lack of PFD use among fall overboard victims (Lincoln and Conway, 1999; Lucas and Lincoln, 2007; National

* Corresponding author at: NIOSH Western States Division, 315 E. Montgomery Avenue, Spokane, WA 99207, USA.

E-mail addresses: tteske@cdc.gov (T.D. Teske), scase@cdc.gov (S.L. Case), dlucas@cdc.gov (D.L. Lucas), cforrester@cdc.gov (C.L. Forrester), jlincoln@cdc.gov (J.M. Lincoln).

[1] New work address: National Institute for Occupational Safety and Health, Office of Agriculture Safety and Health, Room 513, 1150 Tusculum Ave, Cincinnati, OH 45226, USA.

https://doi.org/10.1016/j.jsr.2022.08.022
0022-4375/

Institute for Occupational Safety and Health (NIOSH), 1994; National Institute for Occupational Safety and Health (NIOSH), 1997).

There is a clear need to increase the use of PFDs in the fishing industry, and understanding the barriers that workers have to wearing PFDs is an important first step. A previous NIOSH study of predictors of PFD use among workers in the Alaskan fishing industry concluded that workers are likely to increase their PFD usage if their perceptions of risk and beliefs about PFDs are improved (Lucas et al., 2013). The study found that fishermen with heightened risk perceptions related to falling overboard had significantly higher levels of PFD usage than fishermen with lower risk perceptions. The study also showed a gap between fishermen's perceived efficacy of PFDs to prevent fatalities from falls overboard and their use when working on deck. Of those surveyed, 85% said PFDs were fairly or very effective at preventing fatalities from falls overboard; however, only 33% of those same respondents indicated that they frequently or always wore a PFD when working on the deck of a fishing vessel (Lucas et al., 2013). The study also found that significant barriers to PFD usage included beliefs that wearing a PFD is uncomfortable, interferes with work, and creates an entanglement hazard (Lucas et al., 2013). However, field-testing of PFDs with Alaskan fishermen showed there are commercially available PFDs that overcome some off these barriers and are acceptable to work in (Lucas et al., 2012). Based on these findings, researchers recommended the development and implementation of tailored interventions to improve risk perception and to overcome efficacy barriers to PFD use (Lucas et al., 2013).

Health behavior research has shown that increasing risk perceptions can lead to behavior change when accompanied by increased self-efficacy (Witte, 1992; Rosenstock et al., 1988). Attitudes and beliefs can be influenced, and previous research has demonstrated that when attitudes and beliefs are modified, behavior change may follow (Prochaska and Velicer, 1997). For example, researchers with the Florida Prevention Research Center at the University of South Florida used participatory research, peer-level recommendations, and social marketing to increase the use of safety glasses among citrus industry workers and overcome their belief that safety glasses reduced their productivity (Tovar-Aguilar et al., 2014). In this study, the researchers were able to increase the use of safety glasses among the primary audience by having well-respected workers model the behavior and demonstrate the glasses had no impact on productivity. In a similar fashion, based on NIOSH's previous studies on PFD use among fishermen, it should be possible to influence fishermen's perceptions and beliefs regarding PFDs, overcome their negative attitudes about working in PFDs, and thereby increase PFD use.

Social marketing is an intervention approach that has a strong foundational theory, which suggests that by using a "marketing mix" of product, price, place, and promotion to engage audiences, one can provide an effective channel for motivating behavior change (Kotler et al., 2002).

NIOSH has applied a social marketing approach to underground hard rock miners, another high-risk occupation. NIOSH incorporated adult learning theory, storytelling, and a focus on occupational culture to create valid and authentic communication products for these workers and others in high-risk industries (Cullen and Fein, 2005; Cullen et al., 2008). As Van Maanen and Barley (Van Maanen and Barley, 1984) explained, this shared risk is a bonding attribute in the occupational culture of high-risk industries:

> "Danger…invites work involvement and a sense of fraternity…Recognition that one's work entails danger heightens the contrast between one's own work and the safer work of others, and encourages comparison of self with those who share one's work situation. Attitudes, behaviors, and self-images for coping physically and psychologically with threat become part of an occupational role appreciated best, it is thought, only by one's fellow workers."

Similarly, commercial fishing is a high-risk industry with a strong occupational culture that converts high-risk to an integral part of their occupational culture. Therefore, this combination of adult learning theory, storytelling, and occupational culture could also be important to increasing PFD use among commercial fishing crews. The purpose of this study was to develop a social marketing intervention to increase PFD use among commercial fishermen, and to evaluate message recollection and appeal, motivation towards action, and perceptions of fall overboard risks.

2. Methods

Researchers used the 4 Ps of the social marketing mix and results of past PFD studies to set the parameters of the project as shown below. The social marketing intervention ran from May 2014 to November 2015.

2.1. Product

Social marketing campaigns have two basic tenets to help focus efforts on measurable outcomes: identifying a specific target audience and a targeted behavior to be changed. In the case of this intervention, the product is the behavior fishermen are being encouraged to adopt, wearing a PFD while working on deck. The social marketing intervention targeted two fishing populations in Alaska: Bristol Bay salmon drift gillnet fishermen (hereafter, gillnetters) and Bering Sea and Aleutian Island (BSAI) crab fishermen (hereafter, crabbers). These groups of fishermen were selected because of their inclusion in the original NIOSH PFD study and existence of baseline data on PFD use and fleet-specific attitudes towards PFDs (Lucas et al., 2013; Lucas et al., 2012).

2.2. Price

The price was considered on both a social and monetary level. There was potentially a social cost for adopting a behavior that could run counter to the dominant culture of commercial fishing. Fishermen could be perceived as being too risk adverse to be an effective fisherman, resulting in a cost of loss of social status. The monetary cost arose from the need for fishermen to purchase the PFD from a gear vendor. In some cases, this could be over $200 dollars for the more effective models as indicated in the NIOSH PFD study.

