Interactive Journal of Medical Research

Impact Factor (2022): 2.0
Volume 11 (2022), Issue 1 ISSN 1929-073X

Contents

Viewpoint

Implementation and Experiences of Telehealth: Balancing Policies With Practice in Countries of South Asia, Kuwait, and the European Union

Oommen John[1,2*], MD; Suptendra Nath Sarbadhikari[3*], MBBS, PhD; Thanga Prabhu[4*], MD; Ashvini Goel[5*], MD; Alexander Thomas[6*], MS, MPhil, PGDMLE; Sunil Shroff[7*], MS; Fazilah Allaudin[8*], MD; Chaminda Weerabaddana[9*], MD; Dari Alhuwail[10,11*], PhD; Udaya Koirala[12*], MD; Jayalal Johnrose[13*], MD; Patricia Codyre[14*], MD; Andy Bleaden[15*], BA; Shubnum Singh[16*], MD; Shuchin Bajaj[17*], MD

[1]George Institute for Global Health, University of New South Wales, New Delhi, India
[2]Prasanna School of Public Health, Manipal, India
[3]The George Institute for Global Health India, New Delhi, India
[4]St Johns Health Innovation Foundation, Bengaluru, India
[5]Telemedicine Society of India, Lucknow, India
[6]Association of Healthcare Providers India, New Delhi, India
[7]Madras Medical Mission, Chennai, India
[8]Planning Division, Ministry of Health, Kuala Lumpur, Malaysia
[9]Ministry of Health, Colombo, Sri Lanka
[10]Information Science Department, Kuwait University, Kuwait City, Kuwait
[11]Health Informatics Unit, Dasman Diabetes Institute, Kuwait City, Kuwait
[12]Telemedicine Society of Nepal, Kathmandu, Nepal
[13]Indian Medical Association, New Delhi, India
[14]Digital Health and Innovation, World Health Organization Regional Office for South-East Asia, New Delhi, India
[15]European Connected Health Alliance, Padfield, High Peak, United Kingdom
[16]National Healthcare Council, Confederation of Indian Industry, New Delhi, India
[17]Ujala Cygnus Healthcare Services, New Delhi, India
[*]all authors contributed equally

Corresponding Author:
Suptendra Nath Sarbadhikari, MBBS, PhD
The George Institute for Global Health India
F-BLOCK, 311-312, Third Floor
Jasola Vihar
New Delhi, 110025
India
Phone: 91 9717669574
Email: supten@gmail.com

Abstract

This viewpoint summarizes the discussion that occurred during the “Translating Policy to Practice in Telehealth–Lessons from Global Implementation Experiences” panel that was held virtually at Telemedicon2020, December 18-20, 2020. This panel brought together policy and implementation experts from some countries of South Asia, Kuwait, and the European Union to share their experiences in the development and implementation of telehealth standards and of the scale up of telehealth interventions within health systems. Several common themes arose from the discussion, including the significant role of people; encouragement by respective government policymakers; addressing concerns, particularly related to privacy, confidentiality, and security; and capacity building of human resources. These are discussed in turn, along with the future directions identified by the panelists, which emphasized the need for active encouragement toward the adoption and diffusion of digital health in general and of telehealth in particular. All stakeholders, ranging from governmental policymakers to common citizens, need to come together to build trusting partnerships to realize the advantages offered by telehealth.

(Interact J Med Res 2022;11(1):e30755) doi:10.2196/30755

KEYWORDS

telehealth policy and practice; implementation lessons; challenges in scaling up; capacity building of human resources; data privacy; telehealth; health policy; telemedicine; implementation; challenges; human resources; digital health; data security

Introduction

This paper summarize the discussion that occurred during the "Translating Policy to Practice in Telehealth–Lessons from Global Implementation Experiences" panel held virtually at Telemedicon2020 [1], the annual conference of the Telemedicine Society of India (TSI), from December 18 to December 20, 2020. This panel brought together policy and implementation experts from across the globe to share their experiences in the development and implementation of telehealth standards and of scaling up telehealth interventions within health systems. The panel composition is given in the Table 1. Each panelist was asked to (1) trace the evolution of telemedicine in their respective countries, particularly with reference to the COVID 19 pandemic; (2) describe the current policies guiding telemedicine; (3) describe the actual use and adoption of telemedicine within their countries; and (4) identify future directions in the adoption and diffusion of telehealth.

The COVID-19 pandemic has accelerated the role of various digital health interventions, including telehealth, in supporting health services delivery [2]. Telehealth includes a broader scope of remote health care services than telemedicine. Telemedicine refers specifically to remote clinical services, whereas telehealth can refer to remote nonclinical services (administrative/educational). The Global Strategy for Digital Health (GSDH) 2020-2025 [3] retains this definition of telemedicine as:

> *The delivery of health care services, where distance is a critical factor, by all health-care professionals using information and communications technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and the continuing education of health care workers, with the aim of advancing the health of individuals and communities.*

The GSDH 2020-2025 generally defines digital health as "the field of knowledge and practice associated with the development and use of digital technologies to improve health" [3]. This definition is extended on page 13 of the document as:

> *This definition encompasses eHealth, in line with that in document EB142/20 on mHealth, noted by the Executive Board at its 142nd session (see document EB142/2017/REC/2, summary records of thirteenth meeting, section 2), which stated that "Today the term digital health is often used as a broad umbrella term encompassing eHealth as well as developing areas such as the use of advanced computing sciences (in the fields of big data, genomics and artificial intelligence, for example)"*

Although several countries have launched interim guidance on telemedicine, a sustainable telehealth ecosystem would need to take into consideration standards, interoperability, and regulatory frameworks (see [4-6] for further details).

In the subsequent sections, we provide a brief historical overview of telemedicine, with respect to the periods prior to and during the COVID-19 pandemic, in some of the participating countries, particularly India. We then discuss the current status, with respect to both policy frameworks and actual practice in some of the countries represented herein. We further elaborate on the perspectives, as described by different panelists, and suggest the way forward for the adequate adoption and diffusion of telehealth in countries of South Asia and beyond.

Table 1. Panelists and their affiliations.

Name and role	Country	Affiliation
Ashvini Goel (Chair)	India	Telemedicine Society of India (TSI), Lucknow
Alexander Thomas (Cochair)	India	Association of Healthcare Providers India (AHPI), New Delhi
Oommen John (Moderator-1)	India	George Institute for Global Health, University of New South Wales, New Delhi, and Prasanna School of Public Health, Manipal Academy of Higher Education, Manipal
A Thanga Prabhu (Moderator-2)	India	St Johns Health Innovation Foundation, Bengaluru
Sunil Shroff (Panelist)	India	Madras Medical Mission, Chennai
Fazilah Allaudin (Panelist)	Malaysia	Planning Division, Ministry of Health
Chaminda Weerbaaddana (Panelist)	Sri Lanka	Ministry of Health
Dari Alhuwail (Panelist)	Kuwait	Information Science Department, Kuwait University, and Health Informatics Unit, Dasman Diabetes Institute
Udaya Koirala (Panelist)	Nepal	Telemedicine Society of Nepal, Kathmandu
JA Jayalal (Panelist)	India	Indian Medical Association, New Delhi
Patricia Codyre (Panelist)	SEARO[a], World Health Organization	Digital Health and Innovation, SEARO, New Delhi
Andy Bleaden (Panelist)	European Union	European Connected Health Alliance (ECHAlliance), United Kingdom
SN Sarbadhikari (Panelist)	India	George Institute for Global Health, New Delhi
Shubnum Singh (Panelist)	India	Confederation of Indian Industries, National Healthcare Council, New Delhi
Shuchin Bajaj (Panelist)	India	Ujala Cygnus Healthcare Services, New Delhi

[a]SEARO: World Health Organization Regional Office for South-East Asia.

Brief Historical Overview

Pre-COVID-19 Period

Telemedicine has been helping family physicians by giving them easy access to specialized physicians and helping them in the close monitoring of patients. Various types of telemedicine services such as store and forward, and real-time, remote, and self-monitoring provide various educational, health care delivery and management, disease screening, and disaster management services across South Asian countries. In India, telemedicine had been traditionally led by some like-minded and passionate health care experts from diverse backgrounds and organizations, who, in their quest to bridge the humongous "health care divide" in India, decided to utilize communication technology for the provision of quality health care to care-seekers in underserved and difficult-to-reach areas of the country. They were ably facilitated by the Indian Space Research Organization with offer of their satellites for the purpose, since 2001. In India, several public and private telemedicine projects have already been in place, although fragmented and at small scales [7].

In Nepal, rural and remote health centers have been connected through a telemedicine network for specialist consultation, although issues such as electricity connection and network connectivity have hindered the widespread adoption and diffusion [8].

Sri Lanka has been promoting teleconsultations in response to emergency situations following the tsunami in 2004 [9].

During (and After) COVID-19

The Indian government released telemedicine practice guidelines [10] soon after the global lockdown in 2020 in response to the COVID-19 pandemic to promote teleconsultations. Free teleconsultations are being offered through the government-sponsored *e-sanjeevani OPD* telemedicine platform [11]. The government also partnered with many private organizations to promote teleconsultations.

There is compelling evidence [12] to suggest that telehealth may have a significant effect to advance the health care of the future. Nevertheless, the feasibility and application of telehealth in resource-constrained settings and low- and middle-income countries must be established to avail its potential and transform health care for the global population. As telehealth is advancing rapidly, a global consensus is absolutely necessary for definitions, boundaries, protocols, monitoring, and evaluation, as well as to ensure data privacy.

Telemedicine adoption had accelerated because of COVID-19. From March 2020 onward as the number of cases increased, a WhatsApp group was first used to help prevent the spread of disease. A small telemedicine platform was used to reach more patients. State government accommodations though legal cover came much later. Within 6-7 months, 10,000 volunteers came forward to help without salary or recognition. Real-time triaging of COVID-19 cases has been accomplished in 16 states. Teletriaging has been performed to avoid panic. Fake news and the spread of misinformation (infodemic) also need to be addressed [13,14]. During the onset of the pandemic in India,

less tests were conducted, and sometimes the reports of COVID-19 tests were not shared with the patients in a timely manner [15]. Social isolation, home quarantine, or institutional isolation were implemented and patients were appropriately advised. Plasma use went into the black market and donors were difficult to find initially, although this was solved over time. No charges were collected from users, and some organizations had been running COVID-19 response teams with grants only.

Experts also stressed the importance in addressing interoperability needs beyond technological aspects. Interoperabilities for human and institutional factors such as culture, governance, and policy also need to be kept in mind. For addressing these issues, changes in management principles have to be applied judiciously and continuously. The National Digital Health Mission and the Swasth alliance have come together to help manage COVID-19. The novel coronavirus (SARS-Cov-2) has globally acted as the chief transformation officer, causing massive digital disruption, especially for the health and education sectors. Learning from each other, we should be able to address the bigger problem. Digital health literacy is also badly needed. Digital determinants of health should be addressed [16]. Catching them at a young age would make health care workers more digitally savvy. The capacity building of human resources for health needs to be implemented in an ethical manner to enhance patient safety [17].

In Nepal, the government actively promoted the use of information and communications technology during the pandemic by offering several online consultation apps, and telemedicine practice guidelines were released [18].

In Sri Lanka, a robust primary health care delivery model along with a strong telecommunication network supported health care delivery during COVID-19 [9].

Current Status

Policy

In India, the Telemedicine Practice Guidelines 2020 were formally notified in May 2020 [10]. Although various other countries have been trying to promote telemedicine since the turn of the millennium, digital disruption enforced by the COVID-19 pandemic has hastened and streamlined these efforts. In general, all countries in this region have been supportive and encouraging toward the adoption of telehealth across the continuum of health care delivery.

The oil-rich Gulf Region, including the state of Kuwait, has made huge strides in the adoption of digital health solutions, including telehealth, electronic health records, laboratory information systems, picture archiving and communications systems, radiology information systems, and health information exchange in some countries [19]. However, challenges exist with medical terminologies and adoption of various standards for these digital solutions. With respect to telehealth specifically, it was pointed out that experiences shared by experts from normative agencies such as the World Health Organization (WHO) [20-22] and other groups such as the TSI should be shared with the community for learning purposes. One utterly important fact that was stressed is the safety of digital technology, which is paramount, while keeping patients or consumers of health services in the center of care and empowered to play a key role. The discussion also focused on the importance of establishing legal and ethical frameworks that are respectful of various cultures throughout the care continuum.

Actual Practice

Experts from Malaysia [23] stressed upon the fact that rather than highlighting new technologies only, people and processes need to come together. Telemedicine is only a medium. Privacy and confidentiality are important. Patients' and consumers' digital rights must be respected.

In Sri Lanka, physicians are trained at University of Colombo with an MSc in health informatics, which has given rise to a talent pool that has been very useful to deploy telemedicine. Training of health care workers [24] was essential for success. Chaminda Weerbaaddana from the Ministry of Health, Sri Lanka, stated:

> *Sri Lankans have access to a primary health care organization within a very short distance irrespective of their geographical location of residence. As such, before the pandemic, provider to client telehealth was not seen as a priority. Despite lack of need for telehealth services due to geographical reasons, there was a demand for such services due to diseases associated with stigma, where people would prefer to maintain anonymity when seeking services.*

Legal cover is essential to practice telemedicine. Guidelines have already been put in place by the Ministry of Health. More than 200 physicians have been trained in telemedicine. Two vendors have been identified and services are offered free to the public. Training was conducted online. Thus, Sri Lanka has proven that health care can be delivered via telemedicine, as has been done for COVID-19. Training care providers along with private players, especially with respect to ethics and security issues, is being undertaken.

Although Kuwait [19] is a small nation, it has a huge diversity in its population across socioeconomic status, language, and culture. This diversity needs to be considered when trying to build a standardized telemedicine practice, irrespective of service provider or consumer background. Dari Alhuwail, from Kuwait, stated:

> *Telemedicine should be integrated with national health strategies and those investments need to be made not only in equipment, but in training the workforce and ensuring continuing support. All digital health solutions need to be humanized and a dialogue amongst all stakeholders to tailor the solutions to serve them all is essential.*

It was also suggested that digital health should be part of medical and health sciences educational curricula, and potentially even in the general and higher education systems where consumers of health services can understand how to best leverage these tools and play a more active role in their own health care.

In Nepal, new technology has to be adopted. The human factor may be one of the major barriers for the speedy development and use of telemedicine. Doctor-to-doctor and doctor-to-patient consultations are quite different. Based on a surgeon's experience, it was stated that cameras that are used to cover remote surgery have been found to be useful. Demystification of telemedicine technology should be obtained through vigorous training for all levels of human resources involved. A structured telemedicine curriculum should be introduced to formal medical and information technology education at different levels. Continued technical support at remote sites is necessary for continued service and to avoid unnecessary frustration.

The WHO South East-Asia Regional Office has shared WHO digital health guidelines and specific guidance on telemedicine implementation that have been developed in consultation with member states [20,21]. WHO guidelines aligned to sustainable development goals (SDGs) and digital health platforms for tracking progress on health-related SDGs are also under implementation to track triple-billion targets. A digital health implementation handbook was recently released during the virtual World Health Assembly. These frameworks provide the building blocks for telemedicine. COVID-19 has served as a gentle push to take health care online. Following these recommendations, judicious roadmaps are a key to success. Patient centricity is essential. Improving eHealth awareness is needed. Immediate response to the current pandemic is essential, but we also need to learn from it quickly. Human capacity needs to be built up, and digital tools and telehealth should be leveraged for capacity building.

In India, industry bodies such as the Federation of Indian Chamber of Commerce and Industries and Confederation of Indian Industries, in collaboration with the Health Sector Skills Council, could contribute to the mainstreaming of telehealth in India.

The Indian Medical Association (IMA) is in favor of the widespread use of telehealth in India. India has 1,062,398 modern medicine practitioners registered with medical councils as of December 31, 2017 [22]. Sensitizing and educating such a large number of doctors in India can be a daunting task. To address the issue, a telehealth training course called "Train to Practice" was designed by the TSI [21]. The TSI initiated this training for registered medical practitioners (RMPs) within 2 weeks of the guidelines having been passed. The volunteer members of the society (including some coauthors of this article) trained close to 3000 doctors and sensitized another 25,000 doctors within the next 6 months [25]. This course is now available online for all RMPs [26].

Even in Europe, the value of health care ecosystems is now being understood in the context of the vulnerabilities that the pandemic has exposed. Health and social care players are coming together and silos are being broken. With a transforming health care delivery system, new economic opportunities are emerging. The European Connected Health Alliance (ECHAlliance) shared how connecting the dots [27,28] can help to scale up innovation built in Australia to be rolled out across Scotland and then to Ireland, which can then be showcased back to Australia where they have shared their best practice in a proven deployment. Andy Bleaden from ECHAlliance stated:

> *Our ECHAlliance Ecosystems connected the dots in health care during the pandemic taking solutions from one country and adopting them nationwide in another as the NHSNearMe program in Scotland and then offering this adoption back to the Australian health care market who had not seen it implemented at scale.*

Perspectives

Textbox 1 summarizes the current status of telehealth implementation according to the discussant countries, along with the next steps and barriers to implementation.

Textbox 1. Next steps and barriers to telehealth implementation.

India

Barriers

- Privacy and security
- Perceived ease of use
- Language barriers
- Fraud and abuse
- Questionable quality of care
- Perceived usefulness/preference for face-to-face consultation
- Shortage of tech-savvy workforce
- Inappropriate behavior by patients
- Digital divide
- Institutional, cultural, and governance issues
- Medical errors

Next steps

- The Personal Data Protection (PDP) Bill, currently tabled in the Parliament, may soon be passed to give robust directions for policy and implementation
- Appropriate capacity building for human resources for health and for raising awareness for patients may be undertaken
- With the imminent 5G connectivity in India, the digital divide is likely to be reduced
- Health data literacy and digital health literacy need to be encouraged across the entire health professional education environment
- The National Digital Health Mission is likely to ensure the smoother adoption of digital health

Sri Lanka

Barriers

- Inadequate training of health care providers
- Privacy and security
- Ethical issues

Next steps

- Training the care providers along with private players, especially for ethics and security issues, is being undertaken

Kuwait

Barriers

- Huge diversity in the population across socioeconomic status, language, and culture
- Inadequate training of health care providers
- Lack of user friendliness of digital health solutions

Next steps

- Telemedicine needs to be integrated with national health strategies
- Investments need to be made in training the workforce and ensuring continuing support
- All digital health solutions need to be humanized and a dialogue among all stakeholders to tailor the solutions to serve them all is essential

Malaysia

Barriers

- People and processes need to come together
- Privacy and confidentiality are important
- Patients' and consumers' digital rights must be respected

Next steps

- The Malaysian Medical Council Advisory on Virtual Consultation (2020) defined the clinical, ethical, legal, technical, and operational aspects of telemedicine for health practitioners

Nepal

Barriers

- The human factor may be one of the major barriers for the speedy development and use of telemedicine; doctor-to-doctor and doctor-to-patient consultations are different
- Regular technological support, particularly at remote sites, is often unavailable

Next steps

- Demystification of the telemedicine technology must be obtained through vigorous training for all levels of human resources involved
- A structured telemedicine curriculum should be introduced to formal medical and information technology education at different levels
- Continued technical support at remote sites is necessary for continued service and to avoid unnecessary frustration

European Union

Barriers

- Underestimating the value of health care ecosystems
- Health care and social workers working independently
- Doubts regarding scale and economic outcomes

Next steps

- The value of health care ecosystems is now being understood in the context of the vulnerabilities that the pandemic has exposed
- Health and social care players are coming together and silos are being broken
- With a transforming health care delivery system, new economic opportunities are emerging. The European Connected Health Alliance (ECHAlliance) is connecting the dots to help scale up innovation built in Australia to be rolled out across Scotland, followed by Ireland, and then showcased back to Australia

Conclusions

All panelists emphasized the need for active encouragement toward the adoption and diffusion of digital health in general and of telehealth in particular. All stakeholders, ranging from governmental policymakers to common citizens, have to come together to build trusting partnerships to realize the advantages offered by telehealth. The panelists emphasized the importance of scientific research and evidence-based policy recommendations to improve the use and adoption of effective, efficient, and safe digital health solutions. Various telemedicine applications were referred to, such as those used for diagnosing surgical site infections [29] as well as for rural use in the context of the COVID-19 pandemic [30].

After summarizing the proceedings, the Chair suggested that we can develop appropriate courses for telehealth and introduce them as a part of the curriculum for all health professionals, including physicians, surgeons, dentists, nurses, and allied health care professionals. Like-minded organizations such as the Association of Healthcare Providers of India, TSI, IMA, and reputed academic institutions can come together and influence the regulatory commissions for adopting these recommendations.

Conflicts of Interest

None declared.

References

1. Telemedicon 2020. 2020. URL: https://tsi.org.in/telemedicon-2020/ [accessed 2022-01-28]
2. Sarbadhikari S, Sarbadhikari SN. The global experience of digital health interventions in COVID-19 management. Indian J Public Health 2020 Jun;64(Supplement):S117-S124 [FREE Full text] [doi: 10.4103/ijph.IJPH_457_20] [Medline: 32496240]
3. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2022-01-28]
4. Gudi N, Konapur R, John O, Sarbadhikari SN, Landry M. Telemedicine supported strengthening of primary care in WHO South East Asia region: lessons from the COVID-19 pandemic experiences. BMJ Innov 2021 May 20;7(3):580-585. [doi: 10.1136/bmjinnov-2021-000699]

5. Intan SM, Defi IR. Telemedicine guidelines in South East Asia-a scoping review. Front Neurol 2020;11:581649. [doi: 10.3389/fneur.2020.581649] [Medline: 33519669]
6. Basu A, Kuziemsky C, de Araújo Novaes M, Kleber A, Sales F, Al-Shorbaji N, et al. Telehealth and the COVID-19 pandemic: international perspectives and a health systems framework for telehealth implementation to support critical response. Yearb Med Inform 2021 Aug 21;30(1):126-133 [FREE Full text] [doi: 10.1055/s-0041-1726484] [Medline: 33882598]
7. Chellaiyan V, Nirupama A, Taneja N. Telemedicine in India: where do we stand? J Family Med Prim Care 2019 Jun;8(6):1872-1876 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_264_19] [Medline: 31334148]
8. Siddiquee N, Poudyal A, Pandey A, Shrestha N, Karki S, Subedi R, et al. Telemedicine in resource-limited setting: narrative synthesis of evidence in Nepalese context. Smart Homecare Technol Telehealth 2020 Jan;6:1-14. [doi: 10.2147/shtt.s227854]
9. Kulatunga G, Hewapathirana R, Marasinghe RB, Dissanayake VHW. A review of telehealth practices in Sri Lanka in the context of the COVID-19 pandemic. Sri Lanka J Bio-Medical Informatics 2020 Oct 20;11(1):8. [doi: 10.4038/sljbmi.v11i1.8090]
10. Board of Governors in supersession of the Medical Council of India. Telemedicine practice guidelines 2020. Ministry of Health and Family Welfare, Government of India. 2020 May 14. URL: https://www.mohfw.gov.in/pdf/Telemedicine.pdf [accessed 2022-01-28]
11. Dinesh K, Guleria KSDS, Patiyal N, Negi AK, Kanwar V. Utilization of outpatient eSanjeevani National Teleconsultation Service during COVID- 19 pandemic in a public healthcare institution in North India. Ind J Pharmacy Pharmacol 2021 Jan 15;7(4):265-269. [doi: 10.18231/j.ijpp.2020.045]
12. Doraiswamy S, Abraham A, Mamtani R, Cheema S. Use of telehealth during the COVID-19 pandemic: scoping review. J Med Internet Res 2020 Dec 01;22(12):e24087 [FREE Full text] [doi: 10.2196/24087] [Medline: 33147166]
13. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659 [FREE Full text] [doi: 10.2196/19659] [Medline: 32558655]
14. Purnat TD, Vacca P, Czerniak C, Ball S, Burzo S, Zecchin T, et al. Infodemic signal detection during the COVID-19 pandemic: development of a methodology for identifying potential information voids in online conversations. JMIR Infodemiology 2021 Jul 28;1(1):e30971 [FREE Full text] [doi: 10.2196/30971] [Medline: 34447926]
15. Biswas S. Coronavirus: why is India testing so little? BBC News. 2020 Mar 20. URL: https://www.bbc.com/news/world-asia-india-51922204 [accessed 2022-01-28]
16. An L, Bacon E, Hawley S, Yang P, Russell D, Huffman S, et al. Relationship between coronavirus-related eHealth literacy and COVID-19 knowledge, attitudes, and practices among US adults: web-based survey study. J Med Internet Res 2021 Mar 29;23(3):e25042 [FREE Full text] [doi: 10.2196/25042] [Medline: 33626015]
17. Sarbadhikari SN, Pradhan KB. The need for developing technology-enabled, safe, and ethical workforce for healthcare delivery. Saf Health Work 2020 Dec;11(4):533-536 [FREE Full text] [doi: 10.1016/j.shaw.2020.08.003] [Medline: 32839672]
18. Prasain K. With all hospital beds occupied, telemedicine and digital health services on the rise. Kathmandu Post. 2021 May 08. URL: https://kathmandupost.com/money/2021/05/08/with-all-hospital-beds-occupied-telemedicine-and-digital-health-services-on-the-rise [accessed 2022-01-28]
19. Buabbas A, Alshawaf H. Stud Health Technol Inform 2019 Aug 21;264:1644-1645. [doi: 10.3233/SHTI190576] [Medline: 31438272]
20. 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(2):153. [doi: 10.4258/hir.2012.18.2.153]
21. WHO Guideline: Recommendations on digital health interventions for health system strengthening. World Health Organization. 2019. URL: https://apps.who.int/iris/bitstream/handle/10665/311941/9789241550505-eng.pdf [accessed 2022-01-28]
22. Government of India Minister of State in the Ministry of Health Family Welfare: Lok Sabha. 2018. URL: http://164.100.47.194/loksabha/Questions/QResult15.aspx?qref=69322&lsno=16 [accessed 2022-01-28]
23. Koh D. COVID-19: Malaysia's pandemic approaches and its impact on telehealth. Healthcare IT News. 2020. URL: https://www.healthcareitnews.com/news/asia-pacific/covid-19-malaysia-s-pandemic-approaches-and-its-impact-telehealth [accessed 2022-01-28]
24. Siribaddana P, Hewapathirana R, Jayatilleke A, Sahay S, Dissanayake V. Strengthening health systems through informatics capacity development among doctors in low-resource contexts: the Sri Lankan experience. WHO South East Asia J Public Health 2019 Sep;8(2):87-94 [FREE Full text] [doi: 10.4103/2224-3151.264852] [Medline: 31441443]
25. Shroff S. Feedback for Telemedicine Practice course. Telemedicine Society of India Tamilnadu Chapter. 2020. URL: https://tsi.org.in/learn/ [accessed 2022-01-28]
26. Shroff S, Puhan B, Dorairaj L, Agarwal M, Rajendran M, Modali R, et al. Education and training for ethical practice of telemedicine for registered medical practitioners in India. Telehealth Med Today 2021 Apr 23;6(2):254 [FREE Full text] [doi: 10.30953/tmt.v6.254]

27. Telemedicine in a week! Pop-Up Ecosystem 4th Webinar. ECH Alliance. URL: https://echalliance.com/telemedicine-in-a-week/ [accessed 2022-01-28]
28. NHS in Scotland is now using Near Me video consultation with care homes. ECH Alliance. URL: https://echalliance.com/nhs-in-scotland-is-now-using-near-me-video-consultation-with-care-homes/ [accessed 2022-01-28]
29. Sandberg CEJ, Knight SR, Qureshi AU, Pathak S. Using telemedicine to diagnose surgical site infections in low- and middle-income countries: systematic review. JMIR Mhealth Uhealth 2019 Aug 19;7(8):e13309 [FREE Full text] [doi: 10.2196/13309] [Medline: 31429414]
30. Chu C, Cram P, Pang A, Stamenova V, Tadrous M, Bhatia RS. Rural telemedicine use before and during the COVID-19 pandemic: repeated cross-sectional study. J Med Internet Res 2021 Apr 05;23(4):e26960 [FREE Full text] [doi: 10.2196/26960] [Medline: 33769942]

Abbreviations

ECHAlliance: European Connected Health Alliance
GSDH: Global Strategy for Digital Health
IMA: Indian Medical Association
RMP: registered medical practitioner
SDG: sustainable development goal
TSI: Telemedicine Society of India
WHO: World Health Organization

Edited by A Mavragani; submitted 28.05.21; peer-reviewed by S Doraiswamy, L Suppan, M Ji; comments to author 03.07.21; revised version received 12.08.21; accepted 22.12.21; published 08.02.22.

Please cite as:
John O, Sarbadhikari SN, Prabhu T, Goel A, Thomas A, Shroff S, Allaudin F, Weerabaddana C, Alhuwail D, Koirala U, Johnrose J, Codyre P, Bleaden A, Singh S, Bajaj S
Implementation and Experiences of Telehealth: Balancing Policies With Practice in Countries of South Asia, Kuwait, and the European Union
Interact J Med Res 2022;11(1):e30755
URL: https://www.i-jmr.org/2022/1/e30755
doi:10.2196/30755
PMID:35133279

Review

Approaches to Determine and Manage Sexual Consent Abilities for People With Cognitive Disabilities: Systematic Review

Shaniff Esmail[1*], PhD; Brendan Concannon[2*], MSci

[1]Department of Occupational Therapy, University of Alberta, Edmonton, AB, Canada

[2]Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

[*]all authors contributed equally

Corresponding Author:
Shaniff Esmail, PhD
Department of Occupational Therapy
University of Alberta
Corbett Hall
8205 - 114St 2-64
Edmonton, AB, T6G 2G4
Canada
Phone: 1 (780) 492 0397
Fax: 1 (780) 492 4628
Email: shaniff.esmail@ualberta.ca

Abstract

Background: This review focused on how sexual consent ability was determined, managed, and enhanced in people with cognitive disabilities, with the aim of better understanding the recurring themes influencing the design and implementation of these approaches. If a person's consensual ability becomes compromised, owing to either an early or late-onset cognitive disability, the formal systems involved must establish plans to balance the individual's rights and restrictions on sexual expression. This review identified these plans, focusing on how they promoted the intimacy rights of the individual.

Objective: This study aims to identify approaches that determine sexual consent ability in people with cognitive disabilities, identify the means of managing and enhancing sexual consent ability in people with cognitive disabilities, and note the recurring themes that influence how these approaches and management systems are designed and implemented.

Methods: A systematic literature review was performed using EBSCOhost (Social Gerontology, CINAHL Plus, MEDLINE, and SocINDEX), Embase, PsyInfo, and Scopus to locate reports on terms expanded on sexual consent and cognitive disability.

Results: In all, 47 articles were identified, featuring assessment practices, legal case studies, and clinical standards for managing sexual consent capacity in people with cognitive disabilities. A total of 8 studies (5/8, 63% qualitative and 3/8, 38% quantitative) were included out of the 47 articles identified. Approaches for determining sexual consent included functional capacity and person-centered, integrated, and contextual approaches. Management of sexual consent ability included education, attitude, and advanced directives and support networks. The recurring themes that influenced these approaches included the 3 legal criteria of consent, American Bar Association and American Psychological Association Model, Lichtenberg and Strzepek Instrument, Ames and Samowitz Instrument, Lyden approach, Mental Capacity Act of 2005, and Vancouver Coastal Health Authority of 2009.

Conclusions: Determining sexual consent takes a holistic approach, with individuals judged in terms of their adaptive abilities, capacities, and human rights. The attitudes of those using this holistic approach need to be balanced; otherwise, the sexual rights of assessed people could be moved either in favor or against them. The ideal outcome, after person-centered considerations of those living with cognitive disabilities includes the people themselves being involved in the process of personalizing these approaches used to facilitate healthy intimate relationships.

(Interact J Med Res 2022;11(1):e28137) doi:10.2196/28137

KEYWORDS

sexual consent; capacity; disability; sexual expression; dementia; ethics; long-term care

Introduction

Background

Defining Cognitive Disability, Sexuality, and Consent

Cognitive disabilities are defined as long-term mental impairments, including those of intellectual and developmental order. The terms *cognitive* and *intellectual* are often used interchangeably, with intellectual disabilities defined as a limitation in academic functioning based on standardized intelligence tests and IQs, associated with IQ scores <75; limitations in learning behavior to one or all three skill types—conceptual skills, social skills, and practical skills; and manifestation of the disability before the age of 18 years [1,2]. Noncognitive mental conditions, including psychiatric and psychosocial conditions, are associated with anxiety, mood, and personality disorders.

Sexuality is a holistic concept encompassing sex, gender identity, orientation, eroticism, intimacy, and reproduction [3]. It is influenced by physical, psychological, social, financial, cultural, legal, historical, and spiritual factors [3]. Sexuality includes sexual expression. The central debate surrounding the rights of people with cognitive disabilities, who wish to express their sexuality, may affect the balance between harm reduction and free sexual expression [4]. If a person living with a cognitive disability is proven to have a reduction in consensual capacity, the protection versus empowerment paradox may begin to emerge [5,6]. The legal, clinical, or ethical system needs to find a balance between the 2 competing interests; protect the person from sexual abuse, by restricting their sexual expression; or allow them to express their sexuality, but in limited capacity as a safety measure [5]. An important factor affecting the resolution of the protection versus empowerment paradox is the degree to which an individual can demonstrate their capacity to consent in a sexual relationship.

In the United States, the legal definition of consent is rooted in the 3 legal criteria of consent as reported by Stavis [7] in 1991. A placeholder definition of consent requires that a person communicates a "knowing, intelligent, and voluntary agreement to engage in a given activity" [8,9]. Assessments that measure consensual ability for a sexual relationship are often based on a person being able to satisfy all three of the following criteria [3,7,10,11]:

1. Knowledge—recognition of the other person in the relationship, including who, what, where, and when and safety aspects of the sexual activity in question, such as the ability to identify body parts.
2. Intelligence—also known as rationality or understanding, which includes awareness of potential risks (pros and cons) of sexual engagement, appropriateness, consequences, correct familiarity of partner identity, and the ability to discriminate among fantasy, reality, lies, and truth.
3. Voluntariness—decisional capability to engage or refrain from sexual activity and the ability to take self-protective measures against abuse and exploitation or other unwanted advances. This includes the ability to say "no," either verbally or nonverbally and the ability to remove oneself from the situation when either they or their partner indicates stopping sexual behavior.

These 3 legal criteria of consent are quite controversial, because thresholds vary from basic to complex levels of acceptability, depending on differences in state laws [12]. People with cognitive disabilities may be unable to fulfill requirements in one region of the United States, yet their ability to demonstrate consent could be acceptable in another region. Outside of the United States, consent definitions may differ among countries such as the United Kingdom, Australia, Ireland, and Canada. In Canada, criminal offenses, including those regarding consent, are governed by the *Criminal Code* [13]. Criminal law powers are under the exclusive jurisdiction of the federal government; however, the fulfillment of these laws is often handled by provincial regulations (eg, dealing with sexual assault on university campuses) [14]. The *Criminal Code* defines consent with a capacity element, which checks if an individual understands the nature of sexual activity, identity of those involved, and their ability to communicate the choice itself [13].

The capability of people to satisfy sexual consent criteria is often determined by either medical professionals or neuropsychological experts in the judicial system. Common paradoxes have emerged, owing to the philosophical arguments surrounding people's ability to give consent. These paradoxes include whether people can demonstrate rudimentary versus contextual understanding of the sexual relationship [15] and the degree of flexible versus inflexible behavioral allowances in such a relationship [16]. Power dynamics, regarding those who are legally able to discern a person's consensual ability, are met with arguments of feminism, ableism, and disability rights movements [17,18]. The 3 legal criteria of consent have been accused of failing to consider an individual's values, culture, and life history [10,19]. Outside the United States, the United Kingdom and Wales follow the Mental Capacity Act 2005, which has its own criticisms. Although the Mental Capacity Act has similar themes to the 3 legal criteria of consent, it too is criticized because "its best-interests approach is paternalistic" [20]. The Mental Capacity Act does not exactly guarantee the rights of an individual, instead only working *if practicable to do so* [21]. Some experts argue that an individual's sexual preference is a form of personal expression, not always systematic or organized, including the weighing of risks and benefits, unlike the 3 legal criteria of consent by the medical and judicial systems in the United States [10,19]. Consent capacity is considered a state instead of a trait, meaning it is expected to change over time [11], and it must be determined in the present moment: not a decision made ahead of time [22]. Assessments of a person's ability to demonstrate capacity in 1 or all 3 prongs of consent can be determined by questionnaires such as the Mini-Mental Status Exam (MMSE) to determine rationality or the Tool for the Assessment of Levels of Knowledge Sexuality and Consent [11]. However, the use of these assessments is controversial because of the following:

1. Rudimentary requirements that check for consent capacity may fail to understand the contextual reason to *why* a person with a cognitive disability may wish to consent to sex [18].

2. Complex knowledge of consent may have assessments and protocols that are too difficult for even the general population to pass [23].
3. The *Ice Cream Reference*—a person with a cognitive disability is expected to follow a rigid medicalized or judicial process in order demonstrate their consent capacity; however, decisions about sexual relationships are arguably more related to selecting a flavor of ice cream than, say, a life-or-death surgical treatment [24].
4. Assessments of consent capacity often place the burden of proof on the person with a cognitive disability rather than putting the onus on others to prove otherwise. Having individuals provide predetermined comfort with various levels of intimacy carries an unfair standard, because even the general population may not know what levels of intimacy they are comfortable with before engaging in such behaviors [4].

There is no clear definition, criteria, or standard for determining a person's sexual consent capacity [5,9].

Human Rights and Sexual Abuse

In the late 1960s, the United States Supreme Court declared constitutional rights for people with cognitive disabilities, who were cared for under the powers of the state governments. These constitutional rights were created to protect vulnerable people from harm related to sexual exploitation and abuse, while also upholding their rights to sexual expression. These rights include several categories, including those related to family matters [7] and sexual self-determination [4]. Although it may seem obvious that people have default rights to privacy, sexuality education, and freedom of choice for sexual expression [3,25], these rights can be restricted by either informal or formal control systems. Informal control systems such as civil liability [26], immutable family policy [22], and residential policy [4,27,28] may interfere with a resident's rights to sexual expression, whereas restrictions from formal control systems may be decided by clinical, ethical, and legal issues [26]. The central reason to why these systems may restrict a person's right to sexual expression is based on the theme of consent [1,15,18,29]. New York Penal Law Section 130 states that a lack of consent is an element of every conceivable sexual offense, as written in the article, and adults in a sexual relationship must all be consenting [30]. Since 2012, the United Nations has moved toward an equalization stance on the sexual rights of people living with cognitive disabilities. Article 12 of the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) states that people with cognitive disabilities reserve equal rights to legal capacity as all other people within all aspects of their lives, including the rights to intimate relationships [31]. Canada has accepted and ratified Article 12; however, with reservation, resulting in continued use of substitute decision makers (SDMs) to assist those living with cognitive disabilities in Canada [32]. The United States has been hesitant to fully accept and ratify the UNCRPD statement on the rights of people with disabilities [33]. It is important for legal systems to establish the components of human rights, because this increases the awareness of any potential violations.

Sexual abuse occurs when one person forcefully or covertly performs nonconsensual sexual acts, including touching, kissing, oral sex, and anal or vaginal intercourse [3,34]. Sexual abuse may also involve threatening, coercing, tricking, or manipulating another person into unwanted sexual contact or into such contact in which the other person does not have the capacity to consent [11]. According to the *Criminal Code* of Canada, an alleged case of sexual assault requires a juridical system to check if people take reasonable steps to ascertain consent [13]. Other elements are also checked for, such as the presence of physical force, threats, underaged individuals, fraud, sexual intentions and motivations, recklessness, incapacitation, and chemical impairment and those in positions of authority or trust [13].

People with cognitive disabilities have a greater risk of being sexually abused [3,25]. Although the statistics on sexual abuse are difficult to determine, one report predicts that 39% to 68% of female children and 16% to 30% of male children with cognitive disabilities will be sexually abused before they are aged 18 years in North America [35,36]. After interviewing over 40,000 abused victims in Israel, including children up to the age of 14 years and ranging from minor to severe cognitive disabilities, the incidence of sexual abuse was found to be consistent with the previously stated percentages [37]. A survey of over 5000 adult women living with cognitive disabilities in North Carolina reported that 48% of sexual assaults were committed by people who were currently or previously in an intimate relationship [38]. A 15-year longitudinal study in Ireland determined the following statistics after 118 proven and confirmed episodes of sexual abuse [39]:

1. Most of the perpetrators were men.
2. A percentage (n=66) of the perpetrators had a cognitive disability.
3. In all, 24% (n=28) of the perpetrators were relatives.
4. In all, 9% (n=11) of the perpetrators were agency staff members.
5. In all, 8% (n=9) of the perpetrators were familiar people.
6. The remaining perpetrators were either volunteers, strangers, or unknown.

Society Attitudes on the Sexuality of People With Cognitive Disabilities

The *zeitgeist* to uphold and safeguard the sexual rights of persons with cognitive disabilities may differ from past ideals, which were weighted toward the protective sides of the protection versus empowerment paradox [5,40]. Societal views on the amount and types of sexual expression that people with cognitive disabilities were expected to experience were driven by moral aesthetics, which are beliefs and morals that affect the general public's preference to accept certain behaviors while rejecting others. Thus, people with cognitive disabilities were historically denied the right to express their sexuality, because society may have considered them to be the following:

1. Hypersexual—*oversexed* people who were often seen as a threat to the gene pool and general public, owing to their excessive sexual behavior [8,11]. These people may have been identified as having *super human strength* sex drives [18]. Reported cases of older adults living with dementia may repeatedly approach partners for sex, after forgetting

they had sex earlier [41]. An emerging tendency toward public masturbation is a potential problem for older adults living with dementia [41]. Child masturbation can also be a common form of childhood sexual behavior, which is considered *developmentally normal*, unless inappropriate owing to public occurrence, excessiveness, or when the behavior causes injury [42-44].

2. Asexual—*eternal children* were often seen as potential sexual victims who were deemed to have a major difference in their chronological and mental ages [45] and assumed to not necessarily want sexual relationships or need sexuality education, because it may incite increased interest in the activity or the risk of abuse [3,18,46]. Therefore, some policies thought better to keep *Pandora's box closed* to reduce these risks, which actually increased the vulnerability of these people, due to the lack of education about those who might exploit them in the first place [25,47]. *Pillow Angels* are defined as people with cognitive disabilities who were thought to be incapable, or should be made incapable, of becoming adults and were removed from sexual relationships to be spared the dangers of sexuality, such as pregnancy and sexual exploitation [48,49].
3. Deviant—in the last 20 years of research on the well-being of people who are lesbian, gay, bisexual, transgender, or queer (LGBTQ), there is evidence of victimization among such sexual and gender minorities in both youth and adults [50]. Although initial perceptions of the North American society show a more open and tolerant view of the LGBTQ community, victimization rates and disparities have worsened since the 1990s [50-53]. LGBTQ older adults, living with cognitive disabilities, are often met with *pervasive stigma* by staff in long-term care (LTC) facilities, who have reported to feel disgust or panic, sometimes resorting to denial of such residents' sexuality [22,29,54,55].

The moral aesthetic to control how people with cognitive disabilities express their sexuality are bound to clinical, ethical, and legal issues [26]. There is a potential overlap among these issues.

Social acceptability struggles to find a balance between sexual acts that are safe versus unsafe, normal versus deviant, and legal versus illegal and what role sexual functioning has in the first place [34].

Clinical policies in a LTC facility could be undeveloped or inconsistent with those living with cognitive disabilities and their sexual expression, resulting in the facility facing repercussions if sexual expression is allowed to continue [26,56]. LTC facilities for people with cognitive disabilities default to a *protective care paradigm*, with staff and family members restricting such residents from sexually expressing their behaviors to reduce the risk of potential sexual abuse [29]. The result is that such residents may resort to *opportunistic moments of privacy to act on their sexual desire* [48,57], which may lead to unsafe sexual behaviors [56,58]. Alternatively, they could be affected by iatrogenic loneliness, which is a type of loneliness created by extensive long-term residence policies that prevent them from having privacy and intimacy, resulting in feelings of frustration and unhappiness [28,59]. LTC staff views on sexual expression differ according to the experience levels of staff members, with frontline staff being more accepting of such behaviors than the managers; however, gay residents are more likely to be restricted to such behaviors in general [29]. Ethical views, independent of those with cognitive disabilities, include moral and religious beliefs that others enforce regarding sexual behavior [26]. In theory, a LTC facility must support such a resident's rights to sexual autonomy; however, this obligation is abandoned once the administration, facility staff, or individual's family members oppose the behavior [4,22]. Fear of legal repercussions and public ridicule are potential reasons why such people's sexual interests are downplayed or avoided by family, caregivers, or long-term facility staff [23]. Overall, the community is capable of supporting the sexuality of people with cognitive disabilities, upholding attitudes of community inclusion and opportunity; however, personal belief systems are affected by societal attitudes and are what prevent caregivers from providing experiential guidance [34,60]. Negative attitudes, such as the eugenics movement, were perhaps too difficult for North American society to discard entirely [48,61]. Thus, society *morphed* them to the new era, resulting in *new-genics* or *neogenics* [62]. The intention of determining a person's capacity to consent to sex has remained a *plague* in the societal attitude to desexualize people with cognitive disabilities, under the moral esthetic to either protect such individuals from themselves or to protect the world from them [63].

Sterilization and Eugenics

Sterilization is the process of inhibiting a person's reproductive ability. It inflicts physical and moral injuries to those who do not consent to it [21,64]. The eugenics movement of the late 19th century led to an increase in nonconsensual sterilization practices, sometimes with the use of deception [21]. Eugenicists believed that the human race could be improved by practicing either positive or negative eugenics, which either encouraged the selective breeding of those with desirable traits or the prevention of *defectives* from having offspring [65]. In the early 1960s, 28 US states had sterilization protocols, some made compulsory and executed upon people with cognitive disabilities, without their consent [3].

The justification for sterilization was often influenced by the eugenics movement, which believed that *feebleminded* people would reduce the overall intelligence of the population, especially if they were allowed to reproduce [21]. It was believed that people with cognitive disabilities would threaten the *heritage of intelligence* [40,66,67]. In actuality, the eugenics idea to use selective breeding to control for inherited psychological traits was proven to be false [21,68]. Feeblemindedness was previously used as a *conveniently vague grouping*, used to classify those who were outside the obvious diagnostic labels such as schizophrenia [21]. Some countries sterilized those diagnosed as *feebleminded*; however, it was later realized that many of these people were actually affected by a lack of education [21]. In a report from India that surveyed nearly 20,000 women, higher levels of education increased the likelihood of modern contraceptive use over sterilization; however, the degree of cultural, socioeconomic, educational, and accessibility to modern contraceptives had a profound effect on choice [69]. In India, sterilization was more common in

women living in low socioeconomic classes, especially in socially disadvantaged women with low education levels [69]. Sterilization was sometimes ordered because of the protection versus empowerment paradox. Some systems believed that women with cognitive disabilities would fail to provide adequate care for any children they would have; thus, these potential mothers were prevented from leaving their institutions unless they agreed to become sterilized [70]. With the exception of reducing pregnancy and reducing bodily fluids, sterilization of people with cognitive disabilities was found to be ineffective in achieving any of its goals [71,72]. There is limited evidence available to support sterilization for the management of menstruation, with some experts agreeing that cases involving clinical control of menstrual bleeding are better handled by long-term contraceptive injections [73].

There have been major changes in legislation regarding the practice of nontherapeutic sterilization [21]. In Canada, a major case occurred in 1986. This *E (Mrs) versus Eve* case argued that court-ordered sterilization of people, living with cognitive disabilities, would be an infraction against their rights. The mother requested sterilization of her daughter, Eve, aged 24 years, to avoid the risk of pregnancy. After a contentious appeal, the request was denied. The Supreme Court of Canada ruled in favor of Eve, due to a lack of evidence suggesting that forgoing sterilization would have a detrimental effect on physical or mental health in Eve [74]. The choice to allow nontherapeutic sterilization of people with cognitive disabilities, in which a procedure would leave a person sterile despite having no life-threatening condition to begin with, was deemed a choice the courts could not safely exercise [21,74]. Later reports have claimed that some countries have implemented human rights protection to prevent nonconsensual sterilization practices; however, some countries have no such safeguards in place [21]. The following references contain additional information pertaining to the statistics and country policy on sterilization: Stein and Tepper [3], Rowlands and Amy [21], Braun et al [64], Tilley et al [70], Shea and Kevles [75], and Park and Radford [76].

Benefits of Healthy Sexual Expression

There are psychological and physical benefits of safe sexual expression. Improved self-esteem, cognitive functioning, social relationships, mood, and feelings of independence have been reported as potential benefits [77,78]. Sexual expression may reduce the risk of cancer and cardiovascular disease [77]. There are associations between sexual expression and weight loss, reduced risk of heart disease and stroke, and bolstered immune systems [79]. For older adults with cognitive disabilities, sexual expression has been reported to reduce sensitivity to pain, improve cardiovascular health, reduce the likelihood of depression and loneliness, and improve overall well-being [80]. Sexuality is *central to an individual's health and well-being* [4].

Aim of This Review

This review aims to uncover the used approaches of clinical, legal, and residential systems to determine and manage the sexual consent abilities of people with cognitive disabilities. Recurring themes influencing the shape of these approaches were also identified. Specific audiences for this review include human ecologists, sexuality experts and therapists, forensic neuropsychologists, occupational therapists, sexual educators, health care professionals, service providers, and caregivers.

Objectives

The objectives of this review are as follows:

1. Identify approaches used to determine sexual consent ability in people with cognitive disabilities.
2. Identify means of managing and enhancing sexual consent ability in people with cognitive disabilities.
3. Note the recurring themes affecting how such approaches and management systems are designed and implemented.

Methods

Research Question

This report presents a systematic review of the literature, based on consultation with human ecology and rehabilitation medicine experts, to create the following research question: What are the approaches for determining, managing, and improving sexual consent ability in people with cognitive disabilities?

Search Strategy

After discussing the research question with a university librarian, the following bibliographic databases were searched: EBSCOhost (abstracts in Social Gerontology, CINAHL Plus, MEDLINE, and SocINDEX), Embase, PsyInfo, and Scopus. The search strategy included a combination of subject headings and keywords to combine the concepts of consent in sexuality and cognitive disability. Textbox 1 lists the inclusion and exclusion standards for each article. The full search strategy is provided in Multimedia Appendix 1.

A total of 2 researchers performed the screening process for each article (BJC and recruited researcher, Lyndsay Pinder). Differences among the researchers in terms of accepted and rejected articles were resolved through discussion. All articles indicating topics of sexuality and consent within their titles or abstracts were reserved to complete the first pass of the search process (BJC and Lyndsay Pinder). For the second pass, all reserved reports from the first pass had their full texts screened to confirm the context of the subject (BJC, SE, and Lyndsay Pinder). The methodological quality of the reports featuring experimentation was not formally assessed. There were no data limits.

Textbox 1. Search criteria and terms.

Inclusion criteria

- Article stated a topic, discussion, or approach to determine the consent capacity of people with cognitive disabilities.
- All articles featuring qualitative, quantitative, legal, descriptive, and review reports were accepted.
- Reports were accepted in all languages and in article, dissertation thesis, review, or book format.

Exclusion criteria

- Topic was about physical disability or did not indicate a potential compromise in a person's consensual ability.
- The article briefly mentioned

  consent to sexuality

  or a similar phrase; however, further details were not provided.
- Conference papers, public opinions and non–peer-reviewed articles.

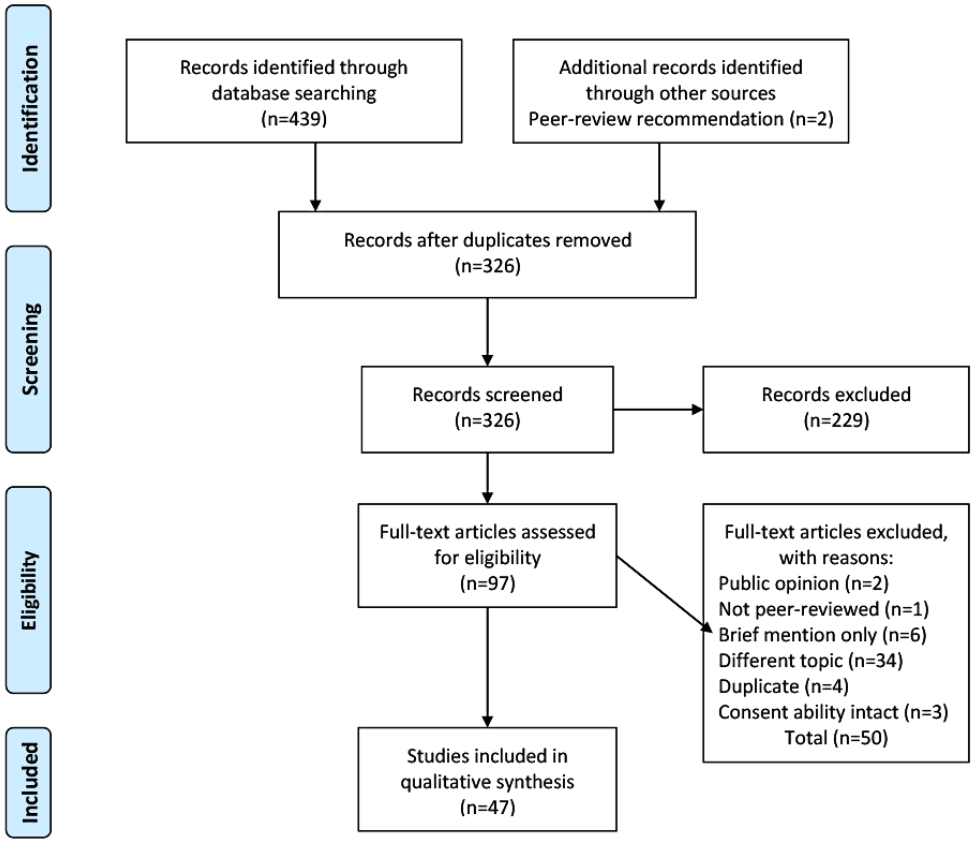
Search terms used

- ([sex* or intima*] adj10 [consent or consensual]) AND ([(intellectual* or mental* or cognitive*) adj4 (impair* or disab* or deficit*)] or long term care or longterm-care or nursing home* or alzheimer* or dementia or autis* or Down* Syndrome).
- These terms were entered into the databases mapped to the following fields: title, abstract, subject heading word, and keyword heading word.

Search Results

The search resulted in 439 articles being identified, of which 2 (0.5%) articles were recommended for inclusion in the peer-review process [12,81]. After the first pass, 22.1% (97/439) articles remained after the titles and abstracts were screened, and 26.2% (115/439) duplicates were removed. During the second screening pass, the full texts of 97 articles were screened to confirm eligibility (Figure 1). This resulted in a net total of 10.7% (47/439) articles being included in this review.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of search results [82].



Results

Overview

The 47 reports included in this review featured assessment practices, legal case studies, and clinical standards for managing sexual consent capacity in people with cognitive disabilities. Most reports were in the form of expert opinions (36/47, 77%). There were 8 studies (5/8, 63% qualitative and 3/8, 38% quantitative) included in this study. The qualitative studies included the following:

1. A survey with vignettes to check the ability of residential facility staff to properly identify safe or unsafe sexual behaviors (nonconsensual sexual behavior) in people with cognitive disabilities and respond accordingly [34].
2. A survey of members of the American Psychological Association to determine which criteria were considered the most important when determining sexual consent capacity in people with cognitive disabilities [9].
3. A survey to determine factors that increase the risk of SDMs to decide an *all-or-none* outcome for a person's sexual autonomy [83].
4. Semistructured interviews in residential mental health treatment facilities to determine what conceptualizes consent to sexual expression from the point of view of administrators, clinical staff, and former clients [84].
5. Semistructured interviews with directors of nursing to identify challenges in managing sexual expression [85].

Quantitative studies focused on educational interventions for the improvement of sexual consent ability in people with cognitive disabilities [40,86]; 1 study performed a validity measure to compare neuropsychological tests with the Sexual Consent and Education Assessment [45]. A total of 4 reports were dedicated to introducing a theme, which would later influence the approaches used to manage or enhance the sexual consent capacity of people with cognitive disabilities. Table 1 provides a summary of each theme, and Table 2 provides a brief description of each approach. The 2 approaches for determining or managing consent were peripheral to these themes [12,81]. For more information about the research studies, refer to Table 3.

Table 1. Themes affecting the approaches for determining sexual consent capacity in people with cognitive disabilities.

Theme	Key components	References
Lichtenberg and Strzepek Instrument	Interdisciplinary characteristics. Client is assessed (MMSE[a]), followed by a same-sex interview to determine these three main criteria: • Awareness of the relationship—patient aware of intent, partner identity, and intimacy comfort level. • Ability to avoid exploitation—patient behavior consistent with former beliefs and able to say no. • Awareness of potential risks—consequences of relationship and awareness of relationship duration. • Interview relayed to interdisciplinary team (nurses, occupational therapists, psychiatrists, etc).	[22,87-89]
3 legal criteria of consent	Legal characteristics; client is required to demonstrate ability in the following: • Knowledge—basic recognition of the other person, relationship, and sexual activity in question. • Intelligence—(rationality and understanding) aware of potential risks in the sexual relationship. • Voluntariness—ability to resist or stop the sexual activity and identify willingness to continue.	[3,5-7,9, 10, 19, 26, 40, 89-93]
Ames and Samowitz instrument	Legal and clinical characteristics based on 3 legal criteria of consent; has 2 categories, A and B; consent determined by communication and behavior. Category B determines client consent ability based on their behavior showing the following: • Voluntariness. • Safety and avoidance of harm. • No exploitation. • No abuse. • Ability to say no. • Socially appropriate time and place.	[5,6,28,94]
Mental Capacity Act 2005	Legal characteristics; based in England and Wales; section 1 of the Act assumes the people have capacity to consent unless proven otherwise; knowledge and resources to aid the person's decisions are encouraged; Includes rules for SDMs[b]. Consent requires the person to understand the following: • Is there understanding of the decision that needs to be made and why? • Does the individual understand the probable consequences when making the decision? • Is the individual capable of understanding, remembering, deliberating, and using information that pertains to the decision? • Is the individual able to communicate his or her decision in any way?	[66,95-97]
Lyden approach	Legal and clinical characteristics; endorses the 3 legal criteria of consent; encourages person-centered and integrated approaches; has important points for individualizing the assessment process, especially for communication. Has three general methods for determining the consent ability of a person with a cognitive disability, including the following: • Review the relevant records (including info on reproductive ability and other disabilities). • Create discussions, including those who know or work with the person being assessed. • Conduct a personal interview to determine knowledge and voluntariness, supplemented with a mental status evaluation.	[5,26,89,92]
ABA/APA[c] model	Legal and clinical characteristics; based on 3 legal criteria of consent, Lyden approach and Lichtenberg and Strzepek Instrument; expands on above models to include steps on how to enhance consent capacity and form comprehensive neuropsychological testing components; recommends LTC[d] facilities to develop policies and procedures for sexual relations that are consistent with state statutes.	[10,22,26,88]
Vancouver Coastal Health Authority 2009	Clinical characteristics; downloadable manual. Provides recommendations for homecare staff and nurses such as the following: • Respect the rights of persons with the capacity to consent to sexual activity. • Do not reveal confidential specifics about the person's sexual activity to those not directly involved in their care (including family members), without the person's expressed consent, if the person has capacity. • Remember that people who do not have capacity to consent to sex are still sexual beings with intimacy needs. • Remember that not every person is heterosexual. • Address one's own attitudes and behavior toward older adults and general sexuality.	[16,98]

[a]MMSE: Mini-Mental State Exam.

[b]SDM: substitute decision maker.

[c]ABA/APA: American Bar Association and American Psychological Association.

[d]LTC: long-term care.

Table 2. The approaches used to determine and manage sexual consent abilities for people with cognitive disabilities.

Approach, type, and details	References
Advance directive	
Managing consent	
Older adults with cognitive disabilities	
Living wills for the continuation or startup of relationships in advance.	[99,100]
Integrated approach	
Determining consent	
Cognitive disabilities	
I-team[a] discussion, client assessments, enforcing client rights and education.	[88,92,97]
Reduce *all-or-none* SDM[b] decision outcomes on client rights.	[83]
Older adults with cognitive disabilities	
I-team, person-centered, interval checkups, and review policy with SDM.	[89,101]
I-team, person-centered, emphasis on client limits and their context.	[10,22]
Inappropriate behavioral disabilities	
Client screening process, semistructured interview, and I-Team discussion.	[93]
Person-centered approach	
Determining consent	
Cognitive disabilities	
Holistic case-by-case, based on needs and policy, and client and staff education.	[4,23,29]
Older adults with cognitive disabilities	
The 4 Ps: prioritize people, practice effectively, preserve safety, and promote trust.	[16]
Committee approach—staff, family, friends, residents, and client discussion.	[26,28]
Education	
Managing consent	
Cognitive disabilities	
Teach awareness of normal sex behavior to both clients and staff.	[27,34,95]
Client education checked by SCEA[c], VABS[d], or IQ tests.	[40,86]
Developmental disabilities	
Consult certified sexuality educators or experts such as AASECT[e] or OWL[f].	[11]
Increase client sex-related knowledge, based on 3 legal criteria of consent.	[3]
Older adults with cognitive disabilities	
Training for professionals and LGBTQ[g] toolkits (info packages) for them.	[94]
Attitude	
Managing consent	
Cognitive disabilities	
Policy feminist disability theory, consent culture, and rely less on assessment.	[17,18,66,102]
Positive liberty, client proactive education, and attention to LGBTQ issues.	[48]
Social reframing. Recognize ability without facilitating pity.	[91]
Older adults with cognitive disabilities	
Request and consult national resources to train teams for clientele.	[85]
Inappropriate behavioral disabilities	
Psychological, social, and facility improvements over drugs. Staff education.	[41]

Approach, type, and details	References
Functional capacity	
Determining consent	
Medical condition (stroke or comatose)	
Consent-Plus with committee input, MMSE[h] (or similar), and interviews.	[19]
Cognitive disabilities	
SSAS[i] assessment, based on the 3 legal criteria of consent.	[90]
Focus on client act-specific action (not partner choice) based on MCA 2005[j].	[96]
Adaptive capacity—correlate client's other abilities to sexual consent.	[1]
Sex consent requires basic, consequential knowledge.	[9]
Older adults with cognitive disabilities	
Assessments (MMSE and IQ), coupled with witness statements and context.	[103]
People with psychiatric conditions (schizophrenia, personality)	
Communicate situational and internal understanding.	[84]
Support network	
Both	
Older adults with cognitive disabilities	
Cognition-plus. Determines consent, managed with family, staff, and SDM.	[81]
Contextual	
Determining consent	
Mild cognitive disabilities	
Consent assessment is kept the same among people and based on context.	[12]

[a]I-Team: interdisciplinary team.
[b]SDM: substitute decision maker.
[c]SCEA: Sexual Consent and Education Assessment.
[d]VABS: Vineland Adaptive Behavior Scale (Interview Edition).
[e]AASECT: American Association of Sexuality Educators, Counselors, and Therapists.
[f]OWL: Our Whole Lives.
[g]LGBTQ: lesbian, gay, bisexual, transgender, transsexual, and queer.
[h]MMSE: Mini-Mental State Exam.
[i]SSAS: Social Sexual Awareness Scale.
[j]MCA 2005: Mental Capacity Act, 2005.

Table 3. Studies on sexual consent and education for people with cognitive disabilities.

Study type	Approach	Aim	Key findings	References
Qualitative	Integrated	Interview facility staff and residents to determine factors that increase risk of SDMs[a] deciding *all-or-none* resolutions of resident consent capacity to sexual relationships instead of allowing partial expression.	Wording of legislation, lack of resources for SDMs and relational dynamics between them and staff increase risks of *all-or-none* decision outcomes. Recommends addressing these factors in integrated approach to reduce risk.	[83]
Qualitative	Attitude and education	Semistructured interview needs assessment of directors of nursing to identify challenges to sexual expression management in LTC[b] setting.	Directors of nursing requested sexual expression to be addressed in a top-down manner, with national organizations' support in resources and training.	[85]
Qualitative	Functional capacity	Interview facility staff and residents to determine key components of sexual consent.	Three key themes participants defined for consent: communication—includes all involved in sexual relationship either verbal or nonverbal, situational understanding—includes ability for all involved to interpret assent of partners, and internal understanding—includes personal understanding of desire for sexual relationship.	[84]
Qualitative	Education and attitude	Survey with vignettes to check residential staff ability to properly identify safe or unsafe sexual behaviors and respond accordingly.	Staff could generally identify the difference between abusive and safe sexual behavior. Increased age of staff correlated with less accuracy in identifying safe or unsafe behavior.	[34]
Qualitative	Functional capacity	Survey of APA[c] to determine important criteria to determine key components of sexual consent.	Key themes defined for consent: basic sexual knowledge, knowledge of the consequences of sexual behavior, and aptitudes related to self-protection.	[9]
Quantitative	Education	Education intervention—*Living Your Life*, twice weekly, 45 minutes per session, 10-week total, to improve sexuality-related decision ability.	SCEA[d] scale showed improved scores after education. Retention showed only slight decay after 6-month follow-up.	[86]
Quantitative	Education and functional capacity	Functional approach cohort study compared sexual consent ability of people living with cognitive disabilities to presumed normal people.	Some people with cognitive disabilities scored high on all measures, including the Sex-Ken-ID[e]. Recommended ongoing education instead of single inoculation model.	[40]
Quantitative	Functional capacity	Cross-sectional validity measure used SCEA to compare neuropsychological tests with IQ, adaptive behavioral age, and sex education on consent ability.	Neuropsychological test battery, especially those measuring executive measures, were found to be more accurate in predicting competency than IQ, adaptive behavior age, and sex education.	[45]

[a]SDM: substitute decision maker.

[b]LTC: long-term care.

[c]APA: American Psychological Association.

[d]SCEA: Sexual Consent and Education Assessment.

[e]Sex-Ken-ID: Sex Knowledge, Experience, and Needs Scale for People with Intellectual Disabilities.

Themes Affecting Approaches for Determining Sexual Consent Capacity

Most reports (n=14) placed the 3 legal criteria of consent themes at the forefront of their approach to determine the sexual consent capacity of people living with cognitive disabilities. Other reports described adapted instruments, such as the Lichtenberg and Strzepek Instrument (n=4) or the Ames and Samowitz Instrument (n=4), which are approaches based on the 3 legal criteria of consent, however, with clinical considerations. The Lichtenberg and Strzepek Instrument mentions the use of an interdisciplinary team for the second part of its assessment process, which incorporated a team of professionals (eg, psychologist, psychiatrist, nurses, recreational therapists, dietitians, and aide staff) to analyze the information obtained by a psychologist or psychiatrist's interview with the *patient* in the first part [87]. The Lyden approach, with person-centered and integrated considerations, has important points for individualizing the assessment process, especially for communication during the interview process. The American Bar Association and American Psychological Association (ABA/APA) model has a handbook, which is based on the 3 legal criteria of consent, Lyden approach, and Lichtenberg and

Strzepek Instrument. It includes comprehensive neuropsychological testing components for determining consent capacity; however, it only summarizes the team-based aspects of determination and care plans [88].

With the 3 legal criteria of consent being based in the United States, some international reports described the Mental Capacity Act of 2005 as their main approach (n=4). The Mental Capacity Act of 2005 uses rules reminiscent of the 3 legal criteria of consent and contains the prefix assumption that a person has the default capacity to consent unless proven otherwise. The Mental Capacity Act of 2005 also contains prefix rules to ensure that knowledge and resources are available for assisting a person to make a consensual decision. Some reports introduced white papers and guides for assisting adult sexual health in LTC facilities (n=2), the most recommended guide being from the Vancouver Coastal Health Authority 2009. This guide provides important reminders to nurses and homecare staff regarding the rights of people under their care, while also recommending support to healthy sexual behavior by providing the means to do so (eg, provision of private spaces to reduce public sexual activity) [98].

Approaches for Determining Sexual Consent Ability

Functional Capacity

Approaches endorsing the use of functional capacity have shifted away from diagnostic-based assessments (eg, IQ and mental age scores) of decision-making ability to alternative identifiers. There was an overall emphasis in the literature to rely less on mental assessment outcomes when determining the sexual decision-making capacity of people with cognitive disabilities [17,28,85,91,102]. The problem with mental assessments is the risk of social or cultural factors, concluding that people with cognitive disabilities are incompetent in society, which results in such people being placed in the *cloak of competence* [104]. Repeated assessment measures show mixed results for the same person, varying between having and not having a sexual consent capacity [3]. Thus, even if people with cognitive disabilities are able to demonstrate sexual consent capacity, social judgment may still suspect such people as being less than capable when compared with normative scores of intelligence and adaptive behavioral ability [91].

Despite the shortcomings of mental assessments, reports vouching for functional capacity approaches recommended to either expand the assessment's ability to check for adaptive behavioral domains [1,19], featuring the use of neuropsychological assessments [88] or relegate such mental assessments to a supporting role [103]. Although a mental state assessment such as the MMSE may provide a basic idea on how well a person with cognitive disabilities makes rational decisions, it was recommended to consider other domains of functional capacity such as literacy skills, self-care skills, independent care ability, and physical upkeep [1]. A secondary diagnosis, including checks for both cognitive and functional ability, should be used when making a capacity judgment [103]. A report by Bogacki et al [90] introduced the Social Sexual Awareness Scale, a scale based on the 3 legal criteria of consent that checks a person's knowledge for contextual and safety factors such as risk of disease, contraceptive use, age of partners, and the handling of unwanted advances [90]. Another recommendation was to assess decision-making capacity on act-specific decisions (eg, being able to make a decision on a sexual relationship, retain that decision, and have a rudimentary understanding of the sexual act) instead of theoretical ability [96]. One report suggested a holistic approach, complete with a progressive committee [19], supplemented with theories of Consensual-Minimalism and Consent-Plus by Wertheimer [105]. Consensual-Minimalism checks for the most straightforward sense of consent among those in a relationship, whereas Consent-Plus pertains to situations where consent is necessary in addition to other factors that are additionally required to make the sexual relationship permissible (eg, social, religious, and cultural factors) [105].

Person-Centered Approach

The person-centered approach is philosophically driven to promote ethical integrity when working with people to determine their consent capacity [4]. Discussions with the people themselves will enable a better understanding of both their sexuality and cognitive disabilities, which are essential for determining their preferences [15,99]. The person-centered approach needs to be flexible to accommodate the specific needs of each individual, promote their dignity and autonomy, and uncover potential contexts that could identify risks associated with sexual expression [16,18]. In terms of philosophical components throughout the literature, the person-centered approach was recommended to include the following (people living with cognitive disabilities, who are to be provided such services, will be referred to as *clients* in this list):

1. Open communication—this factor begins with individualizing the communication process with clients, following key components of the Lichtenberg and Strzepek Instrument, Ames and Samowitz Instrument, Lyden approach, ABA/APA model, and Vancouver Coastal Health Authority themes [5,6,87,98]. At the same time, given the personal biases and stigmas associated with sexuality, everyone including caregivers and staff should discuss such potential issues while planning to address them [23].
2. Committee approach—following open communication, this factor can *diffuse liability exposure and provide enhanced objectivity* [5]. If persistent evidence shows that a client's consent ability has become compromised, any decisions involving their rights to sexual intimacy must involve discussions with family, friends, caregivers, and staff in a *multidisciplinary team* setting [16]. Such a committee can incorporate various perspectives to enhance a client's autonomy, dignity, and rights to sexual expression in addition to determining areas and means for improving their potential shortcomings in the 3 legal criteria of consent themes [5,26]. Lay-witnesses are useful for determining other contextual information, such as the client's adaptive capacity, in addition to determining potential SDMs if allowed [1,89].
3. Capacity assumed—a client's sexual decision-making capacity needs to be assumed intact unless proven otherwise [4,83]. It is unethical and discriminatory to use assessments to prove that a client is incapable of demonstrating consent capacity [17]. A common-sense approach, with staff and

peers observing the client's interactions, nonverbal language, and social cues should instead be considered when determining their sexual decision-making capacity [4].

4. Withhold Bias—people's attitudes, including those of staff members and caregivers, may inadvertently be against a client's sexual preferences and deny them their rights to sexual expression [91]. It was strongly recommended for both clients and staff to receive education programs to discern differences between normative and unhealthy sexual behaviors in addition to reframing their perceptions regarding client sexual preferences and rights [3,23,29,34,88,92].
5. Tracking—a client's sexual preference and decision-making ability is expected to change over time; thus, person-centered approaches should track a client's progress, reconfirming that they retain the capacity to both understand and refuse a sexual interaction when necessary [23]. The impact of educational programs for enhancing a client's sexual consent capacity should also be used and tracked over time [40]. Improving the client's access to materials for practicing safe sex, such as condoms and contraceptives, should be continuously implemented [23]. Safeguards that limit forms of client sexual expression, such as restricting a relationship to kissing and touching only, requires continual monitoring by staff to ensure that such safeguards are not exceeded [5].

One report presented a system for nurses to use, which adheres to some of the previously mentioned philosophical components. The system comprises the 4 themes of the code, Professional Standards of Practice and Behavior for Nurses and Midwives [106]. The 4 themes are prioritize people, practice effectively, preserve safety, and promote professionalism and trust [16]. It includes considerations such as withholding bias, open communication, and the committee approach. A report by Wilkins [107] suggested an emphasis among a substituted judgment, best interest standard, or a mix of the two. The substituted judgment standard emphasizes the use of advanced directives, whereas the best interest standard emphasizes the balance between risks and benefits, allowing some restricted forms of client sexual expression if the potential benefits are worth the risks.

Integrated Approach

This is perhaps the most comprehensive approach in terms of providing a detailed care plan for determining sexual consent capacity in people with cognitive disabilities, while also discussing plans for enhancing consent capacity if necessary. The key features of the integrated approach include an interdisciplinary team discussion process, using aspects of a person-centered approach, complete with other holistic considerations. With the service user's permission, the interdisciplinary team can be comprised an array of practitioners including physicians, occupational therapists, psychologists, social workers, nurses, and legal guardians of the person involved in the discussion [22]. The integrated approach is likely to encompass the following themes:

1. The 3 legal criteria of consent—for legally defining the terms of consent [7,10,22].
2. ABA/APA model—for blueprinting the overall assessment and care plan design process, endorsing person-centered considerations of the evaluated person's sexual values, which endorses the Lichtenberg and Strzepek instruments for both functional capacity and ethical considerations [87,88].
3. Lyden approach—for individualizing the communication and assessment process, encouraging person-centered aspects to the approach [5,22,88].
4. Lichtenberg and Strzepek Instrument—assuming that the assessment aspect of the process is performed with an MMSE [87,88]; however, such assessments were designated as a supplementary role to determine where lacking areas of knowledge could be improved [48].

