BMJ Health & Care Informatics

Network graph representation of COVID-19 scientific publications to aid knowledge discovery

George Cernile,[1] Trevor Heritage,[1] Neil J Sebire ,[2] Ben Gordon,[2] Taralyn Schwering,[1] Shana Kazemlou,[1] Yulia Borecki[1]

To cite: Cernile G, Heritage T, Sebire NJ, *et al*. Network graph representation of COVID-19 scientific publications to aid knowledge discovery. *BMJ Health Care Inform* 2021;**28**:e100254. doi:10.1136/bmjhci-2020-100254

Received 12 October 2020
Revised 01 December 2020
Accepted 11 December 2020

[1]Inspirata, Tampa, Florida, USA
[2]HDRUK, London, UK

Correspondence to
Professor Neil J Sebire;
neil.sebire@hdruk.ac.uk

ABSTRACT

Introduction Numerous scientific journal articles related to COVID-19 have been rapidly published, making navigation and understanding of relationships difficult.

Methods A graph network was constructed from the publicly available COVID-19 Open Research Dataset (CORD-19) of COVID-19-related publications using an engine leveraging medical knowledge bases to identify discrete medical concepts and an open-source tool (Gephi) to visualise the network.

Results The network shows connections between diseases, medications and procedures identified from the title and abstract of 195 958 COVID-19-related publications (CORD-19 Dataset). Connections between terms with few publications, those unconnected to the main network and those irrelevant were not displayed. Nodes were coloured by knowledge base and the size of the node related to the number of publications containing the term. The data set and visualisations were made publicly accessible via a webtool.

Conclusion Knowledge management approaches (text mining and graph networks) can effectively allow rapid navigation and exploration of entity inter-relationships to improve understanding of diseases such as COVID-19.

INTRODUCTION

There is urgency to accelerate research that can help contain the spread of the COVID-19 epidemic, to ensure that those affected are promptly diagnosed and receive optimal care and to support research priorities in a way that leads to the development of global research platforms in preparation for the next disease epidemic, thus allowing for accelerated research, and research and development for diagnostics, therapeutics and vaccines and their timely access. In view of the urgency of this outbreak, the international community is mobilising to find ways to significantly accelerate the development of interventions.[1] Experts have identified key knowledge gaps and research priorities and shared scientific data on ongoing research, thereby accelerating the generation of critical scientific information to contribute to the control of the COVID-19 emergency.[2]

However, the pace and volume of research mean that it is hard to stay up to date with the growing body of new scientific papers about the disease and the novel coronavirus that causes it. To mitigate this, many organisations are hosting digital collections holding thousands of freely available papers that can help researchers quickly find the information they seek, and several studies have described or mapped the rapid evidence generation in this area.[3–5] By one estimate, the COVID-19 literature published since January has reached more than 200 000 papers and is doubling every 30 days, one of the biggest episodes of disease-specific publications of scientific literature ever.[6]

One approach to navigating and searching such knowledge collections is through graph databases, which represent the connections between the semantic concepts with nodes, edges and other properties of the data.[7] This allows semantic queries to search across the data set to find relationships between papers on any set of data points. Such a graph displayed in a visualisation tool gives an interactive overview of the nodes and connections between the concepts across the papers and allows one to move around and focus on what is interesting to the researcher.[8]

The aim of this short report is to demonstrate the feasibility of using a network graph approach for rapid navigation of the COVID-19 literature in a publicly available format and to present an openly available tool for exploring a COVID-19 knowledge data set.

METHODS

The COVID-19 Open Research Dataset (CORD-19) is a rapidly increasing open-source collection of scholarly articles related to the coronavirus which has been designed to facilitate the development of text mining

and information retrieval systems.[9 10] As of 8 August 2020, the data set has 207 311 papers from over 160 000 sources. The articles available include title, abstract, authors, source, publication date and in some cases full text.[11]

We used proprietary natural language processing (NLP) and artificial intelligence (AI) engines, which leverage the heuristic segmentation approach (a fast heuristic search algorithm) and a knowledge-driven approach for concept identification, context determination, inferencing and extraction of corresponding values and units. The engine works with domain-specific knowledge bases of clinical terms, concepts and rules that are tailored to the data to be extracted.[12]

In this study, we used a collection of 10 knowledge bases consisting of a core knowledge base and 9 domain-specific knowledge bases that were built using UMLS (Unified Medical Language System) terms and updated with recently added terms specific to COVID-19: core oncology knowledge base, pharmacological substance (medications) (T121), virus (T005), therapeutic or preventive procedure (T061), sign or symptom (T184), disease or syndrome (T047), gene or genome (T028), immunological factor (T129), finding (T033), and body part, organ or organ component (T023).[13]

The title and abstract sections of all papers in the CORD-19 Dataset were processed against the various knowledge sources to extract discrete data from each paper and were stored in a database. Along with the discrete data, the following metadata were also stored: CORD-19 UID (unique identifier), title, abstract, body text, publication date, URL, authors, journal, knowledge base (which of the 10 available knowledge sources was used to extract the term, term category or question; ie, medication, virus, symptom), paper ID (identification

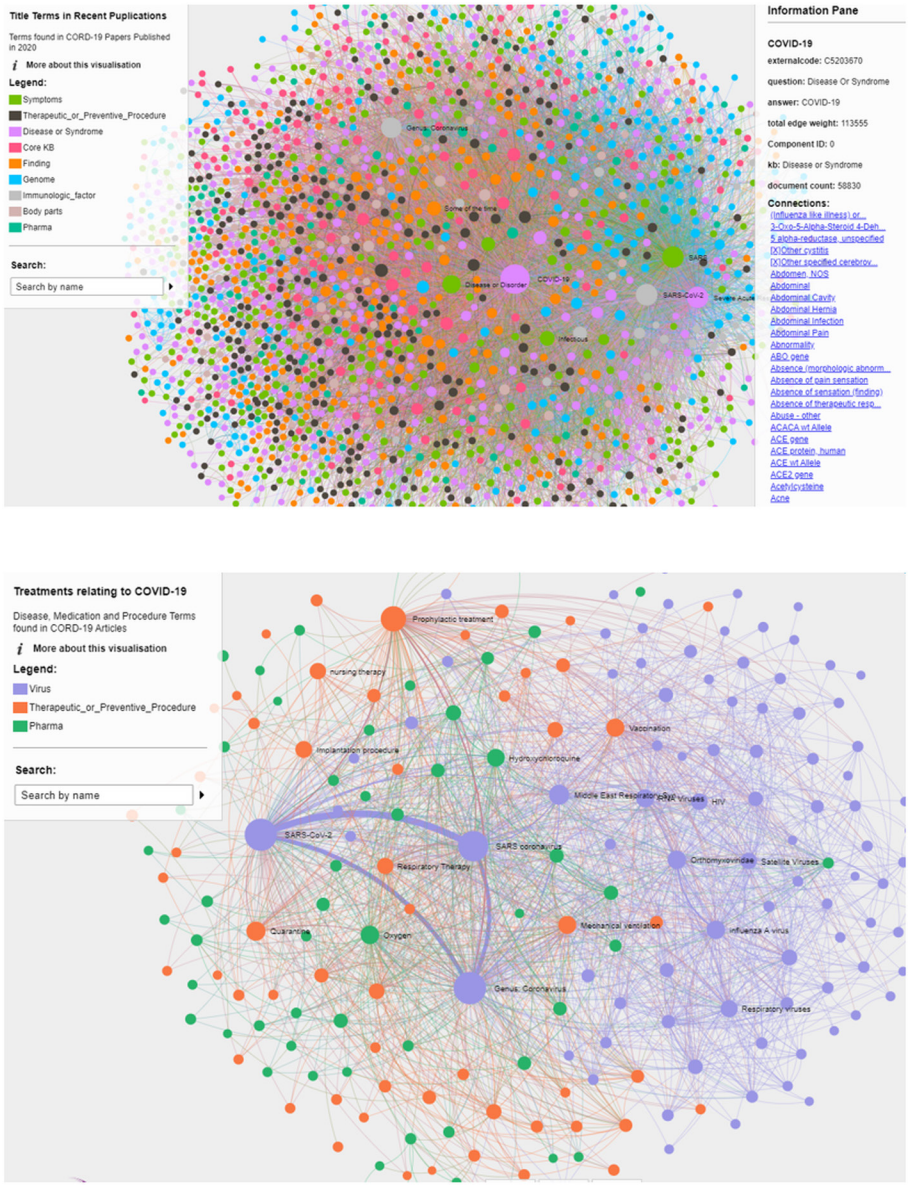

Figure 1 Example of network graphs including high-density network showing concepts associated with COVID-19 (top) and specific query treatment map for COVID-19 (bottom). CORD-19, COVID-19 Open Research Dataset; KB, knowledge base.

Table 1 Extracted concepts from CORD-19 Dataset by knowledge base (semantic type) showing the number of unique terms found and the total number of extracted concepts from each knowledge base, as well as the number of papers containing terms from that knowledge base and the percentage coverage across the entire data set

Knowledge base	Unique terms	Extracted concepts	Papers	Coverage (%)
Body parts	1332	172 438	77 400	37
Core knowledge base	1434	338 552	102 037	49
Disease or syndrome	7195	507 819	152 402	73
Finding	5580	526 433	145 504	70
Genome	9395	419 413	86 073	41
Immunological factor	1845	130 996	45 912	22
Pharmacological substance	2599	58 308	30 494	15
Symptoms and side effects	8883	630 116	144 063	69
Therapeutic or preventive procedure	4923	332 260	111 277	54
Virus	1308	240 993	84 325	41
Total	44 494	3 357 328	195 958	94

Papers may contain multiple extracted concepts, and concepts may be found in multiple papers within the knowledge base; hence, we provide both all extracted concepts using the natural language processing tool in addition to the number of unique terms.
CORD-19, COVID-19 Open Research Dataset.

of the paper in the CORD-19 Dataset from which the term was extracted) and source section (either title or abstract). Generic terms with little significance were determined, for example, 'air', 'water' and 'virus', and these were removed from the set of extracted concepts.

Networks created with the entire set of results and all the knowledge sources are very large with too many terms to visualise details in the data. For this reason, a subset of the data was selected to enable meaningful visual exploration by selecting a subset of the knowledge sources, paper sections and publication year for each network based on specific medical themes, for example, treatments, cardiology and so on. Duplicate terms (same terms found in multiple knowledge sources) were consolidated to remove redundant data. For example, 'obesity' is included in both the 'symptoms and side effects' and the 'disease or syndrome' knowledge sources.

For each term found in a paper, a link was created to every other term on the same paper. The culmination of these links for all papers resulted in a network structure where the weight of a connection between any two terms was determined by the number of papers linking the terms. Additional filtering was performed to refine the scope of the network and removal of noise to aid readability and navigation; for example, links with low weights were removed, as were links with terms that were disconnected from the rest of the network.

The open-source software tool Gephi was used to create a visualisation of the network using the collections of terms and connections that made up the network structure.[14] Network nodes were coloured based on the knowledge source, with the size of the nodes proportional to the frequency of each term and the connection weight (edge thickness) based on the number of associated papers. The networks were exported and visualised in an HTML (hypertext markup language) website using the Sigma JS JavaScript library.

RESULTS

A total of 207 311 publications from the CORD-19 Dataset were processed using the NLP engine. In total 3 357 328 total entities were extracted from 195 958 of these papers, consisting of 44 494 unique terms. Four network graphs were generated using these extracted data: cardiological diseases, lung diseases, title network and treatment network (https://nlp.inspirata.com/networkvisualisations/treatmentnetwork/#) (figure 1). The filters applied to create each of the networks and the number of terms, edges and papers involved in each network are displayed in table 1 and online supplemental table 1.

DISCUSSION

Recently there have been several initiatives to explore knowledge graphs in medical data and with some applied to aspects of COVID-19-associated published literature.[15 16] This study has demonstrated the feasibility of using a graph database approach to create a targeted concept association networks as an interactive way to allow users to easily navigate the rapidly growing COVID-19-related literature, and particularly as a way to understand and explore the relationships between key concepts within this corpus of literature articles, which is potentially widely applicable to other disease areas.

This approach is also applicable to any collection of scientific literature, such as PubMed or ClinicalTrials.gov, or proprietary document management systems. Specific lexical terms and knowledge sources can be used from

the UMLS collection or other publicly available sources and imported for use with NLP/AI engines.

One constraint of this knowledge mining approach is that the network size increases as more knowledge sources are added. As a consequence, methods to simplify the network to enable easier visual exploration are required, such as 'pruning'.[17] The concept is to remove a subset of the 'least important' edges while maintaining the overall graph connectivity, since it becomes more difficult to interactively explore without a priori knowledge of the specific knowledge sources as the network density increases. Another limitation is that the network only shows the first-level connections or the direct connection between papers and concepts. It does not find connections between concepts that span several papers, although this can be achieved by traversing the network visually.

We addressed these limitations of network size and the search for deep connections by implementing a breadth-first search on the network structure.[18] Essentially this approach searches the graph data structure beginning at a root node by exploring all of the adjacent nodes at a given depth before moving to the nodes at the next subsequent level. This search type is efficient and can be applied across very large networks, even when all the knowledge sources are used simultaneously, and can find the shortest path connections (the trail of papers) between any concepts.

This study has demonstrated that an approach using graph databases and network analysis can be developed rapidly and is a useful approach to understanding large volumes of medical literature, quickly grasping the current state of our knowledge, and discovering previously unknown or unnoticed relationships between emerging medical concepts. The unusual circumstances of a global pandemic have given rise to the assembly of an unprecedented volume of medical literature, and this work demonstrates a powerful approach to condensing the literature into insights that help us fight this disease. Further development of this approach will enable ongoing analysis and deep searching of large collections of literature, such as PubMed, and application to other disease areas, as well as for target or biomarker discovery.[19–21]

Contributors TH, GC, TS, SK and YB conceived the study and performed the data analysis and tool development. All authors contributed to the manuscript preparation and writing and critically improved the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests TH, GC, TS, SK and YB are employed by Inspirata, a company specialising in health data management, and carried out the work as part of their employment.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

ORCID iD

Neil J Sebire http://orcid.org/0000-0001-5348-9063

REFERENCES

1 Kambhampati SBS, Vaishya R, Vaish A. Unprecedented surge in publications related to COVID-19 in the first three months of pandemic: a bibliometric analytic report. *J Clin Orthop Trauma* 2020;11:S304–6.
2 Coronavirus disease, 2019. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [Accessed 3 May 2020].
3 Zyoud Sa'ed H, Al-Jabi SW. Mapping the situation of research on coronavirus disease-19 (COVID-19): a preliminary bibliometric analysis during the early stage of the outbreak. *BMC Infect Dis* 2020;20:561.
4 Liu N, Chee ML, Niu C, *et al*. Coronavirus disease 2019 (COVID-19): an evidence map of medical literature. *BMC Med Res Methodol* 2020;20:177.
5 Albahri AS, Hamid RA, Alwan JK, *et al*. Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. *J Med Syst* 2020;44:122.
6 Brainard J. Scientists are drowning in COVID-19 papers. Can new tools keep them afloat? *Science* 2020.
7 Lysenko A, Roznovăț IA, Saqi M, *et al*. Representing and querying disease networks using graph databases. *BioData Min* 2016;9:23.
8 Fensel D, Şimşek U, Angele K. *Knowledge graphs*, 2020.
9 Kaggle. COVID-19 open research dataset challenge (CORD-19). Available: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge [Accessed 12 Oct 2020].
10 Lu Wang L, Lo K, Chandrasekhar Y, *et al*. CORD-19: the Covid-19 open research dataset. *ArXiv* 2020. [Epub ahead of print: 22 Apr 2020].
11 Semantic Scholar. [PDF] CORD-19: The Covid-19 Open Research Dataset. Available: https://www.semanticscholar.org/paper/CORD-19%3A-The-Covid-19-Open-Research-Dataset-Wang-Lo/4a10dffca6dcce9c570cb75aa4d76522c34a2fd4 [Accessed 12 Oct 2020].
12 Inspirata Launches Cloud-Based Cancer and Clinical Data Extraction Software Service. Available: https://www.inspirata.com/inspirata-launches-nlp-on-demand/ [Accessed 12 Oct 2020].
13 Unified medical language system (UMLS). Available: https://www.nlm.nih.gov/research/umls/index.html [Accessed 12 Oct 2020].
14 Gephi - The Open Graph Viz Platform. Available: https://gephi.org/ [Accessed 12 Oct 2020].
15 Domingo-Fernández D, Baksi S, Schultz B, *et al*. COVID-19 knowledge graph: a computable, multi-modal, cause-and-effect knowledge model of COVID-19 pathophysiology. *Bioinformatics* : 2020;12.
16 Das D, Katyal Y, Verma J. Information Retrieval and Extraction on COVID-19 Clinical Articles Using Graph Community Detection and Bio-BERT Embeddings. In: *ACL 2020 work NLP-COVID*, 2020.
17 Zhou F, Mahler S, Toivonen H. Simplification of networks by edge pruning. *Lect Notes Comput Sci* 2012;7250:179–98.
18 Wikipedia. Breadth-first search. Available: https://en.wikipedia.org/wiki/Breadth-first_search [Accessed 27 Nov 2020].
19 Shi L, Li S, Yang X. Semantic health knowledge graph: semantic integration of heterogeneous medical knowledge and services. *Biomed Res Int* 2017.
20 Sharma S, Santosh T, Santra B. Incorporating domain knowledge into medical NLI using knowledge graphs. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, 2020.
21 Xu J, Kim S, Song M, *et al*. Building a PubMed knowledge graph. *Sci Data* 2020;7.