2.3. Place

For the third part of the marketing mix, place, the study focused on two aspects of this as well: the location where the messages would be shared and the channels that would be used to spread the messages. The locations for the dissemination of the messages were the ports of Naknek, AK and Dutch Harbor, AK with additional materials sent to gear vendors around the Northwest coast of the United States including Seattle and Bellingham, Washington and Newport, Oregon. The Northwestern target ports are areas where fishermen purchase gear ahead of their respective fishing season and have it shipped up to Alaska for use when they arrive.

Each of the Alaskan ports see a large influx of fishermen directly before the start of either the salmon season (in Naknek) and crab season (in Dutch Harbor). This concentration of workers prior to the season makes it an ideal location to present messages related to fall overboard safety as the fishermen gear up for the season, purchasing necessary equipment, and readying their vessels. Qualitative data gathered during the previous NIOSH PFD study indicated that local fishing supply shops were the primary place fishermen purchased their personal protective equipment for the season such as rain gear and boots. These locations also carried PFDs and other safety gear for the fleet. Having physical messages in these shops was a critical channel for sharing our PFD safety messages. Other locations targeted in the ports were grocery stores, restaurants and bars, community centers, and health centers.

Two print advertising channels were developed to support these port-based channels. Half-page, full-color ads ran for two years in a widely read commercial fishing trade magazine and direct mail postcards with fishery-specific messages were sent to all permit holders in the salmon drift gillnet and BSAI crab fleets prior to the start of each season.

In addition to physical messages, researchers used web-based channels to promote their messages including the NIOSH website, a campaign-specific website run by a partner, digital display ads on industry websites and blogs, and dedicated social media accounts for the campaign on Facebook and Twitter.

2.4. Promotion

A strategic communication firm was contracted to develop the creative concept for the promotion of the social marketing campaign. The firm reviewed NIOSH research on falls overboard and came back to the agency with three final concepts for consideration: Loss, Steer, and Salty.

The Loss concept focused on highlighting the risks of falls overboard as shown in hypothetical scenarios and pivoting to the social and emotional toll felt by survivors. One of the key messages said, "His boots didn't save him. A PFD would have." The goal as stated by the communication firm was to have fishermen start to see PFDs as standard deck gear like their boots and rain gear, not just as emergency or safety gear. This concept was ultimately rejected by NIOSH as being an overly emotional appeal that was similar to other safety messages used with the audience in the past. It could be dismissed by fishermen because they could overestimate their survival skills or ability to avoid the situations outlined in the scenarios. In essence they could tell themselves, "I'm smarter than that and wouldn't get myself into that situation.".

The second concept, Steer, focused on the vessel skippers as opinion leaders among the target audiences. There are no regulations requiring fishermen to wear a PFD while working on their vessel, however the skipper can choose to institute a PFD use policy on his vessel. The creative concept for Steer featured images of a skipper with text stating, "There are only two acceptable answers on board. "Yes" and "Captain." Secondary copy discussed the skipper's duty to look after the crew's safety by saying, "Put low-profile, purpose-built PFDs on your crew's gear list. And require they wear them while working on deck." While this concept did a good job of reaching out to strong opinion leaders in these audiences, it was not selected because there was no supporting messaging for the larger part of the audience, the crewmembers, who would be ultimately responsible for performing the behavior of wearing PFDs. This type of solution or messaging would also reduce the self-efficacy of the deck crew and could potentially undermine the long-term adoption of PFDs by individuals if they were seen as something being done to them and not something they were choosing for themselves.

The final concept, Salty, featured a fictitious, culturally-relevant spokesman that would look, talk, and act like gillnetters and crabbers. The concept tagline, "Live to be Salty," focused on the idea that by wearing a PFD, crewmembers would reduce their risk of a fatal fall overboard and live to be an experienced, older expert in the industry. The main goal was to make the spokesman memorable, quotable, and different from other safety messages in this industry. This type of message was expected to jolt message recipients out of their standard media consumption habits. When compared to the two other concepts, this one was viewed as the most innovative and most likely to resonate with the audience based on what was known about their attitudes towards PFDs and their occupational culture.

Initial concepts for the spokesman, Angus (Fig. 1), aligned with many cultural norms associated with commercial fishermen, including a desire for independence, bluntness, practicality, respect for successful fishermen (known as highliners), and a master/apprentice style training environment.

Based on the feedback of local fishing experts, the initial design was refined to more fully align with the specific traits of Alaskan gillnetters and crabbers. Stock photography images used in test versions of the concept did not depict the correct types of vessels or settings. These images also showed the subject holding a cigar, which was not acceptable for a public health message. The spokesman character's name was Angus McGilly, a humorous but memorable name that gave license for the spokesman to speak bluntly and sarcastically about PFD use. This name was changed to Angus Iversen to better reflect the Scandinavian heritage of many Northwest and Alaskan fishermen. Finally, the subject was not wearing a PFD and therefore not demonstrating the desired behavior change.

To correct these initial problems, a new version of Angus was photographed on an Alaskan crab vessel and salmon gillnet vessel wearing a PFD (Fig. 2). Photographs were also taken in the wheelhouse of the crab vessel to target messages to vessel captains.

2.5. Refinement with stakeholder input

Posters with revised images and proposed messages were tested with industry stakeholders including commercial fishermen, US Coast Guard marine safety experts, marine safety trainers, and marine supply vendors. A small focus group featuring commercial fishermen and marine safety trainers helped narrow down draft quotes from Angus. Stakeholders were asked to review the concepts and provide feedback, focusing on the messaging to make sure the messages resonated with them and were appropriate for cultural norms. Some reviewers expressed dissatisfaction in some of the phrasing, such as using improper language (e.g., "ain't") (Fig. 2), or in making crude jokes, however they approved of the concept overall. Other reviewers expressed appreciation for the creativity and plain-spoken tone of the messages. Using this information, the messages were revised to better reflect the stakeholder preferences (Fig. 3). Concurrently the draft messages were shared with potential gear vendor partners to get buy in from them as dissemination channels for the physical campaign materials. Researchers also ran workshops with sales staff at the shops to educate them on the results of the NIOSH PFD study and give them information about which PFDs may work best for a fisherman based on their type of fishery and other gear preferences.