Friends and family of the evaluated person are encouraged to play a role in the discussion process [22]. Team input determines the restrictions of sexual expression, if any and medications to be prescribed, if any, while also noting contextual factors of the relationship, such as potential risks of coercion or abuse [22,92]. Overall, the integrated approach focuses on holistic contextual factors throughout the assessment process, including factors such as the person's communication ability, access to privacy, informed consent ability, family involvement, religious beliefs, and social history [22,88].

Contextual Approach

The contextual approach was aimed at individuals with mild cognitive disabilities and has 2 components [12]. First, whenever a judicial system assesses the consent ability of an individual living with a cognitive disability, the ruling must meet the same standard as anyone else. Second, it is recommended that consent ability be focused on the situational context rather than on intellectual attributes. For example, an individual with a cognitive disability may show consent capability in a healthy relationship but not when their partner uses coercion or threats.

The first component becomes especially important when consent definitions require an understanding of tests involving the nature, consequences, and moral dimensions of sexual acts [12]. It is important to keep assessments among individuals the same, whether they have intellectual disabilities, because this focuses on social innerworkings within an intimate relationship [12]. This component also respects the capabilities of individuals with cognitive disabilities. The second component realizes that consent ability is affected by social constructs such as communication, social skills, and community support. A contextual approach reassures the balance of protection toward vulnerable persons while respecting their consensual rights to sexual relationships.

Approaches for Managing Sexual Consent Ability

Education

There was a pattern in the reports explaining how education could improve the sexual decision-making ability of people living with cognitive disabilities. The pattern starts by mentioning the 3 legal criteria of consent components (knowledge, understanding, and voluntariness), followed by a

defined set of basic skill checks to determine whether such people could address these components. Note that the 3 legal criteria of consent have a knowledge component, which defaults to being improved by sex education; however, its other components, such as understanding, may benefit from education as well [3,5,40,86]. The basic skill checks included areas that expanded upon the three legal criteria of consent [3,5,11,40]:

1. Knowledge of body parts and sexual relationships and acts.
2. Knowledge of consequences from sexual relationships.
3. Understanding of appropriate sexual behavior and context for it.
4. Understanding of the voluntary nature of a sexual relationship.
5. Ability to recognize abusive situations.
6. Ability to be assertive in such situations to reject unwanted advances.

The reports endorsing educational approaches described measures that check for these areas, such as the Sexual Consent and Education Assessment [86], Sex Knowledge, Experience and Needs Scale for People with Intellectual Disabilities [40], and Tool for the Assessment of Levels of Knowledge Sexuality and Consent [11]. For educational programs themselves, the recommendations were the *Living Your Life—The Sex Education and Personal Development Resource for Special Education Needs* [86,108] and comprehensible evidence-based programs with simple language [3]. It is important for the aforementioned assessments to be used only for the identification of gaps in a person's knowledge of safe sexual acts, followed by providing educational programs to rectify such gaps if necessary [48,92].

In addition to people with cognitive disabilities, it was encouraged for staff in LTC facilities to receive education to better identify the difference between healthy and unhealthy sexual behaviors and how to resolve such situations accordingly [27,34]. Both families and LTC staff were recommended to receive education to better understand the rights to intimacy, sexuality, and privacy for people with cognitive disabilities [27]. It was suggested that staff use *LGBTQ toolkits*, which are manuals that describe ethical approaches when working with older female adults who have these sexual identities [94].

Criminal justice systems were encouraged to use education and training programs to increase the awareness of communication disorders, while also considering alternative communication platforms and multidisciplinary collaborations with relevant disciplines [109]. Children and people with communication disabilities are at a disadvantage when disclosing their experiences of sexual abuse to a criminal justice system, often because the system's procedures may not be adapted to meet the needs of such people [109-111]. Sexual abuse cases showed improved outcomes when collaborative support was combined with communication awareness, such as for law enforcement and child protection services [111]. Improved access to sexual and gender-based violence education for vulnerable populations, such as refugees, was also recommended in addition to encouraging inclusion in community support programs [109].

Attitude

There were 3 articles that argued for both disability and feminist rights movements to overcome negative attitudes within communities [17,18,102]. Doyle [18] described how social construction theory for feminism relates to people with cognitive disabilities, defining it as culturally set norms, rights, and commitments that detail expectations on how people of differing statuses relate to one another [112]. Doyle [18] also explained that script theory is a form of construction theory that influences a person's sexual behaviors by external and internal factors, defined as society's *mutually shared conventions and norms* and personal motivations, respectively [18,113]. Script theory explains how the sexuality of a person with cognitive disabilities is potentially influenced by critical factors beyond just sex drive and instinct: it has learned behavior characteristics, influenced by social contexts, affecting how people express themselves [18].

Negative cultural attitudes, such as rape culture, should be countered by plans using lifelong sexuality education and policy change with *intermovement collaboration*, addressing the aforementioned internal and external factors [17]. A person with a cognitive disability could be incapable of demonstrating or understanding consent capacity, due to a lack of knowledge or having misaligned sexual scripts [18]; however, educational interventions may allow that person to reach capacity [5]. Note that a strict education approach emphasizes sexuality instruction to fill missing gaps in knowledge and understanding of consent criteria, although the attitude approach is often based on social construction theory, suggesting the use of education to reframe a person's sexual scripts.

The attitude approach also discussed how a person with cognitive disabilities and their external factors such as government, legal systems, administration, practitioners, staff of LTC facilities, and family could be influenced. A study by Syme et al [85] determined that a proactive approach to policy development in LTCs was recommended, in addition to addressing negative staff and family attitudes. Addressing the lack of awareness of sexual expression in people with cognitive disabilities, making necessary environmental changes to ensure privacy, identifying staff or family attitudes, and tracking the person's sexuality with recurrent assessments were the top areas to address in LTC facilities [85]. The study by Syme et al [85] found that all nursing directors (n=20) endorsed the use of sexuality education, with *many* endorsing attitude discussions with staff and family about aging sexuality, changing negative attitudes, increasing people's awareness of their own attitudes, and emphasizing sexual health. A report by Victor [23] mentioned the use of staff attitudes about intimacy and dementia survey proposed by Kuhn [114] to measure staff or caregiver attitudes on this topic. A collaborative reform process was recommended to change legal terms (eg, mentally impaired to vulnerable or protected persons) when reframing government views that hold people to a disproportionately higher standard [92]. For committee approaches, members were tasked with reducing the risk of personal bias within the leading assessor, by providing a *balanced exchange of ideas in a competent and thorough manner*, throughout the consent capacity determination process [5].

Advanced Directives

Advanced directives are contingency plans that allow people to consent to specific sexual acts ahead of time or grant decision-making power to an SDM for an applicable future context [115]. The plan involves the person, who is legally capable of providing sexual consent at the time, the past self, to set up a contingency for an impending period when their sexual decision-making ability may become compromised, the future self. This compromised ability to provide sexual consent may occur owing to impending conditions from events such as dementia, stroke, or brain surgery. There are two types of advanced directives [115]: instructional directives and proxy directives. The instructional directives can be either permissive, allowing permissions to take place when they legally could not or restrictive, halting actions from happening when situations would normally favor them [115]. The proxy directive features the use of SDMs, either as a surrogate decision maker or the power of attorney [115].

A person's ability to consent to sex can change over time, varying across situations in terms of capacity and sexual preference [3,5,11]. Unless legal exceptions have been established, it is illegal for consent to be given by someone else [7]. Advanced directives are noted for upholding a person's core values and religious beliefs when resolving decisions involving sexual relationships, perhaps preserving the sexual preferences of a person living with a cognitive disability [77]; however, the drawback is that the person is *locked into certain conditions that may not coincide with the desires of the future self* [28]. The use of advanced directives in this manner is referred to as the substituted judgment standard [28]. One report presented the *Prior Consent Thesis*, a philosophy-driven argument that states that a competent person can give valid prior consent to a competent partner, in which consent could remain even after mental capacity in one person becomes compromised [100]. Sexual advance directives are not promises that could lead to the promiser being locked in servitude. The advance directive is not about owing a commitment; a person's advance directive merely states that such an encounter be allowed if they token consent [100].

Support Network

Using the cognition-plus test, this approach contains three steps [81]: (1) check if the individual is capable of communicating a desire for an intimate relationship, (2) check if the individual is aware of the nature and consequences of sexual decisions, and (3) determine the adequacy of an individual's decision-making support network.

If step 1 is unfulfilled, the test ends. Individuals who cannot determine an intimate relationship cannot qualify as sexual agents. If steps 1-2 are fulfilled, the individual is deemed to have consent capacity without the need for assistance. Step 3 becomes active only if step 2 is unfulfilled. The determination of an individual's support network is contextual-based, guided by the fiduciary law [81]. The legal system would need to check if the support network is free from conflict of interest, while showing an understanding of the individual's sexual preferences, with contingencies to protect the individual against consequences of sexual encounters, such as pregnancy or sexually transmitted disease [81]. The support network is expected to be different among individuals but may comprise the individual's friends, family, institutional staff, and SDMs. SDMs are not recommended to act alone in a support network [81]. It is possible that people in the support network may have disagreements on the individual's preferences, in which case these disagreements will need to be resolved in a civil manner before this approach is implemented [81]. The legal system should not intervene whenever a civil dispute occurs: its focus should only be on checking the support network's adequacy [81]. It is important to note that support networks should not exist to make decisions for people with cognitive disabilities but only to assist them in achieving their decisions [81].

Discussion

Principal Findings

This review focused on the approaches used to determine and manage sexual consent abilities in people with cognitive disabilities, noting the recurring themes influencing how these approaches were implemented. The literature assumes that such people are capable of having the capacity to desire and consent to healthy intimate relationships; however, some situations may ignore or suppress these capabilities [5]. The movement of medicalized models to be enhanced by social support was also emphasized.

Key Points When Determining Sexual Consent Ability

Review of the literature has established that determining consensual abilities requires a holistic approach, with individuals being considered in terms of their adaptive abilities, capacities, and human rights. An abridged description of such a holistic approach includes identification of the person's sexual identity, beliefs and values, opinions from friends and family, medical records and clinical interviews (person-centered), neuropsychological testing, and functional capacity measures involving adaptive capability skill checks, followed preferably by an interdisciplinary discussion and action plan [88]. Although the 3 legal criteria of consent may appear as the starting point for defining a person's consent capacity, sexual identity, beliefs, values, culture, and life history should be examined first to guide the consent-determining process [10]. This promotes the person-centered approach, especially if the assessed individual is of the LGBTQ community. The use of a committee during person-centered and integrated approaches must maintain a morality balance within the main reviewer; otherwise, the consent determination process can be skewed either in favor of or against the assessed individual [91]. Continual monitoring of an individual's sexual preferences allows caregivers or service providers to offer empathetic maintenance over time, which becomes especially useful during cases of *fading identity*, owing to conditions such as dementia [23]. Educational programs focused specifically on these identified preferences can be provided to improve empathetic maintenance for such service providers.

The functional capacity report by Harris [1] provided examples of evidence for sexual consent ability in people with cognitive disabilities, including those demonstrating skills such as exercising *good judgment* and being able to identity their name

and address correctly. However, the report by Harris [1] also admits that appraising these examples can be an *abstract exercise*, with some forms of evidence carrying a higher amount of probative weight than others, such as previous experience with sex education over IQ scores. A report by Thomas [84] provided elements of consent capacity definitions (communication, situational, and internal understanding), which were recommended to be presented to other residential mental health experts to create and improve the acceptable definition of consent ability. The most important consideration when using the functional capacity approach is to enforce that people with cognitive disabilities are not regulated differently than people without cognitive disabilities. If they should be denied legal right to sexual consent, the conclusion should only be drawn if the functional capacity appraisal process is performed on a truly equal basis for all [66].

Critical requirements of consent culture, which states that "people can have sex only when everyone agrees it is OK" [17], is met with scenarios where a person in a married relationship wants to have sex with their partner who has severe dementia, incapable of showing either signs of assent or refusal. Although it is important for all parties in the relationship to show the ability to refuse a sexual encounter at any point, which many people with cognitive disabilities will be incapable of doing [23], one proposed idea is to affiliate consensual sex with a continuation aspect, assuming consent from a previous loving relationship will remain even after someone loses the ability to consent afterward [100]. This continuation aspect of a sexual relationship with someone living with a cognitive disability is the *Prior Consent Thesis*, which deems current sexual relationships permissible, providing both people in the relationship gave consent to sex with each other before.

Key Points When Managing and Enhancing Sexual Consent Ability

The key aspects to consider when managing and enhancing consensual ability in people with cognitive disabilities starts with attitude change. Some of the recurring attitudinal barriers identified in the literature include internal factors, such as those inflicting the individuals themselves as explained by script theory and external factors, peripheral to the individual affecting their consent ability and rights, examples being care providers, legal systems, family, friends, and supportive decision makers [18]. The attitude approach gives the impression of being the most important approach because of its ability to reframe either the internal or external factors' view of sexual expression for people with cognitive disabilities.

For internal factors, script theory explained that people with cognitive disabilities may show unhealthy sexual behaviors because of unlearned scripts [18], which can be supplemented by education to fortify the knowledge and understanding prongs in the 3 legal criteria of consent theme. Script theory also explains that people may have intact sexual knowledge; however, *vital elements* pertaining to healthy attitudes can be misaligned, which may result in unhealthy behaviors directed at oneself or other people [18]. It has been reported that sexual knowledge alone does not always transfer to safer sexual behaviors [116]; thus, the importance of lived experiences on consensual ability emphasizes the importance of understanding the contextual reason why people with cognitive disabilities may consent to sex [18]. It is important to note that script theory does not explain inherent sexuality within individuals. Sometimes, unhealthy sexual behaviors stem from physical conditions, such as those from affected neurobiological areas in the brain within a person living with dementia [93].

External factors influencing sexual expression in people with cognitive disabilities include both formal and local situations. A report by Arstein-Kerslake and Flynn [102] provided details of a grassroots voice movement, using feminist disability theory to encourage a formal legal system to adopt vital changes to its interpretation of sexual rights for people with cognitive disabilities, further describing the drafting of its convention by a disabled people's organization to reform the rights of such people. Perhaps these details may be enough to convince a similar rights movement to attempt their own reform process; however, the unique details pertaining to each convention's attitudinal barriers will need to be reported in future literature for conventions to strategize and gain a greater level of confidence in attitude change. The more details a system can obtain for establishing attitude change, the louder a new convention can have a voice.

To check and address internal and external factors in the local situation, attitude checks using tools such as the staff attitudes about intimacy and dementia survey were recommended to give a general idea about potential staff, family, and caregiver biases toward sexual preferences in people with cognitive disabilities [23,93,114]. Publicly available guides, such as the Vancouver Coastal Health Authority Guide (2009), provide recommendations about components of staff education, SDMs, and family decision-making, including what to do in case decisions fall upon family members who do not support the client's sexual activity. For situations with decision makers who do not agree with their clients' sexual preferences, example guidelines include reiterating the legal obligations of such decision makers while reassuring them that they do not have to change their personal values—they only need to respect the legal rights of the client, especially after they have been given sex education [98].

The attitude approach is arguably the most important approach to consider, because all efforts to realize a person's sexual consent ability can be lost, should the final decision fall onto someone who does not agree with the sexual preference in the first place. The comprehensive aspects of the integrated approach are not immune to this. The Hillman report [22] described an in-depth integrated approach that was used to assist with the sexual consent ability of a resident in a nursing home; however, the final outcome resulted in the family relocating the resident from there because of a difference in perceptions regarding sexual relationships. Using education to reframe staff and family attitudes is key because to quote one director of nursing from the study by Syme et al [85], "If the families don't buy it, it's gonna fail miserably." The Boni-Saenz [81] report mentioned that support networks may fail because cognition-plus does not force its members to agree with their sexual preferences [81]. This suggests that both education and attitude approaches may ensure that consent plans are fulfilled properly. It was

recommended for newer doctrines to reframe attitudes and counter stereotyping beliefs within those who serve people with cognitive disabilities, using laws to *exert positive expressive pressure on social norms* [81].

Advanced directives have evolved to a point where they may play a role in sexual decision-making. They are best established within third-party systems, especially those equipped to monitor their use, such as those in LTC settings [115]. Advanced directives may improve the decisional accuracy of institutional staff and loved ones who may act on behalf of the person living with a cognitive disability [117] or act as an *essential element* in a legal case against sexual assault [115]. The literature has revealed that the use of SDMs has been controversial. Although an SDM may appear qualified to make a decision on behalf of a person for a future sexual situation, it is possible the SDM may develop a conflict of interest, failing to act with loyalty in their proxy decision-making, resulting in potential *congruence problems* [118]. Having an advanced directive coupled with LTC facility monitoring or other such third-party settings reduces the risk of objective harm [115]. Further controversy stems from the disagreements among systems adopting the use of SDMs and the UNCRPD. The UNCRPD is against the use of SDMs, owing to the motion by Article 12 to establish equal rights for those living with cognitive disabilities [31,32]. The UNCRPD states that a lack of consensual capacity should not restrict a person's ability to make a decision, insisting that said person's *will and preferences* must be acknowledged [31,32]. Several experts have accused this interpretation as unrealistic [32,119-121]. Forgoing the use of SDMs would reduce civil commitment and potentially increase the risk of harm for individuals living with severe cognitive or psychotic disorders [32].

There are philosophical arguments that may prefer either the individual's past or future self to take precedence over the final decision of the advanced directive. It is important for evidence of both the past and future self to show some form of communication to consent, be it verbal or nonverbal. If the past self had a contingency to consent to a sexual relationship and the future self showed a token of interest, such as overtly wanting to hold hands with someone they like, this may show an overlap in interests between the past and future selves. The noticeable overlap in consensual interests between past and future selves is known as the Consensus of Consents [115], which is the key to deciding whether such an advance directive is legally enforceable. Unfortunately, a person with a cognitive disability may show behavior that is less obvious in showing contemporaneous consent. It could be that some people living with dementia will have conditions so dire that they cannot communicate any form of consent at all. In these cases, the ability to express some form of definitive volition is required, either in the form of a verbal *yes* or some behavioral initiation of sexual expression [122]. With this volitional requirement, the voluntariness prong of the 3 legal criteria of consent demonstrates that sexual advance directives have both legal and medical domains, thus requiring dedicated supporters to have knowledge about consent capacity, mental conditions, and the very people living with cognitive disabilities themselves.

Limitations

This review used a systematic method to identify approaches for determining and managing sexual consent capacity in people with cognitive disabilities. There was an emphasis on recognizing the patterns of themes, each influencing how consent-determining and enhancement programs were implemented. Although the literature on the subject may have a diverse array of ideas, acknowledging the views and rights of those who desire intimate relationships, this review emphasizes a convergent style to bring these ideas together. With this review's emphasis on pattern recognition for noting recurring themes, there is a strong possibility that emergent ideas may have been downplayed or missed entirely. This review did not include ideas from conference papers, public opinions, or non–peer-reviewed articles. This may have shifted this review's focus to a stronger understanding of already-established sexual consent themes; however, it could be that newer ideas may change these already-existing themes. Future research may provide emergent ideas with a greater consideration of the subject. In addition, although this review featured reports from both clinical and legal sources, this review predominantly used a clinical search protocol to locate the literature. The search process was not dedicated to the legal databases. Future research on this topic should be performed with a legal background, incorporating the necessary legal databases and journals.

Conclusions

The desire to have an intimate relationship is one of the core elements of sexuality, which is part of what it is to be human. Healthy sexual relationships are driven by consent, which is commonly defined by people's capability to demonstrate sexual knowledge, intelligence, and voluntariness. However, if a person with a cognitive disability has a compromised consent ability, the involved legal, clinical, or ethical systems must determine the balance between permitting and restricting sexual activity to reduce the risk of unhealthy or harmful sexual behavior. It is important for the attitudes of those involved in this process to be balanced; otherwise, the sexual rights of such assessed people could be moved either in favor or against them. The means for determining the sexual consent ability of people with cognitive disabilities include functional capacity and person-centered and integrated approaches. Management of consent ability includes education, attitude, and advanced directive approaches. These approaches seek the ideal outcome where person-centered considerations of those living with cognitive disabilities are understood and they themselves are involved in the process of personalizing the approaches used to facilitate healthy intimate relationships.

Acknowledgments

The authors wish to acknowledge the contributions of the librarian in residence from the Rehabilitation Research Center, Faculty of Rehabilitation Medicine, University of Alberta, who lent their support for the search protocol used in this systematic review.

Authors' Contributions

BJC and SE coconceptualized this review. SE contributed to the design of the sexual consent research question and search parameters. BJC led the manuscript writing process. SE contributed to the writing process and the revisions. All the authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[PDF File (Adobe PDF File), 470 KB - ijmr_v11i1e28137_app1.pdf]

References

1. Harris J. Sexual consent and disability. New York Uni Law Rev 2018;93(3):480-557 [FREE Full text]
2. Schalock RL, Borthwick-Duffy SA, Bradley VJ, Buntinx WH, Coulter DL, Craig EM, et al. Intellectual Disability: Definition, Classification, and Systems of Supports. Eleventh Edition. Washington, DC: American Association on Intellectual and Developmental Disabilities; 2010.
3. Stein S, Tepper M. Sexuality. In: Health Care for People with Intellectual and Developmental Disabilities across the Lifespan. Cham: Springer; 2016:1365-1381.
4. Tarzia L, Fetherstonhaugh D, Bauer M. Dementia, sexuality and consent in residential aged care facilities. J Med Ethics 2012 Oct;38(10):609-613 [FREE Full text] [doi: 10.1136/medethics-2011-100453] [Medline: 22736582]
5. Lyden M. Assessment of sexual consent capacity. Sex Disabil 2007 Feb 20;25(1):3-20. [doi: 10.1007/s11195-006-9028-2]
6. Ames TR, Samowitz P. Inclusionary standard for determining sexual consent for individuals with developmental disabilities. Mental Retard 1995;33(4):264 [FREE Full text]
7. Stavis PF. Harmonizing the right to sexual expression and the right to protection from harm for persons with mental disability. Sex Disabil 1991;9(2):131-141. [doi: 10.1007/bf01101738]
8. Kaeser F. Can people with severe mental retardation consent to mutual sex? Sex Disabil 1992;10(1):33-42 [FREE Full text] [doi: 10.1007/bf01102246]
9. Hill Kennedy C, Niederbuhl J. Establishing criteria for sexual consent capacity. Am J Mental Retard 2001;106(6):503. [doi: 10.1352/0895-8017(2001)106<0503:ecfscc>2.0.co;2]
10. Abellard J, Rodgers C, Bales A. Balancing sexual expression and risk of harm in elderly persons with dementia. J Am Acad Psychiatry Law 2017;45:485-492 [FREE Full text]
11. Stein S, Dillenburger K. Ethics in sexual behavior assessment and support for people with intellectual disability. Int J Disab Hum Dev 2016;16(1):11-17. [doi: 10.1515/ijdhd-2016-0023]
12. Denno DW. Sexuality, rape, and mental retardation. University of Illinois Law Review. 1997. URL: https://heinonline.org/HOL/P?h=hein.journals/unilllr1997&i=327 [accessed 2022-01-25]
13. Criminal code: C C-46. CanLII. 1985. URL: https://www.canlii.org/en/ca/laws/stat/rsc-1985-c-c-46/latest/ [accessed 2022-01-25]
14. Lopes-Baker A, McDonald M, Schissler J. Canada and United States: campus sexual assault law and policy comparative analysis. Can.-U.S. L.J. 2017. URL: https://scholarlycommons.law.case.edu/cuslj/vol41/iss1/8 [accessed 2022-01-25]
15. Curtice M, Mayo J, Crocombe J. Consent and sex in vulnerable adults: a review of case law. Br J Learn Disabil 2012 Jun 28;41(4):280-287. [doi: 10.1111/j.1468-3156.2012.00753.x]
16. Syme M. Supporting safe sexual and intimate expression among older people in care homes. Nurs Stand 2017 Aug 23;31(52):52-63. [doi: 10.7748/ns.2017.e10906] [Medline: 28831892]
17. Moras R. Feminism, rape culture, and intellectual disability: incorporating sexual self-advocacy and sexual consent capacity. In: Emerging Perspectives on Disability Studies. New York: Palgrave Macmillan; 2013:189-207.
18. Doyle S. The notion of consent to sexual activity for persons with mental disabilities. Liverpool Law Rev 2010 Oct 26;31(2):111-135. [doi: 10.1007/s10991-010-9076-7]
19. Lindsay JR. The need for more specific legislation in sexual consent capacity assessments for nursing home residents. J Leg Med 2010 Jul 31;31(3):303-323. [doi: 10.1080/01947648.2010.505835] [Medline: 20830640]
20. Dimopoulos A. Issues in Human Rights Protection of Intellectually Disabled Persons. London, UK: Routledge; 2010.

21. Rowlands S, Amy JJ. Sterilization of those with intellectual disability: evolution from non-consensual interventions to strict safeguards. J Intellect Disabil 2019 Jun;23(2):233-249 [FREE Full text] [doi: 10.1177/1744629517747162] [Medline: 29228865]
22. Hillman J. Sexual consent capacity: ethical issues and challenges in long-term care. Clin Gerontol 2017 May 10;40(1):43-50. [doi: 10.1080/07317115.2016.1185488] [Medline: 28452627]
23. Victor E, Guidry-Grimes L. Relational autonomy in action: rethinking dementia and sexuality in care facilities. Nurs Ethics 2019 Sep 18;26(6):1654-1664. [doi: 10.1177/0969733018780527] [Medline: 30021486]
24. Richardson JP, Lazur A. Sexuality in the nursing home patient. Am Fam Physician 1995 Jan;51(1):121-124. [Medline: 7810464]
25. Greydanus D, Omar HA. Sexuality issues and gynecologic care of adolescents with developmental disabilities. Pediatr Clin North Am 2008 Dec;55(6):1315-35, viii [FREE Full text] [doi: 10.1016/j.pcl.2008.08.002] [Medline: 19041460]
26. Ambrosini D, Hategan A, Bourgeois J. The sexual consent and patients with major neurocognitive disorders. In: On-Call Geriat Psychiatry. Cham: Springer; 2016:209-215.
27. Bradley C. Supporting intimacy in the rehabilitative, assisted, and long-term care settings. Topics Geriat Rehab 2016;32(3):218-220. [doi: 10.1097/TGR.0000000000000113]
28. Wilkins JM. More than capacity: alternatives for sexual decision making for individuals with dementia. Gerontologist 2015 Oct 27;55(5):716-723. [doi: 10.1093/geront/gnv098] [Medline: 26315314]
29. Roelofs T, Luijkx KG, Embregts PJ. Intimacy and sexuality of nursing home residents with dementia: a systematic review. Int Psychogeriatr 2014 Nov 10;27(3):367-384 [FREE Full text] [doi: 10.1017/s1041610214002373]
30. Article 130 - NY Penal Law. New York State Law. 2020. URL: http://ypdcrime.com/penal.law/article130.htm#p130.05 [accessed 2022-01-25]
31. Sexual and reproductive health. World Health Organization. URL: https://www.who.int/reproductivehealth/topics/gender_rights/sexual_health/en/ [accessed 2022-01-25]
32. Dufour M, Hastings T, O'Reilly R. Canada should retain its reservation on the United Nation's convention on the rights of persons with disabilities. Can J Psychiatry 2018 Dec;63(12):809-812 [FREE Full text] [doi: 10.1177/0706743718784939] [Medline: 29925272]
33. Kanter AS. Let's try again: why the United States should ratify the United Nations convention on the rights of people with disabilities. Touro Law Review. 2019. URL: https://tinyurl.com/mrxezvem [accessed 2022-01-25]
34. Crawford D. Assessment of staff perceptions of sexual behavior among persons with intellectual and developmental disabilities. Dissertation, Massachusetts School of Professional Psychology. 2014. URL: https://www.proquest.com/docview/1558180795?pq-origsite=gscholar&fromopenview=true [accessed 2022-01-25]
35. Senn CY. Vulnerable: sexual abuse and people with an intellectual handicap. ERIC Institute of Education Sciences. 1988. URL: https://eric.ed.gov/?id=ED302975 [accessed 2022-01-25]
36. Mahoney A, Poling A. Sexual abuse prevention for people with severe developmental disabilities. J Dev Phys Disabil 2011 May 3;23(4):369-376 [FREE Full text] [doi: 10.1007/s10882-011-9244-2]
37. Hershkowitz I, Lamb ME, Horowitz D. Victimization of children with disabilities. Am J Orthopsychiatry 2007 Oct;77(4):629-635 [FREE Full text] [doi: 10.1037/0002-9432.77.4.629] [Medline: 18194043]
38. Martin S, Ray N, Sotres-Alvarez D, Kupper LL, Moracco KE, Dickens PA, et al. Physical and sexual assault of women with disabilities. Violence Against Women 2006 Sep;12(9):823-837 [FREE Full text] [doi: 10.1177/1077801206292672] [Medline: 16905675]
39. McCormack B, Kavanagh D, Caffrey S, Power A. Investigating sexual abuse: findings of a 15-year longitudinal study. J Appl Res Int Dis 2005 Sep;18(3):217-227 [FREE Full text] [doi: 10.1111/j.1468-3148.2005.00236.x]
40. Murphy G, O'Callaghan A. Capacity of adults with intellectual disabilities to consent to sexual relationships. Psychol Med 2004 Oct;34(7):1347-1357 [FREE Full text] [doi: 10.1017/s0033291704001941] [Medline: 15697061]
41. Benbow S, Jagus CE. Sexuality in older women with mental health problems. Sex Rel Ther 2002 Aug;17(3):261-270 [FREE Full text] [doi: 10.1080/14681990220149068]
42. Hornor G. Sexual behavior in children. J Pediatr Health Care 2004 Mar;18(2):57-64 [FREE Full text] [doi: 10.1016/s0891-5245(03)00154-8]
43. Patterson T, Scott C. Behavioral management of inappropriate masturbation in an 8-year-old girl. Child Fam Behav Ther 2013 Jul;35(3):256-263 [FREE Full text] [doi: 10.1080/07317107.2013.818907]
44. Casteels K, Wouters C, Van Geet C, Devlieger H. Video reveals self-stimulation in infancy. Acta Paediatrica 2004;93(6):844-846 [FREE Full text] [doi: 10.1111/j.1651-2227.2004.tb03029.x]
45. Kennedy C. Legal and psychological implications in the assessment of sexual consent in the cognitively impaired population. Assessment 2003 Dec;10(4):352-358 [FREE Full text] [doi: 10.1177/1073191103258592] [Medline: 14682481]
46. Tissot C. Establishing a sexual identity. Case studies of learners with autism and learning difficulties. Autism 2009 Nov;13(6):551-566 [FREE Full text] [doi: 10.1177/1362361309338183] [Medline: 19933763]
47. Morano JP. Sexual abuse of the mentally retarded patient: medical and legal analysis for the primary care physician. Prim Care Companion J Clin Psychiatry 2001 Jun;3(3):126-135 [FREE Full text] [doi: 10.4088/pcc.v03n0304] [Medline: 15014610]

48. Onstot A. Capacity to consent: policies and practices that limit sexual consent for people with intellectual/developmental disabilities. Sex Disabil 2019 Jul 1;37(4):633-644. [doi: 10.1007/s11195-019-09580-9]
49. Davis LJ. At the same time, out of time: Ashley X. In: The Disability Studies Reader. New York: Routledge; 2016.
50. McKay T, Lindquist CH, Misra S. Understanding (and acting on) 20 years of research on violence and LGBTQ + communities. Trauma Viol Abuse 2019 Dec;20(5):665-678 [FREE Full text] [doi: 10.1177/1524838017728708] [Medline: 29334007]
51. Toomey R, Russell S. The role of sexual orientation in school-based victimization. Youth Society 2013 Apr 08;48(2):176-201 [FREE Full text] [doi: 10.1177/0044118x13483778]
52. Katz-Wise S, Hyde JS. Victimization experiences of lesbian, gay, and bisexual individuals: a meta-analysis. J Sex Res 2012;49(2-3):142-167 [FREE Full text] [doi: 10.1080/00224499.2011.637247] [Medline: 22380586]
53. Friedman M, Marshal MP, Guadamuz TE, Wei C, Wong CF, Saewyc EM, et al. A meta-analysis of disparities in childhood sexual abuse, parental physical abuse, and peer victimization among sexual minority and sexual nonminority individuals. Am J Public Health 2011 Aug;101(8):1481-1494 [FREE Full text] [doi: 10.2105/ajph.2009.190009]
54. Archibald C. Sexuality and dementia in residential care--whose responsibility? Sex Rel Ther 2002 Aug;17(3):301-309 [FREE Full text] [doi: 10.1080/14681990220149103]
55. Doll GM. Sexuality in nursing homes: practice and policy. J Gerontol Nurs 2013 Jul 26;39(7):30-38. [doi: 10.3928/00989134-20130418-01] [Medline: 23614386]
56. Winges-Yanez N. Why all the talk about sex? An authoethnography identifying the troubling discourse of sexuality and intellectual disability. Sex Disabil 2013 Nov 17;32(1):107-116 [FREE Full text] [doi: 10.1007/s11195-013-9331-7]
57. Feely M. Sexual surveillance and control in a community-based intellectual disability service. Sexualities 2016 Jun 02;19(5-6):725-750 [FREE Full text] [doi: 10.1177/1363460715620575]
58. English B, Tickle A, dasNair R. Views and experiences of people with intellectual disabilities regarding intimate relationships: a qualitative metasynthesis. Sex Disabil 2017 Aug 17;36(2):149-173 [FREE Full text] [doi: 10.1007/s11195-017-9502-z]
59. Miles SH, Parker K. Sexuality in the nursing home: iatrogenic loneliness. Gen J Am Society Aging 1999;23(1):36-43 [FREE Full text]
60. Hamilton C. 'Now I'd like to sleep with Rachael' – researching sexuality support in a service agency group home. Disab Society 2009 Apr 14;24(3):303-315 [FREE Full text] [doi: 10.1080/09687590902789461]
61. Stone D. Policy Paradox: The Art of Political Decision Making. 3rd Ed. New York: WW Norton & Company; 2012:271-288.
62. Campbell M. Disabilities and sexual expression: a review of the literature. Sociol Compass 2017 Aug 08;11(9):e12508 [FREE Full text] [doi: 10.1111/soc4.12508]
63. Perlin ML, Lynch AJ. All his sexless patients: persons with mental disabilities and the competence to have sex. Washington Law Review. 2014. URL: https://digitalcommons.law.uw.edu/wlr/vol89/iss2/13/ [accessed 2022-01-25]
64. Braun K, Herrmann SL, Brekke OA. Sterilization policies, moral rehabilitation and the politics of amends. Crit Policy Stud 2014 Jun 19;8(2):203-226 [FREE Full text] [doi: 10.1080/19460171.2013.869179]
65. Ward L. Eugenics, euthenics, and eudemics. Am J Sociol 1913 May;18(6):737-754 [FREE Full text] [doi: 10.1086/212155]
66. Arstein-Kerslake A. Understanding sex: the right to legal capacity to consent to sex. Disab Society 2015 Dec 14;30(10):1459-1473 [FREE Full text] [doi: 10.1080/09687599.2015.1116059]
67. Trent Jr JW. Inventing the Feeble Mind: A History of Mental Retardation in the United States. Berkeley, California: University of California Press; 1994.
68. Wilson R. Eugenics undefended. Monash Bioeth Rev 2019 Oct;37(1-2):68-75 [FREE Full text] [doi: 10.1007/s40592-019-00094-w] [Medline: 31325149]
69. de Oliveira IT, Dias JG, Padmadas SS. Dominance of sterilization and alternative choices of contraception in India: an appraisal of the socioeconomic impact. PLoS One 2014;9(1):e86654 [FREE Full text] [doi: 10.1371/journal.pone.0086654] [Medline: 24489759]
70. Tilley E, Walmsley J, Earle S, Atkinson D. 'The silence is roaring': sterilization, reproductive rights and women with intellectual disabilities. Disab Society 2012 May;27(3):413-426 [FREE Full text] [doi: 10.1080/09687599.2012.654991]
71. Goldman EB, Quint EH. Arguments against sterilization of developmentally disabled minors. J Child Neurol 2011 May 29;26(5):654-656. [doi: 10.1177/0883073811402077] [Medline: 21531911]
72. Di Giulio G. Sexuality and people living with physical or developmental disabilities: a review of key issues. Can J Hum Sexual 2003;12(1):53-68 [FREE Full text]
73. McCarthy M. 'I have the jab so I can't be blamed for getting pregnant': contraception and women with learning disabilities. Women's Stud Int Forum 2009 May;32(3):198-208 [FREE Full text] [doi: 10.1016/j.wsif.2009.05.003]
74. E. (Mrs.) v. Eve. Supreme Court of Canada Judgments. 1986. URL: https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/170/index.do [accessed 2022-01-25]
75. Shea CM, Kevles DJ. In the name of eugenics: genetics and the uses of human heredity. History Edu Q 1986;26(4):621. [doi: 10.2307/369028]
76. Park D, Radford JP. From the case files: reconstructing a history of involuntary sterilisation. Disabil Soc 1998 Jun;13(3):317-342 [FREE Full text] [doi: 10.1080/09687599826669] [Medline: 11660707]

77. Tenenbaum EM. To be or to exist: standards for deciding whether dementia patients in nursing homes should engage in intimacy, sex, and adultery. Indiana Law Review. 2009. URL: https://journals.iupui.edu/index.php/inlawrev/article/view/4005 [accessed 2022-01-25]
78. Hajjar R, Kamel H. Sexuality in the nursing home, part 1: attitudes and barriers to sexual expression. J Am Med Directors Assoc 2003 May;4(3):152-156 [FREE Full text] [doi: 10.1016/s1525-8610(04)70325-4]
79. Lindau S, Gavrilova N. Sex, health, and years of sexually active life gained due to good health: evidence from two US population based cross sectional surveys of ageing. Br Med J 2010 Mar 09;340:c810 [FREE Full text] [doi: 10.1136/bmj.c810] [Medline: 20215365]
80. Syme M. The evolving concept of older adult sexual behavior and its benefits. Generations 2014;38(1):35-41 [FREE Full text]
81. Boni-Saenz AA. Sexuality and incapacity. Ohio State Law. 2015. URL: https://scholarship.kentlaw.iit.edu/fac_schol/846/ [accessed 2022-01-25]
82. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
83. Brassolotto J, Howard L, Manduca-Barone A. Sexual expression in Alberta's continuing care homes: capacity, consent, and co-decision-making. Can J Aging 2020 Jan 24;40(1):156-165. [doi: 10.1017/s0714980819000813]
84. Thomas JG. Capacity to consent to sexual activity among residents of long-term mental health treatment programs: perspectives of clients, staff, and administrators. Dissertation, The Florida State University. 2015. URL: https://www.proquest.com/docview/1694580445?pq-origsite=gscholar&fromopenview=true [accessed 2022-01-25]
85. Syme ML, Lichtenberg P, Moye J. Recommendations for sexual expression management in long-term care: a qualitative needs assessment. J Adv Nurs 2016 Oct 18;72(10):2457-2467 [FREE Full text] [doi: 10.1111/jan.13005] [Medline: 27188413]
86. Dukes E, McGuire BE. Enhancing capacity to make sexuality-related decisions in people with an intellectual disability. J Intellect Disabil Res 2009 Aug;53(8):727-734. [doi: 10.1111/j.1365-2788.2009.01186.x] [Medline: 19527433]
87. Lichtenberg P, Strzepek DM. Assessments of institutionalized dementia patients' competencies to participate in intimate relationships. Gerontologist 1990 Feb;30(1):117-120 [FREE Full text] [doi: 10.1093/geront/30.1.117] [Medline: 2311954]
88. Syme ML, Steele D. Sexual consent capacity assessment with older adults. Arch Clin Neuropsychol 2016 Sep 01;31(6):495-505. [doi: 10.1093/arclin/acw046] [Medline: 27480989]
89. Metzger E. Ethics and intimate sexual activity in long-term care. AMA J Ethics 2017 Jul 01;19(7):640-648 [FREE Full text] [doi: 10.1001/journalofethics.2017.19.7.ecas1-1707] [Medline: 28813236]
90. Bogacki DF, Armstrong DJ, Stuempfle P. he social sexual awareness scale (SSAS): development and validation of an instrument to measure capacity to consent to high risk sexual behavior. Am J Forens Psychol 2004;22(2):5-38 [FREE Full text] [doi: 10.1037/t68816-000]
91. Gill M. Rethinking sexual abuse, questions of consent, and intellectual disability. Sex Res Soc Policy 2010 Apr 16;7(3):201-213. [doi: 10.1007/s13178-010-0019-9]
92. Kramers-Olen A. Quantitative assessment of sexual knowledge and consent capacity in people with mild to moderate intellectual disability. S Afr J Psychol 2017 Sep 22;47(3):367-378. [doi: 10.1177/0081246317726457]
93. Srinivasan S, Glover J, Tampi RR, Tampi DJ, Sewell DD. Sexuality and the older adult. Curr Psychiatry Rep 2019 Sep 14;21(10):97. [doi: 10.1007/s11920-019-1090-4] [Medline: 31522296]
94. Spring L. Older women and sexuality – are we still just talking lube? Sex Rel Ther 2014 May 29;30(1):4-9. [doi: 10.1080/14681994.2014.920617]
95. Eastgate G. Sex, consent and intellectual disability. Aust Fam Physician 2005 Mar;34(3):163-166. [Medline: 15799666]
96. Griffith R, Tengnah C. Assessing capacity to consent to sexual relations: a guide for nurses. Br J Community Nurs 2013 Apr;18(4):198-201. [doi: 10.12968/bjcn.2013.18.4.198] [Medline: 23574912]
97. Kramers-Olen A. Sexuality, intellectual disability, and human rights legislation. S Afr J Psychol 2016 Nov 04;46(4):504-516. [doi: 10.1177/0081246316678154]
98. Breen S, Carlson M, Clements G, Everett B, Young J. Supporting sexual health and intimacy in care facilities: guidelines for supporting adults living in long-term care facilities and group homes in British Columbia, Canada. Vancouver Coastal Health Authority. 2009. URL: http://www.vch.ca/Documents/Facilities-licensing-supporting-sexual-health-and-intimacy-in-care-facilities.pdf [accessed 2022-01-26]
99. Joy M, Weiss KJ. Consent for intimacy among persons with neurocognitive impairment. J Am Acad Psychiatry Law 2018 Sep;46(3):286-294. [doi: 10.29158/JAAPL.003763-18] [Medline: 30368460]
100. Director S. Consent's dominion: dementia and prior consent to sexual relations. Bioethics 2019 Nov 28;33(9):1065-1071. [doi: 10.1111/bioe.12652] [Medline: 31463966]
101. Doyle D, Bisson D, Janes N, Lynch H, Martin C. Human sexuality in long-term care. Can Nurse 1999 Jan;95(1):26-29. [Medline: 10401271]
102. Arstein-Kerslake A, Flynn E. Legislating consent. Soc Legal Stud 2015 Aug 17;25(2):225-248 [FREE Full text] [doi: 10.1177/0964663915599051]

103. Marson DC, Huthwaite JS. An unexpected excursion: sexual consent capacity in a nursing home patient with alzheimer's disease. In: Forensic Neuropsychology Casebook. New York: Guilford Press; 2005:146-164.
104. Edgerton RB. The Cloak of Competence. California: University of California Press; 1993.
105. Wertheimer A. Consent to Sexual Relations. Cambridge, MA: Cambridge University Press; 2003.
106. The code: professional standards of practice and behaviour for nurses, midwives and nursing associates. Nursing and Midwifery Council, London. 2015. URL: https://www.nmc.org.uk/standards/code/ [accessed 2022-01-26]
107. Wilkins J. Ethical considerations regarding capacity assessments for consent to sexual activity in older adults with dementia living in residential care settings. Am J Geriatr Psychiatry 2015 Mar;23(3):S113 [FREE Full text] [doi: 10.1016/j.jagp.2014.12.112]
108. Craft A, Bustard S. Living Your Life: The Sex Education and Personal Development Resource for Special Needs Eduation (Revised Ed). London, UK: The Ann Craft Trust UK Brook Publications; 2003.
109. Marshall JE, Helen B. Understanding sexual and gender-based violence against refugees with a communication disability and challenges to accessing appropriate support: a literature review. Manchester Metropolitan University. 2017. URL: https://e-space.mmu.ac.uk/620167/ [accessed 2022-01-25]
110. Keilty J, Connelly G. Making a statement: an exploratory study of barriers facing women with an intellectual disability when making a statement about sexual assault to police. Disab Society 2010 Jul;16(2):273-291 [FREE Full text] [doi: 10.1080/09687590120035843]
111. Faller KC, Henry J. Child sexual abuse: a case study in community collaboration. Child Abuse Neglect 2000 Sep;24(9):1215-1225 [FREE Full text] [doi: 10.1016/s0145-2134(00)00171-x]
112. Weis K, Borges SS. Victimology and rape: the case of the legitimate victim. Issues Criminol 1973;8(2):71-115 [FREE Full text]
113. McCarthy M, Thompson D. A prevalence study of sexual abuse of adults with intellectual disabilities referred for sex education. J Appl Res Intellect Disabil 1997;10(2):105-124 [FREE Full text] [doi: 10.1111/j.1468-3148.1997.tb00012.x]
114. Kuhn D. Intimacy, sexuality, and residents with dementia. Alzheimer's Care Today. 2002. URL: https://tinyurl.com/yckrscyd [accessed 2022-01-25]
115. Boni-Saenz A. Sexual advance directives. Alabama Law Review. 2016. URL: https://scholarship.kentlaw.iit.edu/fac_schol/878/ [accessed 2022-01-25]
116. Schaafsma D, Kok G, Stoffelen JM, Curfs LM. People with intellectual disabilities talk about sexuality: implications for the development of sex education. Sex Disabil 2017;35(1):21-38 [FREE Full text] [doi: 10.1007/s11195-016-9466-4] [Medline: 28250541]
117. Cahn N, Ziettlow A. Makings things fair: an empirical study of how people approach the wealth transmission system. George Washington University Law School. 2015. URL: https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2357&context=faculty_publications [accessed 2022-01-25]
118. Kohn N. Matched preferences and values: a new approach to selecting legal surrogates. San Diego Law Review. 2015. URL: https://digital.sandiego.edu/cgi/viewcontent.cgi?article=1120&context=sdlr [accessed 2022-01-25]
119. Bartlett P. The United Nations convention on the rights of persons with disabilities and mental health law. Mod Law Rev 2012 Sep 03;75(5):752-778 [FREE Full text] [doi: 10.1111/j.1468-2230.2012.00923.x]
120. Dawson J. A realistic approach to assessing mental health laws' compliance with the UNCRPD. Int J Law Psychiatry 2015;40:70-79 [FREE Full text] [doi: 10.1016/j.ijlp.2015.04.003] [Medline: 25982965]
121. Del Villar K. Should supported decision-making replace substituted decision-making? The convention on the rights of persons with disabilities and coercive treatment under Queensland's Mental Health Act 2000. Laws 2015 May 25;4(2):173-200 [FREE Full text] [doi: 10.3390/laws4020173]
122. Leonhard C. The unbearable lightness of consent in contract law. Case Western Reserve Law Review. 2012. URL: https://scholarlycommons.law.case.edu/caselrev/vol63/iss1/9/ [accessed 2022-01-26]

Abbreviations

ABA/APA: American Bar Association and American Psychological Association
LTC: long-term care
MMSE: Mini-Mental Status Exam
SDM: substitute decision maker
UNCRPD: United Nations Convention on the Rights of Persons with Disabilities

Edited by A Mavragani; submitted 24.02.21; peer-reviewed by T Donley, A Boni-Saenz, J Smith; comments to author 23.10.21; revised version received 05.12.21; accepted 17.12.21; published 04.02.22.

Please cite as:
Esmail S, Concannon B
Approaches to Determine and Manage Sexual Consent Abilities for People With Cognitive Disabilities: Systematic Review
Interact J Med Res 2022;11(1):e28137
URL: https://www.i-jmr.org/2022/1/e28137
doi:10.2196/28137
PMID:35119371

Review

Ethical, Legal, and Sociocultural Issues in the Use of Mobile Technologies and Call Detail Records Data for Public Health in the East African Region: Scoping Review

Juliet Nabbuye Sekandi[1,2], MSc, MD, DrPH; Kenya Murray[1,2], MPH; Corinne Berryman[3], MPH; Paula Davis-Olwell[1,2], BS, MA, PhD; Caroline Hurst[3], BSc; Robert Kakaire[1], BA, MPH, DrPH; Noah Kiwanuka[4], MD, MPH, PhD; Christopher C Whalen[1,2], MSc, MD; Erisa Sabakaki Mwaka[5], MD, PhD

[1]Global Health Institute, College of Public Health, University of Georgia, Athens, GA, United States

[2]Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, GA, United States

[3]Department of Health Promotion and Behavior, College of Public Health, University of Georgia, Athens, GA, United States

[4]Department of Epidemiology and Biostatistics, School of Public Health, Makerere University, Kampala, Uganda

[5]Department of Anatomy, School of Biomedical Sciences, College of Health Sciences, Makerere University, Kampala, Uganda

Corresponding Author:
Juliet Nabbuye Sekandi, MSc, MD, DrPH
Global Health Institute
College of Public Health
University of Georgia
Wright Hall, Room 227B
100 Foster Rd
Athens, GA, 30602
United States
Phone: 1 706 542 5257
Email: jsekandi@uga.edu

Abstract

Background: The exponential scale and pace of real-time data generated from mobile phones present opportunities for new insights and challenges across multiple sectors, including health care delivery and public health research. However, little attention has been given to the new ethical, social, and legal concerns related to using these mobile technologies and the data they generate in Africa.

Objective: The objective of this scoping review was to explore the ethical and related concerns that arise from the use of data from call detail records and mobile technology interventions for public health in the context of East Africa.

Methods: We searched the PubMed database for published studies describing ethical challenges while using mobile technologies and related data in public health research between 2000 and 2020. A predefined search strategy was used as inclusion criteria with search terms such as “East Africa,” “mHealth,” “mobile phone data,” “public health,” “ethics,” or “privacy.” We screened studies using prespecified eligibility criteria through a two-stage process by two independent reviewers. Studies were included if they were (1) related to mobile technology use and health, (2) published in English from 2000 to 2020, (3) available in full text, and (4) conducted in the East African region. We excluded articles that (1) were conference proceedings, (2) studies presenting an abstract only, (3) systematic and literature reviews, (4) research protocols, and (5) reports of mobile technology in animal subjects. We followed the five stages of a published framework for scoping reviews recommended by Arksey and O’Malley. Data extracted included title, publication year, target population, geographic region, setting, and relevance to mobile health (mHealth) and ethics. Additionally, we used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Extension for Scoping Reviews checklist to guide the presentation of this scoping review. The rationale for focusing on the five countries in East Africa was their geographic proximity, which lends itself to similarities in technology infrastructure development.

Results: Of the 94 studies identified from PubMed, 33 met the review inclusion criteria for the final scoping review. The 33 articles retained in the final scoping review represent studies conducted in three out of five East African countries: 14 (42%) from Uganda, 13 (39%) from Kenya, and 5 (16%) from Tanzania. Three main categories of concerns related to the use of mHealth

technologies and mobile phone data can be conceptualized as (1) ethical issues (adequate informed consent, privacy and confidentiality, data security and protection), (2) sociocultural issues, and (3) regulatory/legal issues.

Conclusions: This scoping review identified major cross-cutting ethical, regulatory, and sociocultural concerns related to using data from mobile technologies in the East African region. A comprehensive framework that accounts for the critical concerns raised would be valuable for guiding the safe use of mobile technology data for public health research purposes.

(Interact J Med Res 2022;11(1):e35062) doi:10.2196/35062

KEYWORDS

mobile health; public health; ethics; privacy; call detail records; East Africa; Africa; mobile apps; mHealth

Introduction

The exponential scale and pace of real-time data generated from mobile phones present new opportunities and challenges across multiple sectors, including health care delivery and public health [1]. By 2025, mobile subscribers are projected to grow from 477 million to 617 million [2]. One opportunity arising from the growth trajectory in connected mobile technology devices is the evolving field of mobile health (mHealth). The World Health Organization defines mHealth as medical and public health practice supported by mobile devices such as phones, patient monitors, smartphone apps, and functionalities, including voice or SMS text messages [3]. Globally, mHealth solutions have been deployed to address problems such as a shortage of health care workers [4,5], monitoring medication adherence [6,7], and improving quality of health care [8].

In 2020, the National Institutes of Health (NIH) announced a grand opportunity for health research for new funding, titled "Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa)." The goal is to spur new health discoveries and catalyze innovation in health care, public health, and health research in Africa through data science [9]. Additionally, the DS-I Africa announcement called for research to explore the ethical, legal, and social implications associated with data science [10]. These calls highlight new areas for scientific inquiry and recognize the need to address new ethical challenges when using "big data" streams across Africa.

The use of mobile phone call detail records (CDRs), which are time-stamped activity logs that are also known as billable events, is a relatively new concept in health research [11-14]. Although telecommunication companies have used CDRs for more than three decades for customer billing and targeted marketing purposes, the business sector has only recently innovated more ways to use these data in remote banking and financial service delivery. However, the health sector has sparsely leveraged such enormous amounts of mobile phone data for public health, especially in Africa [3-5]. The Data for Development Challenge, Google Flu Trends, Datathon for Social Good, and the Cairo Transport App Challenge, among others, have used archived CDRs to study human activity [6]. In the field of infectious diseases, innovative use of aggregated anonymized data from CDRs revealed informative patterns of the spatiotemporal transmission of various pathogens in the community, including rubella [15,16], malaria [12,17,18], and Ebola [19,20] in Africa. More recently, the COVID-19 pandemic has accelerated this novel utilization of mobile technologies and CDRs for public health surveillance and research [21-23]. However, a lack of well-established standardized ethical and regulatory frameworks has limited the widespread use of such data in health research, particularly in East Africa [24].

Our ongoing NIH-funded Mapping Tuberculosis Transmission Study (MATTS) in Kampala, Uganda, motivated this scoping review. MATTS takes a unique approach that utilizes individual-level CDRs combined with epidemiologic surveys and molecular data to map potential tuberculosis transmission "hot spots." During the initiation of the study, members of the research team from the University of Georgia and Makerere University met with key stakeholders representing the public and private sectors in Uganda. The private sector players were telecommunication companies, and the public players were government agencies or regulatory entities. The goal of the meetings was to learn about existing ethical standards and policy guidelines that apply to utilizing CDRs and other data from mobile technologies for health research. The meetings led to several questions about data access, sharing, transfer, storage, security, privacy, confidentiality, legal, regulatory, and social concerns. In the end, our team surmised that the government and telecommunication entities had existing policies and procedures on personal data that did not necessarily accommodate the utilization of mobile data such as CDRs for health research. Therefore, our research serves as a logical step to generate evidence that is critical for informing the development of policy frameworks that will facilitate better access to mobile data for public health research.

In utilizing CDRs, there are potential points of ethical breaches as personal information flows from a user's mobile phone to a database for health research purposes. For example, when a user activates a connection to mobile phone networks by calling or texting, "event-driven" personal data are generated in real time and stored as CDRs, whereas "network-driven" data are generated when the cell phone is not in active use (see Multimedia Appendix 1). The event-driven metadata in the records contain the caller and recipient phone numbers, date and time of the call, call duration, and the coordinates of the geographical location of the cell tower that serviced the billable event [25]. Each of the data points generated can potentially result in privacy breaches if the data are accessed without ethical safeguards [26]. The Belmont Report provides standard bioethics guidance for researchers based on the core ethical principles (ie, respect for persons, beneficence, and justice) [27]. Even so, we expect that there are new ethical challenges related to the use of the data generated from mobile phones or other technology/devices that warrant further evaluation.

We situated this scoping review in the East African region because of the cross-cutting similarities in sociocultural context, the landscape and stage of mobile technology infrastructure development, and the likelihood of future unified regulatory policies given shared economic interests. Scoping reviews are useful for examining emerging evidence when it is still unclear what other, more specific, questions can be addressed by a more precise systematic review [28]. The objective of the scoping review was to explore the ethical and related concerns that arise from the use of data from CDRs and mobile technology interventions for public health in the context of East Africa.

Our review findings are used to inform an ongoing primary study to explore ethical, legal, sociocultural, and regulatory concerns from the perspective of key stakeholders in the public and private sectors in Uganda. We expect that this work will further inform the development of comprehensive guidelines relevant to mHealth and public health research in the East African region.

Methods

Design

We followed five of the six stages of the framework for scoping reviews proposed by Arksey and O'Malley [29]: (1) identifying the research question; (2) identifying relevant studies; (3) study selection; (4) charting the data; and (5) collating, summarizing, and reporting the results. The sixth stage of the framework is optional, which was not included in this scoping review. It involves a consultation exercise with stakeholders to gain a more holistic view of the issues under evaluations. Additionally, we used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Extension for Scoping Reviews checklist to guide the presentation of this scoping review [30]. We applied existing definitions to identify common ethical concerns related to informed consent, privacy and confidentiality, and data security and protection. We also used team consensus to classify concerns that fit within legal, regulatory, and sociocultural categories.

Operational Definitions

Privacy of health information refers to an individual's right to control the acquisition, uses, or disclosures of his or her identifiable health data for others [31,32]. Confidentiality refers to the obligations of those who receive information to respect the privacy interests of those to whom the data relate [31,32]. Data security refers to the physical, technological, or administrative safeguards or tools used to protect identifiable health data from unwarranted access or disclosure [32]. Data protection is closely related to data security, which refers to adopting appropriate, reasonable, technical, and organizational measures to prevent loss, damage, or unauthorized destruction or processing of personal data [33]. General ethical concerns refer to research processes that fail to uphold or violate the principles of respect for persons, balance of potential benefits and harms, and equitable sharing of benefits and burdens across research groups [27].

Identification of the Research Question

The primary research question was: *what are the ethical, privacy, and data security concerns pertaining to the use of mHealth technologies and mobile phone data for public health research in the East African region?* The rationale for focusing on the five countries in East Africa was that the ethics and mHealth technology landscapes are likely to be similar in several ways within the same region. Therefore, the findings could inform regional guidelines and policies.

Identification of Relevant Studies

A systematic search of the literature was performed on November 18, 2020, in the PubMed database using the search terms "mobile," "ethics," and "East Africa." The detailed search strategy and terms used are provided in Textbox 1. The search was limited to English and included only studies focused on the East Africa region. We considered only articles published from 2000 to 2020 due to the widespread use of mobile technology during this period in the region of interest.

Textbox 1. Full search strategy of the scoping review in the PubMed database.

("Digital Health" [Title/Abstract] OR "Telemedicine" [MeSH (Medical Subject Heading)] OR "Cell Phone" [MeSH] OR "Smartphone" [MeSH] OR "mHealth" [Title/Abstract] OR "telemedicine" [Title/Abstract] OR "remote monitoring" [Title/Abstract] OR "mobile technology" [Title/Abstract] OR "mobile health" [Title/Abstract] OR "video observed therapy" [Title/Abstract] OR "video observed treatment" [Title/Abstract] OR "video directly observed treatment" [Title/Abstract] OR "video directly observed therapy" [Title/Abstract] OR "Call Detail Records" [Title/Abstract] OR "cell phone" [Title/Abstract] OR "cell phone" [Title/Abstract] OR "mobile phones" [Title/Abstract] OR "smartphone" [Title/Abstract] OR "cellphone data" [Title/Abstract] OR "cell phone data" [Title/Abstract] OR "data privacy" [Title/Abstract] OR "cell phone/ethics" [MeSH] OR "Cell Phone/legislation and jurisprudence" [MeSH] OR "Cell Phone Use/legislation and jurisprudence"[Mesh] OR "Cell Phone Use/therapeutic use"[Mesh]) **AND** ("bioethics" [MeSH] OR "ethics" [Mesh] OR "bioethics" [Title/Abstract] OR "public health ethics" [Title/Abstract] OR "ethics" [Title/Abstract] OR "research ethics" [Title/Abstract] OR "autonomy" [Title/Abstract] OR "personal autonomy" [MeSH] OR "principle-based ethics" [MeSH] OR "personal autonomy" [Title/Abstract] OR "principle-based ethics" [Title/Abstract] OR "privacy" [MeSH] OR "confidentiality" [MeSH] OR "computer security" [MeSH] OR "privacy" [Title/Abstract] OR "confidentiality" [Title/Abstract] OR "computer security" [Title/Abstract] OR "data anonymization" [Title/Abstract] OR "data compromising" [Title/Abstract] OR "compromising of data" [Title/Abstract] OR "data security" [Title/Abstract] OR "information protection" [Title/Abstract] OR "Data Protection" [Title/Abstract] OR "Data Security" [Title/Abstract] OR "information protection" [Title/Abstract] OR "Data encryption" [Title/Abstract] OR "Data encryptions" [Title/Abstract] OR "social justice" [Title/Abstract]) **AND** ("Africa" [MeSH] OR "Africa, South of the Sahara" [MeSH] OR "Africa, Eastern" [MeSH] OR "Tanzania" [MeSH] OR "Rwanda" [MeSH] OR "Burundi" [MeSH] OR "Kenya" [MeSH] OR "Uganda" [MeSH] OR "Tanzania" [ALL] OR "Rwanda" [ALL] OR "Burundi" [ALL] OR "Kenya" [ALL] OR "Uganda" [ALL])

Study Selection

Retrieved studies reporting on mobile technology and health in the East African region were included for further review. Two independent reviewers for each article applied the inclusion and exclusion criteria, with an additional author serving as the tiebreaker in the event of uncertainty regarding whether an article met the inclusion or exclusion criteria. Studies were included if they were (1) related to mobile technology and health, (2) published in English from 2000 to 2020, (3) available in full text, and (4) conducted in the East African region. We excluded articles that (1) were conference proceedings, (2) without full text or presenting an abstract only, (3) systematic and literature reviews, (4) research protocols, and (5) reports of mobile technology in animal subjects.

Charting of the Data

Two authors (CH and CB) performed abstract screening. Three authors (KM, CH, and CB) subsequently reviewed full-text articles, and extracted and stored the data in a Microsoft Excel database developed for this review. From the eligible studies, we extracted the title, publication year, journal name, impact factor, target population, geographic region, setting, and relevance to mHealth and ethics. For each stage of the review, two reviewers (KM and CB) independently reviewed each full-text article. Discrepancies were discussed collectively until consensus was reached.

Collating, Summarizing, and Reporting Results

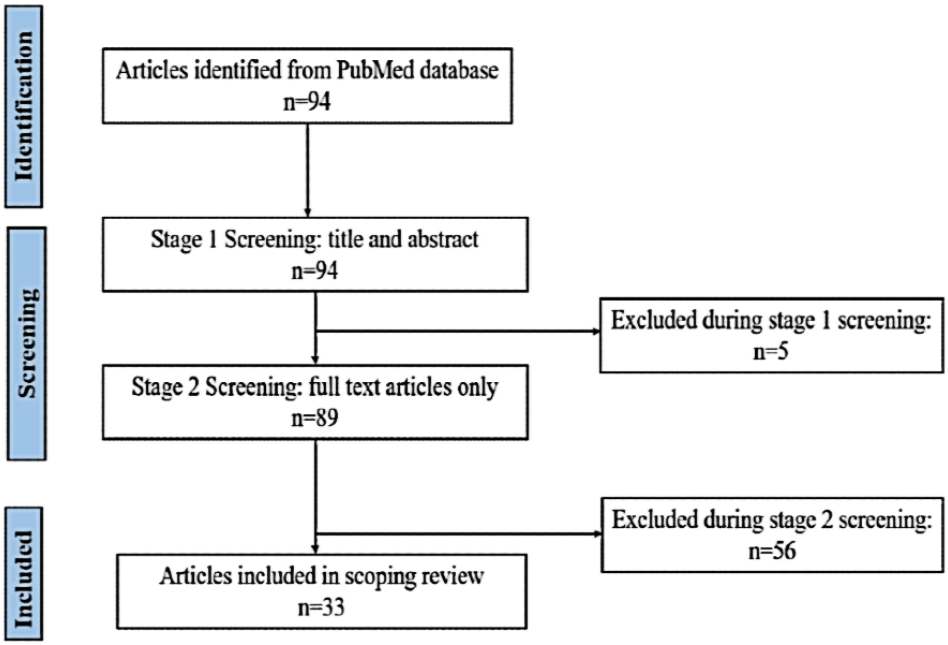
Given the diversity of articles included in the final scoping review, we employed an iterative team discussion approach to the analysis in order to establish the main themes related to ethical and related concerns. We use descriptive statistics to summarize the characteristics of the studies, including country of publication, year of publication, area of focus, study design, and mHealth intervention employed.

Results

Articles Selected

The search identified 94 articles eligible for further review (Figure 1). Of these, 59 did not meet the inclusion criteria and 5 additional studies were excluded based on a review of titles and abstracts. We included 33 studies published in 22 journals in the final scoping review.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram showing article search and screening.



Study Characteristics

The 33 articles retained in the final scoping review represent studies conducted in three out of five East African countries, including 14 (42%) from Uganda and 13 (39%) from Kenya (Table 1). The number of studies published on ethical concerns of mHealth technology use and mobile phone data increased over time, beginning with one in 2009 and climbing to six in 2018. The articles were mostly published in the journal *AIDS and Behavior*, followed by *Journal of Medical Internet Research* and *BMC Public Health*. The majority of mHealth interventions were SMS-based, followed by mobile phone app–based, whereas CDR aggregate data analysis was seldom used among the studies we reviewed. More than half of the studies included were observational in study design, followed by a mixed methods design. Two of the 33 studies included met the eligibility criteria but did not mention ethical concerns beyond compliance with human research and approval from respective institutional review boards.

Three main categories of concerns related to the use of mHealth technologies and mobile phone data can be conceptualized as (1) *ethical issues* (adequate informed consent, privacy and confidentiality, data security and protection), (2) *sociocultural issues*, and (3) *regulatory/legal issues*. These themes are not necessarily mutually exclusive. A detailed summary of each study by theme is provided in Table 2. Ethical concerns were predominantly identified in the included studies, which are discussed in detail below.

Table 1. Summary of articles included in the scoping review for final analysis (N=33).

Characteristics	Studies, n (%)
Country	
Uganda	14 (42)
Kenya	13 (39)
Tanzania	5 (15)
Uganda and Tanzania	1 (3)
Publication year	
2009	1 (3)
2010	1 (3)
2011	1 (3)
2012	2 (6)
2013	4 (12)
2014	1 (3)
2015	3 (9)
2016	4 (12)
2017	5 (15)
2018	6 (18)
2019	5 (15)
mHealth[a] intervention	
SMS messaging	14 (42)
Mobile phone app	5 (15)
Mobile phone	4 (12)
CDR[b] aggregate analysis	2 (6)
Voice and SMS messaging	2 (6)
Mobile phone survey	2 (6)
Tablet	2 (6)
Computer-assisted personal interviewing	1 (3)
Telemedicine	1 (3)
Study design	
Observational	23 (70)
Mixed methods	4 (12)
Descriptive report	2 (6)
Modelling	2 (6)
Quasiexperimental	1 (3)
Randomized controlled trial	1 (3)
Area of focus	
HIV/AIDS	15 (45)
Reproductive health	4 (12)
Infectious disease	3 (9)
Noncommunicable disease	3 (9)
Eye and vision health	2 (6)
Maternal and child health	2 (6)

Characteristics	Studies, n (%)
Data management	1 (3)
Image-based health	1 (3)
Telemedicine	2 (3)

[a]mHealth: mobile health.

[b]CDR: call detail record.

Table 2. Descriptive themes identified regarding ethical, legal, regulatory, and sociocultural concerns of mobile health (mHealth) interventions and mobile phone data use.

Studies and countries	Domain of concern	mHealth intervention	Key recommendation
Adequate informed consent			
Uganda [34,35], Kenya [36]	Cannot consent due to failed understanding of technology	Computer-assisted personal interviewing, mobile phone surveys, telemedicine	Participants' inadequate understanding of the capabilities of mHealth interventions; thus, the question of whether they understood sufficiently to properly consent was raised
Tanzania and Uganda [37]	Consent must be a prerequisite to mHealth interventions	Image-based mHealth app	Not provided
Kenya [38]	Consent needed for different types of prevention of mother-to-child transmission information	SMS text messaging	Not provided
Privacy and confidentiality			
Uganda [39-48]	Password/PIN[a] protection	SMS text messaging, mobile job aid, mobile phone tool	Use of PIN and passwords offers protection of confidentiality. However, the mere presence of passwords may arouse suspicion by intimate partners and others
Kenya [38,49-53]	Phone theft, data breaches	Smartphone app	Not provided
Tanzania [16,18,54,55]	Phone sharing	Smartphone app	Not provided
Data security and protection			
Tanzania [56-58]	Data breaches, phone theft, access rights to protect client data	Smartphone app, mobile job aid, mobile app	mHealth interventions should have an eye toward maternal perception of data security, and with prior and ongoing consultation with community members. Locking phones with a password improves the ability of CHWs[b] to maintain confidentiality of their clients' information, particularly for women who did not want to disclose their use of family planning to their husband or other family members
Kenya [50,59]	Phone theft	SMS text messaging, smartphone ophthalmic exam	Not provided
Uganda [34,35]	Mobile phone numbers linked to national ID cards	Computer-assisted personal interviewing, interactive voice survey	Not provided
Sociocultural			
Tanzania [56,60]	Breach of pregnancy-related information	Mobile phone app	Support from male heads of household may be important in implementing successful mHealth interventions
Uganda [43]	Gender dynamics, delivery of interactive voice survey in voice of opposite sex	Interactive voice survey, SMS messaging	Preference of male or female voices for phone call interventions may vary based on the patient's gender
Regulatory/legal			
Tanzania [56]	Data protection legislation	Mobile phone app	Data protection legislation is needed in regions where local dynamics are important when protecting individuals' health data
Uganda [35]	National ethics guidelines	Interactive voice survey	Not provided

[a]PIN: personal identification number.

[b]CHW: community health worker.

Ethical Concerns

Adequate Informed Consent

Five studies (15%) described adequate consent concerns pertaining to mHealth technology use. In Uganda, participants feared that computer-assisted personal interviews to collect maternal health data also involved audio recording of their responses. To address this concern, the researchers reiterated the need to ensure that participants clearly understand how the technology works and answer any questions during the consent process [34]. In addition, Mwaka and colleagues [35] echoed the importance of understanding consent requirements and how local cultural norms impact participation in mobile phone–based surveys.

In Kenya, researchers questioned the validity of informed consent when there are English words that have no direct equivalent translation in a native local dialect such as Kiswahili. For example, one study [36] found that more than half of the participants did not understand several translated words such as "videoconference," "store and forward," "digital photograph," "wireless," or "email" in either language. Thus, a combination of words was needed to explain telemedicine terms; nevertheless, the participants seemed to have a poor understanding, thereby threatening the validity of the consent process [36]. Jennings and colleagues [38] further pointed out the need to include a standardized technology orientation for prevention of mother-to-child transmission (PMTCT) to ensure adequate informed consent. The technology orientation should be applicable to all users regardless of their specific health condition.

Privacy and Confidentiality

Among the 33 included studies, 22 (67%) described privacy and confidentiality concerns regarding utilizing mHealth technology and mobile phone data. Specifically, two studies were not focused directly on end-user experiences regarding privacy and confidentiality, but highlighted the need for additional research on how mHealth technologies could improve patient privacy [61,62].

Four studies conducted in Uganda explored an SMS-based intervention to support care among people living with HIV/AIDS (PLWHA) and revealed mixed results. Some participants had no privacy concerns related to disclosure of their personal health–related information, including their HIV status [39-42]. For example, participants believed that using personal identification numbers or passwords increased protection of their individual data. End users completing health surveys using interactive voice response (IVR) technology in Uganda also indicated that their confidentiality was increased by using mobile phone–based surveys [43].

Moreover, there were increased privacy concerns in situations where the participants had not disclosed their HIV status to family members or friends, especially when a phone was shared. Three studies in Uganda described privacy and confidentiality concerns pertaining to the use of SMS-based reminders for medication adherence, fearing that they could result in unintended HIV status disclosure when messages are seen by others [44,45,54]. Similar privacy and confidentiality concerns were expressed among pregnant HIV-infected women in Kenya indicating opposition to receiving messages with HIV-related terms such as "infection" or "medication," in fear that it would disclose their HIV status [49]. Additionally, perceptions of data privacy may differ between health care workers and patients. Two studies in Uganda focused on the impact of using mHealth interventions to improve care among PLWHA. The community health workers (CHWs) expressed concerns of threat to data confidentiality if the phone is stolen or multimedia capabilities are misused, while the participants perceived the smartphone to protect their confidentiality [46,47].

In Kenya, end users of mHealth technologies supporting PMTCT, antiretroviral therapy (ART) adherence, and family planning indicated that their personal privacy was protected while discussing sensitive health matters via call or text compared to face-to-face encounters [38,50-52]. Similarly, two other studies in Kenya [55] and Tanzania [53], respectively, found minimal confidentiality concerns among participants using a smartphone to support family planning services and HIV testing. Moreover, participants indicated that they were more likely to answer questions honestly using the mHealth intervention compared to speaking in person with a provider or attending a clinic [53,55].

Aggregated and anonymized CDR data may present minimal ethical and privacy concerns when compared to individual-level or other mobile phone data. Tatem and colleagues [18] analyzed aggregate-level CDRs in Tanzania to inform malaria elimination. The researchers did not explicitly explore privacy and confidentiality concerns, but they noted that aggregated and anonymized CDRs could support analysis without compromising the privacy of the mobile phone users. Wesolowski and colleagues [16] also used aggregated and anonymized mobile phone data in the form of CDRs combined with community health surveys in Kenya to study the impact of travel on the transmission dynamics of malaria. Again, the researchers did not examine privacy concerns directly, but pointed out privacy concerns as a potential barrier to the availability of CDR data for health research.