BMJ Health & Care Informatics

COVID-19 and beyond: virtual consultations in primary care—reflecting on the evidence base for implementation and ensuring reach: commentary article

Freda Mold ,[1] Debbie Cooke,[1] Athena Ip,[1] Parijat Roy,[2] Susan Denton,[3] Jo Armes[1]

To cite: Mold F, Cooke D, Ip A, *et al*. COVID-19 and beyond: virtual consultations in primary care—reflecting on the evidence base for implementation and ensuring reach: commentary article. *BMJ Health Care Inform* 2021;**28**:e100256. doi:10.1136/bmjhci-2020-100256

Received 13 October 2020
Revised 15 December 2020
Accepted 21 December 2020

[1]Faculty of Health and Medical Sciences, School of Health Sciences, University of Surrey, Guildford, UK
[2]River Wey Medical Practice, Farnham, UK
[3]Guildowns Group Practice, Guildford, UK

Correspondence to
Dr Freda Mold;
freda.mold@surrey.ac.uk

INTRODUCTION

COVID-19 has resulted in an unprecedented expansion of virtual consultations in primary and community care services.[1] Although virtual consultations have been available for a long time, they were not widely adopted before COVID-19.[2] There has now been a rapid deployment of virtual consultations and telephone consultations (TCs) in response to COVID-19.[3]

Virtual consultations come in many forms, including synchronous TCs, video, text/image messaging and asynchronous email consultations. Virtual consultations enable communication with a range of healthcare staff and are based on an array of provider platforms (Attend Anywhere, WebGP, accuRx, eCONSULT). Digital (or online) triage systems are often linked to virtual consultations, to determine the priority and urgency of a patient condition to manage demand, and are considered the first step in determining whether a virtual consultation is needed. Digital triage is a workflow management system, such as FootFall and AskMyGP. Although much has been written about triage in primary care,[4] or indeed the use of the telephone in arranging care,[5] relatively little is known about the potential for web-based, real-time (synchronous) communication for some patient groups. Before COVID-19 most practices offered TCs, with few offering video consultations. However, the potential for video consultations, in particular, has still not been realised. In time, these may be considered complementary forms of care delivery.[1 2] COVID-19 has led to the rapid expansion of virtual consultations, in its various forms. This commentary paper focuses primarily on web-based virtual consultations.

There are currently vast avoidable inequalities in healthcare and health outcomes (mortality and morbidity) for different service user groups globally.[6] This includes Black, Asian, other ethnic and older groups who are also at higher risk of contracting COVID-19 and being adversely affected.[7] Evidence collected before COVID-19 indicated the potential of virtual consultations to widen care disparities for specific groups, including people with physical/cognitive disabilities (sensory/communication impairments (audio/sight))[8] and those living with social deprivation, with limited digital access (including vulnerable groups)[9 10] and areas with poor broadband coverage.[11]

COVID-19 clinicians and researchers/academics are moving forward in developing ways to mitigate these disadvantages by developing strategies to enable greater access and engagement for a wider range of service users.[12] Primary care has adapted rapidly to virtual consultations and embraced their use, despite concerns about confidentiality, safety and security, with the view that many professional and organisational lessons can be learnt to improve access and delivery. To improve delivery moving forward during COVID-19, there is a need to reflect on experiences of delivery, examine and build on current evidence, and consider the training needs and competencies of healthcare professionals delivering care virtually.

SERVICE USERS

There is little long-term evidence to support the implementation of virtual consultations due to a lack of knowledge about barriers to use, non-use, and whether these services disproportionately disadvantage vulnerable

and hard-to-reach groups. We need to understand how to promote virtual consultations without worsening existing health inequalities.[10] The current evidence base also reports mixed impact on health outcomes[13 14] and how use of virtual consultation promotes self-help is unclear.[15]

Service users and their carers' ability to consult with general practice clinicians digitally/online improves access for specific groups (younger, women, employed)[10 16 17] by providing an alternative route to care[18 19] or perceptions that it might lead to better follow-up.[16]

Other benefits have also been reported, such as enabling some service users to express themselves more openly on health issues through virtual consultation,[20] the possibility of sharing images, when needed[8 10] or reducing relatives/carers' need to accompany service users to face-to-face appointments,[19] greater opportunities for multi-person interaction with relatives/carers,[19] reducing time off work for appointments[21] and widening opportunities for access[18] including those in geographically remote areas,[22] if housebound[23] or when transport is costly/time-consuming.[19 24] As such service users report greater satisfaction, convenience[9 10] and timeliness of care.[25] Virtual consultations might also empower individuals thereby improving engagement with service providers.[26] Consideration might also be made to service users who are shielding (and housebound) during the COVID-19 pandemic.

Virtual consultations have also raised questions about the suitability of delivery across different service user groups (such as vulnerable/ hard-to-reach and those with sensory and/or learning disabilities) across different health conditions, or at different time points in the service user care journey (newly diagnosed vs management of long-term chronic conditions).[9 10] In exploring virtual consultations, both its barriers and facilitators, we also need knowledge to go beyond simply looking into age, gender or sociodemographic differences of users, and consider other factors.[25] These include the use of proxies in virtual consultations (children/grandchildren offering technical support, use of family translators), safeguarding for specific service user groups, accuracy of medicine prescribing/issues, and confidentiality and data security.[13 27]

There is a need, therefore, to understand and carefully evaluate the process of implementing virtual consultations so that we can learn how to embed this approach as part of routine care, so it is inclusive. This will be invaluable in exploring its use, especially over time, and may also highlight unexpected consequences of use, such as offering alternative access routes or possible disparities in access for specific service user groups. Indeed, further evidence would be invaluable to understand the experience of specific groups especially those who are digitally disadvantaged, such as those living in areas with poor/intermittent broadband coverage, those with no means to use or acquire the technology needed to use virtual consultations or those with low computer/technology literacy. Likewise, practical issues also need consideration when attempting to develop and implement virtual consultation, including insufficient broadband widths for both service users and clinicians.[17]

CLINICIANS/WORKFORCE

Early pre-COVID-19 research indicates primary care staff/clinicians were already concerned about the impact of implementing virtual consultations on workload, and potential to change the length, workflow or structure of the working day.[28 29] Although these concerns might have declined since COVID-19, use of virtual consultations during the pandemic may highlight other documented concerns[30] such as raised clinical risk/medicolegal risk,[8] shifting role responsibilities/greater reliance on general practice administrators[31] and increased need for subsequent follow-up (either via telephone or face-to-face), thereby adding to clinicians' existing workload concerns.[18]

The conversation around virtual consultations also highlights the need to be aware of the very real impact of this new working style/pattern, and its impact on well-being or 'cognitive load' of primary care colleagues.[32] As this type of working uses many more skills than face-to-face communication, there needs to be an acknowledgement of the burden of multitasking and the potential impact this might have on the patient–professional interaction. Virtual consultations, however, might equally improve communication, by building rapport and confidence in openly discussing health issues[20 21] which might not have otherwise been raised. Multiple skills are needed for both clinicians and service users to feel the encounter was successful. This includes having the camera/sound equipment, to begin with, the information technology (IT) skills to use the virtual consultation system and the ability to pick up/provide verbal and visual cues in the conversation (if via video), and overall confidence in the system and security to share personal health information via remote means.

Other benefits of virtual consultations may be to offer new opportunities to extend service provision to a broader variety of service user groups and clinical settings,[33] provide shared learning opportunities[19] and greater joint working across professional groups, including allied health professionals (physiotherapists, pharmacists) and settings,[34–38] and possibly reducing the number of referrals to specialists.[39]

Virtual consultations might also become the new medium for general practitioners (GPs) to use in their additional training roles. This provides opportunities for shared clinics and learning (eg, dermatology, heart failure) using virtual consultations as a learning event[40] but also presents potential challenges when multiple clinicians are needed (eg, joint surgeries). Virtual consultations might also be reframed towards more patient-centred approaches, allowing for different ways of engagement to services.[41]

Training and supporting materials will provide reassurance for clinicians and might be critical to embedding virtual consultations into current general practice. Training for staff to use virtual consultations is essential to familiarise them with the system, the equipment and treatment procedures. Moreover, the lack of staff training has been shown to affect the uptake of virtual consultation.[42] Training material will also need to be grounded on existing, up-to-date national guidance, local primary care policies and governance. The growth of virtual consultation has impacted on the need to include advanced communication training for GPs, which starts from an undergraduate level through to continuing professional development. Education techniques currently include video-based feedback (either prerecorded or in real-time) to improve GP communication skills in consultations.[43 44] COVID-19 has presented challenges for GP trainers in terms of delivering training that ensures safety for all participants in the consultation (trainer/observer, GP, patient/actor) in face-to-face encounters. Solutions for remote training include audio/video three-way consultations between participants, which may open exciting possibilities for training, enabling wider geographical, specialists and joint-working across clinics/settings.

Previous research indicates concerns about local governance[8] and that general practice policies were either not known about or followed.[9] However, this may rapidly change with the publication of guidance on using virtual consultation, such as that produced by the General Medical Council to support ethical decision-making and risk assessment.[45] However, clear guidelines for general practice staff are necessary to support implementation.[8 13] Since COVID-19, it is possible that the rapid development of local policies, together with research, may result in more robust and reassuring care delivery for general practice staff moving forward.

A recent study demonstrated a higher risk of death from COVID-19 among GPs from single-handed general practices in areas of economic deprivation.[46] This increased risk may be explained by the limited implementation of virtual consultations, due to practice size/resource resulting in a greater need for face-to-face care delivery. As such, uptake of virtual consultations within these areas needs to be increased and further support provided to reduce the risk of contracting COVID-19 and further widening of health inequalities across poorer regions.

THE ORGANISATION OF CARE

Before COVID-19 there seemed to be fewer benefits for the use of virtual consultations for general practices, this was mainly based on the lack of effective platforms or limited technology stock for use in primary care.[18] Since COVID-19 the rapid deployment of digital/online platforms and investment (laptops with cameras, duel headsets) has led to improved technical infrastructure.[16] However, implementation is reliant on several factors, from consistent internet provision and reliable wi-fi[7] to continued adjustments and agile approach to working practices to cater to specific service user groups and expectations of all users. This also includes flexibility about redistributing tasks to support implementation. This will include greater reliance on administrative staff to monitor/manage and signpost/support new systems workflow,[31] such as time spent on virtual consultations over other modes of contact (and differences in distribution between staff).[15]

Virtual consultations may offer opportunities for reform in primary care, however, we must be cautious not to consider it to be a golden bullet. A recent study by Farr *et al* shows that although virtual consultations may offer alternative routes to care access for some service users, it may also inadvertently lead to subsequent or more frequent follow-up visits, either via telephone or face-to-face[18 25 47] or an even greater need for care delivery 'downstream' between secondary and primary care providers.[48]

Potential benefits of virtual consultations might be the rapid responsiveness to care, including use of emailed images or use in visual assessments to pick up on visual cues.[17 23] This is especially important when needed to support diagnosis, and when physical presence is deemed too risky, in light of COVID-19, and when socially distanced care provision is preferred.

It may be too soon to evaluate whether there are any cost-savings from virtual consultations[15 30] as these are reliant on robust and long-term evaluation data. However, the opinion of some GPs is that virtual consultations may be cost-effective for those in rural areas.[19] The picture is unclear as to whether they are cost-effective long term, especially as the financial implications of this mode is reliant on many other factors, such as reducing the number of face-to-face appointments, need for follow-up face-to-face/telephone calls/visits—for the same issue, or whether more referrals are made, pushing demand to other parts of the health system (two-way direction between primary and secondary care). Ultimately, cost and organisational considerations may dictate the success and long-term sustainability of this service.

Integration of virtual consultations is an important topic if post-COVID-19 care via this mode is to become more established. However, the need to successfully embed them into the current system while limiting any barriers to its interaction with pre-existing systems is complex. There is a need to fully embed these practices across different platforms and primary, secondary care and tertiary care across regions and services. Likewise, integration of virtual consultation between health and social care organisations, like care homes, could further offer opportunities to improve timely care delivery and prevent hospital admission.[11] The impact of these services is again yet to be explored in greater detail.[11 21 25]

Recent National Health Service (NHS) England guidance illustrates the need to support general practices to scale up and extend the range of digital solutions to meet current care needs.[49] While there remain variations in structural factors (phone lines/equipment issues)[9] and

the need to fully integrate and financially support these systems across differently sized practices, there remains great optimism that by working together, service gaps and the inequalities that might arise in access can be addressed.

FUTURE RESEARCH AND SERVICE DEVELOPMENTS

Research to date into virtual consultations use in primary care has indicated many areas for future research. These include exploration of equality issues associated with use for specific service user groups[10 18] (acceptability, barriers and facilitators), and professional/organisational issues regarding quality, cost and sustainability of virtual consultations over time.

There is good evidence to indicate how routine/low-acuity illness might be managed using virtual consultations.[47] However, virtual consultations do pose challenges to primary care staff in diagnosing and managing more complex/chronic conditions or multimorbidities.[3 30] Although we know some service user groups—such as those with type 2 diabetes—have been early adopters of both digital/online access to medical records and virtual consultations,[50–52] there is still relatively little known about the need for efficiency and sustainability of this type of contact for different service user groups, conditions and across geographically diverse areas.[22]

This research agenda can only be met by co-design studies with the central involvement of clinicians, a range of staff and service users across geographically diverse areas, to prioritise the research agenda, co-create more equitable systems and disseminate these practices.

This research also needs to be conducted by collaborative research teams across the UK, so we can learn from each other, general practice colleagues and service users to improve care provision, ensuring it is person-centred and addresses what is important to whom, at what times/when and for what purpose. In conducting this research, we might also find unintended consequences of virtual consultation use (or non-use), which would be equally valuable in understanding where systems fail to deliver, for whom and why.

Likewise, we may be surprised that barriers to use can be overcome with the implementation of very simple strategies, such as offering IT support at a general practice level, initiating new ways of working between staff members, and providing greater staff training and support and service user empowerment/support.[15 50] However, such strategies can only be implemented and sustained by acknowledging general practice IT, governance and reimbursement needs.

Indeed, pre-COVID-19 recommendations have also indicated the need for more robust evidence to support the large-scale roll-out of virtual consultations in primary care[53] ; however, the pandemic has changed this landscape, accelerating the need for high-quality research across geographical areas and general practice types/sizes.

Evidence-based frameworks are available through which to map and evaluate the success of new systems, including the Non-adoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies[54] and the Candidacy framework,[55] which might provide rich data concerning how service users seek healthcare and the underpinning decisions, behaviours and psychosocial factors which contribute to their care journey.

Future research outputs need to support a range of vulnerable service user groups, based on a range of multidisciplinary teams working across a growing number of integrated services in the community: applied to offer practical solutions (communication toolkits/IT support, professional and service user education, IT equipment at each site); inclusive of a range of service user groups (vulnerable and hard-to-reach groups); and based on an agile/flexible system (which can adapt at a general practice-level and service users circumstances, as and when they occur). Although the Royal College of General Practitioners and NHS England are beginning to develop such resources,[56 57] much work needs to be done to ensure equity and safety of these systems. Research into the acceptability and experience of use during the pandemic would also be valuable to move forward our knowledge and application of this technology.

IMPLICATIONS FOR GENERAL PRACTICE AND EDUCATION

Virtual consultations may now be here to stay, with the UK government calls for greater expansion and implementation moving forward.[58 59] This permanency has massive implications for general practice in terms of workload/flow, what works—for whom and why, what does not work—for whom and why, when to make reasonable adjustments and, how agile primary care is adapting to a hybrid approach of face-to-face and digital/online care delivery. Moreover, this change also has implications for the education of our future workforce, beginning at undergraduate level, provoking the need to advance digital/online communication skills training and education (eye contact, body language, environmental conditions). Furthermore, if virtual consultations are to maintain the quality of care, known to be the cornerstone of primary care, there is also a need to understand how they may impact on relational issues, such as building rapport, aiding communication, and demonstrating empathy and compassion, on which the quality of the patient–professional relationship is grounded.[18 20 23 31] There is also a need to consider care continuity, which both GPs and patients value.[60]

CONCLUSION

We need to maintain person-centred, timely and equitable access while also supporting staff to work safely and efficiently during, and crucially after, the COVID-19 pandemic. While there is research to indicate that virtual

consultations can promote access to care,[61] little is known about the barriers and facilitators of this type of consultation for people from vulnerable groups in primary care[51] and the long-term implications for digital access. Virtual consultation has facilitated access to primary care for many people throughout the COVID-19 pandemic. Future research will show how virtual consultation impacts on those using health services, their care providers and the organisation of care. Until then web-based virtual consultation will complement other care delivery options.