2.6. Development of social marketing intervention materials

Twelve posters were developed, combining Angus images and quotes addressing specific hazards. For example, an image of Angus on a salmon gillnet vessel was combined with the quote, *"You may learn to think like a fish, but you'll never breathe like one."* This quote was supported by a tailored hazard message and language addressing common barriers for salmon fishermen identified in the original PFD study, *"Salmon fishermen have the highest number of man overboard fatalities in Alaska. It doesn't have to be this way. Today's low-profile PFDs are comfortable, don't tangle in gear, and extend survival time in the water. Choose your PFD today at livetobesalty.org.*

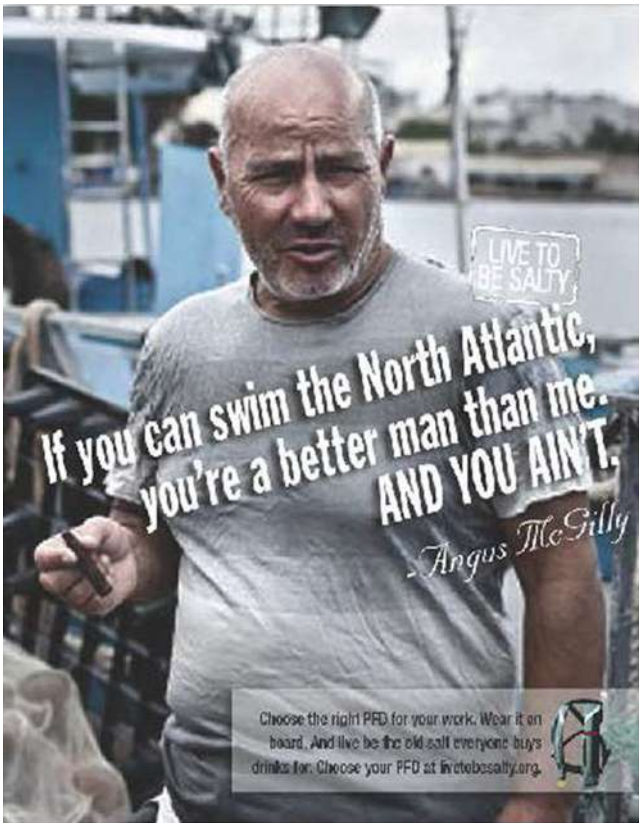

Fig. 1. Original Live to be Salty concept message that needed to be adjusted to meet the needs of the campaign.

Wear it. And Live." Another example focused on captains starts with the quote, "*My boat. My rules. You're wearing one, period.*" This addresses the captain's role as policymaker on the vessel and responsibility for the care of the crew. The message was again supported by additional copy addressing barriers to PFD use, "*Today's options make it easy to find a comfortable, work-friendly PFD that extends survival time in the water. Guide your men to livetobesalty.org. Then make PFDs mandatory on deck.*" Six posters focused on PFD comfort and cold-water hazards, three referenced specific hazards to gillnet fishermen, and three were targeted at vessel captains. This series of messages formed the foundation of all the social marketing intervention messaging.

In addition to the quote and hazard information, each poster contained a call to action with the link to a website (livetobesalty.org) containing information on the results from the NIOSH PFD study. Live to be Salty also featured a Facebook page and Twitter account to disseminate messages and engage participants with timely information and responses.

Posters and other point-of-sale collateral, including stickers, standup cardboard displays and beverage coasters were created. Stickers were placed on packaging of rain gear and deck boots to encourage crews to think of PFDs as standard deck equipment, rather than solely emergency devices (Fig. 4).

2.7. Social marketing intervention rollout

NIOSH researchers and external partners distributed intervention materials in Naknek, AK in the weeks prior to the Bristol Bay salmon season opening in June 2014. This process was repeated in Dutch Harbor, AK in preparation for two Bering Sea Aleutian Island crab seasons in October 2014 and January 2015. Materials were distributed to other ports and coastal communities in Alaska

Fig. 2. Revised message concept with correct imagery.

and the Pacific Northwest to engage fishermen during their off-seasons, which ran from July 2014 through May 2015 for gillnetters and March 2015 to September 2015 for crabbers.

2.8. Evaluation

A cross-sectional survey was administered during 2014 and 2015 at the beginning of the summer gillnet season and the fall crabbing season. The survey instrument included 20 closed-ended questions repeated from the 2008/2009 survey (Lucas et al., 2013; Lucas et al., 2012) to measure perceptions of the risk of falling overboard; attitudes about PFD efficacy and comfort, 6 multi-part questions to measure the social marketing intervention, and 4 demographic questions. The survey in 2014 allowed researchers to track any change in PFD use among the populations of fishermen between the survey in 2008/2009 and the start of the Live to be Salty campaign. The 2015 survey allowed researchers to track changes in PFD use and intended behavior change based on exposure to and recollection of the Live to be Salty messages. Skip patterns were used in the survey. For example, if the respondent did not recognize Angus Iversen or the Live to be Salty campaign messages, they did not answer subsequent questions regarding actions taken. The survey was approved by the NIOSH Human Subjects Review Board and the Office of Management and Budget.

The survey methodology was the same as for the previous PFD survey conducted in Alaska in 2008/2009. Researchers arrived in Naknek and Dutch Harbor several days prior to the start of salmon

Fig. 3. - Bering Sea message with revised copy removing improper English.

and crab seasons, respectively. Potential survey respondents were identified in ports and boat yards and were invited to participate. Paper-based surveys were completed on-the-spot with fishermen that volunteered and met the inclusion criteria. Surveys were administered until 100 responses had been collected in each fishery. Surveys were conducted in Naknek for gillnetters in June 2014 and June 2015, and in Dutch Harbor for crabbers in October 2014 and October 2015.