Data Security and Protection

Seven of the 33 studies (21%) described data security and protection concern among users. Two studies in Tanzania described mixed levels of data security and protection concerns between CHWs and their female clients during the use of mHealth interventions to improve the quality of maternal and family planning services [56,57]. CHWs had positive views when using smartphones to collect reproductive health data from their clients. They noted that using mobile technology improved data security and protection of clients' privacy. However, some of the female clients expressed concerns about potential data breaches if personal information related to their pregnancy is stored on the smartphone and accidentally shared. In Kenya, a few data security and protection concerns were reported when researchers analyzed the perceptions of HIV-infected patients toward a mobile phone messaging intervention to support ART adherence [50]. In rural Uganda, Mercader and colleagues [34] assessed the acceptability of computer-assisted personal interviewing for maternal, newborn,

and child health surveys, and concluded that respondents perceived personal health data recorded on tablets as more secure than paper surveys.

In Uganda, the new requirement to link the National Identification Number (NIN) with one's personal phone number is perceived as a threat to data privacy and security. Mwaka and colleagues [35] explored the views of 14 key stakeholders regarding the potential challenges for obtaining consent during implementation of mobile phone surveys for IVR-based noncommunicable disease research. Some key informants highlighted overlapping concerns regarding data security and privacy where the identity of the respondent can be traced through the linkage between the NIN and phone number. In Kenya, health care providers and stakeholders expressed data security and protection concerns in using a smartphone-based ophthalmic exam, recommending a secure data encryption system to protect personal data [59].

In Tanzania, Steiner and colleagues [58] reviewed four field applications of a mobile device deployed to support low-resourced countries with data entry and project monitoring. The researchers created a data collection platform with user-defined access rights to ensure the data security and protection of clients [58]. During a 3-day workshop focused on addressing ethical issues of safety and privacy among mHealth developers and users in low-resource settings, the authors gathered solutions from 27 mHealth stakeholders from various geographical regions, including Uganda and Tanzania [37]. The stakeholders concluded that patient authorization and informed consent must remain prerequisites to mHealth interventions, specifically during the implementation and scale-up stages.

Sociocultural Issues

Three mHealth studies (9%) described concerns related to complex gender dynamics and sociocultural beliefs. In Tanzania, two studies revealed gender-based power imbalances among women and men. Female participants reported low mobile phone ownership and expressed concerns about spouses being suspicious that if a wife owned a phone, it may facilitate infidelity in the relationship [56,60]. In Uganda, female participants expressed sociocultural concerns pertaining to using mHealth interventions to support health. For example, some participants were reluctant to use IVR for health purposes because of community beliefs that there are "*evil spirits*" that could claim the life of the recipient if one received unknown calls. In addition, women participants were particularly apprehensive about IVR surveys being sent to their cell phones if they were being administered in a male voice; they feared that this would raise suspicion of cheating by their spouses and could spark domestic violence [43]. Phone sharing is particularly common in countries within Africa, thus creating another potential source of privacy breaches and sociocultural concerns.

Regulatory and Legal Issues

Two studies raised regulatory and legal concerns regarding using mHealth technology. These studies reflect the perspectives of the researchers governing the studies and the key stakeholders surveyed. In Tanzania, researchers concluded that the findings of their study on reproductive health provided impetus for stronger data protection legislation in regions such as rural Tanzania [56]. Protecting individuals' health data is particularly important given the inherent sociocultural beliefs surrounding the secrecy of health information [56]. In Uganda, researchers indicated that the processes for mobile phone surveys should follow standard international regulatory guidelines for sharing personal information while also ensuring alignment with existing local laws and policies [35].

Discussion

Principal Findings

The aim of this scoping review was to synthesize the current state of evidence on the ethical, sociocultural, legal, and regulatory concerns related to the use of mHealth technologies and mobile phone data for public health research in the East Africa region. Our review builds on a growing body of work examining concerns with mHealth data security, privacy, and confidentiality [24]. To our knowledge, this is the first study to synthesize end-user and public health research concerns pertaining to mobile phone data use in the East Africa region. Five interrelated themes emerged as key concerns with the use of mobile phone data and mHealth interventions: adequate informed consent, privacy and confidentiality, data security and protection, sociocultural issues, and regulatory/legal issues. Of equal concern is the collection, use, and sharing of personal information such as CDRs to third parties without the notice or consent of consumers. Countries need to put in place legislation to protect the privacy and secure these personal data while facilitating their use for the good of the populations. In the future, these themes and others could inform the formulation of a framework for ethical regulatory policies using mHealth in the East African Community (EAC) region.

African countries have rapidly adopted the use of mobile technologies to support their daily needs such as social, economic, education, and travel needs, among many others. Specifically, the East African region shows a dominance in the adoption of mHealth programs to overcome some structural barriers in the health system [63]. Studies in high-income countries deploying mHealth interventions have revealed concerns mostly related to data privacy, security, storage, and transmission [64-66]. In contrast, most studies in the East African region highlight concerns about the potential disclosure of personal health information at the user interface, such as reading text messages. The pace of development and access to mobile technology could partly explain these differences in concerns between high- and low-income settings.

Sociocultural issues seem to influence the differential levels of positive and negative perceptions around data privacy and confidentiality concerns by gender and user group across the East African region (see Multimedia Appendix 2). For example, privacy and confidentiality concerns in Uganda and Tanzania highlighted gender norms and cultural beliefs that exist around secrecy about pregnancy and childbirth due to a fear of supernatural powers that could cause harm to the mother and unborn child. These beliefs could inhibit the willingness to share pregnancy-related personal information via the phone [67,68]. In addition, the issues of phone sharing due to limited mobile

phone ownership among women could precipitate sociocultural tensions. For example, spousal conflicts and violence, especially directed toward female mHealth users, have been reported when women receive various forms of health support from male health care workers. Although mHealth interventions are beneficial, implementers must be mindful to anticipate the issues that may result from phone sharing in the cultural context of Africa [69].

mHealth data have the potential to enhance or jeopardize privacy. For example, the HIV-related stigma and discrimination among PLWHA in Africa could be attenuated if disclosure of HIV status is kept confidential. The findings of this scoping review suggest that among PLWHA mHealth interventions, supporting patient care should be accompanied by unique personal identification numbers or passwords to facilitate protection of privacy and confidentiality of HIV-related information. Several studies from Kenya and Uganda revealed that populations experienced difficulty in understanding the capability of mHealth interventions, which in turn raises questions about the informed consent process [34-36].

The lack of comprehensive ethical frameworks to guide the use of mHealth in public health research and practice remains a challenge across countries in the EAC region [70]. Very few studies included in this review explored the legal concerns of data use beyond pointing to the need for legal and regulatory policies for data protection and privacy of personal data. In 2018, MEASURE Evaluations published a report of guidelines developed for Kenya and Tanzania to help mHealth program managers and Ministry of Health officials systematically address mHealth data privacy and security issues [24]. The authors acknowledged that the guidelines were limited in scope, only serving as a building block that would allow stakeholders to be informed, and guide the teams responsible for developing and implementing responsible data practices, especially data security and privacy.

Legal regulation of personal health information within the EAC is not uniformly developed. The Republic of Uganda [33] and the Republic of Kenya [71] passed their data protection laws in 2019, whereas Rwanda passed a similar law in 2021 [72]. Tanzania currently has a draft data protection bill that has not yet passed into law, whereas Burundi does not have a law that specifically regulates personal data protection. Based on the experience of our health research team in Uganda, the new law is not yet operationalized into clear policy guidance, particularly in the context of public health research and practice. As such, we have encountered barriers to access CDR data from telecommunication companies even when we present written and signed consent from the cell phone owners. To the greatest extent possible, each nation needs to cooperate to set up regional mobile technology standards reflecting its local situations and government policy regulations [65]. The results of our scoping review serve to highlight concerns and gaps that must be addressed to create an enabling policy and regulatory environment for public health researchers. Additionally, public views on the use of CDRs in health research must also be explored.

Future Implications for Public-Private Partnerships

Our work has implications for public-private partnerships, given the potential mutual benefits from CDRs, a largely untapped emerging data type that is collected by telecommunication companies in the EAC region. First, the public and private sectors could share expert human capital financial and infrastructural resources to catalyze the growth of these entities within the region. For example, the public health research enterprise could quickly access a large quantity of rich data on spatial and temporal mobility patterns of the population, which are routinely collected but underutilized by telecommunication companies. This approach to collecting specific data would likely be far more efficient and cost-effective than traditional methods. Second, the private sector could gain new business insights about the populations they serve based on a range of geospatial or data-mining analyses and interpretations generated by public researchers and data scientists. In Africa, many lessons can be learned from major innovations in banking and agriculture that have already spawned from the extensive digital infrastructure. Lastly, we envision a unique opportunity for the public health and private sectors to engage jointly with regulatory policymakers to advance data governance policies for mutual benefit. It is important to note that this process will likely be dynamic given the rapidly evolving mobile technology landscape.

Limitations

This scoping review has some limitations. First, we acknowledge that we might have missed studies relevant to mHealth and ethical, legal, or privacy concerns in the East Africa region if they were published in electronic databases other than PubMed, in a language other than English, or outside of the period of our study. Second, studies published from Rwanda and Burundi may have been excluded from this review specifically because English was specified as the only language for articles that met other inclusion criteria. In this case, any published articles written in French could have been excluded; however, during our search, we did not identify any such articles in the initial screening. Overall, we believe that the majority of the public health research articles are accessible through PubMed. Therefore, we do not expect that our findings and conclusions were significantly influenced by any papers we might have missed.

Conclusions

This scoping review identified major cross-cutting ethical, regulatory, and sociocultural concerns related to use of data from mobile technologies in the East African region. A comprehensive framework that accounts for ethical, sociocultural, legal, and regulatory concerns in the cultural context of the EAC region is needed to guide the safe use of mobile technology data for public health research purposes.

Acknowledgments

We acknowledge the contributions of Abirami Balajee in coordination of the initial PubMed search, literature review, and assistance in the initial data abstraction process. This study was funded through the National Institutes of Health/National Institute of Allergy and Infectious Diseases administrative supplement (2R01AI093856-06AI).

Authors' Contributions

Conceptualization: JNS, CCW, RK, ESM, NK; Data Review and Abstraction: CB, CH, KM; Methodology and Formal Analysis: KM, CB, CH, JNS, PDO; Supervision: JNS, PD-O; Writing original draft: KM, JNS; Writing reviewing editing to final manuscript: JNS, KM, CB, PD-O, CH, RK, NK, CCW, ESM. All authors approved the final version of the manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data sharing pathways for mobile phone records in mHealth and public health research.
[DOCX File , 157 KB - ijmr_v11i1e35062_app1.docx]

Multimedia Appendix 2
Positive and negative perceptions about privacy, confidentiality, and safety of data collected with mobile phones.
[DOCX File , 137 KB - ijmr_v11i1e35062_app2.docx]

References

1. Schuttner L, Sindano N, Theis M, Zue C, Joseph J, Chilengi R, et al. A mobile phone-based, community health worker program for referral, follow-up, and service outreach in rural Zambia: outcomes and overview. Telemed J E Health 2014 Aug;20(8):721-728 [FREE Full text] [doi: 10.1089/tmj.2013.0240] [Medline: 24926815]
2. Parasuraman S, Sam A, Yee SK, Chuon BC, Ren L. Smartphone usage and increased risk of mobile phone addiction: a concurrent study. Int J Pharma Investig 2017;7(3):125. [doi: 10.4103/jphi.jphi_56_17]
3. The Mobile Economy sub-Saharan Africa. GSMA. 2020. URL: https://www.gsma.com/mobileeconomy/sub-saharan-africa/ [accessed 2021-05-11]
4. WHO Global Observatory for eHealth. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. Geneva: World Health Organization; 2011.
5. Källander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten Asbroek AHA, et al. Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries: a review. J Med Internet Res 2013 Jan 25;15(1):e17 [FREE Full text] [doi: 10.2196/jmir.2130] [Medline: 23353680]
6. Dawood N, Mahomed Asmail F, Louw C, Swanepoel DW. Mhealth hearing screening for children by non-specialist health workers in communities. Int J Audiol 2021 Apr 12;60(sup1):S23-S29. [doi: 10.1080/14992027.2020.1829719] [Medline: 33043733]
7. Sekandi JN, Buregyeya E, Zalwango S, Dobbin KK, Atuyambe L, Nakkonde D, et al. Video directly observed therapy for supporting and monitoring adherence to tuberculosis treatment in Uganda: a pilot cohort study. ERJ Open Res 2020 Jan 06;6(1):00175-2019 [FREE Full text] [doi: 10.1183/23120541.00175-2019] [Medline: 32280670]
8. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug 06;1(2):160-171. [doi: 10.9745/ghsp-d-13-00031]
9. The NIOH. Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) Research Hubs (U54–Clinical Trial Optional). National Institutes of Health Department of Health and Human Services. URL: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-015.html [accessed 2021-05-11]
10. Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa): Ethical, Legal, and Social Implications Research (U01 – Clinical Trial Not Allowed). National Institutes of Health Department of Health and Human Services. URL: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-017.html [accessed 2021-05-11]
11. Milusheva S. Managing the spread of disease with mobile phone data. J Dev Econ 2020 Nov;147:102559 [FREE Full text] [doi: 10.1016/j.jdeveco.2020.102559] [Medline: 33144750]
12. Pindolia DK, Garcia AJ, Wesolowski A, Smith DL, Buckee CO, Noor AM, et al. Human movement data for malaria control and elimination strategic planning. Malar J 2012 Jun 18;11(1):205 [FREE Full text] [doi: 10.1186/1475-2875-11-205] [Medline: 22703541]
13. Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, Snow RW, et al. Quantifying the impact of human mobility on malaria. Science 2012 Oct 12;338(6104):267-270 [FREE Full text] [doi: 10.1126/science.1223467] [Medline: 23066082]

14. Jones KH, Daniels H, Heys S, Ford DV. Challenges and potential opportunities of mobile phone call detail records in health research: review. JMIR Mhealth Uhealth 2018 Jul 19;6(7):e161 [FREE Full text] [doi: 10.2196/mhealth.9974] [Medline: 30026176]
15. The Mobile Economy West Africa. GMSA. 2018. URL: https://tinyurl.com/2nzc3c69 [accessed 2021-05-11]
16. Wesolowski A, Metcalf CJE, Eagle N, Kombich J, Grenfell BT, Bjørnstad ON, et al. Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. Proc Natl Acad Sci U S A 2015 Sep 01;112(35):11114-11119 [FREE Full text] [doi: 10.1073/pnas.1423542112] [Medline: 26283349]
17. Buckee CO, Wesolowski A, Eagle NN, Hansen E, Snow RW. Mobile phones and malaria: modeling human and parasite travel. Travel Med Infect Dis 2013 Jan;11(1):15-22 [FREE Full text] [doi: 10.1016/j.tmaid.2012.12.003] [Medline: 23478045]
18. Tatem AJ, Qiu Y, Smith DL, Sabot O, Ali AS, Moonen B. The use of mobile phone data for the estimation of the travel patterns and imported Plasmodium falciparum rates among Zanzibar residents. Malar J 2009 Dec 10;8(1):287 [FREE Full text] [doi: 10.1186/1475-2875-8-287] [Medline: 20003266]
19. Leidig JT, Thelsen C, Vogel N, Scripps J, Wolffe G. Modeling Ebola virus diffusion in Senegal using mobile phone datasets and agent-based simulation. 2015 Presented at: NetMob 2015; April 8-10, 2015; Boston, MA.
20. Kramer AM, Pulliam JT, Alexander LW, Park AW, Rohani P, Drake JM. Spatial spread of the West Africa Ebola epidemic. R Soc Open Sci 2016 Aug;3(8):160294 [FREE Full text] [doi: 10.1098/rsos.160294] [Medline: 27853607]
21. Attipoe-Dorcoo S, Delgado R, Gupta A, Bennet J, Oriol NE, Jain SH. Mobile health clinic model in the COVID-19 pandemic: lessons learned and opportunities for policy changes and innovation. Int J Equity Health 2020 May 19;19(1):73 [FREE Full text] [doi: 10.1186/s12939-020-01175-7] [Medline: 32429920]
22. Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, et al. Mobile health apps on COVID-19 launched in the early days of the pandemic: content analysis and review. JMIR Mhealth Uhealth 2020 Sep 16;8(9):e19796 [FREE Full text] [doi: 10.2196/19796] [Medline: 32609622]
23. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, De Nadai M, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv 2020 Jun 05;6(23):eabc0764 [FREE Full text] [doi: 10.1126/sciadv.abc0764] [Medline: 32548274]
24. Spigel L, Wambugu S, Villella C. mHealth Data Security, Privacy, and Confidentiality: Guidelines for Program Implementers and Policymakers. Measure Evaluation. 2018. URL: https://www.measureevaluation.org/resources/publications/ms-17-125a.html [accessed 2021-05-11]
25. Oliver N, Matic A, Frias-Martinez E. Mobile network data for public health: opportunities and challenges. Front Public Health 2015 Aug 07;3:189. [doi: 10.3389/fpubh.2015.00189] [Medline: 26301211]
26. Letouzé E, Vinck P, World Bank Group, D4D team. The law, politics and ethics of cell phone data analytics. White Paper series: Data-Pop Alliance. 2015. URL: https://tinyurl.com/4rn2eav3 [accessed 2011-05-11]
27. The Belmont report: ethical principles and guidelines for the protection of human subjects of research. US Department of Health & Human Services. 1978. URL: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html [accessed 2021-05-11]
28. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
29. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Method 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
30. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Tunçalp, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 02;169(7):467-473. [doi: 10.7326/m18-0850]
31. Arora S, Yttri J, Nilse W. Privacy and security in mobile health (mHealth) research. Alcohol Res 2014;36(1):143-151 [FREE Full text] [Medline: 26259009]
32. Cohn SP, National Committee on Vital and Health Statistics. Privacy and confidentiality in the Nationwide Health Information Network. American Health Information Management Association. 2006 Jun 22. URL: https://library.ahima.org/doc?oid=75960#.YoJAPJPMKi5 [accessed 2021-05-11]
33. The Data Protection and Privacy Act 2019. Uganda Ministry of Information and Communications Technology. 2019. URL: https://ict.go.ug/wp-content/uploads/2019/03/Data-Protection-and-Privacy-Act-2019.pdf [accessed 2021-05-11]
34. Mercader HFG, Kabakyenga J, Katuruba DT, Hobbs AJ, Brenner JL. Female respondent acceptance of computer-assisted personal interviewing (CAPI) for maternal, newborn and child health coverage surveys in rural Uganda. Int J Med Inform 2017 Feb;98:41-46. [doi: 10.1016/j.ijmedinf.2016.11.009] [Medline: 28034411]
35. Mwaka E, Nakigudde J, Ali J, Ochieng J, Hallez K, Tweheyo R, et al. Consent for mobile phone surveys of non-communicable disease risk factors in low-resource settings: an exploratory qualitative study in Uganda. Mhealth 2019 Aug;5:26. [doi: 10.21037/mhealth.2019.07.05] [Medline: 31559271]
36. Odhiambo R, Mars M. Patients' understanding of telemedicine terms required for informed consent when translated into Kiswahili. BMC Public Health 2018 May 03;18(1):588 [FREE Full text] [doi: 10.1186/s12889-018-5499-1] [Medline: 29720139]

37. Laflamme L, Chipps J, Fangerau H, Juth N, Légaré F, Sawe HR, et al. Targeting ethical considerations tied to image-based mobile health diagnostic support specific to clinicians in low-resource settings: the Brocher proposition. Glob Health Action 2019 Sep 18;12(1):1666695 [FREE Full text] [doi: 10.1080/16549716.2019.1666695] [Medline: 31532350]
38. Jennings L, Ong'ech J, Simiyu R, Sirengo M, Kassaye S. Exploring the use of mobile phone technology for the enhancement of the prevention of mother-to-child transmission of HIV program in Nyanza, Kenya: a qualitative study. BMC Public Health 2013 Dec 05;13(1):1131 [FREE Full text] [doi: 10.1186/1471-2458-13-1131] [Medline: 24308409]
39. Campbell JI, Aturinda I, Mwesigwa E, Burns B, Santorino D, Haberer JE, et al. The Technology Acceptance Model for Resource-Limited Settings (TAM-RLS): a novel framework for mobile health interventions targeted to low-literacy end-users in resource-limited settings. AIDS Behav 2017 Nov 18;21(11):3129-3140 [FREE Full text] [doi: 10.1007/s10461-017-1765-y] [Medline: 28421356]
40. Siedner MJ, Haberer JE, Bwana MB, Ware NC, Bangsberg DR. High acceptability for cell phone text messages to improve communication of laboratory results with HIV-infected patients in rural Uganda: a cross-sectional survey study. BMC Med Inform Decis Mak 2012 Jun 21;12(1):56 [FREE Full text] [doi: 10.1186/1472-6947-12-56] [Medline: 22720901]
41. Siedner MJ, Santorino D, Haberer JE, Bangsberg DR. Know your audience: predictors of success for a patient-centered texting app to augment linkage to HIV care in rural Uganda. J Med Internet Res 2015 Mar 24;17(3):e78 [FREE Full text] [doi: 10.2196/jmir.3859] [Medline: 25831269]
42. Kunutsor S, Walley J, Katabira E, Muchuro S, Balidawa H, Namagala E, et al. Using mobile phones to improve clinic attendance amongst an antiretroviral treatment cohort in rural Uganda: a cross-sectional and prospective study. AIDS Behav 2010 Dec 11;14(6):1347-1352. [doi: 10.1007/s10461-010-9780-2] [Medline: 20700644]
43. Ssemugabo C, Rutebemberwa E, Kajungu D, Pariyo GW, Hyder AA, Gibson DG. Acceptability and use of interactive voice response mobile phone surveys for noncommunicable disease behavioral risk factor surveillance in rural Uganda: qualitative study. JMIR Form Res 2019 Dec 03;3(4):e15000 [FREE Full text] [doi: 10.2196/15000] [Medline: 31793889]
44. Musiimenta A, Atukunda EC, Tumuhimbise W, Pisarski EE, Tam M, Wyatt MA, et al. Acceptability and feasibility of real-time antiretroviral therapy adherence interventions in rural Uganda: mixed-method pilot randomized controlled trial. JMIR Mhealth Uhealth 2018 May 17;6(5):e122. [doi: 10.2196/mhealth.9031] [Medline: 29773527]
45. Rana Y, Haberer J, Huang H, Kambugu A, Mukasa B, Thirumurthy H, et al. Short message service (SMS)-based intervention to improve treatment adherence among HIV-positive youth in Uganda: focus group findings. PLoS One 2015 Apr 16;10(4):e0125187 [FREE Full text] [doi: 10.1371/journal.pone.0125187] [Medline: 25881059]
46. Chang LW, Kagaayi J, Arem H, Nakigozi G, Ssempijja V, Serwadda D, et al. Impact of a mHealth intervention for peer health workers on AIDS care in rural Uganda: a mixed methods evaluation of a cluster-randomized trial. AIDS Behav 2011 Nov 8;15(8):1776-1784 [FREE Full text] [doi: 10.1007/s10461-011-9995-x] [Medline: 21739286]
47. Chang LW, Njie-Carr V, Kalenge S, Kelly JF, Bollinger RC, Alamo-Talisuna S. Perceptions and acceptability of mHealth interventions for improving patient care at a community-based HIV/AIDS clinic in Uganda: a mixed methods study. AIDS Care 2013 May 08;25(7):874-880 [FREE Full text] [doi: 10.1080/09540121.2013.774315] [Medline: 23452084]
48. Hermans S, Elbireer S, Tibakabikoba H, Hoefman B, Manabe Y. Text messaging to decrease tuberculosis treatment attrition in TB-HIV coinfection in Uganda. Patient Prefer Adherence 2017;11:1479-1487. [doi: 10.2147/PPA.S135540] [Medline: 28919720]
49. Ronen K, Unger JA, Drake AL, Perrier T, Akinyi P, Osborn L, et al. SMS messaging to improve ART adherence: perspectives of pregnant HIV-infected women in Kenya on HIV-related message content. AIDS Care 2018 Apr 18;30(4):500-505 [FREE Full text] [doi: 10.1080/09540121.2017.1417971] [Medline: 29254362]
50. Kinyua F, Kiptoo M, Kikuvi G, Mutai J, Meyers AF, Muiruri P, et al. Perceptions of HIV infected patients on the use of cell phone as a tool to support their antiretroviral adherence; a cross-sectional study in a large referral hospital in Kenya. BMC Public Health 2013 Oct 21;13(1):987 [FREE Full text] [doi: 10.1186/1471-2458-13-987] [Medline: 24143931]
51. Vahdat HL, L'Engle KL, Plourde KF, Magaria L, Olawo A. There are some questions you may not ask in a clinic: providing contraception information to young people in Kenya using SMS. Int J Gynaecol Obstet 2013 Nov 02;123(Suppl 1):e2-e6 [FREE Full text] [doi: 10.1016/j.ijgo.2013.07.009] [Medline: 24012514]
52. Velloza J, Ngure K, Kiptinness C, Quame-Amaglo J, Thuo N, Dew K, et al. A clinic-based tablet application to support safer conception among HIV serodiscordant couples in Kenya: feasibility and acceptability study. Mhealth 2019 Feb;5:4. [doi: 10.21037/mhealth.2019.01.04] [Medline: 30976596]
53. Dev R, Woods NF, Unger JA, Kinuthia J, Matemo D, Farid S, et al. Acceptability, feasibility and utility of a Mobile health family planning decision aid for postpartum women in Kenya. Reprod Health 2019 Jul 08;16(1):97 [FREE Full text] [doi: 10.1186/s12978-019-0767-9] [Medline: 31286989]
54. Atukunda EC, Musiimenta A, Musinguzi N, Wyatt MA, Ashaba J, Ware NC, et al. Understanding patterns of social support and their relationship to an ART adherence intervention among adults in rural southwestern Uganda. AIDS Behav 2017 Feb 26;21(2):428-440 [FREE Full text] [doi: 10.1007/s10461-016-1559-7] [Medline: 27671479]
55. Conserve DF, Muessig KE, Maboko LL, Shirima S, Kilonzo MN, Maman S, et al. Mate Yako Afya Yako: Formative research to develop the Tanzania HIV self-testing education and promotion (Tanzania STEP) project for men. PLoS One 2018 Aug 27;13(8):e0202521 [FREE Full text] [doi: 10.1371/journal.pone.0202521] [Medline: 30148846]

56. Hackett KM, Kazemi M, Sellen DW. Keeping secrets in the cloud: mobile phones, data security and privacy within the context of pregnancy and childbirth in Tanzania. Soc Sci Med 2018 Aug;211:190-197. [doi: 10.1016/j.socscimed.2018.06.014] [Medline: 29960170]
57. Braun R, Lasway C, Agarwal S, L'Engle K, Layer E, Silas L, et al. An evaluation of a family planning mobile job aid for community health workers in Tanzania. Contraception 2016 Jul;94(1):27-33. [doi: 10.1016/j.contraception.2016.03.016] [Medline: 27039033]
58. Steiner A, Hella J, Grüninger S, Mhalu G, Mhimbira F, Cercamondi CI, et al. Managing research and surveillance projects in real-time with a novel open-source eManagement tool designed for under-resourced countries. J Am Med Inform Assoc 2016 Sep;23(5):916-923. [doi: 10.1093/jamia/ocv185] [Medline: 26911813]
59. Lodhia V, Karanja S, Lees S, Bastawrous A. Acceptability, usability, and views on deployment of Peek, a mobile phone mHealth intervention for eye care in Kenya: qualitative study. JMIR Mhealth Uhealth 2016 May 09;4(2):e30 [FREE Full text] [doi: 10.2196/mhealth.4746] [Medline: 27160779]
60. Thobias J, Kiwanuka A. Design and implementation of an m-health data model for improving health information access for reproductive and child health services in low resource settings using a participatory action research approach. BMC Med Inform Decis Mak 2018 Jun 25;18(1):45 [FREE Full text] [doi: 10.1186/s12911-018-0622-x] [Medline: 29941008]
61. Kivuti-Bitok LW, McDonnell G, Pokhariyal GP, Roudsari AV. Self-reported use of internet by cervical cancer clients in two National Referral Hospitals in Kenya. BMC Res Notes 2012 Oct 09;5(1):559 [FREE Full text] [doi: 10.1186/1756-0500-5-559] [Medline: 23046538]
62. Henry JV, Winters N, Lakati A, Oliver M, Geniets A, Mbae SM, et al. Enhancing the supervision of community health workers with WhatsApp mobile messaging: qualitative findings from 2 low-resource settings in Kenya. Glob Health Sci Pract 2016 May 26;4(2):311-325. [doi: 10.9745/ghsp-d-15-00386]
63. Lee S, Cho Y, Kim S. Mapping mHealth (mobile health) and mobile penetrations in sub-Saharan Africa for strategic regional collaboration in mHealth scale-up: an application of exploratory spatial data analysis. Global Health 2017 Aug 22;13(1):63 [FREE Full text] [doi: 10.1186/s12992-017-0286-9] [Medline: 28830540]
64. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
65. Whittaker R. Issues in mHealth: findings from key informant interviews. J Med Internet Res 2012 Oct 02;14(5):e129 [FREE Full text] [doi: 10.2196/jmir.1989] [Medline: 23032424]
66. Schnall R, Higgins T, Brown W, Carballo-Dieguez A, Bakken S. Trust, perceived risk, perceived ease of use and perceived usefulness as factors related to mHealth technology use. Stud Health Technol Inform 2015;216:467-471 [FREE Full text] [Medline: 26262094]
67. Spangler SA, Bloom SS. Use of biomedical obstetric care in rural Tanzania: the role of social and material inequalities. Soc Sci Med 2010 Aug;71(4):760-768. [doi: 10.1016/j.socscimed.2010.05.025] [Medline: 20579797]
68. Mgata S, Maluka SO. Factors for late initiation of antenatal care in Dar es Salaam, Tanzania: a qualitative study. BMC Pregnancy Childbirth 2019 Nov 12;19(1):415 [FREE Full text] [doi: 10.1186/s12884-019-2576-0] [Medline: 31718586]
69. Jones KH, Daniels H, Heys S, Ford DV. Toward an ethically founded framework for the use of mobile phone call detail records in health research. JMIR Mhealth Uhealth 2019 Mar 22;7(3):e11969 [FREE Full text] [doi: 10.2196/11969] [Medline: 30900996]
70. mHealth Regulation Impact Assessment: Africa. Mobile for Development mHealth. GSMA. 2015. URL: https://tinyurl.com/yz5nmn3f [accessed 2021-05-11]
71. The Data Protection Act, 2019. Office of the Data Protection Commissioner Kenya. 2019. URL: https://www.odpc.go.ke/dpa-act/ [accessed 2021-05-11]
72. Rwanda: Data protection law published in the Official Gazette. DataGuidance. 2021 Oct 18. URL: https://www.dataguidance.com/news/rwanda-data-protection-law-published-official-gazette [accessed 2022-05-16]

Abbreviations

ART: antiretroviral therapy
CDR: call detail record
CHW: community health worker
DS-I Africa: Harnessing Data Science for Health Discovery and Innovation in Africa
EAC: East African Community
IVR: interactive voice response
MATTS: Mapping Tuberculosis Transmission Study
mHealth: mobile health
NIH: National Institutes of Health
NIN: National Identification Number
PLWHA: people living with HIV/AIDS

PMTCT: prevention of mother-to-child transmission
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses

Edited by T Leung; submitted 19.11.21; peer-reviewed by D Lizotte, U Bork; comments to author 09.12.21; revised version received 17.02.22; accepted 14.04.22; published 02.06.22.

Please cite as:
Sekandi JN, Murray K, Berryman C, Davis-Olwell P, Hurst C, Kakaire R, Kiwanuka N, Whalen CC, Mwaka ES
Ethical, Legal, and Sociocultural Issues in the Use of Mobile Technologies and Call Detail Records Data for Public Health in the East African Region: Scoping Review
Interact J Med Res 2022;11(1):e35062
URL: https://www.i-jmr.org/2022/1/e35062
doi:10.2196/35062
PMID:35533323

Review

Interventions in Chinese Undergraduate Students' Mental Health: Systematic Review

Yi Shan[1], Prof Dr; Meng Ji[2], PhD; Wenxiu Xie[3], MPhil; Rongying Li[4], BM; Xiaobo Qian[5], BEng; Xiaomin Zhang[6], PhD; Tianyong Hao[5], Prof Dr

[1]School of Foreign Studies, Nantong University, Nantong, China
[2]School of Languages and Cultures, University of Sydney, Sydney, Australia
[3]Department of Computer Science, City University of Hong Kong, Hong Kong, China
[4]School of Artificial Intelligence, South China Normal University, Guangzhou, China
[5]School of Computer Science, South China Normal University, Guangzhou, China
[6]Department of Linguistics, Macquarie University, Sydney, Australia

Corresponding Author:
Yi Shan, Prof Dr
School of Foreign Studies
Nantong University
No 9 Seyuan Rd
Nantong, 226019
China
Phone: 86 15558121896
Email: victorsyhz@hotmail.com

Abstract

Background: Over 30% of university students from 8 countries were afflicted with mental distress according to a World Health Organization survey. Undergraduate students in increasing numbers in China have also been reported to suffer from different mental problems. Various psychological distresses significantly impact their academic and daily life, thereby causing role impairments and unsatisfactory academic achievements. While the prevalence of, diverse underlying factors for, and interventions of social support in college students' mental health have extensively been investigated in China, there is no study exclusively focusing on the impact of interventions on their psychological well-being.

Objective: The aim of this review was to identify and synthesize the interventions in the mental health concerns of Chinese undergraduate students studying in China reported in the literature to inform educational authorities, college and university management, students' affairs counselors, and mental health providers.

Methods: We performed a systematic review and reported the research findings of previous studies according to the protocol of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement. First, based on the predefined search strategy, keyword searches were performed in the PubMed and ProQuest databases to retrieve relevant studies. Subsequently, we screened the candidate articles based on predefined inclusion and exclusion criteria. Finally, we analyzed the included papers for qualitative synthesis.

Results: We retrieved a total of 675 studies from the PubMed and ProQuest databases using the search strategy on March 15, 2022. Among these candidate studies, 15 that were not written in English, 76 duplicates, and 149 studies of other document types were removed before screening. An additional 313 studies were excluded in the screening process, with 73 articles ruled out for being not relevant to interventions, not related to mental health, or not focused on undergraduate students in the full-text review. As a result, 49 papers were eligible and included in this systematic review. In the qualitative synthesis, we divided the interventions reported in the selected studies into two categories: (1) social support from government authorities, university authorities, students' affairs counselors and teachers, family members, health care authorities and professionals, and the media (various online platforms), and (2) various coping strategies adopted by undergraduate students themselves. We identified further research on mental health interventions that may be delivered by digital medical platforms, conversational agents (eg, chatbots), and researchers.

Conclusions: This was the first systematic review of interventions to address the mental health concerns of Chinese undergraduate students studying in China. The categorization of reported interventions and the identification of new intervention channels can

effectively inform stakeholders. Interventions for undergraduate students' mental health is a research topic worth further investigation.

(Interact J Med Res 2022;11(1):e38249) doi:10.2196/38249

KEYWORDS

systematic review; intervention; mental health; depression; anxiety; stress; Chinese undergraduate students

Introduction

Background

Over 30% of university students from 8 countries were afflicted with mental distress according to a World Health Organization survey [1]. Undergraduate students in increasing numbers in China have also been reported to suffer from different mental problems [2], including depression, compulsion, anxiety, and interpersonal sensitivity [3-6]. Students experienced different degrees of depressive symptoms in different parts of China, with an incidence of 9.7% in Eastern and Western China, 11.7% in Harbin, 11.8% in 6 universities in Wuhan, 16.8% in Anhui, and 32.82% in Western Liaoning [7-11]. Transitioning from adolescence to adulthood while leaving home to attend colleges requires facing many challenges independently [12], which causes an increase in symptoms of depression, anxiety, and stress [13-15]. Various psychological distresses significantly impact the academic and daily lives of students, thereby causing role impairments and unsatisfactory academic achievements [16-18].

Given the prevalence of mental disorders in 28.4% of Chinese college students [19], studies have been performed to identify the various underlying contributing factors such as interpersonal relationships [6]; multiple factors from individuals, families, schools, and society [12]; gender and income [20,21]; academic stress and load, financial difficulty, departure from home, unstable family, and bullying on campus [3,4,22,23]; personal behaviors and social settings [24]; and lack of physical activities [25]. Pinpointing these factors contributing to college students' mental distress facilitates developing effective mental health interventions to reduce potential adverse impacts on their psychological well-being. Mental health promotion and prevention are needed to improve the mental health condition of college students who are especially vulnerable to pressure and other mental health issues [26], which will contribute to their overall well-being [27]. While the prevalence of, diverse underlying factors for, and interventions of social support in college students' mental health have been extensively investigated in China, there is no study exclusively addressing the interventions for their mental health concerns.

Interventions in the Literature

Social support has proven to be one of the most critical and effective interventions to mitigate mental health risks imposed on college students [28-30]. Social support is a form of mutual communication and connection network, including emotional support, instrumental support, and informational support [31]. Such support has been strongly associated with mental health among college students [32]. When obtaining robust social support from friends, family, and teachers, university students had better mental health to sustain themselves against crises and stress [32]. Social support was proven to moderate the relationship between stress and depression [33]. High levels of social support could buffer mental health concerns [34].

Various therapies are also effective mental health interventions. Several therapies have showed effects on par with those of pharmacological treatment [35], including psychotherapy [36], interpersonal therapy [37], problem-solving therapy [38], supportive therapy [39], psychoeducation [40], and exercise/physical activity [41].

Universities have traditionally been providing mental health services in clinical settings, such as face-to-face individual or group-based consultations [42]. However, the available resources of many universities are too limited to support comprehensive approaches to students' mental health, and students are frequently unwilling to visit traditionally structured student counseling centers for help [42,43]. Therefore, it is necessary to identify effective mental health interventions that can be delivered to students in virtual settings and cover the spectrum of interventions from prevention to treatment [44].

Mowbray et al [44] developed internet-based interventions that were designed to promote mental health help-seeking, including a mental health literacy/destigmatization intervention, a feedback intervention, and a help-seeking list intervention. Web-based depression and anxiety interventions have been confirmed to be effective for treating common mental disorders [45,46].

Objective

The aim of this review was to identify and synthesize the interventions addressing the mental health concerns of Chinese undergraduate students studying in China reported in the literature. The synthesized interventions are expected to inform stakeholders, including educational authorities, college and university management, students' affairs counselors, and mental health providers.

Methods

Study Design

To analyze and synthesize the interventions focused on the mental health of Chinese undergraduate students studying in China, we performed a systematic review and reported the research findings according to the protocol of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement [47]. We applied keyword searches to retrieve publications related to this research topic in two databases (PubMed and ProQuest), and screened the candidate articles based on the predefined inclusion and exclusion criteria. Finally, we analyzed the included papers for qualitative synthesis.

Search Strategy

We searched the PubMed and ProQuest databases to identify relevant studies. Based on the studies mentioned in the Introduction, we defined undergraduates' mental health concerns as issues related to depression, anxiety, stress, and disorder. Drawing on these keywords, we designed the following search strategy for this review: ((college student [Title/Abstract]) OR (university student [Title/Abstract]) OR (student [Title/Abstract])) AND ((digit* [Title/Abstract]) OR (online [Title/Abstract])) AND ((mental health [Title/Abstract]) OR (mental disorder [Title/Abstract]) OR (depression* [Title/Abstract]) OR (anxiety [Title/Abstract]) OR (disorder [Title/Abstract]) OR (stress [Title/Abstract])) AND (intervention) AND ((Chinese) OR (non-English) OR (Chinese-speaking)). Considering that many relevant articles were not published in peer-reviewed journals, we considered both peer-reviewed and nonpeer-reviewed articles in this systematic review. We did not impose any restrictions on the publication date of the candidate articles to retrieve all papers related to this topic. On March 15, 2022, we retrieved the candidate studies from the PubMed and ProQuest databases using the above search strategy. The keywords of the strategy were searched in the titles and abstracts of the candidate articles.

Study Selection Criteria

Papers that were not written in English were excluded because translation of the articles was not feasible or reliable. Additionally, we only included journal articles and excluded other article types (eg, review papers, letters, reports, and editorials).

This systematic review focused on the interventions in the mental health of Chinese undergraduate students in China. Therefore, the following criteria had to be satisfied in the selection of eligible papers: (1) the target population is undergraduate students; (2) the undergraduate students are Chinese who study in colleges and universities in China; and (3) the candidate articles need to be related to interventions for undergraduate students' mental health or mental disorder problems, including depression, anxiety, and stress. Articles that were focused on undergraduate students' mental health or mental disorder problems but involved no interventions, and those focused on interventions addressing the mental health or mental disorder problems of Chinese undergraduate students studying in countries other than China were excluded from this review.

Screening and Article Selection

We used Microsoft Excel to collect and manage the data of the candidate papers, including author, year of publication, country, target population (participants), study design/method, interventions, and limitations. The screening of the eligible studies was performed in the following two steps. First, two reviewers (YS and YC) reviewed the titles and abstracts of the candidate articles, excluding those that were not related to interventions addressing the mental health concerns of Chinese undergraduate students studying in China. Papers whose eligibility was unclear were retained for the full-text review. Second, six reviewers (YS, YC, XQ, RL, XW, and TL) reviewed the full texts of the remaining articles independently. Any controversies were resolved through discussing and consulting with two additional authors (MJ and WX) to make final decisions at the panel meeting of all members of the research team.

Data Extraction

We extracted data in light of our research objective to collect key information from eligible articles. The data extraction was performed by six researchers (YS, YC, XQ, RL, XW, and TL) independently, and reviewed and cross-checked by two researchers (MJ and WX). Any discrepancies were addressed through a consensus discussion at the panel meeting of all members of the research team.

Data Analysis and Synthesis

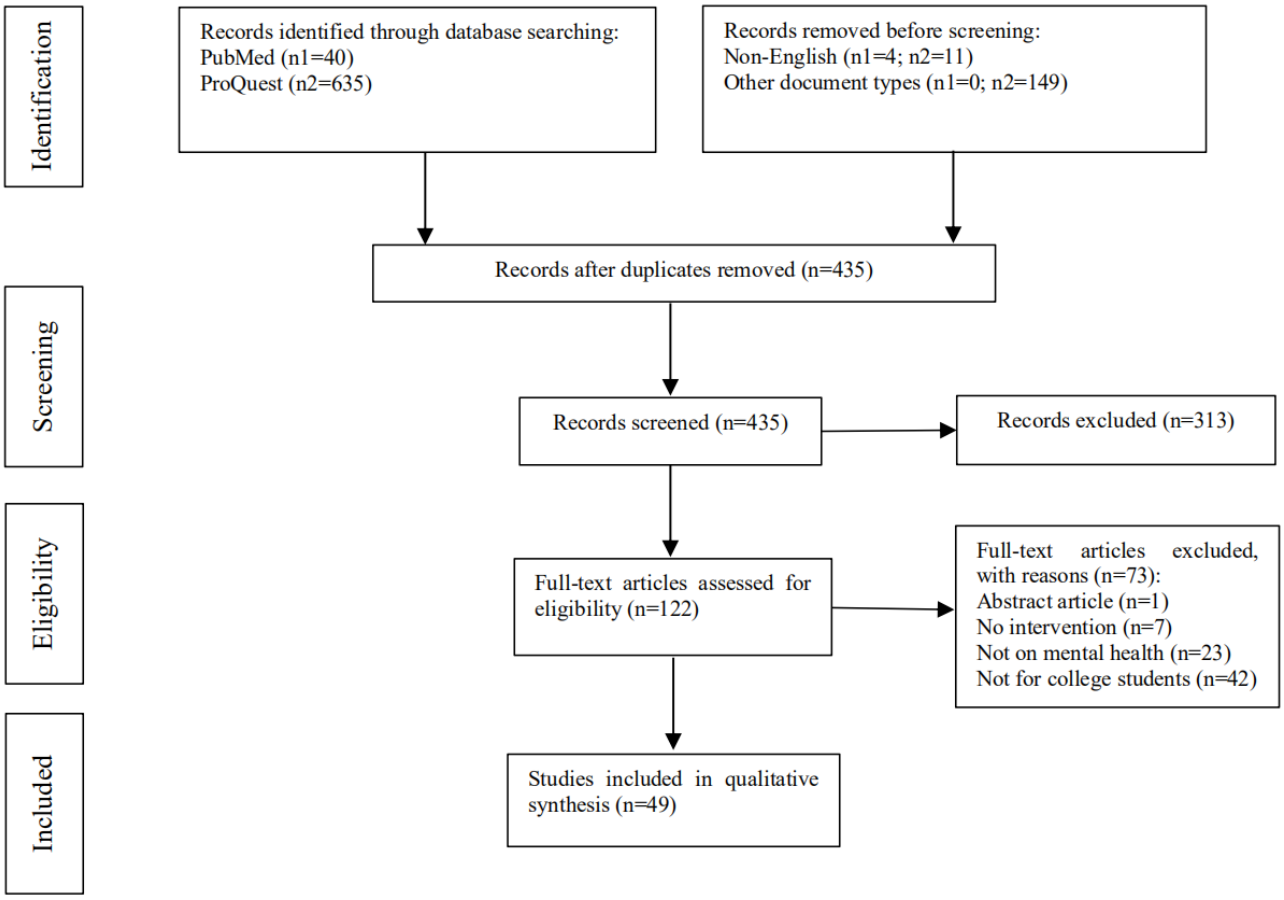
It was not feasible to conduct a meta-analysis owing to the expected variety of study designs, mental health interventions, and study limitations. As such, a descriptive analysis was carried out to summarize the data extracted from the included papers.

Results

Search Results

We retrieved a total of 675 studies from the PubMed and ProQuest databases using the search strategy. Among these candidate studies, 15 that were not written in English, 76 duplicates, and 149 studies of other document types were removed before screening. An additional 313 studies were excluded in the screening process, with 73 articles ruled out for being not relevant to interventions, not related to mental health, or not including undergraduate students in the full-text review. As a result, 49 papers were deemed to be eligible and included in this systematic review. The PRISMA flowchart of the screening and reviewing processes is provided in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the selection of eligible studies.



Characteristics of Included Studies

The following items of the extracted data were obtained: publication information (name of authors, country, and publication year), target population (participants), study design (research methods), interventions, and limitations. Table S1 in Multimedia Appendix 1 describes these characteristics of the 49 included studies.

Interventions in the mental health of Chinese undergraduate students studying in China have only recently attracted scholarly attention, as evidenced by the fact that all 49 included studies [2,12,26,48-93] were published between 2020 and 2022, including 11 (22%) published in 2020, 33 (67%) published in 2021, and 2 (4%) published in 2022. This indicates the worsening mental health condition of this particular population in recent years, especially in the context of the repeated resurgences of the COVID-19 pandemic, which increased the stress of college students and exposed them to new frustrating stressors that caused various mental health concerns [2]. Given college students' high stress and anxiety levels during the COVID-19 pandemic [88], 35 (71%) of the 49 studies [2,12,48,49,52-58,61-64,67-69,72-77,81-89,91,92] investigated the COVID-19–induced mental health concerns of Chinese undergraduate students and proposed specific interventions in response to the severity of the psychological impact of COVID-19 [87].

The included studies were primarily conducted as cross-sectional surveys via online questionnaires. The main limitation of these studies is that self-reported online questionnaires are likely to result in a certain degree of recall bias and response bias due to the stigma attached to mental health conditions.

Undergraduates' Mental Health Concerns

In the 49 included studies, we identified the following forms of mental health concerns: depression, anxiety, stress, interpersonal sensitivity, fear, distress, psychological disorders, self-harm/suicidal, insomnia, obsessive interpersonal sensitivity, trauma, negative emotions, and insecurity.

Categorization of Mental Health Interventions

Overview

The interventions proposed in the included studies can be divided into two broad categories: social support and coping strategies. Social support was provided by government authorities [49,52-54,61,62,68,72,84,86-88]; university authorities [12,48,50,51,53-55,57-59,62,64,65,67,69,91,93]; students' affairs counselors and teachers [54,75,83-85,87,93]; family members [52,54,57,83,94]; health care authorities and professionals [52,63,66,75,77-79,87]; researchers [70,85]; and media-, internet-, and smartphone-based interventions [26,53,55,56,60,71,73,75,77,80,81]. Positive coping strategies were adopted by undergraduate students themselves [52,82,83,85,86,89,92,93]. In addition, negative coping strategies were also reported in a few studies [52,59,74,75,89]. All of the interventions are listed in Table 1.

Table 1. Interventions reported in the 49 included studies.

Reference	Interventions, conclusions, and recommendations
Huang et al [2]	Active coping strategies helped improve their psychological well-being; family support was particularly important for maintaining mental health and ameliorating mental health challenges in this major health crisis; suitable psychointervention, routine screening for risk behaviors, and provision of further social support are needed for undergraduate students in the COVID-19 pandemic or other emergency public health events
Lei et al [12]	Universities should develop a care culture and environment that supports the life adjustment of college students, promotes cultural and sports activities, and facilitates the expansion of social networks; mental health education and psychological counseling services should be strengthened, including hotlines offering timely help to address students' urgent needs; early detection and effective management of mental health problems can effectively reduce serious mental health disorders
Mak et al [26]	Internet-based cognitive behavioral and mindfulness training programs are effective, which can be easily incorporated into existing service provision portfolios that promote mental health and reduce psychological distress to ultimately promote mental health among college students and young working adults
Yu et al [48]	Screening for ACEs[a], and strength-based, trauma-informed interventions on fostering resilience are needed to promote mental well-being among Chinese young adults
Zhang et al [49]	In practical interventions, authorities (eg, governments and universities) should first focus on improving efficacy appraisal by providing psychological support to gain the trust of college students so that they believe in and comply with scientific prevention and control measures. By inviting psychiatrists to deliver lectures, authorities can reasonably and effectively enhance the public information of COVID-19–related knowledge and scientific prevention and control measures.
Shen et al [50]	It is important to address ADHD[b] symptoms among students with anxiety; it is of importance to screen medical students for anxiety disorders to better promote the mental health and well-being of this population and better prevent suicidal behaviors
Sze et al [51]	Given its associations with negative emotions and other aspects of health, screening and management of EE[c] may improve multiple areas of health and well-being
Zhao and Zhou [52]	It is critical for policymakers, public health agencies, parents, psychologists, and health care staff to remain sensitive to the potential negative consequences of ubiquitous social media exposure; the general public, especially those who have been directly or indirectly traumatized by COVID-19, could be advised to avoid excessive social media use and learn effective emotion regulation strategies (eg, reappraisal) to reduce negative emotions induced by news coverage
Yu et al [53]	The government can open free psychological hotline consultations to help college students solve their psychological problems; the media should release correct information in a timely manner and prevent the spread of rumors; universities can actively organize health education activities and encourage college students to arrange their time reasonably and take the initiative to find a suitable way to relieve stress during home quarantine
Li and Peng [54]	Adopting positive coping strategies may enhance social support that in turn relieves anxiety. The effect of social support, especially family and counselor support, can decrease anxiety in coping with the COVID-19 pandemic cognitively and behaviorally. Policymakers and school administrators should encourage meaningful communication between family members and activate effective counseling services to maintain positive mental health
Wang et al [55]	Reducing SNS[d] addiction and mental problems by conducting interventions using cognitive behavioral approaches; screening for and addressing excessive SNS use are needed to prevent SNS addiction and mental distress among young people.
Sit et al [56]	Evidence-based digital mental health interventions
Liang et al [57]	University campuses should develop and implement effective screening procedures to closely monitor students' exposure to stressors and mental health status; psychological intervention programs should be designed to address fear and fully utilize psychological assistance hotlines to help college students better adjust themselves; performing psychological help-seeking intervention, strengthening the dissemination of mental health knowledge, and improving the level of mental health perception are effective ways to improve help-seeking attitudes and increase the probability that college students will seek psychological help
Nurunnabiet al [58]	University authorities should be aware of students' coping strategies. In particular, students who live without parents or relatives should be taken care of properly during the outbreak. To help students cope with the mental pressure, university authorities may consider arranging or organizing programs such as an online experience-sharing competition, and encourage students by offering rewards or financial aids. Required food and health care materials should be supplied to ensure the students' safety
Wu et al [59]	Sleep hygiene, mobile phone and internet use hygiene, mental health education courses, professional psychological counseling, and other interventions should be considered and implemented. Appropriate interventions that target problematic smartphone use could potentially reduce anxiety and depression levels, which will in turn provide a buffer against the negative impact of poor sleep quality on eating disorder symptoms
Chen et al [60]	Web-based intervention for subclinical depression (MoodBox) informed by evidence-based psychological interventions, including CBT[e], IPT[f], and mindfulness meditation
Yu et al [61]	Various cognitive, behavioral, and psychosocial responses to COVID-19 showed both direct and indirect effects (via mental distress due to COVID-19) on depression. Thus, interventions to improve such multidimensional factors might reduce mental distress during the initial COVID-19 outbreak period

Reference	Interventions, conclusions, and recommendations
Zhang et al [62]	Relevant education and psychological counseling to parents during the outbreak to help them understand their children's mental state, with universities providing psychological counseling and psychological interventions to students, focusing on college students who are most severely affected by the epidemic
Li et al [63]	Recommends providing long-term psychological services for students; the results could help health care professionals identify college students at high risk of mental health problems so that appropriate interventions can be targeted against them
Tang et al [64]	Recommends providing psychological interventions for quarantined college students to help them reduce fear and improve sleep duration. Universities need to consider planning acute and long-term psychological services for more vulnerable students, graduates, and students living in the most severely affected areas
Zhou et al [65]	One week of positive mental imagery training can help to improve negative emotions and anxiety in depression; further exploration of this training program is suggested
Shen et al [66]	Providing mental health care and counseling services to students of high-risk groups in medical schools; the early diagnosis and treatment of ADHD may have a suicide prevention effect
Li et al [67]	Parents strengthen communication with their children and provide psychological support to their children. Universities carry out relevant online mental health courses and implement psychological intervention measures to improve students' psychological adaptability
Yang et al [68]	The government, school administrators, and society strengthen operability research to provide coping strategies, implement psychological interventions, and conduct relevant training
Liu et al [69]	Universities should adopt a web-based PPI[g] to improve the mental health of college students
Carciofo [70]	Longitudinal studies of these variables may establish causal relationships and may inform interventions to treat psychological distress and disorders
Yen et al [71]	Depressed college students have less hostility after entering the internet, suggesting that the internet as a useful medium to provide treatment for people with depression
Zheng et al [72]	Recommends adequate social support and long-term targeted psychological interventions for college students. More serious mental health problems seen among fourth-year students, proposing to specifically increase their employment opportunities and develop mental health rehabilitation programs
Song et al [73]	Online or smartphone-based psychoeducation and psychological interventions that will also reduce the risk of virus transmission by foregoing face-to-face therapy
Tao et al [74]	Unsupervised, self-initiated interventions against mental and sleep disorders of students can lead to more disastrous outcomes
Jia et al [75]	Public health education from health authorities in various governments is needed for dissemination of the importance of preventive measures during COVID-19. Psychological health services should be implemented to alleviate the adverse effects of this pandemic under national social distancing. Psychological interventions could also be carried out through online platforms under national social distancing during COVID-19. Teachers should also pay attention to strengthening the dissemination of COVID-19 knowledge and preventive measures to reduce the level of anxiety and depression in the student population
Dun et al [76]	Interventions to decrease sedentary time and improve mental health may be warranted to mitigate weight gain during the lockdown period and reverse the weight gain in youth after the COVID-19 pandemic
Yu et al [77]	Grief counseling and online sacramental ceremonies should be implemented for this group to prevent negative emotional difficulties; mindfulness meditation and CBT can reduce students' anxiety and depression
Pan and Zhuang [78]	Integration of cognitive behavioral intervention and adventure training in a class setting might be an effective and feasible approach for the mental health counseling of university students
Auyeung and Mo [79]	PPI
Zhao et al [80]	Mediating effect of online social support was stronger among college students with lower perceived social support than those with higher perceived social support
Xin et al [81]	Online brief interventions need to be made available, including screening of mental distress, counseling hotlines, emotional regulation and coping skills, and promotion of positive psychology
Liang et al [82]	Compared to meeting no guidelines, meeting the sleep guideline (alone or in combination with other guidelines) was associated with significantly lower levels of depression and anxiety; meeting both SB[h] and MVPA[i] guidelines was also associated with a significantly lower level of depression. Hence, meeting more guidelines, especially adhering to a healthy sleep routine, may play an important role in promoting the mental health of young adults
Sun et al [83]	Perceived available peer support negatively contributed to depressive symptoms. Both negative and positive indicators of emotional well-being mediated the association between perceived available peer support and depressive symptoms, and advanced the practical needs for preventive efforts and accessible care to support the psychological and emotional needs of young people during the COVID-19 pandemic.

Reference	Interventions, conclusions, and recommendations
Li et al [84]	Over 50% of the participants had obvious fear and anxiety symptoms at 61.64% and 58.39%, respectively. Conformity (49.49%), invulnerability (26.11%), insensitivity (21.49%), and rebelliousness (12.41%) symptoms also appeared. Senior students experienced more anxiety than freshmen. Psychological symptoms (except for insensitivity) had no significant difference with respect to gender, residence, and annual household income in one-way analysis of variance
Zhu et al [85]	Association between mental health and emotion regulation, which will help direct a psychological intervention that relieves these issues during the pandemic
Zhuo et al [86]	Back-to-school students who are certain and uncertain that COVID-19 will rebound again were significantly more anxious and depressed than those with optimistic attitudes. Government departments should pay high attention to the mental health problems evoked by intolerance of uncertainty (IU). Social support as a moderator could buffer the relationship between IU and mental health, including anxiety and depression during unprecedentedly uncertain times
Li et al [87]	Mental health services reducing PTSD[j] should be provided; students who have lost loved ones and suffered family financial loss should be given particular care
Zhan et al [88]	Education departments should attach great importance to the mental health of college students, and it is necessary to provide precise psychological interventions for groups experiencing greater pressure levels and marked anxiety and depression
Ding et al [89]	Three coping styles were all significantly correlated with psychological distress in Chinese college students during the early stage of the COVID-19 pandemic. Adaptive emotion-focused coping was negatively associated with perceived stress and psychological distress. Emotion-focused coping was positively associated with perceived stress and distress. Individuals who use specific reactive emotion-focused coping strategies more often, such as focusing on emotions, denial, seeking emotional social support, and disengaging, experience more stress.
Zhou et al [90]	Intercultural cooperation should be promoted to develop a cross-culturally valid concept of stigma against psychological help that could be used as the basis for intercultural comparison and developing interventions to reduce stigma
Liang et al [91]	Guiding postgraduate students to correctly understand their mental health status and individual differences in mental tolerance, and encouraging postgraduate students to seek help if they experience psychological problems so as to help them adjust their goals and plans according to reality, and avoid the development of other problems such as PTSD; establishing an early warning system for the mental health of postgraduate students during the pandemic and improve online and offline psychological counseling service systems; considering the characteristics and situation of different postgraduate groups for postgraduate student management to develop targeted mental health education programs and adopting objective measures, so as to improve postgraduate mental health and nurture both physical and mental health to facilitate China's modernization; developing and maintaining conditions to improve the communication between postgraduate students and advisors during the pandemic and create a new postgraduate guidance mode to relieve the psychological problems of postgraduate students
Wen et al [92]	Enhancing positive self-beliefs such as hope and self-efficacy helps to buffer the effects of insecurity on stress. Physical and psychological exercises that enhance hope can be effective interventions to help university students buffer the impacts of insecurity and alleviate stress during the outbreak. Improving positive self-beliefs can help to relieve the pressure of university students during the outbreak. University students can improve their self-efficacy by participating in movement-based courses, including Pilates and Tai Chi, so as to improve their positive mood and relieve stress. They can also effectively improve their hope levels by setting personal goals and conducting goal-pursuit exercises, which can also contribute to the reduction of insecurity and stress
Lin et al [93]	Dental schools and educators promote stress-coping strategies and modify teaching curricula to reduce students' stress. Stress management efforts such as time management, encouragement from advisors, and regular exercise are recommended

[a]ACE: adverse childhood experience.

[b]ADHD: attention deficit and hyperactivity disorder.

[c]EE: emotional eating.

[d]SNS: social networking site.

[e]CBT: cognitive behavioral therapy.

[f]IPT: interpersonal psychotherapy.

[g]PPI: positive psychology intervention.

[h]SB: sedentary behavior.

[i]MVPA: moderate-to-vigorous physical activity.

[j]PTSD: posttraumatic stress disorder.

Social Support

Main Categories

In the 49 included studies, we identified various types of social support from government authorities; university authorities; students' affairs counselors and teachers; family members; health care authorities and professionals; researchers; and the media-, internet-, and smartphone-based interventions.

Government Authorities

Governments need to join hands with school administrators and various social parties to strengthen feasibility research to offer coping strategies, perform psychological interventions, and conduct relevant training [68]. Specifically, government

authorities at all levels need to (1) specially improve efficacy appraisal through providing psychological backup for undergraduate students, by inviting psychiatrists to deliver relevant lectures [49]; (2) offer free psychological counseling via hotlines to help undergraduate students solve their psychological problems [53]; (3) develop interventions to improve undergraduate students' various cognitive, behavioral, and psychosocial responses to public health emergencies such as COVID-19 [61]; (4) provide parents with relevant education and psychological counseling to help them understand their children's mental state [62]; (5) advocate adequate social support and long-term targeted psychological intervention to provide more employment opportunities and develop mental health rehabilitation programs for the fourth-year undergraduate students who suffered more mental problems [72,84]; (6) pay special attention to mental health concerns induced by intolerance of uncertainty [86]; and (7) particularly care for students who have lost loved ones and experienced family financial losses [87].

Policymakers and public health agencies need to (1) be sensitive to the potential adverse effects of omnipresent exposure to social media [52] and (2) encourage effective communication among family members and activate effective psychological counseling services [54].

Education departments ought to provide precise psychological interventions for those suffering greater pressure and marked anxiety and depression among undergraduate students [88].

University Authorities

Previous studies proposed university authorities to make the following interventions in undergraduate students' mental health concerns, including (1) developing a caring culture and ambience, which backs up undergraduate students to make life adjustments, promoting cultural and sports activities, facilitating the expansion of social networks, strengthening mental health education and counseling (eg, hotlines providing timely help for those in urgent need), early detection and effectively managing mental health concerns [12]; (2) screening for adverse childhood experiences and providing strength-based, trauma-informed interventions on fostering resilience [48]; (3) screening for anxiety disorders [50,57]; (4) screening and managing of emotional eating [51]; (5) organizing health education activities, and encouraging undergraduate students to arrange their time reasonably and find a proper approach to alleviate stress [53]; (6) encouraging timely, effective communication between family members [54]; (7) conducting interventions using cognitive behavioral approaches and screening for and addressing excessive social networking service use [55]; (8) developing psychological help-seeking interventions, strengthening the dissemination of mental health knowledge, and improving the level of mental health perception to improve help-seeking attitudes and increase the probability of seeking psychological help [57]; (9) being aware of undergraduate students' coping strategies, organizing programs such as an online experience-sharing competition, encouraging students by offering rewards or financial aids, and supplying necessary food and health care materials [58]; (10) implementing various interventions such as sleep hygiene, mobile phone and internet use hygiene, mental health education courses, and professional psychological counseling [59,62,64,67,91]; (11) providing positive mental imagery training [65]; (12) offering relevant online mental health courses [67]; (13) adopting a web-based positive psychology intervention [69]; (14) establishing an early warning system for mental health concerns, guiding students to correctly understand their mental health status and individual differences in mental tolerance while encouraging them to seek help if necessary, developing targeted mental health education programs, and creating a new guidance mode to improve the communication between students and counselors [91]; and (15) promoting stress-coping strategies [93].

Students' Affairs Counselors and Teachers

Student counselors need to assume the responsibilities to provide support and positive coping strategies [54,93] along with psychological health services [75,85], in particular accessible care [83] for senior students who experienced more anxiety [84] and students who have lost loved ones and suffered family financial loss [87], emotion regulation guidance [85], and encouragement [93]. Teachers should pay close attention to disseminating COVID-19 knowledge and preventive measures [75], and modify teaching curricula [93] to reduce undergraduate students' academic pressure in the face of public health emergencies and natural disasters.

Family Members

One study argued that the role of family support in maintaining undergraduate students' mental health must be emphasized [94]. We found 4 studies reporting interventions delivered by family members. Family support can effectively decrease anxiety cognitively and behaviorally [54]. Therefore, parents need to strengthen communication with their children and provide psychological support for them [57], provide accessible care to support the psychological and emotional needs of their children [83], and maintain sensitivity to the potential negative consequences on their children brought about by ubiquitous social media coverage [52].

Health Care Authorities and Professionals

Health care professionals should identify, diagnose, and treat undergraduate students at high risk of mental health concerns early so that proper interventions, including long-term psychological services, can be tailored for them [63,66]. Grief counseling, mindfulness meditation, and cognitive behavioral therapy are recommended to reduce students' anxiety and depression [77], along with positive psychological intervention [79], and authorities should remain sensitive to the potential negative consequences of ubiquitous social media exposure [52].

Health authorities should provide public health education, implement psychological health services online and offline [75], integrate cognitive behavioral intervention and adventure training in a class setting [78], and offer mental health services reducing posttraumatic stress disorder [87].

Researchers

Researchers need to conduct longitudinal studies of morning affect, eveningness, and amplitude distinctness to establish

causal relationships between these factors and negative emotionality, and thus inform interventions to treat psychological distress and disorders [70]. They should also study the association between mental health and emotion regulation to help direct psychological interventions [85].

Media-, Internet-, and Smartphone-Based Interventions

Evidence-based digital mental health interventions were found to be useful in improving mental health concerns [56,60], because depressed undergraduate students became less hostile when logging onto the internet, suggesting the internet as a useful medium to provide treatment for people with depression [71]. In the context of face-to-face intervention delivery hindered by the COVID-19 pandemic, internet-delivered interventions such as internet-delivered cognitive behavioral therapy can be considered to address effects of social networking service addiction on the mental health status of Chinese university students [55]. It is effective to use internet-based cognitive behavioral and mindfulness training programs, which can easily be integrated into existing service provision portfolios that promote mental health and reduce psychological distress to promote the mental health of undergraduate students [26]. Online brief interventions need to be made available, including screening of mental distress, counseling hotlines, emotional regulation and coping skills, and promotion of positive psychology [60,73,75,81], whose mediating effects were proven to be stronger among undergraduate students with lower perceived social support [80]. Online sacramental ceremonies should be implemented to prevent negative emotional difficulties [77]. The media need to release correct information in a timely manner and curb the spread of rumors [53], which may aggravate the mental health concerns of psychologically vulnerable undergraduate students.

Coping Strategies

A recent study mentioned positive coping strategies as intervention measures to help improve college students' psychological well-being [2]. Some of the 49 included studies reported positive coping strategies, including regulating emotions effectively [52,85]; meeting the sleep, sedentary behavior, and moderate-to-vigorous physical activity guidelines [82]; developing optimistic attitudes [86]; adopting problem-focused and adaptive emotion-focused coping [89]; enhancing positive self-beliefs (eg, hope and self-efficacy) and physical and psychological exercises [92]; and managing stress through time management and regular exercise [93]. Moreover, undergraduate students need to provide mutual peer support and accessible care, because perceived peer support and care alleviated depressive symptoms and met young people's psychological and emotional needs [83]. All of these positive coping strategies were proven to be effective in promoting undergraduate students' mental health [82].

However, negative coping should be avoided, which can lead to more disastrous consequences [75]. The reported negative coping strategies include excessive use of social media, which can be counteracted with emotion regulation (eg, reappraisal) to reduce negative emotions induced by news coverage [52]; problematic smartphone use [59]; unsupervised, self-initiated intervention [74]; and maladaptive emotion-focused coping [89].

Discussion

Principal Findings

The great uncertainty about the pandemic, the abrupt transition to and participation in online classes, and the COVID-19–related impacts on life all frequently contributed to the increased stress and anxiety of undergraduate students [95-97], in addition to their inability to tackle problems concerning interpersonal relationships, academic challenges, and career development due to the lack of life and social experience [98,99]. Therefore, 35 (71%) of the 49 included studies that investigated Chinese undergraduates' mental health concerns induced more or less by the COVID-19 pandemic. The proposed interventions to counteract the severe psychological influence of the pandemic were synthesized into two broad categories: social support and coping strategies. Social support has proven to be an effective protective factor for mental health in previous studies [100-103]. Social support means providing practical help, emotional backup, and information assistance by those around individuals in mental distress [104]. Social support was found to be negatively correlated with adverse mental health outcomes (anxiety, depression, and insomnia), which were aggravated by COVID-19–induced intolerance of uncertainty [86]. Social support also served as a moderator buffering the relationship between intolerance of uncertainty and mental health, including anxiety and depression [86].

Coping strategies have proven to be effective protective factors for mental health in previous studies [100-103]. All of these interventions are crucially important in the context of the COVID-19 pandemic and other future public health crises or natural disasters, which can cause long-term mental disorders in various populations [99]. Therefore, the effective mental health interventions in the forms of various social support and coping strategies reported in the 49 included studies can surely shed light on the interventions in undergraduate students' mental health issues.

Moreover, we identified some mental health interventions that were not reported in the included studies but are potentially effective and robust, including those delivered by digital medical platforms, conversational agents (eg, chatbots), and researchers.

Digital Medical Platforms

Two of the most popular digital medical platforms, *haodaifu* ("The Good Doctor") and *zuoshouyisheng* ("The Left-handed Doctor"), in China were not recommended in the selected studies. These two platforms, among others, should proactively be advocated as effective mental health interventions for undergraduate students, especially given the widely acknowledged stigma attached to mental health concerns [50]. In saving undergraduate students' face and protecting their privacy, digital medical platforms should be popularized among undergraduate students and the general public. Internet-based interventions have been confirmed to be effective for treating common mental disorders [44-46].

Conversational Agents

The chatbot, as the most popular type of conversational agent, simulates human conversations to provide medical and health care interventions. These human-like, empathetic chatbots are capable of monitoring people's health [105]. Chatbots and conversational agents display many advantages unmatched by other health consultation alternatives, such as easing the overburdened contact centers and decreasing health risks caused through personal contact [106], providing the only possible solution to catering to the unprecedented demand for health-related information given the lack of professional human agents [107,108], providing timely services at any time [109], ensuring consistent quality services [110], and avoiding moral judgment of user information [111]. Chatbots need to be recommended as an effective mental health intervention tool to undergraduate students, who can try an app on *zuoshouyisheng* that is equipped with the chatbot function.

Researchers

Only two of the included studies mentioned the role of researchers [70,85]. In fact, researchers can effectively intervene in undergraduate students' mental health by synthesizing interventions proposed in previous studies and by reporting novel intervention strategies. These synthesized and novel interventions can inform various stakeholders, including educational authorities, college and university management, students' affairs counselors, and mental health providers. Although playing a role of indirect intervention, their part should never be overlooked.

Comparison With Previous Work

In this review, we synthesized the interventions for addressing Chinese undergraduates' mental health concerns into two categories of social support and coping strategies, confirming the findings of these two effective protective factors for mental health highlighted in previous studies [100-103]. Social support from family members, friends, colleagues, relatives, and neighbors [104], and from educational authorities, college and university management, students' affairs counselors, and mental health providers proposed in many of the 49 included studies can deliver critical and effective interventions, mitigating mental health risks imposed on college students [28-30], sustaining them against crises and stress [32], moderating the relationship between stress and depression [33], and buffering mental health concerns [34]. Previous studies showed that younger adults and people with greater social strain but less social support suffer worse mental health, and that perceived social support impacts the overall depression outcome and the recovery from affective disorders [104,112-114]. The importance of various forms of social support can never be overemphasized in the interventions in Chinese undergraduates' mental health concerns.

Eight of the 49 included studies reported positive coping strategies [52,82,83,85,86,89,92,93], which effectively improved undergraduates' mental health conditions. Four of the 49 publications proposed negative coping strategies [52,59,74,75,89], which may cause disastrous consequences. Coping strategies refer to the thoughts and behaviors used by individuals to manage the internal and external demands of stressful events [115]. Although various forms of social support turned out to play an essential role in mitigating Chinese undergraduates' mental health concerns, these stakeholders' own coping strategies are more essential. When people face challenging or intricate negative events, the coping style they adopt, be it positive or negative, is crucially important, which will influence their psychosocial outcomes and especially their mental health [115]. People adopting positive coping strategies were afflicted with less emotional distress, whereas those adopting negative coping strategies suffered more emotional distress [116].

Eleven of the 49 included studies described evidence-based internet-delivered digital mental health interventions such as internet-based cognitive behavioral and mindfulness training programs, online brief interventions, online sacramental ceremonies, internet-delivered cognitive behavioral therapy, and others [26,53,55,56,60,71,73,75,77,80,81]. These reported internet- or technology-based interventions can easily be incorporated into existing mental health service provision portfolios. The internet is likely to be an ideal channel to provide promising interventions for students in tertiary education [43,117]. Web-based depression and anxiety interventions were proven effective in treating common mental disorders [45,46]. An internet-based intervention complex, comprising a mental health literacy/destigmatization intervention, a feedback intervention, and a help-seeking list intervention, was proposed by Mowbray et al [44] to promote mental health help-seeking, improve mental health–related knowledge, and decrease the stigma attached to mental health concerns. Online interventions have the potential to fulfill a function in decreasing the depression and anxiety prevalence in the target populations [44]. According to a very recent Canadian study, organizations have been expanding the use of virtual care and digital mental health interventions such as web-based programs, apps, and websites [118]. This is a prevalent trend worldwide. The rapid shift from the use of traditional interventions to the use of digital mental health services and interventions [119] can inform stakeholders of the necessity to adopt digital interventions to support undergraduates' mental health.

The quality and accessibility of mental health can be improved through mobile apps [120]. Chatbots are one of the main mobile apps for mental health [121]. Chatbots were used to deliver psychological service for medical professionals and the public in China [122,123]. However, none of the 49 included studies proposed chatbot-delivered interventions in undergraduates' mental health issues. Chatbots have been pervasively used in the field of mental health [121], contributing to addressing the shortage of mental health care [124]. They promise to be an ideal tool that delivers interventions to people suffering from mental health concerns, especially those reluctant to seek mental health advice due to stigmatization. Stakeholders concerned about Chinese undergraduates' mental health should adopt and popularize this new technology.

Limitations

This review had several limitations. First, we merely searched two databases (PubMed and ProQuest) for eligible publications. Therefore, there are possibly articles left uncovered in this

review. In further research, we will consider more databases, including Embase, CINAHL, PsycInfo, ACM Digital Library, and others. Second, some principal findings may have a low generalizability, considering that most interventions were only reported in only one or two selected articles. Third, we did not compare the findings of this review with other systematic reviews, as this review was the first of its kind.