Contributors FM was the lead author, conceptualised the article and wrote the first outline draft of the manuscript. FM, DC, AI, SD and JA made multiple revisions to the manuscript, edited and reviewed it. AI undertook rapid literature review, data extraction and added to referencing. PR and SD reviewed, edited and contributed to the team discussion about all aspects of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD

Freda Mold http://orcid.org/0000-0002-6279-5537

REFERENCES

1 British Medical Association. COVID-19: steps for GP practices to take, 2020a. Available: https://www.bma.org.uk/advice-and-support/covid-19/practical-guidance/covid-19-steps-for-gp-practices-to-take [Accessed 13 Apr 2020].
2 British Medical Association. COVID-19: video consultations and homeworking, 2020b. Available: https://www.bma.org.uk/advice-and-support/covid-19/adapting-to-covid/covid-19-video-consultations-and-homeworking [Accessed 20 Aug 2020].
3 Royal College of General Practitioners. *Press release. RCGP survey provides snapshot of how GP care is accessed in latest stages of pandemic health business news*, 2020.
4 Holt TA, Fletcher E, Warren F, *et al*. Telephone triage systems in UK general practice: analysis of consultation duration during the index day in a pragmatic randomised controlled trial. *Br J Gen Pract* 2016;66:e214–8.
5 Bardsley M, Steventon A, Doll H. Impact of telehealth on general practice contacts: findings from the whole systems demonstrator cluster randomised trial. *BMC Health Serv Res* 2013;13:395.
6 World Health Organisation. Social determinants of health. key concepts Webpage. Available: https://www.who.int/social_determinants/thecommission/finalreport/key_concepts/en/ [Accessed 8 Oct 2020].
7 Intensive Care National Audit & Research Centre (ICNARC). COVID-19 report source: ICNARC case mix programme database, 2020. Available: https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports [Accessed 13 Apr 2020].
8 Brant H, Atherton H, Ziebland S, *et al*. Using alternatives to face-to-face consultations: a survey of prevalence and attitudes in general practice. *Br J Gen Pract* 2016;66:e460–6.
9 Atherton H, Brant H, Ziebland S, *et al*. Alternatives to the face-to-face consultation in general practice: focused ethnographic case study. *Br J Gen Pract* 2018;68:e293–300.
10 Mold F, Hendy J, Lai Y-L, *et al*. Electronic consultation in primary care between providers and patients: systematic review. *JMIR Med Inform* 2019;7:e13042.
11 Wade V, Whittaker F, Hamlyn J. An evaluation of the benefits and challenges of video consulting between general practitioners and residential aged care facilities. *J Telemed Telecare* 2015;21:490–3.
12 NHS Implementation Plan. Implementing phase 3 of the NHS response to the COVID-19 pandemic, Ref: 001559, 2020. Available: https://www.england.nhs.uk/wp-content/uploads/2020/08/C0716_Implementing-phase-3-v1.1.pdf [Accessed 8 Oct 2020].
13 Chang F, Paramsothy T, Roche M, *et al*. Patient, staff, and clinician perspectives on implementing electronic communications in an interdisciplinary rural family health practice. *Prim Health Care Res Dev* 2017;18:149–60.
14 Levine DM, Dixon RF, Linder JA. Association of structured virtual visits for hypertension follow-up in primary care with blood pressure control and use of clinical services. *J Gen Intern Med* 2018;33:1862–7.
15 Cowie J, Calveley E, Bowers G, *et al*. Evaluation of a digital consultation and self-care advice tool in primary care: a Multi-Methods study. *Int J Environ Res Public Health* 2018;15:896.
16 Zanaboni P, Fagerlund AJ. Patients' use and experiences with e-consultation and other digital health services with their general practitioner in Norway: results from an online survey. *BMJ Open* 2020;10:e034773.
17 Hammersley V, Donaghy E, Parker R, *et al*. Comparing the content and quality of video, telephone, and face-to-face consultations: a non-randomised, quasi-experimental, exploratory study in UK primary care. *Br J Gen Pract* 2019;69:e595–604.
18 Farr M, Banks J, Edwards HB, *et al*. Implementing online consultations in primary care: a mixed-method evaluation extending normalisation process theory through service co-production. *BMJ Open* 2018;8:e019966.
19 Johansson AM, Lindberg I, Söderberg S. Healthcare personnel's experiences using video consultation in primary healthcare in rural areas. *Prim Health Care Res Dev* 2017;18:73–83.
20 Fagerlund AJ, Holm IM, Zanaboni P. General practitioners' perceptions towards the use of digital health services for citizens in primary care: a qualitative interview study. *BMJ Open* 2019;9:e028251.
21 Donaghy E, Atherton H, Hammersley V, *et al*. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. *Br J Gen Pract* 2019;69:e586–94.
22 Bradford NK, Caffery LJ, Smith AC. Telehealth services in rural and remote Australia: a systematic review of models of care and factors influencing success and sustainability. *Rural Remote Health* 2016;16:3808.
23 Randhawa RS, Chandan JS, Thomas T, *et al*. An exploration of the attitudes and views of general practitioners on the use of video consultations in a primary healthcare setting: a qualitative pilot study. *Prim Health Care Res Dev* 2019;20:e5.
24 Chudner I, Goldfracht M, Goldblatt H, *et al*. Video or In-Clinic consultation? selection of attributes as preparation for a discrete choice experiment among key stakeholders. *Patient* 2019;12:2019:69–82.
25 Carter M, Fletcher E, Sansom A, *et al*. Feasibility, acceptability and effectiveness of an online alternative to face-to-face consultation in general practice: a mixed-methods study of webGP in six Devon practices. *BMJ Open* 2018;8:e018688.
26 Shaw S, Wherton J, Vijayaraghavan S. *Advantages and limitation of virtual online consultations in a NHS acute trust: the vocal mixed-methods study*. Southampton (UK): NIHR Journals Library, 2018.
27 Care Quality Commission. The state of care in independent online primary health services. Findings from CQC's programme of comprehensive inspections in England, 2018. Available: https://www.cqc.org.uk/sites/default/files/20180322_state-of-care-independent-online-primary-health-services.pdf[Accessed 7 Oct 2020].
28 Caffery LJ, Smith AC. A literature review of email-based telemedicine. *Stud Health Technol Inform* 2010;161:20–34.
29 Atherton H, Pappas Y, Heneghan C, *et al*. Experiences of using email for general practice consultations: a qualitative study. *Br J Gen Pract* 2013;63:e760–7.
30 Edwards HB, Marques E, Hollingworth W, *et al*. Use of a primary care online consultation system, by whom, when and why: evaluation of a pilot observational study in 36 general practices in South West England. *BMJ Open* 2017;7:e016901.
31 Casey M, Shaw S, Swinglehurst D. Experiences with online consultation systems in primary care: case study of one early adopter site. *Br J Gen Pract* 2017;67:e736–43.
32 Ambrose L. Remote consulting: recognising the cognitive load. *Br J Gen Pract* 2020;70:295.
33 Jiménez-Rodríguez D, Santillán García A, Montoro Robles J, *et al*. Increase in video consultations during the COVID-19 pandemic: healthcare professionals' perceptions about their implementation and adequate management. *Int J Environ Res Public Health* 2020;17:5112–4.

34 Liddy C, Afkham A, Drosinis P, *et al*. Impact of and satisfaction with a new eConsult service: a mixed methods study of primary care providers. *J Am Board Fam Med* 2015;28:394–403.
35 Liddy C, McKellips F, Armstrong CD, *et al*. Improving access to specialists in remote communities: a cross-sectional study and cost analysis of the use of eConsult in Nunavut. *Int J Circumpolar Health* 2017;76:1323493.
36 Liddy C, Moroz I, Mihan A, *et al*. A systematic review of asynchronous, Provider-to-Provider, electronic consultation services to improve access to specialty care available worldwide. *Telemed J E Health* 2019;25:184–98.
37 McKellips F, Keely E, Afkham A, *et al*. Improving access to allied health professionals through the Champlain BASE™ eConsult service: a cross-sectional study in Canada. *Br J Gen Pract* 2017;67:e757–63.
38 NHS England. First contact physiotherapist (FCP) London: NHS England, 2020. Available: https://www.england.nhs.uk/gp/expanding-our-workforce/first-contact-physiotherapists/ & https://www.england.nhs.uk/wp-content/uploads/2019/05/elective-care-high-impact-interventions-first-contact-practitioner-msk-services-specification.pdf [Accessed 26 Nov 2020].
39 Cravo Oliveira T, Barlow J, Bayer S. The association between general practitioner participation in joint teleconsultations and rates of referral: a discrete choice experiment. *BMC Fam Pract* 2015;16:50.
40 Gallagher J, James S, Keane C, *et al*. Heart failure virtual consultation: bridging the gap of heart failure care in the community - a mixed-methods evaluation. *ESC Heart Fail* 2017;4:252–8.
41 Greenhalgh T, Vijayaraghavan S, Wherton J, *et al*. Virtual online consultations: advantages and limitations (vocal) study. *BMJ Open* 2016;6:e009388.
42 Peel NM, Russell TG, Gray LC. Feasibility of using an in-home video conferencing system in geriatric rehabilitation. *J Rehabil Med* 2011;43:364–6.
43 Dohms MC, Collares CF, Tibério IC. Video-based feedback using real consultations for a formative assessment in communication skills. *BMC Med Educ* 2020;20:57.
44 Eeckhout T, Gerits M, Bouquillon D, *et al*. Video training with peer feedback in real-time consultation: acceptability and feasibility in a general-practice setting. *Postgrad Med J* 2016;92:431–5.
45 General Medical Council. Remote consultations. Webpage, 2020. Available: https://www.gmc-uk.org/ethical-guidance/ethical-hub/remote-consultations [Accessed 12 Nov 2020].
46 Fisher R, Asaria M. How might COVID-19 affect the number of GPs available to see patients in England? 2020. Available: https://www.health.org.uk/publications/long-reads/how-might-covid-19-affect-the-number-of-gps-available-to-see-patients-in-england?utm_campaign=11732193_GP%20supply%20paper%20%20August%202020%20 %20Stakeholders%20and%20WILMINGTON&utm_medium=email&utm_source=The%20He alth%20Foundation&dm_i=4Y2,6ZGM9,HFJPOF,S4Q2S,1 [Accessed 27 Aug 2020].
47 Hertzog R, Johnson J, Smith J, *et al*. Diagnostic accuracy in primary care E-Visits: evaluation of a large integrated health care delivery system's experience. *Mayo Clin Proc* 2019;94:976–84.
48 Jacklin PB, Roberts JA, Wallace P, *et al*. Virtual outreach: economic evaluation of joint teleconsultations for patients referred by their general practitioner for a specialist opinion. *BMJ* 2003;327:84.
49 NHS England. Updated version. securing excellence in primary care (GP) digital services. The primary care (GP) digital services operating model 2019-2021, 2020. Available: https://www.england.nhs.uk/wp-content/uploads/2019/10/gp-it-operating-model-v4-sept-2019.pdf [Accessed 4 Sep 2020].
50 Huygens MWJ, Swinkels ICS, Verheij RA, *et al*. Understanding the use of email consultation in primary care using a retrospective observational study with data of Dutch electronic health records. *BMJ Open* 2018;8:e019233.
51 Mold F, Raleigh M, Alharbi NS, *et al*. The impact of patient online access to computerized medical records and services on type 2 diabetes: systematic review. *J Med Internet Res* 2018;20:e235.
52 Quinn LM, Davies MJ, Hadjiconstantinou M. Virtual consultations and the role of technology during the COVID-19 pandemic for people with type 2 diabetes: the UK perspective. *J Med Internet Res* 2020;22:e21609.
53 Health Education England. The future of primary care: creating teams for tomorrow. Report by the primary care workforce Commission, 2015. Available: https:// www.hee.nhs.uk/sites/default/files/documents/WES_The-future-of-primarycare.pdf [Accessed 27 Apr 2016].
54 Greenhalgh T, Wherton J, Papoutsi C, *et al*. Beyond adoption: a new framework for theorizing and evaluating Nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017;19:e367.
55 Dixon-Woods M, Cavers D, Agarwal S, *et al*. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Med Res Methodol* 2006;6:35.
56 Royal College of General Practitioners. Online consultations in general practice: the questions to ask. London: RCGP, 2018. Available: www.rcgp.org.uk/-/media/Files/Policy/A-Z-policy/2018/RCGP-online-consultations-questions-to-ask-march-2018.ashx?la=en [Accessed 1 Nov 2020].
57 NHS England. Using online consultations in primary care: implementation toolkit. London: NHS England, 2020. Available: www.england.nhs.uk/wp-content/uploads/2020/01/online-consultations-implementation-toolkit-v1.1-updated.pdf [Accessed 1 Nov 2020].
58 Department of Health and Social Care. Coronavirus (COVID-19): The future of healthcare. Matt Hancock's speech about future of healthcare (lesson 3), 2020. Available: https://www.gov.uk/government/speeches/the-future-of-healthcare [Accessed 15 Aug 2020].
59 NHS Long-Term Plan. Digital-first primary care by 2023-24. Available: https://www.england.nhs.uk/gp/digital-first-primary-care/ [Accessed 4 Sept 2020].
60 Gray DP, Freeman G, Johns C, *et al*. Covid 19: a fork in the road for general practice. *BMJ* 2020;370:m3709.
61 Popeski N, McKeen C, Khokhar B, *et al*. Perceived barriers to and facilitators of patient-to-provider e-mail in the management of diabetes care. *Can J Diabetes* 2015;39:478–83.

BMJ Health & Care Informatics

Assessment of automated clinical trial recruitment and enrolment using patient-facing technology

Naomi S Bardach ,[1] Regina Lam,[2] Carolyn B Jasik[3,4]

To cite: Bardach NS, Lam R, Jasik CB. Assessment of automated clinical trial recruitment and enrolment using patient-facing technology. *BMJ Health Care Inform* 2021;**28**:e100076. doi:10.1136/bmjhci-2019-100076

Received 31 May 2019
Revised 13 March 2020
Accepted 12 June 2020

[1]Pediatrics, University of California San Francisco, San Francisco, California, USA
[2]School of Medicine, Mayo Clinic School of Medicine, Rochester, Minnesota, USA
[3]Research Unit, Omada Health, San Francisco, California, USA
[4]Pediatrics, University of California San Francisco School of Medicine, San Francisco, California, USA

Correspondence to
Dr Naomi S Bardach;
naomi.bardach@ucsf.edu

ABSTRACT

Objective Interactive patient care systems (IPCS) at the bedside are becoming increasingly common, but evidence is limited as to their potential for innovative clinical trial implementation. The objective of this study was to test the hypothesis that the IPCS could feasibly be used to automate recruitment and enrolment for a clinical trial.

Methods In medical-surgical units, we used the IPCS to randomise, recruit and consent eligible subjects. For participants not interacting with IPCS study materials within 48 hours, study staff-initiated recruitment in-person. Eligible study population included all caregivers and any patients >6 years old admitted to medical-surgical units and oncology units September 2015 to January 2016. Outcomes: randomisation assessed using between-group comparisons of patient characteristics; recruitment success assessed by rates of consent; paperless implementation using successful acquisition of electronic signature and email address. We used χ^2 analysis to assess success of randomisation and recruitment.

Results Randomisation was successful (n=1012 randomised, p>0.05 for all between-group comparisons). For the subset of eligible, randomised patients who were recruited, IPCS-only recruitment (consented: 2.4% of n=213) was less successful than in-person recruitment (61.4% of n=87 eligible recruited, p<0.001). For those consenting (n=61), 96.7% provided an electronic signature and 68.9% provided email addresses.

Conclusions Our results suggest that as a tool at the bedside, the IPCS offers key efficiencies for study implementation, including randomisation and collecting e-consent and contact information, but does not offer recruitment efficiencies. Further research could assess the value that interactive technologies bring to recruitment when paired with in-person efforts, potentially focusing on more intensive user-interface testing for recruitment materials.

Trial registration number NCT02491190.

Summary

What is already known?

- Interactive patient care systems are becoming more popular and are used for a variety of patient-oriented interactions, including not only entertainment, but also delivery of patient education videos, survey questions, food ordering and communication with providers.

What does this paper add?

- This paper reports on the use of an interactive patient care system to automate clinical trial tasks, focusing particularly on success of randomisation, recruitment and collection of consent and contact information. We found that the system was successful for randomisation and collection of consent and contact information, and had poor performance as a stand-alone recruitment method.