2.9. Statistical analysis

Descriptive statistics were calculated to explore patterns among variables and differences between crabbers and gillnetters working during the 2014 and 2015 seasons. Comparisons were made to examine overall differences between crabbers and gillnetters, as well as within-fishery differences before and after the intervention. Missing data were excluded from percent distributions. Statistics describing the recognition of Angus and the intervention were calculated from the 2015 survey only, since the campaign had not yet started in the baseline year of 2014. Significance tests were not performed since the sample sizes were small, especially when stratifying on certain groups (such as fishermen who recalled seeing an Angus message).

3. Results

3.1. Demographics and PFD use

A total of 401 surveys were completed in four trips ahead of the start of salmon and crab fishing seasons in 2014 and 2015 (Table 1). Characteristics such as age, sex, crew position, and fishing experience were similar from year to year for each group.

Among crabbers, self-reported PFD use (measured on a scale of never, sometimes, frequently, or always) remained high in both years, at 83.7% sometimes, frequently, or always wearing a PFD in 2014 and 85.9% in 2015. The proportion of crabbers who reported always wearing a PFD remained high in 2015 at 45.5%, and in all years was much higher than gillnetters.

Among gillnetters, any PFD use (responses of sometimes, frequently, or always wearing a PFD) increased from 36.7% in 2014 to 44.8% in 2015. However, very few gillnetters reported always wearing their PFD in both 2014 and 2015 (4.4% and 5.2%, respectively).

3.2. Message recall

Table 2 shows unaided and aided message recall among respondents. Participants were first asked about general PFD safety messages without any additional prompts or aids (e.g., photos of the ads). Over half of crabbers (56, 56.6%) and gillnetters (52, 51.5%) recalled hearing or seeing PFD safety messages within the previous month. Among crabbers, recognition of the intervention increased with additional prompting, with 56.7% recognizing the "Live to be Salty" slogan, 62.0% recognizing Angus from a photograph, and 72.7% recognizing a complete ad involving Angus and a PFD message. Recall with intervention-specific prompts among gillnetters

Fig. 4. "Apparel Stickers" used in marine supply stores to remind fishermen to purchase a PFD while buying their "standard" deck workwear.

Table 1
Demographic characteristics and PFD use by fleet and survey year.

	Bering Sea Crabbers						Bristol Bay Drift Gillnetters					
	2014 (N = 100)			2015 (N = 100)			2014 (N = 100)			2015 (N = 101)		
Continuous Variables	n	Mean	SD	n	Mean	SD	N	Mean	SD	n	Mean	SD
Age (yrs)	100	36.7	10.7	100	35.9	11.1	98	36.9	14.0	101	35.4	14.4
Experience (yrs)	100	16.0	10.6	100	15.6	11.0	97	15.8	13.3	100	14.6	12.7
Season (months)	99	7.3	2.1	99	7.1	1.9	97	3.6	2.4	97	3.9	2.8
Vessel Length (ft)	100	123.8	17.8	100	123.0	18.9	95	31.9	1.5	100	32.3	3.9
Crew Size (# workers)	100	6.8	1.3	100	6.6	1.4	96	3.4	0.6	100	3.5	0.7
Categorical Variables	Freq	%		Freq	%		Freq	%		Freq	%	
Sex (male)	100	100.0		99	100.0		94	95.0		92	91.1	
Position												
Captain	16	16.0		15	15.2		47	48.5		41	40.6	
Deckhand	74	74.0		72	72.7		48	49.5		54	53.5	
Other	10	10.0		12	12.1		2	2.1		6	5.9	
Missing	0	–		1	–		3	–		0	–	
PFD Usage												
Never	16	16.3		14	14.1		57	63.3		53	55.2	
Sometimes	23	23.5		28	28.3		23	25.6		26	27.1	
Frequently	8	8.2		12	12.1		6	6.7		12	12.5	
Always	51	52.0		45	45.5		4	4.4		5	5.2	
Missing	2	–		1	–		10	–		5	–	

Table 2
Intervention recall by fleet, 2015.

	Bering Sea Crabbers (N = 100)		Bristol Bay Drift Gillnetters (N = 101)	
	n	%	N	%
Recalled any PFD safety ads				
Yes	56	56.6	52	51.5
No	43	43.4	49	48.5
Missing	1	–	–	–
Heard "Live to be Salty" slogan				
Yes	55	56.7	26	26.8
No	42	43.3	71	73.2
Missing	3	–	4	–
Recognized Angus				
Yes	62	62.0	48	48.0
No	38	38.0	52	52.0
Missing	–	–	1	–
Recognized campaign ads				
Yes	72	72.7	50	50.0
No	27	27.3	50	50.0
Missing	1	–	1	–

also increased, but overall recognition of the ads was lower than among crabbers, at 50.0%.

Participants were also asked to select all channels where they saw the ads (Table 3). Magazines and posters were the most commonly identified channels from which respondents recalled seeing the ads. Of the 72 crabbers who recognized the ads, 46 (63.9%) saw the magazine ads and 33 (45.8%) recalled the intervention posters. Stickers that had been placed at local gear shops were also frequently identified (21, 29.2%). These results were similar among the 50 gillnetters, who also selected magazine ads (26, 52.0%) and posters (22, 44.0%), although stickers were less common (6, 12.0%).

The intervention involved internet components, including a website and social media pages, as well as a mailing component. These channels experienced low responses in both groups. Only five crabbers (6.9%) and two gillnetters (4.0%) identified "internet" as the source from which they saw the ad. Postal mail was the least commonly identified channel in both groups, with only two crabbers and no gillnetters selecting this response.