Conclusions

Considering that colleges and universities in China have reported unprecedented numbers of students in psychological distress in recent years [2], we performed the first systematic review of interventions addressing the mental health concerns of Chinese undergraduate students studying in China. We found that it is necessary to review this topic systematically, considering the deteriorating mental health of Chinese undergraduate students, especially in the context of COVID-19 resurgences. We divided all the interventions reported in the selected studies into two categories: (1) social support from government authorities, university authorities, students' affairs counselors and teachers, family members, health care authorities and professionals, and the media (various online platforms), and (2) various coping strategies adopted by undergraduate students themselves. We identified further research on mental health interventions through digital medical platforms, conversational agents (eg, chatbots), and researchers. These interventions combined can provide important implications for practical interventions in the mental health concerns of college students. The intervention in undergraduate students' mental health concerns is a research topic worth further investigation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of the eligible studies.
[DOCX File , 82 KB - ijmr_v11i1e38249_app1.docx]

References

1. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, WHO WMH-ICS Collaborators. WHO World Mental Health Surveys International College Student Project: prevalence and distribution of mental disorders. J Abnorm Psychol 2018 Oct;127(7):623-638 [FREE Full text] [doi: 10.1037/abn0000362] [Medline: 30211576]
2. Huang Y, Su X, Si M, Xiao W, Wang H, Wang W, et al. The impacts of coping style and perceived social support on the mental health of undergraduate students during the early phases of the COVID-19 pandemic in China: a multicenter survey. BMC Psychiatry 2021 Oct 27;21(1):530 [FREE Full text] [doi: 10.1186/s12888-021-03546-y] [Medline: 34706690]
3. Said D, Kypri K, Bowman J. Risk factors for mental disorder among university students in Australia: findings from a web-based cross-sectional survey. Soc Psychiatry Psychiatr Epidemiol 2013 Jun;48(6):935-944. [doi: 10.1007/s00127-012-0574-x] [Medline: 22945366]
4. Huckins JF, DaSilva AW, Hedlund EL, Murphy EI, Rogers C, Wang W, et al. Causal factors of anxiety and depression in college students: longitudinal ecological momentary assessment and causal analysis using Peter and Clark momentary conditional independence. JMIR Ment Health 2020 Jun 10;7(6):e16684 [FREE Full text] [doi: 10.2196/16684] [Medline: 32519971]
5. Cheng S, Jia C, Wang Y. Only children were associated with anxiety and depressive symptoms among college students in China. Int J Environ Res Public Health 2020 Jun 05;17(11):4035 [FREE Full text] [doi: 10.3390/ijerph17114035] [Medline: 32517044]
6. Li J, Li J, Jia R, Wang Y, Qian S, Xu Y. Mental health problems and associated school interpersonal relationships among adolescents in China: a cross-sectional study. Child Adolesc Psychiatry Ment Health 2020 Mar 30;14(1):12 [FREE Full text] [doi: 10.1186/s13034-020-00318-6] [Medline: 32256690]
7. Chen L, Wang L, Qiu XH, Yang XX, Qiao ZX, Yang YJ, et al. Depression among Chinese university students: prevalence and socio-demographic correlates. PLoS One 2013 Mar 13;8(3):e58379 [FREE Full text] [doi: 10.1371/journal.pone.0058379] [Medline: 23516468]
8. Sun L, Sun L, Sun Y, Yang L, Wu H, Zhang D, et al. Correlations between psychological symptoms and social relationships among medical undergraduates in Anhui Province of China. Int J Psychiatry Med 2011 Nov 21;42(1):29-47. [doi: 10.2190/pm.42.1.c] [Medline: 22372023]
9. Sun X, Niu G, You Z, Zhou Z, Tang Y. Gender, negative life events and coping on different stages of depression severity: A cross-sectional study among Chinese university students. J Affect Disord 2017 Feb;209:177-181. [doi: 10.1016/j.jad.2016.11.025] [Medline: 27923194]
10. Hou H, Feng X, Li Y, Meng Z, Guo D, Wang F, et al. Suboptimal health status and psychological symptoms among Chinese college students: a perspective of predictive, preventive and personalised health. EPMA J 2018 Dec 28;9(4):367-377 [FREE Full text] [doi: 10.1007/s13167-018-0148-4] [Medline: 30538788]

11. Gorsic M, Novak D. Design and pilot evaluation of competitive and cooperative exercise games for arm rehabilitation at home. Annu Int Conf IEEE Eng Med Biol Soc 2016 Aug;2016:4690-4694 [FREE Full text] [doi: 10.1109/EMBC.2016.7591774] [Medline: 28269319]
12. Lei X, Liu C, Jiang H. Mental health of college students and associated factors in Hubei of China. PLoS One 2021 Jul 2;16(7):e0254183 [FREE Full text] [doi: 10.1371/journal.pone.0254183] [Medline: 34214107]
13. Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, Clarahan M, et al. The prevalence and correlates of depression, anxiety, and stress in a sample of college students. J Affect Disord 2015 Mar 01;173:90-96. [doi: 10.1016/j.jad.2014.10.054] [Medline: 25462401]
14. Kontoangelos K, Tsiori S, Koundi K, Pappa X, Sakkas P, Papageorgiou C. Greek college students and psychopathology: new insights. Int J Environ Res Public Health 2015 Apr 29;12(5):4709-4725 [FREE Full text] [doi: 10.3390/ijerph120504709] [Medline: 25938913]
15. Wang J, Zou J, Luo J, Liu H, Yang Q, Ouyang Y, et al. Mental health symptoms among rural adolescents with different parental migration experiences: A cross-sectional study in China. Psychiatry Res 2019 Sep;279:222-230. [doi: 10.1016/j.psychres.2019.03.004] [Medline: 30878304]
16. Alonso J, Mortier P, Auerbach RP, Bruffaerts R, Vilagut G, Cuijpers P, WHO WMH-ICS Collaborators. Severe role impairment associated with mental disorders: Results of the WHO World Mental Health Surveys International College Student Project. Depress Anxiety 2018 Sep 30;35(9):802-814 [FREE Full text] [doi: 10.1002/da.22778] [Medline: 29847006]
17. Alonso J, Vilagut G, Mortier P, Auerbach RP, Bruffaerts R, Cuijpers P, WHO WMH-ICS Collaborators. The role impairment associated with mental disorder risk profiles in the WHO World Mental Health International College Student Initiative. Int J Methods Psychiatr Res 2019 Jun 06;28(2):e1750 [FREE Full text] [doi: 10.1002/mpr.1750] [Medline: 30402985]
18. Bruffaerts R, Mortier P, Kiekens G, Auerbach RP, Cuijpers P, Demyttenaere K, et al. Mental health problems in college freshmen: prevalence and academic functioning. J Affect Disord 2018 Jan 01;225:97-103 [FREE Full text] [doi: 10.1016/j.jad.2017.07.044] [Medline: 28802728]
19. Gao L, Xie Y, Jia C, Wang W. Prevalence of depression among Chinese university students: a systematic review and meta-analysis. Sci Rep 2020 Sep 28;10(1):15897. [doi: 10.1038/s41598-020-72998-1] [Medline: 32985593]
20. Ngasa SN, Sama C, Dzekem BS, Nforchu KN, Tindong M, Aroke D, et al. Prevalence and factors associated with depression among medical students in Cameroon: a cross-sectional study. BMC Psychiatry 2017 Jun 09;17(1):216 [FREE Full text] [doi: 10.1186/s12888-017-1382-3] [Medline: 28599624]
21. Blanco C, Okuda M, Wright C, Hasin DS, Grant BF, Liu S, et al. Mental health of college students and their non-college-attending peers: results from the National Epidemiologic Study on Alcohol and Related Conditions. Arch Gen Psychiatry 2008 Dec;65(12):1429-1437. [doi: 10.1001/archpsyc.65.12.1429] [Medline: 19047530]
22. Arayeshgari M, Tapak L, Roshanaei G, Poorolajal J, Ghaleiha A. Application of group smoothly clipped absolute deviation method in identifying correlates of psychiatric distress among college students. BMC Psychiatry 2020 May 04;20(1):198 [FREE Full text] [doi: 10.1186/s12888-020-02591-3] [Medline: 32366242]
23. Plexousakis SS, Kourkoutas E, Giovazolias T, Chatira K, Nikolopoulos D. School bullying and post-traumatic stress disorder symptoms: the role of parental bonding. Front Public Health 2019;7:75. [doi: 10.3389/fpubh.2019.00075] [Medline: 31024876]
24. Sayers J. The world health report 2001 — Mental health: new understanding, new hope. Bull World Health Organ 2001;79(11):1085 [FREE Full text] [doi: 10.1590/S0042-96862001001100014]
25. Wu X, Tao S, Zhang Y, Zhang S, Tao F. Low physical activity and high screen time can increase the risks of mental health problems and poor sleep quality among Chinese college students. PLoS One 2015;10(3):e0119607 [FREE Full text] [doi: 10.1371/journal.pone.0119607] [Medline: 25786030]
26. Mak WW, Chio FH, Chan AT, Lui WW, Wu EK. The efficacy of internet-based mindfulness training and cognitive-behavioral training with telephone support in the Enhancement of mental health among college students and young working adults: randomized controlled trial. J Med Internet Res 2017 Mar 22;19(3):e84 [FREE Full text] [doi: 10.2196/jmir.6737] [Medline: 28330831]
27. Promoting mental health: concepts, emerging evidence, practice: summary report. World Health Organization. 2004. URL: https://apps.who.int/iris/handle/10665/42940 [accessed 2022-02-07]
28. Tol WA, Barbui C, Galappatti A, Silove D, Betancourt TS, Souza R, et al. Mental health and psychosocial support in humanitarian settings: linking practice and research. Lancet 2011 Oct 29;378(9802):1581-1591 [FREE Full text] [doi: 10.1016/S0140-6736(11)61094-5] [Medline: 22008428]
29. Xiang Y, Dong X, Zhao J. Effects of envy on depression: the mediating roles of psychological resilience and social support. Psychiatry Investig 2020 Jun;17(6):547-555 [FREE Full text] [doi: 10.30773/pi.2019.0266] [Medline: 32517416]
30. Zhao J, Peng X, Chao X, Xiang Y. Childhood maltreatment influences mental symptoms: the mediating roles of emotional intelligence and social support. Front Psychiatry 2019 Jun 21;10:415. [doi: 10.3389/fpsyt.2019.00415] [Medline: 31316399]
31. Taylor SE. Social support: a review. In: Friedman HS, editor. The Oxford Handbook of Health Psychology. Oxford, UK: Oxford University Press; 2011.
32. Tahmasbipour N, Taheri A. A survey on the relation between social support and mental health in students Shahid Rajaee University. Procedia Soc Behav Sci 2012;47:5-9. [doi: 10.1016/j.sbspro.2012.06.603]

33. Wang X, Cai L, Qian J, Peng J. Social support moderates stress effects on depression. Int J Ment Health Syst 2014;8(1):41 [FREE Full text] [doi: 10.1186/1752-4458-8-41] [Medline: 25422673]
34. Li J, Liang W, Yuan B, Zeng G. Internalized stigmatization, social support, and individual mental health problems in the public health crisis. Int J Environ Res Public Health 2020 Jun 23;17(12):4507 [FREE Full text] [doi: 10.3390/ijerph17124507] [Medline: 32585910]
35. Wahle F, Bollhalder L, Kowatsch T, Fleisch E. Toward the design of evidence-based mental health information systems for people with depression: a systematic literature review and meta-analysis. J Med Internet Res 2017 May 31;19(5):e191 [FREE Full text] [doi: 10.2196/jmir.7381] [Medline: 28566267]
36. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 01;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
37. de Mello MF, de Jesus Mari J, Bacaltchuk J, Verdeli H, Neugebauer R. A systematic review of research findings on the efficacy of interpersonal therapy for depressive disorders. Eur Arch Psychiatry Clin Neurosci 2005 Apr 12;255(2):75-82. [doi: 10.1007/s00406-004-0542-x] [Medline: 15812600]
38. Bell AC, D'Zurilla TJ. Problem-solving therapy for depression: a meta-analysis. Clin Psychol Rev 2009 Jun;29(4):348-353. [doi: 10.1016/j.cpr.2009.02.003] [Medline: 19299058]
39. Brent DA, Holder D, Kolko D, Birmaher B, Baugher M, Roth C, et al. A clinical psychotherapy trial for adolescent depression comparing cognitive, family, and supportive therapy. Arch Gen Psychiatry 1997 Sep 01;54(9):877-885. [doi: 10.1001/archpsyc.1997.01830210125017] [Medline: 9294380]
40. Cuijpers P. A psychoeducational approach to the treatment of depression: A meta-analysis of Lewinsohn's “coping with depression” course. Behav Ther 1998;29(3):521-533. [doi: 10.1016/S0005-7894(98)80047-6]
41. Ströhle A. Physical activity, exercise, depression and anxiety disorders. J Neural Transm 2009 Jun;116(6):777-784. [doi: 10.1007/s00702-008-0092-x] [Medline: 18726137]
42. Drum DJ, Denmark AB. Campus suicide prevention: bridging paradigms and forging partnerships. Harv Rev Psychiatry 2012;20(4):209-221. [doi: 10.3109/10673229.2012.712841] [Medline: 22894730]
43. Farrer L, Gulliver A, Chan JKY, Batterham PJ, Reynolds J, Calear A, et al. Technology-based interventions for mental health in tertiary students: systematic review. J Med Internet Res 2013 May 27;15(5):e101 [FREE Full text] [doi: 10.2196/jmir.2639] [Medline: 23711740]
44. Mowbray CT, Megivern D, Mandiberg JM, Strauss S, Stein CH, Collins K, et al. Campus mental health services: recommendations for change. Am J Orthopsychiatry 2006 Apr;76(2):226-237. [doi: 10.1037/0002-9432.76.2.226] [Medline: 16719642]
45. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
46. Reger MA, Gahm GA. A meta-analysis of the effects of internet- and computer-based cognitive-behavioral treatments for anxiety. J Clin Psychol 2009 Jan;65(1):53-75. [doi: 10.1002/jclp.20536] [Medline: 19051274]
47. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
48. Yu Z, Wang L, Chen W, Perrin N, Gross D. Childhood adversity and mental health among Chinese young adults: the protective role of resilience. J Adv Nurs 2021 Dec 07;77(12):4793-4804. [doi: 10.1111/jan.15070] [Medline: 34617614]
49. Zhang W, Xiong S, Zheng Y, Wu J. Response efficacy and self-efficacy mediated the relationship between perceived threat and psychic anxiety among college students in the early stage of the COVID-19 pandemic. Int J Environ Res Public Health 2022 Feb 28;19(5):2832 [FREE Full text] [doi: 10.3390/ijerph19052832] [Medline: 35270528]
50. Shen Y, Zhang Y, Chan BSM, Meng F, Yang T, Luo X, et al. Association of ADHD symptoms, depression and suicidal behaviors with anxiety in Chinese medical college students. BMC Psychiatry 2020 Apr 22;20(1):180 [FREE Full text] [doi: 10.1186/s12888-020-02555-7] [Medline: 32321462]
51. Sze KYP, Lee EKP, Chan RHW, Kim JH. Prevalence of negative emotional eating and its associated psychosocial factors among urban Chinese undergraduates in Hong Kong: a cross-sectional study. BMC Public Health 2021 Mar 24;21(1):583 [FREE Full text] [doi: 10.1186/s12889-021-10531-3] [Medline: 33761930]
52. Zhao N, Zhou G. Social media use and mental health during the COVID‐19 pandemic: moderator role of disaster stressor and mediator role of negative affect. Appl Psychol Health Well‐Being 2020 Sep 17;12(4):1019-1038. [doi: 10.1111/aphw.12226]
53. Yu M, Tian F, Cui Q, Wu H. Prevalence and its associated factors of depressive symptoms among Chinese college students during the COVID-19 pandemic. BMC Psychiatry 2021 Jan 29;21(1):66 [FREE Full text] [doi: 10.1186/s12888-021-03066-9] [Medline: 33514336]
54. Li Y, Peng J. Does social support matter? The mediating links with coping strategy and anxiety among Chinese college students in a cross-sectional study of COVID-19 pandemic. BMC Public Health 2021 Jul 02;21(1):1298 [FREE Full text] [doi: 10.1186/s12889-021-11332-4] [Medline: 34215244]

55. Wang T, Wong JYH, Wang MP, Li ACY, Kim SS, Lee JJ. Effects of social networking service (SNS) addiction on mental health status in Chinese university students: structural equation modeling approach using a cross-sectional online survey. J Med Internet Res 2021 Dec 08;23(12):e26733 [FREE Full text] [doi: 10.2196/26733] [Medline: 34889760]
56. Sit HF, Hong IW, Burchert S, Sou EKL, Wong M, Chen W, et al. A feasibility study of the WHO Digital Mental Health Intervention step-by-step to address depression among Chinese young adults. Front Psychiatry 2021 Jan 7;12:812667. [doi: 10.3389/fpsyt.2021.812667] [Medline: 35069297]
57. Liang S, Chen R, Liu L, Li X, Chen J, Tang S, et al. The psychological impact of the COVID-19 epidemic on Guangdong college students: the difference between seeking and not seeking psychological help. Front Psychol 2020 Sep 4;11:2231. [doi: 10.3389/fpsyg.2020.02231] [Medline: 33013582]
58. Nurunnabi M, Hossain SFAH, Chinna K, Sundarasen S, Khoshaim HB, Kamaludin K, et al. Coping strategies of students for anxiety during the COVID-19 pandemic in China: a cross-sectional study. F1000Res 2020 Sep 10;9:1115 [FREE Full text] [doi: 10.12688/f1000research.25557.1] [Medline: 33274049]
59. Wu R, Guo L, Rong H, Shi J, Li W, Zhu M, et al. The role of problematic smartphone uses and psychological distress in the relationship between sleep quality and disordered eating behaviors among Chinese college students. Front Psychiatry 2021 Dec 13;12:793506. [doi: 10.3389/fpsyt.2021.793506] [Medline: 34966312]
60. Chen X, Zhang X, Zhu X, Wang G. Efficacy of an internet-based intervention for subclinical depression (MoodBox) in China: study protocol for a randomized controlled trial. Front Psychiatry 2020 Jan 12;11:585920. [doi: 10.3389/fpsyt.2020.585920] [Medline: 33510658]
61. Yu Y, She R, Luo S, Xin M, Li L, Wang S, et al. Factors influencing depression and mental distress related to COVID-19 among university students in China: online cross-sectional mediation study. JMIR Ment Health 2021 Feb 22;8(2):e22705 [FREE Full text] [doi: 10.2196/22705] [Medline: 33616541]
62. Zhang Y, Zhao J, Xi J, Fan B, Wang Q, Yao Z, et al. The prevalence and determinant of PTSD symptoms among home-quarantined Chinese university students during the COVID-19 pandemic. Healthcare 2021 Oct 16;9(10):1383 [FREE Full text] [doi: 10.3390/healthcare9101383] [Medline: 34683063]
63. Li N, Fan L, Wang Y, Wang J, Huang Y. Risk factors of psychological distress during the COVID-19 pandemic: the roles of coping style and emotional regulation. J Affect Disord 2022 Feb 15;299:326-334 [FREE Full text] [doi: 10.1016/j.jad.2021.12.026] [Medline: 34920036]
64. Tang W, Hu T, Hu B, Jin C, Wang G, Xie C, et al. Prevalence and correlates of PTSD and depressive symptoms one month after the outbreak of the COVID-19 epidemic in a sample of home-quarantined Chinese university students. J Affect Disord 2020 Sep 01;274:1-7 [FREE Full text] [doi: 10.1016/j.jad.2020.05.009] [Medline: 32405111]
65. Zhou W, Li P, Lei X, Yuan H. Envision a bright future to heal your negative mood: a trial in China. Cogn Ther Res 2021 Aug 11;46(2):343-357. [doi: 10.1007/s10608-021-10246-0]
66. Shen Y, Chan BSM, Huang C, Cui X, Liu J, Lu J, et al. Suicidal behaviors and attention deficit hyperactivity disorder (ADHD): a cross-sectional study among Chinese medical college students. BMC Psychiatry 2021 May 18;21(1):258 [FREE Full text] [doi: 10.1186/s12888-021-03247-6] [Medline: 34006240]
67. Li D, Zou L, Zhang Z, Zhang P, Zhang J, Fu W, et al. The psychological effect of COVID-19 on home-quarantined nursing students in China. Front Psychiatry 2021 Mar 26;12:652296. [doi: 10.3389/fpsyt.2021.652296] [Medline: 33897502]
68. Yang K, Wang L, Liu H, Li L, Jiang X. Impact of coronavirus disease 2019 on the mental health of university students in Sichuan Province, China: An online cross-sectional study. Int J Ment Health Nurs 2021 Aug;30(4):875-884 [FREE Full text] [doi: 10.1111/inm.12828] [Medline: 33704896]
69. Liu K, Duan Y, Wang Y. The effectiveness of a web-based positive psychology intervention in enhancing college students' mental well-being. Soc Behav Pers 2021 Aug 04;49(8):1-13. [doi: 10.2224/sbp.10459]
70. Carciofo R. Morning affect, eveningness, and amplitude distinctness: associations with negative emotionality, including the mediating roles of sleep quality, personality, and metacognitive beliefs. Chronobiol Int 2020 Nov 13;37(11):1565-1579. [doi: 10.1080/07420528.2020.1798978] [Medline: 32787687]
71. Yen J, Yen C, Wu H, Huang C, Ko C. Hostility in the real world and online: the effect of internet addiction, depression, and online activity. Cyberpsychol Behav Soc Netw 2011 Nov;14(11):649-655. [doi: 10.1089/cyber.2010.0393] [Medline: 21476897]
72. Zheng X, Guo Y, Ma W, Yang H, Luo L, Wen L, et al. A longitudinal study on the mental health of college students in Jinan during the peak stage of the COVID-19 epidemic and the society reopening. Biomed Hub 2021 Oct 12;6(3):102-110 [FREE Full text] [doi: 10.1159/000519586] [Medline: 34950671]
73. Song B, Zhao Y, Zhu J. COVID-19-related traumatic effects and psychological reactions among international students. J Epidemiol Glob Health 2021 Mar;11(1):117-123 [FREE Full text] [doi: 10.2991/jegh.k.201016.001] [Medline: 33605116]
74. Tao X, Chen D, Fan Y, Zhang L, Shan H, Wei Y, et al. A cross-sectional study for the mental health status and sleep quality among college students in Macao during the COVID-19 pandemic. PeerJ 2021;9:e12520. [doi: 10.7717/peerj.12520] [Medline: 34966578]
75. Jia Y, Qi Y, Bai L, Han Y, Xie Z, Ge J. Knowledge-attitude-practice and psychological status of college students during the early stage of COVID-19 outbreak in China: a cross-sectional study. BMJ Open 2021 Feb 05;11(2):e045034 [FREE Full text] [doi: 10.1136/bmjopen-2020-045034] [Medline: 33550266]

76. Dun Y, Ripley-Gonzalez JW, Zhou N, You B, Li Q, Li H, et al. Weight gain in Chinese youth during a 4-month COVID-19 lockdown: a retrospective observational study. BMJ Open 2021 Jul 22;11(7):e052451 [FREE Full text] [doi: 10.1136/bmjopen-2021-052451] [Medline: 34301671]
77. Yu J, Yang Z, Wu Y, Ge M, Tang X, Jiang H. Prevalence of and factors associated with depressive symptoms among college students in Wuhan, China during the normalization stage of COVID-19 prevention and control. Front Psychiatry 2021 Oct 15;12:742950. [doi: 10.3389/fpsyt.2021.742950] [Medline: 34721111]
78. Pan J, Zhuang X. Effectiveness of an adventure-based cognitive behavioral program for Hong Kong university students. Res Soc Work Pract 2021 Oct 19;32(2):227-237. [doi: 10.1177/10497315211042096]
79. Auyeung L, Mo PKH. The efficacy and mechanism of online positive psychological intervention (PPI) on improving well-being among Chinese university students: a pilot study of the Best Possible Self (BPS) intervention. J Happiness Stud 2018 Nov 26;20(8):2525-2550. [doi: 10.1007/s10902-018-0054-4]
80. Zhao C, Xu H, Lai X, Yang X, Tu X, Ding N, et al. Effects of online social support and perceived social support on the relationship between perceived stress and problematic smartphone usage among Chinese undergraduates. Psychol Res Behav Manag 2021;14:529-539. [doi: 10.2147/PRBM.S302551] [Medline: 33976576]
81. Xin M, Luo S, She R, Yu Y, Li L, Wang S, et al. Negative cognitive and psychological correlates of mandatory quarantine during the initial COVID-19 outbreak in China. Am Psychol 2020;75(5):607-617. [doi: 10.1037/amp0000692] [Medline: 32673008]
82. Liang K, de Lucena Martins CM, Chen S, Clark CCT, Duncan MJ, Bu H, et al. Sleep as a priority: 24-hour movement guidelines and mental health of Chinese college students during the COVID-19 pandemic. Healthcare 2021 Sep 06;9(9):1166 [FREE Full text] [doi: 10.3390/healthcare9091166] [Medline: 34574940]
83. Sun Y, Lin S, Chung KKH. University students' perceived peer support and experienced depressive Symptoms during the COVID-19 pandemic: the mediating role of emotional well-being. Int J Environ Res Public Health 2020 Dec 12;17(24):9308 [FREE Full text] [doi: 10.3390/ijerph17249308] [Medline: 33322751]
84. Li Y, Qin L, Shi Y, Han J. The psychological symptoms of college student in China during the lockdown of COVID-19 epidemic. Healthcare 2021 Apr 11;9(4):447 [FREE Full text] [doi: 10.3390/healthcare9040447] [Medline: 33920411]
85. Zhu Y, Wang H, Wang A. An evaluation of mental health and emotion regulation experienced by undergraduate nursing students in China during the COVID-19 pandemic: A cross-sectional study. Int J Ment Health Nurs 2021 Oct 13;30(5):1160-1169 [FREE Full text] [doi: 10.1111/inm.12867] [Medline: 33848056]
86. Zhuo L, Wu Q, Le H, Li H, Zheng L, Ma G, et al. COVID-19-related intolerance of uncertainty and mental health among back-to-school students in Wuhan: the moderation effect of social support. Int J Environ Res Public Health 2021 Jan 22;18(3):981 [FREE Full text] [doi: 10.3390/ijerph18030981] [Medline: 33499409]
87. Li X, Fu P, Fan C, Zhu M, Li M. COVID-19 stress and mental health of students in locked-down colleges. Int J Environ Res Public Health 2021 Jan 18;18(2):771 [FREE Full text] [doi: 10.3390/ijerph18020771] [Medline: 33477595]
88. Zhan H, Zheng C, Zhang X, Yang M, Zhang L, Jia X. Chinese college students' stress and anxiety levels under COVID-19. Front Psychiatry 2021 Jun 10;12:615390. [doi: 10.3389/fpsyt.2021.615390] [Medline: 34177635]
89. Ding Y, Fu X, Liu R, Hwang J, Hong W, Wang J. The impact of different coping styles on psychological distress during the COVID-19: the mediating role of perceived stress. Int J Environ Res Public Health 2021 Oct 18;18(20):10947 [FREE Full text] [doi: 10.3390/ijerph182010947] [Medline: 34682693]
90. Zhou Y, Lemmer G, Xu J, Rief W. Cross-cultural measurement invariance of scales assessing stigma and attitude to seeking professional psychological help. Front Psychol 2019;10:1249. [doi: 10.3389/fpsyg.2019.01249] [Medline: 31214074]
91. Liang Z, Kang D, Zhang M, Xia Y, Zeng Q. The impact of the COVID-19 pandemic on Chinese postgraduate students' mental health. Int J Environ Res Public Health 2021 Nov 03;18(21):11542 [FREE Full text] [doi: 10.3390/ijerph182111542] [Medline: 34770056]
92. Wen F, Zhu J, Ye H, Li L, Ma Z, Wen X, et al. Associations between insecurity and stress among Chinese university students: the mediating effects of hope and self-efficacy. J Affect Disord 2021 Feb 15;281:447-453. [doi: 10.1016/j.jad.2020.12.047] [Medline: 33360366]
93. Lin X, Zhang C, Yang S, Hsu M, Cheng H, Chen J, et al. Stress and its association with academic performance among dental undergraduate students in Fujian, China: a cross-sectional online questionnaire survey. BMC Med Educ 2020 Jun 03;20(1):181 [FREE Full text] [doi: 10.1186/s12909-020-02095-4] [Medline: 32493378]
94. Caravaca-Sánchez F, Aizpurua E, Stephenson A. Substance use, family functionality, and mental health among college students in Spain. Soc Work Public Health 2021 Feb 17;36(2):221-231. [doi: 10.1080/19371918.2020.1869134] [Medline: 33463401]
95. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
96. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Res 2020 May;287:112934 [FREE Full text] [doi: 10.1016/j.psychres.2020.112934] [Medline: 32229390]

97. Wang X, Hegde S, Son C, Keller B, Smith A, Sasangohar F. Investigating mental health of US college students during the COVID-19 pandemic: cross-sectional survey study. J Med Internet Res 2020 Sep 17;22(9):e22817 [FREE Full text] [doi: 10.2196/22817] [Medline: 32897868]
98. Zou P, Sun L, Yang W, Zeng Y, Chen Q, Yang H, et al. Associations between negative life events and anxiety, depressive, and stress symptoms: A cross-sectional study among Chinese male senior college students. Psychiatry Res 2018 Dec;270:26-33. [doi: 10.1016/j.psychres.2018.09.019] [Medline: 30243129]
99. Mak IWC, Chu CM, Pan PC, Yiu MGC, Chan VL. Long-term psychiatric morbidities among SARS survivors. Gen Hosp Psychiatry 2009 Jul;31(4):318-326 [FREE Full text] [doi: 10.1016/j.genhosppsych.2009.03.001] [Medline: 19555791]
100. Yalcin I. Relationships between well-being and social support: a meta analysis study. Turkish J Psych 2015;26(1):21-32. [doi: 10.5080/u7769]
101. Zhu S, Zhuang Y, Ip P. Impacts on children and adolescents' lifestyle, social support and their association with negative impacts of the COVID-19 pandemic. Int J Environ Res Public Health 2021 Apr 29;18(9):4780 [FREE Full text] [doi: 10.3390/ijerph18094780] [Medline: 33947146]
102. Meshi D, Ellithorpe ME. Problematic social media use and social support received in real-life versus on social media: associations with depression, anxiety and social isolation. Addict Behav 2021 Aug;119:106949. [doi: 10.1016/j.addbeh.2021.106949] [Medline: 33934007]
103. Fu C, Wang G, Shi X, Cao F. Social support and depressive symptoms among physicians in tertiary hospitals in China: a cross-sectional study. BMC Psychiatry 2021 Apr 29;21(1):217 [FREE Full text] [doi: 10.1186/s12888-021-03219-w] [Medline: 33926402]
104. Pössel P, Burton SM, Cauley B, Sawyer MG, Spence SH, Sheffield J. Associations between social support from family, friends, and teachers and depressive symptoms in adolescents. J Youth Adolesc 2018 Feb 10;47(2):398-412. [doi: 10.1007/s10964-017-0712-6] [Medline: 28695369]
105. Rahman M, Amin R, Liton M, Hossain N. Disha: an implementation of machine learning based Bangla healthcare chatbot. 2019 Presented at: 22nd International Conference on Computer and Information Technology (ICCIT); December 18-20, 2019; Bangladesh. [doi: 10.1109/iccit48885.2019.9038579]
106. Hao K. The pandemic is emptying call centers. AI chatbots are swooping in. MIT Technology Review. 2020 May 14. URL: https://www.technologyreview.com/2020/05/14/1001716/ai-chatbots-take-call-center-jobs-during-coronavirus-pandemic/ [accessed 2022-02-07]
107. Judson T, Odisho A, Young J, Bigazzi O, Steuer D, Gonzales R, et al. Implementation of a digital chatbot to screen health system employees during the COVID-19 pandemic. J Am Med Inform Assoc 2020 Jul 01;27(9):1450-1455 [FREE Full text] [doi: 10.1093/jamia/ocaa130] [Medline: 32531066]
108. Miner AS, Laranjo L, Kocaballi AB. Chatbots in the fight against the COVID-19 pandemic. NPJ Digit Med 2020 May 4;3(1):65. [doi: 10.1038/s41746-020-0280-0] [Medline: 32377576]
109. Mittal A, Agrawal A, Chouksey A, Shriwas R, Agrawal S. A comparative study of chatbots and humans. Int J Adv Res Comput Commun Eng 2016;5(3):1055-1057. [doi: 10.17148/IJARCCE.2016.53253]
110. AbuShawar B, Atwell E. Usefulness, localizability, humanness, and language-benefit: additional evaluation criteria for natural language dialogue systems. Int J Speech Technol 2016 Jan 4;19(2):373-383. [doi: 10.1007/s10772-015-9330-4]
111. Følstad A, Skjuve M. Chatbots for customer service: user experience and motivation. 2019 Presented at: 1st International Conference on Conversational User Interfaces (CUI 2019); August 22-23, 2019; Dublin, Ireland. [doi: 10.1145/3342775.3342784]
112. Escobar DFSS, Noll PRES, Jesus TFD, Noll M. Assessing the mental health of Brazilian students involved in risky behaviors. Int J Environ Res Public Health 2020 May 22;17(10):3647 [FREE Full text] [doi: 10.3390/ijerph17103647] [Medline: 32455911]
113. Leskelä U, Rytsälä H, Komulainen E, Melartin T, Sokero P, Lestelä-Mielonen P, et al. The influence of adversity and perceived social support on the outcome of major depressive disorder in subjects with different levels of depressive symptoms. Psychol Med 2006 Mar 28;36(6):779-788. [doi: 10.1017/s0033291706007276]
114. Hallgren M, Lundin A, Tee FY, Burström BO, Forsell Y. Somebody to lean on: Social relationships predict post-treatment depression severity in adults. Psychiatry Res 2017 Mar;249:261-267. [doi: 10.1016/j.psychres.2016.12.060] [Medline: 28131948]
115. Folkman S. Stress, coping, and hope. Psychooncology 2010 Sep 26;19(9):901-908. [doi: 10.1002/pon.1836] [Medline: 20799373]
116. Yan L, Gan Y, Ding X, Wu J, Duan H. The relationship between perceived stress and emotional distress during the COVID-19 outbreak: effects of boredom proneness and coping style. J Anxiety Disord 2021 Jan;77:102328 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102328] [Medline: 33160275]
117. Reavley N, Jorm A. Prevention and early intervention to improve mental health in higher education students: a review. Early Interv Psychiatry 2010 May;4(2):132-142. [doi: 10.1111/j.1751-7893.2010.00167.x] [Medline: 20536969]
118. Davenport TA, Cheng VWS, Iorfino F, Hamilton B, Castaldi E, Burton A, et al. Flip the clinic: a digital health approach to youth mental health service delivery during the COVID-19 pandemic and beyond. JMIR Ment Health 2020 Dec 15;7(12):e24578 [FREE Full text] [doi: 10.2196/24578] [Medline: 33206051]

119. Kemp J, Chorney J, Kassam I, MacDonald J, MacDonald T, Wozney L, et al. Learning about the current state of digital mental health interventions for Canadian youth to inform future decision-making: mixed methods study. J Med Internet Res 2021 Oct 19;23(10):e30491 [FREE Full text] [doi: 10.2196/30491] [Medline: 34665141]
120. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. Mhealth 2018 Mar;4:6. [doi: 10.21037/mhealth.2018.03.02] [Medline: 29682510]
121. Abd-Alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M. An overview of the features of chatbots in mental health: a scoping review. Int J Med Inform 2019 Dec;132:103978. [doi: 10.1016/j.ijmedinf.2019.103978] [Medline: 31622850]
122. Liu S, Yang L, Zhang C, Xiang Y, Liu Z, Hu S, et al. Online mental health services in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e17-e18 [FREE Full text] [doi: 10.1016/S2215-0366(20)30077-8] [Medline: 32085841]
123. Zhu Y, Janssen M, Wang R, Liu Y. It is me, chatbot: working to address the COVID-19 outbreak-related mental health issues in China. User experience, satisfaction, and influencing factors. Int J Hum–Comput Interact 2021 Nov 01:1-13. [doi: 10.1080/10447318.2021.1988236]
124. Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. J Med Internet Res 2019 Apr 05;21(4):e12887 [FREE Full text] [doi: 10.2196/12887] [Medline: 30950796]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by T Leung; submitted 24.03.22; peer-reviewed by D Xia, Y Yu; comments to author 25.04.22; revised version received 16.05.22; accepted 27.05.22; published 15.06.22.

Please cite as:
Shan Y, Ji M, Xie W, Li R, Qian X, Zhang X, Hao T
Interventions in Chinese Undergraduate Students' Mental Health: Systematic Review
Interact J Med Res 2022;11(1):e38249
URL: https://www.i-jmr.org/2022/1/e38249
doi:10.2196/38249
PMID:35704383

Review

Existing Funding Sources in Degenerative Cervical Myelopathy Research: Scoping Review

Henry Bestwick[1*], BA; Jye Quan Teh[1*], BA; Oliver Mowforth[2], MA, MBBCHIR; Ben Grodzinski[1], BA; Mark Kotter[2], MD, MPhil, PhD; Benjamin Davies[2], BSc, MPhil, MBChB

[1]Clinical School, University of Cambridge, Cambridge, United Kingdom

[2]Academic Neurosurgery Unit, Department of Clinical Neurosurgery, University of Cambridge, Cambridge, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Benjamin Davies, BSc, MPhil, MBChB
Academic Neurosurgery Unit
Department of Clinical Neurosurgery
University of Cambridge
Academic Neurosurgery Unit
University of Cambridge
Cambridge, CB2 0QQ
United Kingdom
Phone: 44 779 478 0568
Email: bd375@cam.ac.uk

Abstract

Background: Degenerative cervical myelopathy (DCM) is a common, disabling condition of symptomatic cervical spinal cord compression that requires significant research advances to improve patient outcomes. A James Lind Alliance Partnership recently identified the top research priorities for DCM. To effectively address these priorities, appropriate funding of DCM research is essential.

Objective: The aim of this paper is to review current funding in DCM research and highlight future research funding opportunities.

Methods: A systematic search of Web of Science for “cervical AND myelopathy” was conducted. Papers exclusively studying DCM with declared funding and published between January 1, 1995, and March 21, 2020, were considered eligible. Funding sources were classified by country of origin and organization type. A grant search was also conducted using Dimensions.ai (Digital Science Ltd).

Results: A total of 621 papers were included, with 300 unique funding bodies. The top funders were AO Spine (n=87); National Institutes of Health, USA (n=63); and National Natural Science Foundation, China (n=63). Funding sources in the USA (n=242) supported the most DCM research, followed by China (n=209) and Japan (n=116). Funding in the USA was primarily provided by corporate or nonprofit organizations (146/242, 60.3%), while in China, the majority of funding was from institutions (208/209, 99.5%). Dimensions.ai gives an estimate for the total declared grant funding awards for DCM-specific research. Data here showed 180 grants awarded specifically for DCM research, with a total value of US $45.6 million since 1996.

Conclusions: DCM funding appears to be predominantly from the USA, China, and Japan, aligning with areas of high DCM research activity and underpinning the importance of funding to increasing research capacity. The existing funding sources differ from medical research in general, representing opportunities for future investment in DCM.

(Interact J Med Res 2022;11(1):e36194) doi:10.2196/36194

KEYWORDS

cervical cord; myelopathy; spondylosis; stenosis; disc herniation; ossification posterior longitudinal ligament; degeneration; research funding; systematic review; spinal cord; patient and public involvement

Introduction

Degenerative cervical myelopathy (DCM), often previously referred to as cervical spondylotic myelopathy, is a progressive, slow motion, spinal cord injury caused by degenerative changes that lead to narrowing of the spinal canal [1]. It is the most common nontraumatic cause of spinal cord impairment [2], with recent estimates suggesting that as many as 1 in 50 adults could be affected in their lifetime [3,4].

DCM can cause a range of symptoms, including loss of manual dexterity, imbalance and falls, and incontinence and pain [1]. The mainstay of treatment is decompressive surgery [5]. Although this has been demonstrated to offer the most meaningful benefit, recovery is rarely complete and most people are left with life-long disabilities [1,6]. In a recent comparison of quality of life in chronic disease, people with DCM were found to have one of the lowest 36-Item Short Form Survey (SF-36) scores of any chronic disease [6]. Research leading to improved outcomes is urgently required.

To formally address this problem, a consensus initiative was established to improve research efficiency in DCM. AO Spine Research Objectives and Common Data Elements for DCM (RECODE-DCM) is an international, multistakeholder partnership between surgeons, health care professionals, and patients [7]. A National Institute for Health Research (NIHR) James Lind Alliance priority setting partnership established the top 10 DCM research priorities, including raising awareness, developing new treatments and diagnostic tools, and acquiring a better understanding of pathophysiology [8].

To enable these questions to be addressed, research funding targeting these priorities is urgently needed. The main aim of this study is to characterize the funding of existing DCM research and identify potential future funding organizations. Within this, our objectives are as follows: to characterize which countries, organizations, and type of organizations fund the majority of DCM research and to provide an overview of the estimated total grant funding in DCM.

Methods

To characterize the funding of existing DCM research, we used 2 methods. The first was most closely aligned with a scoping review and involved formulating a research question, identifying relevant studies, and further categorizing and analyzing the results as is standard in a scoping review methodology [9].

Identifying the Research Question

The aim of this study was to characterize the funding of existing DCM research. Although reviews may normally examine the results of individual papers, our search was focused on extracting the funding information of papers and focusing only on whether the article was on DCM.

To estimate the number of funders in DCM research, we attempted to extract funding details from literature databases. Of the existing medical literature databases, the only database allowing extraction of funding details is Web of Science [10,11].

Identifying Relevant Studies

There are many different terms used for DCM around the world [12]. Therefore, to ensure the search was comprehensive, the search terms “cervical AND myelopathy” were used. All papers from January 1, 1995, to March 21, 2020, were eligible for inclusion. These were then filtered by the presence of funder details to exclude entries without explicit funding sources.

Study Selection

The search output was uploaded to Rayyan, a systematic review web platform [13]. Titles and abstracts were then screened independently by 2 authors (JQT and HB). The inclusion criteria for literature were the following: all languages, primary research and systematic or narrative reviews, preclinical and clinical studies, and DCM-related spinal conditions. Meanwhile, the exclusion criteria were the following: corrections, letters, editorials, commentaries, proposals, technical notes, and conference papers; myelopathy not caused by DCM; cervical spinal surgery not specific to DCM; radiculopathy only; and thoracic or lumbar myelopathy. Any conflicts or undecided papers were resolved by discussion between JQT and HB until consensus was reached.

Estimating DCM Grants Using Dimensions.ai

To supplement our study, we undertook an additional search.

Using a grant-searching function on Dimensions.ai (Digital Science Ltd), we gathered funding information from 1996 to the present day using keywords for DCM [14]. Dimensions.ai is a platform that can be used to search grants awarded for specific research. Dimensions.ai provides information on the research title and abstract, investigator, funding amount, and over what period the research is expected to be completed. However, it does not include information regarding whether an author has been supported by a general scholarship and undertaken research in a given area, nor does it give information if a grant awarded for another project coincidentally funded research in another field. It therefore gives an overview of estimated value for total grant funding awards specifically requested for DCM research alone.

The grant information we gathered included the total number of grants on Dimensions.ai, total grant funding, average grant awarded, and date and amount of earliest grant shown on Dimensions.ai.

As DCM has only recently been proposed as an umbrella term [12], a search was completed using the following DCM-related terms: “degenerative cervical myelopathy,” “cervical spondylotic myelopathy,” “ossification posterior longitudinal ligament,” “ossification ligamentum flavum,” “cervical myelopathy,” “cervical,” and “myelopathy” [15,16]. The search results were then screened manually to identify those specific to DCM. Any irrelevant research was excluded.

Charting the Results

Typical paper-specific information that could be recorded in a scoping review (such as aims, methodology, and results) was not necessary to answer our question on the sources of funding for DCM research. As the aim was to examine the funding

landscape of DCM research, more information regarding the characteristics of funding organizations was required.

Collating and Summarizing the Results

The total number of papers with a funding body was recorded. These funding bodies were then ordered with regard to how many papers they supported. The funding bodies were further delineated into their country of origin and the type of sector.

Funders were classified by country of origin by 2 authors (JQT and HB). Identifying countries associated with each funder required criteria to classify a funder: the funding body had to be a university, national funding body, provincial or state funding body, or organization or company; the funding body needed to have headquarters in a specific country; and the funding body could not have a country of origin that was unclear.

Some organizations are international without specific association with any particular country and were labeled as such, for example AO Spine. AO Spine is a global organization with headquarters in Switzerland. It is funded by the AO Foundation, with funding distributed globally. It was therefore felt to be best defined as an organization without a specific country of origin. Entries that did not satisfy the criteria were labeled as "unclear."

To further investigate funders, we classified them into 2 categories: institutional and corporate/charitable. The corporate/charitable group was further classified into for-profit and not-for-profit organizations.

Funders were classified as institutional if they were any of the following: a regional or central governmental funding body, a university, a research institution, or a hospital associated with a university or research institution.

Alternatively, they were classified as corporate/charitable if they were any of the following: a charitable or not-for profit organization or a for-profit organization or corporation.

The number of institutional organizations compared to corporate/charitable organizations was compared on a global and country level. If the funder did not satisfy either set of criteria or if it was unclear which category they would fit into, they were labeled "unclear."

Data Analysis

Data cleaning and visualization was conducted using Python [17-19].

Results

Study Selection

Of the 6757 papers returned from Web of Science, 621 papers acknowledged funding and survived passed application of our inclusion and exclusion criteria (Figure 1).

We identified 300 unique funding bodies that supported DCM research (Table 1). Many research papers had more than 1 funding body: there were a total of 920 references of funding from 300 unique funders.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

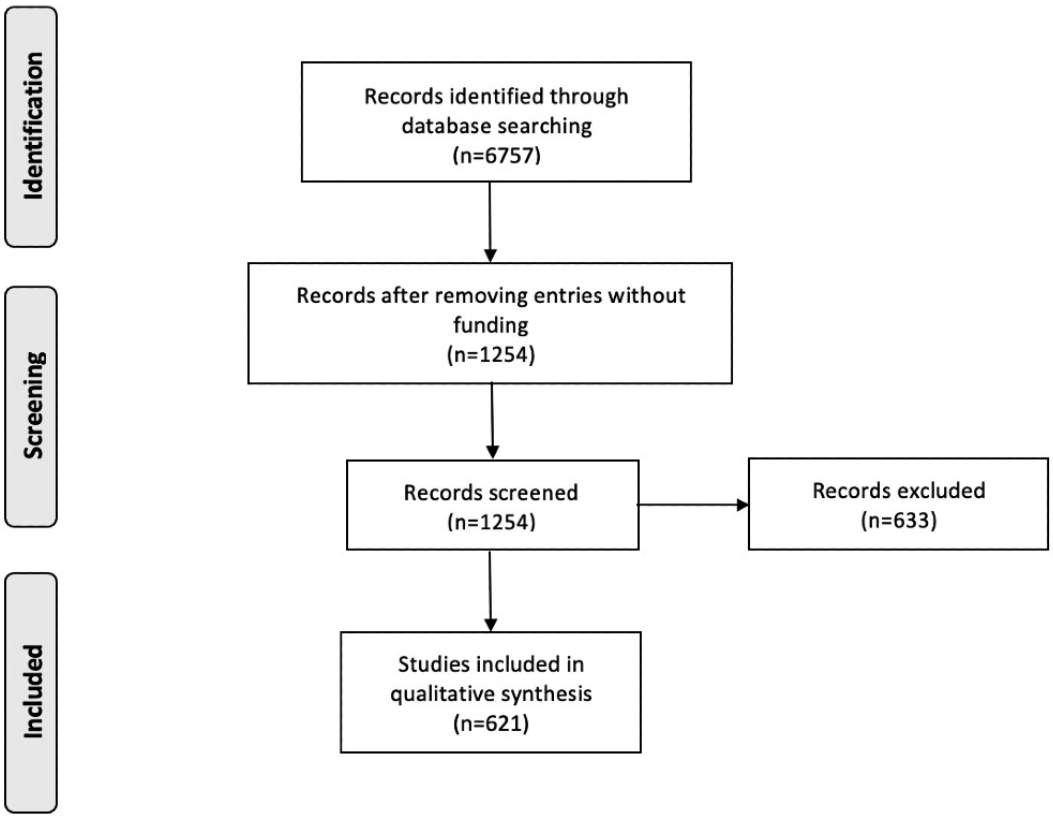


Table 1. Top 10 funding organizations for health research (annual figure, 2013) and DCM-specific research funding.

Rank	Top 10 funding organizations for all health research expenditure, millions[a]	Top 10 funders for DCM[b]-specific research by research output, number of papers
1	National Institutes of Health, 26,081.3	AO Spine, 87
2	European Commission, 3717.7	National Institutes of Health, 63
3	UK Medical Research Council, 1321.5	National Natural Science Foundation of China, 63
4	Institut national de la santé et de la recherche médicale, 1041.2	Ministry of Health, Labour and Welfare Japan, 47
5	United States Department of Defense, 1017.7	Ministry of Education, Culture, Sports, Science and Technology Japan, 25
6	Wellcome Trust, 909.1	DePuy Synthes, 22
7	Canadian Institutes of Health Research, 883.6	Cervical Spine Research Society, 18
8	Australian National Health and Medical Research Council, 777.6	DeZwirek Family Foundation, 18
9	Howard Hughes Medical Institute, 752.0	Gerald and Tootsie Halbert Chair in Neural Repair and Regeneration, 18
10	Deutsche Forschungsgemeinschaft/German Research Foundation, 630.6	National Research Foundation of Korea, 16

[a]In US dollars.

[b]DCM: degenerative cervical myelopathy.

Top Funders for DCM Research

The 300 funding bodies were then ordered according to the number of papers they supported. Table 1 shows this data for the top 10 DCM-specific funding organizations by research output and compares it to the top 10 funding organizations for general health research.

Funded DCM Research by Country

DCM research funding had a global distribution (Figure 2). The top 3 countries for number of funded DCM papers were the United States, China, and Japan, followed by Canada and the United Kingdom (Table 2 and Table 3). There were 112 papers without a specific country of origin, including 87 funded by AO Spine.

Figure 2. World heat map of degenerative cervical myelopathy funding sources. This map excludes funders that were not clearly associated with a specific country. The greater the number of funders of degenerative cervical myelopathy research, the hotter the colour of the country.

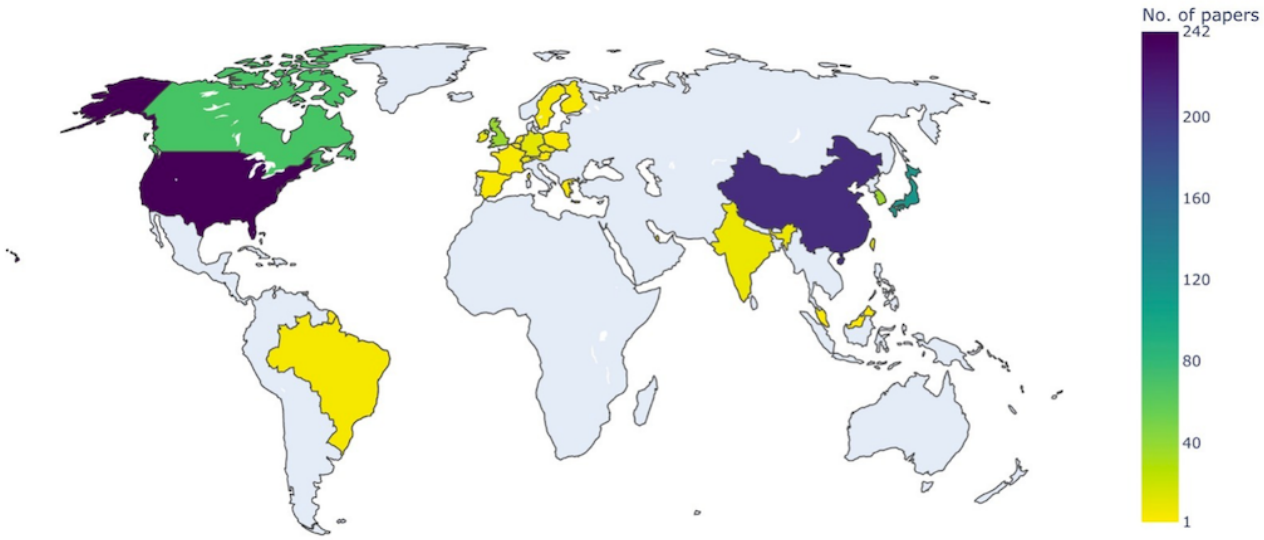


Table 2. Top 10 funding countries for health research and DCM-specific research funding.

Rank	Top 10 countries for research and development expenditure as a percentage[a] of that country's GDP[b]	Top 10 countries for DCM[c]-specific research funding, number of papers
1	Israel, 4.95	USA, 242
2	South Korea, 4.81	China, 209
3	Switzerland, 3.37	Japan, 116
4	Sweden, 3.34	Canada, 69
5	Japan, 3.26	United Kingdom, 38
6	Austria, 3.17	South Korea, 37
7	Germany, 3.09	Hong Kong, 18
8	Denmark, 3.06	Germany, 11
9	United States, 2.84	Switzerland, 8
10	Belgium, 2.82	Ireland, 7

[a]Total values not available.

[b]GDP: gross domestic product.

[c]DCM: degenerative cervical myelopathy.

Table 3. Top 10 countries by number of DCM papers that received research funding. The minimum percentage of papers from each country that was supported by research funding is estimated with reference to the total number DCM papers published during this time period from each country [20]. Raw data were requested directly from the author.

Country	Number of papers supported by funding as the percentage of total DCM[a] papers from the country, n/N (%)
United States	242/314 (77.1)
China	209/409 (51.1)
Japan	116/633 (18.3)
Canada	69/136 (50.7)
United Kingdom	38/60 (63.3)
South Korea	37/122 (30.3)
Hong Kong	18[b]
Germany	11/82 (13.4)
Switzerland	8/17 (47.0)
Ireland	7/9 (77.8)

[a]DCM: degenerative cervical myelopathy.

[b]Full data unavailable.

Funder Sectors

In total, 598/920 (65%) funding sources were institutional, 318/920 (34.6%) were a corporate/charitable source, and 4/920 (0.4%) were unclear. Of the 318 corporate/charitable sources, 229/318 (72%) were not-for-profit or charitable organizations, and 89/318 (28%) were for-profit corporations. Many funders supported more than 1 paper. The proportion of research funding from institutional and corporate or charitable funders varied across countries (Figure 3). China was the country with the greatest number of papers funded by institutional sources (n=208), followed by Japan (n=96) and the United States (n=96; Table 4). The Unites States was the country with the greatest number of papers funded by corporate or charitable funders (n=146), followed by Japan (n=20) and the United Kingdom (n=14).

Figure 3. Breakdown of funded papers in the top 10 countries. DCM: degenerative cervical myelopathy.

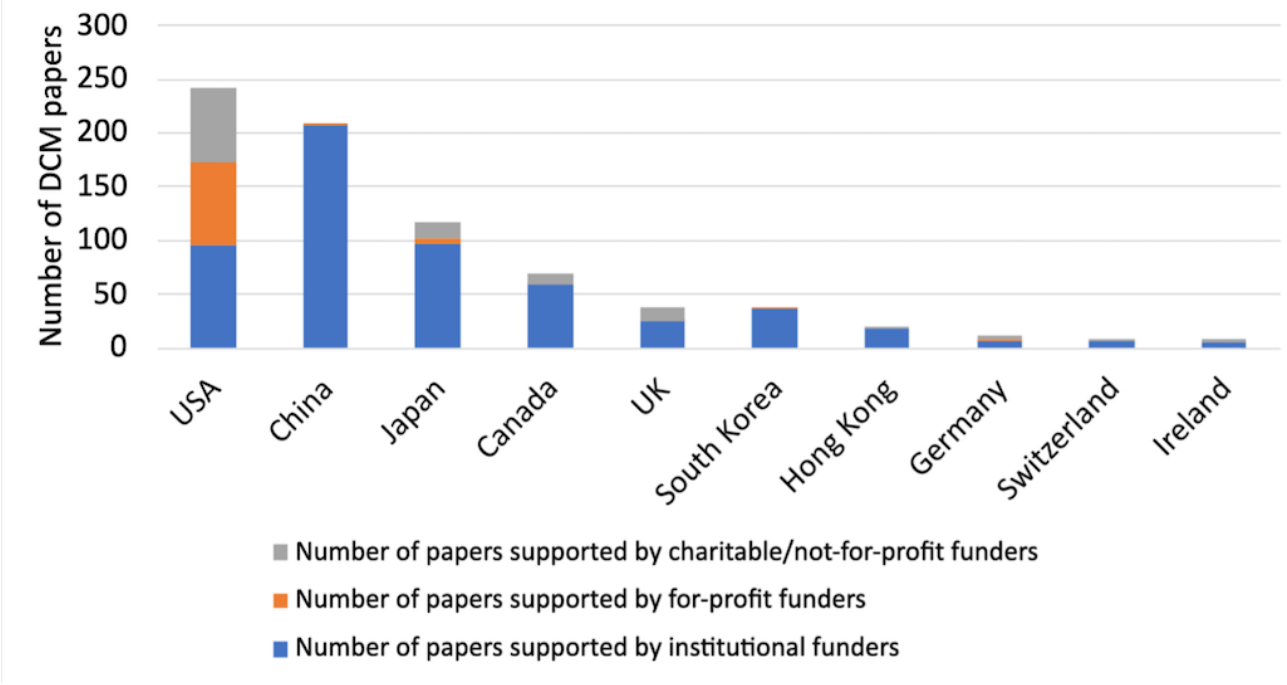


Table 4. Top 5 countries for papers supported by institutional and corporate/charitable funders.

Funder by country	Portion supported by funder, n/N (%)
Institutional funders	
China	208/209 (99.5)
Japan	96/116 (82.8)
United States	96/242 (39.7)
Canada	58/69 (84.1)
South Korea	36/37 (97.3)
Corporate/charitable funders	
United States	146/242 (60.3)
Japan	20/116 (17.2)
United Kingdom	14/38 (36.8)
Canada	11/69 (15.9)
Germany	5/11 (45.5)

Grant Funding Awards

The second part of the results relate to the search on Dimensions.ai. A total of US $45.6 million in grant funding for DCM papers was identified in Dimensions.ai (Table 5). Different results were obtained using variations of the search terms, including differing numbers of grants and a different percentage of grants that were manually verified to be DCM-specific.

Table 5. Portion grants that were found to be DCM-specific following manual verification by search term.

Search terms	DCM[a]-specific grants, n/N (%)	Total DCM funding, millions[b]
Degenerative cervical myelopathy	24/24 (100)	11.4
Cervical spondylotic myelopathy	48/63 (76)	14.1
Cervical myelopathy	104/128 (81)	20.0
Ossification posterior longitudinal ligament	4/81 (5)	0.137
Ossification ligamentum flavum	0/20 (0)	0

[a]DCM: degenerative cervical myelopathy.

[b]In US dollars.

Discussion

Principal Results

Our study identified 300 unique funding bodies for DCM research. A total of 621 papers acknowledged funding, largely provided by 4 organizations. These, aside from AO Spine, are associated with Japan, China, and the USA. Funding bodies originate from the corporate, charitable, and institutional sectors, but these are distributed unequally across different countries, and research is primarily supported by institutional bodies (598/920, 65%). Our Dimensions.ai research showed a minimum of US $45.6 million dollars of grant funding awards specifically for DCM from 1974 to 2020.

DCM Research Has Relied On a Small Number of Funders

Viergever and Hendriks [21] identified the top 10 funders for health research globally. They identified the US National Institutes of Health (NIH) as the largest funder of all, but many other leading providers were unrepresented in our DCM review, such as the European Commission, UK Medical Research Council, and the Wellcome Trust, the largest philanthropic funding body for health research [21]. Furthermore, only 5 out of 10 countries in our list of the top 10 countries that support DCM research were included in the top 10 of the World Bank's 2018 analysis of research and development research expenditure as a percentage of gross domestic product [22]. These countries include the USA, Japan, South Korea, Germany, and Switzerland. Interestingly, while Israel spent the most on research and development per gross domestic product, our analysis did not identify any funding bodies from Israel. Taken together, this suggests many unused funders and identifies opportunities for DCM research.

Corporate and Charity Sectors Are Underrepresented in DCM

DCM research has relied on institutional organizations, constituting 598 (64.1%) of the listed 920 reported funding sources. Although there are exceptions [11], this contrasts research funding as a whole, which is estimated to be 60% corporate, 30% institutional, and 10% from non-profit organizations [23]. For DCM, the corporate or non-profit sectors remain a challenge. The existing corporate sector for DCM is focused on medical devices and may not be best placed to support the full breadth of research priorities identified by AO Spine RECODE-DCM. For example, DePuy Synthes (a subsidiary of the Johnson & Johnson family of companies) funded 22/621 (3.5%) papers and are solely an orthopedic and neurosurgical device company [24].

Furthermore, while AO Spine has been a notable supporter, DCM does not have a specific funding organization comparable to ones like the Motor Neurone Disease Association or the Multiple Sclerosis Society [25,26]. Charitable organizations are not just significant research funders: they are essential for advocacy. In the United Kingdom for example, the charitable sector contributes £1.6 billion (US $1.7 billion) to medical research [27] and also acts as a lobby group [28]. With Myelopathy.org, DCM now has a dedicated charitable organization, with medical research funding being among its charitable aims [29].

Stakeholders in AO Spine RECODE-DCM have been cognizant of these challenges, which is reflected in the raising of awareness being established as the leading research priority and with understanding the disease burden and socioeconomic impact being among the other priorities.

Begum et al [11], however, demonstrated that burden of disease has been a relatively unimportant driver of research investment or activity. In an analysis of research funding by the US NIH, disease burden correlated poorly with research investment [30]. Instead, funding decisions may be more significantly informed by political influences, public interest, and transmissibility risk [31]. In oncological research, there is a relative paucity of research output for certain cancers, such as lung, esophageal, and pancreatic cancers, despite their increasing burden and poor prognosis [32]. This reinforces the importance of raising awareness for DCM to facilitate funding for research.

Comparison to Prior Work

In a comparison of these results to a study capturing all DCM papers published in the past 25 years [20], our data suggests that at least 27.46% (621/2261) of DCM research has specific funding.

The location of funding aligns with the location of DCM research output, which has been dominated by the USA, China, Japan, and Canada over the past 20 years [33,34]. This was expected and in keeping with other global health care research investment [10]. It highlights the importance of securing investment to accelerate advances in research outcomes. This is, therefore, now a critical part of ensuring that the aims of the AO Spine RECODE-DCM research priorities are met.

The Global Context

We identified a lack of DCM research funding originating in many low-and-middle-income countries (LMICs), including no funding from the entire African continent. This is common to many health care fields [23]. Yusuf et al [35] identified potential causes of lack of neuroscience research in Africa, among which insufficient funding was one. This is notable for 2 reasons. First, DCM is a global problem [1,36]. Spinal cord disorders such as DCM will increase with a globally aging population, and the prevalence and mortality of spinal disorders, particularly the cervical spine, are increasing in LMICs [37]. Second, from a funding perspective, there is increasing investment in health care research and development in LMICs [38]. Notable examples include organizations such as the NIH [39], NIHR [40], and Canadian Institutes of Health Research (CIHR) [41], as well as philanthropic organizations, such as the Bill and Melinda Gates Foundation. Although much funding is targeted for specific priorities or diseases, much is also investigator-led. The driving force and overall aims behind this increased global investment is multifaceted [42-44] but nevertheless represents an opportunity for DCM.

Maximizing Investment and Future Directions

Despite relatively little investment, DCM research has made significant progress, with the number of published papers increasing year on year and many conducted without research funding [33]. This has contributed to many advances in DCM research [45]. This also highlights the fact that investment and research activity may not always be a linear relationship. For example, in a review of global research activity within esophageal cancer, it was identified that the USA published relatively little compared to their overall research expenditures, while Japan published relatively more [10].

This calls for reduced system inefficiencies to maximize the return of research investment [46], for example, by ensuring research aligns with community needs [47] and is conducted in a robust and transparent manner [48,49] such that its findings can be effectively used. Addressing inefficiencies is the aim of AO Spine RECODE-DCM [50-54]. In addition to setting research priorities, it has agreed to a standardized definition and name for the condition and for a minimum data set to be measured in all research studies [7,52,53,55].

Limitations

There were some limitations to this review. First, information on the funding of DCM research was extracted from the funding metadata in a single research paper database and the acknowledgement sections of published articles therein. Lack of funding information in other common databases prohibited their use. Nonetheless, a database of research funding grants was searched in parallel and the data considered together [56,57]. Although this approach was innovative, systematic, and able to identify a significant amount of data, it is unlikely to have been comprehensive, thus representing a minimum estimate of funding. Research on funding is a largely unexplored area, and the systems in place to document funding sources and tools to support interrogating these systems remain limited and inconsistent.

Second, the funding of published research papers is only a surrogate for research investment [58,59]: it does not quantify the specific amount or role of funding, nor does it account for unpublished research. Moreover, the discrepancies we identified in the results of similar search terms in Dimensions.ai highlights the inconsistency in terminology in this field. However, using papers gives the general overview that our study aims to provide and is a useful and pragmatic way to understand how research is broadly supported.

Third, we categorized funders into institutional, corporate, or charitable groups. However, this may be too simplistic. In reality, organizations are complex and interconnected, with institutions receiving charitable funding [60] and charities receiving corporate backing [28]. Despite this, our study does give a broad understanding over how DCM research is supported by these sectors.

Finally, we note that our review contained studies mainly in English, 1 in German, but none in Chinese. The contribution of Chinese-language papers to global research is significant; Xie and Freeman [61] attribute 37% of global citations in scientific articles to China, compared to our 34% of papers with a Chinese funding origin.

Chinese language papers were not explicitly excluded by our review, but none were identified in the results. We note that Web of Science, our required platform due to its unique ability to extract funding information, searches a relatively small population of Chinese-language papers [61]. There may be a population of DCM papers with a funder originating in China and written in Chinese which has not been included in this study. Although we might have underestimated the total contribution of Chinese funding to DCM research, we still show a substantial contribution. Thus, our study provides as useful, pragmatic, and comprehensive snapshot as is currently feasible.

Conclusions

This is the first review to attempt a global synthesis of the funding landscape of global DCM research, which highlights opportunities for future DCM research. AO Spine has been the leading funder of DCM research, while on a country-specific basis, DCM research has predominantly been funded by the USA, China, and Japan. As this aligns with areas of high research output, it reaffirms the importance of research investment for accelerating advances in DCM. The paucity of investment from major funding organizations and countries with leading research and development expenditure, alongside the increasing investment in global health research, represents opportunities for DCM.

Acknowledgments

MK gratefully acknowledges support by the Cambridge National Institute for Health Research (NIHR) Brain Injury MedTech Cooperative and is funded by a NIHR Clinician Scientist Award (#CS-2015-15-023). BD is supported by a Royal College of Surgeons Research Fellowship and an NIHR Clinical Doctoral Research Fellowship. OM is supported by an Academic Clinical Fellowship. AO Spine RECODE-DCM study was organized and funded by AO Spine through the AO Spine Knowledge Forum Spinal Cord Injury, a focused group of international spinal cord injury experts. AO Spine is a clinical division of the AO Foundation, which is an independent medically guided not-for-profit organization. Study support was provided directly through the AO Spine Research Department.

Data Availability

This article uses data available on Web of Science and Dimensions.ai.

Disclaimer

This report is independent research arising from a Clinician Scientist Award (#CS-2015-15-023) and a Clinical Doctoral Research Fellowship supported by the NIHR. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NHIR, or the Department of Health and Social Care.

Authors' Contributions

HB and JQT contributed equally to conception, data collection and analysis, and writing of the first draft. BG and OM contributed to conception and draft editing. BD contributed to conception, data collection and analysis, and draft editing. MK contributed to draft editing.

Conflicts of Interest

BD and MK are members of the AO Spine, Knowledge Forum Spinal Cord Injury and founders of Myelopathy.org.