INTRODUCTION

Consumer-facing health technology has the potential to revolutionise care, re-orienting traditional provider-centric models of care delivery.[1 2] In the inpatient setting, interactive patient care systems (IPCS) at the bedside such as GetWell Network, myStation and OneView provide personal health information, educational materials and patient engagement features to optimise patient–provider communication, in addition to on-demand entertainment.[3–6] These systems have a substantial national presence, implemented in almost 40 000 beds in 2013.[7 8] IPCS adoption will likely increase as hospitals seek to meet federal and local demands for deeper and more meaningful patient engagement.[9]

With the increasing pressure to integrate point-of-care patient engagement technologies into the clinical workflow,[1] our need for data on their usefulness is becoming more urgent.[10] To gather this data, it is possible that we may be able to leverage the technologies to automate trial implementation, realising efficiencies over the traditional in-person research staffing. Though limited, prior research suggests that technology platforms can streamline research processes and perform as well as, if not better than, paper methods.[11] For instance, tablet-delivered digital multimedia study materials have improved understanding of clinical trials during paediatric patient recruitment[12] and electronic health record platforms have been used for recruitment.[13] However, no study to

our knowledge has examined the use of IPCS to automate the multiple aspects of a clinical study.

In this report, we present data on our experience with a pragmatic test of IPCS clinical trial implementation. The parent trial assessed the effect of a patient-facing and family-facing educational video on patient experience metrics. For implementation of the parent trial, we tested the IPCS for patient identification, randomisation, recruitment and consenting, with a secondary plan for in-person recruitment and consenting for patients who did not interact with the IPCS study materials. The objective of this study was to assess feasibility of using the IPCS to automate aspects of a randomised clinical trial, including: (1) identifying and (2) randomising eligible patients and (3) recruiting and (4) consenting participants, including gathering electronic signatures and disseminating consent forms via email.

METHODS

Setting

The participating hospital opened in February 2015 with an IPCS at 183 beds at the new site. We worked with the IPCS vendor, OneView Healthcare, to plan the workflow (figure 1) and features needed for a study assessing the effect of a patient engagement video.

Technology

The IPCS had several features already in place that supported the study implementation: a working interface with the hospital's admission, discharge and transfer system; a patient education portal, available on the home page, which displayed a flag for assigned education until the education had been viewed; the ability, within the portal, to serve videos or weblinks; the ability to auto-assign patient education based on patient criteria.

IPCS features created for this study included: (1) automated identification of eligible participants using complex criteria (admission date, no prior admission during study period, hospital unit); (2) automated randomisation of eligible participants to intervention or control; (3) delivery of study recruitment and consent materials to eligible participants, with passing of a patient identifier into the consenting forms; (4) browser adjustments to enable web-collected e-signatures.

Study population

Eligible population: all caregivers and any patients >6 years old admitted to medical-surgical units and oncology units 16 September 2015 to 9 January 2016. Exclusion criteria: no parent or guardian available, non-English speaking, being in foster care, prior admission during the study period.

Data

Data from the IPCS: user engagement with the IPCS standard features, engagement with the study-specific materials, length of stay and number of admissions.

Data from the interactive study materials and from the consent process were collected and managed using REDCap electronic data, a secure web-based data capture tool, hosted at the University of California San Francisco.[14] Study staff recorded recruitment attempts and reasons for exclusion for patients they approached.

Automated clinical trial implementation features

Patient identification

The IPCS assigned study information in the education portal for all eligible patients. For these patients, the IPCS home page displayed a flag in the education portal until the study information website was opened.

Randomisation

The IPCS randomised patients 1:1 to the educational video intervention or to control. All eligible patients were randomised at admission, due to the technological limitation of communicating consent information from REDCap to the IPCS. Once randomised, the fidelity to the

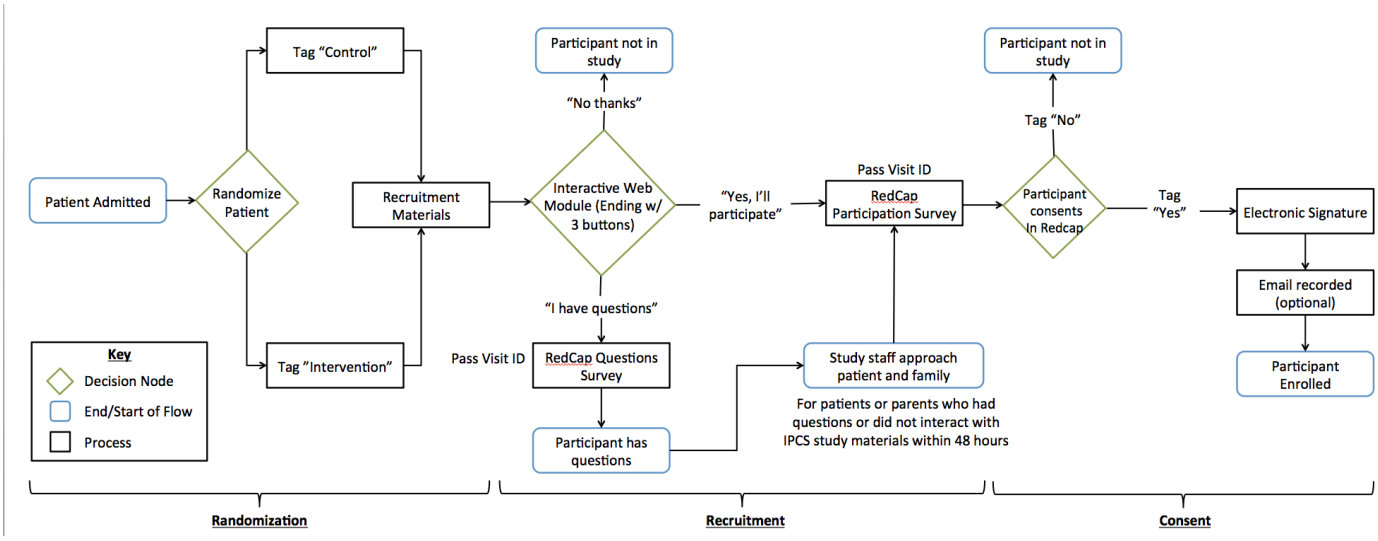

Figure 1 Workflow for automated clinical trial implementation.

protocol was driven by the IPCS programming that made the intervention video available to those randomised to the intervention. Hence, fidelity of delivery of availability of the video was 100% (verified by random period checks of individual patient video assignments during the course of the study).

Recruitment

The study information website launched an interactive slide-deck presentation, made using Articulate software, describing the study. At the end of the presentation, viewers were asked to click one of three options: opt out, continue to consent or ask questions.

For patients admitted for >48 hours who had not interacted with IPCS recruitment materials, and who were available (eg, not off-unit, parent or guardian available, not busy with clinical staff), study staff used an in-person recruitment protocol during weekdays using a standard consenting process. Participants who received in-person recruitment still reviewed the IPCS recruitment materials with the in-person facilitator. Recruiting staff were blinded regarding allocation.

Consent and e-signature

Clicking on one of the options at the end of the interactive material (I'm interested, no thanks, or I have questions) opened one of three web-based REDCap surveys, which recorded the patient ID and the response. Those who were interested were then screened and consented using the RedCap survey. Consent forms were available for parents or age-eligible children, according to branching logic. We used the e-signature feature within REDCap and the IPCS bedside tablet interface to gather signatures for parents agreeing to release medical records (figure 2). The REDCap survey also optionally collected caregiver email addresses to send consent copies electronically.

Measures

Randomisation success

In order to assess for adequate randomisation, we compared study groups on available variables from the IPCS: number of interactions with the IPCS, interaction with study recruitment materials, average length of stay and mean number of admissions. We chose the average length of stay and the mean number of admissions because we hypothesised that they may be positively associated with patient or family member interactions with the study recruitment materials and with the intervention video, and therefore wanted to verify that they were balanced between the groups.

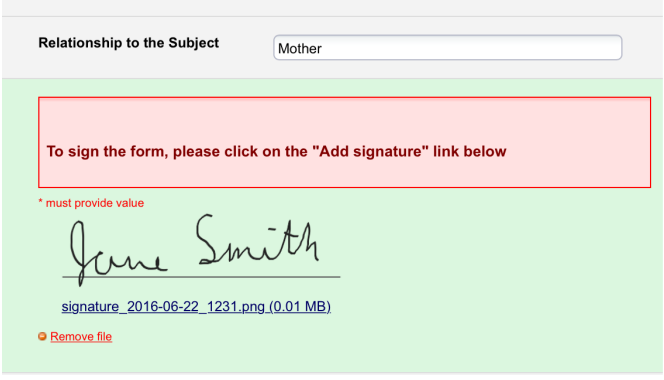

Figure 2 E-signature interface.

Recruitment success

We assessed the success of recruitment modality (IPCS vs IPCS with in-person facilitation) by comparing enrolment rates by modality.

Feasibility of electronic signature and email

Analysis

Statistical analysis focused on the success of automated randomisation and recruitment.

Binary outcomes were compared using χ^2 analysis, or Fisher's exact for cell sizes <10. Analyses were conducted using Stata V.13.

RESULTS

Randomisation

There were 1012 patients admitted to the eligible units during the study period, with 502 randomised to intervention (patient education video and recruitment materials) and 510 randomised to control (recruitment materials only). The randomisation was adequate, with no statistically significant differences between groups in characteristics from IPCS data (table 1).

Recruitment, consent and e-signature collection

Figure 3 depicts a study recruitment flow diagram. Of 1012 patients, 21.0% (n=213) opened the study materials only through the IPCS patient education portal. Of those, 8.5% (n=18) completed the interactive materials, with five consenting to participate (29.4% of eligible patients completing interactive materials; 2.4% of those opening materials). Of those who did not open the materials in the IPCS, who were subsequently recruited in-person (n=176), 90 were ineligible due to being non-English speakers (n=58) or due to not having a guardian present (n=32). Of those recruited in person who were eligible (n=87), a larger proportion consented to participate than the patients only opening study materials through the patient education portal (64.4% vs 2.4%, p<0.001 for comparison; figure 3).

Of consented participants (n=59), 96.6% (n=57) gave an electronic signature to release medical records and 71.2% (n=42) of participating parents opted to give their email address.

DISCUSSION

This study offers the first look at the potential of IPCS for supporting clinical trial implementation via automated methods. An IPCS system was successful in identifying eligible subjects, randomisation, collecting electronic signatures for

Table 1 Randomisation success assessed through comparison of parent trial group participants

	Parent trial intervention (N=502)	Parent trial control (N=510)	P value
Ever interacted with IPCS, n (%)*	492 (98.0%)	501 (98.2%)	0.79
Consented to research, n (%)	24 (4.8%)	28 (5.5%)	0.84
Length of stay (mean days, 95% CI)	4.49 (3.82 to 5.16)	5.18 (4.27 to 6.10)	0.23
Total admissions (mean number, 95% CI)	1.32 (1.23 to 1.40)	1.25 (1.19 to 1.31)	0.16

We report on data only from the IPCS, which do not include demographics, as not all patients consented to releasing medical record information.
*The IPCS is used for other tasks (eg, ordering food, watching movies), so these measures whether patients ever interacted with the IPCS during their hospital stay, for comparison of levels of IPCS engagement between the intervention and control groups.
IPCS, interactive patient care system.

medical records release and capturing email addresses. Study enrolment via a stand-alone IPCS process was successful for only a fraction of potential subjects. Failures were due to either a complete lack of or very minimal engagement with IPCS-assigned materials. Ultimately a staff member was required to complete consent and study enrolment.

Prior studies have assessed the success of using multimedia interactive materials to improve the informed consent process.[15–21] The evidence is mixed regarding their success, with most studies finding improved comprehension of study materials,[15–17 19 21] a preference for multimedia materials,[15 20 21] generally increased time spent on informed consent[20] and mixed effects on patient enrolment and retention.[17 18 21] While these studies demonstrate the potential benefit of interactive systems for informed consent, our study expanded the use of the system to identify eligible patients, randomise them, and recruit and consent participants.

We found that the IPCS identified eligible patients and adequately randomised them, as illustrated by the similarities across groups in recruitment and consent rates as well as across measured characteristics (table 1). This suggests that randomising IPCS features to assess their effects is feasible and could be considered for future studies. If implemented, blinding study staff to allocation assignment is necessary to avoid potential post-randomisation selection bias.[22]

Our results suggest that the IPCS alone is not sufficient for patient recruitment. We saw IPCS recruitment failures at three points: (1) participant opening assigned study materials; (2) participant completion of interactive study materials and (3) participant enrolment after viewing the materials. Failures at the first two stages imply that enhanced patient engagement features could potentially improve study recruitment success. Enhancements to improve interaction with materials might include more noticeable indicators of the presence of study materials (eg, interrupting of programming or more prominent visual notifications to the user (eg, blinking notification, banner on screen once a day or at routine intervals until study completion, text message reminders, etc). Changing the study materials (we used an interactive slide deck but there are other potential modalities such as whiteboard animation or including videos of patient participants) may have improved participant completion of interactive study materials. Testing these other options was beyond the scope of the study, would best be done with more intensive user-interface testing, and could be the focus of future work. Complementary qualitative research to explore barriers and facilitators to success would provide greater detail and context-specific information to explain successes and failures.

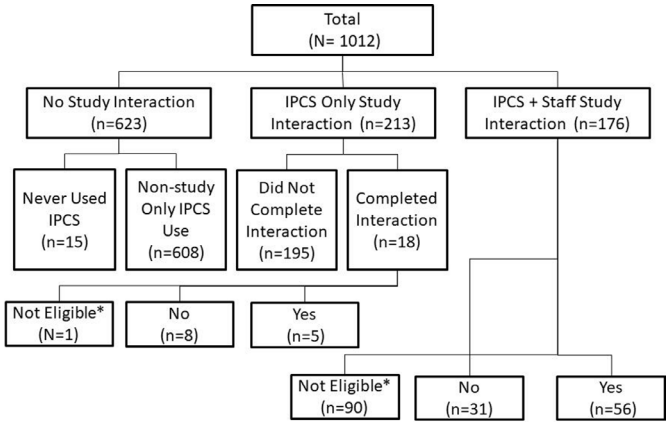

Figure 3 Recruitment and consent flow chart. IPCS, interactive patient care systems.

Our findings also suggest that when leveraging an IPCS technology in trial implementation, staff for in-patient recruitment should not be eliminated. In contrast, efficiencies from IPCS may be realised in identifying eligible patients, randomisation and data collection. For example, patients who were readmitted were excluded automatically from the study, and randomisation was built into the technology, eliminating that step. The IPCS was an efficient platform for gathering electronic consent, with 96% of parents providing e-signatures and 71% providing email addresses for optional follow-up. Finally, the electronic consenting process eliminated the potential for lost paper forms, decreasing the risk of privacy loss. The electronic consent branching logic allowed for a tailored and shortened consent and data validation decreased errors. The assessment of potential cost-effectiveness of using the automated functions for clinical research implementation could be the focus of further study.

Limitations

While our use of the IPCS to give study information and enrol participants yielded low recruitment, the causes of low recruitment may not be directly attributable to the IPCS

technology. IPCS study recruitment materials that were less easy to ignore (we used a passive reminder) may have led to different IPCS recruitment rates. Different interactive materials may have yielded different effects on recruitment. Generalisability should be understood within the context of specific IPCS software and implementation. Our Oneview Healthcare software included a randomiser function; others may not. We did not have sociodemographic data to explore whether IPCS interactions and recruitment success might have differed by patient characteristics. Finally, qualitative observations and more intensive user interface testing in the future can give greater understanding about how to better adapt IPCS patient engagement features for recruitment.

CONCLUSION

Interactive patient care systems are innovative new tools with the potential for supporting inpatient research. This study illustrates that technology, while potentially adding value in the healthcare context, does not inevitably replace human interactions. Our results suggest that as an electronic communication tool at the bedside, the IPCS offers key efficiencies for study implementation, including patient identification, randomisation and collecting e-consent and study contact information, but that it is limited in its ability to inform and recruit. Further research assessing whether patient engagement enhancements to the IPCS improve recruitment rates will better illuminate the potential value that interactive technologies bring when paired with in-person efforts.

Twitter Naomi S Bardach @naomibardach

Acknowledgements Seth Bokser from Oneview HealthCare for assistance with implementation, Stephen McGowan and Philip Urrea from OneView HealthCare, for work on the development of the features in the interactive patient care system; Omada Health for allowing Dr Jasik to participate in this project. Drs Alice Ainsworth and Nikki Suarez, for their work on the development of the patient education video.