3.3. Appeal

Opinions of Angus were largely positive. Among the 62 crabbers who recognized Angus, the majority agreed or strongly agreed with statements that he seemed "like a seasoned fisherman" (42, 71.2%), "smart" (35, 60.3%), and "funny" (33, 58.9%). These responses were similar among 48 gillnetters, who also indicated he was "like a seasoned fisherman" (35, 77.8%), "smart" (32, 72.7%), and "funny" (24, 54.5%). Most respondents did not perceive Angus as a peer, with only 42.1% of crabbers and 46.5% of gillnetters agreeing or strongly agreeing that he was "like me.".

Message appeal was also favorable. Overall, respondents indicated they liked the PFD message. The majority of the 72 crabbers who had seen the ads agreed or strongly agreed that the message "was meant for fishermen like me" (65, 94.2%), "grabbed my attention" (60, 87.0%), "was convincing" (56, 83.6%), and "said something important" (56, 81.2%). Likewise, gillnetters responded that the message "was meant for fishermen like me" (47, 95.9%), "grabbed my attention" (45, 91.8%), "said something important" (42, 85.7%), and "was convincing" (40, 81.6%).

Table 3
Intervention channels identified and actions taken based on the intervention by fleet, 2015.*

	Bering Sea Crabbers (N = 72)		Bristol Bay Drift Gillnetters (N = 50)	
	n	%	n	%
Intervention Channels				
Newspaper	6	8.3	2	4.0
Magazine	46	63.9	26	52.0
Billboard	8	11.1	5	10.0
Poster	33	45.8	22	44.0
Postcard	2	2.8	1	2.0
Internet	5	6.9	2	4.0
Email	3	4.2	1	2.0
Postal Mail	2	2.8	0	0.0
Sticker	21	29.2	6	12.0
Actions Taken				
Looked for more information about PFD models	15	20.8	14	28.0
Visited the Live to be Salty website	4	5.6	1	2.0
Shared PFD message with others	19	26.4	9	18.0
Tried on a PFD	20	27.8	15	30.0
Purchased a new PFD	19	26.4	10	20.0
Wore PFD more often	18	25.0	11	22.0
Planned to take one or more actions	32	44.4	26	52.0

* Based on responses from those who recalled the ads. Responses not mutually exclusive.

3.4. Risk perception

Fall overboard risk perceptions were examined to determine any relationship between recalling the ads and those perceptions. Respondents indicated, on a scale from 0 to 100%, their perceived likelihood of ever falling overboard during their career. In general, respondents felt they were at low risk with little variation by ad recall status. Among crabbers, those who had seen the ads reported a median 12.5% (IQR = 45.0) chance of falling overboard, similar to 15.0% (IQR = 50.0) among crabbers who did not recall seeing the ads. Gillnetters who recalled the ads reported slightly higher likelihood at a median 35.0% (IQR = 40.0), compared to those who had not seen the ads (median = 16.0; IQR = 45.0).

3.5. Changes in behaviors

To measure the intervention's effect on behavior, respondents were asked if the ads prompted them to take action in a number of ways (Table 3). Trying on a PFD was the leading action taken among both groups. Crabbers then most frequently purchased a new PFD (19, 26.4%) while gillnetters sought more information on PFD models (14, 28.0%). Visiting the Live to be Salty website rarely occurred in either group. Overall, 32 crabbers (44.4%) and 26 gillnetters (52.0%) indicated they planned to take some type of action based on seeing the ads.

4. Discussion

Commercial fishermen work on wet, rolling vessels and are at risk of falling overboard while working on an open deck. Increased wear of properly sized and fitted PFDs would help prevent further deaths in the industry. Federal regulations require all fishing vessels to carry-one US Coast Guard-approved PFD per person on board; however, there are no requirements for them to be worn (46 CFR §28.110). In Alaska's cold-water fisheries, the requirement is to carry immersion suits (buoyant, full-body suit worn in the event of a vessel evacuation) (46 CFR §28.110), and many vessels do not also carry PFDs that could be worn while working. This, in conjunction with fishermen's general resistance to additional regulations, means it is critical to continue efforts to promote voluntary adoption of PFDs (Weil et al., 2016).

The findings from the 2014/2015 surveys showed that PFD use varied greatly between gillnetters and crabbers. In both years of the study, nearly half of crabbers reported they always wore PFDs, in contrast to over half of gillnetters reporting they never wore PFDs. The increase in crabbers reporting they always wear a PFD since the original survey in 2008/2009 is encouraging. Crabbers reported a substantial increase in always wearing a PFD, from 22.3% during the 2008/2009 survey (Lucas et al., 2013) to 52.0% in 2014. Crabbing vessels are larger, averaging 90-120′ in length, and are often company-owned. Because the high wear rate in the fleet occurred prior to the 2014/2015 surveys, the increase was not primarily due to the intervention, but instead was likely due to changes in company and vessel policies. In contrast, Bristol Bay drift gillnet boats are smaller, about 32′ in length, and tend to be independently owned and operated and therefore not subject to widespread company policies. The findings of continued low PFD use among this fleet is consistent with the 2008/2009 survey finding of 4.7% saying they always wear a PFD when working on deck (Lucas et al., 2013). The consistent low PFD wear rate in this fleet is of continued concern, particularly as they experience some of the highest numbers of loss of life due to falls overboard in the country (Case et al., 2018). It is clear that additional efforts targeted to workers in this fishery are a priority, including increasing awareness of commercially available PFDs that alleviate concerns of discomfort and potential for entanglement. Additionally, manufacturers should incorporate feedback from workers in the design and development of new, innovative devices. For example, based on the 2008/2009 survey, Kent Safety Products conducted market research to obtain feedback on a prototype PFD. In turn, adjustments were made to the PFD resulting in a lightweight, inherently buoyant vest thin enough to wear under rain gear (National Institute for Occupational Safety and Health (NIOSH), 2014). This process should be adopted by other manufacturers to make workable, wearable products for fishermen.