References

1. Davies BM, Mowforth OD, Smith EK, Kotter MR. Degenerative cervical myelopathy. BMJ 2018 Feb 22;360:k186 [FREE Full text] [doi: 10.1136/bmj.k186] [Medline: 29472200]
2. Kalsi-Ryan S, Karadimas SK, Fehlings MG. Cervical spondylotic myelopathy: the clinical phenomenon and the current pathobiology of an increasingly prevalent and devastating disorder. Neuroscientist 2013 Aug;19(4):409-421. [doi: 10.1177/1073858412467377] [Medline: 23204243]
3. Kovalova I, Kerkovsky M, Kadanka Z, Kadanka, Nemec M, Jurova B, et al. Prevalence and Imaging Characteristics of Nonmyelopathic and Myelopathic Spondylotic Cervical Cord Compression. Spine (Phila Pa 1976) 2016 Dec 15;41(24):1908-1916. [doi: 10.1097/BRS.0000000000001842] [Medline: 27509189]
4. Smith SS, Stewart ME, Davies BM, Kotter MRN. The Prevalence of Asymptomatic and Symptomatic Spinal Cord Compression on Magnetic Resonance Imaging: A Systematic Review and Meta-analysis. Global Spine J 2021 May;11(4):597-607 [FREE Full text] [doi: 10.1177/2192568220934496] [Medline: 32677521]
5. Fehlings MG, Tetreault LA, Riew KD, Middleton JW, Aarabi B, Arnold PM, et al. A Clinical Practice Guideline for the Management of Patients With Degenerative Cervical Myelopathy: Recommendations for Patients With Mild, Moderate, and Severe Disease and Nonmyelopathic Patients With Evidence of Cord Compression. Global Spine J 2017 Sep;7(3 Suppl):70S-83S [FREE Full text] [doi: 10.1177/2192568217701914] [Medline: 29164035]
6. Oh T, Lafage R, Lafage V, Protopsaltis T, Challier V, Shaffrey C, et al. Comparing Quality of Life in Cervical Spondylotic Myelopathy with Other Chronic Debilitating Diseases Using the Short Form Survey 36-Health Survey. World Neurosurg 2017 Oct;106:699-706. [doi: 10.1016/j.wneu.2016.12.124] [Medline: 28065875]
7. Davies BM, Khan DZ, Mowforth OD, McNair AGK, Gronlund T, Kolias AG, et al. RE-CODE DCM (search Objectives and ommon ata lements for egenerative ervical yelopathy): A Consensus Process to Improve Research Efficiency in DCM, Through Establishment of a Standardized Dataset for Clinical Research and the Definition of the Research Priorities. Global Spine J 2019 May;9(1 Suppl):65S-76S [FREE Full text] [doi: 10.1177/2192568219832855] [Medline: 31157148]
8. Mowforth OD, Starkey ML, Kotter MR, Davies BM. Letter to the Editor. The need for research prioritization in cervical myelopathy. J Neurosurg Spine 2020 Jan 03:1-3. [doi: 10.3171/2019.11.SPINE191300] [Medline: 31899877]
9. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
10. Klingelhöfer D, Zhu Y, Braun M, Brüggmann D, Schöffel N, Groneberg DA. A world map of esophagus cancer research: a critical accounting. J Transl Med 2019 May 10;17(1):150 [FREE Full text] [doi: 10.1186/s12967-019-1902-7] [Medline: 31077194]

11. Begum M, Lewison G, Sommariva S, Ciani O, Tarricone R, Sullivan R. European diabetes research and its funding, 2002-2013. Diabet Med 2017 Oct;34(10):1354-1360. [doi: 10.1111/dme.13411] [Medline: 28636762]
12. Nouri A, Tetreault L, Singh A, Karadimas SK, Fehlings MG. Degenerative Cervical Myelopathy: Epidemiology, Genetics, and Pathogenesis. Spine (Phila Pa 1976) 2015 Jun 15;40(12):E675-E693. [doi: 10.1097/BRS.0000000000000913] [Medline: 25839387]
13. Faster systematic reviews. Rayyan QCRI. URL: https://rayyan.qcri.org/welcome/ [accessed 2020-08-15]
14. Dimensions.ai. URL: https://www.dimensions.ai/ [accessed 2020-11-29]
15. Davies BM, Goh S, Yi K, Kuhn I, Kotter MRN. Development and validation of a MEDLINE search filter/hedge for degenerative cervical myelopathy. BMC Med Res Methodol 2018 Jul 6;18(1):597-607. [doi: 10.1186/s12874-018-0529-3]
16. Wu ZK, Zhao Q, Tian J, Qian Y, Zhou Y, Yang F, et al. Anterior versus posterior approach for multilevel cervical spondylotic myelopathy. Cochrane Database Syst Rev 2016 Sep 28 [FREE Full text] [doi: 10.1002/14651858.CD012365]
17. Numpy. URL: https://numpy.org/ [accessed 2020-11-29]
18. Panda Data Analysis Library. URL: https://pandas.pydata.org/ [accessed 2020-11-29]
19. Choropleth maps in Python. Plotly. URL: https://plotly.com/python/choropleth-maps/ [accessed 2020-11-29]
20. Grodzinski B, Bestwick H, Bhatti F, Durham R, Khan M, Partha Sarathi CI, et al. Research activity amongst DCM research priorities. Acta Neurochir 2021 Feb 24;163(6):1561-1568. [doi: 10.1007/s00701-021-04767-6]
21. Viergever RF, Hendriks TCC. The 10 largest public and philanthropic funders of health research in the world: what they fund and how they distribute their funds. Health Res Policy Sys 2016 Feb 18;14(1):1-15. [doi: 10.1186/s12961-015-0074-z]
22. Government/agency RTWB. Research and development expenditure (% of GDP) | Data. The World Bank. URL: https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?end=2018&most_recent_value_desc=true&start=1996&view=chart [accessed 2020-11-29]
23. Røttingen J, Regmi S, Eide M, Young AJ, Viergever RF, Årdal C, et al. Mapping of available health research and development data: what's there, what's missing, and what role is there for a global observatory? The Lancet 2013 Oct;382(9900):1286-1307. [doi: 10.1016/s0140-6736(13)61046-6]
24. Johnson & Johnson Medical Devices. URL: https://www.jnjmedicaldevices.com/en-US/companies/depuy-synthes [accessed 2021-01-13]
25. Research. MS Society. URL: https://www.mssociety.org.uk/research [accessed 2021-01-13]
26. MND Association. URL: https://www.mndassociation.org/research/ [accessed 2021-01-13]
27. News PAOMRC. Charities' funding contributes to UK medical research excellence. Association of Medical Research Charities. URL: https://www.amrc.org.uk/news/charities-funding-contributes-to-uk-medical-research-excellence [accessed 2021-01-13]
28. Government/agency R. UK Parliament. House of Commons - Health - Minutes of Evidence. URL: https://publications.parliament.uk/pa/cm200405/cmselect/cmhealth/42/4112502.htm [accessed 2021-01-13]
29. Research projects. Myelopathy.org. URL: https://myelopathy.org/myelopathy-info/research-projects/ [accessed 2021-01-13]
30. Gillum LA, Gouveia C, Dorsey ER, Pletcher M, Mathers CD, McCulloch CE, et al. NIH Disease Funding Levels and Burden of Disease. PLoS ONE 2011 Feb 24;6(2):e16837. [doi: 10.1371/journal.pone.0016837]
31. Institute OM. Scientific Opportunities and Public Needs: Improving Priority Setting and Public Input at the National Institutes of Health. Washington, DC: The National Academies Press; 1998:978.
32. Begum M, Lewison G, Lawler M, Sullivan R. Mapping the European cancer research landscape: An evidence base for national and Pan-European research and funding. Eur J Cancer 2018 Sep;100:75-84 [FREE Full text] [doi: 10.1016/j.ejca.2018.04.017] [Medline: 30014883]
33. Mowforth OD, Davies BM, Goh S, O'Neill CP, Kotter MRN. Research Inefficiency in Degenerative Cervical Myelopathy: Findings of a Systematic Review on Research Activity Over the Past 20 Years. Global Spine J 2020 Jun;10(4):476-485 [FREE Full text] [doi: 10.1177/2192568219847439] [Medline: 32435569]
34. Yin M, Xu C, Ma J, Ye J, Mo W. A Bibliometric Analysis and Visualization of Current Research Trends in the Treatment of Cervical Spondylotic Myelopathy. Global Spine Journal 2020 Sep 01;11(6):988-998. [doi: 10.1177/2192568220948832]
35. Yusuf S, Baden T, Prieto-Godino LL. Bridging the Gap: establishing the necessary infrastructure and knowledge for teaching and research in neuroscience in Africa. Metab Brain Dis 2013 Oct 30;29(2):217-220. [doi: 10.1007/s11011-013-9443-x]
36. Beaglehole R, Bonita R. What is global health? Global Health Action 2010 Apr 06;3(1):5142-5143. [doi: 10.3402/gha.v3i0.5142]
37. Waheed MA, Hasan S, Tan LA, Bosco A, Reinas R, ter Wengel PV, et al. Cervical spine pathology and treatment: a global overview. J Spine Surg 2020 Mar;6(1):340-350. [doi: 10.21037/jss.2020.01.12]
38. Conalogue DM, Kinn S, Mulligan J, McNeil M. International consultation on long-term global health research priorities, research capacity and research uptake in developing countries. Health Res Policy Sys 2017 Mar 21;15(1):1-6. [doi: 10.1186/s12961-017-0181-0]
39. Government/agency RKFF. Kaiser Family Foundation: the US Government and Global Health. Kaiser Family Foundation. 2019. URL: https://www.kff.org/global-health-policy/fact-sheet/the-u-s-government-and-global-health/ [accessed 2021-01-13]
40. Global health research. NIHR. URL: https://www.nihr.ac.uk/explore-nihr/funding-programmes/global-health.htm [accessed 2021-01-13]

41. CIHR: Global Health Research. CIHR. URL: https://cihr-irsc.gc.ca/e/31562.html [accessed 2021-01-13]
42. Abimbola S, Negin J, Martiniuk A. Charity begins at home in global health research funding. The Lancet Global Health 2017 Jan;5(1):e25-e27. [doi: 10.1016/s2214-109x(16)30302-3]
43. Annual Report of the Chief Medical Officer. Government/agency report: APS Group for the Department of Health and Social Care. 2019. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/832948/annual-report-of-the-chief-medical-officer-2019.pdf [accessed 2021-01-13]
44. Frenk J, Gómez-Dantés O, Moon S. From sovereignty to solidarity: a renewed concept of global health for an era of complex interdependence. The Lancet 2014 Jan;383(9911):94-97. [doi: 10.1016/s0140-6736(13)62561-1]
45. Badhiwala JH, Ahuja CS, Akbar MA, Witiw CD, Nassiri F, Furlan JC, et al. Degenerative cervical myelopathy — update and future directions. Nat Rev Neurol 2020 Jan 23;16(2):108-124. [doi: 10.1038/s41582-019-0303-0]
46. Research: increasing value, reducing waste. The Lancet. URL: https://www.thelancet.com/series/research [accessed 2021-01-13]
47. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. Lancet 2014 Jan 11;383(9912):156-165. [doi: 10.1016/S0140-6736(13)62229-1] [Medline: 24411644]
48. Chan A, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. The Lancet 2014 Jan;383(9913):257-266. [doi: 10.1016/s0140-6736(13)62296-5]
49. Ioannidis JPA, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet 2014 Jan 11;383(9912):166-175 [FREE Full text] [doi: 10.1016/S0140-6736(13)62227-8] [Medline: 24411645]
50. Davies BM, McHugh M, Elgheriani A, Kolias AG, Tetreault LA, Hutchinson PJA, et al. Reported Outcome Measures in Degenerative Cervical Myelopathy: A Systematic Review. PLoS One 2016;11(8):e0157263 [FREE Full text] [doi: 10.1371/journal.pone.0157263] [Medline: 27482710]
51. Davies BM, McHugh M, Elgheriani A, Kolias AG, Tetreault L, Hutchinson PJA, et al. The reporting of study and population characteristics in degenerative cervical myelopathy: A systematic review. PLoS One 2017;12(3):e0172564 [FREE Full text] [doi: 10.1371/journal.pone.0172564] [Medline: 28249017]
52. KHAN DZ, DAVIES BM, KOTTER MR. Spinal Research — A Field in Need of Standardization. J Rheumatol 2020 Feb 15;47(4):633-634. [doi: 10.3899/jrheum.191225]
53. Ahmed S, Hirayama Y, Khan D, Mowforth O, Kotter MR, Davies B. Letter to the Editor: The Need for Standardization of Terminology in Spinal Research. Spine (Phila Pa 1976) 2020 Oct 01;45(19):E1286-E1287. [doi: 10.1097/BRS.0000000000003634] [Medline: 32925823]
54. Grodzinski B, Mowforth O, Tetreault LA, Davies BM. Improving the Quality of Systematic Reviews in Spinal Surgery Requires Community-Wide Engagement and Pragmatism. Global Spine Journal 2020 Nov 04;10(8):1078-1079. [doi: 10.1177/2192568220952758]
55. Khan DZ, Khan MS, Kotter MR, Davies BM. Tackling Research Inefficiency in Degenerative Cervical Myelopathy: Illustrative Review. JMIR Res Protoc 2020 Jun 11;9(6):e15922. [doi: 10.2196/15922]
56. Herzog C, Hook D, Konkiel S. Dimensions: Bringing down barriers between scientometricians and data. Quantitative Science Studies 2020 Feb;1(1):387-395. [doi: 10.1162/qss_a_00020]
57. Hook DW, Porter SJ, Herzog C. Dimensions: Building Context for Search and Evaluation. Front. Res. Metr. Anal 2018 Aug 23;3. [doi: 10.3389/frma.2018.00023]
58. Payne A, Siow A. Does Federal Research Funding Increase University Research Output? The B. E. Journal of Economic Analysis & Policy 2003;3(1):1-24. [doi: 10.2202/1538-0637.1018]
59. Sandström U. Research quality and diversity of funding: A model for relating research money to output of research. Scientometrics 2008 Nov 13;79(2):341-349. [doi: 10.1007/s11192-009-0422-2]
60. Funding, finance, and operations. Universities UK. URL: https://www.universitiesuk.ac.uk/policy-and-analysis/Pages/funding.aspx [accessed 2021-01-13]
61. Xie Q, Freeman RB. Bigger Than You Thought: China's Contribution to Scientific Publications and Its Impact on the Global Economy. China & World Economy 2019 Jan 15;27(1):1-27. [doi: 10.1111/cwe.12265]

Abbreviations

CIHR: Canadian Institutes of Health Research
DCM: degenerative cervical myelopathy
LMIC: low-and-middle-income country
NIH: National Institutes of Health
NIHR: National Institute for Health Research
RECODE-DCM: Research Objectives and Common Data Elements for Degenerative Cervical Myelopathy
SF-36: 36-Item Short Form Survey

Edited by T Leung; submitted 05.01.22; peer-reviewed by C Visper, J Hartvigsen; comments to author 29.03.22; revised version received 29.04.22; accepted 17.05.22; published 30.06.22.

Please cite as:
Bestwick H, Teh JQ, Mowforth O, Grodzinski B, Kotter M, Davies B
Existing Funding Sources in Degenerative Cervical Myelopathy Research: Scoping Review
Interact J Med Res 2022;11(1):e36194
URL: https://www.i-jmr.org/2022/1/e36194
doi:10.2196/36194
PMID:35771617

Original Paper

Machine Learning Approaches for Predicting Difficult Airway and First-Pass Success in the Emergency Department: Multicenter Prospective Observational Study

Syunsuke Yamanaka[1], MD, PhD; Tadahiro Goto[2], MD, MPH, PhD; Koji Morikawa[3], PhD; Hiroko Watase[4], MD, MPH; Hiroshi Okamoto[5], MD, MPH; Yusuke Hagiwara[6], MD, MPH; Kohei Hasegawa[7], MD, MPH

[1]Department of Emergency Medicine & General Internal Medicine, The University of Fukui, Fukui, Japan

[2]Department of Clinical Epidemiology & Health Economics, School of Public Health, The University of Tokyo, Tokyo, Japan

[3]Connect Inc, Tokyo, Japan

[4]Department of Surgery, University of Washington, Seattle, WA, United States

[5]Department of Intensive Care, St. Luke's International Hospital, Tokyo, Japan

[6]Department of Pediatric Emergency and Critical Care Medicine, Tokyo Metropolitan Children's Medical Center, Tokyo, Japan

[7]Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, United States

Corresponding Author:
Tadahiro Goto, MD, MPH, PhD
Department of Clinical Epidemiology & Health Economics
School of Public Health
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo, 113-8655
Japan
Phone: 81 3 5841 1887
Email: tag695@mail.harvard.edu

Abstract

Background: There is still room for improvement in the modified LEMON (look, evaluate, Mallampati, obstruction, neck mobility) criteria for difficult airway prediction and no prediction tool for first-pass success in the emergency department (ED).

Objective: We applied modern machine learning approaches to predict difficult airways and first-pass success.

Methods: In a multicenter prospective study that enrolled consecutive patients who underwent tracheal intubation in 13 EDs, we developed 7 machine learning models (eg, random forest model) using routinely collected data (eg, demographics, initial airway assessment). The outcomes were difficult airway and first-pass success. Model performance was evaluated using c-statistics, calibration slopes, and association measures (eg, sensitivity) in the test set (randomly selected 20% of the data). Their performance was compared with the modified LEMON criteria for difficult airway success and a logistic regression model for first-pass success.

Results: Of 10,741 patients who underwent intubation, 543 patients (5.1%) had a difficult airway, and 7690 patients (71.6%) had first-pass success. In predicting a difficult airway, machine learning models—except for k-point nearest neighbor and multilayer perceptron—had higher discrimination ability than the modified LEMON criteria (all, $P \leq .001$). For example, the ensemble method had the highest c-statistic (0.74 vs 0.62 with the modified LEMON criteria; $P < .001$). Machine learning models—except k-point nearest neighbor and random forest models—had higher discrimination ability for first-pass success. In particular, the ensemble model had the highest c-statistic (0.81 vs 0.76 with the reference regression; $P < .001$).

Conclusions: Machine learning models demonstrated greater ability for predicting difficult airway and first-pass success in the ED.

(Interact J Med Res 2022;11(1):e28366) doi:10.2196/28366

KEYWORDS

intubation; machine learning; difficult airway; first-pass success

Introduction

In the emergency department (ED), achieving successful tracheal intubation at the initial attempt (ie, first-pass success) is essential [1]. The literature has shown that repeated intubation attempts are associated with a higher rate of adverse events [2-4]. However, recent studies have also reported first-pass success rates of 74%-84% in the ED [5,6], suggesting that there are still occasions where repeated intubation attempts are required. To improve ED airway management, the development of effective risk stratification and prediction tools is instrumental.

A widely used prediction tool for difficult airway is the modified LEMON (look, evaluate, Mallampati, obstruction, neck mobility) criteria [7], which has been validated [8]. Although the criteria have good prediction ability (eg, sensitivity 86%, specificity 48% for direct laryngoscope) [8], there remains room for improvement. Besides, no prediction tool accurately predicts first-pass success (or failure) in the ED. The recent advent of machine learning approaches has enabled clinicians and researchers to accurately predict various diseases and conditions, such as sepsis [9], acute asthma [10], and ED triage [11,12]. Compared with conventional prediction tools and regression approaches, modern machine learning approaches have several advantages, such as incorporating high-order, nonlinear interactions between predictors and mitigating overfitting [13]. Despite the clinical and research importance, no study has yet applied modern machine learning approaches to predict a difficult airway in advance of preparing for airway management or to predict first-pass success once the intubation strategy has been determined in the ED.

To address this significant knowledge gap in the literature, using data from a prospective, multicenter study of ED airway management, we aimed to develop machine learning models that accurately predict difficult airway and first-pass success and to compare their performance with conventional approaches.

Methods

Study Design, Setting, and Participants

This study analyzes data from a multicenter, prospective study of emergency airway management—the second Japanese Emergency Airway Network (JEAN-2) study. The details of the study design, setting, participants, methods of data measurement, and definitions of variables have been reported elsewhere [14]. In brief, the JEAN-2 study is a consortium of 13 academic and community EDs, including 10 level I and 3 level II equivalent trauma centers. These EDs are located in different geographic regions across Japan. The median ED census is 29,000 patients per year (range of 16,000 to 67,000 annual visits). These ED are affiliated with an emergency medicine residency training program. Attending physicians or resident physicians who are under the supervision of the attending physician perform intubations. In this observational study, patients were managed at the discretion of treating physicians. The institutional review board at each participating center approved the waiver of informed consent before data collection. This study used data from consecutive (both children and adults) patients who underwent ED management at one of the participating EDs from January 1, 2010 through December 31, 2018. Patients who underwent surgical intubations at the first attempt were excluded.

Outcomes

The outcomes of interests were difficult airway and first-pass success. According to the American Society of Anesthesiologists (ASA) guidelines, a difficult airway was defined as multiple intubation attempts by emergency physicians or anesthesiologists according to the ASA guidelines [15]. First-pass success was defined as intubation success at the initial attempt of each encounter [16]. Intubation success was defined as the proper placement of a tracheal tube through the vocal cord, confirmed by the use of quotative or end-tidal CO_2 monitoring [17]. An intubation attempt was defined as a single insertion of the laryngoscope past the teeth [18].

Predictors of Machine Learning Models

To develop machine learning models for the difficult airway outcome, we used the following variables that are *routinely* obtained in advance of the actual intubation attempt: patient demographics (age, sex, estimated height and body weight, BMI), components of the modified LEMON criteria, pre-intubation vital signs (pulse rate, systolic blood pressure, respiratory rate, oxygen saturation), and Glasgow coma scale. To develop models that predict the first-pass success outcome (once the intubation strategy has been determined), we used all available intubation-related information—in addition to the aforementioned predictors—such as type of day (weekend/weekday), medications, intubation methods, intubation devices, intubator's post-graduate year, and intubator's specialty.

Statistical Analysis

Summary statistics were used to describe the characteristics of patients and airway management. After performing imputations for missing continuous variables (most predictors had <10% missingness; Multimedia Appendix 1) using random forest [19], we conducted predictor preprocessing, such as one-hot encoding (ie, creation of dummy variables), normalization, and standardization. The nonlinear predictors included in the developed models were age, body weight, height, BMI, and pre-intubation vital signs. In the training set (80% random sample), for each outcome, we developed 7 machine learning models: (1) logistic regression model with elastic-net (penalized logistic regression) [20], (2) random forest [21], (3) gradient boosting decision tree [22], (4) multilayer perceptron [23], (5) k-point nearest neighbor [24], (6) XGBoost [25], and (7) ensemble model (ridge regression and the random forest with an equal weight) [26]. For the difficult airway outcome, the modified LEMON criteria model was used as the reference model. For the first-pass success outcome, a (nonpenalized) logistic regression model was used as the reference model. We performed stratified 5-fold cross-validation to determine the optimal hyperparameters with the highest c-statistic (ie, the area under the receiver operating characteristic [ROC] curve).

In the test set (the remaining 20% of the random sample), we measured the performance of reference and machine learning models. We estimated the c-statistic of each model and examined the following association measures: sensitivity,

specificity, positive and negative predictive values, and positive and negative likelihood ratios. The c-statistic is the probability that, given 2 individuals (one who experiences the outcome of interest and the other who does not), the model estimates a higher probability for the first patient than for the second [27]. We determined the threshold of perspective prediction (cut-off) results based on the ROC curve from the Youden method [28]. For the model with the highest c-statistic among the 7 machine learning models, we computed the variable importance—how strongly each of the predictors improved the c-statistic. We also examined calibration plots of the best-performing machine learning model for each of the outcomes. Data were analyzed using python (version 3.7.3) and R (version 3.6.2). Two-sided *P* values <.05 were considered statistically significant.

Results

Patient Characteristics

During the 108-month study period, the JEAN-2 study recorded data for 10,816 patients (capture rate, 96%) who underwent emergency airway management at one of the 13 participating EDs. Of these, 75 patients who underwent surgical intubation at their first attempt were excluded; the remaining 10,741 patients comprised the analytic cohort. The patient characteristics, details of airway management, and intubation outcomes are shown in Table 1. The median age was 71 (IQR 56 -81) years, 2.8% (304/10,741) were children, and 38.0% (4079/10,741) were female. Overall, 5.1% (543/10,741) of patients had a difficult airway outcome, while 71.6% (7690/10,741) had first-pass success. An aborted intubation attempt occurred for 39 patients.

Table 1. Patient characteristics, airway management, and outcomes in 10,741 patients who underwent tracheal intubation in the emergency department.

Variables	Results
Age (years), median (IQR)	71 (56-81)
Children (18 years), n (%)	304 (2.8)
Female gender, n (%)	4079 (38.0)
Estimated height (cm), median (IQR)	160 (153-170)
Estimated body weight (kg), median (IQR)	60 (50-67)
BMI (kg/m^2), median (IQR)	22.0 (19.5-24.3)
Primary indication, n (%)	
Medical cardiac arrest	3785 (35.2)
Traumatic cardiac arrest	438 (4.1)
Medical indication	5440 (50.6)
Airway problem (eg, obstruction)	289 (2.7)
Breathing problem (eg, respiratory failure)	1673 (15.6)
Circulation problem (eg, shock)	1080 (10.1)
Altered mental status	2036 (19.0)
Others	360 (3.4)
Traumatic indication	1080 (10.1)
Modified LEMON[a] criteria, n (%)	
Look externally	583 (5.0)
3-3-(2) rule	3620 (33.7)
Obstruction	774 (7.2)
Neck mobility	1101 (10.3)
Any criterion met in the modified LEMON criteria	4709 (43.8)
Intubation outcomes, n (%)	
Difficult airway	543 (5.1)
First-pass success	7690 (71.6)

[a]LEMON: look, evaluate, Mallampati, obstruction, neck mobility.

Prediction Performance for Difficult Airway Outcomes

Table 2 summarizes the performance of the modified LEMON criteria (reference) and 7 machine learning models when predicting a difficult airway outcome in the ED. Compared with the modified LEMON criteria, the discrimination ability of machine learning models—except for the k-point nearest-neighbor model and multilayer perceptron model—were significantly greater ($P \leq .001$). Among the 7 machine learning models, the ensemble model had the highest c-statistic (0.74, 95% CI 0.67-0.79; Figure 1A), with a sensitivity of 0.67 (95% CI 0.65-0.69), specificity of 0.70 (95% CI 0.68-0.72), positive predictive value of 0.09 (95% CI 0.08-0.11), and negative predictive value of 0.98 (95% CI 0.97-0.98). Compared with the modified LEMON criteria, which had a specificity of 0.57 (95% CI 0.56-0.58), all machine learning models had higher specificity, with the multilayer perceptron model achieving a specificity of 0.92 (95% CI 0.90-0.93). The calibration plot (Figure 2A)—which indicates how far the predicted risk from the ensemble model deviated from the actual risk—showed that the ensemble model overestimated the risk of the outcome, while there was a positive relationship between the predicted and actual risks, largely due to the class imbalance (ie, difficult airway outcome occurred only in 5% of the sample).

Table 2. Performance of 7 machine learning models and modified LEMON (look, evaluate, Mallampati, obstruction, neck mobility) criteria when predicting difficult airway outcome in the emergency department.

Models	C-statistic[a] (95% CI)	*P* value	Sensitivity (95% CI)	Specificity (95% CI)	PPV[b] (95% CI)	NPV[c] (95% CI)	PLR[d] (95% CI)	NLR[e] (95% CI)
Modified LEMON criteria (reference)	0.62 (0.60-0.64)	Reference[f]	0.67 (0.66-0.68)	0.57 (0.56-0.58)	0.08 (0.07-0.08)	0.97 (0.97-0.97)	1.57 (1.48-1.68)	0.57 (0.54-0.61)
Penalized logistic regression	0.73 (0.68-0.79)	<.001	0.66 (0.64-0.68)	0.68 (0.66-0.70)	0.09 (0.08-0.10)	0.98 (0.97-0.98)	2.05 (1.75-2.40)	0.51 (0.43-0.59)
Random forest	0.72 (0.67-0.77)	<.001	0.58 (0.56-0.60)	0.74 (0.72-0.75)	0.09 (0.08-0.11)	0.97 (0.97-0.98)	3.84 (2.50-5.90)	0.84 (0.55-1.29)
Gradient boost	0.72 (0.66-0.77)	.001	0.77 (0.75-0.79)	0.58 (0.56-0.60)	0.08 (0.07-0.09)	0.98 (0.98-0.99)	1.84 (1.63-2.08)	0.39 (0.35-0.44)
Multilayer perceptron	0.57 (0.50-0.63)	.14	0.19 (0.17-0.20)	0.92 (0.90-0.93)	0.09 (0.08-0.11)	0.96 (0.95-0.97)	2.24 (1.44-3.48)	0.89 (0.57-1.38)
K-point nearest neighbor	0.54 (0.49-0.61)	.02	0.39 (0.36-0.41)	0.70 (0.68-0.72)	0.06 (0.05-0.07)	0.96 (0.95-0.97)	1.30 (1.00-1.68)	0.87 (0.67-1.14)
XGBoost	0.72 (0.67-0.77)	<.001	0.69 (0.67-0.71)	0.60 (0.58-0.62)	0.07 (0.06-0.09)	0.98 (0.97-0.98)	1.70 (1.47-1.97)	0.52 (0.45-0.61)
Ensemble model[g]	0.74 (0.67-0.79)	<.001	0.67 (0.65-0.69)	0.70 (0.68-0.72)	0.09 (0.08-0.11)	0.98 (0.97-0.98)	2.21 (1.89-2.58)	0.48 (0.41-0.56)

[a]C-statistic in the modified LEMON was evaluated using 95% CIs.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

[d]PLR: positive likelihood ratio.

[e]NLR: negative likelihood ratio.

[f]Comparison of the area under the curve of the reference model (modified LEMON) with that of each machine learning model using the DeLong test.

[g]Ensemble prediction model using these machine learning models (that combined ≥2 models).

Figure 1. Discrimination ability of the ensemble model and logistic regression (reference) model in predicting intubation outcomes, including (A) difficult airway outcomes and (B) first-pass success outcomes. mLEMON: modified look, evaluate, Mallampati, obstruction, neck mobility model.

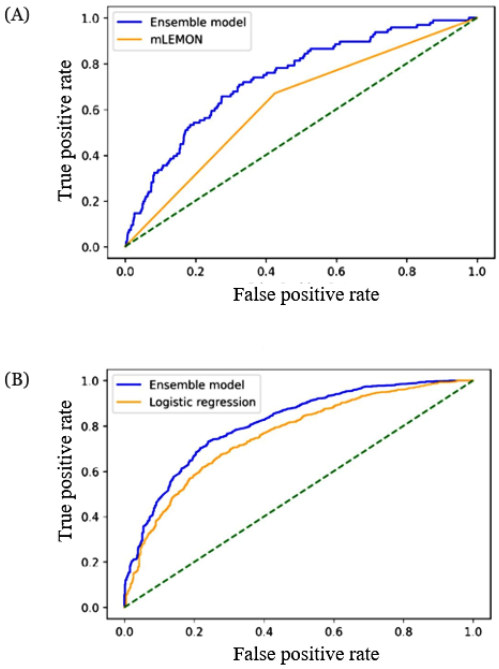
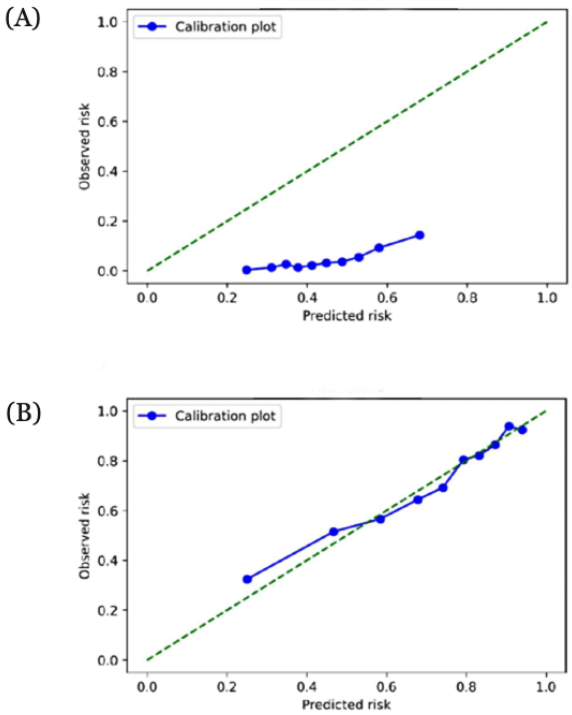


Figure 2. Calibration plots of ensemble models in predicting intubation outcomes, including (A) difficult airway outcomes and (B) first-pass success outcomes.



Prediction Performance for First-Pass Success Outcomes

Table 3 summarizes the performance of the reference model and 7 machine learning models when predicting the first-pass success outcome in the ED. Compared with the reference model, the discrimination ability of machine learning models—except for the random forest and k-point nearest neighbor models—was significantly higher (all $P<.05$). Among the 7 machine learning models, the ensemble model had the highest c-statistic (0.81, 95% CI 0.79-0.83; Figure 1B). Compared with the reference model, the ensemble model had a higher sensitivity (0.79, 95% CI 0.77-0.81) and specificity (0.67, 95% CI 0.65-0.69), with a PPV of 0.85 (95% CI 0.84-0.87) and NPV of 0.57 (95% CI 0.55-0.59). Compared with the reference model, which had a specificity of 0.36 (95% CI 0.34-0.38), most machine learning models had higher specificity, with the random forest model achieving a specificity of 0.70 (95% CI 0.68-0.72). In the calibration plot of the ensemble model (Figure 2B), the

model-predicted probability was well-matched with the observed probabilities.

Table 3. Performance of 7 machine learning models and reference model when predicting first-pass success outcome in the emergency department.

Models	C statistic (95% CI)	*P* value	Sensitivity (95% CI)	Specificity (95% CI)	PPV[a] (95% CI)	NPV[b] (95% CI)	PLR[c] (95% CI)	NLR[d] (95% CI)
Logistic regression (reference)	0.76 (0.74-0.78)	(Reference)[e]	0.91 (0.89-0.92)	0.36 (0.34-0.38)	0.78 (0.76-0.79)	0.61 (0.59-0.63)	1.42 (1.33-1.51)	0.26 (0.24-0.27)
Penalized logistic regression	0.81 (0.79-0.83)	.001	0.79 (0.77-0.80)	0.70 (0.68-0.72)	0.86 (0.85-0.88)	0.57 (0.55-0.59)	2.59 (2.29-2.92)	0.31 (0.27-0.35)
Random forest	0.78 (0.76-0.81)	.12	0.78 (0.76-0.79)	0.64 (0.62-0.66)	0.84 (0.83-0.86)	0.54 (0.52-0.56)	2.16 (1.94-1.2.41)	0.35 (0.31-0.39)
Gradient boost	0.80 (0.78-0.82)	.005	0.92 (0.91-0.94)	0.40 (0.38-0.43)	0.79 (0.77-0.81)	0.69 (0.67-0.71)	1.55 (1.45-1.66)	0.19 (0.17-0.20)
Multilayer perceptron	0.81 (0.79-0.83)	.002	0.92 (0.91-0.93)	0.44 (0.42-0.46)	0.80 (0.78-0.82)	0.69 (0.67-0.71)	1.64 (1.53-1.76)	0.18 (0.17-0.19)
K-point nearest neighbor	0.75 (0.73-0.77)	.60	0.98 (0.97-0.98)	0.18 (0.16-0.20)	0.74 (0.73-0.76)	0.78 (0.76-0.80)	1.19 (1.15-1.24)	0.12 (0.11-0.12)
XGBoost	0.81 (0.79-0.83)	<.001	0.94 (0.93-0.95)	0.38 (0.36-0.40)	0.79 (0.77-0.81)	0.73 (0.71-0.75)	1.53 (1.43-1.62)	0.15 (0.14-0.16)
Ensemble model[f]	0.81 (0.79-0.83)	<.001	0.79 (0.77-0.81)	0.67 (0.65-0.69)	0.85 (0.84-0.87)	0.57 (0.55-0.59)	2.39 (2.13-2.67)	0.31 (0.28-0.35)

[a]PPV: positive predictive value.

[b]NPV: negative predictive value.

[c]PLR: positive likelihood ratio.

[d]NLR: negative likelihood ratio.

[e]Comparison of the area under the curve of the reference model with that of each machine learning model using the DeLong test.

[f]Ensemble prediction model using these machine learning models (that combined ≥2 models).

Variable Importance

Table 4 shows the variable importance of the best performance model (the ensemble model) for each outcome. For the difficult airway prediction, the most contributing predictor was age, followed by any criterion met in the modified LEMON criteria and hyoid mental distance ≥3 fingers. For the first-pass success prediction, the best contributing predictor was the use of laryngeal pressure, followed by lifting force required for laryngeal deployment and Cormack grade of 3.

Table 4. Importance of each predictor of ensemble model when predicting difficult airway and first-pass success outcomes.

Predictors	Δ c-statistics[a]
Predictors for difficult airway outcome	
Age	0.093
Any modified LEMON[b] criterion met	0.093
Hyoid mental distance ≥3 fingers	0.091
Interincisor distance of 3 fingers	0.084
BMI	0.080
Body weight	0.073
>80 years old	0.070
Hyoid mental distance of 2 fingers	0.053
Airway obstruction	0.049
Neck mobility	0.048
Predictors for first-pass success outcome	
Use of laryngeal pressure	0.118
Lifting force required for laryngeal deployment	0.108
Cormack grade of 3	0.099
Any modified LEMON criterion met	0.094
Cormack grade of 1	0.094
Intubator's post-graduation year of 1 or 2	0.090
Neuromuscular blocking agent (rocuronium)	0.077
Rapid sequence intubation	0.076
Video laryngoscope (C-MAC)	0.075
Video Cormack grade of 1	0.074
Interincisor distance ≥3 fingers	0.079

[a]The variable importance of a predictor is agnostic regarding the direction of the association.

[b]LEMON: look, evaluate, Mallampati, obstruction, neck mobility.

Discussion

Principal Findings

In this analysis of multicenter prospective data from 10,741 ED patients, we applied modern machine learning models to predict intubation-related outcomes in the ED. Specifically, compared with conventional approaches (ie, modified LEMON criteria and nonpenalized logistic regression model), most machine learning models demonstrated superior discrimination performance when predicting both difficult airway and first-pass success outcomes. Additionally, these machine learning models also achieved higher specificity when predicting these 2 outcomes. To the best of our knowledge, this is the first study that has investigated the performance of modern machine learning models when predicting clinically important intubation outcomes in the ED setting.

Consistent with our findings, the following has been reported as predictors for first-pass success in the ED: patient characteristics (eg, restricted mouth opening, restricted neck extension, and swollen tongue), high Cormack grade, intubators' characteristics (eg, clinical experience and working department), the use of rapid-sequence intubation, and the use of video laryngoscope at the first attempt [6,29-31].

The importance of accurate prediction for difficult airways has been emphasized in ED airway management [8]. Although the modified LEMON criteria (and the LEMON criteria) have been validated as an indicator for difficult airways, their prediction ability is suboptimal for clinical use [7,8]. In the operating room setting, a couple of studies have reported a potential benefit of machine learning models for predicting difficult airways [32,33]. For example, in a single-center study of 80 patients, a deep learning approach using data from the patients' facial images had high discrimination ability for difficult airways—defined as multiple attempts by an intubator with at least 12 months of anesthesia experience, grade 3 or 4 laryngoscopic view, need for a second intubator, or nonelective use of an alternative airway device [32]. Our multicenter study—with a sample size that is many times larger than the prior studies on this topic—builds on these earlier reports and extends them by demonstrating that modern machine learning models outperform conventional approaches for predicting intubation outcomes in the ED.

The observed improvement in prediction ability by machine learning approaches may be explained by several reasons. First, the machine learning approaches account for high-order interactions between predictors and nonlinear relationships with an outcome, which traditional modeling approaches cannot address [34]. Second, the modified LEMON criteria may be too parsimonious (ie, the use of a limited number of predictors), while the applied machine learning models could use a larger number of predictors. Third, the modern machine learning approaches enabled us to minimize overfitting, such as lasso and ridge penalizations (ie, elastic net model and cross-validation). In addition to these strengths, modern machine learning models also are scalable for further improvement by integration with recently developed techniques such as image analysis of patients' faces and necks [32,35].

Although the machine learning models achieved a more significant predictive ability, their performance remained imperfect. This may be explained, at least partially, by the limited set of predictors (eg, lack of detailed information on the intubation competency and experience of each intubator) and data measurement errors. Additionally, one may surmise that the modified LEMON criteria are simpler and easier to use in the ED. Despite the known trade-off between parsimonious models and more complex models with a larger number of predictors, the use of modern machine learning models has advantages in the era of health information technology, including automated data entry through voice recognition, natural language processing, continuous sophistication of models through sequential extractions of electronic health records, and reinforcement learning [36,37]. Our findings and the recent advent of machine learning approaches collectively support cautious optimism that machine learning may enhance the clinician's ability—as assistive technology—to predict patient outcomes in the ED. The resulting accurate prediction of intubation outcomes has several important implications in airway practice in the ED. For example, early identification of difficult airways should help ED providers develop individualized and optimal management strategies and prepare for rescue airways [14,38]. Besides, the accurate estimation of the probability of first-pass success given the conditions (eg, the airway management strategies and intubator to be used) would not only increase the opportunity for clinical training (eg, which patient can be safely intubated by the intubator) but also improve patient safety.

To implement our developed machine learning models, a web-based application or integrated emergency department information system is needed. The rapid development of health information technology (eg, web-based artificial intelligence application with the model) enables us to implement the developed model into the real clinical setting. Furthermore, the current models can be used not only for practice but as an educational tool. For example, in simulation-based intubation training, supervisors can evaluate the trainee's intubation strategy by indicating the actual probability of difficult airway and first-pass success.

Limitations

Several potential limitations of this study should be noted. First, our data may be subject to self-reporting and measurement bias (eg, underreporting difficult airways). However, the study was conducted by investigators using a standardized protocol [6], which led to the high capture rate (96%) and low proportion of missingness in the predictors and outcomes (Multimedia Appendix 1). Second, we did not have detailed information on the procedural competency of each intubator, as this factor is also challenging to define and measure in real-world settings. To address this issue, we used years of experience and specialty, which are readily available in most ED settings, as a proxy for the competency. Third, machine learning models have a common limitation in the interpretability of models. Fourth, because of the small samples of children (2.8%), our model may not have optimal prediction ability in pediatric populations. Finally, our models may not be generalizable to other practice settings, although the study sample consisted of a geographically diverse patient across Japan.

Conclusions

In summary, based on the extensive multicenter, prospective data from 10,741 ED intubations, we developed modern machine learning models to predict clinically essential intubation outcomes. Using routinely available data as the predictors, we found that the machine learning models had a greater ability to predict difficult airways and first-pass success than conventional approaches. Although formal validation is required, this study lends support to the application of machine learning models for the prediction of intubation-related outcomes, which will, in turn, improve airway management practice and outcomes of critically ill patients in the ED.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Proportion of missingness in predictors used in machine learning models.
[DOCX File , 17 KB - ijmr_v11i1e28366_app1.docx]

References

1. Goto T, Goto Y, Hagiwara Y, Okamoto H, Watase H, Hasegawa K. Advancing emergency airway management practice and research. Acute Med Surg 2019 Oct;6(4):336-351 [FREE Full text] [doi: 10.1002/ams2.428] [Medline: 31592072]

2. Hasegawa K, Shigemitsu K, Hagiwara Y, Chiba T, Watase H, Brown CA, Japanese Emergency Medicine Research Alliance Investigators. Association between repeated intubation attempts and adverse events in emergency departments: an analysis of a multicenter prospective observational study. Ann Emerg Med 2012 Dec;60(6):749-754.e2. [doi: 10.1016/j.annemergmed.2012.04.005] [Medline: 22542734]
3. Walls RM, Brown CA, Bair AE, Pallin DJ, NEAR II Investigators. Emergency airway management: a multi-center report of 8937 emergency department intubations. J Emerg Med 2011 Oct;41(4):347-354. [doi: 10.1016/j.jemermed.2010.02.024] [Medline: 20434289]
4. Pallin DJ, Dwyer RC, Walls RM, Brown CA, NEAR III Investigators. Techniques and trends, success rates, and adverse events in emergency department pediatric intubations: a report from the National Emergency Airway Registry. Ann Emerg Med 2016 May;67(5):610-615.e1. [doi: 10.1016/j.annemergmed.2015.12.006] [Medline: 26921968]
5. Park L, Zeng I, Brainard A. Systematic review and meta-analysis of first-pass success rates in emergency department intubation: Creating a benchmark for emergency airway care. Emerg Med Australas 2017 Feb;29(1):40-47. [doi: 10.1111/1742-6723.12704] [Medline: 27785883]
6. Goto Y, Goto T, Hagiwara Y, Tsugawa Y, Watase H, Okamoto H, Japanese Emergency Medicine Network Investigators. Techniques and outcomes of emergency airway management in Japan: An analysis of two multicentre prospective observational studies, 2010-2016. Resuscitation 2017 May;114:14-20. [doi: 10.1016/j.resuscitation.2017.02.009] [Medline: 28219617]
7. Reed MJ, Dunn MJG, McKeown DW. Can an airway assessment score predict difficulty at intubation in the emergency department? Emerg Med J 2005 Feb;22(2):99-102 [FREE Full text] [doi: 10.1136/emj.2003.008771] [Medline: 15662057]
8. Hagiwara Y, Watase H, Okamoto H, Goto T, Hasegawa K, Japanese Emergency Medicine Network Investigators. Prospective validation of the modified LEMON criteria to predict difficult intubation in the ED. Am J Emerg Med 2015 Oct;33(10):1492-1496. [doi: 10.1016/j.ajem.2015.06.038] [Medline: 26166379]
9. Islam MM, Nasrin T, Walther BA, Wu C, Yang H, Li Y. Prediction of sepsis patients using machine learning approach: A meta-analysis. Comput Methods Programs Biomed 2019 Mar;170:1-9. [doi: 10.1016/j.cmpb.2018.12.027] [Medline: 30712598]
10. Goto T, Camargo CA, Faridi MK, Yun BJ, Hasegawa K. Machine learning approaches for predicting disposition of asthma and COPD exacerbations in the ED. Am J Emerg Med 2018 Sep;36(9):1650-1654. [doi: 10.1016/j.ajem.2018.06.062] [Medline: 29970272]
11. Goto T, Camargo CA, Faridi MK, Freishtat RJ, Hasegawa K. Machine learning-based prediction of clinical outcomes for children during emergency department triage. JAMA Netw Open 2019 Jan 04;2(1):e186937 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.6937] [Medline: 30646206]
12. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care 2019 Feb 22;23(1):64 [FREE Full text] [doi: 10.1186/s13054-019-2351-7] [Medline: 30795786]
13. Kotsiantis S, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. In: Maglogiannis I, Karpouzis K, Wallace BA, Soldatos J, editors. Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies. Amsterdam, Netherlands: IOS Press; 2007:3-27.
14. Goto T, Gibo K, Hagiwara Y, Morita H, Brown DF, Brown CA, Japanese Emergency Medicine Network Investigators. Multiple failed intubation attempts are associated with decreased success rates on the first rescue intubation in the emergency department: a retrospective analysis of multicentre observational data. Scand J Trauma Resusc Emerg Med 2015 Jan 16;23(1):5 [FREE Full text] [doi: 10.1186/s13049-014-0085-8] [Medline: 25700237]
15. Apfelbaum J, Hagberg C, Caplan R, Blitt C, Connis R, Nickinovich D, American Society of Anesthesiologists Task Force on Management of the Difficult Airway. Practice guidelines for management of the difficult airway: an updated report by the American Society of Anesthesiologists Task Force on Management of the Difficult Airway. Anesthesiology 2013 Feb;118(2):251-270 [FREE Full text] [doi: 10.1097/ALN.0b013e31827773b2] [Medline: 23364566]
16. Goto T, Gibo K, Hagiwara Y, Okubo M, Brown DFM, Brown CA, et al. Factors associated with first-pass success in pediatric intubation in the emergency department. West J Emerg Med 2016 Mar;17(2):129-134 [FREE Full text] [doi: 10.5811/westjem.2016.1.28685] [Medline: 26973736]
17. Link MS, Berkow LC, Kudenchuk PJ, Halperin HR, Hess EP, Moitra VK, et al. Part 7: adult advanced cardiovascular life support. Circulation 2015 Nov 03;132(18_suppl_2):S444-S464. [doi: 10.1161/cir.0000000000000261]
18. Hasegawa K, Hagiwara Y, Chiba T, Watase H, Walls RM, Brown DF, Japanese Emergency Medicine Research Alliance. Emergency airway management in Japan: Interim analysis of a multi-center prospective observational study. Resuscitation 2012 Apr;83(4):428-433. [doi: 10.1016/j.resuscitation.2011.11.027] [Medline: 22155701]
19. Stekhoven D. Nonparametric missing value imputation using random forest. 2012. URL: https://ui.adsabs.harvard.edu/abs/2015ascl.soft05011S [accessed 2021-01-08]
20. Warton DI. Penalized normal likelihood and ridge regularization of correlation and covariance matrices. Journal of the American Statistical Association 2012 Jan 01;103(481):340-349. [doi: 10.1198/016214508000000021]

21. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci 2003 Nov 04;43(6):1947-1958. [doi: 10.1021/ci034160g] [Medline: 14632445]
22. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot 2013;7:21 [FREE Full text] [doi: 10.3389/fnbot.2013.00021] [Medline: 24409142]
23. Pal S, Mitra S. Multilayer perceptron, fuzzy sets, and classification. IEEE Trans Neural Netw 1992;3(5):683-697. [doi: 10.1109/72.159058] [Medline: 18276468]
24. Bay S. Nearest neighbor classification from multiple feature subsets. Intelligent Data Analysis 1999 Sep;3(3):191-209. [doi: 10.1016/s1088-467x(99)00018-9]
25. Chen T, He T, Benesty M, Khotilovich V, Tang Y. Xgboost: extreme gradient boosting. R package version 2015:04 [FREE Full text]
26. Ruta D, Gabrys B. Classifier selection for majority voting. Information Fusion 2005 Mar;6(1):63-81. [doi: 10.1016/j.inffus.2004.04.008]
27. Pencina MJ, D'Agostino RB. Evaluating discrimination of risk prediction models: the c statistic. JAMA 2015 Sep 08;314(10):1063-1064. [doi: 10.1001/jama.2015.11082] [Medline: 26348755]
28. Taner T, Antony J. The assessment of quality in medical diagnostic tests: a comparison of ROC/Youden and Taguchi methods. Int J Health Care Qual Assur Inc Leadersh Health Serv 2000;13(6-7):300-307. [doi: 10.1108/09526860010378744] [Medline: 11484648]
29. Sagarin MJ, Barton ED, Chng Y, Walls RM, National Emergency Airway Registry Investigators. Airway management by US and Canadian emergency medicine residents: a multicenter analysis of more than 6,000 endotracheal intubation attempts. Ann Emerg Med 2005 Oct;46(4):328-336. [doi: 10.1016/j.annemergmed.2005.01.009] [Medline: 16187466]
30. Jung W, Kim J. Factors associated with first-pass success of emergency endotracheal intubation. Am J Emerg Med 2020 Jan;38(1):109-113. [doi: 10.1016/j.ajem.2019.09.001] [Medline: 31843066]
31. Taboada M, Soto-Jove R, Mirón P, Martínez S, Rey R, Ferreiroa E, et al. Evaluation of the laryngoscopy view using the modified Cormack-Lehane scale during tracheal intubation in an intensive care unit. A prospective observational study. Rev Esp Anestesiol Reanim (Engl Ed) 2019 May;66(5):250-258. [doi: 10.1016/j.redar.2019.01.004] [Medline: 30862397]
32. Connor C, Segal S. Accurate classification of difficult intubation by computerized facial analysis. Anesth Analg 2011 Jan;112(1):84-93. [doi: 10.1213/ANE.0b013e31820098d6] [Medline: 21081769]
33. Moustafa MA, El-Metainy S, Mahar K, Mahmoud Abdel-magied E. Defining difficult laryngoscopy findings by using multiple parameters: A machine learning approach. Egyptian Journal of Anaesthesia 2019 May 17;33(2):153-158. [doi: 10.1016/j.egja.2017.02.002]
34. Kuhn M, Johnson K. Applied predictive modeling. New York, NY: Springer; 2013.
35. An G, Omodaka K, Tsuda S, Shiga Y, Takada N, Kikawa T, et al. Comparison of machine-learning classification models for glaucoma management. J Healthc Eng 2018;2018:6874765 [FREE Full text] [doi: 10.1155/2018/6874765] [Medline: 30018755]
36. Freeman MB. Method and apparatus for automated data entry. Justia Patents. 2001. URL: https://patents.justia.com/patent/6299063 [accessed 2022-01-14]
37. Thanaki J. Python natural language processing. Birmingham, United Kingdom: Packt Publishing Ltd; 2017.
38. Lockey D, Crewdson K, Weaver A, Davies G. Observational study of the success rates of intubation and failed intubation airway rescue techniques in 7256 attempted intubations of trauma patients by pre-hospital physicians. Br J Anaesth 2014 Aug;113(2):220-225 [FREE Full text] [doi: 10.1093/bja/aeu227] [Medline: 25038154]

Abbreviations

ASA: American Society of Anesthesiologists
ED: emergency department
JEAN-2: second Japanese Emergency Airway Network
ROC: receiver operating characteristic

Edited by A Mavragani; submitted 03.03.21; peer-reviewed by A Miller, A Beane, Y Katayama; comments to author 16.04.21; revised version received 07.06.21; accepted 06.12.21; published 25.01.22.

Please cite as:
Yamanaka S, Goto T, Morikawa K, Watase H, Okamoto H, Hagiwara Y, Hasegawa K
Machine Learning Approaches for Predicting Difficult Airway and First-Pass Success in the Emergency Department: Multicenter Prospective Observational Study
Interact J Med Res 2022;11(1):e28366
URL: https://www.i-jmr.org/2022/1/e28366
doi:10.2196/28366
PMID:35076398

Original Paper

Self-Reported Data and Physician-Reported Data in Patients With Eosinophilic Granulomatosis With Polyangiitis: Comparative Analysis

Irena Doubelt[1,2], MD; Jason M Springer[3], MD, MSc; Tanaz A Kermani[4], MD; Antoine G Sreih[5], MD; Cristina Burroughs[6], MA; David Cuthbertson[6], MSc; Simon Carette[1,2], MD, MPhil; Nader A Khalidi[7], MD; Curry L Koening[8], MD, MSc; Carol Langford[9], MD, MHS; Carol A McAlear[5], MA; Larry W Moreland[10], MD; Paul A Monach[11], MD; Dianne G Shaw[12], MA; Philip Seo[13], MD; Ulrich Specks[14], MD; Kenneth J Warrington[15], MD; Kalen Young[12], MA; Peter A Merkel[5], MD, MPH; Christian Pagnoux[1,2], MD, MSc, MPH[‡]

[1]Vasculitis Clinic, Mount Sinai Hospital, Toronto, ON, Canada

[2]Division of Rheumatology, University of Toronto, Toronto, ON, Canada

[3]Division of Rheumatology and Immunology, Medical Center, Vanderbilt University, Nashville, TN, United States

[4]Division of Rheumatology, University of California, Los Angeles, Los Angeles, CA, United States

[5]Division of Rheumatology, University of Pennsylvania, Philadelphia, PA, United States

[6]Health Informatics Institute, University of South Florida, Tampa, FL, United States

[7]Division of Rheumatology, McMaster University and St. Joseph's Healthcare, Hamilton, ON, Canada

[8]Division of Rheumatology, University of Utah Hospital, Salt Lake City, UT, United States

[9]Division of Rheumatology, Cleveland Clinic, Cleveland, OH, United States

[10]Division of Rheumatology, University of Pittsburgh, Pittsburgh, PA, United States

[11]Division of Rheumatology, Veterans Affairs Boston Healthcare System, Boston, MA, United States

[12]Vasculitis Foundation, Kansas City, MO, United States

[13]Division of Rheumatology, Johns Hopkins University, Baltimore, MD, United States

[14]Division of Pulmonary and Critical Care Medicine, College of Medicine and Science, Mayo Clinic, Rochester, MN, United States

[15]Division of Rheumatology, College of Medicine and Science, Mayo Clinic, Rochester, MN, United States

[‡]Vasculitis Clinical Research Consortium and Vasculitis Patient-Powered Research Network

Corresponding Author:
Christian Pagnoux, MD, MSc, MPH
Vasculitis Clinic
Mount Sinai Hospital
60 Murray Street, Ste 2-220, Box 8
Toronto, ON, M5T 3L9
Canada
Phone: 1 416 586 4800 ext 8549
Email: christian.pagnoux@sinaihealth.ca

Abstract

Background: Patient-based registries can help advance research on rare diseases such as eosinophilic granulomatosis with polyangiitis (EGPA), a complex multiorgan form of antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis.

Objective: The aim of this study is to compare patient-reported and physician-reported data on manifestations, treatments, and outcomes for patients with EGPA.

Methods: We completed a comparative analysis of patients ≥18 years with EGPA in Canada and the United States from the following 2 cohorts: (1) The Vasculitis Patient-Powered Research Network (VPPRN), a self-enrolled secure portal with patient-entered data updated quarterly (2014-2019) and (2) the Vasculitis Clinical Research Consortium (VCRC) observational studies, a physician-entered database (2003-2019) of patients who fulfilled the 1990 American College of Rheumatology classification criteria for EGPA. The studied parameters included demographic characteristics, clinical manifestations, ANCA status, treatments, and relapses.

Results: Data from 195 patients with a validated diagnosis of EGPA in the VPPRN and 354 patients enrolled in the VCRC were analyzed. Compared to the VCRC cohort, the patients in the VPPRN cohort were more likely to be female (135/195, 69.2% compared to 209/354, 59%; *P*=.02) and younger at diagnosis (47.3 compared to 50.0 years; *P*=.03); both cohorts reported similar frequencies of asthma (177/184, 96.2% in the VPPRN cohort compared to 329/354, 92.9% in the VCRC cohort; *P*=.13) and cardiac manifestations (44/153, 28.8% compared to 75/354, 21.2%; *P*=.06), but the VPPRN cohort reported less frequent lung manifestations other than asthma and more frequent disease manifestations in all other organ systems. The ANCA positivity was 48.9% (64/131) in the VPPRN patients compared to 38.9% (123/316; *P*=.05) in the VCRC cohort. Relapsing disease after study enrollment was reported in 32.3% (63/195) of patients in the VPPRN compared to 35.7% (99/277) of patients in the VCRC. Most therapies (GC, cyclophosphamide, mepolizumab) were used at similar frequencies in both groups, except for rituximab with VPPRN patients reporting more use than the VCRC cohort (47/195, 24.1% compared to 29/277, 10.5%; *P*<.001).

Conclusions: Overall, patients and physicians report manifestations of EGPA at similar frequencies. However, observed differences between patient and physician reports imply the potential occurrence of selection biases. These results support the use of patient-reported data in EGPA but also the need for careful consideration of disease-specific definitions for the study of EGPA and how patient- and physician-reported data are collected.

Trial Registration: ClinicalTrials.gov NCT00315380, https://clinicaltrials.gov/ct2/show/NCT00315380; ClinicalTrials.gov NCT01241305, https://clinicaltrials.gov/ct2/show/NCT01241305

(Interact J Med Res 2022;11(1):e27273) doi:10.2196/27273

KEYWORDS

eosinophilic granulomatosis with polyangiitis; patient-reported outcomes measures; clinical outcomes; granulomatosis; patient outcomes; digital health; health network; health databases; research network

Introduction

Vasculitides are rare, heterogeneous, multisystem diseases causing inflammation of blood vessels [1]. Physicians determine disease activity, damage, and prognosis in vasculitis using various clinical, laboratory, or radiological parameters or tools, such as the Birmingham Vasculitis Activity Score [2] or the Vasculitis Damage Index [3]. Many of these parameters may at times differ from the patient's subjective disease experiences. There is growing interest in increasing patient engagement in health care research to improve the alignment of patients' and physicians' perspectives in the diagnosis, management, and assessment of outcomes and burden of disease; more and more action has been taken to achieve this. [4].

The Vasculitis Patient-Powered Research Network (VPPRN), an international, internet-based, prospective longitudinal registry of patient- or caregiver-reported information, was launched in 2014 to support people with any form of vasculitis by involving them in clinical research [4-6]. With over 3000 members enrolled to date, mostly from North America, it maintains a secure web-based registry where patients provide clinical data about themselves and their condition regarding demographic characteristics, diagnosis, disease extent, medications, and outcomes [4].

The Vasculitis Clinical Research Consortium (VCRC), established in 2003, has been collecting longitudinal data in patients with various vasculitides across 8 US and 2 Canadian sites. VCRC site investigators collected similar clinical information as the VPPRN.

This study aimed to compare patient self-reported and physician-reported clinical manifestations, treatments, and outcomes in patients with eosinophilic granulomatosis with polyangiitis (EGPA), one of the antineutrophil cytoplasmic antibody (ANCA)–associated vasculitides. EGPA is a complex multisystem disorder that can involve any combination of many manifestations, especially including asthma, rhinosinusitis, eosinophilia, and vasculitis in various organs. Research on EGPA has not been as extensively conducted compared to several other forms of vasculitis. A better understanding of the utility of patient-based research registries could help advance research on this rare disease.

Methods

Vasculitis Patient-Powered Research Network: Patient-Driven Cohort

The VPPRN provides a secure portal through which patients self-enroll and self-report information longitudinally using the internet-based platform, as previously described [6]. Data from patients in the VPPRN (2014-2019) who self-identified as being age 18 years or older, living in Canada or the United States, and having EGPA were used for this analysis. For validation of the diagnosis, patients were excluded if they indicated that the diagnosis of EGPA was not made by a doctor and/or if they reported never having used systemic glucocorticoids (GC).

Standardized questions were used to obtain data on demographic characteristics (age, sex, ethnicity), signs and symptoms of vasculitis at any time after disease onset, diagnostic tests, prescribed treatments, and outcomes from patients, with quarterly updates by email reminders. Questions related to disease manifestations were asked using lay terms, as listed in Table 1. Patients could select responses of yes, no, or I don't know; blank responses were excluded from the data analysis.

Vasculitis Clinical Research Consortium: Physician-Driven Cohort

Data from patients with EGPA entered into the VCRC database, as part of either the VCRC Longitudinal Study (LS; NCT00315380) [7] or the One-Time DNA (OT) Study

(NCT01241305; conducted 2013-2019) [8], were used for this analysis. Patients were ages 18 years or older at enrollment and met the 1990 American College of Rheumatology classification criteria for EGPA [9]. In the VCRC-LS observational cohort, participants had in-person assessments with site investigators at either quarterly or annual visits, based on each patient's preference and availability, with data collection of clinical and laboratory information. Patients in the VCRC-OT were assessed only at a single study visit (at diagnosis or later). All study visits involved the completion of standardized forms that collected information on patient demographic characteristics, disease characteristics, relapse(s) prior to enrollment, and, for LS only, treatments received, relapses after enrollment, and disease-related damage (from the disease itself or treatment).

Patients may be enrolled in both databases (VPPRN and VCRC); however, at present, to comply with regulations protecting health information, the databases are not linked.

Data Elements

Demographic characteristics, main clinical manifestations of EGPA (from disease diagnosis to data extraction), ANCA status, follow-up duration, relapses (from diagnosis and/or after enrollment), and all treatments ever received (GC and other immunosuppressive drugs) were analyzed for both cohorts.

Ethics Approval

The VCRC study protocols were approved by the local hospital research ethics board committees at all participating VCRC sites. The VPPRN data collection protocol was approved by the institutional review board at the University of South Florida (Pro00018514_CR000001). All subjects in both the VCRC and VPPRN provided consent for their participation prior to enrollment. All research was performed in accordance with the ethical principles that have their origin in the Declaration of Helsinki.

Statistical Analyses

Descriptive statistics were computed by calculating the mean and SD for quantitative variables and count (percent) for categorical variables. Quantitative variables were compared using unequal variances *t* tests; categorical variables were compared using chi-square tests. Statistically significant differences were defined as those with *P* values ≤.05. Statistical analyses were performed using Stata (version 12; StataCorp).

Results

Vasculitis Patient-Powered Research Network: Patient Characteristics

At the time of data extraction, a total of 208 patients were registered in the VPPRN with a diagnosis of EGPA. A total of 13 patients were excluded as they indicated they had not been diagnosed by a doctor and/or had never been treated with systemic GC. Of the final 195 patients, 69.2% (n=135) were female; the average age at diagnosis was 47.3 years; 176 (90.3%) were White, 4 (2.1%) were Asian, and 1 (0.5%) was Black or African American. Additional cohort-specific details are presented in Table 1.

For the 195 patients, the methods by which their diagnoses of EGPA were determined comprised of the following: symptoms (n=178, 91.3%), laboratory testing (n=157, 80.5%), biopsy results (n=98, 50.3%), and imaging (n=98, 50.3%).

Vasculitis Clinical Research Consortium: Patient Characteristics

At the time of data extraction, the VCRC cohort included 354 patients (277 LS and 77 OT). Of the 354 patients, 59% (n=209) were female; the average age at diagnosis was 50.0 years; 309 (87.3%) were White, 20 (5.6%) were Asian, and 7 (2%) were Black or African American. Additional cohort-specific details are presented in Table 1.

Comparisons Between Patients with EGPA in the VPPRN and VCRC

Comparisons of demographic characteristics, ANCA status, clinical manifestations (from disease diagnosis to data extraction), relapses, and all treatments ever used for patients in the VPPRN and VCRC are shown in Table 1. Compared to the patients in the VCRC, patients in the VPPRN were younger at the time of diagnosis and reported similar frequencies of asthma and cardiac manifestations, less frequent lung manifestations other than asthma, and more frequent disease manifestations in all other organ systems. Relapse rates post enrollment were similar between the 2 cohorts.

Table 1. A comparison of the clinical characteristics of the Vasculitis Patient-Powered Research Network and Vasculitis Clinical Research Consortium eosinophilic granulomatosis with polyangiitis cohorts.

Characteristics	VPPRN[a] cohort (N=195)	VCRC[b] cohort (N=354)	*P* value
Sex, n (%)			
Female	135 (69.2)	209 (59)	.02
Male	60 (30.8)	145 (40.9)	.02
Age at diagnosis in years, mean (SD)	47.3 (14.3)	50.0 (14.2)	.03
Positive test for ANCA[c], n (%)[d]	64 (48.9)	123 (38.9)	.05
Manifestations, n (%)[e,f]			
Asthma	177 (96.2)	329 (92.9)	.13
Coughed up blood or bleeding in the lungs/alveolar hemorrhage	24 (14.4)	21 (5.9)	.001
Problems with your lungs/lung involvement	126 (72.4)	296 (83.6)	.003
Problems with your nose or sinuses/nasal involvement	165 (92.2)	292 (82.5)	.003
Fever	82 (55.4)	62 (17.5)	<.001
Weight loss	95 (55.6)	106 (29.9)	<.001
Severe joint pain or swelling/arthralgia(s)	116 (67.1)	140 (39.5)	<.001
Rash/skin	125 (70.6)	106 (29.9)	<.001
Inflammation of the heart lining/cardiac	44 (28.8)	75 (21.2)	.06
Problems with your kidneys/renal disease	39 (22.4)	36 (10.2)	<.001
Numbness, tingling, trouble moving arms, hands, legs, or feet, or other forms of nerve damage/neurological	155 (87.6)	214 (60.5)	<.001
Inflammation in one or both eyes that required treatment/eye disease	42 (26.1)	31 (8.8)	<.001
Loss of blood supply to intestines or perforation/mesenteric ischemia	10 (6.1)	7 (2)	.02
Thrombosis	24 (14)	24 (6.8)	.007
Follow-up time in years, mean (SD)			
From diagnosis	8.0 (6.8)	7.0 (6.2)	.08
From enrollment	2.2 (1.1)	3.6 (3.5)	<.001
Relapses, n (%)			
Total since diagnosis	N/A	175 (49.4)	N/A[g]
After enrollment	63 (32.3)	99 (35.7)[h]	.44
Deaths, n (%)	0 (0)[i]	11 (4)[h]	N/A
Treatments ever received, n (%)[j]			
Systemic glucocorticoids	195 (100)	354 (100)	.99
Cyclophosphamide	79 (40.5)	115 (41.5)[h]	.83
Mepolizumab	20 (10.3)	25 (9)[h]	.65
Rituximab	47 (24.1)	29 (10.5)[h]	<.001

[a]VPPRN: Vasculitis Patient-Powered Research Network.

[b]VCRC: Vasculitis Clinical Research Consortium.

[c]ANCA: antineutrophil cytoplasmic antibody.

[d]For this category, N=131 for the VPPRN cohort and N=316 for the VCRC cohort. Percentages have been calculated accordingly.

[e]For the VPPRN cohort, the N value for each category is the number of patients who responded yes or no (the response of "I don't know" was excluded). The N values are as follows: asthma (N=184); coughed up blood or bleeding in the lungs (N=166); problems with your lungs (N=174); problems with your nose or sinuses (N=179); fever (N=148); weight loss (N=171); severe joint pain or swelling (N=173); rash (N=177); inflammation of the heart lining (N=153); problems with your kidneys (N=174); numbness, tingling, trouble moving arms, hands, legs, or feet, or other forms of nerve damage (N=177); Inflammation in one or both eyes that required treatment (N=161); loss of blood supply to intestines or perforation (N=165); thrombosis

(N=171). The percentages have been calculated accordingly.

[f]For this category, some items are presented as follows: phrasing in VPPRN database/phrasing in VCRC database.

[g]N/A: not applicable.

[h]Data were available for 277 patients in the VCRC–Longitudinal Study (VCRC-LS). These percentages have been calculated accordingly.

[i]All patients in the VPPRN logged into the portal and completed at least 1 form within 24 months prior to data extraction with none documented as lost to follow-up; captured follow-up losses and deaths in VPPRN are limited due to the inherent nature of the study design (privacy concerns associated with contacting treating physicians, family members, etc.).

[j]Additional treatments for patients in the VCRC-LS (N=277) included azathioprine (n=145, 52.3%), methotrexate (n=109, 39.4%), and mycophenolate mofetil (n=25, 9%).

Discussion

Principal Findings

In this study, we compared patient-reported and physician-reported outcomes in 2 cohorts of patients with EGPA; both patients and physicians reported a spectrum of outcomes and relative frequencies of manifestations, relapse rates, and medication use that are quite consistent with what is expected for this heterogeneous multisystem disease. However, some interesting differences in how this disease was reported were also observed between the 2 cohorts, raising the possibility of selection biases impacting data collection in these 2 registries.

With the exception of asthma and cardiac involvement, patients reported, for example, higher frequencies of almost all manifestations of EGPA compared to physicians.

These differences could be due to any combination of several reasons, including the following: (1) patients over-report manifestations, some of which may not be due to EGPA; (2) physicians under-report manifestations and/or do not validate patients' reports of problems prior to evaluating them; (3) patients and physicians have a different understanding of specific manifestations; and (4) the VPPRN and VCRC involved 2 separate cohorts due to selection and inclusion biases.

Whereas the 2 cohorts did appear similar overall to what is expected for a large group of patients with EGPA, it is possible that the patients in the VPPRN had, at least initially, more severe disease or a different disease phenotype. The VPPRN cohort had a higher proportion of ANCA-positive patients, who were shown in a few previous cohort studies to present more often with surrogates of vasculitis, such as renal or cutaneous involvement [10-13]. ANCA-positive patients may also be treated with rituximab more frequently compared to the ANCA-negative EGPA population. Due to regulations protecting health information, direct linkage and comparisons of the patient-reported and physician-reported data for those "shared" participants (patients enrolled in both studies regardless of ANCA status) were not possible.

Certain subgroups may have been overrepresented due to biased sampling, as in other patient-driven registries. The mode of survey via internet-based participation for the VPPRN cohort may have enrolled more younger and female patients and more patients with ready access to the internet [14]. Such distinctions in patient- and physician-based disease features have also been observed in other rheumatic conditions [15-18]. However, previous studies using data in the VPPRN on the 2 other types of ANCA-associated vasculitis (granulomatosis with polyangiitis and microscopic polyangiitis) provided evidence for the clinical validity of the VPPRN data, thus the likely limited impact of any selection biases [6].

Misinterpretation or the use of different definitions of disease manifestations may also have contributed to some differences or inaccurate results in both cohorts. For instance, almost all patients with EGPA have asthma, whereas 70-80% of patients responded yes to "problems with [your] lungs/ lung involvement." Asthma is considered a comorbidity or underlying manifestation of EGPA. For disease scoring and in studies on EGPA, "lung manifestations" thus usually involve other nonasthma symptoms or manifestations, separately, such as lung infiltrates, nodules, or alveolar hemorrhage. This distinction is not always applied by physicians in studies and may be even more difficult to understand by patients. The way data and information are collected in registries is crucial and clear wording is essential to deal with such aspects of the disease, especially for studies with patient-reported data. Investigators should consider all these factors when researching EGPA, both through internet-based mechanisms and traditional clinic-based approaches.

This study provides new and valuable information from both a patient and physician perspective and has several notable strengths including the sizes of the geographically diverse study cohorts, the extent of the data, overlapping leadership of the 2 research cohorts, and the long duration of follow-up for data collection. Highly experienced investigators in the field of vasculitis selected, designed, and collected the data elements using standardized forms in the VCRC and VPPRN cohorts, and patients had direct input during each stage of the process for the development of the VPPRN forms.

Conclusions

There is a clear mandate and many benefits for health care professionals to incorporate patient perspectives into the assessment and management of vasculitis [19]. How patients' perspectives may correspond to physicians' assessments remains a subject of study and uncertainty. This study comparing patient-reported and physician-reported characteristics and outcomes in patients with EGPA should encourage rethinking and refinement of how patients are recruited and how their data are collected for complex diseases. These results support the use of patient-reported data in EGPA but also the need for careful consideration of these 2 types of registries, how patient- and physician-reported data are collected, as well as disease-specific definitions for the study of EGPA. Establishing a common set of disease-specific items and outcomes in EGPA and using similar definitions or wording is advised when seeking to combine data from both patients and physicians.

Acknowledgments

The VPPRN has received support from the Patient-Centered Outcomes Research Institute (grant PPRN-1306-04758) and the VCRC. The VCRC is part of the Rare Diseases Clinical Research Network, an initiative of the Office of Rare Diseases Research at the National Center for Advancing Translational Science (NCATS). The VCRC has received funding from the NCATS, the National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant U54 AR057319), the National Center for Research Resources (grant U54 RR019497), and GlaxoSmithKline plc.

Authors' Contributions

All authors were involved in substantial contributions to study conception and design, acquisition of data, and approval of the final manuscript version. ID, CB, DC, PA Merkel, and CP had full access to the data in the study and take responsibility for the integrity of the data, accuracy of the data analysis, data interpretation, and drafting of the article.

Conflicts of Interest

CP reports receiving fees for serving on advisory boards for ChemoCentryx Inc, GlaxoSmithKline plc, Sanofi SA, F. Hoffman-La Roche AG, InflaRx GmbH, and AstraZeneca plc; he also reports receiving lecture fees and research grant support from F. Hoffman-La Roche AG and GlaxoSmithKline plc. PA Merkel reports receiving funds for the following activities in the past 2 years: consulting for AbbVie Inc, AstraZeneca plc, Biogen Inc, Boehringer Ingelheim, Bristol Myers Squibb Co, Celgene Corp, ChemoCentryx Inc, CSL Behring, Dynacure SA, EMD Serono, Forbius, Genentech Inc/F. Hoffman-La Roche AG, Genzyme Corp/Sanofi SA, GlaxoSmithKline plc, InflaRx GmbH, Insmed Inc, Janssen Inc, Kiniksa Pharmaceuticals Ltd, Kyverna Therapeutics Inc, Magenta Therapeutics Inc, Neutrolis Inc, Novartis AG, Pfizer Inc, Sparrow Pharmaceuticals Inc, Takeda Pharmaceutical Company Ltd, Talaris Therapeutics Inc; research support from AbbVie Inc, AstraZeneca plc, Boehringer Ingelheim, Bristol Myers Squibb Co, Celgene Corp, ChemoCentryx Inc, Eicos, Forbius, Genentech Inc/F. Hoffman-La Roche AG, Genzyme Corp/Sanofi SA, GlaxoSmithKline plc, InflaRx GmbH, Sanofi SA, and Takeda Pharmaceutical Company Ltd; and royalties from UpToDate Inc. AGS is an employee of Bristol Myers Squibb Co. NAK is part of the advisory board/speaker bureau at F. Hoffman-La Roche AG and also receives research grant support from F. Hoffman-La Roche AG, Bristol Myers Squibb Co, Sanofi, AbbVie Inc, and AstraZeneca plc.

References

1. Jennette JC, Falk RJ, Bacon PA, Basu N, Cid MC, Ferrario F, et al. 2012 revised International Chapel Hill Consensus Conference nomenclature of vasculitides. Arthritis Rheum 2013 Jan;65(1):1-11 [FREE Full text] [doi: 10.1002/art.37715] [Medline: 23045170]
2. Mukhtyar C, Lee R, Brown D, Carruthers D, Dasgupta B, Dubey S, et al. Modification and validation of the Birmingham Vasculitis Activity Score (version 3). Ann Rheum Dis 2009 Dec;68(12):1827-1832. [doi: 10.1136/ard.2008.101279] [Medline: 19054820]
3. Exley AR, Bacon PA, Luqmani RA, Kitas GD, Gordon C, Savage CO, et al. Development and initial validation of the Vasculitis Damage Index for the standardized clinical assessment of damage in the systemic vasculitides. Arthritis Rheum 1997 Feb;40(2):371-380. [doi: 10.1002/art.1780400222] [Medline: 9041949]
4. Young K, Kaminstein D, Olivos A, Burroughs C, Castillo-Lee C, Kullman J, Vasculitis Patient-Powered Research Network, et al. Patient involvement in medical research: what patients and physicians learn from each other. Orphanet J Rare Dis 2019 Jan 24;14(1):21 [FREE Full text] [doi: 10.1186/s13023-018-0969-1] [Medline: 30678705]
5. PCORnet PPRN Consortium, Daugherty SE, Wahba S, Fleurence R. Patient-powered research networks: building capacity for conducting patient-centered clinical outcomes research. J Am Med Inform Assoc 2014;21(4):583-586 [FREE Full text] [doi: 10.1136/amiajnl-2014-002758] [Medline: 24821741]
6. Springer JM, Kermani TA, Sreih A, Shaw DG, Young K, Burroughs CM, et al. Clinical characteristics of an internet-based cohort of patient-reported diagnosis of granulomatosis with polyangiitis and microscopic polyangiitis: observational study. J Med Internet Res 2020 Jul 20;22(7):e17231 [FREE Full text] [doi: 10.2196/17231] [Medline: 32459634]
7. Longitudinal study for eosinophilic granulomatosis with polyangiitis (Churg-Strauss). ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/record/NCT00315380?view=record [accessed 2022-03-22]
8. One-time DNA study for vasculitis. ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/NCT01241305 [accessed 2022-03-22]
9. Masi AT, Hunder GG, Lie JT, Michel BA, Bloch DA, Arend WP, et al. The American College of Rheumatology 1990 criteria for the classification of Churg-Strauss syndrome (allergic granulomatosis and angiitis). Arthritis Rheum 1990 Aug;33(8):1094-1100. [doi: 10.1002/art.1780330806] [Medline: 2202307]
10. Cottin V, Bel E, Bottero P, Dalhoff K, Humbert M, Lazor R, Groupe d'Etudes et de Recherche sur les Maladies Orphelines Pulmonaires (GERM"O"P). Revisiting the systemic vasculitis in eosinophilic granulomatosis with polyangiitis (Churg-Strauss): a study of 157 patients by the Groupe d'Etudes et de Recherche sur les Maladies Orphelines Pulmonaires

and the European Respiratory Society Taskforce on eosinophilic granulomatosis with polyangiitis (Churg-Strauss). Autoimmun Rev 2017 Jan;16(1):1-9. [doi: 10.1016/j.autrev.2016.09.018] [Medline: 27671089]
11. Sablé-Fourtassou R, Cohen P, Mahr A, Pagnoux C, Mouthon L, Jayne D, French Vasculitis Study Group. Antineutrophil cytoplasmic antibodies and the Churg-Strauss syndrome. Ann Intern Med 2005 Nov 01;143(9):632-638. [doi: 10.7326/0003-4819-143-9-200511010-00006] [Medline: 16263885]
12. Comarmond C, Pagnoux C, Khellaf M, Cordier J, Hamidou M, Viallard J, French Vasculitis Study Group. Eosinophilic granulomatosis with polyangiitis (Churg-Strauss): clinical characteristics and long-term followup of the 383 patients enrolled in the French Vasculitis Study Group cohort. Arthritis Rheum 2013 Jan;65(1):270-281 [FREE Full text] [doi: 10.1002/art.37721] [Medline: 23044708]
13. Sinico RA, Di Toma L, Maggiore U, Bottero P, Radice A, Tosoni C, et al. Prevalence and clinical significance of antineutrophil cytoplasmic antibodies in Churg-Strauss syndrome. Arthritis Rheum 2005 Sep;52(9):2926-2935 [FREE Full text] [doi: 10.1002/art.21250] [Medline: 16142760]
14. Abdou NI, Kullman GJ, Hoffman GS, Sharp GC, Specks U, McDonald T, et al. Wegener's granulomatosis: survey of 701 patients in North America. Changes in outcome in the 1990s. J Rheumatol 2002 Feb;29(2):309-316. [Medline: 11838848]
15. Tomasson G, Farrar JT, Cuthbertson D, McAlear CA, Ashdown S, Cronholm PF, Vasculitis Clinical Research Consortium. Feasibility and construct validation of the patient reported outcomes measurement information system in systemic vasculitis. J Rheumatol 2019 Aug;46(8):928-934 [FREE Full text] [doi: 10.3899/jrheum.171405] [Medline: 30824648]
16. Kirwan JR, Boonen A, Harrison MJ, Hewlett SE, Wells GA, Singh JA, et al. OMERACT 10 Patient Perspective Virtual Campus: valuing health; measuring outcomes in rheumatoid arthritis fatigue, RA sleep, arthroplasty, and systemic sclerosis; and clinical significance of changes in health. J Rheumatol 2011 Aug;38(8):1728-1734. [doi: 10.3899/jrheum.110393] [Medline: 21807793]
17. Hewlett S, Sanderson T, May J, Alten R, Bingham CO, Cross M, et al. 'I'm hurting, I want to kill myself': rheumatoid arthritis flare is more than a high joint count--an international patient perspective on flare where medical help is sought. Rheumatology (Oxford) 2012 Jan;51(1):69-76. [doi: 10.1093/rheumatology/keq455] [Medline: 21565901]
18. Neogi T, Xie H, Felson DT. Relative responsiveness of physician/assessor-derived and patient-derived core set measures in rheumatoid arthritis trials. J Rheumatol 2008 May;35(5):757-762 [FREE Full text] [Medline: 18412313]
19. Robson JC, Jayne D, Merkel PA, Dawson J. Systemic vasculitis and patient-reported outcomes: how the assessment of patient preferences and perspectives could improve outcomes. Patient Relat Outcome Meas 2019;10:37-42 [FREE Full text] [doi: 10.2147/PROM.S163601] [Medline: 30804691]

Abbreviations

ANCA: antineutrophil cytoplasmic antibody
EGPA: eosinophilic granulomatosis with polyangiitis
GC: glucocorticoids
LS: longitudinal study
NCATS: National Center for Advancing Translational Science
OT: one-time DNA
VCRC: Vasculitis Clinical Research Consortium
VPPRN: Vasculitis Patient-Powered Research Network

Edited by G Eysenbach; submitted 21.01.21; peer-reviewed by G Klein, D Ichikawa; comments to author 11.05.21; revised version received 03.07.21; accepted 30.11.21; published 25.05.22.