Contributors NSB conceptualised, designed and carried out the study, drafted the initial manuscript, and approved the final manuscript as submitted. RL drafted the initial manuscript and approved the final manuscript as submitted. CBJ contributed to design and conduct of the study, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Funding This project was supported by the National Centre for Advancing Translational Sciences, National Institutes of Health, through UCSF-CTSI Grant Number UL1 TR000004. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. NSB was supported by the National Institute of Child Health and Human Development (K23HD065836).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The Human Research Protection Program and Institutional Review Board at the University of California San Francisco approved this study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iD

Naomi S Bardach http://orcid.org/0000-0002-5178-0719

REFERENCES

1 Singh A, Rhee KE, Brennan JJ, *et al*. Who's my doctor? using an electronic tool to improve team member identification on an inpatient pediatrics team. *Hosp Pediatr* 2016;6:157–65.
2 Tang C, Lorenzi N, Harle CA, *et al*. Interactive systems for patient-centered care to enhance patient engagement. *J Am Med Inform Assoc* 2016;23:2–4.
3 Interactive Patient System, Cerner, 2016. Available: www.cerner.com/solutions/medical_devices/interactive_patient_system/ [Accessed April 6, 2016].
4 GetWell Network. Interactive experience for patient and family engagement, 2016. Available: www.getwellnetwork.com/solutions/patient-experience/pediatric [Accessed 6 Apr 2016].
5 OneView. Revolutionizing the patient experience through innovative software for healthcare facilities, 2016. Available: www.oneviewhealthcare.com/about-us/ [Accessed 6 Apr 2016].
6 Collins SA, Rozenblum R, Leung WY, *et al*. Acute care patient portals: a qualitative study of stakeholder perspectives on current practices. *J Am Med Inform Assoc* 2017;24:e9–17.
7 FreshTracks Capital. GetWell Acquired By Welsh, Carson, Anderson & Stowe, 2013. Available: www.freshtrackscap.com/getwell-acquired-by-welsh-carson-anderson-stowe [Accessed 6 Apr 2016].
8 Wicklund E. Hospitals find new benefits in 'Interactive patient systems': Healthcare IT News, 2011. Available: www.healthcareitnews.com/news/hospitals-find-new-benefits- [Accessed 6 Apr 2016].
9 HealthIT.gov. Step 5: achieve meaningful use stage 2, 2014. Available: www.healthit.gov/providers-professionals/step-5-achieve-meaningful-use-stage-2 [Accessed 6 Apr 2016].
10 Rudin RS, Bates DW, MacRae C. Accelerating innovation in health it. *N Engl J Med* 2016;375:815–7.
11 Chalil Madathil K, Koikkara R, Obeid J, *et al*. An investigation of the efficacy of electronic consenting interfaces of research permissions management system in a hospital setting. *Int J Med Inform* 2013;82:854–63.
12 Tait AR, Voepel-Lewis T, Levine R. Using digital multimedia to improve parents' and children's understanding of clinical trials. *Arch Dis Child* 2015;100:589–93.
13 Mastellos N, Bliźniuk G, Czopnik D, *et al*. Feasibility and acceptability of transform to improve clinical trial recruitment in primary care. *Fam Pract* 2016;33:186–91.
14 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
15 Jimison HB, Sher PP, Appleyard R, *et al*. The use of multimedia in the informed consent process. *J Am Med Inform Assoc* 1998;5:245–56.
16 Rhodes KV, Lauderdale DS, Stocking CB, *et al*. Better health while you wait: a controlled trial of a computer-based intervention for screening and health promotion in the emergency department. *Ann Emerg Med* 2001;37:284–91.
17 Flory J, Emanuel E. Interventions to improve research participants' understanding in informed consent for research: a systematic review. *JAMA* 2004;292:1593–601.
18 Gesualdo P, Ide L, Rewers M, *et al*. Effectiveness of an informational video method to improve enrollment and retention of a pediatric cohort. *Contemp Clin Trials* 2012;33:273–8.
19 O'Lonergan TA, Forster-Harwood JE. Novel approach to parental permission and child assent for research: improving comprehension. *Pediatrics* 2011;127:917–24.
20 Schlechtweg PM, Hammon M, Giese D, *et al*. iPad-based patient briefing for radiological examinations-a clinical trial. *J Digit Imaging* 2014;27:479–85.
21 Synnot A, Ryan R, Prictor M, *et al*. Audio-visual presentation of information for informed consent for participation in clinical trials. *Cochrane Database Syst Rev* 2014;5:CD003717.
22 Viera AJ, Bangdiwala SI. Eliminating bias in randomized controlled trials: importance of allocation concealment and masking. *Fam Med* 2007;39:132–7.

BMJ Health & Care Informatics

Artificial intelligence-based prediction of transfusion in the intensive care unit in patients with gastrointestinal bleeding

Riccardo Levi ,[1] Francesco Carli,[2] Aldo Robles Arévalo,[3] Yuksel Altinel,[4] Daniel J Stein,[5] Matteo Maria Naldini,[6] Federica Grassi,[7] Andrea Zanoni,[8] Stan Finkelstein,[9,10] Susana M Vieira,[3] João Sousa,[3] Riccardo Barbieri,[1] Leo Anthony Celi [11,12]

To cite: Levi R, Carli F, Arévalo AR, *et al*. Artificial intelligence-based prediction of transfusion in the intensive care unit in patients with gastrointestinal bleeding. *BMJ Health Care Inform* 2021;**28**:e100245. doi:10.1136/bmjhci-2020-100245

Received 24 September 2020
Revised 26 October 2020
Accepted 27 November 2020

For numbered affiliations see end of article.

Correspondence to
Mr Riccardo Levi;
riccardo.levi@mail.polimi.it

ABSTRACT

Objective Gastrointestinal (GI) bleeding commonly requires intensive care unit (ICU) in cases of potentialhaemodynamiccompromise or likely urgent intervention. However, manypatientsadmitted to the ICU stop bleeding and do not require further intervention, including blood transfusion. The present work proposes an artificial intelligence (AI) solution for the prediction of rebleeding in patients with GI bleeding admitted to ICU.

Methods A machine learning algorithm was trained and tested using two publicly available ICU databases, the Medical Information Mart for Intensive Care V.1.4 database and eICU Collaborative Research Database using freedom from transfusion as a proxy for patients who potentially did not require ICU-level care. Multiple initial observation time frames were explored using readily available data including labs, demographics and clinical parameters for a total of 20 covariates.

Results The optimal model used a 5-hour observation period to achieve an area under the curve of the receiving operating curve (ROC-AUC) of greater than 0.80. The model was robust when tested against both ICU databases with a similar ROC-AUC for all.

Conclusions The potential disruptive impact of AI in healthcare innovation is acknowledge, but awareness of AI-related risk on healthcare applications and current limitations should be considered before implementation and deployment. The proposed algorithm is not meant to replace but to inform clinical decision making. Prospective clinical trial validation as a triage tool is warranted.

Summary

What is already known?

- Gastrointestinal bleeding is a severe event that requires admission to the ICU.
- Many patients in the ICU for gastrointestinal bleeding undergo only increased monitoring without intervention.
- ICU stay is associated with increased cost and morbidity.

What does this paper add?

- An algorithmic approach using artificial intelligence on readily available electronic data can accurately predict ICU transfusion need.
- Using this approach to identify patients at low risk for ongoing bleeding and transfusion could be validated prospectively to identify patients who may not require ICU-level care.

INTRODUCTION

Gastrointestinal (GI) haemorrhage is a common condition that frequently requires hospitalisation, often in the intensive care unit (ICU)[1] with considerable associated morbidity. In particular, ICU admission is associated with increased costs and a greater rate of complications and poor outcomes compared with ward admission.[2–4] Some patients are initially admitted to the ICU for haemodynamic instability but stabilise without further intervention and are discharged to the ward the following day.

Previous instruments, such as the Rockall or the Blatchford score[5] have been applied to triage patients based on the likelihood of mortality, recurrent/ongoing bleeding, need for hospitalisation and requirement for endoscopic intervention. However, these models are validated only for upper GI bleeding with a focus on endoscopic intervention and mortality and do not assist in informing level of monitoring for hospitalised patients. Currently, there is no model to assist in triaging patients with GI bleeding including those with an undifferentiated source to an appropriate acuity of care.

We identified the need for blood transfusion as a surrogate for persistent bleeding. Previous prospective studies have shown that up to half of patients with GI bleeding may not require transfusion.[6] We used an ICU database to train a prediction model but

focused on the first few hours on arrival as a proxy of the patient's state in the emergency department.

The use of artificial intelligence (AI) represents an opportunity for more effective and efficient care delivery by predicting disease trajectory and complications.[7–12] Previous work in GI bleeding has used methods such as artificial neural networks,[13 14] support vector machines[13] to predict the need for intervention; and fuzzy models[15] to identify which lab test is likely to contribute information gain and influence clinical management of patients with GI bleeding in the ICU. This study focused on using machine learning to predict transfusion to better identify those patients who continue to bleed.

METHODS

This study is reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology statement.[16]

Database description

Data were collected from the Medical Information Mart for Intensive Care-III (MIMIC-III) V.1.4[17] and in the eICU Collaborative Research Database V.2.0 (eICU-CRD).[18] Both databases contain information from patients admitted to the ICU. The MIMIC-III database collects detailed haemodynamic and clinical parameters from all ICU patients admitted to a single major academic medical centre between 2008 and 2014, whereas the eICU-CRD is a multicentre database with high granularity data for over 200 000 admissions to ICUs monitored by an eICU[19] across the USA.

Ethical approval

Both databases are previously de-identified and have been reviewed by the institutional review boards (IRB) of their hosting organisations and determined to be exempt from subsequent IRB.

Definition of outcome

The outcome of this study is ongoing GI bleeding after admission to the ICU. Since this outcome variable is not encoded, blood transfusions were used as surrogate marker.

Software

Models were developed in Python V.3.7 using data science packages including pandas V.0.25.3 (data wrangling),[20] NumPy V.1.17.5 (computations),[21] SciPy V.1.4.1 (hypothesis testing),[22] Scikit-learn V.0.22.1 (modelling)[23] and Hyperopt V.0.2.3 (hyperparameter optimisation).[24]

Data preparation

We included non-pregnant adult patients (≥18 years old) admitted to the ICU and diagnosed with GI bleeding based on the International Classification of Diseases (ICD-9) codes (see table A1, online supplemental digital content 1,). For patients with multiple ICU admissions within a single hospitalisation event, only the first ICU stay was considered. The inclusion criteria for each database are further detailed in figure 1.

Missing records were imputed with the last observation available carried forward. Patients missing their first value were imputed with the intra-subject median. In order to take into account the dynamics of the observed features within the training window (eg, increasing, decreasing trends), we adopted a feature engineering approach (see text, online supplemental digital content 2). Also, non-normally distributed features (skewness >3) were log-transformed[25] in order to obtain a normal distribution for improved model performance.

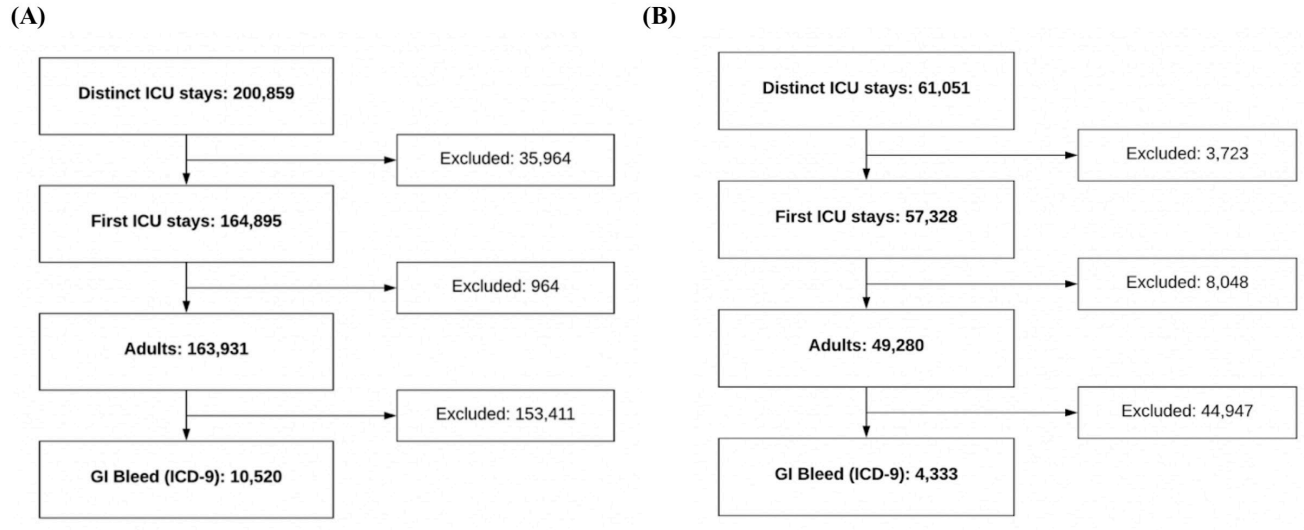

Figure 1 Inclusion criteria for the cohort extracted from the (A) eICU-CRD and (B) MIMIC-III. eICU-CRD, eICU Collaborative Research Database; ICD-9, International Classification of Diseases-9; ICU, intensive care unit; GI, gastrointestinal; MIMIC-III, Medical Information Mart for Intensive Care-III.

Table 1 List of covariates, the output variable and demographic information for each cohort. Continuous variables are stated as mean (IQR), otherwise are the number of occurrences. only a subset of these variables (selected by recursive feature elimination procedure) enters in the final models.

	MIMIC-III (n=4314)	eICU-CRD (n=10 306)
Demographics		
Age at admission (years)	83.5 (56–81)	76.7 (56–79)
Gender (n)		
Male	2491	5927
Female	1823	4379
Output variable (transfusion)		
Transfused patients (n, % wrt total number of patients)	2077 (48.15%)	2712 (26.31%)
Covariates		
Heart rate (bpm)	92.9 (79.0–105.7)	94.0 (79.9–106.5)
Mean blood pressure (mm Hg)	78.9 (68.5–87.8)	78.4 (67.6–87.5)
Systolic blood pressure (mm Hg)	114.5 (99.0–129.0)	108.1 (93–121)
Diastolic blood pressure (mm Hg)	60.3 (54.7–65.2)	62.6 (56.0–68.2)
Respiratory rate (breaths/min)	21.2 (18.0–24.0)	21.9 (17.8–24.4)
Haematocrit (%)	28.4 (23.8–32.6)	26.5 (20.7–31.6)
Haemoglobin (g/L)	97 (80–112)	87 (67–104)
White blood cell ($\times 10^9$/L)	11.8 (7.2–14.1)	11.7 (7.4–14.4)
Platelet ($\times 10^9$/L)	227.5 (137.0–286.0)	207 (129.0–263.0)
Creatinine (mg/dL)	1.79 (0.85–1.88)	1.73 (0.80–1.90)
Blood urea nitrogen (mg/dL)	39.5 (19.0–51.0)	39.2 (19.0–51.0)
Potassium (mEq/L)	4.34 (3.80–4.70)	4.38 (3.80–4.80)
Bicarbonate (mEq/L)	22.6 (20.0–26.0)	22.7 (20.0–26.0)
Amount blood transfused (mL)	601.0 (375.0–750.0)	571.9 (324.0–700.0)
Glucose (mg/dL)	160.2 (106.0–174.0)	153.2 (105.0–176.0)
Albumin (g/dL)	3.17 (3.2–3.2)	2.96 (2.8–3.1)
Temperature (°C)	36.3 (36.0–36.7)	36.4 (36.4–36.5)
Partial thromboplastin time (s)	37.3 (26.1–37.9)	35.3 (26.0–37.0)

eICU-CRD, eICU Collaborative Research Database; ICU, intensive care unit; MIMIC-III, Medical Information Mart for Intensive Care-III.

Feature selection has been performed by recursively discarding features that do not reduce accuracy performance when eliminated. This procedure is called recursive feature elimination (RFE), a method used to remove non-predictive covariates with a greedy approach[26] (see text, online supplemental digital content 3). Final input datasets gather 4333 first ICU admissions from the MIMIC-III database and 10520 first ICU admissions from the eICU-CRD along with 20 covariates. Input variables include several laboratory analyses and demographic information that are available in each database. Detailed information of these features is described in table 1.

Prediction time windows

Several time windows were assessed for data extraction of the training/testing data and the data for the output variable (blood transfusion) that was predicted. Four different time windows starting from ICU admission (hour 0) were evaluated: training time from 0 to 3 hours/prediction time 4–24 hours, training time 0–4 hours/prediction time 5–24 hours, training time 0–5 hours/prediction time 6–24 hours, training time 0–6 hours/prediction time 7–24 hours. The training timeframe contains the covariates recorded during that time frame for each ICU stay. All training time windows include information recorded prior to the ICU admission (up to −1 hour). The prediction time window is when the surrogate variable (blood transfusion) was recorded (see figure 2).

This analysis helped us to find the optimal training/prediction time windows. The selected time windows were those that achieved the best predictive performance. In addition to that, the best training time window is the one that gathered the highest amount of data before a blood transfusion. Except from that, there is no other contextual detail that was considered during this analysis.

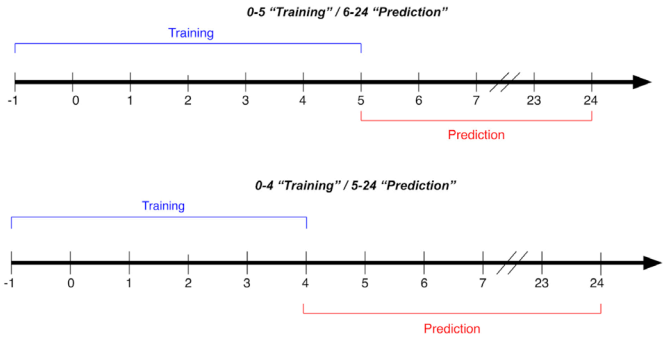

Figure 2 Graphical schema of the time windows.