Our findings also showed that Angus was a memorable spokesman for the safety messages. Overall recognition of Angus, the slogan, and the ads was higher among crabbers than gillnetters. Because some crabbers also participate in salmon tendering during the summer months, it is plausible that the crabbers were more exposed to the messaging by seeing posters and other materials in both ports. Study results also suggest that the channel selection put Angus in places where he would be seen by the audiences. For these remote workers, print channels proved to be the most successful, especially the repetition of large ads in a popular trade magazine. While this should be expanded in future interventions where possible, this was also the most expensive channel to use and could not be easily expanded without an increase in resources. The partnerships with local gear vendors proved to be an effective channel as well, as the posters and apparel stickers used in these locations were seen and remembered by the fishermen. In contrast, the online resources were not commonly used among these remote workers. Most fishermen in Naknek and Dutch Harbor have little or no access to the internet while in port, so directing them to a website or social media was found to be of little value.

The findings show an interesting result regarding Angus' cultural appropriateness. While the majority of respondents from both groups agreed with the statement that Angus seems "like a seasoned fisherman," fewer felt that he was "like me." This may have to do with the age demographics of our respondents. The average respondent was in their mid-30 s, much younger than Angus who appears to be in his late 50 s or early 60 s. Despite the age difference, respondents showed collegial feelings for Angus through comments while the survey was being administered: "Who is that old bastard? I've seen him in Pacific Fishing [Magazine]." "I've seen the ads with the crusty old guy." "Angus Iversen? Sounds like a guy who goes to the bar, gets drunk, then loses a fight with himself in the parking lot." While these sentiments may sound harsh or dismissive, it is reassuring to see them remembering Angus, treating him like a peer, and not dismissing him outright.

The overarching goal of the intervention was to facilitate behavior change and improve PFD use in these fleets, and the survey results showed that a number of respondents took or planned to take some form of action because of the Live to be Salty ads. Trying on a PFD was the most common action taken, and nearly a quarter of respondents reported wearing their PFD more often because of the intervention. Although further work is needed to continue to improve consistent PFD use in these fisheries, we consider the intervention to be successful at motivating workers to take action, from gathering and sharing safety information to incorporating a PFD into their standard work gear. However, fishermen reported low self-perceived risk of falling overboard. This potentially helps explain the overall low rate of PFD wear among gillnetters. As discussed by Lucas et al. (Lucas et al., 2013), efforts to increase concern over man overboard risks may lead to increased PFD use by changing workers' attitudes about the value and utility of PFDs.

The overall reaction by the target fisheries to this campaign are encouraging and researchers should consider the further use of social marketing to encourage behavior change in commercial fishing through the use of fishery-specific campaigns. Social marketing has been shown to be a successful method occupational safety and health intervention in other fisheries and industries around the Unites States. Researchers from the Northeast Center for Occupational Health and Safety (NEC) have conducted successful social marketing interventions with lobstermen in New England related to PFD adoption (Sorensen et al., 2021) and farmers throughout the Northeast to encourage installation of rollover protection systems (ROPS) on farm tractors (Sorensen et al., 2011).

5. Limitations

This study is subject to a number of limitations. The first is the sampling methodology. The researchers used convenience sampling to identify and approach fishermen. The sample size for each year for each fishery was 100 respondents, however since the surveys were anonymous there was no way to determine if fishermen had taken the survey before. Additionally, the populations of these audiences are predominantly male, with the crab fishery being likely 100% male, so capturing female respondents' attitudes about the campaign spokesman was virtually impossible. The salmon gillnet fleet does have more gender diversity, but it is still predominantly male. Future research could be conducted to evaluate whether Angus Iversen is a suitable spokesperson for fishing workers not identifying as male.

Another limitation is the inability to generalize the results to the rest of the commercial fishing industry. Since the messages were targeted specifically at the two Alaskan fisheries included in the study, it is unknown if these messages would be effective in other fisheries or areas of the country. Anecdotal evidence during the dissemination of campaign materials showed that other fisheries did see some value in the messages with requests for materials coming from the east coast and Gulf coast.

6. Conclusion

The survey showed that Angus was a culturally appropriate spokesman with a strong message that could be recalled and acted upon. Further message development should consider focusing on occupational culture to create valid and authentic communication products for fishermen and others in high-risk industries.

While the printed channels (e.g., posters, ads, stickers) were successful in delivering the messages and reaching our audiences, the development and implementation of this intervention was extremely resource intensive, and the images and messages may not be easily transferrable to other fisheries or regions. Future research may examine the utility of Angus as a messenger for other hazards in commercial fishing or could include the development of a new spokesman using the similar PFD safety messages for other fisheries with high fall overboard fatality risks around the country. Additionally, because of the small number of female respondents, we could not examine differences in PFD perceptions or opinions of Angus based on sex. Study of these topics among females in the fishing industry may reveal unique differences and could provide valuable insight needed to improve future interventions.

Conflict of interest statement

I, Theodore D. Teske, have no conflicts of interest related to the research described in this manuscript.

This work was conducted as part of my job at NIOSH using internal funding.

Acknowledgements

The authors express their gratitude to the commercial fishermen who participated in the study. The authors also wish to thank the many partners who assisted in the development and implementation of the intervention, including the Alaska Marine Safety Education Association, US Coast Guard, Commercial Fishing Safety Advisory Committee, and local gear shops in Naknek, Dutch Harbor, Bellingham, and Seattle.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health. Mention of company names or products does not imply endorsement by the National Institute for Occupational Safety and Health. This study was funded internally by the Centers for Disease Control and Prevention's National Institute for Occupation Safety and Health (CDC/

NIOSH). There were no external grants or contracts provided to me to conduct the research.