Please cite as:
Doubelt I, Springer JM, Kermani TA, Sreih AG, Burroughs C, Cuthbertson D, Carette S, Khalidi NA, Koening CL, Langford C, McAlear CA, Moreland LW, Monach PA, Shaw DG, Seo P, Specks U, Warrington KJ, Young K, Merkel PA, Pagnoux C
Self-Reported Data and Physician-Reported Data in Patients With Eosinophilic Granulomatosis With Polyangiitis: Comparative Analysis
Interact J Med Res 2022;11(1):e27273
URL: https://www.i-jmr.org/2022/1/e27273
doi:10.2196/27273
PMID:35612893

Original Paper

The Logistics of Medication and Patient Flow in Video-Based Virtual Clinics During a Sudden COVID-19 Outbreak in Taiwan: Observational Study

Ying-Hsien Chen[1,2*], MD; Hui-Wen Wu[1,2*], BA; Ching-Chang Huang[1,2], MD; Jen-Kuang Lee[1,2], MD; Li-Tan Yang[1,2], MD; Tse-Pin Hsu[1,2], MA; Chi-Sheng Hung[1,2], MD, PhD; Yi-Lwun Ho[1,2], MD, PhD

[1]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

[2]National Taiwan University Hospital Telehealth Center, National Taiwan University Hospital, Taipei, Taiwan

[*]these authors contributed equally

Corresponding Author:
Yi-Lwun Ho, MD, PhD
Department of Internal Medicine
National Taiwan University Hospital
No 7 Chung Shan S Rd
Taipei, 100225
Taiwan
Phone: 886 02 23123456 ext 262152
Email: ylho@ntu.edu.tw

Abstract

Background: The COVID-19 pandemic was well controlled in Taiwan until an outbreak in May 2021. Telemedicine was rapidly implemented to avoid further patient exposure and to unload the already burdened medical system.

Objective: To understand the effect of COVID-19 on the implementation of video-based virtual clinic visits during this outbreak, we analyzed the logistics of prescribing medications and patient flow for such virtual visits at a tertiary medical center.

Methods: We retrospectively collected information on video-based virtual clinic visits and face-to-face outpatient visits from May 1 to August 31, 2021, from the administrative database at National Taiwan University Hospital. The number of daily new confirmed COVID-19 cases in Taiwan was obtained from an open resource.

Results: There were 782 virtual clinic visits during these 3 months, mostly for the departments of internal medicine, neurology, and surgery. The 3 most common categories of medications prescribed were cardiovascular, diabetic, and gastrointestinal, of which cardiovascular medications comprised around one-third of all medications prescribed during virtual clinic visits. The number of virtual clinic visits was significantly correlated with the number of daily new confirmed COVID-19 cases, with approximately a 20-day delay (correlation coefficient 0.735; $P<.001$). The patient waiting time for video-based virtual clinic visits was significantly shorter compared with face-to-face clinic visits during the same period (median 3, IQR 2-6 min vs median 20, IQR 9-42 min; rank sum $P<.001$). Although the time saved was appreciated by the patients, online payment with direct delivery of medications without the need to visit a hospital was still their major concern.

Conclusions: Our data showed that video-based virtual clinics can be implemented rapidly after a COVID-19 outbreak. The virtual clinics were efficient, as demonstrated by the significantly reduced waiting time. However, there are still some barriers to the large-scale implementation of video-based virtual clinics. Better preparation is required to improve performance in possible future large outbreaks.

(Interact J Med Res 2022;11(1):e37880) doi:10.2196/37880

KEYWORDS

COVID-19; telemedicine; video-based virtual clinic

Introduction

COVID-19, caused by SARS-CoV-2, rapidly escalated to a worldwide pandemic from late 2019 to early 2020 [1,2]. More than 150 million confirmed cases had been reported worldwide at the end of April 2021. Social distancing, mask wearing, and even lockdowns have been used worldwide to try and control the pandemic. Vaccination for SARS-CoV-2 was first approved in December 2020, and by the end of April 2021, 25 million doses of vaccine had been administered worldwide [3]. To continue medical care and minimize unnecessary contact, many countries have implemented virtual clinics for patients with chronic conditions during the pandemic [4-11]. The platform and workflow of such virtual clinics continues to evolve to match the demands of medical needs during different phases of the pandemic [12].

In Taiwan, thanks to the timely use of case-based and population-based interventions, including border control, enhanced surveillance, contact tracing, travel restrictions, and quarantine, the number of confirmed COVID-19 cases has been well controlled, with only 1129 cases reported at the end of April 2021 [13,14]. However, a community outbreak occurred in early May 2021 that resulted in a rapid increase in the number of cases from 1100 to over 15,000 in less than 3 months. The Taiwan government raised the nationwide epidemic alert level from II to III on May 15, 2021. The vaccination rate was also suboptimal at that time. The first AstraZeneca vaccine was administered on March 3, 2021, and at the end of April 2021, fewer than 60,000 individuals had received at least one dose of vaccine.

During the outbreak, most patients with COVID-19 in Taiwan were hospitalized for isolation, monitoring, and treatment. One consequence of the rapid increase in COVID-19 hospitalizations was that medical resources were directed to pandemic control as opposed to regular medical work. For example, around 1100 negative-pressure isolation rooms around Taiwan were directly controlled by the government [15]. From January to March 2020, the number of medical visits fell by 6.8% compared to the same period in 2019 [16]. Meanwhile, emerging data showed that the virus could be spread by patients who were asymptomatic and that viral RNA can be detected in nasal and throat swabs up to 11 days after contact in a patient who is asymptomatic [17]. Therefore, regular outpatient visits by patients with chronic diseases decreased substantially. To deal with this situation, the Taiwan government reduced restrictions on telehealth to allow physicians to speak to patients with stable chronic conditions via virtual clinics, either by video teleconferencing or telephone call. The Bureau of National Health Insurance (NHI) in Taiwan also started to reimburse for virtual clinic visits from May 15, 2021, during the pandemic [18]. Regular refilling of prescriptions for this group of patients was also allowed during a virtual clinic visit. Virtual clinics were then gradually initiated nationwide, from community clinics to medical centers. This process was carefully controlled over 3 months [19]. The use of virtual clinics also decreased gradually 3 months after the outbreak. This is the first time that reimbursement for a virtual clinic has been granted in Taiwan.

The aim of this study was to present the logistics of medication and patient flow for video-based virtual clinics during the May 2021 outbreak in Taiwan at a tertiary hospital. Although there have been some previous reports on the use of virtual clinics during the COVID-19 pandemic worldwide [4-11], our study reports the implementation of virtual clinics in a country with few confirmed cases and good infection control for 16 months from January 2020. We analyzed video-based virtual clinic visits from May to August 2021 at National Taiwan University Hospital (NTUH) in Taipei. We also discuss potential improvements that can be made in preparation for the next outbreak, as well as opportunities to further develop the use of telehealth in the future.

Methods

Ethics

This study was approved by the Institutional Review Board of National Taiwan University Hospital (Taipei, Taiwan; 202202037RINC).

Study Design

We designed this retrospective study to evaluate video-based virtual clinic visits during the May 2021 COVID-19 outbreak in NTUH. The numbers of patients with chronic conditions who used the video-based virtual clinic and who attended face-to-face clinic visits were obtained from our electronic health database. *International Classification of Diseases, Tenth Revision, Clinical Modification* (*ICD-10-CM*) codes were used to extract the diagnoses from the database. The *ICD-10-CM* codes used in this study are as follows: diabetes mellitus (E08, E09, E10, E11, E13), hypertension (I10, I15), coronary arterial disease (I20, I21, I22, I23, I24, I25), hypertensive heart disease without heart failure (I11), atrial fibrillation (I48), enlarged prostate with lower urinary tract symptoms (N40.1), hyperlipidemia (E78), Parkinson disease (G20), and congestive heart failure (I50). The characteristics of patients who attended the video-based virtual clinic were also analyzed. Furthermore, associations between the number of daily confirmed COVID-19 cases and the numbers of patients who visited our virtual or face-to-face clinics were also analyzed. The care of patients who were hospitalized with COVID-19 in Taiwan is directed and coordinated by the Taiwan Centers for Disease Control, based on location and the availability of negative-pressure rooms and trained personnel. As the number of cases at a single hospital may not reflect the severity of the outbreak, we used the total number of cases collected at all places around Taiwan to reflect the situation.

Video-Based Virtual Clinic Implementation at Our Institute

Regulations

Regulations on the implementation of telemedicine were revised in Taiwan in 2018. Hospitals or clinics are only permitted to set up virtual clinics when following special regulations. However, the ability to prescribe medications via virtual clinics is not routinely granted and can only be performed in a few emergency situations. NTUH, a tertiary teaching hospital in

Taipei, has provided a video-based virtual clinic service after obtaining approval in 2019, before the COVID-19 pandemic. The software, platform, and hardware for the implementation of video-based virtual clinics was therefore well established. However, the service fee for video-based virtual clinic appointments was not covered by the NHI in Taiwan before the COVID-19 pandemic. After the May 2021 COVID-19 outbreak, the Bureau of NHI agreed to cover the fee for the virtual clinics (telephone or video-based) during the pandemic. The prescription of medications during virtual clinic visits was also allowed.

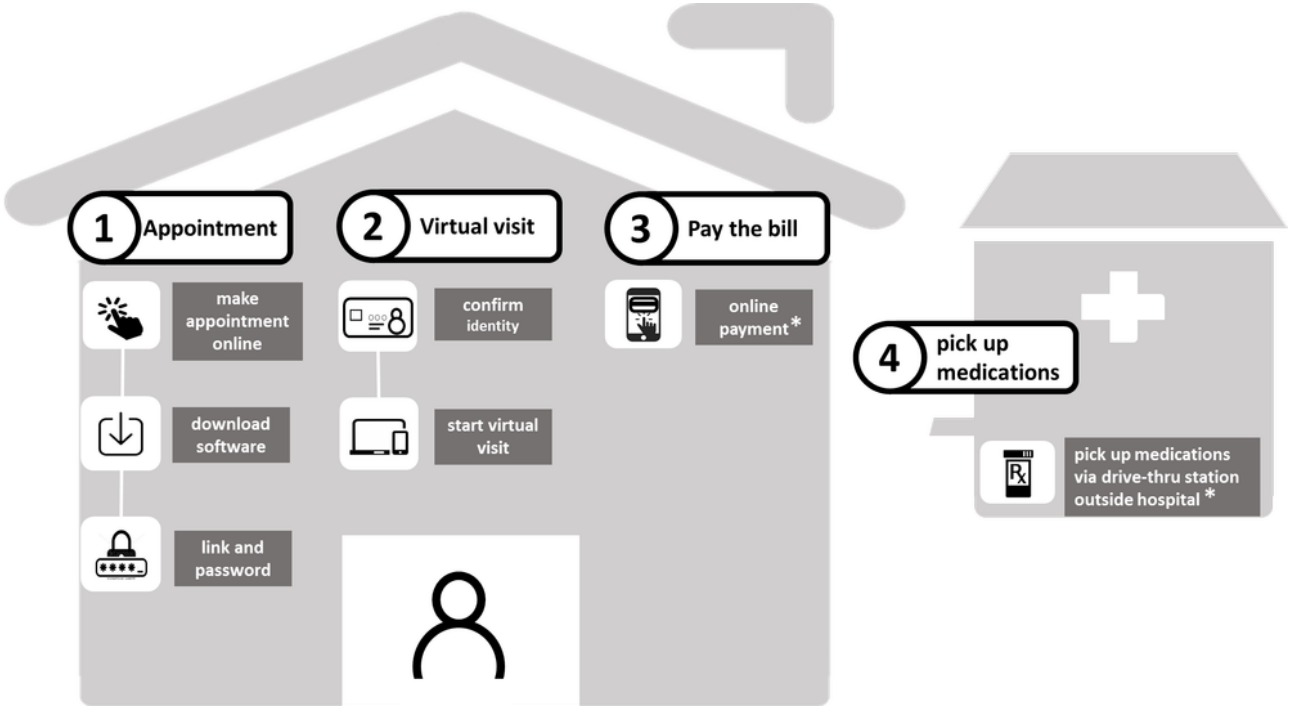
Payment

Reimbursement by the NHI for virtual clinic visits during the pandemic was confined to patients with stable chronic illnesses, such as hypertension, diabetes, and heart failure with a stable condition. Before the pandemic, these patients received long-term follow-up at our hospital every 3 months with medication prescriptions.

Appointment

Patients who agreed to use video-based virtual clinics could make an appointment online with the telehealth center in NTUH. Webex software, a teleconference link, use instructions, and a temporary username and password were then automatically emailed to the patient a day before their appointment. The workflow of the virtual clinic system is shown in Figure 1.

Figure 1. The workflow of our video-based virtual visit. (1) Appointment: eligible patients could make an appointment online and download the necessary software in advance. (2) Virtual visit: patients enter the virtual waiting room before the appointment time. Patients show their National Health Insurance card to the camera to confirm their identity. The physician then starts the virtual visit. (3) Pay the bill: patients or caregivers can pay the bill online (*or inside the hospital, according to their choice). (4) Medication pick-up: patients or caregivers pick up their medications via a drive-through station served by a pharmacist outside the hospital (*or inside the hospital, according to their choice).



Video-Based Clinic Visit

Patients entered the virtual waiting room before their virtual visit. To confirm their identity, a video-based clinic visit was used instead of a telephone call. Patients provided their NHI card (mostly with a photo) at the beginning of the video conference to confirm their identity. The physicians in our telehealth center then interviewed the patient via videoconference with full-HD quality in an isolated clinic to maximize privacy. A record of the interview was documented in our health information system during the visit by the physician.

Medication Logistics

After the video-based clinic visit, the patients or caregivers should come to the hospital to collect their prescribed medications and pay the bill. The prescribed medications were reviewed by the pharmacist after a video-based clinic visit. A rapid medication collection station that was served by a pharmacist was set up outside the hospital to minimize the exposure of those visiting. The patients or caregivers collected the prescribed medications from the station outside the hospital, in a drive-through–like service, or inside the hospital according to their choice. An online bill-paying service was provided by the hospital from around July 2021. Patients could choose to pay online or inside the hospital according to their choice. The number of pills of the prescribed medications were obtained from our electronic health database, and the medications were classified into major clinical categories.

Waiting Time and Visit Time

We calculated the time used during the virtual clinic visits by our patients. The patient waiting time for a virtual clinic visit was defined as the time from when they entered the virtual

waiting room to the initiation of the physician visit. These time points were recorded by the Webex software and retrieved from the electronic health database, retrospectively. The physician visit time was defined as the time between the initiation and the end of the physician visit. We also calculated the time used during face-to-face clinic visits during the same time period and across the same 9 specialties at our hospital during the study period. The waiting time for face-to-face clinic visits was defined as the time from when the patient registered with their NHI card after arriving for the physician visit, as recorded in our electronic medical record system. The physician visit time for face-to-face clinic visits was defined as the time between the initiation and the end of the physician visit.

Statistical Analysis

Descriptive statistics were used in this study to analyze the data. Continuous variables are expressed as mean and SD. Cross-correlation was used to illustrate the association between the number of daily confirmed COVID-19 cases in Taiwan, and daily face-to-face clinic visits and daily virtual clinic visits at our hospital. The patient waiting time and physician waiting time in our video-based clinic were compared with those in face-to-face clinics using the Wilcoxon rank sum test. A *P* value of .05 or less was considered to indicate a statistically significant difference.

Results

A total of 782 patients used the video-based virtual clinic between May and August 2021. Their mean age was 63.1 (SD 17.9) years, and 302 (38.6%) were male. The mean number of diagnoses for each patient was 3.8 (SD 2). Nine clinical departments in our hospital were involved in the video-based virtual clinic project (Table 1). The diagnosis listed was not mutually exclusive.

The medications prescribed in the virtual clinic are shown in Table 2. The internal medicine department had the highest number of virtual clinic visits, followed by the neurology and surgery departments. The medications most commonly prescribed in the virtual clinic were cardiovascular, diabetic, and gastrointestinal medications, of which cardiovascular medications comprised around one-third of all medications prescribed during this period.

The daily average number of virtual clinic visits was 13.5 (IQR 10.3-16.7) during the study period. During the same period, an average of 1886 (IQR 1750-2022) patients with chronic conditions visited our face-to-face clinic daily for long-term prescriptions. The number of face-to-face clinic visits was 2455 in the same period in 2020.

Figure 2 demonstrates the association between the number of virtual clinic patients (black line), number of face-to-face clinic patients (blue dotted line), and the number of new daily COVID-19 cases (black dashed line) in Taiwan. The outbreak began on May 12, plateaued in mid-May to early June 2021, and then subsided gradually in August 2021. The Central Epidemic Command Center raised the nationwide pandemic alert from level 2 to level 3 on May 15, 2021, and finally downgraded it to level 2 on July 27, 2021.

Table 1. Baseline characteristics of the patients visiting the virtual clinic between May and August 2021.

Demographic	Patients (N=782)
Sex (male), n (%)	302 (38.6)
Age (years), mean (SD)	63.1 (17.9)
Department of virtual clinic, n (%)	
Internal medicine	438 (56.0)
Neurology	89 (11.4)
Surgery	68 (8.7)
Family medicine	47 (6.0)
Psychiatry	43 (5.5)
Ophthalmology	40 (5.1)
Obstetrics and gynecology	31 (4.0)
Urology	23 (2.9)
Orthopedics	3 (0.4)
Most common diagnosis, n (%)	
Diabetes mellitus	40 (5.1)
Hypertension	40 (5.1)
Coronary arterial disease	38 (4.9)
Hypertensive heart disease without heart failure	24 (3.1)
Atrial fibrillation	15 (1.9)
Enlarged prostate with lower urinary tract symptoms	15 (1.9)
Hyperlipidemia	13 (1.7)
Parkinson disease	13 (1.7)
Congestive heart failure	11 (1.4)

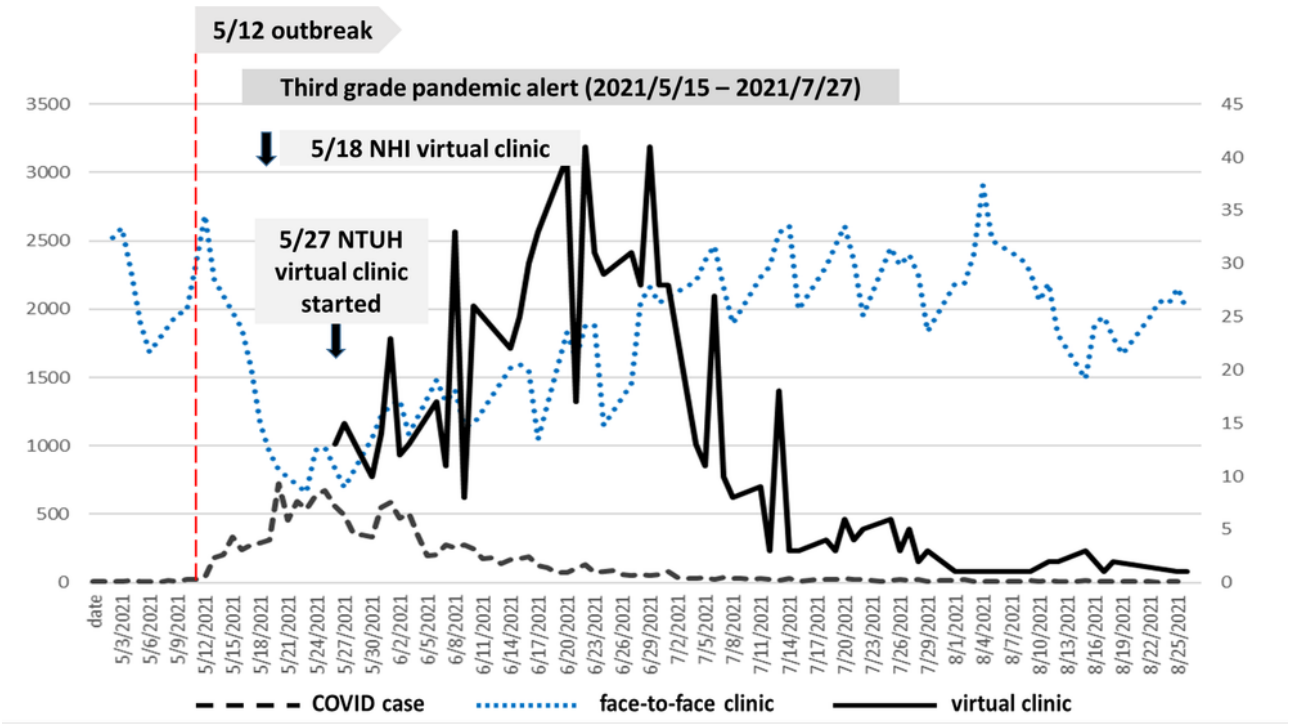
Table 2. Medications prescribed during video-based virtual clinic visits.

Category	Pill prescribed (N=123,745), n (%)
Cardiovascular	35,453 (31.5)
α-Blocker	924 (0.8)
β-Blocker	6242 (5.6)
Vasodilator	280 (0.2)
ACEI[a]	154 (0.1)
ARB[b]	7336 (6.5)
Calcium channel blocker	6749 (6.0)
Antiadrenergic agent	168 (0.1)
Antiarrhythmia	3018 (2.7)
Digoxin	68 (0.1)
Diuretic	2552 (2.3)
Lipid-lowering agent	6408 (5.7)
Other cardiovascular	1554 (1.4)
Hematologic	10,202 (9.1)
Endocrine and metabolic	17,466 (15.5)
Antihistamine	979 (0.9)
Respiratory	1917 (1.7)
Gastrointestinal	14,017 (12.5)
Immunosuppressive	684 (0.6)
Anti-infectious	1352 (1.2)
Urologic	2044 (1.8)
Dermatological	54 (0.0)
Ophthalmological	150 (0.1)
Combination pill	76 (0.1)
Insulin pen needle	1092 (1.0)
Nutritional	2806 (2.5)

[a]ACEI: angiotensin-converting enzyme inhibitor.

[b]ARB: angiotensin II receptor blocker.

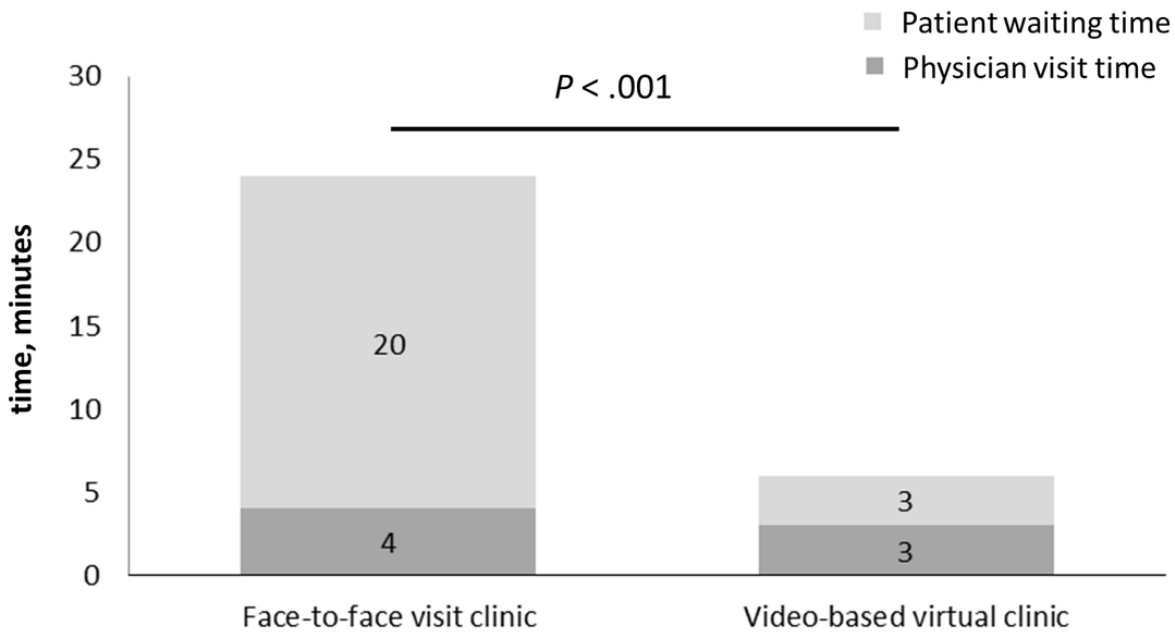
Figure 2. Time curve of daily new COVID-19 cases (gray dashed line), NTUH face-to-face clinic visits (blue dotted line), and NTUH virtual clinic visits (black line; the scale is shown on the secondary y-axis on the right side) between May 1 and August 31, 2021. After the community outbreak around May 12, the number of face-to-face clinic visits decreased immediately, while the number of virtual clinic visits increased much later, with a 20-day delay. The Central Epidemic Command Center announced a level 3 pandemic alert from May 15 to July 27, 2021. NHI: National Health Insurance; NTUH: National Taiwan University Hospital.



The number of face-to-face clinic visits at our hospital decreased immediately after the outbreak. However, there was a delay of 20 days before the number of virtual clinic visits increased. The cross-correlation between the number of face-to-face clinic visits and new daily COVID-19 cases was significant and without a lag (correlation coefficient –0.79; *P*<.001), while the cross-correlation between the number of virtual clinic visits and new daily COVID-19 cases was significant with a lag of 20 days (correlation coefficient 0.735; *P*<.001).

The median patient waiting time for the video-based virtual clinic was 3 (IQR 2-6) minutes. However, among the specialties with a higher virtual visit volume (over 80 during this period), the waiting time was significantly less compared with specialties that had a lower volume (rank sum *P*<.05). The median physician visit time was 3 (IQR 2-5) minutes. We also collected the waiting and visit duration times during the same study period and the same 9 specialties for the face-to-face clinic visits (n=99,806). The median waiting time for the face-to-face clinic was 20 (IQR 9-42) minutes, while the median visit time was 4 (IQR 2-9) minutes. The patient waiting time for video-based virtual clinic visits was significantly shorter (median 3 min vs 20 min; *P*<.001, rank sum test) compared with that for face-to-face clinic visits, while the physician visit time was also significantly shorter for video-based virtual clinic visits (median 3 min vs 4 min; *P*<.001, rank sum test) (Figure 3). This reduction in patient waiting time was significant across all 9 specialties (all *P*<.05, rank sum test).

Figure 3. The patient waiting times and physician visit times for video-based virtual clinics and face-to-face clinic visits. The median patient waiting time was 20 (IQR 9-42) minutes for face-to-face clinic visits and 3 (IQR 2-6) minutes for video-based virtual clinic visits (rank sum *P*<.001). The median physician visit time was 4 (IQR 2-9) minutes for face-to-face clinic visits and 3 (IQR 2-5) minutes for video-based virtual clinic visits (rank sum *P*<.001).



Discussion

Principal Findings

Our retrospective analysis revealed several major findings: video-based virtual clinics can be implemented rapidly after the outbreak for patients with chronic illnesses who require regular follow-up, the number of video-based virtual clinic visits at our hospital increased with a 20-day delay after the COVID-19 outbreak and was correlated with the number of daily new COVID-19 cases, and video-based virtual clinics shortened the patient waiting time significantly compared with face-to-face clinic visits. The implementation of video-based virtual clinics during the outbreak decreased unnecessary exposure of patients with chronic illnesses, who are a relatively vulnerable group. The time and cost of travel to face-to-face clinic visits can also be saved.

The implementation of virtual clinics during the COVID-19 pandemic has been reported in different countries and in different disease specialties [4-11]. The scale, rate, and barriers for the implementation of virtual clinics in different countries may not be the same. Gilbert et al [9] reported on the experience of virtual clinic implementation in a UK tertiary orthopedic center during the COVID-19 pandemic. A COVID-19 action team was established in their hospital to implement the virtual clinic visits, which were conducted via a telephone call or a videoconference. Considerable administrative, clinical, and technical resources were directed to the virtual clinic, and over 90% of the visits could be completed via the virtual clinic 1 month after its implementation. The majority of virtual clinic visits were performed via the telephone (telephone vs video, around 9 vs 1). In feedback from the patients after a virtual clinic visit, 94% using the telephone service expressed that they would like to use the virtual clinic next time, compared with only 44% of those using the videoconference. This report showed the experience of scaling up the amount of virtual clinic appointments during a severe and prolonged outbreak. However, the barriers for the implementation of videoconferencing as the method for virtual clinic visits warrant further studies.

Kim et al [11] also reported the results of teleconsultation, which was temporarily allowed in Korea during the COVID-19 pandemic. The researchers obtained teleconsultation data from the NHI claims database. From February to June 2020, a total of 567,390 cases received teleconsultation, of which 46.4% were claimed by clinics. Teleconsultations accounted for only 0.25% of the total medical use during this period. The most common major disease category for these consultations was circulatory diseases, followed by endocrine, nutritional, and metabolic diseases. From the patient number versus time curve, the increase in the number of virtual clinic visits in tertiary and teaching hospitals was delayed compared with the immediate increase in the number of appointments at primary health clinics. After the COVID-19 situation had become stable, the number of teleconsultations in hospitals decreased, but those in primary health clinics persisted. The researchers also noted that teleconsultations were provided most often to 3 types of patients: (1) those scheduled for follow-up visits, (2) those with chronic illnesses such as diabetes and hypertension, and (3) those living in high-infection areas. These findings are consistent with our results. However, the COVID-19 outbreak was more severe in Korea compared with Taiwan, which may partially explain the differences observed in the number of teleconsultations used between Korea and Taiwan.

In our study, there was a 20-day lag cross-correlation between the number of virtual clinic visits and the number of daily new COVID-19 cases. Several factors, including demand and supply, may have contributed to this phenomenon. With regards to supply, hospitals needed time to adjust to the workflow and implement the virtual clinics. The information and technology team had to be activated to redesign the processes, including appointment booking, physician-patient interviews, electronic medical records, billing, and pharmacy. With regards to demand, the patients with chronic illnesses could have obtained medications from pharmacies or primary care clinics. They may also have waited for the outbreak to subside. Hence, they may have postponed the regular follow-up visits. From our data, it is possible that the use of virtual clinics at our hospital during the outbreak was mainly driven by the severity of the outbreak. After the number of daily new COVID-19 cases decreased, the use of the virtual clinics also decreased, even though NHI reimbursement persisted.

Although our findings suggest that it is possible to implement video-based virtual clinics for the care of patients with chronic illnesses during the pandemic, the number of video-based virtual clinic visits was still low compared with the number of face-to-face clinic visits. Part of the delay in the increase in video-based virtual clinic visits may be because most patients were waiting for a rapid decrease in the daily number of new COVID-19 cases. However, other potential barriers to prevent the prompt conversion to and uptake of virtual clinics still exist [20,21]. A guide to overcoming these barriers and to facilitate the implementation of virtual clinics has been proposed [22]. Here, we discuss some of the potential barriers observed during the May 2021 outbreak at a tertiary medical center in Taiwan.

The service fee for the virtual clinic visits was reimbursed after the May 2021 outbreak. The registration fee (for the appointment) and copayment for medications from virtual clinic visits are the same as those for face-to-face clinic visits. The NHI payment therefore largely reduced the economic barrier. However, the patients needed to have equipment and access to the internet to be able to use the virtual clinics. Patients in extreme economic conditions may therefore not have been able to afford a virtual clinic visit.

To confirm the identity of a patient for a virtual clinic visit, we only allowed videoconferencing and not telephone calls. Videoconferencing has the advantage of the patient being able to confirm their identity. Furthermore, through teleconference software, the patient also has the autonomy to enter the virtual waiting room before their appointment time. In contrast, patients can simply wait for a call from the clinic if using the telephone as the modality of a virtual visit. Our data showed a shorter waiting time compared with conventional face-to-face visits. A retrospective study also showed that video visits were associated with fewer 90-day emergency department visits or hospitalizations compared with telephone visits as the modality of telemedicine for patients with heart failure, after adjusting for multiple predictors [23].

However, video-based virtual clinics require a minimum internet speed and camera resolution. Moreover, patients with chronic illnesses are mostly older adults, who may have lower levels of digital literacy. The use of teleconference equipment is not always straightforward for this patient group. Patients with hearing or cognitive dysfunction also have more difficulties in using these technologies [24]. The nurses in our telehealth center had to spend a lot of time instructing patients on how to use the software to perform the videoconference. Moreover, family members of the patients frequently had to ask for leave to stay at home and operate the video software and hardware to complete the virtual clinic visit. Video-based virtual clinics may also raise concerns regarding the security of patient information.

Performing a physical examination during a telephone-based virtual visit is almost impossible; however, it is still difficult during a video-based virtual clinic visit. A patient-assisted clinical examination guide has been proposed to help physicians evaluate physical signs during a videoconference [25]. However, only inspection and vital sign measurements can be reliably performed in a routine video-based virtual clinic visit. Novel technologies such as an electronic stethoscope, pulse oximeter, or wearable electrocardiography can only be used in certain conditions. This barrier may preclude the use of virtual clinic visits for the initial visit of a new patient [26].

Regulations in Taiwan require that a pharmacist must deliver medications directly to patients and explain their use to avoid medication error [27]. During the COVID-19 pandemic, this regulation has been upheld. Therefore, after a virtual clinic visit, patients or caregivers still had to go to the hospital to collect prescriptions and medications. In Taiwan, several hospitals including ours provide a drive-through collection method to help minimize contact between patients and pharmacists when collecting their medications. Novel pharmacy and drug delivery services are urgently needed in the near future to maximize the benefits of telehealth care during the COVID-19 pandemic [28,29].

There were several limitations to this study. First, the retrospective observational design of the study precludes any causal inference. Second, the data were collected at a tertiary medical center, which has accommodated many confirmed cases of COVID-19. The use of virtual clinics in a community hospital and primary care clinic cannot be determined from our data. Further analysis of the use of virtual clinics in other institutes is warranted to better understand the resource use and to facilitate better planning for future outbreaks. Finally, the study period was only 3 months, so the effects of current virtual clinic implementation on health outcomes over a longer time period are still not clear.

Conclusion

Our study demonstrated the implementation of video-based virtual clinics during the COVID-19 pandemic at a tertiary teaching hospital in Taiwan. Video-based virtual clinics minimized contact between patients and health workers, and saved the patients' time both in waiting time and travel to the hospital. Video-based virtual clinics can be rapidly adopted at a medical center to provide continuous care during disease outbreaks. An increase in the number of confirmed COVID-19 cases heralded a surge in the use of the virtual clinic. However, efforts are still required to reduce the barriers for virtual clinic implementation.

Authors' Contributions

YHC, HWW designed and performed the study; CCH, JKL, LTY, performed the study; CSH, performed the statistical analysis and wrote the manuscript; TPH supervised the study; YLH, commented on the manuscript and supervised the work. YHC and HWW contributed equally to the work. CSH and YLH are equal corresponding authors.

Conflicts of Interest

None declared.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733 [FREE Full text] [doi: 10.1056/NEJMoa2001017] [Medline: 31978945]
2. Coronavirus disease 2019 (COVID-19) situation report – 51. World Health Organization. 2020. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10 [accessed 2021-12-05]
3. Number of cumulative cases of coronavirus (COVID-19) worldwide from January 22, 2020 to May 15, 2022, by day. Statista. URL: https://www.statista.com/statistics/1103040/cumulative-coronavirus-covid19-cases-number-worldwide-by-day/ [accessed 2021-12-05]
4. Berlin A, Lovas M, Truong T, Melwani S, Liu J, Liu ZA, et al. Implementation and outcomes of virtual care across a tertiary cancer center during COVID-19. JAMA Oncol 2021 Apr 01;7(4):597-602 [FREE Full text] [doi: 10.1001/jamaoncol.2020.6982] [Medline: 33410867]
5. Puttasiddaiah PM, Morris S, Dwamena S, Sanu A. Attend anywhere clinic: a virtual outpatient clinic experience in otolaryngology during the COVID-19 pandemic. Laryngoscope Investig Otolaryngol 2021 Jun;6(3):586-589 [FREE Full text] [doi: 10.1002/lio2.557] [Medline: 34195381]
6. Popova D, Young K, Hobday D, Welman T, Pahal GS. Hand trauma virtual clinics prove effective during the COVID-19 pandemic. Hand (N Y) 2021 May 31:15589447211017210. [doi: 10.1177/15589447211017210] [Medline: 34053322]
7. 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 Aug 28;22(8):e21609 [FREE Full text] [doi: 10.2196/21609] [Medline: 32716898]
8. Zhao M, Qin D, Cataldo G, Sharma K, Dandwate N, Orencole MP, et al. Virtual multidisciplinary care for heart failure patients with cardiac resynchronization therapy devices during the Coronavirus Disease 2019 pandemic. Int J Cardiol Heart Vasc 2021 Jun;34:100811 [FREE Full text] [doi: 10.1016/j.ijcha.2021.100811] [Medline: 34095452]
9. Gilbert AW, Billany JCT, Adam R, Martin L, Tobin R, Bagdai S, et al. Rapid implementation of virtual clinics due to COVID-19: report and early evaluation of a quality improvement initiative. BMJ Open Qual 2020 May;9(2):e000985 [FREE Full text] [doi: 10.1136/bmjoq-2020-000985] [Medline: 32439740]
10. AlShehery MZ, Duraisamy B, Zaidi ARZ, AlShehry NF, Zaidi FZ, Rababah AA, et al. COVID-19 and palliative care services: comparative patterns of inpatient, outpatient, and consultation services in a tertiary care center in Riyadh. Cureus 2020 Dec 09;12(12):e11996 [FREE Full text] [doi: 10.7759/cureus.11996] [Medline: 33437551]
11. Kim H, Lee H, Park CS, Kim S, Cho SA, Yoo SM, et al. Preliminary results of teleconsultations temporarily allowed during the COVID-19 pandemic. Yonsei Med J 2021 Sep;62(9):850-857 [FREE Full text] [doi: 10.3349/ymj.2021.62.9.850] [Medline: 34427072]
12. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]
13. Cheng H, Chueh Y, Chen C, Jian S, Lai S, Liu D. Taiwan's COVID-19 response: timely case detection and quarantine, January to June 2020. J Formos Med Assoc 2021 Jun;120(6):1400-1404 [FREE Full text] [doi: 10.1016/j.jfma.2020.10.023] [Medline: 33191093]
14. Ng T, Cheng H, Chang H, Liu C, Yang C, Jian S, et al. Comparison of estimated effectiveness of case-based and population-based interventions on COVID-19 containment in Taiwan. JAMA Intern Med 2021 Jul 01;181(7):913-921 [FREE Full text] [doi: 10.1001/jamainternmed.2021.1644] [Medline: 33821922]
15. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. JAMA 2020 Apr 14;323(14):1341-1342. [doi: 10.1001/jama.2020.3151] [Medline: 32125371]
16. Lee Y, Hu H, Yen Y, Chu D, Yang N, Chou S, et al. Impact of the COVID-19 pandemic on the utilization of medical and dental services in Taiwan: a cohort study. J Dent Sci 2021 Oct;16(4):1233-1240 [FREE Full text] [doi: 10.1016/j.jds.2021.02.001] [Medline: 33613855]

17. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. N Engl J Med 2020 Mar 19;382(12):1177-1179 [FREE Full text] [doi: 10.1056/NEJMc2001737] [Medline: 32074444]
18. Taiwan can help - National Health Insurance's contribution in combating COVID-19. Ministry of Health and Welfare. URL: https://covid19.mohw.gov.tw/en/cp-4778-53691-206.html [accessed 2021-12-05]
19. Taiwan epidemic report. COVID-19 Dashboard. URL: https://covid-19.nchc.org.tw/dt_005-covidTable_taiwan.php [accessed 2021-12-05]
20. Alkureishi MA, Choo Z, Rahman A, Ho K, Benning-Shorb J, Lenti G, et al. Digitally disconnected: qualitative study of patient perspectives on the digital divide and potential solutions. JMIR Hum Factors 2021 Dec 15;8(4):e33364 [FREE Full text] [doi: 10.2196/33364] [Medline: 34705664]
21. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, Van Spall HGC. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74 [FREE Full text] [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]
22. Smith WR, Atala AJ, Terlecki RP, Kelly EE, Matthews CA. Implementation guide for rapid integration of an outpatient telemedicine program during the COVID-19 pandemic. J Am Coll Surg 2020 Aug;231(2):216-222.e2 [FREE Full text] [doi: 10.1016/j.jamcollsurg.2020.04.030] [Medline: 32360960]
23. Sammour Y, Spertus JA, Shatla I, Gupta SK, Nassif ME, Magalski A, et al. Comparison of video and telephone visits in outpatients with heart failure. Am J Cardiol 2021 Nov 01;158:153-156. [doi: 10.1016/j.amjcard.2021.08.008] [Medline: 34470705]
24. Kalicki AV, Moody KA, Franzosa E, Gliatto PM, Ornstein KA. Barriers to telehealth access among homebound older adults. J Am Geriatr Soc 2021 Sep;69(9):2404-2411 [FREE Full text] [doi: 10.1111/jgs.17163] [Medline: 33848360]
25. Benziger CP, Huffman MD, Sweis RN, Stone NJ. The telehealth ten: a guide for a patient-assisted virtual physical examination. Am J Med 2021 Jan;134(1):48-51 [FREE Full text] [doi: 10.1016/j.amjmed.2020.06.015] [Medline: 32687813]
26. Lee PS, Koo S, Panter S. The value of physical examination in the era of telemedicine. J R Coll Physicians Edinb 2021 Mar;51(1):85-90. [doi: 10.4997/JRCPE.2021.122] [Medline: 33877145]
27. Pharmaceutical Affairs Act enforcement rules. Laws & Regulations Database of the Republic of China (Taiwan). URL: https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=L0030002 [accessed 2021-12-05]
28. Peláez Bejarano A, Villar Santos P, Robustillo-Cortés M, Sánchez Gómez E, Santos Rubio MD. Implementation of a novel home delivery service during pandemic. Eur J Hosp Pharm 2021 Nov;28(Suppl 2):e120-e123 [FREE Full text] [doi: 10.1136/ejhpharm-2020-002500] [Medline: 33115800]
29. Unni EJ, Patel K, Beazer IR, Hung M. Telepharmacy during COVID-19: a scoping review. Pharmacy (Basel) 2021 Nov 11;9(4):183 [FREE Full text] [doi: 10.3390/pharmacy9040183] [Medline: 34842823]

Abbreviations

ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification
NHI: National Health Insurance
NTUH: National Taiwan University Hospital

Edited by T Leung; submitted 10.03.22; peer-reviewed by J Li; comments to author 01.04.22; revised version received 26.04.22; accepted 20.05.22; published 10.06.22.

Please cite as:
Chen YH, Wu HW, Huang CC, Lee JK, Yang LT, Hsu TP, Hung CS, Ho YL
The Logistics of Medication and Patient Flow in Video-Based Virtual Clinics During a Sudden COVID-19 Outbreak in Taiwan: Observational Study
Interact J Med Res 2022;11(1):e37880
URL: https://www.i-jmr.org/2022/1/e37880
doi:10.2196/37880
PMID:35687404

Original Paper

The Effects of Intravenous Iron Infusion on Preoperative Hemoglobin Concentration in Iron Deficiency Anemia: Retrospective Observational Study

Guy Nicholls[1], MBBS; Rajan Mehta[1], BSc, MBBS; Karen McVeagh[2], BSc; Matthew Egan[1], MBBS

[1]Anaesthetics and Critical Care, Southend University Hospital, Southend on Sea, United Kingdom

[2]East of England Ambulance Service, Essex, United Kingdom

Corresponding Author:
Guy Nicholls, MBBS
Anaesthetics and Critical Care
Southend University Hospital
Prittlewell Chase
Southend on Sea, SS0 0RY
United Kingdom
Phone: 44 7379867037
Email: Guy.a.nicholls91@gmail.com

Abstract

Background: An iron infusion pathway using Ferrinject (ferric carboxymaltose) was implemented at Southend University Hospital for preoperative surgical patients with iron deficiency anemia undergoing major surgery. This was based on a treatment algorithm proposed by Munting and colleagues according to the international consensus statement on perioperative management of anemia and the UK National Institute for Health and Care Excellence (NICE) guidelines. These guidelines state that intravenous iron is indicated when oral iron is poorly tolerated or ineffective, there is insufficient time to surgery, or due to a functional iron deficiency.

Objective: The objective of this study was to evaluate the change in adult hemoglobin (Hb) concentration (g/L) after Ferrinject infusion at the time of surgery.

Methods: Data were retrospectively collected on all surgical patients that received an iron infusion preoperatively for iron deficiency anemia from July 2019 to April 2020. Nonsurgical, obstetric, and pediatric patients, and those without a postinfusion Hb level measurement were excluded. Data collected included the Hb, ferritin, and transferrin levels pre and postinfusion; correct dose of intravenous iron received; and any adverse reactions noted.

Results: Thirty-two surgical patients with iron deficiency anemia received intravenous iron between July 2019 and April 2020 prior to surgery. The average pre and post iron infusion Hb concentration across the cohort was 97 g/L and 114 g/L, respectively (18% increase; *P*=.001). Two (6%) patients had a posttransfusion Hb level ≥130 g/L prior to surgery after infusion. Nine patients had both a pre and postinfusion ferritin level recorded, which showed an increase from 12 ng/mL preinfusion to 94 ng/mL postinfusion (*P*=.02). Twenty-three (72%) patients did not receive the full dose of intravenous iron based on their Hb level and weight. Twenty-four (75%) patients received an iron infusion >2 weeks prior to surgery and the other 8 (25%) patients received the infusion <2 weeks before their surgery. There was an average increase in Hb of 22% (21 g/L, 95% CI 13-28) and 5% (5 g/L, 95% CI 1-10), respectively, across the two groups (*P*=.03). There were no documented adverse reactions to intravenous iron.

Conclusions: Intravenous iron is an effective intervention to improve the Hb concentration in patients with iron deficiency anemia despite the majority of patients not receiving the full dose based on their baseline Hb level and weight. Increasing the interval time between infusion and surgery was associated with a greater increase in Hb, with only a minimal increase observed if given less than 2 weeks prior to surgery.

(Interact J Med Res 2022;11(1):e31082) doi:10.2196/31082

KEYWORDS

anemia; perioperative medicine; anesthetics; preoperative; perioperative; surgery; hemoglobin; hemoglobin concentration; iron deficiency; intravenous; blood

Introduction

Preoperative optimization is fundamental to improving surgical outcomes, with correction or improvement in iron deficiency anemia (IDA) being an important aspect in the preoperative management of patients undergoing major elective surgery [1]. This has resulted in a large drive across the United Kingdom for developing perioperative iron infusion pathways. Large studies have shown that approximately 40% of patients presenting for major surgery were anemic, with iron deficiency accounting for 80% of these cases [2].

We conducted a preliminary audit to determine the prevalence of IDA in patients undergoing major surgery at Southend University Hospital (Southend on Sea, UK) in 2018. We identified that 32% (52/187) of patients were anemic, when defined as a hemoglobin (Hb) level of <130 g/L for both men and women. Out of the 15 patients with available hematinics in the audit, 80% were shown to have IDA.

This preliminary audit in conjunction with the body of evidence of the benefits of correcting anemia prior to surgery led to the development of a local preoperative iron infusion pathway in 2019 for IDA. This was guided by the international consensus statement on the perioperative management of anemia and National Institute for Health and Care Excellence (NICE) guidelines [3,4]. We acknowledge that since the setting up of the pathway, the release of the results of the PREVENTT trial has raised a debate as to whether iron infusions to increase Hb levels improve clinical outcomes, although the average increase in Hb levels in the study was only 5 g/L [5,6].

In this local pathway, patients who were identified as having IDA during preoperative assessment were given an intravenous iron infusion in the form of Ferrinject (ferric carboxymaltose). The dose of Ferrinject, in grams, is based on weight and the starting Hb level (g/L). For an Hb level <100 g/L, the dose is 1.5 grams if the body weight is <70 kilograms and the dose is 2 grams if the body weight is >70 kilograms. If the Hb level is >100 g/L, the dose is 1 gram if the body weight is <70 kilograms and the dose is 1.5 grams if the body weight is >70 kilograms. A maximum dose of 1 gram can be given per infusion [6].

The objective of this study was to assess the impact of this formalized IDA preoperative pathway on the change in Hb concentration for patients with IDA undergoing major elective surgery.

Methods

Adult surgical patients with IDA who received an intravenous iron infusion preoperatively through the pathway following its introduction in 2019 over a 9-month period were included in the study. Nonsurgical, obstetric, and pediatric patients were excluded. Patients were also excluded if there were no documented post iron infusion Hb levels.

Data were collected using electronic clinical records and laboratory results. Data collected included: patient demographics, pre and post iron infusion blood results (Hb, ferritin, and transferrin saturations), timing of infusion prior to surgery, whether the correct dose of Ferrinject was administered (as per the guideline), and if any adverse reactions to the infusion occurred. Preassessment proformas were assessed to determine if patients received oral iron in the preoperative period and the hospital's blood bank checked to see if any blood transfusions were administered in the period leading up to surgery.

The primary objective was to assess the change in Hb concentration following an intravenous iron infusion of Ferrinject for patients with IDA undergoing major elective surgery. This was further analyzed to measure the change in Hb with respect to the length of time that the infusion was administered prior to surgery to assess the effects of earlier administration.

The secondary measures were the change in ferritin concentration postinfusion, proportion of patients who received the correct dose of intravenous iron (Ferrinject), proportion of patients with an Hb level >130 g/L at the time of surgery, and whether there were any adverse effects to the infusion.

Statistical analysis was performed using a paired Student *t* test following assessment for normality using the Kolmogorov-Smirnov test and the 95% CI. Statistical significance was defined as $P<.05$.

This study was approved by the hospital Trust's Research and Development department. As no identifiable patient information was collected, patient consent was not required.

Results

Thirty-five adult surgical patients received an intravenous iron infusion through the preoperative IDA pathway between July 2019 and April 2020. Three were excluded due to unavailable data. Baseline characteristics are summarized in Table 1. The median age was 65 years and the majority were men. The most common type of surgery performed by specialty was general surgery, followed by urology, gynecology, orthopedics, and combined general surgery/gynecology. There was no documentation of any patient receiving oral iron or a blood transfusion in the preoperative period.

Table 1. Baseline characteristics of the patients.

Characteristics	Value
Number of infusions, n	35
Cases excluded, n	3
Total cases analyzed, n	32
Age (years), mean (range)	65 (25-88)
Baseline Hb[a] (g/L), mean	97
Male, n (%)	17 (53)
Pre and posttransfusion ferritin available, n	9
Preoperative oral iron, n	0
Preoperative blood transfusion, n	0
Cases by specialty, n (%)	
General surgery	18 (56)
Gynecology	5 (16)
General surgery/gynecology combined	1 (3)
Urology	6 (19)
Orthopedics	2 (6)

[a]Hb: hemoglobin.

For the primary outcome, the mean Hb level was increased from pre to post iron infusion by 18% on average for all cases. Significant increases were also found in the subanalyses for the <2 weeks group, 2 to 8 weeks group, and >8 weeks group. Overall, 75% of patients received an iron infusion >2 weeks prior to their surgery date, with an average increase in the Hb level of 22% compared to an increase of only 5% for those that received the iron infusion <2 weeks prior to surgery, representing a significant difference (Table 2).

With respect to the secondary outcomes (Table 3), preinfusion and postinfusion ferritin values were only available for 9 patients, demonstrating an overall mean increase. Three of the 32 patients exhibited a decrease in Hb at the time of surgery. Nine out of the 32 patients received the correct (full) dose of Ferrinject, and most of these patients were in the 2-8 weeks group, followed by the <2 weeks group and the >8 weeks group. Two patients had an Hb level of >130 g/L at the time of surgery post iron infusion, and there were no documented adverse effects to any of the infusions.

Table 2. Results for the primary outcome.

Cases	Preinfusion Hb[a] (g/L)	Postinfusion Hb (g/L)	Change in Hb (95% CI)	*P* value
All cases	97	114	17 (11-23)	<.001
<2 weeks	103	108	5 (1-10)	.03
2-8 weeks	99	112	13 (2-24)	.03
>8 weeks	91	121	30 (21-39)	<.001
>2 weeks	95	116	21 (13-28)	<.001
>2 weeks vs >8 weeks	N/A[b]	N/A	N/A	.03

[a]Hb: hemoglobin.

[b]N/A: not applicable.

Table 3. Results for the secondary outcomes (N=32).

Outcome	Value
Ferritin level (ng/mL), mean (95% CI)[a]	
Preinfusion	12 (1.3-22)
Postinfusion	94 (32-156)
Increase	82 (20-145)
Timing of intravenous iron before surgery, n (%)	
<2 weeks	8 (25)
2 to 8 weeks	13 (41)
>8 weeks	11 (34)
>2 weeks	24 (75)
Patients receiving a full dose of intravenous iron, n (%)	
Overall	9 (28)
<2 weeks	2 (25)
2 to 8 weeks	5 (39)
>8 weeks	2 (18)
Patients with a decrease in Hb[b], n	3 (9)
Patients with Hb >130 g/L postinfusion, n (%)	2 (6)

[a]*P*=.02; data available for 9 patients.

[b]Hb: hemoglobin.

Discussion

Principal Findings

Although we were only able to identify 35 infusions over the 9-month study period, our results demonstrate that earlier administration of intravenous iron results in a far greater rise in Hb levels before surgery than when the infusion is given closer to the surgery date, in particular less than 2 weeks before surgery. Although this result appears to be intuitive, we felt it was important to assess our service to monitor the potential expected change in the Hb levels of our patients. When given less than 2 weeks before surgery, if only a 5 g/L rise is expected, then this is very unlikely to provide a benefit in the perioperative period compared to the 21 g/L rise found for the >2 weeks group, which has a much higher chance of providing early perioperative benefits and may also benefit the patient postoperatively.

As intravenous iron infusions have gained motion in becoming a cornerstone in patient blood management, the PREVENTT trial has raised important questions as to whether there are any improvements in clinical outcomes in treating anemia before surgery with intravenous iron. The study, published in October 2020 [5], showed no decrease in the need for blood transfusions or mortality compared to placebo, although the median time from infusion to surgery was only 15 days and only showed a small increase in Hb levels of 5 g/L. This may not be a sufficient increase to translate into a clinical benefit, particularly as studies into improving Hb levels preoperatively in other specialties have shown benefits, such as in orthopedics and cardiac surgery [7,8]. It is also important to note that patients with an Hb level of less than 90 g/L were excluded from this study and the outcomes were only available for major abdominal surgery.

Patients that receive their infusion less than 2 weeks before their surgery are most likely to be urgent surgical cancer patients. Although the PREVENTT trial may have shown no decrease in mortality or units of red cells transfused, these patients may benefit postoperatively from improved iron stores and the ability to increase their Hb level. One of the positive signals from PREVENTT was that improving Hb levels perioperatively may decrease the risk of being readmitted to hospital; however, further studies are needed to validate this finding. Moreover, if a larger rise in Hb is required to provide an early perioperative benefit, then patients awaiting elective surgery—such as in elective orthopedics where waiting times far exceed 8 weeks—may show a positive benefit, as evidenced by the increase in Hb of approximately 30 g/L for the group who received an infusion over 8 weeks before surgery.

There were three patients that showed a drop in Hb at the time of surgery. All of these patients were undergoing major cancer surgery (two were resections for bowel cancer and the other was a cystectomy); thus, we can only postulate as to whether this was due to failure of treatment, ongoing blood loss, or another underlying mechanism.

In our hospital, we have been using Ferrinject for our intravenous iron infusions, which has a maximal dose of 1 gram per infusion. Dosing is based on the starting Hb concentration and ideal body weight [9]. The majority of patients will require two infusions to receive the correct dose of intravenous iron with Ferrinject. Only a patient with a starting Hb >100 g/L and a body weight <70 kilograms can receive the full dose from one

infusion as it would be 1 gram. A solution to this would be to switch to using ferric derisomaltose (Monofer) where a dose of up to 2 grams can be given per infusion [10]. Thus, switching to Monofer, in which the full intravenous iron dose can be given in a single administration, would provide cost savings to our service through not requiring two sessions to administer the full dose, thereby freeing up clinic time to provide additional treatments and easing the burden on patients in not having to come back for a second infusion. The raw cost of a vial of Ferrinject and Monofer is very similar. Calculations from the British National Formulary estimate a dose of 1, 1.5, and 2 grams costing £154.23 (≈US $230) vs £169.50 (≈US $230), £250 (≈US $338) vs £254 (≈US $344), and £309 (≈US $418) vs £339 (≈US $458), respectively [11].

Despite concerns raised regarding the safety profile of intravenous iron in the past, it was a positive finding that there no were documented adverse effects in any of the 32 infusions. Although our study is too small to make any meaningful conclusion on safety, a meta-analysis published in 2015 including 10,000 patients treated with different forms of iron replacement also found no increased risk of serious adverse effects with intravenous iron [12].

Out study does have several limitations, in particular the smaller than expected sample size after the preliminary audit. There is also the possibility of confounding factors; although we did attempt to remove these by checking if patients were taking oral iron or received a preoperative blood transfusion from the hospital records and prescriptions, we cannot completely rule out that they were not taking any iron supplementation that had not been prescribed, changed their dietary intake, or received a blood transfusion at another hospital.

Conclusions

With this study, we believe that we have shown that intravenous iron is an effective treatment option for increasing Hb levels in IDA when given an appropriate timeframe to take effect and despite the majority of patients not receiving the full dose. An iron infusion should ideally be given greater than 2 weeks before surgery to achieve a clinically significant increase in the preoperative Hb concentration. At our hospital, a large proportion of surgical patients with IDA are not referred for intravenous iron and those that are tend to be underdosed when using Ferrinject. Our service will likely benefit from switching to Monofer or by improving pathways to allow patients to receive two infusions before their surgery.

Further study is required to fully quantify the rate of Hb increase and to determine whether increasing Hb levels preoperatively and increasing postoperative iron stores effectively decreases the number of blood transfusions required and improves postoperative outcomes.

Conflicts of Interest

None declared.

References

1. Musallam K, Tamim H, Richards T, Spahn DR, Rosendaal FR, Habbal A, et al. Preoperative anaemia and postoperative outcomes in non-cardiac surgery: a retrospective cohort study. Lancet 2011 Oct 15;378(9800):1396-1407. [doi: 10.1016/S0140-6736(11)61381-0] [Medline: 21982521]
2. Munting K, Klein AA. Optimisation of pre-operative anaemia in patients before elective major surgery - why, who, when and how? Anaesthesia 2019 Jan;74(Suppl 1):49-57. [doi: 10.1111/anae.14466] [Medline: 30604424]
3. Muñoz M, Acheson A, Auerbach M, Besser M, Habler O, Kehlet H, et al. International consensus statement on the peri-operative management of anaemia and iron deficiency. Anaesthesia 2017 Feb;72(2):233-247. [doi: 10.1111/anae.13773] [Medline: 27996086]
4. Blood Transfusion NICE Guideline NG24.: National Institute For Health and Care Excellence; 2015 Nov 18. URL: https://www.nice.org.uk/guidance/ng24 [accessed 2020-07-01]
5. Richards T, Baikady R, Clevenger B, Butcher A, Abeysiri S, Chau M, et al. Preoperative intravenous iron to treat anaemia before major abdominal surgery (PREVENTT): a randomised, double-blind, controlled trial. Lancet 2020 Oct;396(10259):1353-1361. [doi: 10.1016/S0140-6736(20)31539-7]
6. Vifor Pharma. Ferinject (ferric carboxymaltose) - Summary of Product Characteristics. emc. 2021 Jul 27. URL: https://www.medicines.org.uk/emc/product/5910 [accessed 2021-07-27]
7. Kotzé A, Carter L, Scally AJ. Effect of a patient blood management programme on preoperative anaemia, transfusion rate, and outcome after primary hip or knee arthroplasty: a quality improvement cycle. Br J Anaesth 2012 Jun;108(6):943-952 [FREE Full text] [doi: 10.1093/bja/aes135] [Medline: 22593128]
8. Rineau E, Chaudet A, Chassier C, Bizot P, Lasocki S. Implementing a blood management protocol during the entire perioperative period allows a reduction in transfusion rate in major orthopedic surgery: a before-after study. Transfusion 2016 Mar;56(3):673-681. [doi: 10.1111/trf.13468] [Medline: 26748489]
9. Ferinject (ferric carboxymaltose) dosing and administration summary. Surrey and Sussex NHS. URL: https://surreyandsussex.nhs.uk/wp-content/uploads/2013/04/Ferinject.pdf [accessed 2021-07-07]
10. Pharmacosmos UK Ltd. Monofer - Solution for injection/infusion - Summary of Product Characteristics. emc. 2020 Sep 22. URL: https://www.medicines.org.uk/EMC/medicine/23669 [accessed 2021-07-01]

11. Ferric Derisomaltose. British National Formulary. 2021 Nov 26. URL: https://bnf.nice.org.uk/drug/ferric-derisomaltose.html [accessed 2021-07-07]
12. Avni T, Bieber A, Grossman A, Green H, Leibovici L, Gafter-Gvili A. The safety of intravenous iron preparations: systematic review and meta-analysis. Mayo Clin Proc 2015 Jan;90(1):12-23. [doi: 10.1016/j.mayocp.2014.10.007] [Medline: 25572192]

Abbreviations

Hb: hemoglobin
IDA: iron deficiency anemia
NICE: National Institute for Health and Care Excellence

Edited by A Mavragani; submitted 08.06.21; peer-reviewed by B Giles, M Salimi; comments to author 09.07.21; revised version received 30.09.21; accepted 13.12.21; published 03.02.22.

Please cite as:
Nicholls G, Mehta R, McVeagh K, Egan M
The Effects of Intravenous Iron Infusion on Preoperative Hemoglobin Concentration in Iron Deficiency Anemia: Retrospective Observational Study
Interact J Med Res 2022;11(1):e31082
URL: https://www.i-jmr.org/2022/1/e31082
doi:10.2196/31082
PMID:35113024

Original Paper

Occurrence Patterns of Traumatic Brain Injury Within the Emergency Department and Internal Screening Process Efficacy During the COVID-19 Pandemic: Retrospective Analysis

Tapasvini Anmol Paralkar[1*], BS; Phoebe Lay[1*], BSA; Sawyer Stubbs[1,2], BSA; Syed Hadi Ahmed[3], BBA; Minha Ghani[4], BS; Nico Osier[2,5], BSN, BS, PhD

[1]College of Natural Sciences, The University of Texas at Austin, Austin, TX, United States

[2]Department of Neurology, Dell Medical School, Austin, TX, United States

[3]McCombs School of Business, The University of Texas at Austin, Austin, TX, United States

[4]College of Liberal Arts, The University of Texas at Austin, Austin, TX, United States

[5]School of Nursing, The University of Texas at Austin, Austin, TX, United States

[*]these authors contributed equally

Corresponding Author:
Nico Osier, BSN, BS, PhD
School of Nursing
The University of Texas at Austin
1710 Red River Street
Austin, TX, 78712
United States
Phone: 1 (512) 471 7913
Email: nicoosier@utexas.edu

Abstract

Background: Traumatic brain injury (TBI) is one of the leading causes of death in pediatric patients. Continued recruitment of pediatric TBI participants into a biobank amidst the COVID-19 pandemic not only necessitates adaptive changes to traditional recruitment methods but also requires an evaluation of emergency department (ED) utilization by TBI-presenting patients.

Objective: The primary objective of this exploratory retrospective study was to evaluate pediatric TBI-related ED utilization during the pandemic. The secondary objective was to appraise the efficacy of the research team's internal screening processes.

Methods: Potential participants (ie, individuals who met all inclusion criteria and would be approached by a consenter) were screened from an ED's electronic health record system. Data regarding their visit were recorded in a Health Insurance Portability and Accountability Act–compliant manner, which were cleaned through Google Sheets. Cleaned data were then coded as either a screening variable or a hospital utilization variable to examine the effects of the pandemic on internal operations and hospital utilization patterns. The variables were compared between select months during the pandemic in 2020 to analogous months in 2019 in the R programming language via the two-sample Student *t* test and the Mann-Whitney-Wilcoxon rank-sum test.

Results: The sample (N=2321) consisted of 1245 entries from 2019 and 1076 entries from 2020. A significantly greater proportion of potential participants (P<.001) were identified in 2020 (222/633, 35.1%) than in 2019 (195/908, 21.4%). A significantly greater proportion of potential participants (P<.001) had a visit reason indicative of a TBI in 2020 (181/222, 81.5%) than in 2019 (103/195, 52.8%). A significantly greater proportion of these injuries (P=.02) occurred inside (39/181, 21.5%) in 2020 than in 2019 (11/103, 10.7%). No significant difference was found across the mechanism of injury categories reported for potential participants between 2019 and 2020. Potential participants were significantly older (P=.006) in 2019 (mean 8.93 years) than in 2020 (mean 7.31 years). Screeners spent significantly longer (P=.03) to identify potential participants in March 2020 (55 minutes) than in March 2019 (32 minutes), but spent significantly less time (P=.01) to do so in July 2020 (22 minutes) than in July 2019 (42 minutes). Screening coverage was significantly lower (P<.001) in March 2020 (241.8 hours) than in March 2019 (346.5 hours). Screening coverage was significantly greater (P<.001) in April 2020 (611.5 hours) and July 2020 (513.5 hours) than in April 2019 (470.5 hours) and July 2019 (404.3 hours), respectively.

Conclusions: There was a significant increase in the rate of incoming TBI cases to the ED during the COVID-19 pandemic, warranting continued enrollment with added safety measures. Additionally, refinement of internal processes improved the accuracy

of data collection. As demonstrated in this study, researchers can leverage ongoing data collection to facilitate process improvements and evaluate the impact of unexpected global events on their research.

(Interact J Med Res 2022;11(1):e29513) doi:10.2196/29513

KEYWORDS

COVID-19; coronavirus; pandemic; clinical recruitment; traumatic brain injury; children; participant-focused; recruitment; enrollment; digital screening; brain; EHR; electronic health record; database

Introduction

Traumatic brain injury (TBI) is defined as an alteration in brain function or pathology caused by an external force [1]. TBIs most frequently result from contact sport injuries, falls, and motor vehicle accidents (MVAs) [2]. The clinical diagnosis of TBI is made through observation of symptoms, including neurological deficits, loss of consciousness, memory loss surrounding the event, or alteration in mental state at the time of injury [2]. TBIs are more common in the age groups of 0-4 years, 15-19 years, and ≥75 years, indicating that a large majority of TBI occurrences are found in the pediatric population; notably, TBI is the leading cause of death or other negative outcomes in the pediatric population [3].

Over 812,000 emergency department (ED) visits for pediatric TBI occurred in 2014, suggesting that many parents or guardians may bring their child to the ED first, instead of making an appointment with their primary health care provider [4]. Additionally, significant increases in ED visits for pediatric TBI suggest that there are more focused efforts in referring children with suspected TBI primarily to the ED [5]. However, there remains an important gap regarding how to maximize recruitment of ED patients into pediatric TBI biobanks. This knowledge gap has only widened with the development of COVID-19. The parent study, which focuses on enrollment of pediatric TBI patients into biobanks, initially halted recruitment because in-person enrollment was critical for proper consent and biospecimen procurement; however, the research team has maintained existing digital screening efforts to identify potential participants, as described in the Methods section below. It is critical to continue TBI research in pediatric populations while additionally examining the effects of the pandemic on health system utilization.

Emerging research is exploring the effect of the pandemic on health system utilization. One recent study showed that the pandemic delayed access to pediatric care in Italy, with parental hesitation surrounding viral exposure being a commonly reported deterrent [6]. Similarly, an Austrian observational study noted an unexpected decline in hospital admissions for people with acute coronary syndrome during the COVID-19 pandemic [7]. The authors of that study posited that the observed decrease was influenced by several factors, including fear of infection and strict stay-at-home orders. The reduced use of medical services during a public health crisis is not a new phenomenon. The impact of COVID-19 on medical admissions and research recruitment parallels that of the experience during the 2003 severe acute respiratory syndrome (SARS) outbreak. A Taiwanese study reported large reductions in inpatient care expenditures at the height of the 2003 SARS outbreak, followed by a return to usual levels toward the end of the pandemic, ultimately suggesting that the fear of disease influenced the degree to which people sought care [8]. However, despite these studies, there are insufficient data regarding how the COVID-19 pandemic is specifically impacting the rates of TBI-presenting patients in the ED in the United States.

It is important to understand the effects of the pandemic on ED utilization because pediatric TBI research is a tremendously understudied field. In the parent study, pediatric TBI patients were recruited into a biobank. The biomarkers of these participants were analyzed following blood sample collection. Recruitment of patients occurred within the ED and the research team was unsure how the pandemic would affect the presentation of eligible cases. Furthermore, the COVID-19 pandemic may impact the occurrence of TBI in ways that are not yet understood. For example, the cancellation of school sporting events could lead to a reduction in TBI occurrence (eg, cancellation of contact sports); however, children staying at home due to school closures could lead to an increase in TBI occurrence (eg, rough-housing). There remains a gap in the current literature studying the aforementioned issue surrounding ED utilization in the context of TBI occurrence patterns. To address this gap, the purpose of this study was two-fold. The primary objective was to evaluate changes in ED utilization operationalized by (1) occurrence, (2) location of injury, (3) mechanism of injury (MOI), and (4) age of TBI-presenting patients through relevant electronic health records (EHRs) screened during the pandemic (2020) and the year prior (2019). The secondary objective was to evaluate the efficacy of the research team's internal screening processes amidst the pandemic.

Methods

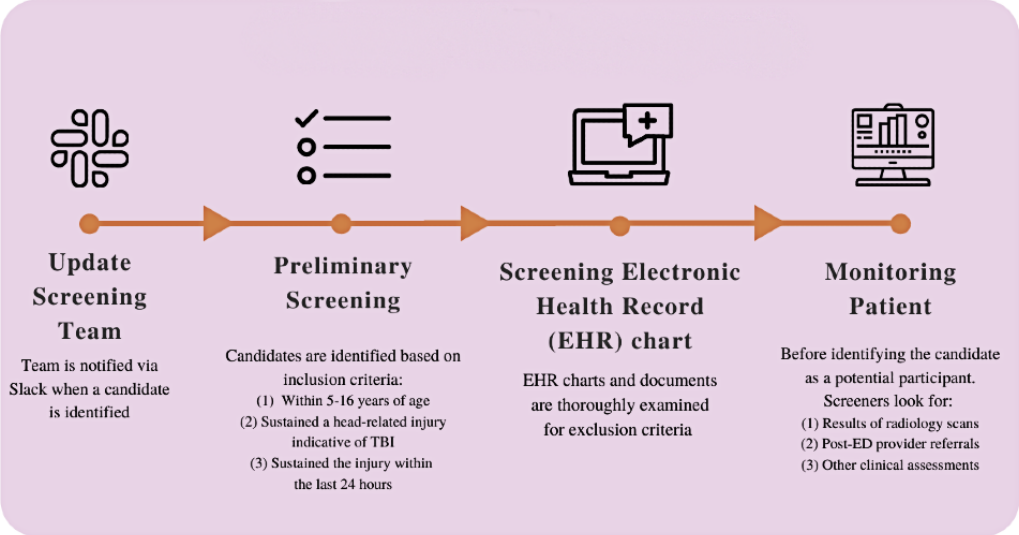
Ethics Committee Approval

Approval was obtained from The University of Texas at Austin Institutional Review Board (protocol number 2018-04-0018).