Training and testing partitions

Several training/testing partitions and strategies were designed in order to fully exploit the information contained in both datasets. Specifically, both datasets are randomly divided into a test (25% of records) and training set (75% of records). A model is fitted on each of the training sets and on a combination of the two. All training subsets were split to perform 10-fold cross validation and to optimise model's hyperparameters. The testing subsets had data that were not used for training/validation.

Three different training sets were considered: (1) including MIMIC-III data only (n=3235); (2) including eICU-CRD data only (n=7729) and (3) a training set composed by 29.17% of MIMIC-III and 70.83% of eICU-CRD (n=10 964). The performance of the models is then gauged on both the test sets, allowing for an external validation of the classifiers for a total of three models per each considered time window:

1. Train on MIMIC-III, internal validation on MIMIC, external validation on eICU-CRD.
2. Train on eICU-CRD, internal validation on eICU-CRD, external validation on MIMIC-III.
3. Train on MIMIC-III and eICU-CRD, internal validation on MIMIC-III and eICU-CRD.

Predictive models

In order to improve the performance of individual machine learning models, the final classifier is determined as an ensemble of machine learning models combined together. To select the models for this ensemble, we assessed several classifiers. Hyperparameter tuning was performed through Bayesian optimisation[27] with a stratified 10-fold cross validation, where class imbalance is taken into account in the parameters of the models. This tuning is carried out with a customised loss function that takes into account accuracy and F1 score (see text, online supplemental digital content 4). This delivers a model based (and hence non arbitrary) procedure to find cut-off-thresholds that optimise jointly the accuracy, specificity and sensitivity of the model. By specifying the weights of F1 score and accuracy inside the custom loss function the model could be oriented to avoid false negative predictions (higher F1 score and recall) with a high accuracy. However, since the model also provides the probability that a patient will bleed the physician could in principle perform standard sensitivity–specificity trade-off decisions.

Given that eICU-CRD exhibits target imbalance (26% transfused patients against 74% non-transfused patients) classifiers trained on this dataset are imbalance-aware in order not to skew predictions towards the majority class (ie, predicting all patients as low risk patients, which is not desirable).

Permutation feature importance[28] of the five most important covariates is estimated for each model. Moreover, the partial dependence function[29] function of the outcome with respect to the most important variable is estimated (see text, online supplemental digital content 5).

In order to assess the goodness of the classifier during testing, we estimated the model's accuracy, sensitivity (recall or true classification positive rate), specificity (true negative classification rate) and area under the curve of the receiving operating curve (ROC-AUC).

To conclude, models are calibrated through Platt's scaling[30 31] to obtain reliable probability estimates. The effects of the calibration can be diagnosed visually with the calibration curves (see text, online supplemental digital content 6).

RESULTS

The best results are achieved when the models are trained on the MIMIC-III dataset (see table A2, online supplemental digital content 7), and the lowest values are observed in the models trained on the eICU-CRD data (see table A3, online supplemental digital content 8). When both datasets are merged (see table A4, online supplemental digital content 9), the performance does not improve considerably, but we can observe a significant improvement in terms of sensitivity. Of note, the sensitivity obtained in the models trained with MIMIC-III is the highest among all other models; which indicates that it is better to detect true positive cases or patients that would require transfusion.

It is also interesting to highlight that the models trained on MIMIC-III (see table A2, online supplemental digital content 7) have a greater discriminative power on the eICU-CRD testing set than the models trained only on the eICU-CRD data (see table A3, online supplemental digital content 8) and even if these are tested on the same database. Thus, a model trained on MIMIC-III is capable of generalising better to patients that the model does not train on.

These observations could be explained by the fact the MIMIC-III input dataset is not skewed (48.14% of the entries required transfusion) as the input dataset from the eICU-CRD (26.31% of the entries required transfusion). This imbalance could skew the model predictions towards the majority class (the most frequent label in the population) that are the patients that did not bleed (not required transfusion).

Table 2 Results for the time window composed by the pair training time of 0–5 hours/prediction time 6–24 hours

	Testing sets							
	ROC-AUC		Accuracy		Specificity		Sensitivity	
Training sets	MIMIC-III	eICU-CRD	MIMIC-III	eICU-CRD	MIMIC-III	eICU-CRD	MIMIC-III	eICU-CRD
MIMIC-III	0.8141	0.7634	0.7470	0.5021	0.6482	0.3502	0.8536	0.9277
eICU-CRD	0.8017	0.7858	0.7470	0.7060	0.7982	0.6872	0.6917	0.7581
MIMIC-III+eICU-CRD	0.8035	0.7908	0.7488	0.6884	0.7143	0.6535	0.7861	0.7861

eICU-CRD, eICU Collaborative Research Database; MIMIC-III, Medical Information Mart for Intensive Care-III; ROC-AUC, area under the curve of the receiving operating curve.

To avoid these misclassifications, the decision threshold was tuned during the optimisation procedure. In case the models were optimised only in terms of accuracy, it could have pushed the model to predict the majority class (non-transfused). By using the customised loss function, it was forced to jointly maximise precision and the recall of the final model notwithstanding the accuracy.

Looking at the results reported in table 2, online supplemental tables A5–A7 (see tables A, online supplemental digital content 10–12) we notice that the performances of all the time windows are satisfying and the overall best ones are obtained when the training phase is performed with data collected in the time window 0–5 hours and the prediction time window is from 6 to 24. Hence, in the following, we will mainly focus on this subdivision.

The models achieve greater ROC-AUC values when they are tested on the MIMIC-III dataset (>0.80) compared with the models tested on the eICU-CRD (0.76–0.79) as shown in table 2. Only accuracy and specificity improve when the models are trained in the eICU-CRD, but no improvement is detected in terms of sensitivity. The highest true classification positive rate is achieved in the models trained on the MIMIC-III, a critical metric being that it indicates how good are the models to predict the need of transfusions (true positives). We remark that this behaviour was expected since the eICU-CRD dataset has a larger variety of patients and hospitals than on the MIMIC-III. Therefore, adding more training data with different characteristics is beneficial for the former but not for the latter.

The highest value of ROC-AUC is achieved when the model is both trained and tested in the MIMIC-III (0.81) as verified in figure 3 as well. When the same model is tested in the eICU-CRD dataset, we observed lower ROC-AUC values. This metric is improved (0.79) when the model is trained with both datasets, but tested in the same dataset. In terms of the ability to predict transfusion, the model trained in MIMIC-III and tested on the eICU-CRD dataset achieves the best sensitivity (0.93).

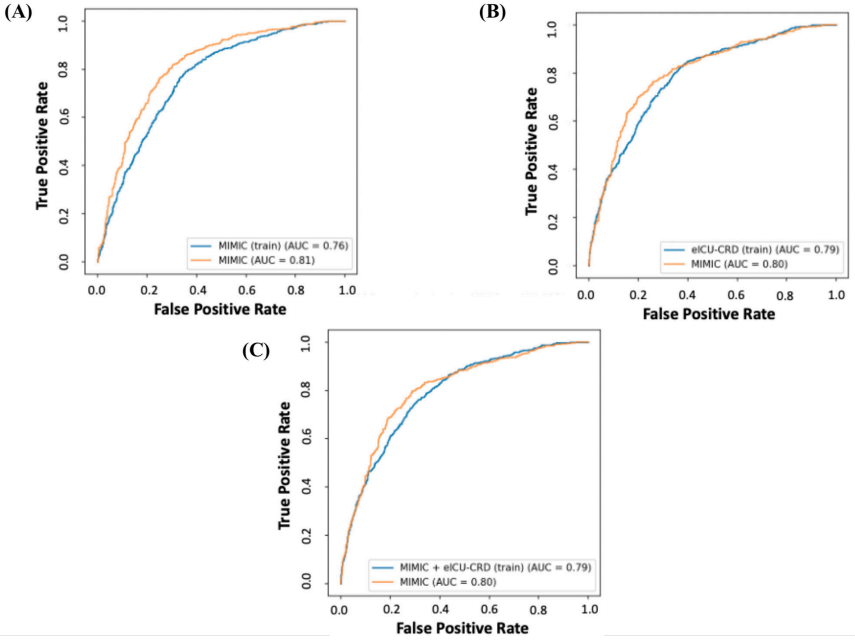

Figure 3 ROC plot for all the test sets. Model is trained on (A) the MIMIC-III training set, (B) the eICU-CRD and (C) on the training set that contains both the MIMIC-III and the eICU-CRD. AUC, area under the curve; eICU-CRD, eICU Collaborative Research Database; ICU, intensive care unit; MIMIC-III, Medical Information Mart for Intensive Care-III; ROC, receiving operating curve.

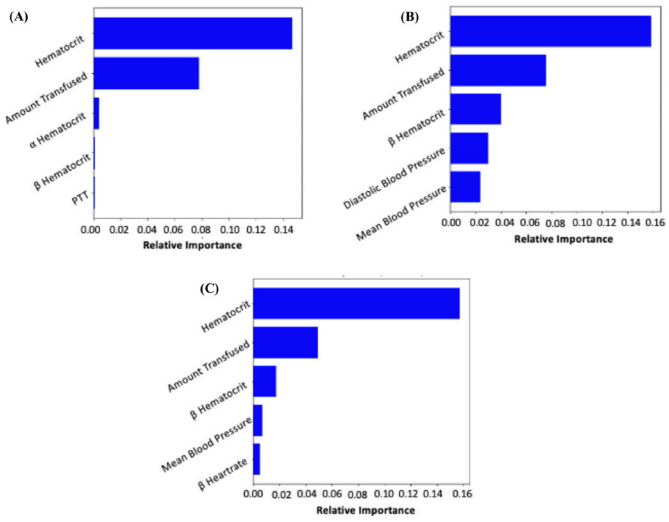

Figure 4 Feature importance plots for all the training sets. Model is trained on (A) the MIMIC-III training set, (B) the eICU-CRD and (C) on the training set that contains both the MIMIC-III and the eICU-CRD. eICU-CRD, eICU Collaborative Research Database; ICU, intensive care unit; MIMIC-III, Medical Information Mart for Intensive Care-III.

The most important features (see figure 4) to predict the need of transfusion are the haematocrit and the amount of blood already transfused during the training time window (0–5 hours) with the corresponding time pattern features (slope and intercept of haematocrit). Because of the importance of haematocrit, the interaction between this feature and the output variable was assessed visually in the partial dependence plots shown in figure 5.

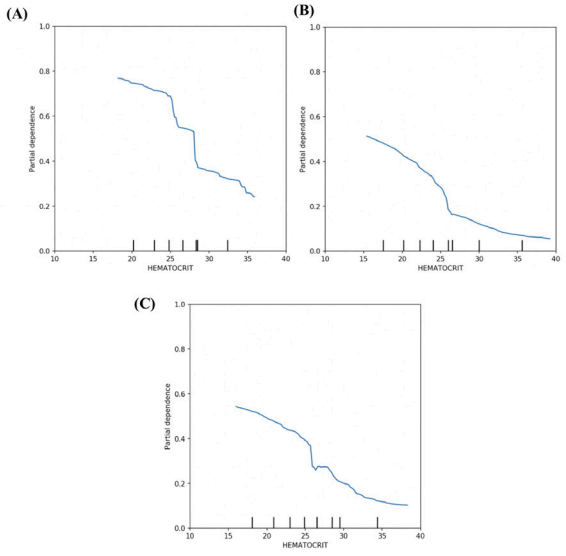

Figure 5 Partial dependence plot of the need of transfusion on haematocrit for all the training sets. Model is trained on (A) the MIMIC-III training set, (B) the eICU-CRD and (C) on the training set that contains both the MIMIC-III and the eICU-CRD. eICU-CRD, eICU-CRD, eICU Collaborative Research Database; ICU, intensive care unit; MIMIC-III, Medical Information Mart for Intensive Care.

Despite the three plots do not have identical shapes, the same trend is verified in the three plots: haematocrit is inversely proportional to the output variable. That implies that if values of haematocrit decreases, the probability of needing blood transfusion increases. Moreover, the partial dependence function shown in figure 5 highlights the presence of a discriminative threshold in the model with respect to haematocrit. It indicates that if the value of haematocrit is greater than this threshold, the probability of bleeding increases substantially. We remark that the value for this threshold seems to be dependent on the dataset that is used for training, where this shift is more noticeable (figure 5A).

DISCUSSION

GI bleeding remains a common reason for ICU admission. In a dataset consisting of over 10 000 patients admitted to the ICU with GI haemorrhage (both upper and lower), under half require transfusion during their ICU admission.[32] We present a model based on observations from the first 5 hours of ICU admission to predict the need for transfusion in the next 24 hours of admission with a high level of accuracy (overall AUC of 0.80). The patient's vital signs and laboratory test findings during the first few hours in the ICU are a good proxy of the measurements in the emergency department.

In the clinical setting, the need for transfusion has been an outcome of interest for GI haemorrhage. Prior work from Villanueva *et al*[6] found that even in active upper GI bleeding, up to half of patients do not require transfusion. Furthermore, it has been established that while the minority of patients with upper GI bleeding require hospitalisation, this can be a significant driver of costs. By identifying patients who will no longer require transfusion, it is possible to safely triage these patients to a regular ward, or even discharged to home if ambulatory monitoring can be provided.

Previous work in this area has focused either on upper or lower GI bleeding separately. In a 2016 analysis by Robertson *et al*,[32] the Rockall, AIMS65 and Glasgow-Blatchford Score (GBS) were all used to predict outcomes for upper GI bleeding. In their population, a total of 62% of the patients required a blood transfusion. They found the GBS to be the best predictor with an ROC-AUC of 0.90. Both the AIMS65 (ROC-AUC 0.72) and full (ROC-AUC 0.68)/pre-endoscopy (ROC-AUC 0.66) Rockall scores were considerably less accurate. However, the use of these scores to predict the need for transfusion has limitations. First, the only score with an ROC-AUC over 0.8, the GBS was validated only on upper GI bleeding (primarily ulcer-related in the initial validation). Furthermore, relying on clinical data input from the healthcare providers, for example, presence of melena, presentation with syncope, presence of heart failure, introduces opportunities for error and bias. Attempts to generalise the use of GBS to lower GI bleeding have found some success but focuses primarily on the prediction of mortality and

need for an intervention instead of transfusion, and with suboptimal accuracy.

The sensitivity, or recall, of the models trained on MIMIC-III is the highest among all other models. A high recall means the algorithm identifies the majority of patients who require transfusion. For the use case presented, sensitivity is more important than precision, or the true positivity rate. When several models have similar ROC-AUC, sensitivity should be prioritised over precision. The consequence of missing patients who eventually bleed and sending them to the regular floor or even discharging them home is worse than over-calling potential persistent bleeders and getting them admitted to the ICU. The context in which the algorithm will be used and for what purpose are crucial to the model building.

Even when models are externally validated in another dataset, there is no guarantee that it will perform well in another patient population. External validation does not circumvent the need to evaluate algorithms trained elsewhere using local data prior to deployment. The performance of any predictive model is dependent on the database used to train the algorithm, and thus, the features available as candidate variables. The relationship between the features and the output of an algorithm is influenced by local practice patterns. In addition, model performance should be continuously monitored after deployment as accuracy almost always wanes over time, requiring model re-calibration.[33]

We submit the potential disruptive impact of AI-based technologies in precision medicine and in clinical decision-support systems. Nonetheless, we are aware of AI-related risks on healthcare applications and the pitfalls that have occurred in the past.[34] Although we reduced the risk of misclassification in the design of our models, we propose a human in the loop system for decision support. A final decision still rests on the healthcare provider after a careful clinical assessment which now includes input from the algorithm. Moreover, before implementation to a real clinical setting, the algorithm requires regulatory approval, human factors engineering to incorporate it into the workflow and prospective evaluation of its impact on hard clinical endpoints including patient harm from false negative predictions.

There are key strengths to the model we presented. First, the calculation can be completely automated without clinician input of symptoms and past medical history. Furthermore, it does not require identification of the source of bleeding–upper versus lower. The model performed well on held out test sets from two different databases, one of them collected from more than 200 hospitals across the USA.

Despite model validation on two databases, the algorithm is not guaranteed to perform accurately in a different institution. We present a reproducible methodology that other hospitals can employ to develop their own algorithm, as different patient demographics and practice patterns would undoubtedly modify the relationship of the features with the outcome being predicted, that is, the need for blood transfusion. At the very least, medical AI algorithms require evaluation on data from the local population prior to prospective evaluation using hard clinical endpoints.

Going forward, this work presents a methodology to build a clinical AI-based model that potentially can be implemented for prediction of the need for transfusion. The algorithm is not meant to replace but to inform decision making, specifically around identification of patients who may not benefit from an ICU-level of care. A prospective trial is warranted to assess the utility of this model in clinical usage.