References

Bureau of Labor Statistics. Civilian occupations with high fatal work injury rates (2019). Washington (DC): US Department of Labor, Bureau of Labor Statistics. Available from: https://www.bls.gov/charts/census-of-fatal-occupational-injuries/civilian-occupations-with-high-fatal-work-injury-rates.htm.

Case, S. L., Lincoln, J. M., & Lucas, D. L. (2018). Fatal falls overboard in commercial fishing – United States, 2000–2016. *MMWR, 67*(16), 465–469.

Cullen, E., Fein, A. (2005). Tell me a story: why stories are essential to effective safety training. Spokane (WA): US Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health. (DHHS [NIOSH] Publication No. 2005-152-163 [RI 9664]).

Cullen, E., Matthews, L., & Teske, T. (2008). Use of occupational ethnography and social marketing strategies to develop a safety awareness campaign for coal miners. *Social Marketing Q, 14*(4), 27–46.

Kotler, P., Roberto, N., & Lee, N. (2002). *Social marketing: Improving the quality of life*. Thousand Oaks, (CA): Sage.

Lincoln, J. M., & Conway, G. A. (1999). Preventing commercial fishing deaths in Alaska. *Occupational and Environmental Medicine, 56*, 691–695.

Lincoln, J. M., & Lucas, D. L. (2010). Occupational fatalities in the United States commercial fishing industry, 2000–2009. *Journal of Agromedicine, 15*(4), 343–350.

Lucas, D. L., & Case, S. L. (2018). Work-related mortality in the US fishing industry during 2000–2014: New findings based on improved workforce exposure estimates. *American Journal of Industrial Medicine, 61*, 21–31.

Lucas, D. L., & Lincoln, J. M. (2007). Fatal falls overboard on commercial fishing vessels in Alaska. *American Journal of Industrial Medicine, 50*, 962–968.

Lucas, D. L., Lincoln, J. M., Carozza, S. E., Bovbjerg, V. E., Kincl, L. D., Teske, T. D., Somervell, P. D., & Anderson, P. J. (2013). Predictors of personal flotation device (PFD) use among workers in the Alaska commercial fishing industry. *Safety Science, 53*, 177–185.

Lucas, D., Lincoln, J., Somervell, P., & Teske, T. (2012). Worker satisfaction with personal flotation devices (PFDs) in the fishing industry: Evaluations in actual use. *Applied Ergonomics, 43*, 747–752.

National Institute for Occupational Safety and Health (NIOSH) (1994). Request for assistance in preventing drownings of commercial fishermen. Cincinnati (OH): US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health (DHHS[NIOSH] Publication No. 94-107).

National Institute for Occupational Safety and Health (NIOSH) (1997). Commercial fishing fatalities in Alaska: Risk factors and prevention strategies. Anchorage (AK): US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health (DHHS[NIOSH] Publication No. 97-163 [CIB 58]).

National Institute for Occupational Safety and Health (NIOSH) (2014). A story of impact: PFD Manufacturer adopts NIOSH research into product development process. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health (DHHS [NIOSH] Publication No. 2015-119).

National Institute for Occupational Safety and Health (NIOSH) (2021). Commercial Fishing Incident Database [Date file]. Anchorage (AK): US Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health.

Prochaska, J., & Velicer, W. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion, 12*(1), 38–48.

Rosenstock, I., Strecher, V., & Becker, M. (1988). Social learning theory and the health belief model. *Health Education Quarterly, 15*(2), 175–183.

Sorensen, J., Jenkins, P., Emmelin, M., Stenlund, H., Weinehall, L., Earle-Richardson, G., & May, J. (2011). The social marketing of safety behaviors: A quasi-randomized controlled trial of tractor retrofitting incentives. *American Journal of Public Health, 101*(4), 678–684.

Sorensen, J., Weil, R., Echard, J., Tinc, P., Hirabayashi, L., Roome, A., Lally, E., & Jenkins, P. (2021). Evaluation of the Lifejackets for Lobstermen program: A social marketing initiative to increase lifejacket use in the Northeast lobster fishing industry. *Safety Science, 142*.

Tovar-Aguilar, J. A., Monaghan, P. F., Bryant, C. A., Esposito, A., Wade, M., Ruiz, O., & McDermott, R. J. (2014.). Improving eye safety in citrus harvest crews through the acceptance of personal protective equipment, community-based participatory research, social marketing, and community health workers. *Journal of Agromedicine, 19*(2), 107–116.

Van Maanen, J., & Barley, S. R. (1984). Occupational communities: Culture and control in organizations. *Research in Organizational Behavior, 6*, 287–365.

Weil, R., Pinto, K., Lincoln, J., Hall-Arber, M., & Sorensen, J. (2016). The use of personal flotation devices in the northeast lobster fishing industry: An examination of the decision-making process. *American Journal of Industrial Medicine, 59*, 73–80.

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communications Monographs, 59*(4), 329–349.

Theodore Teske is a health communication specialist for the National Institute for Occupational Safety and Health (NIOSH) Western States Division (WSD). He has worked for NIOSH since 1999 developing communication-based interventions for the commercial fishing, mining, oil and gas, and aviation industries. His research is focused on improving the process of bringing NIOSH research into practice, including conducting health communication and research translation projects with focuses on occupational culture, social marketing, and technology transfer. He received his BA in Broadcasting and MA in Communication and Leadership Studies from Gonzaga University in Spokane, WA.

Dr. Jennifer Lincoln is an Injury Epidemiologist currently serving as the Associate Director for the Office of Agriculture, Forestry and Fishing Safety and Health, National Institute for Occupational Safety and Health. She received her PhD from Johns Hopkins, School of Public Health. Dr. Lincoln's career has focused on scientific research and leadership to develop tailored risk-reduction interventions for high-risk work, especially in the prevention of traumatic injuries among workers in the commercial fishing industry. In 2007, she created the NIOSH Commercial Fishing Safety Research and Design Program and in 2015 established the Center for Maritime Safety and Health Studies.