Recruitment

This study is a secondary analysis from a larger parent study; the detailed recruitment process is described in Section 2.1 in a previous publication [9] and is resummarized schematically in Figure 1. The parent study, which actively recruited participants from July 2018 to March 2020, focuses on advancing the pediatric TBI knowledge base by analyzing biomarkers in the blood samples of individuals 5-16 years of age who have sustained a TBI and uses pediatric orthopedic injuries as controls.

Figure 1. Schematic outlining the 4-part screening process: (1) communication between screener and consenter via Slack, (2) preliminary screening of emergency department (ED) census, (3) screening the EHR, and (4) continued monitoring for relevant updates (eg, radiology scans, referrals). TBI: traumatic brain injury.



Enrollment entails a two-step process: screening and consenting. Screeners and consenters obtain Health Insurance Portability and Accountability Act (HIPAA) and Collaborative Institutional Training Initiative certifications prior to accessing patient records. The visit reasons and individual patient charts from the ED census and individual EHRs were used to screen 2321 candidates between March 2019 and August 2020. Relevant data collected from the months of March, April, and July were compared across the 2 years. Screening inclusion criteria included (1) aged 5-16 years, (2) sustained a head injury, and (3) the injury occurred within the last 24 hours. If a candidate did not meet these criteria, they were excluded from any subsequent steps of the recruitment process, as outlined in detail in the previous publication [9]. The charts of candidates deemed potentially eligible were opened in compliance with the HIPAA. A form is completed for screened individuals that reflects when and why the chart was opened. These form responses are aggregated and made available for data abstraction as needed.

Once a candidate is identified, the screener updates any other on-call screeners of the newly identified candidate via Slack, a channel-based messaging platform. Communication via Slack allows for any uncertainties to be resolved, prevents miscommunication with other screeners on-call, and avoids duplicate responses. Under normal circumstances, the consenter on-call will also be notified of a candidate via a screener's Slack message. The screener and consenter may continue a private-messaging conversation to exchange any further necessary communication in a HIPAA-compliant manner. Subsequently, under normal circumstances, the consenter begins to collect data for the parent study. This consists of three time points: (1) baseline, the day participants were consented and enrolled into the study; (2) 3 months postinjury and enrollment; and (3) 6 months postinjury and enrollment. At each time point, the participant and their parent or guardian answered questions from the same five Patient-Reported Outcomes Measurement Information System (PROMIS) scales and only the participants additionally answered questions from three Quality of Life in Neurological Disorders (Neuro-QoL) scales. In addition, the participant provided a blood sample once at baseline and a saliva sample once at 6 months. However, the COVID-19 pandemic required an adapted recruitment strategy, namely the continuation of screening and the halting of consenting practices and participant enrollment.

COVID-19 Implications

COVID-19 was designated a global pandemic by the World Health Organization (WHO) on March 11, 2020. The recruitment site affiliated with the parent study halted in-person research activities that did not provide a direct benefit to participants on March 13, 2020. In addition, the National Institutes of Health recommended limiting research-related study visits and nonessential travel [10]. These guidelines contributed to the research team's decision to halt the consenting process to (1) reduce overall exposure and (2) allow pediatric health care teams to focus on emergent issues without extraneous research activities taking place. Screening, an inherently remote process via the online EHR system, continued with slight modifications. Under normal circumstances, the clinical research team would optimize screener-consenter shift coverage.

Although the team was not consenting candidates, screening served as an accessible avenue for the research study to progress amidst the pandemic and to evaluate the efficacy of newly implemented screening protocols and training. An updated screening training module was created to clarify the scope of the study by refining the inclusion criteria. One such clarification was that if a candidate meets the inclusion criteria but is out of the age range, the screening Google Form should be recorded but the EHR chart should remain unopened. A question was added to the form to identify whether or not a candidate was screened during the COVID-19 pandemic.

Operationalization and Processing of Variables

Nine variables were examined to determine the effects of the pandemic on internal operations and hospital utilization patterns. The variables used to examine effects on hospital utilization patterns for potential participants were: (1) type of injury, (2)

location of injury, (3) MOI, and (4) age. The variables used to examine effects on internal operations were: (1) screened charts, (2) screening coverage, (3) identified candidates, (4) identified potential participants, and (5) identification time. Next, records that did not indicate a potential participant were discarded for the scope of the analyses. The ED visit reasons for the remaining records were broken down into two major categories: (1) "TBI," representing records that directly indicated a TBI (eg, "TBI," "Closed Head Injury [CHI]," "Head Trauma"); and (2) "TBI-related orthopedic injuries," representing records that indicated orthopedic injuries that *may* have been related to a TBI (eg, "Jaw Injury," "Forehead Contusion"). From these two major categories, the mechanisms of injury were narrowed down to six of the most commonly reported categories: (1) "Fall," (2) "MVA," (3) "Strike," (4) "Assault," (5) "Unknown," and (6) "Other."

The mechanisms of injury written by screeners also included a description of the injury, from which information about the location could be extracted. This was a secondary variable that was created and split into three categories: (1) "Outdoor Injuries," (2) "Indoor Injuries," and (3) "Unknown." The categories were determined by certain key words included in the description. "Inside" injuries for 2019 were classified as being inside at any location (eg, school, office space, private residence). "Outside" injuries for 2019 were classified as any injury that did not occur inside a building (eg, trampoline fall in backyard, MVA). For example, a description of "CHI on playground while running" indicated an outdoor injury, as playgrounds are located outside, and a description of "Slipped and hit head on bed" would indicate an indoor injury, as beds are located inside a building. The breakdown of each variable and their operational definitions are provided in Table 1 and Table 2.

Table 1. Screening variables.

Variable[a]	Operationalization	Formula
Screened charts	The total number of screened charts during a given time period	N/A[b]
Screening coverage	The number of hours spent screening per month as a proportion of the total number of hours in a given month[c]; the higher the percentage, the better the screening coverage	(Total number of hours spent screening/total number of hours in a given month)×100
Identified candidates	A person who presented to the ED[d] whose census data suggest a potential traumatic brain injury and justify opening a screened chart under Health Insurance Portability and Accountability Act guidelines	Proportion of candidates=total number of opened charts/total number of screened charts
Identified potential participants	An individual who, following further chart review, continues to meet the inclusion criteria and would be approached by a consenter	Proportion of potential participants=total number of potential participants/total number of opened charts
Identification time	Length of stay in the ED of a potential participant recorded by the screener at the time of identification, which can reflect how quickly a screener can identify a potential participant; a shorter length reflects a quicker identification time	N/A

[a]Screening variables are defined as variables that provide information about screening patterns from data collected from the emergency department census and individual electronic health records.

[b]N/A: not applicable.

[c]March and July have 31 days, whereas April has 30 days; thus, the total proportion of hours screened was calculated using 744 and 72 total monthly hours for these months, respectively.

[d]ED: emergency department.

Table 2. Hospital utilization variables.

Variable[a]	Operationalization	Notes
Age	The average age of potential participants during a given time period	Age was treated as a continuous variable since average age was calculated between 2019 and 2020; therefore, no groupings were used
Type of injury	ED[b] visit reasons classified into the following two subcategories: (1) indication of possible TBI[c] and (2) indication of orthopedic injury possibly relating to a TBI	(1) examples include closed head, injury, syncope, and headache; (2) examples include jaw injury, forehead or facial contusion, and cervical spine injury
Mechanism of injury	Collected from the description of the injury and classified into the following five subcategories: (1) fall, (2) motor vehicle accident (MVA), (3) strike, (4) assault, and (5) unknown	A mechanism of injury classified as "unknown" is defined as occurring in an unspecified manner due to lack of details in the EHR[d]
Location of injury	Collected from the description of the injury and classified into the following three subcategories: (1) inside, (2) outside, and (3) unknown	(1) defined as occurring inside at any location (eg, office space, school, private residence); (2) defined as occurring at any defined location other than those defined above as "inside," such as office space, school, private residence (eg, MVA, riding a bike); (3) defined as occurring in an unspecified location due to the lack of details in the EHR (eg, punch to the head, hit wall, fall from syncope, hit in the back of head by elbow)

[a]Hospital utilization variables are defined as variables that provide information about hospital utilization patterns with data collected only from the potential participants' electronic health record.

[b]ED: emergency department.

[c]TBI: traumatic brain injury.

[d]EHR: electronic health record.

Analysis

Screening data were deidentified according to HIPAA guidelines before being recorded and collected through Google Forms, which compiled each screened individual's collected information into a single "record." These records were displayed as rows on the imported Google Sheets corresponding to the responses collected through the Google Form. Information not pertaining to data analyses, such as "Screener Name" or "Glasgow Coma Scale Score," were omitted. Sex was omitted in the analysis because previous screening forms did not collect the participant's sex before all inclusion criteria were met. Thus, if inclusion criteria were not met, sex, ethnicity, and race were not recorded in the screening form, resulting in incomplete records, which were not used. Screening form responses were downloaded from Google Forms and imported into Google Sheets. The values were then standardized and cleaned of erroneous values, missing data, and duplicate form responses through data-matching methods. Form response time stamps were compared to the potential participant's time of arrival, MOI, and age to ensure duplicates were removed and missing data were then completed. The Control+F function served to find responses that had similar characteristics so that the data could be manually reviewed one final time. Thus, the records used for analysis contained no missing data and analyses were performed using complete data sets. The sample size for the analyses was determined after (1) discarding all incomplete records and (2) including only those records that indicated potential participants and candidates, as needed for the analysis. For the purposes of this study, a *candidate* is defined as a person who presented to the ED whose census data suggest a potential TBI and justify opening a screened chart under HIPAA guidelines. A *potential participant* is defined as an individual who, following further chart review, continues to meet inclusion criteria and would be approached by a consenter.

Google Sheets was also used to tabulate the frequency of the following variables pertaining only to potential participants: (1) type of injury, (2) MOI, and (3) location of injury. Since the WHO declared COVID-19 a pandemic in March of 2020, data analysis was restricted to the 6-month period between March and August of 2020; analogous data from 2019 were used to compare cumulative responses. The second comparison was performed between specific months corresponding to state-wide mandates from 2020 and 2019, as shown in Table 3. This comparison was performed for the specific months of March, April, and July between 2020 and 2019, as these months corresponded to state-wide COVID-19 mandates, as referenced in Table 1. An α value of .05 was used as the a priori cutoff for statistical significance.

All comparisons were assessed in the R programming language [11], using the two-sample Student *t* test and the Mann-Whitney-Wilcoxon rank-sum test to determine any significant differences in the following variables between the years 2019 and 2020: (1) screened charts, (2) identification time, and (3) average age of potential participants. Screening coverage was calculated, by hand, using Google Calendar to track and record screening hours. A two-proportion *z* test was used to compare the hours spent screening and the rates of candidates and potential participants between 2019 and 2020. For potential participants whose visit reason was indicative of a TBI, a two-proportion *z* test was also used to compare the frequencies of MOI, types of injury, and location of injury between 2019 and 2020. The Shapiro-Wilk test was used to determine normality of the two samples and to select the appropriate statistical test (ie, Mann-Whitney-Wilcoxon rank-sum test for samples not normally distributed and Student *t* test for normally distributed samples). Further, an F-test was performed to confirm equal variance of normally distributed samples. Unequal

variances were addressed by performing the Welch *t* test. Google Sheets and Canva were used to create data visualizations.

Table 3. Timeline of mandates issued by the state of Texas regarding COVID-19 [12].

Date	Protocol implemented
March 24, 2020	The Stay at Home or Place of Residence order became effective as of 11:59 PM. The first executive order was signed, which banned gatherings of 10+ people, closed dine-in restaurants and schools, and limited visitations to long-term care centers.
April 17, 2020	Texas Governor Greg Abbott issued an executive order that in part calls for schools to remain closed for the remainder of the academic year [13].
May 1, 2020	The reopening process began with 25% capacity at most stores and restaurants.
June 26, 2020	Lockdown orders were reimplemented, with capacity being dropped to 50% at most locations and bars being shut down. Six days later, a mask mandate was instituted.

Results

In total, 1245 screening form entries from 2019 and 1076 screening form entries from 2020 were analyzed. Table 4 provides comparisons of different screening variables between analogous months. The proportion of candidates in 2019 was significantly greater than that in 2020 (*P*<.001). Specifically, the proportion of candidates in April 2019 was significantly greater than that in April 2020 (*P*=.01). The proportion of potential participants in 2019 was significantly lower than that in 2020 (*P*<.001). The proportions of potential participants across all measured time points in April, March, and July were significantly lower in 2019 than in 2020, as seen further in Table 4. The average age of potential participants was significantly older in 2019 by 1.62 years (*P*=.006).

Table 4. Screening and hospital utilization patterns.[a]

Time period	Proportion of screening hours[b]	Total screened charts, n	Total opened charts, n	Proportion of candidates, %	Proportion of potential participants, %	Total length of stay (minutes)	Mean age of potential participants (years)
All months (March, April, and July)							
2019	2392	1245	908	72.9	21.4	37	8.93
2020	2783.8	1076	633	58.8	35.1	35	7.31
P value	.09	.38	.05	<.001	<.001	.48	.006
March							
2019	46.6	284	218	76.8	23.9	32	10
2020	32.5	66	46	69.7	43.5	55	10
P value	<.001	.07	.08	.30	.006	.03	>.99
April							
2019	65.3	214	157	73.4	17.2	36	10.22
2020	84.9	188	117	62.2	29.9	33	9.63
P value	<.001	.09	<.001	.01	.01	.77	.43
July							
2019	54.3	134	91	67.9	26.4	42	8.23
2020	69	209	125	59.8	39.2	22	9.43
P value	<.001	.22	.42	.16	.07	.01	.37

[a]See Table 1 and Table 2 for definitions and formulas for each variable.

[b]Presented as total raw numbers screened for the respective analyzed months for the All Months category and as percentages (raw number/total number of hours) for the individual month categories.

Table 5 displays a comparison of the different types of injuries that were reported for potential participants through the ED census. There was a significantly greater number of potential participants who had visit reasons explicitly indicating a TBI (*P*<.001) and those with an orthopedic injury possibly related to a TBI (*P*=.002) in 2020 than for those who came into the ED for the same reasons in 2019, respectively. Table 5 also displays the general location that potential participants were injured in (ie, outside, inside). There were significantly more injuries that occurred inside in 2020 than those in 2019 (*P*=.02). When examining the various MOI, no significant differences were found between 2019 and 2020. There was a higher frequency of visit reasons indicating an MOI of the Fall and Strike categories in 2020 than in 2019; however, these differences

were not statistically significant. A detailed comparison of the MOI during the months of interest across the 2 years can be found in Figure 2.

One key finding was that approximately 55 minutes passed before a screener identified a candidate in March 2020, which was significantly slower than the 32 minutes that passed in March 2019 (P=.03). Approximately 22 minutes passed before a screener identified a candidate in July 2020, which was significantly faster than the 42 minutes that passed in July 2019 (P=.01). A total of 241.8 hours were utilized for screening in March 2020, which was significantly less than the 346.5 hours utilized for screening in March 2019 (P<.001). In April 2020, 611.5 hours were utilized for screening and 513.5 hours were utilized in July 2020. Both of these monthly totals were significantly greater than the total hours utilized for screening in the analogous months in 2019: 470.5 hours in April (P<.001) and 404.3 hours in July (P<.001).

Table 5. Emergency department (ED) visit reason, mechanism of injury, and location of injury of potential participants.

Characteristic	2019[a] frequency, n (%)	2020[a] frequency, n (%)	*P* value
ED visit reason			
Total participants[b], n	195	222	N/A[c]
TBI[d]	103	181	<.001
TBI-related orthopedic injury[e]	3	18	.002
Mechanism of injury			
Total participants, n	103	181	N/A
Fall	47 (45.6)	104 (57.5)	.07
MVA[f]	4 (3.9)	13 (7.2)	.39
Strike	32 (31.1)	37 (20.4)	.06
Assault	3 (2.9)	3 (1.7)	.78
Unknown	17 (16.5)	23 (12.7)	—[g]
Other	0 (0)	1 (0.5)	—
Location of Injury			
Total participants, n	103	181	N/A
Inside injuries	11 (10.7)	39 (21.6)	.02
Outside injuries	49 (47.6)	77 (42.5)	.38
Unknown location	43 (41.7)	65 (35.9)	—

[a]The years of 2019 and 2020 in this subanalysis are defined to be the months of March, April, and July of each year, respectively.

[b]Indicates the total number of flagged potential participants, including those in the non-TBI-related orthopedic injury and unknown categories, which were not included for comparison.

[c]N/A: not applicable.

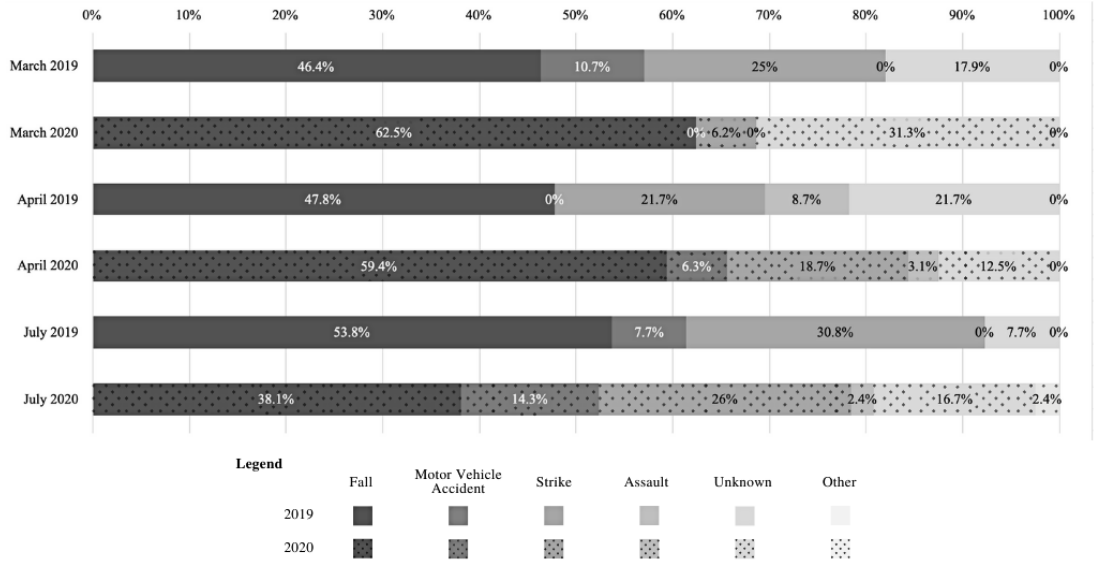
[d]TBI: traumatic brain injury.

[e]These visit reasons include orthopedic injuries that may be related to a TBI (eg, jaw injury, forehead contusion).

[f]MVA: motor vehicle accident.

[g]Statistical analysis was not possible for unknown groups since this was an indefinite category.

Figure 2. Mechanism of injury classifications of potential participants in March, April, and July of 2019 vs 2020.



Discussion

Principal Findings

Regarding ED utilization patterns, significantly older patients were flagged as potential participants in 2020 than in 2019. Significantly more potential participants came into the ED with TBI or TBI-related injuries in 2020 than in 2019, and the MOIs for these potential participants were not significantly different between the two years. There were varying outcomes among the analyzed months regarding evaluation of the efficacy of the internal screening process. Generally, positive effects of the change in internal screening protocols were evident, as measured by analyzed variables such as total screened charts, total opened charts, length of stay (LOS), and total hours spent screening. These findings are discussed in detail further.

The novelty of this study lies in the quick adaptation of the research team's previously established ED tracking protocol to identify potential participants, allowing for examination of the novel effects of the pandemic on ED utilization. Although the research team was no longer consenting and enrolling participants, the internal screening processes proved to be beneficial during the pandemic as the screening could occur concomitantly amidst state-regulated mandates (eg, stay-at-home orders, social distancing recommendations). Although many research studies halted procedures during the pandemic, the preexisting structure of the parent study allowed the research team to examine ED utilization patterns and evaluate the adaptability of the established screening processes with minimal changes to the existing modality.

ED Utilization Patterns

The average age of a potential participant was significantly older in 2020 (8.93 years) than in 2019 (7.31 years) by approximately 1.5 years (P=.006). Notably, in 2019 and 2020, the average age fell in the middle childhood range (6-8 years) as defined by the Centers for Disease Control and Prevention (CDC) [14]. One study noted that children who suffered a TBI during this critical developmental stage were vulnerable to poorer cognitive outcomes as compared to children in other developmental stages [15]. This finding could be related to the level of health care avoidance that is seen in parents during the pandemic, as several studies have supported this hypothesis [16,17]. Although further details about the profiles of such parents are yet to be studied in the United States, it can be hypothesized that parents may be more hesitant to bring a younger child into the ED, which might be perceived as a high-risk environment.

Significantly more potential participants came into the ED with either a TBI (P<.001) or TBI-related orthopedic injuries (P=.002) in 2020 than in 2019, as seen in Table 5. In response to an executive order issued by Texas Governor Greg Abbott calling for all schools to remain closed for the remainder of the 2020 academic year, all nine independent school districts in the Austin area transitioned to online learning, forcing many students to conduct their education at home [13]. Child sports play activity also decreased after lockdown measures, leading to increased home-based activities [18]. It is hypothesized that quarantine measures may have increased TBI or TBI-related orthopedic injuries that occurred at home. This hypothesis is supported by the significant increase in injuries that occurred inside in 2020 when compared with those in 2019 (P=.02).

The proportion of the various MOIs seen in Table 5 from potential participants who came in for a TBI did not significantly differ between years. The lack of significance in the proportion of MVAs reported as the MOI in 2020 when compared to 2019 is interesting, as there were fewer MVAs reported in Austin according to one study [19] and notably less traffic occurring in 2020 overall. There was 40% less traffic congestion recorded in response to the Declaration of State Disaster issued on March 13, 2020, as an effort to "mitigate the spread of COVID-19," when compared to the same day in 2019 [20]. The more severe order of restriction, Stay at Home or Place of Residence, became effective March 24, 2020, resulting in 70% less traffic per day on average, which continued to be true throughout April (76.46%) and July (71.84%), following the significant events described in Table 3.

Evaluating the Efficacy of Internal Screening Processes

The lack of significance for total screened charts recorded by screeners across 2019 and 2020 may be explained in part by the fact that the recruitment site is not a COVID-19 treatment facility, possibly limiting the impact of the pandemic on ED admissions relating to the study's inclusion criteria. However, given the potentially serious nature of a head injury, any concerns about contracting the COVID-19 virus might have been disregarded. The utilization of COVID-19 treatment facilities as compared to non-COVID-19 treatment facilities remains to be empirically studied.

However, of the total screened charts, there were significantly less charts opened in April 2020 (117 charts) when compared to those opened in April 2019 (157 charts), which may be related to screening protocol changes (P<.001). Examination of previous data showed common visit reasons that screeners would pursue (eg, facial lacerations), but ultimately did not result in TBI once the chart was further examined. Based on reports from the CDC, screening training was updated to ensure that screeners would only examine charts with injuries more likely to be indicative of a TBI (eg, fall). According to the CDC, the leading causes of TBI include falls, strikes, MVA, and intentional self-harm [21].

The significant increase in the LOS screening variable during March 2020 (55 minutes) compared with that in March 2019 (32 minutes) could be explained by the difference in screening coverage between the two analogous months (P=.03). If there is consistent coverage throughout the day, there is a higher probability of a candidate being identified quickly and thus a shorter LOS reported by the screener. Since a large majority of the research volunteers are undergraduate students, screening coverage tends to decrease during spring break; however, the impact was greater during 2020 due to the extension of the spring break, as mandated by the university in an effort to adapt to the pandemic. The proportion of total hours spent screening during March 2020 (346.5 hours) was significantly less than the proportion of total hours spent screening during March 2019 (241.8 hours; P<.001). This discrepancy could be attributed to the time it took for the research team to modify its recruitment protocols to adapt to the ongoing changes to safety regulations. The research team does not require its screeners to continue screening throughout academic breaks, thus contributing to the decrease in screening coverage seen in March 2020.

After the recruitment site and the university halted human-subjects research on March 13, 2020, and March 15, 2020, respectively, the research team expeditiously decided to convert all consenting shifts into screening shifts. This decision positively impacted screening coverage in April 2020. There was a significant increase in the proportion of hours spent screening in April 2020 (611.5 hours) when compared to April 2019 (470.5 hours; P<.001). Additionally, there was no significant difference in the LOS between April 2020 (33 minutes) and April 2019 (36 minutes; P=.77), demonstrating the research team's efficient transition of recruitment practices.

There was a significant increase in the total number of hours spent screening in July 2020 (513.5 hours) when compared to July 2019 (404.3 hours). The research team recruited a new cohort of research assistants in May 2020, all of whom were trained to screen quickly and efficiently. With enrollment efforts still halted, the research team encouraged the new cohort to begin screening in the summer, thus increasing screening coverage during a time that otherwise experiences a decrease in coverage. This also positively impacted the LOS during July, as there was a significant decrease in LOS in July 2020 (22 minutes) compared to July 2019 (42 minutes; P=.01).

In addition to improved screening coverage, there was greater accuracy with respect to more charts being opened for further screening where the individual was deemed eligible to approach. Although there was a greater proportion of candidates (charts opened) in 2019 than in 2020 (P<.001), there was a greater proportion of potential participants (cases deemed approachable) in 2020 than in 2019 (P<.001). The proportion of potential participants was significantly greater overall in 2020 than in 2019 (P<.001) and at every measured time point. This further reflects the need for continued enrollment efforts and highlights the impact of the pandemic on the biobanking efforts of the parent study. The improvement in screening accuracy may be due to changes in training and laboratory protocols; however, this remains to be empirically studied.

Strengths and Limitations

Due to the unprecedented nature of the pandemic, the study was executed using the parent study's previously established protocols. Consequently, one study limitation is the inability to sufficiently evaluate the effects of the pandemic, as the data collection methods were originally not tailored to the scope of this study's research questions. The research team could have adapted the methods of data collection to better identify the effects of the pandemic. Collecting data on the length of time that elapsed between injury and the ED visit, whether or not parents were aware that the hospital was not treating COVID-19 patients, and whether or not parents felt comfortable bringing their children in are a few questions that may elucidate how parental behavior may have been impacted by the pandemic. Additionally, collecting data on demographic variables (eg, race, ethnicity) of candidates who visit the ED can help to explore the relationship between hospital utilization rates during a pandemic and other social factors. This is an important trend to study as reports of health disparities and inequities during the pandemic across specific ethnic and race groups become more evident [22].

Although efforts were made to clarify the scope of the study and to refine the inclusion criteria, the study is limited by the precision of the screeners monitoring the EHR and decisions made to open a potential participant's chart. Imprecision may lead to a greater number of potential participants to be included that did not actually meet the inclusion criteria or fewer potential participants, depending on whether a particular screener was more lenient or stringent when determining potential participants. This could have led to inflation or underestimation of the reported values. Additionally, intraobserver bias is possible due to differences in screener experience of each research assistant.

Conclusions and Future Directions

Findings from this study indicate significantly higher ED visits for a TBI during the COVID-19 pandemic during 2020 when compared to the analogous months of 2019 ($P<.001$). Based on these findings, it is plausible for the parent study's enrollment efforts to continue, albeit with added precautions, including: (1) socially distant consenting protocols or virtual consenting, (2) smaller teams of consenters with alternating shifts, (3) precautions such as daily symptom tracking and biweekly COVID-19 tests required for on-site consenters, and (4) requiring on-site consenters to receive the COVID-19 vaccine before returning to their roles. This information is relevant to the research community from a number of perspectives, including funding, tenure, publications, and deliverables during a pandemic. If these precautions cannot be met, the research team must consider remote data collection from clinical studies to continue a steady workflow amidst global change.

As the research team continues to systematically gather data for the parent study, future directions can focus on developing dynamic protocols to reflect relevant public health issues as they evolve, including: (1) modifying data collection protocols, (2) employing effective methods of tracking health care system utilization, (3) understanding the factors that influence a participant's decision to enroll, and (4) understanding the factors that affect participant retention. For existing records collected prior to this analysis, former participants might be contacted to request additional information, thus filling in missing information. Another avenue to reduce risk of infection among in-person consenters might be to postpone sample collection (eg, saliva) until state regulations suggest specimens can be collected safely. This study illuminates the importance of continuous data collection, allowing the research team to adapt existing protocols to combat unexpected changes such as the COVID-19 pandemic.

Acknowledgments

First and foremost, we would like to thank the frontliners and researchers who have worked tirelessly to end the pandemic. We would also like to thank and acknowledge our affiliated institutions, including The University of Texas at Austin School of Nursing, Department of Neurology at Dell Medical School, and Dell Children's Medical Center. Thank you to all of the clinicians and employees at Dell Children's Medical Center that allowed us to work alongside them and help us whenever needed, especially Linda Wasson, the volunteer coordinator, and Kimberly Lewis for helping us secure site approval. We would also like to thank Marilyn Farmer for reviewing and editing the manuscript, as well as Ramsha Jinnah, Anjali Raghavan, Izzy Obias, Janet Le, John McDonald, Marcus Hunter, and Samuel Lin for helping clean the data. We would also like to recognize the volunteers of the Osier Laboratory who screen potential participant charts each and every day, helping us to collect data. NO received the funding for this study. This research was supported by St David's Center for Health Promotion and Disease Prevention Research in Underserved Populations (grant 30-2142-4358) and Heilbrunn Family Center for Research Nursing (grant 26-6820-5550).

Conflicts of Interest

None declared.

References

1. Menon V. Developmental cognitive neuroscience of arithmetic: implications for learning and education. ZDM 2010 Oct;42(6):515-525 [FREE Full text] [doi: 10.1007/s11858-010-0242-0] [Medline: 22003371]
2. Get the facts about TBI. US Department of Health and Human Services Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/traumaticbraininjury/get_the_facts.html [accessed 2020-12-01]
3. Report to Congress: Traumatic brain injury epidemiology and rehabilitation: recommendations for addressing critical gaps. Centers for Disease Control and Prevention. 2015 Mar. URL: https://www.cdc.gov/traumaticbraininjury/pdf/tbi_report_to_congress_epi_and_rehab-a.pdf [accessed 2020-12-01]
4. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/traumaticbraininjury/data/tbi-ed-visits.html [accessed 2020-12-01]
5. Currie D, Snedden T, Pierpoint L, Comstock RD, Grubenhoff JA. Factors influencing primary care follow-up after pediatric mild traumatic brain injury. J Head Trauma Rehabil 2019;34(4):E11-E19 [FREE Full text] [doi: 10.1097/HTR.0000000000000461] [Medline: 30608309]
6. Lazzerini M, Barbi E, Apicella A, Marchetti F, Cardinale F, Trobia G. Delayed access or provision of care in Italy resulting from fear of COVID-19. Lancet Child Adolesc Health 2020 May;4(5):e10-e11 [FREE Full text] [doi: 10.1016/S2352-4642(20)30108-5] [Medline: 32278365]
7. Metzler B, Siostrzonek P, Binder RK, Bauer A, Reinstadler SJ. Decline of acute coronary syndrome admissions in Austria since the outbreak of COVID-19: the pandemic response causes cardiac collateral damage. Eur Heart J 2020 May 14;41(19):1852-1853 [FREE Full text] [doi: 10.1093/eurheartj/ehaa314] [Medline: 32297932]
8. Chang H, Huang N, Lee C, Hsu Y, Hsieh C, Chou Y. The impact of the SARS epidemic on the utilization of medical services: SARS and the fear of SARS. Am J Public Health 2004 Apr;94(4):562-564. [doi: 10.2105/ajph.94.4.562] [Medline: 15054005]

9. Lay P, Paralkar T, Ahmed SH, Ghani M, Muneer S, Jinnah R, et al. A novel clinical research modality for enrolling diverse participants using a diverse team. Brain Sci 2020 Jul 08;10(7):434 [FREE Full text] [doi: 10.3390/brainsci10070434] [Medline: 32650502]
10. Coronavirus Disease 2019 (COVID-19): Information for NIH applicants and recipients of NIH funding. National Institutes of Health Central Resource for Grants and Funding Information. 2019. URL: https://grants.nih.gov/policy/natural-disasters/corona-virus.htm [accessed 2020-12-01]
11. The R Project for Statistical Computing. 2020. URL: https://www.r-project.org/ [accessed 2021-12-30]
12. Adler S. Stay home - work safe, order 2024-007 by the mayor of the city of Austin. Texas Tribune. 2020 Mar 24. URL: https://static.texastribune.org/media/files/db509d0ae441ea82d8e1fcf3106c530a/3%2024%20order.pdf [accessed 2020-12-01]
13. Governor Abbott Issues Executive Order Establishing Strike Force To Open Texas. Office of the Texas Governor. 2020 Apr 17. URL: https://gov.texas.gov/news/post/governor-abbott-issues-executive-order-establishing-strike-force-to-open-texas [accessed 2020-12-01]
14. Middle childhood (6-8 years of age). Child Development Centers for Disease Control and Prevention. URL: https://www.cdc.gov/ncbddd/childdevelopment/positiveparenting/middle.html [accessed 2020-11-30]
15. Crowe LM, Catroppa C, Babl FE, Rosenfeld JV, Anderson V. Timing of traumatic brain injury in childhood and intellectual outcome. J Pediatr Psychol 2012 Aug;37(7):745-754. [doi: 10.1093/jpepsy/jss070] [Medline: 22669504]
16. McDonnell T, McAuliffe E, Barrett M, Conlon C, Cummins F, Deasy C, et al. CUPID COVID-19: emergency department attendance by paediatric patients during COVID-19 - project protocol. HRB Open Res 2020;3:37 [FREE Full text] [doi: 10.12688/hrbopenres.13066.2] [Medline: 32666040]
17. Nicholson E, McDonnell T, Conlon C, Barrett M, Cummins F, Hensey C, et al. Parental hesitancy and concerns around accessing paediatric unscheduled healthcare during COVID-19: a cross-sectional survey. Int J Environ Res Public Health 2020 Dec 11;17(24):9264 [FREE Full text] [doi: 10.3390/ijerph17249264] [Medline: 33322332]
18. State of Play 2020: Pandemic Trends. The Aspen Institute Project Play. 2020. URL: https://www.aspenprojectplay.org/state-of-play-2020/pandemic-trends [accessed 2020-12-01]
19. Boserup B, McKenney M, Elkbuli A. The impact of the COVID-19 pandemic on emergency department visits and patient safety in the United States. Am J Emerg Med 2020 Sep;38(9):1732-1736 [FREE Full text] [doi: 10.1016/j.ajem.2020.06.007] [Medline: 32738468]
20. Austin traffic full-year historical traffic data. TomTom. 2020. URL: https://www.tomtom.com/en_gb/traffic-index/austin-traffic/ [accessed 2020-12-01]
21. Surveillance Report of Traumatic Brain Injury-related Emergency Department Visits, Hospitalizations, and Deaths. Centers for Disease Control and Prevention. 2014. URL: https://www.cdc.gov/traumaticbraininjury/pdf/TBI-Surveillance-Report-FINAL_508.pdf [accessed 2020-12-01]
22. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. JAMA 2020 Jun 23;323(24):2466-2467. [doi: 10.1001/jama.2020.8598] [Medline: 32391864]

Abbreviations

CDC: Centers for Disease Control and Prevention
CHI: closed head injury
ED: emergency department
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
LOS: length of stay
MOI: mechanism of injury
MVA: motor vehicle accident
Neuro-QoL: Quality of Life in Neurological Disorders
PROMIS: Patient-Reported Outcomes Measurement Information System
SARS: severe acute respiratory syndrome
TBI: traumatic brain injury
WHO: World Health Organization

Edited by T Leung; submitted 15.04.21; peer-reviewed by Y Chu, S Kim; comments to author 12.10.21; revised version received 30.12.21; accepted 22.02.22; published 23.03.22.

Please cite as:
Paralkar TA, Lay P, Stubbs S, Ahmed SH, Ghani M, Osier N
Occurrence Patterns of Traumatic Brain Injury Within the Emergency Department and Internal Screening Process Efficacy During the COVID-19 Pandemic: Retrospective Analysis
Interact J Med Res 2022;11(1):e29513
URL: https://www.i-jmr.org/2022/1/e29513
doi:10.2196/29513
PMID:35225820

Original Paper

Perceived Impact of Outdoor Swimming on Health: Web-Based Survey

Heather Massey[1], BSc, MSc, PhD; Paul Gorczynski[2], PhD, CPsychol; C Mark Harper[3,4], BSc, MBBS, PhD; Lisa Sansom[5,6], BSc, MBBS, MRGP; Kieren McEwan[2], MA, PGCE, PhD; Alla Yankouskaya[7], BSc, PhD; Hannah Denton[8], MSc, MRes, PGDip

[1]Extreme Environments Laboratory, School of Sport, Health and Exercise Science, University of Portsmouth, Portsmouth, United Kingdom

[2]School of Sport, Health and Exercise Science, University of Portsmouth, Portsmouth, United Kingdom

[3]Department of Anaesthesia, Royal Sussex County Hospital, Brighton, United Kingdom

[4]Department of Anaesthesia, Sørlandet Sykehus, Kristiansand, Norway

[5]Public Health Department, East Sussex County Council, Lewes, United Kingdom

[6]NHS Primary Care, Hastings, United Kingdom

[7]Department of Psychology, Faculty of Science and Technology, Bournemouth University, Bournemouth, United Kingdom

[8]West Recovery Team, Mill View Hospital, Sussex Partnership NHS Foundation Trust, Hove, United Kingdom

Corresponding Author:
Heather Massey, BSc, MSc, PhD
Extreme Environments Laboratory, School of Sport, Health and Exercise Science
University of Portsmouth
Spinnaker building
Cambridge road
Portsmouth, PO1 2ER
United Kingdom
Phone: 44 02392843545
Email: heather.massey@port.ac.uk

Abstract

Background: Outdoor swimming in lakes, lidos (outdoor pools), rivers, and the sea has grown in popularity in many countries, including the United Kingdom. Many anecdotal accounts indicate improvements in medical conditions, which are considered a consequence of outdoor swimming.

Objective: The aim of this study is to better understand outdoor swimmers' perceptions of their health and the extent to which participation impacted their existing self-reported symptoms.

Methods: A survey was conducted to investigate outdoor swimming behaviors and reports of any diagnosed medical conditions. Medical conditions were coded into categories, and descriptive statistics were generated regarding the outdoor swimmers' behaviors and the effect that outdoor swimming had on their medical symptoms if any. The medical categories were clustered into five larger categories based on their prevalence in the current sample: mental health; musculoskeletal and injury; neurological; cardiovascular and blood disease; and *other*, which comprises inflammatory, immune, endocrine, and respiratory conditions.

Results: In total, 722 outdoor swimmers responded, of whom 498 (68.9%) were female. The probability of outdoor swimming having *some positive impact* on health across all medical categories was 3.57 times higher compared with *no* impact (*B*=1.28, 95% CI 0.63-1.91; *P*<.001), 44.32 times higher for the mental health category (*B*=3.79, 95% CI 2.28-5.30; *P*<.001), 5.25 times higher for musculoskeletal and injury category (*B*=1.66, 95% CI 0.52-2.79; *P*=.004), and 4.02 times higher for the *other* category (*B*=1.39, 95% CI 0.27-2.51; *P*=.02). Overall, outdoor swimming was associated with perceived reductions in symptoms of poor mental health (χ^2_2=25.1; *P*<.001), musculoskeletal and injury (χ^2_2=8.2; *P*=.04), cardiovascular and blood (χ^2_2=14.7; *P*=.006), and *other* conditions (χ^2_2=18.2; *P*<.001).

Conclusions: Physical activity in the form of outdoor swimming is perceived to have positive impacts on health and is associated with perceived symptom reductions in mental health, musculoskeletal and injury, and cardiovascular and blood conditions. This study cannot provide causal relationships or provide mechanistic insights. However, it does provide a starting point for more

targeted prospective intervention research into individual conditions or categories of conditions to establish the impact in those who choose to start outdoor swimming.

(Interact J Med Res 2022;11(1):e25589) doi:10.2196/25589

KEYWORDS

open water swimming; blue space; blue gym; mental health; physical health

Introduction

Background

Swimming outdoors is an increasingly popular recreational physical activity both in the United Kingdom and abroad [1,2], offering opportunities to be physically active in a range of facilities from natural water sources such as ponds, lakes, rivers, and the sea to man-made outdoor facilities, such as open-air pools or lidos. These locations differ from indoor pool swimming as they are based in natural environments, with lower water temperatures and fresh or salt water without chlorine treatment. These different attributes provide opportunities for those who prefer a more natural environment, those who cannot or prefer not to swim in chlorinated water, and those with limited access to indoor facilities.

Most research into exercise for health and well-being has been land-based. However, findings from a small number of swimming studies suggest that psychological effects are similar to exercise on land [3]. In addition, it has been suggested that bodies of difference can be enabled through immersion in water [4]. This can lead to a transformation of the *unhealthy* land body, for example, *large* and *middle-aged*, into a healthy sea body [5] and enable older people to challenge perceptions of burden and dependency [6].

The expansion of outdoor swimming has also been mirrored by the increased volume of research on the potential benefits of activity in blue spaces [7], thus, highlighting not only the need to remain physically active but also the potential of the natural environment to support improvements in health and well-being [8]. At this stage, there are many accounts suggesting that outdoor swimming can promote healthy aging and improve health [9,10]. The evidence remains at an anecdotal, case report or expert opinion level in accordance with evidence-based medicine criteria [11]. The accounts frequently discuss similar themes of transformation, connectedness, and reorientation, which have been well described by Denton and Aranda [12].

Objective

The concept of cold water swimming or cold water spa treatments is not new; Hippocrates claimed that water therapy reduced lassitude (ie, lethargy) [13]. Considering the increasing volume of contemporary anecdotal evidence, we are no closer to establishing which medical conditions, if any, may be improved through regular outdoor swimming, how much improvement can be made, and by what mechanisms improvements occur. This research aims to provide a small step in that process by surveying outdoor swimmers to establish the medical conditions they have been diagnosed with and if they have perceived any change in their symptoms since starting to swim outdoors. Therefore, it is hypothesized that the type of medical condition can reliably predict the perceived health impact of outdoor swimming.

Methods

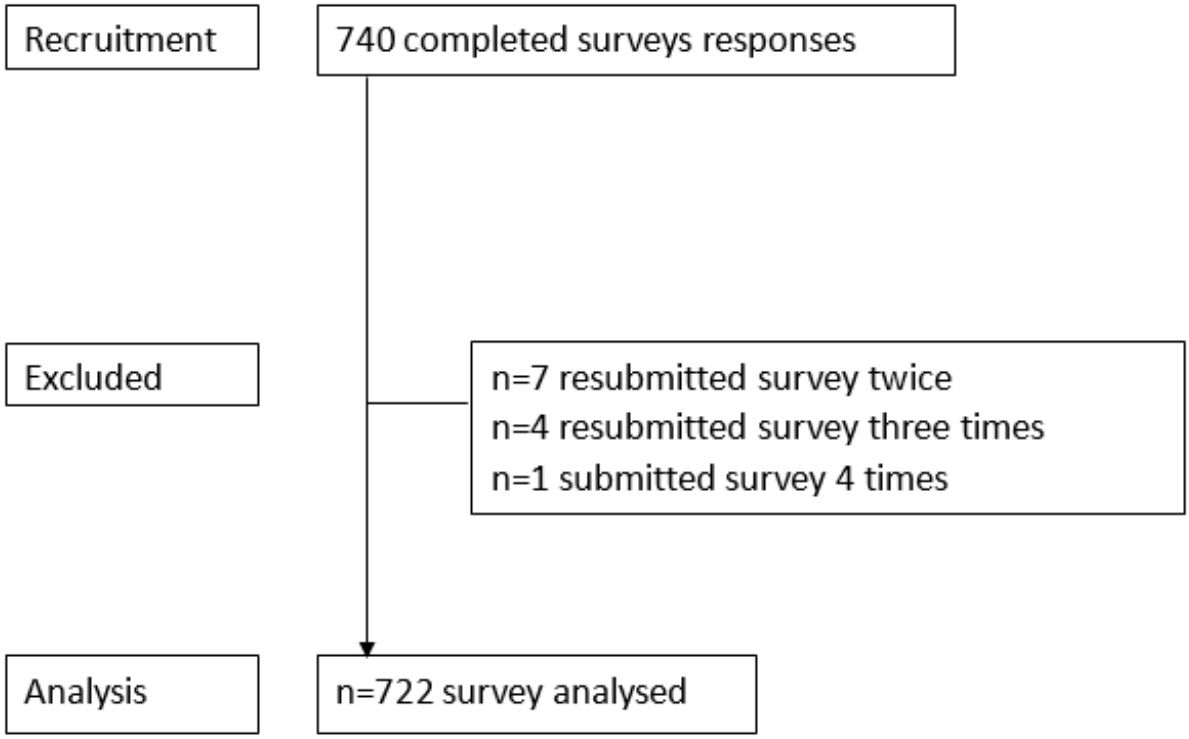
Survey Methods

The survey and manuscript were prepared in accordance with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) guidance [14]. A 14-item web-based cross-sectional survey was completed by 722 people who swim outdoors (duplicate and incomplete forms were removed) following ethical approval from the University of Portsmouth Science Faculty Research ethics committee (SFEC 2018-120), using the JISC web-based survey platform.

Participants and Procedures

Participants were 722 outdoor swimmers (498/722, 68.9% female; 159/722, 22% male; and 65/722, 9% did not identify sex; outdoor swimmers are described here as participants who swim or immerse themselves in natural water environments, such as the sea, rivers, or lakes, or open-air pools). They gave their informed consent to allow the anonymous use of their data for the explicit purpose of establishing what medical conditions outdoor swimmers have and whether they gain any relief or symptom reduction from the activity. All participants freely volunteered to participate without incentive and were recruited using a snowball sampling approach using extended contact networks and social media [15]. Participants were told of the length of the survey, and the research lead's name and contact details were displayed on the first and final pages. Figure 1 shows the pathway used in the study. The data were stored on password-protected servers at the University of Portsmouth.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram.



Instrument

The full survey was available on the web between November 1, 2018, and March 30, 2019, and advertised in Outdoor Swimmer magazine, which is a magazine for people who enjoy outdoor swimming. Once the survey had closed, duplicate responses were removed by using IP addresses and time stamps (30 surveys) and then data were deidentified from the IP address.

Participants voluntarily completed a web-based open survey hosted on a secure web-based survey platform with automatic back end data capture. Participants were asked to answer open-ended questions related to their participation in outdoor swimming. They identified where they swam, whether in outdoor facilities (lido or open-air pool) or in a natural setting (eg, the sea, lakes, and rivers), when they swam outdoors, their regularity of outdoor swimming, and their initial motives for taking up open water swimming. They were then asked to identify the physical and mental health conditions they had been diagnosed with, the impact and change in symptoms that were experienced while engaged in outdoor swimming, whether their symptoms had changed since outdoor swimming, and how long their symptoms changed for. There was no randomization of the survey items. Participants were able to press a back button to review their answers or resubmit their answers. A nonresponse option was also included in each question.

Pilot Testing

A pilot survey (n=10) was conducted between September 15, 2018, and October 20, 2018, to evaluate the instrument length, language, and logic. Modifications were made to the language, and additional answer options were added to closed questions to improve participant understanding and response rate.

Data and Statistical Analyses

Once duplicate responses had been removed, the data set was deidentified. Medical conditions were retrospectively coded using the health categories from the UK Clinical Research Collaboration [16]. Coding was initially undertaken by the corresponding author and cross-referenced with 2 practicing medical doctors (MH and LS). All conditions were coded, leading to some participants having multiple codes because of comorbidities. In a number of cases, no conditions were reported, or the conditions could not be categorized. In these cases, 2 additional categories were used: *no condition* and *not categorized*. The latter occurred primarily because of a vague description of symptoms (out of the 1084 conditions categorized, it occurred in a small number; n=30, 2.77%).

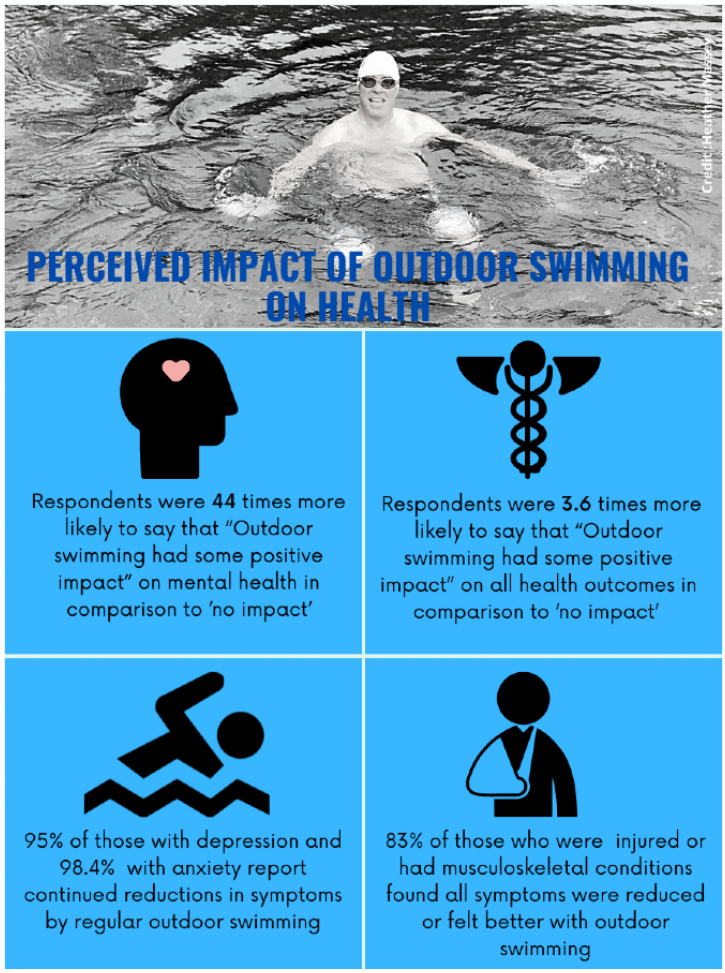
The 21 health condition categories were clustered into five main areas: (1) mental health; (2) injury, accidents, and musculoskeletal injuries; (3) neurological; (4) cardiovascular and blood; and (5) other. These 5 groups were chosen as they encompassed most of the responses. Many who were classified in the injury and accidents group were also classified into the musculoskeletal injuries group. Similarly, cardiovascular conditions were combined with blood disorders, as participants were frequently categorized in both separately. The *other* category contains the inflammatory and immune, metabolic and endocrine, and respiratory conditions clustered together because of the small number in the individual categories. Although there may be associations with this category as a whole, this may not translate into associations or impacts in all conditions within this group. The medical conditions contained within the *other* category are diverse, and for many of the conditions in this group (cancer, congenital disorders, and renal conditions), there are no anecdotal reports that the authors are aware of that

indicate improvements in symptoms as a consequence of outdoor swimming.

The perceived impact of outdoor swimming was measured on a 4-point scale (1=a lot of impact, 2=some impact, 3=little impact, and 4=no impact). Logistic regression analysis was used to test whether participants' medical categories could predict the perceived impact of outdoor swimming. The medical categories were clustered into 5 larger categories based on their prevalence in the current sample: mental health, musculoskeletal and injury, neurological, cardiovascular and blood disease, and *other*. These categories were entered into the analysis as predictors, and regression coefficients were estimated based on a bootstrapping procedure with 5000 successful replicates [17]. Logistic regression analyses were performed using the R-based statistical software JASP (JASP version 0.13.1). A detailed analysis and model fit check can be found in Figure 2.

Other statistical analyses were conducted using SPSS (version 25; IBM Corp). The frequency of responses per medical category was tabulated, and further chi-square analyses were performed in those medical categories that had at least 5% of the participants' self-reports. Chi-square tests of association analyzed the gender; regularity; impact of outdoor swimming on their conditions (both reduced and increased severity of symptoms); the change in symptoms, if any, resulting from outdoor swimming; and finally, whether the regularity of outdoor swimming was associated with changes in symptoms. For all tests of association, statistical significance was defined as $P<.05$. The strength of association was considered using Cramer *V*, based on the following thresholds: small=0.1, moderate=0.3, and large=0.5 [18].

Figure 2. Study infographic. Image credit: Heather Massey.



Results

Participant Characteristics

A total of 722 separate participants entered the survey analysis (Figure 1). Most of the participants were female (498/722, 68.9%), with further participants being male (159/722, 22%) or not stating sex (65/722, 9%). All participants swam outdoors in open water (lakes, rivers, the sea, lochs, quarries, lidos, and reservoirs). They reported swimming all year round (487/722, 67.5%) or seasonally in the summer and autumn (151/722, 20.9%) or only in the winter (1/722, 0.1%). A further group had recently taken up outdoor swimming and were not sure how long they would continue swimming once the water started to cool (58/722, 8%). The main reason for starting outdoor swimming included training for an event or challenge (138/722, 19.1%), liking to feel connected with nature (106/722, 13.7%), improved well-being (98/722, 13.6%), started as a child (70/722, 9.7%) and by a friend's invite (64/722, 8.9%), a change from pool swimming (36/722, 4.9%), and not sure (37/722, 5.1%).

Medical Category Prevalence

The number of respondents also reported multiple medical conditions, and the frequency of reports of single and comorbid conditions are reported in Table 1. A frequency table of the medical categories reported by the participants is shown in Table 2. Most medical conditions reported were in the mental health category (399/722, 55.3%), followed by musculoskeletal (153/722, 21.2%), no medical condition (76/722, 10.5%), neurological conditions (66/722, 9.1%), cardiovascular (61/722, 8.4%), respiratory (52/722, 7.2%), metabolic and endocrine (45/722, 6.2%), and finally, inflammatory or immune conditions (38/722, 5.3%). These medical categories are the main focus of further analysis.

Sex was significantly associated with the categories *no medical condition* (χ^2_2=11.2; *P*=.004) and neurological conditions (χ^2_2=7.3; *P*=.03). More males were associated with no conditions and more females with neurological conditions. No other medical categories had any association with sex.

The duration of symptoms before outdoor swimming was associated with mental health (χ^2_6=84.6; *P*<.001), musculoskeletal and endocrine (χ^2_6=19.4; *P*=.004), and respiratory categories (χ^2_6=19.5; *P*=.003). All were associated with long-term symptoms exceeding 10 years.

Table 1. Frequencies, reporting a single medical category, multiple medical conditions, or no medical condition (N=722).

Conditions	Values, n (%)
No condition	76 (10.5)
Single condition	401 (55.5)
Comorbid	245 (33.9)

Table 2. Frequency of medical conditions reported, percentage of responses, and percentage of participants reporting each medical category (from the UK Clinical Research Collaboration [16]).

Conditions	Responses (N=1084), n (%)	Participants (N=722[a]), n (%)
Not able to categorize	30 (2.8)	30 (4.2)
None	76 (7)	76 (10.5)
Blood	7 (0.6)	7 (1)
Cancer	13 (1.2)	13 (1.8)
Cardiovascular	61 (5.6)	61 (8.4)
Congenital disorder	7 (0.6)	7 (1)
Ear	2 (0.2)	2 (0.3)
Eye	4 (0.4)	4 (0.6)
Infection	14 (1.3)	14 (1.9)
Inflammatory and immune system	38 (3.5)	38 (5.3)
Injury and accidents	18 (1.7)	18 (2.5)
Mental health	399 (36.9)	399 (55.3)
Metabolic and endocrine	45 (4.2)	45 (6.2)
Musculoskeletal	153 (14.2)	153 (21.2)
Neurological	66 (6.1)	66 (9.1)
Oral gastrointestinal	17 (1.6)	17 (2.4)
Renal and urogenital	6 (0.6)	6 (0.8)
Reproductive health and childbirth	19 (1.8)	19 (2.6)
Respiratory	52 (4.8)	52 (7.2)
Skin	26 (2.4)	26 (3.6)
Stroke	0 (0)	0 (0)
Generic health relevance	0 (0)	0 (0)
Other	31 (2.9)	31 (4.3)

[a]The participant percentage will exceed 100% because of a number of participants reporting comorbidities.

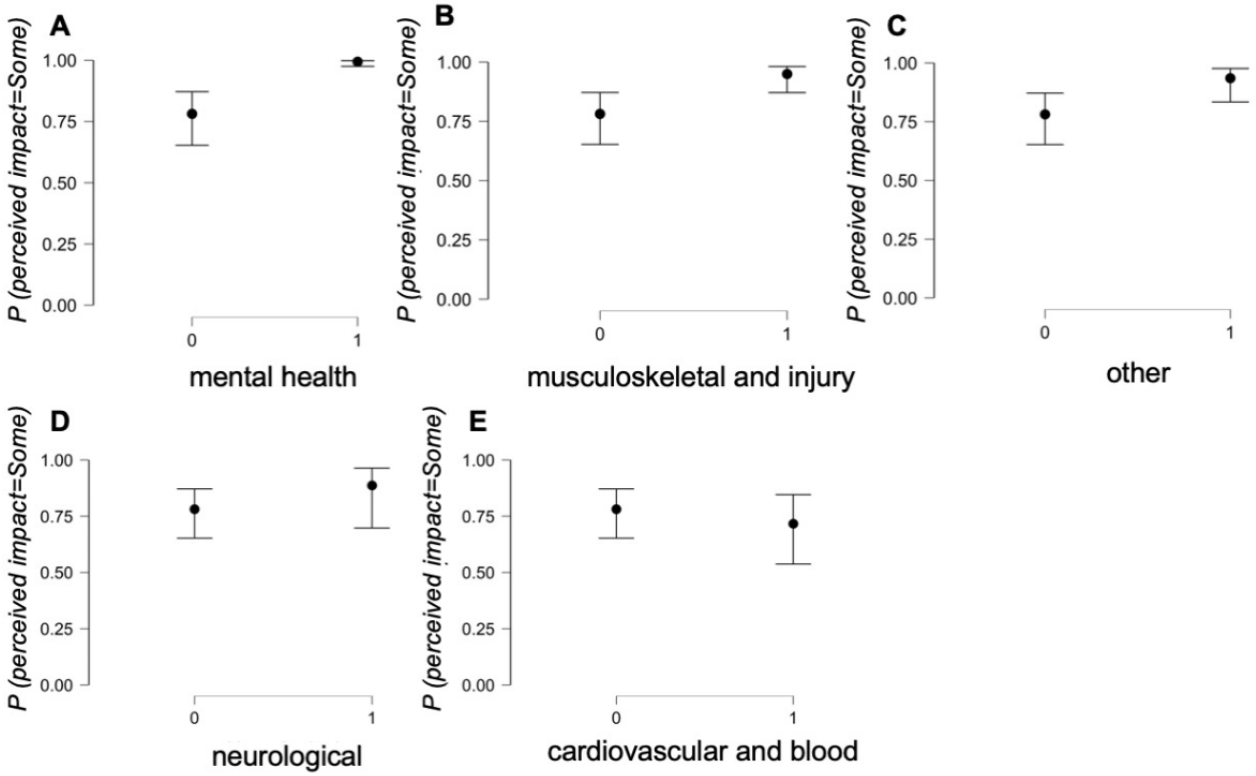
Impact of Outdoor Swimming

Approximately 89.6% (647/722) of participants reported at least one medical condition; their responses to the question "has open water swimming had any impact on your medical symptoms" established that 95.5% (618/647) found *some positive impact*, and 4.5% (29/647) found *no impact*. The logistic regression model was statistically significant (χ^2_{637}=60.8; *P*<.001; McFadden R^2=0.26). The model's intercept indicated that the probability of having *some positive impact* of outdoor swimming on health across all medical categories was 3.57 times higher than that of *no* impact (*B*=1.27, 95% CI 0.63-1.91; Wald statistics=15.13, degree of freedom=1; *P*<.001). The probability of having *some positive impact* of outdoor swimming on health was 44.32 times higher for the mental health category (*B*=3.79, 95% CI 2.28-5.30; Wald statistics=24.36, degree of freedom=1; *P*<.001), 5.25 times higher for the musculoskeletal and injury category (*B*=1.66, 95% CI 0.52-2.79; Wald statistics=8.26, degree of freedom=1; *P*=.004), and 4.02 times higher for the *other* category (*B*=1.39, 95% CI 0.27-2.51; Wald statistics=5.90, degree of freedom=1; *P*=.02; Figure 2; Table 3). For neurological (*B*=0.78, 95% CI −0.50 to 2.08; Wald statistics=1.42, degree of freedom=1; *P*=.23) and cardiovascular and blood (*B*=−0.34, 95% CI −1.27 to 0.58; Wald statistics=0.54, degree of freedom=1; *P*=.46) categories, the estimates were not statistically significant (Figures 2 and 3).

Table 3. Frequency of medical conditions reported into collapsed medical categories, percentage of responses, and percentage of participants reporting each medical category (N=917).

Conditions	Responses, n (%)	Participants, n (%)
Mental health	399 (43.5)	399 (62.7)
Injury and accidents and musculoskeletal	165 (18)	165 (25.9)
Neurological	66 (7.2)	66 (10.4)
Cardiovascular and blood	67 (7.3)	67 (10.5)
Other	200 (24)	200 (34.6)

Figure 3. Condition estimate plots. The probability of having impact of outdoor swimming on health (y-axis) plotted against responses no impact (denoted as 0) and some impact (denoted as 1) for 5 categories of medical conditions. Black dots represent the probability of impacts. The bars represent 95% CIs.



Characterizing Impact

The reasons given for starting outdoor swimming were numerous; however, associations were found for the mental health and musculoskeletal and injury categories. The respondents' reporting of mental health conditions indicated an association with starting outdoor swimming to train for an event or challenge (χ^2_{10}=36.1; *P*<.001). In contrast, those with musculoskeletal conditions or injuries initially participated in outdoor swimming as injuries prevented them from taking part in other forms of activity (χ^2_{10}=31.9; *P*<.001).

On further questioning, the perceived impact was a reduction in symptoms because of outdoor swimming in the following categories: mental health (χ^2_3=25.1; *P*<.001), musculoskeletal and injury (χ^2_3=8.2; *P*=.04), cardiovascular and blood (χ^2_3=14.7; *P*=.006), and *other* categories (χ^2_3=18.2; *P*<.001; Table 4). Furthermore, trends for a reduction in symptoms after open water swimming were reported in the mental health (χ^2_1=2.9; *P*=.08) and *other* categories (χ^2_1=2.7; *P*=.08); these symptom reductions in the mental health category were short-lived (χ^2_4=35.4; *P*<.001), lasting from several hours to 2 days after swimming outdoors. No other associations were found between the duration of reduced symptoms and the medical categories.

Considering the *dose of swimming or cold water*, no associations between swimming frequency and the impact or a change in symptoms were found in any medical category. The same is true of the choice of swimming attire (wetsuit or swimming costume). However, more respondents with cardiovascular and blood conditions wore wetsuits (χ^2_2=7.6; *P*=.21). In contrast, there was a trend for association in the *other* category, indicating that greater numbers of respondents found that their symptoms were affected if they swam all year round (χ^2_3=8.4; *P*=.10). In addition, for many in the musculoskeletal and injury (χ^2_3=9.8; *P*=.02), mental health (χ^2_3=10.0; *P*=.02), and *other* categories (χ^2_3=13.3; *P*=.04), the duration of swimming was primarily governed by the water temperature.

Table 4. Characteristics of the impact of outdoor swimming on the 5 combined medical categories. Percentages represent responses within each medical category (N=722).

Characteristics	Medical categories				
	Mental health (n=395)	Injury, accident, and musculoskeletal (n=161)	Neurological (n=63)	Cardiovascular and blood (n=62)	Other (n=207)
Symptom changes because of swimming (n=622)					
All symptoms reduced or feel better, n (%)	360 (91.1)	134 (83.2)	50 (79.4)	44 (71)	167 (80.7)
No change, n (%)	27 (6.8)	14 (8.7)	10 (15.9)	12 (19.4)	22 (10.6)
All symptoms increased or feel worse, n (%)	0 (0)	3 (1.9)	0 (0)	1 (1.6)	3 (1.4)
Some symptoms reduced some increased, n (%)	8 (2)	10 (6.2)	3 (4.8)	5 (8.1)	15 (7.2)
Chi-square (*df*)	25.1 (3)	8.2 (3)	4.8 (3)	14.7 (3)	18.2 (3)
P value	<.001[a]	.04[a]	.16	.006[a]	<.001[a]
Cramer *V*	0.201	0.115	0.088	0.154	0.171
Continued reduction in symptoms by open water swimming (n=594)					
Yes, n (%)	362 (96.5)	151 (95.6)	58 (65.1)	55 (98.2)	187 (93.5)
No, n (%)	13 (3.5)	7 (4.4)	3 (4.9)	1 (1.8)	13 (6.5)
Chi-square (*df*)	2.7 (1)	0.0 (1)	0.0 (1)	1.1 (1)	2.7 (1)
P value	.08[b]	.57	.54	.26	.08[b]
Cramer *V*	0.068	0.003	0.006	0.043	0.067
If you have continued symptom reduction how long does this last for? (n=576)					
Several hours, n (%)	46 (12.3)	25 (16.3)	11 (18.3)	4 (7.7)	21 (11.4)
1 to 2 days, n (%)	187 (50)	61 (39.9)	18 (30)	18 (34.6)	73 (39.7)
1 week, n (%)	31 (8.3)	8 (5.2)	3 (5)	5 (9.6)	11 (6)
>1 week, n (%)	49 (13.1)	38 (24.8)	13 (21.7)	14 (26.9)	41 (22.3)
Cannot say, n (%)	61 (16.3)	21 (13.7)	15 (25)	11 (21.2)	38 (20.7)
Chi-square (*df*)	35.4 (4)	8.4 (4)	6.9 (4)	5.7 (4)	5.0 (4)
P value	<.001[a]	.08[b]	.14	.22	.29
Cramer *V*	0.248	0.121	0.109	0.1	0.093
Regularity of swimming (n=634)					
Daily, n (%)	39 (9.8)	16 (9.7)	10 (16.1)	6 (9.1)	23 (10.6)
At least once a week, n (%)	317 (80.6)	133 (81.2)	43 (69.4)	55 (83.3)	171 (79.2)
Less than once a week or infrequently, n (%)	37 (9.4)	15 (9.2)	9 (14.5)	5 (7)	22 (10.2)
Chi-square (*df*)	0.2 (2)	0.0 (2)	5.9 (2)	0.3 (2)	0.7 (2)
P value	.88	.99	.05[a]	.89	.71
Cramer *V*	0.02	0.007	0.097	0.021	0.034

[a]Statistically significant at $P \leq .05$.

[b]Trend to statistical significance ($P \leq .10$).

Prevalence of Mental Health Conditions and Impact of Outdoor Swimming

Respondents reporting mental health conditions were the largest cohort in this survey. Table 5 contains a breakdown of the main mental health conditions reported by the participants. Of interest, 30.8% (123/399) participants reported a diagnosis of depression and anxiety, with a further 20.1% (80/399) diagnosed with anxiety alone and 41.1% (164/399) with depression alone. Furthermore, depression was associated with a longer period since diagnosis (χ^2_4=11.7; *P*=.03). Despite a large numerical

response in all conditions within the mental health category, no significant associations of the impact of outdoor swimming on symptoms of these conditions were found, except that a reduction in symptoms was associated with anxiety (χ^2_3=7.0; *P*=.03) and a trend for association with depression (χ^2_3=5.4; *P*=.09). In addition, reports of a continued reduction in symptoms were associated with diagnoses of anxiety or depression (anxiety χ^2_1=4.2, *P*=.04 and depression χ^2_1=5.7, *P*=.03). Therefore, 95% (251/264) of respondents with depression and 98.4% (188/191) of respondents with anxiety reported continued reductions in symptoms through regular outdoor swimming.

Table 5. Cross tabulations between mental health conditions and outdoor swimming survey item responses. Percentages have been calculated across all mental health conditions and responses.

Characteristics	Medical condition within the mental health category				
	Depression (n=287)	Anxiety (n=203)	Posttraumatic stress disorder (n=25)	Bipolar (n=11)	Other (n=40)
Gender (n=399)					
Female, n (%)	212 (53.2)	149 (37.4)	14 (3.5)	9 (2.3)	28 (7)
Male, n (%)	53 (13.3)	34 (8.5)	9 (2.3)	2 (0.5)	9 (2.3)
None given, n (%)	22 (5.5)	20 (5)	2 (0.5)	0 (0)	3 (0.8)
Chi-square (*df*)	0.8 (2)	3.7 (2)	4.7 (2)	1.0 (2)	0.3 (2)
P value	.68	.16	.09[b]	.70	.92
Cramer *V*	0.045	0.097	0.109	0.051	0.027
Symptom duration (n=399)					
<1 year, n (%)	3 (0.8)	5 (1.3)	0 (0)	0 (0)	2 (0.5)
1 to 4.9 years, n (%)	44 (11)	32 (8)	3 (0.8)	0 (0)	10 (2.5)
5 to 9.9 years, n (%)	18 (4.5)	21 (5.3)	1 (0.3)	1 (0.3)	1 (0.3)
≥10 years, n (%)	197 (49.4)	135 (33.5)	20 (5)	9 (2.3)	24 (6)
Not sure, n (%)	24 (6)	10 (2.5)	1 (0.3)	1 (0.3)	3 (0.8)
Chi-square (*df*)	11.7 (4)	9.2 (4)	2.8 (4)	2.5 (4)	6.7 (4)
P value	.03[a]	.08[b]	.63	.66	.24
Cramer *V*	0.171	0.152	0.084	0.08	0.129
Impact of swimming on condition (n=397)					
Yes, n (%)	284 (71.4)	201 (50.8)	25 (6.3)	11 (2.8)	39 (9.8)
No, n (%)	2 (0.5)	1 (0.3)	0 (0)	0 (0)	0 (0)
Chi-square (*df*)	0.8 (1)	<0.1 (1)	0.1 (1)	<0.1 (1)	0.2 (1)
P value	.59	.74	.88	.95	.81
Cramer *V*	0.044	0.001	0.018	0.012	0.024
Symptom changes because of swimming (n=395)					
All symptoms reduced or feel better, n (%)	255 (64.6)	188 (47.7)	24 (6.1)	11 (2.8)	36 (9.1)
No change, n (%)	24 (6.1)	7 (1.8)	0 (0)	0 (0)	3 (0.8)
All symptoms increased or feel worse, n (%)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Some symptoms reduced some increased, n (%)	5 (1.3)	4 (1)	1 (0.3)	0 (0)	0 (0)
Chi-square (*df*)	5.4 (3)	7.0 (3)	2.4 (3)	1.1 (3)	0.1 (3)
P value	.08[b]	.03[a,]	.31	.70	.75
Cramer *V*	0.106	0.134	0.078	0.053	0.049
Continued reduction in symptoms by regular open water swimming (n=391)					
Yes, n (%)	251 (69.3)	188 (50.3)	24 (6.4)	9 (2.4)	38 (10.2)
No, n (%)	13 (3.5)	3 (0.8)	0 (0)	0 (0)	0 (0)
Chi-square (*df*)	5.6 (1)	4.2 (1)	0.9 (1)	0.3 (1)	1.5 (1)
P value	.03[a]	.04[a]	.42	.72	.24
Cramer *V*	0.123	0.106	0.05	0.03	0.064
If you continued swimming, how long were symptoms reduced? (n=373)					
Hours, n (%)	127 (34)	76 (20.4)	9 (2.4)	1 (1)	11 (2.9)

Characteristics	Medical condition within the mental health category				
	Depression (n=287)	Anxiety (n=203)	Posttraumatic stress disorder (n=25)	Bipolar (n=11)	Other (n=40)
1 to 2 days, n (%)	73 (19.5)	60 (16.1)	9 (2.4)	2 (0.5)	12 (3.2)
1 week, n (%)	18 (4.8)	18 (4.8)	2 (0.5)	2 (0.5)	4 (1.1)
>1 week, n (%)	10 (2.7)	9 (2.4)	0 (0)	0 (0)	1 (0.3)
Cannot say, n (%)	42 (11.2)	27 (7.2)	5 (1.3)	3 (0.8)	10 (2.7)
Chi-square (*df*)	5.7 (4)	9.6 (4)	2.2 (4)	2.9 (4)	5.5 (4)
P value	.22	.04[a,]	.70	.51	.23
Cramer *V*	0.123	0.161	0.078	0.089	0.121

[a]Statistically significant at $P \leq .05$.

[b]Trend to statistical significance, $P \leq .10$.

Discussion

Principal Findings

The main findings suggest that physical activity in the form of outdoor swimming affects health and is associated with perceived improvements in some medical categories (mental health, musculoskeletal and injury, and *other*) but not all. The level of evidence at present remains anecdotal [11]. However, by performing the survey, it is clear which categories of medical conditions may have perceived benefits from outdoor swimming. Consequently, these data can be used to focus future research efforts in the conditions where anecdotal support was apparent to establish causality and, if so, potential mechanisms for symptom reduction.

Just over half (399/722, 55.3%) of the survey respondents reported a diagnosed mental health condition. This proportion is higher in comparison with the general population (before the COVID-19 lockdown, 9.7% and after the COVID-19 lockdown, 19.7% [19]), and it is unknown whether this is representative of the population of outdoor swimmers. The use of a single category to describe mental health conditions is limiting. Further subcategorization was possible in this study, with associations between a perceived reduction in symptoms in respondents' reporting of anxiety and depression. In contrast, there is insufficient evidence to suggest a reduction in symptoms of other mental health conditions, which may be because of the smaller number of respondents with these conditions in the present survey. Case report evidence exists of outdoor swimming as a nonmedicalized activity, supporting recovery from depression and anxiety [9]. It is also unclear whether outdoor swimming would result in symptom reduction in all forms of depression or the potential mechanisms at play [20]. In addition, it may also be that symptoms are not changed by the act of outdoor swimming, but the participants' perceptions of their symptoms or feelings of well-being are temporarily changed. However, it is important to separate well-being from mental health rather than them occupying the same continuum [21]. Therefore, it may be that outdoor swimming does not reduce their symptoms, but they do have a sense of greater well-being. Consequently, this study and qualitative aspects of the survey data that are yet to be published (McEwan, personal communication) provide a focus to build research designs that can establish causality and potentially the mechanisms involved.