Author affiliations
[1]Department of Electronic, Information and Bioengineering, Politecnico di Milano, Milano, Italy
[2]Department of Informatics, Università degli Studi di Torino, Torino, Piemonte, Italy
[3]IDMEC, Instituto Superior Tecnico, Universidade de Lisboa, Lisbon, Portugal
[4]General Surgery Department, Istanbul Bagcilar Training and Research Hospital, Istanbul, Turkey
[5]Department of Internal Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
[6]San Raffaele Telethon Institute for Gene Therapy, Milano, Lombardia, Italy
[7]School of Medicine and Surgery, Università degli Studi di Milano-Bicocca, Milano, Lombardia, Italy
[8]Institute of Mathematics, Ecole Polytechnique Federale de Lausanne, Lausanne, VD, Switzerland
[9]Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
[10]Division of Clinical Informatics, Beth Israel Deaconess Medical Center, Boston, MA, USA
[11]Laboratory for Computational Physiology, Harvard-MIT Division of Health Sciences and Technology, Cambridge, Massachusetts, USA
[12]Division of Pulmonary Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

Twitter Leo Anthony Celi @MITCriticalData

Acknowledgements Part of the analysis was performed during the fall course HST.953, Collaborative Data Science in Medicine, at the Massachusetts Institute of Technology and at the first European Society of Intensive Care Medicine datathon in Milan, Italy.

Contributors All authors contributed to writing the manuscript. FC, RL, ARA collaborated on the data extraction, modeling, visualisation and analysis. YA, AZ, MMN, FG guided data extraction. DJS and LAC interpreted, validated results, design of the work and supervised data extraction. SMV, JS, RB and SF reviewed the paper and supervised the work.

Funding ARA is supported by Fundação da Ciência e da Tecnologia (FCT-Portuguese National Science Foundation) through the PhD fellowship PD/BD/114107/2015. ARA, SMV and JS are supported by FCT, through IDMEC, under LAETA, project UIDB/50022/2020; and by FEDER under Programa Operacional Regional de Lisboa, project LISBOA-01-0145-FEDER-031474. LAC is funded by the National Institute of Health through NIBIB R01 EB017205. The project commenced during the HST.953 course (Collaborative Data Science) in Medicine) at the Massachusetts Institute of Technology.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The Massachusetts Institute of Technology (No. 0403000206) and Beth Israel Deaconess Medical Center (2001-P-001699/14) both approved the use of the MIMIC-III for research. The eICU-CRD was certified as meeting Safe Harbor standards by an independent privacy expert (Privacert, Cambridge, Massachusetts, USA) (Health Insurance Portability and Accountability Act Certification No. 1031219-2).

Data availability statement The data that support the findings of this paper employs data from MIMIC-III and eICU-CRD databases. The access to these datasets is controlled and researchers should request access on the PhysioNet

website (https://physionet.org/about/database/). The code is available upon reasonable request contacting R. Levi at riccardo.levi@mail.polimi.it.

ORCID iDs
Riccardo Levi http://orcid.org/0000-0001-9030-0071
Leo Anthony Celi http://orcid.org/0000-0001-6712-6626

REFERENCES

1 Rahman SI-U, Saeian K. Nonvariceal upper gastrointestinal bleeding. *Crit Care Clin* 2016;32:223–39.
2 Esrailian E, Gralnek IM. Nonvariceal upper gastrointestinal bleeding: epidemiology and diagnosis. *Gastroenterol Clin North Am* 2005;34:589–605.
3 Minne L, Abu-Hanna A, de Jonge E. Evaluation of SOFA-based models for predicting mortality in the ICU: a systematic review. *Crit Care* 2008;12:R161.
4 Akaraborworn O, Chaiwat O, Chatmongkolchart S, *et al*. Prediction of massive transfusion in trauma patients in the surgical intensive care units (THAI-SICU study). *Chin J Traumatol* 2019;22:219–22.
5 Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for upper-gastrointestinal haemorrhage. *Lancet* 2000;356:1318–21.
6 Villanueva C, Colomo A, Bosch A, *et al*. Transfusion strategies for acute upper gastrointestinal bleeding. *N Engl J Med* 2013;368:11–21.
7 The Lancet Digital Health. Walking the tightrope of artificial intelligence guidelines in clinical practice. *Lancet Digit Health* 2019;1:e100.
8 Cosgriff CV, Celi LA, Stone DJ. Critical care, critical data. *Biomed Eng Comput Biol* 2019;10:117959721985656–7.
9 Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA* 2018;319:1317.
10 Kindle RD, Badawi O, Celi LA, *et al*. Intensive care unit telemedicine in the era of big data, artificial intelligence, and computer clinical decision support systems. *Crit Care Clin* 2019;35:483–95.
11 Gholami B, Haddad WM, Bailey JM. Ai in the ICU: in the intensive care unit, artificial intelligence can keep Watch. *IEEE Spectr* 2018;55:31–5.
12 Ruffle JK, Farmer AD, Aziz Q. Artificial Intelligence-Assisted Gastroenterology— promises and pitfalls. *Off J Am Coll Gastroenterol* 2019;114.
13 Chu A, Ahn H, Halwan B, *et al*. A decision support system to facilitate management of patients with acute gastrointestinal bleeding. *Artif Intell Med* 2008;42:247–59.
14 Das A, Ben-Menachem T, Farooq FT, *et al*. Artificial neural network as a predictive instrument in patients with acute nonvariceal upper gastrointestinal hemorrhage. *Gastroenterology* 2008;134:65–74.
15 Cismondi F, Celi LA, Fialho AS, *et al*. Reducing unnecessary lab testing in the ICU with artificial intelligence. *Int J Med Inform* 2013;82:345–58.
16 Vandenbroucke JP, von Elm E, Altman DG, *et al*. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007;4:e297.
17 Johnson AEW, Pollard TJ, Shen L, *et al*. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.
18 Pollard TJ, Johnson AEW, Raffa JD, *et al*. The eICU Collaborative research database, a freely available multi-center database for critical care research. *Sci Data* 2018;5:180178.
19 Celi LA, Hassan E, Marquardt C, *et al*. The eICU: It's not just telemedicine. *Crit Care Med* 2001;29:N183–9 http://journals.lww.com/00003246-200108001-00007
20 McKinney W. Data structures for statistical computing in python. proC. 9th python SCI. *Conf* 2010.
21 der WSvan, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng* 2011;13:22–30.
22 Virtanen P, Gommers R, Oliphant TE, *et al*. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat Methods* 2020;17:261–72.
23 Pedregosa F, Varoquaux G, Gramfort A. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011.
24 Bergstra J. Distributed Asynchronous Hyperparameter Optimization [Internet]. Available: https://pypi.org/project/hyperopt/ [Accessed cited 2020 May 30].
25 Yeo I-K, Johnson RA. A new family of power transformations to improve normality or symmetry. *Biometrika* 2000;87:954–9.
26 Guyon I, Weston J, Barnhill S, *et al*. Gene selection for cancer classification using support vector machines. *Mach Learn* 2002;46:389–422.
27 Bergstra J, Yamins D, Cox DD. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. In: *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28. JMLR.org*, 2013: p. I–115–I–123.
28 Breiman L. Random forests. *Mach Learn* 2001;45:5–32.
29 Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning*. New York, NY: Springer New York, 2009.
30 Niculescu-Mizil A, Caruana R. *Predicting good probabilities with supervised learning. In: Proceedings of the 22nd international conference on Machine learning - ICML '05*. New York, USA: ACM Press, 2005: 625–32.
31 Platt J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv Large Margin Classif* 2000;10.
32 Robertson M, Majumdar A, Boyapati R, *et al*. Risk stratification in acute upper Gi bleeding: comparison of the AIMS65 score with the Glasgow-Blatchford and Rockall scoring systems. *Gastrointest Endosc* 2016;83:1151–60.
33 Futoma J, Simons M, Panch T. The myth of generalizability in clinical research and machine learning in healthcare. *Lancet Digit Heal*. In Press 2020.
34 Topol EJ. High-Performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56.

BMJ Health & Care Informatics

Accuracy of periocular lesion assessment using telemedicine

Swan Kang ,[1] Mohammad Dehabadi,[1] Dawn A Sim,[2] Peter B M Thomas ,[2] Doris Appiah Ewusi,[1] Daniel Ezra[1]

To cite: Kang S, Dehabadi M, Sim DA, *et al*. Accuracy of periocular lesion assessment using telemedicine. *BMJ Health Care Inform* 2021;**28**:e100287. doi:10.1136/bmjhci-2020-100287

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjhci-2020-100287).

Received 19 November 2020
Revised 15 December 2020
Accepted 24 December 2020

[1]Adnexal Department, Moorfields Eye Hospital NHS Foundation Trust, London, UK
[2]NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK

Correspondence to
Dr Swan Kang;
swan.kang1@nhs.net

ABSTRACT

Objectives To assess the agreement in diagnosis and management plans reached between clinicians reviewing eyelid lesions remotely and in face-to-face clinics.

Methods In this single-centre observational case series, data were prospectively collected on 50 consecutive adults referred with eyelid lesions suitable to be seen by a nurse. A proforma was completed to gather salient information. A nurse specialist saw patients in face-to-face clinics and collected information using the proforma, devising a diagnosis and management plan. Photographs of the eyelid lesions were taken by a medical photographer. A subsequent remote review was completed by an oculoplastic consultant using the proforma information and photographs in the absence of the patient. The diagnosis and management plan constructed by the nurse specialist were compared with those reached by the consultant.

Results Complete data were available for 44 consecutive cases. There was an overall 91% agreement (40 cases out of 44) between the diagnoses reached by the nurse specialist, and the remote reviewer; kappa coefficient 0.88 (95% CI 0.76 to 0.99). There was an overall 82% agreement (36 out of 44 cases) in the management plans devised by the nurse-led clinic and remote reviewer; kappa coefficient 0.74 (95% CI 0.58 to 0.90). The average time taken for a remote reviewer to reach a diagnosis and management plan was 1 min and 20 s.

Conclusions This study evaluated the feasibility of assessing eyelid lesions using asynchronous telemedicine. There was overall a high rate of concordance in the diagnosis reached, and management devised between the clinic and remote review.

Summary

What is already known?

- The evidence base for the use of telemedicine in oculoplastics is limited.
- Telemedicine provides benefits such as improved outcomes, efficiency and access to healthcare.

What does this paper add?

- Eyelid lesion assessment using store-and-forward telemedicine is comparable to face-to-face evaluation.
- Remote eyelid lesion assessment can be used to optimise patient care pathways.

INTRODUCTION

Telemedicine is the use of electronic information and communication technologies to deliver healthcare services at a distance[1] and is well established in ophthalmology, particularly in the subspecialty areas of medical retina and glaucoma.[2–5] Prior to the COVID-19 pandemic, the reported applications of telemedicine in oculoplastics were largely limited to settings where the access to healthcare remains a challenge[6–9] and only few studies evaluated the utility of synchronous telemedicine in assessment of oculoplastic conditions.[10–12] Since the emergence of COVID-19, the use of telemedicine, particularly video consultations, has increased exponentially. The literature reports varying degree of effectiveness of video consultations in assessment of eyelid lesions.[13 14] While the evidence base for the use of video consultations in oculoplastics is growing following the COVID-19 outbreak, the literature on the application of asynchronous or store-and-forward telemedicine in the assessment of eyelid lesions remains scarce.[15 16] This contrasts significantly with the successful implementation and scaling of a closely related specialty—teledermatology, which has matured over two decades and is now a widely accepted form of service delivery.[17]

The hospital eye service is experiencing a severe shortage of resources to safely cope with demand and it is predicted that the demand will increase by 30%–40% over the next 20 years.[18] In our oculoplastic service patients with eyelid lesions make up more than 50% of new referrals. Optimising pathways for these patients is vital to provide an efficient service and to reduce waiting times particularly when identification of malignant lesions is time sensitive. Multiple initiatives have been implemented to offer prompt diagnosis and treatment of eyelid lesions. A nurse-led eyelid lesions service has been shown to provide comparable diagnostic accuracy compared with a doctor-led service.[19] One-stop minor surgery lists have been set up in order to provide patients with same

day surgery to expedite the delivery of their treatment and to reduce the need for multiple hospital visits. The National Health Service (NHS) long-term plan and NHS England service transformation plans for ophthalmology suggest remote care as a means to cope with a surge in patient demand.[20 21] The recent pandemic along with the increasing access and use of the internet and digital technology as well as growing acceptance of remote care among clinicians and patients have accelerated the move towards telemedicine.

We investigated the use of asynchronous telemedicine to remotely diagnose and formulate management plans for eyelid lesions. To enhance the diagnostic accuracy of such an approach, templates for structured and pertinent data collection including patient history, along with photographs of eyelid lesions were designed and used. The data and photographs were reviewed by a clinician in the absence of the patient. The agreement between the diagnosis and management plan reached by clinician reviewing the data remotely with those devised by clinician reviewing patients in face-to-face clinics was assessed.

METHODS

This was a single-centre prospective observational case series conducted in a tertiary ophthalmic specialist hospital. Verbal consent was obtained from study participants.

Fifty consecutive adult patients, who were referred with benign eyelid lesions based on the information provided by the referrer, were included in the study. We excluded suspected skin cancer referrals. Data were prospectively collected on patients seen between November 2019 to January 2020 in nurse-led clinics as benign eyelid lesion cases are seen in nurse-led clinics at our institution. A bespoke structured proforma was designed to characterise the eyelid lesions and gather relevant information including history of skin lesions, ophthalmic, medical and drug history (online supplemental material). A nurse specialist saw patients in face-to-face clinics and collected information using the proforma and devised a diagnosis and management plan. Photographs of the eyelid lesions were taken at the end of the clinic appointment by a qualified medical photographer using a Canon EOS 7D camera with 5184×3456 pixels resolution. A subsequent remote review of collected data and photographs was completed by an oculoplastic consultant in the absence of the patient. All cases were assessed by the same nurse specialist and the same consultant remote reviewer. Data, including patient demographics, number of days patients waited between the date of referral and the review, time taken for the consultant to review the cases remotely and referral sources as well as histopathological diagnoses where available, were collected. The diagnosis and management plan constructed by the nurse specialist in the face-to-face clinic and those reached by the consultant via remote review were compared by an independent assessor (oculoplastic fellow).

Table 1 Diagnoses and management plans of the excluded cases

Excluded cases	Face-to-face diagnosis	Face-to-face management plan
A	Chalazion	Incision and curettage
B	Chalazion	Incision and curettage
C	Chalazion	Discharge
D	Epidermoid cyst	Follow-up
E	Epidermoid cyst	Discharge
F	Epidermoid cyst	Excision biopsy

Kappa coefficient and 95% CIs were used to evaluate the agreement between the face-to-face clinic and remote review. Pearson's χ^2 test was used compare the distribution of outcomes. Tests with p values less than 0.05 were considered statistically significant. Data analysis was performed using R software (R Core Team, 2017).

RESULTS

Complete data were available for 44 consecutive cases. Six cases were excluded from the study as these patients did not wait to have their photographs taken after the face-to-face consultations, thus not allowing remote review to take place. The details of the excluded cases are listed in table 1. No case was excluded due to the insufficient photograph quality.

The mean age of patients reviewed was 47.3 years (range:18 to 72 years). Twenty-seven (61.4%) patients were female, and 17 (38.6%) patients were male. Patients waited an average of 49 days (range: 18–97) from the date of referral to be seen in the face-to-face clinic. Twenty-eight (64%) patients were referred by general practitioners while 12 (27%) were by ophthalmologists who do not specialise in oculoplastics, and 4 (9%) by optometrists. An average time taken to review a case remotely by means of assessing the collected data and photographs was measured to be 1 min 20 s (range: 20–120 s). The baseline characteristics and key metrics are summarised in table 2.

Table 2 Baseline characteristics of the study participants and key metrics

Characteristics	Total no, n=44
Age—mean (SD) in years	47.3 (14.5)
Sex—female n(%)	27 (61.4)
No of days patients waited between the date of referral and the review—mean (SD) in days	49 (18.0)
Referral source n(%)	
General practitioners	28 (64)
Non-oculoplastic ophthalmologists	12 (27)
Optometrists	4 (9)

Table 3 Outcomes of face-to-face clinic and remote review

	Face-to-face outcome n (%)	Remote review outcome n (%)
Discharge	19 (43.2)	17 (38.6)
Incision and curettage	13 (29.5)	14 (31.8)
Excision biopsy	7 (15.9)	11 (25)
Follow-up	5 (11.4)	2 (4.5)

The primary diagnoses were (as per the standard outpatient care, face-to-face clinic): chalazion n=19 (43.2%), epidermoid cyst n=9 (20.5%), papilloma n=9 (20.5%), hidrocystoma n=2 (4.5%), naevus n=2 (4.5%), xanthelasma n=1 (2.3%), conjunctival granuloma n=1 (2.3%) and lipoma n=1 (2.3%). The outcomes of face-to-face clinic were discharge n=19 (43.2%), incision and curettage n=13 (29.5%), excision biopsy n=7 (15.9%) and follow-up n=5 (11.4%). The outcomes of remote review were discharge n=17 (38.6%), incision and curettage n=14 (31.8%), excision biopsy n=11 (25%) and follow-up n=2 (4.5%) (table 3). There was no statistically significant difference between the distribution of outcomes between face-to-face and remote review (p=0.21).