Devin Lucas is an injury epidemiologist at the National Institute for Occupational Safety and Health (NIOSH), Western States Division, stationed in Anchorage, Alaska. Dr. Lucas has a PhD in Occupational and Environmental Health from Oregon State University and 15 years of experience designing epidemiologic studies of work-related injuries. Dr. Lucas leads research projects that apply epidemiologic methods to identify and characterize occupational hazards and focuses on developing, testing, and promoting practical and scalable injury prevention solutions.

Samantha L. Case is an epidemiologist at the National Institute for Occupational Safety and Health (NIOSH), Western States Division in Anchorage, AK. She earned her MPH from the University of Alaska Anchorage and is currently a PhD student in Safety Sciences at Indiana University of Pennsylvania. She has worked with NIOSH since 2014 to conduct research on safety and injury prevention in the US commercial fishing industry.

Christy L. Forrester, PhD, MS is a health scientist with the National Institute for Occupational Safety and Health (NIOSH), Communication and Research to Practice (r2p) Office. She leads the NIOSH r2p team in developing and adapting innovative strategies and solutions to bridge gaps in the translation of research findings into practical workplace use to improve the safety and health of workers. Dr. Forrester earned her PhD in communication with a focus on organizational and risk communication from George Mason University and MS in epidemiology from the University of Cincinnati.

Journal of Safety Research 83 (2022) 418

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

In Memoriam: Anne Taylor McCartt

The National Safety Council and the editorial team of the *Journal of Safety Research* extend our deepest sympathies to family and friends of Anne McCartt. Anne was a dedicated *Journal of Safety Research* editorial board member for over 17 years and served on the National Safety Council Board of Directors since 2016. Her contributions of time and expertise to safety research cannot be repaid and we will miss her immensely. Below you will find her beautiful obituary, provided by her loving husband Michael.

Anne Taylor McCartt, PhD, 73, of Ballston Spa, New York, died peacefully at her home on July 23, 2022. Anne was born in Durham, NC on December 12, 1948, to the late James Spurgeon and Virginia Ann Taylor McCartt. After a lengthy and fruitful national and international career in highway safety, Anne retired in 2016 as Senior Vice President, Research, at the Insurance Institute for Highway Safety (IIHS) in Arlington, Virginia. She directed a multidisciplinary research staff whose work centered on finding ways to change driver behavior, improve highway design, and make vehicles safer. Anne authored more than 230 technical reports and scientific papers on alcohol-impaired driving, crash avoidance technologies, teenage drivers, distracted driving, and other vehicle safety topics. Anne frequently appeared on multiple media outlets to present research findings to the public. She regularly presented research internationally and was widely respected as a highway safety research mentor.

Anne loved her family deeply and was eagerly awaiting the birth of her third grandchild, Olivia Holland Taylor Curcio, first child of Christopher and Katie Curcio. She was known by her family and friends for her love of beauty and the created world, her passion for the New York Yankees and the Duke Blue Devils, and her ability to quickly complete the New York Times Crossword Puzzle in ink. Anne also loved growing up in the foothills of the Great Smoky Mountains with her mother's incredible baking, her beautiful home and gardens in Ballston Spa, and travel. She had traveled throughout Europe and to Australia and New Zealand multiple times.

Beloved wife of Michael Curcio for 33 years. Remembered in love by all of her family, including her daughter Sara Elizabeth Scoles (Matt) and her son Christopher Michael Curcio (Katie); her sister and brother Peg Hess (Howard) and Jimmy McCartt; her brothers and sisters in-law Joseph Curcio (Tina), Jerry Curcio (Diane), Anthony Curcio (Gina); her granddaughters Emilia and Molly Scoles; her nephews and nieces Jeremy Hess (Leili Besharat) and Kristen Hess (John Fumbanks); Casey Curcio (Kim), Karen Murdoch (Joseph), Tom Curcio (Holly), Trisha Curcio (Leo Munoz), Marisa Curcio (Megan Coiley). Her great nieces and nephews Clementine and Bennett Hess, Taylor Goldstein, Lexi and Rachel Fumbanks, Joe and Isabella Curcio, Brodey and Tyler Curcio, and Alissa Murdoch. Also loved by her aunts and uncles, Katherine McDaniel (Edgar, deceased),Betty Lee Thompson (Charlie, deceased), Hugh M. Taylor, deceased (Willie Love,deceased), Marie Mirro (Joe, deceased), and Pat Mirro (Jerry, deceased) and cousins, Linda Baber, Betsy Boyer, Kathy Love Erikson, Jana Jensen, David Thompson and their families.

Anne grew up in East Tennessee where her father, The Reverend J.S. McCartt, served as an ordained minister in the United Methodist Church. She graduated from Fulton High School in Knoxville as valedictorian and received a B.A magna cum laude from Duke University (1970). She then moved to Albany, NY where she completed a MA in Secondary Mathematics Education in 1972; a MLS in 1975; and a PhD in public administration and policy (1988), all from SUNY Albany. She began her career in highway safety in 1982, serving as Deputy Director at the Institute for Traffic Management and Research at the State University of New York at Albany (SUNY Albany) and an associate research professor at the Rockefeller College of Public Affairs and Policy. After working as a Senior Associate with Preusser Research Group, Inc., she joined IIHS in 2001. Anne served on the boards of MADD (Mothers Against Drunk Driving) and the National Safety Council and was a longtime member and served as President of the Association for the Advancement of Automotive Medicine.

A gift may be made in Anne's name to one of Anne's favorite causes: Doctors Without Borders (doctorswithoutborders.org), American Society for the Prevention of Cruelty to Animals (ASPCA.org), Caring for a Cure at Massachusetts General Hospital (https:because.massgeneral.org/campaign/caring-for-a-cure/c112468).

https://doi.org/10.1016/j.jsr.2022.09.017