The full list of cases where there were disagreements of diagnosis and/or management plan between those reached by the face-to-face clinician and the remote reviewer is outlined in table 4.

There was an overall 91% agreement (40 cases out of 44) between the diagnoses made by the nurse specialist, and the remote reviewer. Kappa coefficient for diagnostic agreement between face-to-face and remote review was 0.88 (95% CI 0.76 to 0.99). Three out of the four cases with disagreements in diagnosis resulted in different management plans. In one case (case 5 of table 4; hereinafter all numerical case numbers refer to those listed in table 4), the nurse specialist diagnosed the lesion as a benign cyst and discharged the patient whereas the remote reviewer suspected a basal cell carcinoma due to the presence of central ulceration and listed for a biopsy. The lesion subsequently spontaneously resolved without intervention indicating that the lesion was of benign origin. In two cases, the lesions (a chalazion and a lipoma, cases 3 and 9, respectively) were not demonstrated well on photographs as these were subcutaneous, the undulation caused by the lesions was subtle, and there were no overlying skin changes. The lesions in question were not the focal point in the photographs probably due to the photographer being uncertain of the location of the lesions of concern. In case 3, the remote reviewer diagnosed an enlarged caruncle, where the patient was referred with a small chalazion in the lower lid near the punctum which did not display well in the photographs. The remote reviewer listed the patient for a biopsy of the caruncle whereas the nurse booked the patient for an incision and curettage of the chalazion. In case 9, the remote reviewer noted and diagnosed a papilloma which was adjacent to the lipoma and listed for a biopsy whereas the nurse practitioner brought the patient back for a review of the suspected lipoma. In case 4, a diagnosis of an epidermoid cyst was made by the nurse specialist whereas the remote reviewer diagnosed it as a chalazion and despite the discrepancy in the diagnosis, both the nurse practitioner and the remote reviewer discharged the patient.

There was an overall 82% agreement (36 out of 44 cases) in the management plans devised by the nurse-led clinic and remote reviewer. Kappa coefficient for management agreement between face-to-face and remote review was 0.74 (95% CI 0.58 to 0.90). In two cases of chalazia (cases 1 and 2), the remote reviewer chose to list for an incision and curettage whereas the nurse specialist discharged the patients. In two cases of papilloma (cases 5 and 6), the remote reviewer either discharged the patient or listed for excision biopsy, however, the nurse specialist booked follow-up appointments in 3 months. In the case of a naevus (case 8), the remote reviewer opted for a biopsy whereas the nurse specialist arranged a follow-up appointment in a clinic.

Seven patients were listed for excision biopsy by the face-to-face clinician and histopathological diagnoses were available for six as one patient did not contact the

Table 4 List of cases where there were disagreements of diagnosis and management plan between those reached by the face-to-face clinician and the remote reviewer

Case no	Face-to-face diagnosis	Face-to-face management plan	Remote diagnosis	Remote management plan
1	Chalazion	Discharge*	Chalazion	Incision and curettage*
2	Chalazion	Discharge*	Chalazion	Incision and curettage*
3	Chalazion*	Incision and curettage*	Enlarged caruncle*	Follow-up*
4	Epidermoid cyst*	Discharge	Chalazion*	Discharge
5	Epidermoid cyst*	Discharge*	Basal cell carcinoma*	Excision biopsy*
6	Papilloma	Follow-up*	Papilloma	Excision biopsy*
7	Papilloma	Follow-up*	Papilloma	Discharge*
8	Naevus	Follow-up*	Naevus	Excision biopsy*
9	Lipoma*	Follow-up*	Papilloma*	Excision biopsy*

*Indicates disagreements between face-to-face and remote reviewer.

Table 5 Clinical diagnoses reached by face-to-face clinician and remote reviewer where histopathological diagnoses were available

Case no	Face-to-face diagnosis	Remote diagnosis	Histopathological diagnosis
a	Epidermoid cyst	Epidermoid cyst	Epidermoid cyst
b	Epidermoid cyst	Epidermoid cyst	Epidermoid cyst
c	Benign intradermal naevus	Benign intradermal naevus	Benign intradermal naevus
d	Papilloma	Papilloma	Benign intradermal naevus
e	Papilloma	Papilloma	Papilloma
f	Granuloma	Granuloma	Granuloma

hospital to arrange a date for surgery. In five cases, clinical diagnose reached by both face-to-face and remote reviewers were in agreement and were confirmed by histopathological diagnoses. In one case, both face-to-face and remote reviewers diagnosed the lesion as a papilloma and the histological diagnosis demonstrated a benign intradermal naevus (table 5).

DISCUSSION

The benefits of telemedicine including improved outcomes, cost and time-efficiency, and increasing access to healthcare are well documented in ophthalmology.[22–24] Telemedicine played a vital role in allowing continued provision of patient care while mitigating the risk of viral transmission during the recent COVID-19 pandemic and it has accelerated rapid integration of remote care into routine clinical practice. Oculoplastics is particularly well suited to telemedicine due to the highly visual nature of clinical assessment which can be evaluated without the need of a specialist equipment, however, the evidence base to support wider application of telemedicine is limited.[13]

Our study assessed the ability to establish accurate diagnoses and management plans of eyelid lesions remotely using a data collection proforma and photographs. Diagnostic and management plan agreements were good between face-to-face consultations and remote review at 91% and 82%, respectively. Histopathological diagnoses, where available, were compared with clinical impressions made by face-to-face assessor and remote reviewer. Both face-to-face assessor and remote reviewer correctly diagnosed eyelid lesions in five out of six cases (83.3%) with both clinicians diagnosing the same lesion as a papilloma where the histopathology demonstrated a benign intradermal naevus. It is possible that the diagnostic discrepancy demonstrated in our study is not indicative of the disagreements between different modes of review but rather represents the diagnostic inaccuracies that would be observed in face-to-face clinic settings. Previous studies reported the diagnostic accuracy of 70%–96%.[25–28] It has therefore been recommended that all excised eyelid tissues should be sent for histopathological analysis as clinical assessment alone will not warrant accurate diagnosis and malignant eyelid lesions may masquerade as clinically benign conditions. It is not uncommon to observe a diversity of clinical opinion among oculoplastic clinicians on the optimal management options for individual patients. In our study, although not statistically significant, the remote reviewer had a lower threshold to opt for excision biopsy whereas the face-to-face clinician was more likely to adopt a ‘wait and see’ approach. In one case, the remote reviewer suspected a basal cell carcinoma and chose to list the patient for a biopsy taking a more cautious approach.

Although the evaluation of eyelid lesions is akin to that of skin lesions, the inherent two-dimensional character of photographic evaluation poses particular challenges for eyelid lesions. The multilamellar anatomical construction of the eyelids means that the lesions can originate from or extend to deeper anatomical structures such as the tarsus or conjunctiva and may not be demonstrated well on photographs. Furthermore, eyelid lesions are likely to be smaller than those found on other parts of the body and may require more specialised lenses for adequate image resolution. A dynamic examination of eyelid lesions provides additional information as it is possible to ascertain, for example, if the lesion is tethered to the underlying structure, which is likely to narrow down the differential diagnoses.

Our study is one of the few that examined the utility of asynchronous telemedicine in oculoplastics. The proforma used in this study has been specifically designed to collect pertinent information to risk stratify the eyelid lesions by incorporating non-genetic risk factors such as age, history of previous skin malignancies and drug history.[29–33]

Our study has several limitations. A small number of patients at a single centre with a limited range of diagnoses was enrolled and suspected malignant cases were excluded, thus limiting generalisability of the study. Six cases were excluded as clinical photographs were not captured. The face-to-face evaluation was performed by a nurse while the remote review was conducted by a doctor. The discrepancy between face-to-face consultations and remote review could be attributed to the inherent difference in training received. Having more than one face-to-face and remote reviewers and assessing intraobserver and interobserver agreement would have allowed more

in-depth exploration and reduce bias. In some cases, the lesion in question was not the focal point of some clinical photographs and this made it difficult for the remote assessor to evaluate the case.

The use of telemedicine in eyelid lesion assessment has a potential to reduce outpatient appointment and surgery waiting times, permit more accurate triaging of eyelid lesion cases, obviate the need for face-to-face appointments, and allow enhanced monitoring of eyelid lesions. This approach can be adopted for (1) referral refinement which has been shown to successfully reduce unnecessary or inappropriate referrals thus reducing waiting times; (2) image-based triage to risk stratify patients to allow more accurate assessment of the clinical urgency; (3) electronic consultation where the remote assessment can obviate the need for face-to-face appointments and (4) remote monitoring of eyelid lesions where photograph-based measurement has been shown to be more accurate than traditional face-to-face clinical evaluation.[34–37]

A multicentre study including more subjects with a wider range of pathologies and histopathological diagnoses needs to be performed to further assess the accuracy of the use of asynchronous eyelid lesion assessment and to evaluate generalisability of the results. Involving more than one face-to-face and remote reviewers is important to minimise bias. Further research into patients' and clinicians' views on the use of telemedicine in oculoplastics should be carried out in order to design a service that addresses their needs and concerns.

Twitter Dawn A Sim @dawnasim and Peter B M Thomas @pbmthomas

Contributors Initials of the authors involved in each of the following step of writing this paper are listed below. Planning: SK, MD, DAS, PBMT, DAE; conducting: SK, MD, DAE, DE; reporting: SK, MD, PBMT, DAS, DE; guarantor: SK.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was performed in full compliance with the Declaration of Helsinki. Permission to perform this study was granted by Moorfields Eye Hospital NHS Foundation Trust (reference number: CA/AD/675).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Swan Kang http://orcid.org/0000-0002-1303-9188
Peter B M Thomas http://orcid.org/0000-0003-1681-7659

REFERENCES

1 Ryu S. Telemedicine: opportunities and developments in member states: report on the second global survey on eHealth 2009 (global Observatory for eHealth series, volume 2). *Healthc Inform Res* 2012;18:153–5.
2 Scanlon PH. The English national screening programme for diabetic retinopathy 2003-2016. *Acta Diabetol* 2017;54:515–25.
3 Wei JC, Valentino DJ, Bell DS, *et al*. A web-based telemedicine system for diabetic retinopathy screening using digital fundus photography. *Telemed J E Health* 2006;12:50–7.
4 Taylor CR, Merin LM, Salunga AM, *et al*. Improving diabetic retinopathy screening ratios using telemedicine-based digital retinal imaging technology: the Vine Hill study. *Diabetes Care* 2007;30:574–8.
5 Boucher MC. Organisation of seamless teleophthalmology reliable management of diabetic retinopathy screening through. *Can J Diabetes* 2009;33:284.
6 Misra N, Khanna RC, Mettla AL, *et al*. Role of teleophthalmology to manage anterior segment conditions in vision centres of South India: EyeSmart study-I. *Indian J Ophthalmol* 2020;68:362–7.
7 BenZion I, Helveston EM. Use of telemedicine to assist ophthalmologists in developing countries for the diagnosis and management of four categories of ophthalmic pathology. *Clin Ophthalmol* 2007;1:489–95.
8 Mines MJ, Bower KS, Lappan CM, *et al*. The United States Army Ocular Teleconsultation program 2004 through 2009. *Am J Ophthalmol* 2011;152:126–32.
9 Verma M, Raman R, Mohan RE. Application of tele-ophthalmology in remote diagnosis and management of adnexal and orbital diseases. *Indian J Ophthalmol* 2009;57:381–4.
10 Rayner S, Beaconsfield M, Kennedy C, *et al*. Subspecialty adnexal Ophthalmological examination using telemedicine. *J Telemed Telecare* 2001;7 Suppl 1:29–31.
11 Hwang CJ, Eftekhari K, Schwarcz RM, *et al*. The aesthetic Oculoplastic surgery video Teleconference consult. *Aesthet Surg J* 2019;39:714–8. 21.
12 Threlkeld AB, Fahd T, Camp M, *et al*. Telemedical evaluation of ocular adnexa and anterior segment. *Am J Ophthalmol* 1999;127:464–6.
13 Kang S, Thomas PBM, Sim DA, *et al*. Oculoplastic video-based telemedicine consultations: Covid-19 and beyond. *Eye* 2020;34:1193–5.
14 Jamison A, Diaper C, Drummond S, *et al*. Telemedicine in Oculoplastics: the real-life application of video consultation clinics. *Ophthalmic Plast Reconstr Surg* 2020. doi:10.1097/IOP.0000000000001852. [Epub ahead of print: 04 Nov 2020].
15 Aoki L, Pereira IC, Matayoshi S. Comparative study between conventional camera images and smartphone images for eyelid tumor telediagnosis. *Rev Col Bras Cir* 2019;46:e2083.
16 Rossato LA, Carneiro RC, Miyazaki A. *Accuracy of clinical examination in the diagnosis of eyelid lesions*. 73, 2014.
17 Mehrtens SH, Shall L, Halpern SM. A 14-year review of a UK teledermatology service: experience of over 40 000 teleconsultations. *Clin Exp Dermatol* 2019;44:874–81.
18 NHS Digital. Hospital Outpatient Activity 2018-19 [Internet]. Available: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-outpatient-activity/2018-19 [Accessed 12 Jul 2020].
19 Mohite AA, Johnson A, Rathore DS, *et al*. Accuracy of clinical diagnosis of benign eyelid lesions: is a dedicated nurse-led service safe and effective? *Orbit* 2016;35:193–8.
20 Plan NLT. NHS Long Term Plan [Internet]. NHS Long Term Plan. Available: https://www.longtermplan.nhs.uk [Accessed 02 Dec 2020].
21 NHS England ». Transforming elective care services ophthalmology [Internet]. Available: https://www.england.nhs.uk/publication/transforming-elective-care-services-ophthalmology/ [Accessed 02 Dec 2020].
22 Jani PD, Forbes L, Choudhury A, *et al*. Evaluation of diabetic retinal screening and factors for ophthalmology referral in a telemedicine network. *JAMA Ophthalmol* 2017;135:706–14.
23 Avidor D, Loewenstein A, Waisbourd M, *et al*. Cost-Effectiveness of diabetic retinopathy screening programs using telemedicine: a systematic review. *Cost Eff Resour Alloc* 2020;18:16.
24 Ting DSW, Tan GSW. Telemedicine for diabetic retinopathy screening. *JAMA Ophthalmol* 2017;135:722–3.
25 Kersten RC, Ewing-Chow D, Kulwin DR, *et al*. Accuracy of clinical diagnosis of cutaneous eyelid lesions. *Ophthalmology* 1997;104:479–84.

26 Deokule S, Child V, Tarin S, *et al*. Diagnostic accuracy of benign eyelid skin lesions in the minor operation theatre. *Orbit* 2003;22:235–8.
27 Margo CE. Eyelid tumors: accuracy of clinical diagnosis. *Am J Ophthalmol* 1999;128:635–6.
28 Hillson TR, Harvey JT, Hurwitz JJ, *et al*. Sensitivity and specificity of the diagnosis of periocular lesions by oculoplastic surgeons. *Can J Ophthalmol* 1998;33:377–83.
29 Clouser MC, Harris RB, Roe DJ, *et al*. Risk group, skin lesion history, and sun sensitivity reliability in squamous cell skin cancer progression. *Cancer Epidemiol Biomarkers Prev* 2006;15:2292–7.
30 Belbasis L, Stefanaki I, Stratigos AJ, *et al*. Non-genetic risk factors for cutaneous melanoma and keratinocyte skin cancers: an umbrella review of meta-analyses. *J Dermatol Sci* 2016;84:330–9.
31 Etzkorn JR, Parikh RP, Marzban SS, *et al*. Identifying risk factors using a skin cancer screening program. *Cancer Control* 2013;20:248–54.
32 Drugge ED, Okundaye OI, Sarac RM, *et al*. Melanoma screening using patient self-assessed risk and total body photography. *Dermatol Online J* 2019;25:13030.
33 Fisher NM, Schaffer JV, Berwick M, *et al*. Breslow depth of cutaneous melanoma: impact of factors related to surveillance of the skin, including prior skin biopsies and family history of melanoma. *J Am Acad Dermatol* 2005;53:393–406.
34 Borooah S, Grant B, Blaikie A, *et al*. Using electronic referral with digital imaging between primary and secondary ophthalmic services: a long term prospective analysis of regional service redesign. *Eye* 2013;27:392–7.
35 Hanson C, Tennant MTS, Rudnisky CJ. Optometric referrals to retina specialists: evaluation and triage via teleophthalmology. *Telemed J E Health* 2008;14:441–5.
36 Kelly SP, Wallwork I, Haider D, *et al*. Teleophthalmology with optical coherence tomography imaging in community optometry. Evaluation of a quality improvement for macular patients. *Clin Ophthalmol* 2011;5:1673–8.
37 Soto JL, Vrcek I, Ozgur O, *et al*. Evaluating the growth of eyelid lesions: comparison of clinical evaluation and a software-based model. *Int Ophthalmol* 2017;37:1257–61.