The benefits of physical activity are well-documented and include a reduced risk of developing both physical and mental health problems and can support the treatment of pre-existing health conditions [22,23]. It is also acknowledged that exercise close to coastal areas or *blue spaces* may affect well-being [24] and mood [25]. Similar to terrestrial activity, swimming appears to positively affect markers of ill health [26], and swimming also offers the opportunity for reduced load bearing on muscles and joints [27]. This may support an explanation for the perception of reduced symptoms or pain in those with musculoskeletal conditions such as arthritis and back pain who do not feel able to be active on land [28]. Greater movement and pain reduction were also reported during immersion in cold water [29]. Therefore, for participants with musculoskeletal problems, outdoor swimming may allow greater movement, activity level, and respite from pain than is possible on land; however, further research is needed to confirm this hypothesis.

Outdoor swimming, for many of the survey respondents has become a lifestyle choice, opting to be frequently physically active (at least once a week) in outdoor water environments. Overall, there was no association between the frequency of swims, swim attire, or swim duration and the impact or reports of symptom reduction. However, there was a trend for association between having an impact and swimming all year round in the *other* category. In addition, the duration of swimming was associated with the water temperature, weather, and sea state. Longer duration swims take place in warmer, calmer water, and shorter swims take place in cooler, rougher water. Therefore, the dose of outdoor swimming cannot be easily prescribed and will depend on the individuals' physiological responses to immersion in cold water, the location, sea state or moving water, and weather conditions. Cold water swimming is not a risk-free activity; individuals with underlying cardiac and cardiovascular conditions may be at elevated risk of adverse cardiac events upon initial immersion in cold water, and those who are unable to keep their airway clear of the water are at risk of drowning because of the cold shock response [30]. Therefore, consultation with a general practitioner is recommended for those wanting to try outdoor swimming and have an underlying medical condition. It is also recommended

that those new to outdoor swimming join a group or swim with a trained open water swimming coach with good local knowledge of the environment and are able to convey the knowledge and skills required to swim safely outdoors.

In this survey, respondents provided information on their motives for starting to swim outdoors. These motives ranged from training for an event to looking for a change from a swimming pool, with some individuals not being sure why they started. Some people clearly stated that it was to improve the elements of their mental health. As seen in other research studies [12], there are several reasons why people begin to and maintain swimming outdoors. Although we may better understand some of these motives, we still do not know whether swimming outdoors has medical utility. Before any recommendations can be made as to the health effects of swimming outdoors, the evidence must clearly demonstrate both association and causation with swimming outdoors and improved health [31]. This research forms part of the development process and our continuing research efforts to better understand and demonstrate evidence of the health impact of outdoor swimming.

Limitations

It is acknowledged that the survey was conducted on people currently swimming in open water, and therefore, those who are likely to have a very positive viewpoint about the activity. Although the views expressed in this paper and in the sister paper have been commonly heard by the authors, it is not clear if this is representative of all outdoor swimmers or those who do not continue to swim outdoors. However, the research was conducted as an internet-based open survey; therefore, no coercion took place for swimmers to participate or give particular answers. In addition, duplicate submissions were removed to prevent the overrepresentation of one person's viewpoint. Furthermore, there may be participants who feel compelled to support outdoor swimming for health improvement. However, most swimmers did not start outdoor swimming to find a means of improving health, but many self-reported health improvements as a consequence of their swimming.

The survey grouped medical conditions into categories, which has limitations in terms of specificity to individual conditions. However, the vast range of potential disease conditions made this the most practical step to take. There were generally a small number of conditions that were more common; for example, in the mental health category, diagnoses of depression and anxiety were common and have been interrogated more deeply than other categories. Similarly, migraine was the most frequent condition in the neurological category. In medical categories with small numbers, this should not rule out any impact that the activity may have on their symptoms, and the number of participants with these conditions was too small to provide a statistical association. The survey as an initial investigation is adequate to gather together anecdotal self-reports; however, further, more in-depth investigation is required. Potential avenues for progressing the research include corroborating patient-reported outcomes with independent medical judgment and inclusion of condition-specific surveys to assess the patient-reported outcomes in a more sensitive, specific, and systematic manner. However, this approach now seems more feasible and can be better targeted, given the information in this paper.

The type of medical condition may limit some of the given responses; for instance, migraine attacks are sporadic and not easy for people to determine how severe or long an attack would last or even if swimming in cold water caused a reduction in symptoms. However, the participants may still have future migraines; therefore, outdoor swimming has not *cured* the condition. Therefore, although they may report some acute symptom reduction, it is unclear how frequently they might or will have these attacks, and without detailed logs of attacks and severity scoring, it can be challenging to establish whether the intervention, in this case, outdoor swimming, is affecting their symptoms.

Opportunities for Further Study

It is well-established that people's perceptions of their illness are crucial to health care providers to evaluate and support the patient's empowerment and self-care ability. An individual's perception of factors related to their illness or symptoms influences their coping behavior [32], trust in themselves, beliefs that they can manage their illness and prevent it from becoming worse [33]. Therefore, an individual's experience in noticing and interpreting their experience of bodily changes during self-care activities such as engaging in outdoor swimming is important information for health care providers about the efficiency of such activities in promoting self-care. Furthermore, a person's perception of improvement in health has a positive effect [34]. This relationship is conceptualized in the Common Sense Model of Illness Perception, which proposes that the positive beliefs that a person holds regarding their illness directly lead to better mental well-being and the development of active coping strategies [35]. However, validation of perceived improvements in health would need clinical interventions that would provide invaluable information about alternative methods for treating many physical and mental health conditions. Therefore, further research is needed to establish whether swimmers' perceptions of improved health translate into outdoor swimming being effective in reducing symptoms of poor mental and physical health. Potential research should look to establish the clinical and cost-effectiveness of outdoor swimming as a treatment intervention. In addition, it would also need to fall in line with current policies such as the National Health Service Long-Term Plan [36], Mental Health Implementation Plan [37], the Community Mental Health Framework for Adults and Older Adults [38], and the Garside report [39]. These policies recommend the need for personalized care and patient choice and provide a framework for delivery. In particular, the Garside report [39] describes the limited evidence base for nature-based activities. So far, the main focus has been on land-based activities, with few water-based interventions. Therefore, further studies to establish the impact on mental and physical health conditions could use clinical trial methodologies. For instance, these may include randomized control trials with mixed method approaches to establish whether the outdoor swimming intervention is acceptable and has a therapeutic effect. If a therapeutic effect is found, for whom does it have an effect (in terms of a medical condition, patient age, and socioeconomic

status), and how does that effect occur (through psychological, sociological, or physiological mechanisms, or more likely a combination of all 3)? Further avenues for research into the perception of impact on health conditions may be considered, such as *why participants perceive improvements*. Such research may include the use of innovative qualitative study methods; one such method was explored with great effect by Denton and Aranda [12].

Conclusions

In conclusion, physical activity in the form of outdoor swimming was perceived to have a positive impact on health and is associated with perceived improvements in some medical conditions, namely mental health, musculoskeletal, and cardiovascular conditions. For many, not just those reporting a reduction in symptoms of a medical condition, outdoor swimming has become a lifestyle choice to be physically active in cold water. Although this study cannot provide causal relationships, unpick the reason for symptom reduction, or provide mechanistic insight, it does provide a starting point for more targeted research into individual conditions or categories of conditions in those who choose or would like to start outdoor swimming.

Acknowledgments

The authors wish to thank the study participants for their support in completing the survey and those who shared social media advertisements.

Authors' Contributions

HM, CMH, and HD conceived and designed the study; HM, collated the data; HM, LS, PG, CMH, and AY analyzed the data; and HM, HD, and KM drafted the manuscript. All authors contributed to the interpretation of the data and critically revised the manuscript.

The data that support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

None declared.

References

1. Active lives. Sport England. 2018. URL: https://www.sportengland.org/research/active-lives-survey/ [accessed 2021-12-07]
2. Nuckton TJ, Claman DM, Goldreich D, Wendt FC, Nuckton JG. Hypothermia and afterdrop following open water swimming: the Alcatraz/San Francisco Swim Study. Am J Emerg Med 2000 Oct;18(6):703-707. [doi: 10.1053/ajem.2000.16313] [Medline: 11043627]
3. Berger BG, Owen DR. Mood alteration with swimming--swimmers really do "feel better". Psychosom Med 1983 Oct;45(5):425-433. [doi: 10.1097/00006842-198310000-00006] [Medline: 6635105]
4. Foley R. Swimming in Ireland: immersions in therapeutic blue space. Health Place 2015 Sep;35:218-225. [doi: 10.1016/j.healthplace.2014.09.015] [Medline: 25456012]
5. Throsby K. 'If I go in like a Cranky Sea Lion, I Come out like a Smiling Dolphin': marathon swimming and the unexpected pleasures of being a body in water. Feminist Rev 2013 Mar 01;103(1):5-22. [doi: 10.1057/fr.2012.23]
6. Pike EC. Aquatic antiques: swimming off this mortal coil? Int Rev Sociol Sport 2011 Mar 11;47(4):492-510. [doi: 10.1177/1012690211399222]
7. Pasanen TP, White MP, Wheeler BW, Garrett JK, Elliott LR. Neighbourhood blue space, health and wellbeing: the mediating role of different types of physical activity. Environ Int 2019 Oct;131:105016 [FREE Full text] [doi: 10.1016/j.envint.2019.105016] [Medline: 31352260]
8. White M, Pahl S, Wheeler B, Fleming L, Depledge M. The 'Blue Gym': what can blue space do for you and what can you do for blue space? J Mar Biolog Assoc UK 2016 Jan 19;96(1):5-12. [doi: 10.1017/s0025315415002209]
9. van Tulleken C, Tipton M, Massey H, Harper CM. Open water swimming as a treatment for major depressive disorder. BMJ Case Rep 2018 Aug 21 [FREE Full text] [doi: 10.1136/bcr-2018-225007] [Medline: 30131418]
10. Costello L, McDermott M, Patel P, Dare J. 'A lot better than medicine' - Self-organised ocean swimming groups as facilitators for healthy ageing. Health Place 2019 Nov;60:102212 [FREE Full text] [doi: 10.1016/j.healthplace.2019.102212] [Medline: 31610442]
11. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, et al. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) levels of evidence. Centre for Evidence-Based Medicine, University of Oxford. 2011. URL: https://www.cebm.ox.ac.uk/resources/levels-of-evidence/explanation-of-the-2011-ocebm-levels-of-evidence [accessed 2021-12-07]
12. Denton H, Aranda K. The wellbeing benefits of sea swimming. Is it time to revisit the sea cure? Qual Res Sport Exerc Health 2019 Sep 02;12(5):647-663. [doi: 10.1080/2159676x.2019.1649714]
13. Parr S. The Story of Swimming. Heaton Moor, Stockport, UK: Dewi Lewis Media Ltd; 2011:1-192.

14. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Dec 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
15. Seale C. Researching Society and Culture (3rd ed). Thousand Oaks, CA: SAGE Publications Ltd; 2012:1-656.
16. Health Research Classification System. UK Clinical Research Collaboration. 2018. URL: http://hrcsonline.net/wp-content/uploads/2018/01/HRCS_Main_Handbook_v2_Feb2018.pdf [accessed 2021-12-07]
17. Adjei IA, Karim R. An application of bootstrapping in logistic regression model. OALib 2016;03(09):1-9. [doi: 10.4236/oalib.1103049]
18. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd Ed. New York: Routledge; 1988.
19. Coronavirus and depression in adults, Great Britain: June 2020. Office for National Statistics. 2020. URL: https://tinyurl.com/ycx36hhz [accessed 2021-12-07]
20. Tipton M, Collier N, Massey H, Corbett J, Harper M. Cold water immersion: kill or cure? Exp Physiol 2017 Sep 21;102(11):1335. [doi: 10.1113/ep086283]
21. Atkinson S. Beyond components of wellbeing: the effects of relational and situated assemblage. Topoi (Dordr) 2013 May 17;32(2):137-144 [FREE Full text] [doi: 10.1007/s11245-013-9164-0] [Medline: 25821285]
22. Depression in adults: recognition and management. NICE Clinical guideline [CG90]. 2009. URL: https://www.nice.org.uk/guidance/cg90 [accessed 2021-12-07]
23. Physical activity: walking and cycling. NICE Public health guideline [PH41]. 2012. URL: https://www.nice.org.uk/guidance/ph41 [accessed 2021-12-07]
24. Shellock R. The wellbeing and human health benefits of exposure to the marine and coastal environment. Evidence statement 7. Department for Environment, Food and Rural Affairs. 2019. URL: https://tinyurl.com/ycyu3db3 [accessed 2021-12-07]
25. Massey H, Kandala N, Davis C, Harper M, Gorczynski P, Denton H. Mood and well‐being of novice open water swimmers and controls during an introductory outdoor swimming programme: a feasibility study. Lifestyle Med 2020 Nov 10;1(2). [doi: 10.1002/lim2.12]
26. Lazar JM, Khanna N, Chesler R, Salciccioli L. Swimming and the heart. Int J Cardiol 2013 Sep 20;168(1):19-26. [doi: 10.1016/j.ijcard.2013.03.063] [Medline: 23602872]
27. Moffat F. The individual physical health benefits of swimming: a literature review. In: Cumming I, editor. The Health and Wellbeing Benefits of Swimming: A Collection of Seven Academic Reviews on the Health and Wellbeing Benefits of Swimming. Leicestershire: Swim England; 2017:1-156.
28. Hall J, Swinkels A, Briddon J, McCabe CS. Does aquatic exercise relieve pain in adults with neurologic or musculoskeletal disease? A systematic review and meta-analysis of randomized controlled trials. Arch Phys Med Rehabil 2008 May;89(5):873-883. [doi: 10.1016/j.apmr.2007.09.054] [Medline: 18452734]
29. Mole TB, Mackeith P. Cold forced open-water swimming: a natural intervention to improve postoperative pain and mobilisation outcomes? BMJ Case Rep 2018 Feb 12 [FREE Full text] [doi: 10.1136/bcr-2017-222236] [Medline: 29440051]
30. Tipton M. The initial responses to cold-water immersion in man. Clin Sci (Lond) 1989 Dec;77(6):581-588. [doi: 10.1042/cs0770581] [Medline: 2691172]
31. Sallis JF, Owen N, Fotheringham MJ. Behavioral epidemiology: a systematic framework to classify phases of research on health promotion and disease prevention. Ann Behav Med 2013 Jul 30;22(4):294-298. [doi: 10.1007/bf02895665]
32. Leventhal H, Nerentz DR, Steele DR. Illness perception and coping with health threat. In: Handbook of Psychology and Health, Volume IV. Hillsdale, NJ: Erlbaum Associates; 1984:219-252.
33. Alsén P, Brink E, Persson L. Patients' illness perception four months after a myocardial infarction. J Clin Nurs 2008 Mar;17(5A):25-33. [doi: 10.1111/j.1365-2702.2007.02136.x] [Medline: 18298753]
34. Hurt CS, Burn DJ, Hindle J, Samuel M, Wilson K, Brown RG. Thinking positively about chronic illness: an exploration of optimism, illness perceptions and well-being in patients with Parkinson's disease. Br J Health Psychol 2014 May 19;19(2):363-379. [doi: 10.1111/bjhp.12043] [Medline: 23510498]
35. Leventhal H, Brissette I. The common-sense model of self-regulation of health and illness. In: Cameron L, Leventhal H, editors. The Self-Regulation of Health and Illness Behaviour. London, UK: Routledge; 2002:42-65.
36. The NHS Long Term Plan. NHS England. 2019. URL: https://www.longtermplan.nhs.uk/publication/nhs-long-term-plan/ [accessed 2021-12-07]
37. NHS Mental Health Implementation Plan 2019/20 – 2023/24. NHS England. 2019. URL: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/07/nhs-mental-health-implementation-plan-2019-20-2023-24.pdf [accessed 2021-12-07]
38. The community mental health framework for adults and older adults. The National Collaborating Centre for Mental Health. 2019. URL: https://www.rcpsych.ac.uk/improving-care/nccmh/care-pathways/community-framework [accessed 2021-12-07]
39. Garside R, Orr R, Short R, Lovell R, Husk K, McEachan R, et al. Therapeutic Nature: nature-based social prescribing for diagnosed mental health conditions in the UK. Final Report for DEFRA. 2020. URL: https://beyondgreenspace.files.wordpress.com/2021/06/15138_therapeuticnature-finalreport-1.pdf [accessed 2021-12-07]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys

Edited by G Eysenbach; submitted 07.11.20; peer-reviewed by R Kretschmann, IV George; comments to author 25.08.21; revised version received 19.10.21; accepted 17.11.21; published 04.01.22.

Please cite as:
Massey H, Gorczynski P, Harper CM, Sansom L, McEwan K, Yankouskaya A, Denton H
Perceived Impact of Outdoor Swimming on Health: Web-Based Survey
Interact J Med Res 2022;11(1):e25589
URL: https://www.i-jmr.org/2022/1/e25589
doi:10.2196/25589
PMID:34982711

Original Paper

Digital Scientific Platform for Independent Content in Neurology: Rigorous Quality Guideline Development and Implementation

Daniel Kantor[1,2*], MD; Martin Farlow[3*], MD; Albert Ludolph[4,5*], MD; Joan Montaner[6*], MD; Raman Sankar[7*], MD; Robert N Sawyer Jr[8*], MD; Fabrizio Stocchi[9,10*], MD; Agnès Lara[11*], PhD; Sarah Clark[12*], PharmD; Karine Deschet[12*], PhD; Loucif Ouyahia[12*], PharmD; Yacine Hadjiat[12*], MD

[1]Florida Atlantic University, Boca Raton, FL, United States
[2]Nova Southeastern University, Fort Lauderdale, FL, United States
[3]Department of Neurology, School of Medicine, Indiana University, Indianapolis, IN, United States
[4]Department of Neurology, University of Ulm, Ulm, Germany
[5]Deutsches Zentrum für Neurodegenerative Erkrankungen, Ulm, Germany
[6]Department of Neurology, Hospital Universitario Virgen Macarena, Seville, Spain
[7]Division of Neurology, Department of Pediatrics, UCLA Mattel Children's Hospital, Los Angeles, CA, United States
[8]Department of Neurology, University at Buffalo, State University of New York, New York, NY, United States
[9]University San Raffaele Roma, Rome, Italy
[10]Istituto di Ricovero e Cura a Carattere Scientifico San Raffaele, Roma, Italy
[11]Medicom Concept, Occitanie, France
[12]Biogen Digital Health, Cambridge, MA, United States
[*]all authors contributed equally

Corresponding Author:
Yacine Hadjiat, MD
Biogen Digital Health
225 Binney Street
Cambridge, MA, 02142
United States
Phone: 1 781 464 2000
Email: yacine.hadjiat@biogen.com

Abstract

Background: Digital communication has emerged as a major source of scientific and medical information for health care professionals. There is a need to set up an effective and reliable methodology to assess and monitor the quality of content that is published on the internet.

Objective: The aim of this project was to develop content quality guidelines for Neurodiem, an independent scientific information platform dedicated to neurology for health care professionals and neuroscientists. These content quality guidelines are intended to be used by (1) content providers as a framework to meet content quality standards and (2) reviewers as a tool for analyzing and scoring quality of content.

Methods: Specific scientific criteria were designed using a 5-point scale to measure the quality of curated and original content published on the website: for Summaries, (1) source reliability and topic relevance for neurologists, (2) structure, and (3) scientific and didactic value; for Congress highlights, (1) relevance of congress selection, (2) congress coverage based on the original program, and (3) scientific and didactic value of individual abstracts; for Expert points of view and talks, (1) credibility (authorship) and topic relevance for neurologists, (2) scientific and didactic value, and (3) reliability (references) and format. The criteria were utilized on a monthly basis and endorsed by an independent scientific committee of widely recognized medical experts in neurology.

Results: Summary content quality for the 3 domains (reliability and relevance, structure, and scientific and didactic value) increased in the second month after the implementation of the guidelines. The domain *scientific and didactic value* had a mean score of 8.20/10. Scores for the domains *reliability and relevance* (8-9/10) and *structure* (45-55/60) showed that the maintenance of these 2 quality items over time was more challenging. Talks (either in the format of interviews or slide deck–supported scientific

presentations) and expert point of view demonstrated high quality after the implementation of the content quality guidelines that was maintained over time (15-25/25).

Conclusions: Our findings support that content quality guidelines provide both (1) a reliable framework for generating independent high-quality content that addresses the educational needs of neurologists and (2) are an objective evaluation tool for improving and maintaining scientific quality level. The use of these criteria and this scoring system could serve as a standard and reference to build an editorial strategy and review process for any medical news or platforms.

(Interact J Med Res 2022;11(1):e35698) doi:10.2196/35698

KEYWORDS

digital health; Neurodiem; neurology; neuroscience; eHealth; methodology; content quality; guideline; platform; development; implementation; quality; brain; communication; health information; health care professional; assessment; monitoring

Introduction

Over the last three decades, digital communication has emerged as a major source of medical information for health care professionals in a wide range of specialties, including neurology [1-3]. Web-based medical content is a valuable resource to provide updates on the latest medical news, share information with peers, and support education [4,5]. In light of the rapid growth in the number of medical websites with various content streams, either curated or original, the quality of scientific information that is disseminated through these tools has become a critical issue [6,7]. Thus, there is a need to ensure that digitally published scientific information dedicated to health care professionals is accurate, credible, relevant, and unbiased.

No single standard exists to objectively evaluate the quality of medical information available on the internet [8], although frameworks and measurement tools that provide users, including clinicians and researchers, with quality assessment options when navigating on an information platform have been proposed [9-17]. A simple and reliable methodology for screening and assessing the quality of web-based information could be a first step in developing optimal editorial guidelines that can be applied to digital scientific content and allow the production and sharing of reliable and accurate materials for the medical community A quality process for web-based medical communication should be equivalent to those established for peer-reviewed scientific journals and include scoring grids of criteria to assess the quality, originality, and relevance of published content.

Neurodiem (Biogen Inc) is a free multilingual and multicountry digital platform for independent information and education—the latest news and literature in neuroscience—dedicated to health care professionals and scientists Neurodiem is nonpromotional and provides strictly independent, impartial, and unbiased scientific content (in particular, with respect to Biogen's drug portfolio and therapeutic areas of interest). To meet this requirement, content published on Neurodiem is selected or generated exclusively by third-party publishers based upon the advice of advisory committees of neurologists and the published information covers all subspecialties of neurology in a balanced manner.

In addition, content on Neurodiem is monitored by a scientific steering committee of expert neurologists (who oversaw the development of the platform and who contributed to the development of these content quality guidelines).

The platform publishes content such as (1) summaries of curated peer-reviewed journal papers (extractions of the most relevant information from a published paper), (2) coverage of and comments on communications from neurology conferences, and (3) presentations and talks on currently debated topics by expert neurologists, and also provides (4) access to a selection of full-text papers from top-tier journals in neuroscience.

The Neurodiem steering committee developed a robust approach to evaluate scientific quality standards for content, while maintaining the independence of delivered information—scientific and medical content quality guidelines for Neurodiem content were first released in January 2020 (Multimedia Appendix 1) and described not only governance, roadmap and workflows for the quality process, but also, proposed custom criteria and a scoring system for the analysis and monitoring of scientific and medical various content stream quality.

The objective of this paper is to respond to an unmet need for content quality control with a new methodology for the assessment of platform content quality.

Methods

Content Quality Criteria and Rating System

The Neurodiem steering committee developed content quality guidelines in order to maintain and improve the quality of scientific content published on the website. We shared these content quality guidelines with stakeholders involved in the content quality process. First, content providers required specifications for scientific and medical quality standards adapted to and expected by a neurology audience, and second, scientific reviewers received an objective quality criteria and rating system which identifies areas of improvement, which can be shared with content providers on an ongoing basis.

A grid with specific quality criteria was designed for each content stream—summaries, congress highlights, expert points of view, and expert talks—published on Neurodiem (except for licensed content, ie, full-text journal papers that have already gone through a peer-review process).

Quality criteria were defined and organized into domains to support granular analysis of content quality based on scientific

relevance as well as editorial and journalistic standards featuring information dedicated to a specialized health care professional audience.

A 5-point scale from 1 (lowest score) to 5 (highest score) was used to rate each criterion. Each quality domain rating was multiplied by a criticality coefficient, and the products were summed for the total quality score; each domain was validated by the steering committee. This quality evaluation method was developed using the Neurodiem platform as an example, and it is most appropriate to Neurodiem content and is specific to each section of the platform. Using generalized evaluation grids could be less accurate and not applicable for all sections, which is why a new flexible method was needed for Neurodiem and similar platforms.

Review Process

In order to analyze and rate the scientific quality of Neurodiem content, 4 countries (France, Germany, Spain, and Italy) were selected for pilot implementation of the content quality guidelines for a 1-year duration. Reviewers had to (1) be a native speaker of a language represented on the Neurodiem local platform, (2) have a high-level scientific and medical profile (including a solid background in neuroscience or neurology), (3) have experience writing, reviewing, and editing scientific content, and (4) commit to a minimum of 1 year, in order to ensure content review homogeneity and be able to perform long-term assessment of the process.

In order to avoid any influence of the review process on content selection and production, the review process was performed after publication (usually within 1 month). Scientific content quality was assessed on a monthly basis for 1 year (10 months; no reviews were performed in August and December).

Due to the high volume of content published on Neurodiem, the quality review process was only performed on a representative sample—papers and talks were selected from the 18 neurology topics on Neurodiem (cognition, critical care, dementia, epilepsy, genetics, headache, imaging, movement disorders, multiple sclerosis, neuro-oncology, neuro-ophthalmology, neuromuscular, neurosurgery, pain, pediatric neurology, rehabilitation, sleep, and stroke) for comprehensive and balanced coverage. The content submitted to review was selected either randomly or if subject matter experts identified the topic as particularly challenging from a scientific accuracy or complexity perspective (eg, cutting-edge imaging, biotechnology, genetics-related content).

We monitored the quality of content published on the Neurodiem website over a 1-year period. The defined target for the monitoring rate was 20% of published content. The review process for Summary content was carried out by a single reviewer to ensure evaluation homogeneity over the period. Baseline data (month 1) were collected when Neurodiem content quality guidelines were not yet in place. The content quality guidelines were implemented in month 2, and analytics and quality improvement objectives were shared with content providers on a monthly basis.

Methodology Endorsement by a Scientific Committee

After 1 year, an independent committee, which consisted of 8 international neurology experts (based on experience and subspecialties in neurology, willingness to work in a digital field, experience using Neurodiem, and their availability) from Germany, Italy, Spain, Canada, and the United States, was involved in providing guidance and evaluating the scientific validity of the quality review process. The experts could not be involved in any commercial activities with Biogen over their period of engagement (to ensure their independence from the project sponsor). The scientific committee members were asked to review selected sections of Neurodiem content quality guidelines (dementia, epilepsy, movement disorders, multiple sclerosis, neuromuscular disorders and rare diseases, neurovascular diseases, and pediatric neurology), and then, were individually interviewed to ascertain their feedback and suggestions. Neurologists' advice and proposals related to definitions, wording, and validity of quality criteria, as well as scoring used for rating, were synthesized in a group meeting. An updated version of the Neurodiem content quality guidelines was released in January 2021 (Multimedia Appendix 1).

Scientific Quality Specifications and Measurement Tool

Summaries

Overview

Paper summaries cover curated full-text papers published in peer-reviewed journals in the field of neuroscience. The scientific quality of a summary depends on the accurate and succinct articulation of the content from the source paper in line with the predefined format for this content type (Multimedia Appendix 1). Three quality domains were defined: (1) reliability and relevance (of the source content and of the topic) with respect to an audience comprising neurologists, (2) structure of the summary, and (3) the scientific and didactic value of the summary.

Reliability and Relevance

Criteria are listed in Table 1.

Table 1. Reliability and relevance criteria.

Item	Criterion description
Journal	Papers curated for generating summaries should be originally published in high-impact factor or renowned peer-reviewed journals in neurology or neuroscience, to target scientific information primarily validated by a board of editors and reviewers. The journal quality assessment is based on H-index classification used in the field of Clinical Neurology, which was recommended by the scientific steering committee. The score reflects a neurologist's quality assessment of the journal: 1 for a paper not curated from a peer-reviewed scientific journal; 2 for a journal not classified in the SCImago Journal Rank; 3 for H-index values <30, 4 for H-index values 30-69, and 5 for H-index values ≥70.
Topic	Selected papers should be representative of current and major and scientific news at the forefront of information in each neurology subspecialty. According to the needs and interests of Neurodiem audience, selected topics should preferentially have direct impact on clinical practice or translate into major changes of the research and development landscape in neurology. The topic is scored on a scale from 1 to 5, based on the contribution in the neurology field or direct or the immediate impact in the clinical practice based on the author's conclusion.

Structure

Items in the *structure* domain (Table 2) were given a score between 1 (strongly disagree) and 5 (strongly agree), based on accuracy and the informative nature, ability to be understood, and attractive value of the original paper's content. Item scores were summed to generate a domain score out of 25 (Multimedia Appendix 1).

Table 2. Structure criteria.

Item and subitems	Description
Title and teaser text	The title and teaser text are the entry points to the paper summary on the home page and, thus, require particular attention. These 2 elements were evaluated on the basis of accuracy of the information, attractivity, clarity, and conciseness. The title should reflect the actual and main findings of the original paper. The teaser text should be distinct from the title text and provide more information while leaving the readers curiosity opened to explore the paper.
Take away	This section should contain 1 to 2 sentences to summarize the main findings of the source paper. Considered to be independent from the rest of the summary, this section is evaluated according to the clarity and relevance of the main results supporting the authors' conclusions.
Why this matters	In the format of 2 bullet points, this section is evaluated based on whether or not the structure and information on the clinical practice included in the original paper are respected. This section should (1) provide contextual information about the state of the art prior to the study and why it was interesting to explore the subject and (2) highlight study results' critical clinical implications or impact, in terms of disease mechanisms or pathophysiological paradigm changes, candidate molecule development, anticipated switch of clinical practices and content quality guidelines, in neuroscience.
Study design	This section is evaluated based on whether the main material and methods used in the study, focused on the key elements in relation to the study results, are summarized, complete and accurate. Layout features for this paragraph should imperatively include
Study objective	• The primary endpoint; when relevant, the secondary endpoints • The characteristics and size of the analyzed population, subpopulations, if applicable; animal model of pathologies will also be defined if needed • The study design, in particular, groups being compared • The follow-up duration and critical time points of analysis • The description of the procedures, clinical scales, or parameters being measured as well as the rationale of these measurements (ie, the expected outcomes. Synthetic background information on investigations performed may be provided when dealing with cutting-edge technologies not obviously known by any subspecialists in neurology)
Key results	The *Key results* section should provide an adequate and concise description of the major findings of the work. Primary endpoint–related results should be prioritized. Secondary endpoints may be included when extending the field of knowledge or paving the way for new hypotheses to be tested. It is important to illustrate remarkable results by providing numerical data with exact *P* values together with confidence intervals to illustrate the effect size and clinical value of data. It is recommended that the study's key take-home messages, including medical interpretations, be summarized at the end of this section in order to facilitate the retention of this information by readers.
Limitations	The *Limitations* section should briefly summarize study methodological characteristics potentially impacting findings interpretation and usually addressed in the discussion part of the source paper.

Scientific and Didactic Value

The domain *scientific and didactic* comprises 2 items (Table 3) scored from 1 (strongly disagree) to 5 (strongly agree), which were summed to generate the domain score (Multimedia Appendix 1).

Table 3. Scientific and didactic value criteria.

Item	Description
Accuracy	The paper summary must represent the original curated content with (1) accurate scientific glossary, abbreviations, and the numerical and statistical data described in the original full-text paper; (2) the accurate and relevant summary of the methodology, results, interpretation, conclusion, and impact in the clinical setting of the original full-text paper.
Didactic dimension	The didactic dimension is evaluated based on whether or not the Summary is clear, succinct, and comprehensible at first reading. It is important to provide enough background information, including context and scientific and medical definitions that are not common or shared among the neurologist community (eg, gene and protein functions, mode of action of new molecules, expected outcomes from emerging technologies). Understandability and readability of the Summary should be also supported by critical data that are presented logically and coherently.

Overall Quality

To obtain the overall quality score for paper Summaries, each domain subscore was weighted by a coefficient: 1, for reliability and relevance; 2, for structure; and 4, for scientific and didactic value. The products were summed to generate a total score out of 100.

Congress Highlights

Overview

Congress highlights present coverage of posters or oral communications from international and national top-tier conferences in neurology in an abstract format. The Neurodiem editorial team proposes a mean coverage of 1 conference per month.

Quality domains considered for reviewing Congress highlights include (1) congress selection relevance, (2) topic selection, and (3) the scientific quality of generated abstracts.

Congress Selection Relevance

This content covers the main international and national conferences in the neurology field. The congress should address topics related to one or multiple subspecialties in neurology (scored out of 5, where 0 is not relevant and 5 is highly relevant).

Congress Coverage

The objective of congress coverage is to provide medical news that faithfully reflects the original congress' program, spirit, and potential scientific and medical breakthroughs. Thus, congress highlights should be characterized by 3 criteria (Textbox 1). Each criterion received a score out of 5; scores were summed for an overall congress coverage score out of 15.

Textbox 1. Coverage criteria.

Criteria

- Coverage of both scientific and clinical-oriented topics. According to the clinician audience targeted for Neurodiem, selected topics should have a direct impact on clinical practice or translate at some point into clinical development or evolution in clinical practice.
- Coverage of hot topics, scientific or clinical highlights and late-breaking news sessions, representative of major and most topics expected to be presented during the congress.
- Coverage of both posters and oral communications, prioritizing oral communication with more validated and impactful outcomes (no more than 10 or 15 per 100 posters).

Scientific and Didactic Value (Individual Abstracts)

Structure

The *structure* domain for conference abstracts is similar to that for summaries. In addition, it is recommended that a short comment from an expert neurologist that identifies implications for clinical practice and clinical research milestones achieved or to be further defined be included.

Accuracy and Didactic Dimension

When applicable, scientific quality assessment should be based upon whether the main scientific content of the original congress communications was respected.

The didactic dimension of the abstract should be assessed on the ability to highlight new concepts and translate the findings into clinical practice.

Each item is scored out of 5, from 1 (strongly disagree) to 5 (strongly agree); scores were summed for an overall domain score out of 10.

Overall Quality

Each domain subscore was weighted by a coefficient: 1, for congress selection relevance; 2, for congress coverage; 1, for structure; 4, for accuracy and didactic dimension; the products were summed to generate a total score out of 100 (Multimedia Appendix 1).

Expert Points of View and Talks

Overview

Expert points of view and talks were developed by the Neurodiem editorial team exclusively for this platform in order to offer a synthesis on a current neurology topic by a recognized medical expert in the field. Expert points of view and expert talks are intended to offer a synthesis on a neurology current topic by a recognized medical expert in the field to allow neurologists to get expert opinions or overviews on emerging, state-of-the-art, or hot topics in neurology. Expert points of view and talks were assessed by an independent reviewer.

Subscores for 3 quality domains were weighted: a coefficient of 3 for credibility and relevance, a coefficient of 4 for scientific and didactic value, and a coefficient of 1 for reliability and format; the products were summed to generate a total score out of 100.

Credibility and Relevance

Authors or speakers and topic (Table 4) scores were summed to generate a domain subscore out of 10, which was weighted by a coefficient of 3.

Table 4. Credibility and relevance criteria.

Item	Description
Authors or speakers	Authors or speakers who have been selected to share their expert point of view should be key medical experts in neurology subspecialties (neurologist or neuroscientist). Presenters should meet quality standards in terms of academic seniority (Assistant Professor degree or equivalent), reputation among their peers and long experience (score out of 5; on a scale from 1, strongly disagree, to 5, strongly agree).
Topic	The topic should be related to recent advances or debated issues in the neurology or neuroscience. The subject should be of interest to the neurologist community; hence, content should have a valuable and original contribution to the field and a significant clinical impact (scored out of 5; 1, not related to neurology; 5, relevant and is a major contribution to the field).

Scientific and Didactic Value

Structure, accuracy and didactic dimension, and writing or speech quality (Table 5) were each assessed out of 5, on a scale from 1 (strongly disagree) to 5 (strongly agree), and summed for a domain subscore out of 15, which was then weighted by a coefficient of 4.

Table 5. Scientific and didactic value.

Item	Description
Structure	Structure for expert points of view and talks presentations were evaluated with a score from 1 to 5 based on the inclusion, accuracy, and the chronology order of • An introduction that includes (1) scientific background information, (2) a rationale for topic selection based the current state of scientific and clinical knowledge, and (3) a presentation overview • Scientific and clinical evidence supporting the topic including numerical key data • An overview of why these results have a scientific and medical impact in the neuroscience/neurology landscape • A summary of take-home messages and conclusions relating to anticipated milestones in neuroscience research, direct implications for clinical practice and/or updates to this content quality guidelines Importantly, the expert point-of-view structure will be supported by occurrence of relevant and meaningful subheadings for each paper's section or by titles corresponding to the different sections or main messages of the presentation. Overall, attention should be paid to logical development of scientific arguments and evidence.
Accuracy and didactic dimension	• Accurate and concise background information and research or clinical context • Relevant selection of specific arguments, scientific evidence, and illustrations for supporting expert demonstration
Writing or speech quality	• Authors or speakers should display the ability to synthesize ideas and provide simplified explanations of cutting-edge techniques or complex concepts • Writing or speech style: a neutral, factual, and formal tone should be used • Clarity and coherence: logical links between arguments and sections supporting scientific discussion • The quality of English or local language and grammar should be appropriate

Reliability and Format

Each of these 2 items (Table 6) were scored out of 5 based on duration of the presentation (1, for too long or too short in duration; 5, strongly agree for those approximately 5 minutes) and summed to generate a domain subscore out of 10 (Multimedia Appendix 1).

Table 6. Reliability and format.

Item	Description
References	References cited in expert point of view or talks should be focused, and source of references should be reliable and consequently selected exclusively from (1) high-impact factor or recognized peer-reviewed journals in the neurology or neuroscience field (2) validated and up-to-date clinical guidelines.
Format	Owing to the summarized format targeted for expert point of view, core content of the paper should not exceed 1500 words. Likewise, informal talks should not exceed 5 minutes (to respect technical feasibility because big files cannot be uploaded to the platform), audience expectations (scientific community has short time to watch presentations and videos)) while more academic presentations should be between 5 and 10 minutes.

Results

Summaries

The overall quality score increased from 75 in month 1 to 85 in month 2. The scores stabilized at a high level (score range between 70 and 80/100) (Figure 1). The domain *scientific and didactic value* had a mean score of 8.20/10.Those for *reliability and relevance* (8-9/10) and *structure* (45-55/60) showed that maintenance of these quality items over time was more challenging.

The ratings for the domain *scientific and didactic value* for the 12-month analysis period showed a trend similar to that of the structure domain an increase until month 4, followed by steady state, with a brief drop at month 8 (Figure 2).

Figure 1. Analysis and monitoring of scientific quality applied to Summaries published on Neurodiem over a 12-month period.

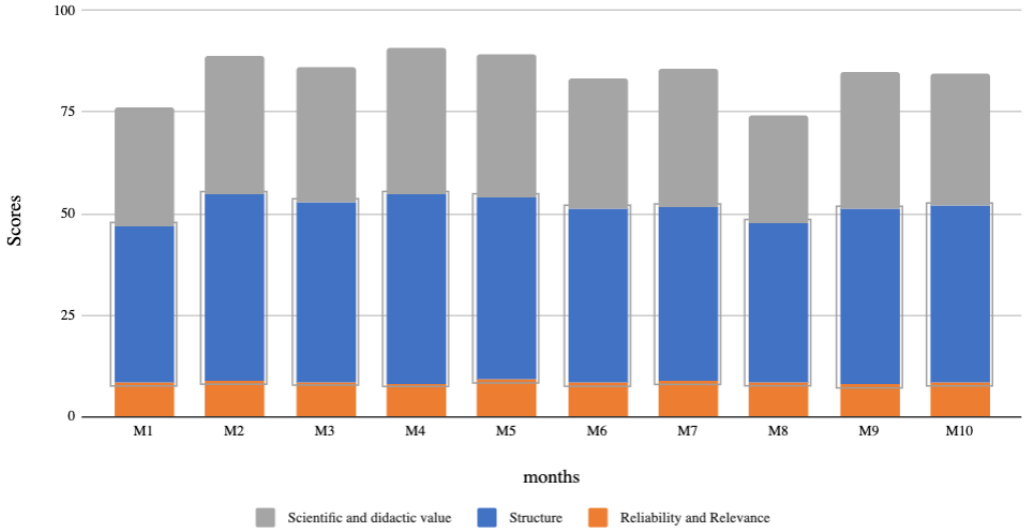
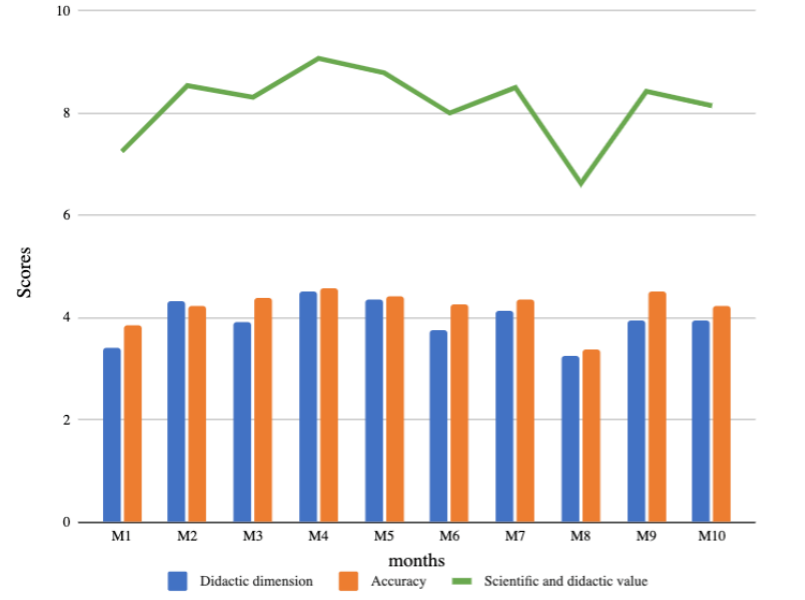


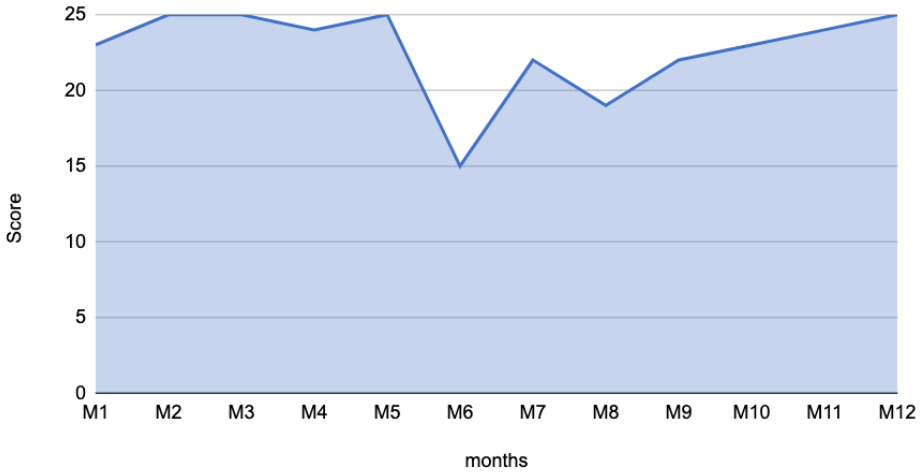
Figure 2. Analysis and monitoring of scientific and didactical quality applied to article Summaries published on Neurodiem over a 12-month period.



Expert Points of View and Talks

Talks were always provided by medical experts without revisions or changes and were maintained in their original format for publication in Neurodiem Scores were high, which was sustained over time (Figure 3). Similar to those for expert talks, scores for expert point of view were high, which was sustained throughout the 12-month period.

Figure 3. Analysis and monitoring of scientific and didactical quality applied to Expert points of view published on Neurodiem over a 12-month period.



Discussion

These evaluation scores and criteria showed that the chosen methodology was appropriate to evaluate content quality shared on internet scientific platform. The development and use of content quality guidelines, with editorial specifications and a measurement tool, ensure that the scientific and medical content on Neurodiem is high quality. The quality criteria and scoring system facilitate the delivery of relevant, reliable, and current information in neurology while respecting the editorial independence of third-parties and expert panel who are involved in the selection and production of content published on the website. This approach has multiple uses. This framework is intended for content providers and medical writers in order to define content quality guidelines and standards for reaching scientific and medical excellence of the content. The content quality guidelines provide a simple criteria checklist and evaluation scale to deliver an objective analysis and quantitative assessment of content, and monthly monitoring of papers and talks provides regular quality reports and objectives to content providers, which leads to the improvement of quality over time.

For the Summaries, a parallel evolution between the structure and Scientific and didactic value could be explained by overlapping assessment objectives of some quality items in both quality domains. Interestingly, the rating decrease at month 8 was observed to be concomitant to some turnover in the medical writers' team of content providers. Hypothetically, the drop could be due to delay in application of the content quality guidelines by the medical writers who recently joined the medical writers' team.

Several eHealth information quality evaluation tools exist to answer the needs of different profiles of internet resource users, including patients and health care professionals [18,19]. These tools have features that overlap with those in our approach, such as assessing the credibility of digital content through experts scientific trust value (authors, speakers), the content, and content reliability via source checking of validity and up-to-date references. The importance for the target audience and their scientific needs are also highly represented in the checklist of both quality measurement systems. The accuracy of scientific content is addressed through the analysis of the strength and value of the scientific evidence provided. Moreover, content readability, such as appropriate language use, clarity of expression, and the logical flow of arguments are common quality criteria in digital medical information evaluations. Although not formally assessed in our approach, disclosures and conflicts of interest are systematically displayed in the author section on Neurodiem to ensure the website's editorial independence. The didactic dimension is an original and key feature of the Neurodiem quality approach. Based upon feedback from the scientific steering committee, the weighting of didactic value was increased, in order to highlight the importance of strong clinical relevance for Neurodiem content.

Our content quality guidelines are used to evaluate of scientific and medical content dedicated to clinicians, and although it was developed for use on a neurology-specific platform, the tool could be easily translated to any medical specialty. In addition, the Neurodiem content criteria grids are adapted to the format of web-based content (the most important sections allowed higher scoring). This methodology could be used by content creators or providers to support the production and review of content and information published on web-based scientific platforms. Alternatively, these content quality guidelines could also be of value to the medical community as a rapid and effective method of appraising the quality of content when consulting medical education websites.

There are some limitations to our quality assessment system. First, although our quality measurement tool provides a relatively strict framework for an objective rating, its application is nevertheless likely subject to inter-reviewer variability (the score attributed to each section may vary from reviewer vision to another). The difference in scores attributed by each reviewer during their assessment may be particularly pronounced for the evaluation of items such as didactic dimension or quality of speech and writing, which can be quite subjective and linked directly to the reviewer own interpretation. In order to reduce heterogeneity in content quality assessment, we propose that there be a standardized training session for reviewers, aimed to educate them on the adequate and consistent use of the scoring system. Second, some global and general conferences in

neurology (eg, European Academy of Neurology, American Academy of Neurology) are characterized by a substantial communication program with a broad panel of scientific and clinical topics in distinct neurology subspecialties; as a result, we must recognize that assessments of whether congress coverage on the website is faithful to the original congress program are challenging. Thus, while our system is well suited to granular review of scientific and medical content at the single paper or topic level, improvements are needed for the assessment of content with a breadth of topics, such as a congress program. In future iterations of the content quality guidelines, automation and artificial intelligence technology could address this issue [20-22].

Our content quality guidelines are an editorial and quality evaluation system for information on Neurodiem that was developed to preserve editorial independence. Our methodology consists of a simple and short set of criteria to be used by content providers or reviewers to objectively assess the scientific and medical excellence of content, with special emphasis on impact and applicability in clinical practice. This standardized approach could be used on any biomedical news and resource digital platforms beyond the initial scope of neurology. These content quality guidelines support the implementation of a content quality strategy in the content creation phase as well as in the review process, which is the cornerstone of a high-quality digital communication platform [23].

Acknowledgments

We would like to thank all experts, stakeholders, and contributors for careful and critical review of the original content quality guidelines and their significant contributions during this medical content quality project.

Conflicts of Interest

DK consults for Biogen, Bristol Myers Squibb, and Janssen and receives research support from Bridtol Myers Squibb. AL consults for Biogen Digital Health.

Multimedia Appendix 1
Content quality measurement guideline.
[PDF File (Adobe PDF File), 678 KB - ijmr_v11i1e35698_app1.pdf]

References

1. Al-Shahi R, Sandercock PA. Internet resources for neurologists. J Neurol Neurosurg Psychiatry 2003 Jun 30;74(6):699-703 [FREE Full text] [doi: 10.1136/jnnp.74.6.699] [Medline: 12754333]
2. Moccia M, Brigo F, Tedeschi G, Bonavita S, Lavorgna L. Neurology and the internet: a review. Neurol Sci 2018 Jun 30;39(6):981-987. [doi: 10.1007/s10072-018-3339-9] [Medline: 29594831]
3. Thomson S, Phillips N. Internet resources for neurosurgeons and neuropathologists. J Neurol Neurosurg Psychiatry 2003 Feb 28;74(2):154-157 [FREE Full text] [doi: 10.1136/jnnp.74.2.154] [Medline: 12531938]
4. Busis NA. Neurology in the electronic information age. Eur J Neurol 1999 Jul 31;6(4):385-414. [doi: 10.1046/j.1468-1331.1999.640385.x] [Medline: 10362892]
5. McLean R, Richards BH, Wardman JI. The effect of web 2.0 on the future of medical practice and education: Darwikinian evolution or folksonomic revolution? Med J Aust 2007 Aug 06;187(3):174-177. [doi: 10.5694/j.1326-5377.2007.tb01180.x] [Medline: 17680746]
6. Devine T, Broderick J, Harris LM, Wu H, Hilfiker SW. Making quality health websites a national public health priority: toward quality standards. J Med Internet Res 2016 Aug 02;18(8):e211 [FREE Full text] [doi: 10.2196/jmir.5999] [Medline: 27485512]
7. Cline R, Haynes K. Consumer health information seeking on the internet: the state of the art. Health Educ Res 2001 Dec 31;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
8. Fahy E, Hardikar R, Fox A, Mackay S. Quality of patient health information on the internet: reviewing a complex and evolving landscape. Australas Med J 2014;7(1):24-28 [FREE Full text] [doi: 10.4066/AMJ.2014.1900] [Medline: 24567763]
9. Gagliardi A, Jadad A. Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination. BMJ 2002 Mar 09;324(7337):569-573 [FREE Full text] [doi: 10.1136/bmj.324.7337.569] [Medline: 11884320]
10. Kim P, Eng TR, Deering MJ, Maxfield A. Published criteria for evaluating health related web sites: review. BMJ 1999 Mar 06;318(7184):647-649 [FREE Full text] [doi: 10.1136/bmj.318.7184.647] [Medline: 10066209]
11. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb 28;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
12. Eysenbach G, Thomson M. The FA4CT algorithm: a new model and tool for consumers to assess and filter health information on the internet. Stud Health Technol Inform 2007;129(Pt 1):142-146. [Medline: 17911695]

13. Närhi U, Pohjanoksa-Mäntylä M, Karjalainen A, Saari JK, Wahlroos H, Airaksinen MS, et al. The DARTS tool for assessing online medicines information. Pharm World Sci 2008 Dec;30(6):898-906. [doi: 10.1007/s11096-008-9249-9] [Medline: 18791806]
14. Robillard JM, Jun JH, Lai J, Feng TL. The QUEST for quality online health information: validation of a short quantitative tool. BMC Med Inform Decis Mak 2018 Oct 19;18(1):87 [FREE Full text] [doi: 10.1186/s12911-018-0668-9] [Medline: 30340488]
15. Sandvik H. Health information and interaction on the internet: a survey of female urinary incontinence. BMJ 1999 Jul 03;319(7201):29-32 [FREE Full text] [doi: 10.1136/bmj.319.7201.29] [Medline: 10390457]
16. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the internet: caveant lector et viewor—let the reader and viewer beware. JAMA 1997 Apr 16;277(15):1244-1245. [Medline: 9103351]
17. Keselman A, Arnott Smith C, Murcko AC, Kaufman DR. Evaluating the quality of health information in a changing digital ecosystem. J Med Internet Res 2019 Feb 08;21(2):e11129 [FREE Full text] [doi: 10.2196/11129] [Medline: 30735144]
18. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. J Med Internet Res 2019 May 02;21(5):e12522 [FREE Full text] [doi: 10.2196/12522] [Medline: 31045507]
19. Zhang Y, Sun Y, Xie B. Quality of health information for consumers on the web: a systematic review of indicators, criteria, tools, and evaluation results. J Assn Inf Sci Tec 2015 Apr 29;66(10):2071-2084. [doi: 10.1002/asi.23311]
20. Boyer C, Dolamic L. Automated detection of HONcode website conformity compared to manual detection: an evaluation. J Med Internet Res 2015 Jun 02;17(6):e135 [FREE Full text] [doi: 10.2196/jmir.3831] [Medline: 26036669]
21. Smeros P, Castillo C, Aberer K. SciLens: evaluating the quality of scientific news articles using social media and scientific literature indicators. 2019 Presented at: World Wide Web Conference; May 13-17, 2019; San Francisco, California. [doi: 10.1145/3308558.3313657]
22. Afsana F, Kabir MA, Hassan N, Paul M. Automatically assessing quality of online health articles. IEEE J Biomed Health Inform 2021 Feb 28;25(2):591-601. [doi: 10.1109/JBHI.2020.3032479] [Medline: 33079686]
23. Jackson JL, Srinivasan M, Rea J, Fletcher KE, Kravitz RL. The validity of peer review in a general medicine journal. PLoS One 2011;6(7):e22475 [FREE Full text] [doi: 10.1371/journal.pone.0022475] [Medline: 21799867]

Edited by T Leung; submitted 14.12.21; peer-reviewed by DLG Rodrigues; comments to author 08.01.22; revised version received 14.01.22; accepted 26.04.22; published 09.06.22.

Please cite as:
Kantor D, Farlow M, Ludolph A, Montaner J, Sankar R, Sawyer Jr RN, Stocchi F, Lara A, Clark S, Deschet K, Ouyahia L, Hadjiat Y
Digital Scientific Platform for Independent Content in Neurology: Rigorous Quality Guideline Development and Implementation
Interact J Med Res 2022;11(1):e35698
URL: https://www.i-jmr.org/2022/1/e35698
doi:10.2196/35698
PMID:35485280

Original Paper

Quantifying the Impact of COVID-19 on Telemedicine Utilization: Retrospective Observational Study

Emily Louise Vogt[1,2], BA; Brandon M Welch[3,4], PhD; Brian E Bunnell[4,5], PhD; Janelle F Barrera[4,5], MPH, CPH, CHES; Samantha R Paige[4], MPH, PhD, CHES; Marisa Owens[6], RHIA; Patricia Coffey[6], MSHIA, RHIA; Nancy Diazgranados[2], MD, MSc; David Goldman[2,7], MD

[1]University of Michigan Medical School, Ann Arbor, MI, United States

[2]Office of the Clinical Director, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD, United States

[3]Biomedical Informatics Center, Medical University of South Carolina, Charleston, SC, United States

[4]Doxy.me Research, Doxy.me, LLC, Rochester, NY, United States

[5]Department of Psychiatry and Behavioral Neurosciences, University of South Florida, Tampa, FL, United States

[6]Clinical Center, National Institutes of Health, Bethesda, MD, United States

[7]Lab of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Rockville, MD, United States

Corresponding Author:
David Goldman, MD
Lab of Neurogenetics
National Institute on Alcohol Abuse and Alcoholism
National Institutes of Health
5625 Fishers Lane
Rm 3S-32: MSC 9412
Rockville, MD, 20892
United States
Phone: 1 301 443 0059
Email: davidgoldman@mail.nih.gov

Abstract

Background: While telemedicine has been expanding over the past decade, the COVID-19–related restrictions regarding in-person care have led to unprecedented levels of telemedicine utilization. To the authors' knowledge, no studies to date have quantitatively analyzed both national and regional trends in telemedicine utilization during the pandemic, both of which have key implications for informing health policy.

Objective: This study aimed to investigate how trends in telemedicine utilization changed across the course of the COVID-19 pandemic.

Methods: Using data from doxy.me, the largest free telemedicine platform, and the NIH (National Institutes of Health) Clinical Center, the largest clinical research hospital in the United States, we assessed changes in total telemedicine minutes, new provider registrations, monthly sessions, and average session length from March to November 2020. We also conducted a state-level analysis of how telemedicine expansion differed by region.

Results: National telemedicine utilization peaked in April 2020 at 291 million minutes and stabilized at 200 to 220 million monthly minutes from May to November 2020. Surges were strongest in New England and weakest in the South and West. Greater telemedicine expansion during the COVID-19 pandemic was geographically associated with fewer COVID-19 cases per capita. The nature of telemedicine visits also changed, as the average monthly visits per provider doubled and the average visit length decreased by 60%.

Conclusions: The COVID-19 pandemic led to an abrupt and subsequently sustained uptick in telemedicine utilization. Regional and institute-level differences in telemedicine utilization should be further investigated to inform policy and procedures for sustaining meaningful telemedicine use in clinical practice.

(Interact J Med Res 2022;11(1):e29880) doi:10.2196/29880

KEYWORDS

telemedicine; COVID-19; utilization; impact; retrospective; observational; trend; telehealth; health policy; policy

Introduction

In the past three decades, telemedicine—defined by the National Institutes of Health (NIH) as the use of telecommunication and technological services to provide and support medical care at a distance—has been widely adopted by health care providers and systems across the world [1,2]. Telemedicine can include synchronous modalities (ie, real-time audio or audiovisual interaction), asynchronous modalities (ie, messages or images exchanged via a patient portal), or remote patient monitoring. However, synchronous audiovisual communication attempts to closely replicate ordinary patient-provider or provider-provider interaction to the maximum extent possible given the limitations in audiovisual communication caused by vision or hearing impairments, both of which are prevalent among patients and providers [3,4].

The main advantages of telemedicine include improved accessibility of care, particularly for rural and underserved communities; flexibility of scheduling; greater continuity of care; reduced cost of care in certain situations; and enhanced collaboration between medical providers [5-7]. The information technology revolution, including the rapid expansion of electronic health record (EHR) usage and sharing, has accelerated the expansion of telemedicine, with more than 60% of US health care systems and 40% to 50% of US hospitals employing some form of telehealth [5]. Due to the COVID-19 pandemic in 2020, telemedicine abruptly became the safer form of care in many cases and the only practically allowable form in others.

Care delivery using telemedicine, however, is not without behind-the-scenes complexities, including coordination across EHRs, patient portals, e-prescribing platforms, other scheduling- or monitoring-related applications; access to secure, effective audiovisual communication software; and acquisition of sufficient computer hardware and high-speed internet. Due to these nuances of telemedicine, a large and growing body of research has emerged to study telemedicine's efficacy across a diverse set of care delivery settings, patient populations, medical specialties, and geographic regions.

Telemedicine research has been an area of broad interest and development since the 1990s; in both 2018 and 2019, there were nearly 3000 publications each year related to telemedicine or telehealth [8]. These studies, largely in behavioral health-related settings, have demonstrated that the consumer experience is equivalent or superior to in-person encounters across a range of diagnoses, patient populations, and health care settings [7,9-11]. Moreover, under certain circumstances telemedicine is more cost-effective than in-person care and can allow health care entities to provide health care where it may otherwise be unavailable, either due to physical distance between the provider and patient or distance between a group of patients who can interact remotely but not in the same physical setting [7,12-14]. However, there has historically been a gap between the increasing research interest in telemedicine and its broad-scale implementation and acceptance by health systems due to reimbursement limitations, technical barriers, physician attitudes, and lack of physician education [15,16]. Physician skills may themselves need to be adapted and optimized for virtual care settings [15,17,18].

During the COVID-19 pandemic, telemedicine has become vital, or mandatory, for many purposes for which it was formerly convenient or optional. Due to the serious and ongoing safety risks associated with the spread of COVID-19 in in-person health care settings, widespread telemedicine adoption has become a necessary substitute for everything from routine health maintenance visits or exams to COVID-19–related issues. In March 2020 alone, the Cleveland Clinic reported 60,000 telehealth visits, a 1700% increase from the previous monthly average [19]. Similarly, a retrospective analysis of January to March 2020 from the Centers for Disease Control and Prevention (CDC) showed a 154% increase in telehealth visits in the last week of March, with COVID-19–related visits comprising 15.2% of these visits [20]. More recent data from the Veterans Affairs health administration indicated that trends have persisted well into the later stages of the pandemic, with a tally of more than 300 million virtual visits in the month of June 2020 alone [21]. Moreover, a recent survey of multispecialty physicians revealed that although only 12% of those surveyed had used telemedicine prior to the pandemic, 91% planned to continue offering telemedicine services following COVID-19 [22]. Thus, understanding the nuances of these trends in telemedicine usage has significant implications for not only the COVID-19 era, but for the future of health care [1,23-25].

While many patients and physicians initially resorted to familiar platforms such as FaceTime for virtual medical appointments, it was quickly recognized that these platforms did not assure security, privacy, or quality, which pushed many health care providers and systems to consider using existing telemedicine platforms. Doxy.me is a national telemedicine platform launched in 2013 primarily to serve mental health providers that has since been scaled to cover all types of medical specialties. While there are other commercial telemedicine platforms in the rapidly expanding telehealth industry, doxy.me is the largest platform that is both free and HIPAA (Health Insurance Portability and Accountability Act)-compliant, making it a popular choice for providers looking for a swift transition to virtual care amidst the COVID-19 pandemic [26-28]. Even as other telemedicine platforms emerged, trends in the usage of doxy.me can therefore reveal national changes in the use of telemedicine occurring in response to the pandemic due to its widespread, national utilization.

In contrast, some institutions like the NIH Clinical Center, known as "America's research hospital" [29], elected to develop their own internal telemedicine programs. While these programs took longer to jumpstart given that they had not been established prior to the pandemic, they allowed institutions like the NIH Clinical Center to ensure full integration with their existing

hospital and research infrastructure, as well as the sustainability of the program beyond the pandemic. The aim of this study was to comparatively analyze trends in telemedicine utilization during COVID-19 from both doxy.me's national platform and the NIH Clinical Center's program to better quantify how and where the COVID-19 pandemic has most influenced telemedicine usage. Synchronous audiovisual visits, and the ancillary factors that enable them, were the focus of this study.

Methods

Study Design

We conducted a retrospective observational study of trends in telemedicine utilization from two different stakeholders: doxy.me, a national platform that has been supporting telemedicine encounters prior to COVID-19; and the NIH Clinical Center, the nation's largest clinical research hospital, which developed a telemedicine program in response to COVID-19. Doxy.me's workflow is designed to be as simple as possible while still being familiar to both patients and doctors, with a check-in feature, a waiting room, and a patient queue. The virtual platform attempts to greet patients with a routine that feels familiar to the clinical experience. No patient information is stored by doxy.me, all calls are encrypted end to end by default, and the platform is built on top of an open-source standard for real-time communication over the internet. The platform is built to only operate within trusted web browsers provided by Mozilla, Apple, Google, and Microsoft, who update their products on a rolling cycle every 6 to 9 months to ensure they are patched and up to date. Our study analyzed doxy.me usage from January to November 2020, the period in which the pandemic and, consequently, the need for virtual encounters accelerated most rapidly.

The NIH Clinical Center's telemedicine program is markedly different from doxy.me, being institution-specific in scope and having been developed directly in response to COVID-19 for secure videoconferencing with past and present patients at the NIH Clinical Center. It is based on Microsoft Teams (Microsoft Corp), a platform compliant with the NIH's privacy and security policies. Information about the NIH Clinical Center's program was provided to individual NIH institutes via Medical Executive Committee meetings and a web-based telehealth resources section. This program was launched in April 2020, and this analysis encompasses all telehealth visits between April and November 2020 at the NIH Clinical Center.

Both doxy.me and the NIH Clinical Center track metrics for quality assessment and quality improvement purposes. Doxy.me tracks registered providers, sessions, and minutes as a measure of growth, and conducts annual risk assessments and updates per HIPAA and Health Information and Technology for Economic and Clinical Health (HITECH) policies accordingly. Updates include vulnerability and patching updates, backup and business continuity plans, encryption of data stored at rest or transfer per the recommendations of the National Institute of Standards and Technology, and regular access and error log audits via an intrusion detection system. Notably, any provider using the platform can generate a Business Associate Agreement signed by doxy.me, to assist in maintaining HIPAA/HITECH compliance across the board. Doxy.me also allows for the storage of emergency forms with the primary provider's contact information for each patient. In contrast, the NIH Clinical Center's data tracking includes linkage of all data contained in the EHR. To initiate an NIH Clinical Center telehealth encounter, care teams enter an electronic appointment request in the EHR. The Health Information Management Department conducts a documentation review for each patient encounter. Completed, documented appointments are marked as "arrived" and subsequently counted as telehealth visits.

Statistical Analysis

Trend analysis was performed on both data sets. For doxy.me data, regression analysis was performed to illuminate state-level trends in telemedicine versus COVID-19 case rates. State-by-state COVID-19 case rates were derived from the CDC's database [30]. Additionally, US census data were used to normalize levels of telemedicine utilization and COVID-19 cases across different states. Data analysis and graphical depictions were performed using Microsoft Office (Microsoft Corp).

Results

Nationwide Telemedicine Trends

Doxy.me's telemedicine volume surged from just over 5.5 million monthly minutes in February 2020 to more than 89 million monthly minutes in March 2020 (Figure 1), a 29-fold increase compared to usage data from March 2019. Increases in the number of telemedicine sessions per provider were also observed, as was a highly significant decrease in the average length of telemedicine sessions during March and April 2020 compared to prepandemic times (Figure 2). This finding indicates a change in the average telemedicine encounter, the nature of which is not fully understood.

Doxy.me's provider registrations—tracking the number of new providers signing up on their platform—revealed that the intensified use of telemedicine was mainly driven by new users. March yielded the greatest number of new provider registrations at 299,324, a more than 3-fold jump in doxy.me's total provider base in just 1 month. Predictably, this uptick in provider registrations occurred in tandem with the initial declaration of COVID-19 as a global health emergency by the World Health Organization and with stay-at-home directives in the United States. However, since May 2020, new provider registrations have stabilized at a much lower level, fluctuating between 12,000 and 19,000 new registrations per month.

Similarly, total minutes of telemedicine utilization accelerated rapidly in March and April 2020 but then stabilized at a new, markedly higher (48×) level as compared to prepandemic times, approaching 300 million minutes/month in April and settling at just over 200 million minutes/month (Figure 1).

Figure 1. The total number of session minutes per month logged on the doxy.me platform from January to November 2020. These data demonstrate that while monthly minutes peaked in April at 291 million minutes, the high level of utilization largely plateaued from May to November.

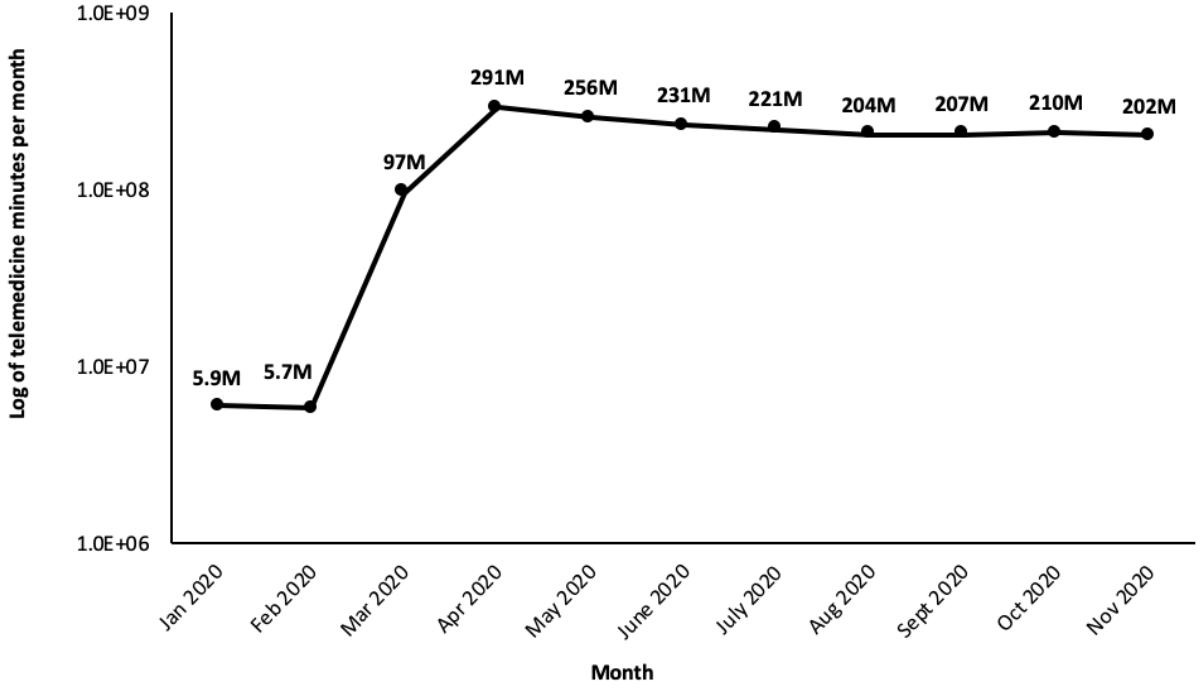
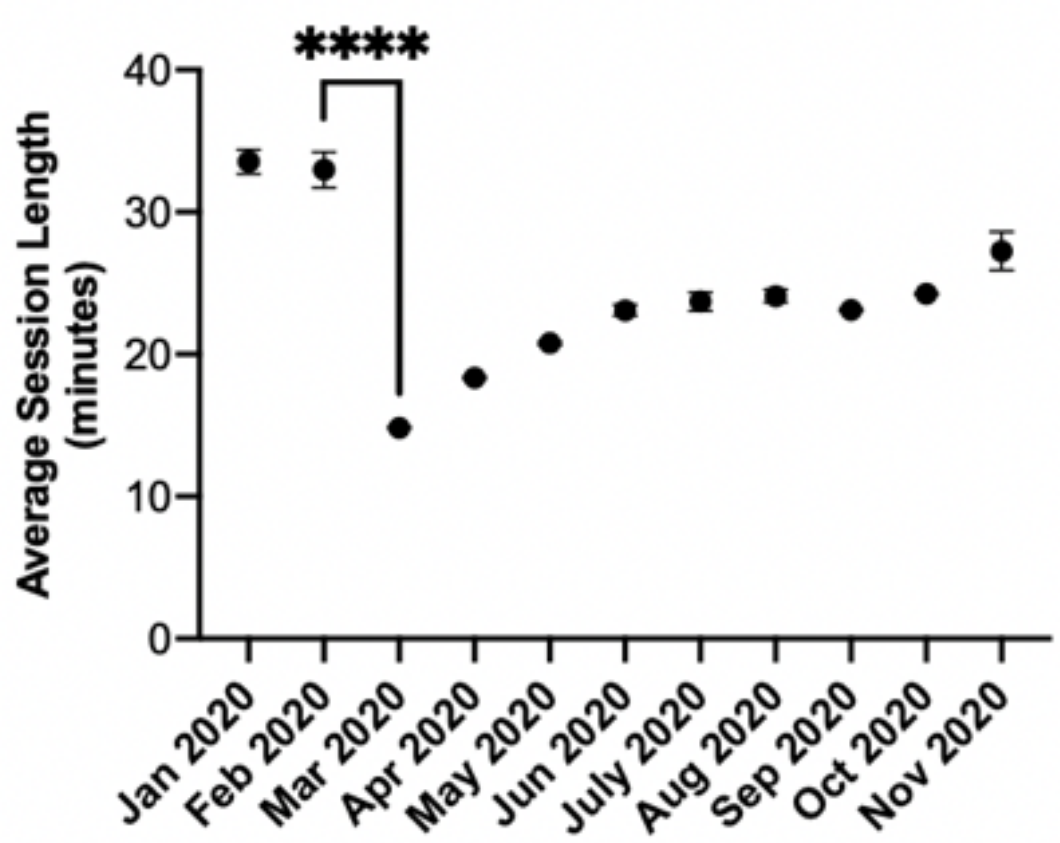


Figure 2. The average session length per doxy.me provider per month from January to November 2020. These results elucidate the dramatic decrease in telemedicine session length (P<.001) at the onset of the pandemic, from February to March. From April to November, the average telemedicine session length normalized back to pre–COVID-19 levels with the average session length during November not significantly lower than in January (P=.10). Asterisk indicates a statistically significant difference.



Regional Telemedicine Trends

Telemedicine utilization across states and over the course of the pandemic revealed that the surge in the utilization of telemedicine was highly correlated with the earlier expansion of telemedicine service capacity in certain states (m=4.18; R^2=0.96), presumably driven by discrepancies in the supply of existing telemedicine providers in certain states versus others (Figure 3). In other words, states with more providers who were already licensed and practicing telemedicine prior to the pandemic had a greater capacity for offering telemedicine services at the onset of the pandemic compared to states with few to no trained telemedicine providers. Surprisingly, states with the greatest expansion of telemedicine provider registrations also tended to have lower aggregate rates of COVID-19 cases per capita (m=–0.0031; R^2=0.29) (Figure 4). Regional trends in the expansion of telemedicine were evident (Figures 3 and 4). New England (including Massachusetts, Connecticut, New Hampshire, Maine, Rhode Island, and Vermont; see note in Figure 3 for Vermont) had the strongest telemedicine expansion early in the pandemic, both in more rural and metropolitan areas. These New England states also had the lowest rates of COVID-19 per capita by November (see Figures 3 and 4). The Mid-Atlantic states (including New York, Pennsylvania, New Jersey, Delaware, Maryland, Virginia, and the District of Columbia) experienced a moderate expansion of telemedicine between April and November, and moderate to low aggregate COVID-19 cases per capita by November, despite initial COVID-19 hotspots in New York City and New Jersey (Figures 3 and 4).

The southern and northwestern states, most notably Mississippi, Alabama, Wyoming, and the Dakotas, had the slowest expansion of telemedicine possibly because of a delay in the upswing in cases and a lag in the implementation of social distancing measures and mask mandates (Figure 3). These states also logged some of the highest rates of COVID-19 cases per capita by November (Figure 4). Thus, it seems that the expansion of telemedicine services across the country seemed to cluster with states that also had a stricter mask mandate and social distancing policies, suggesting the potential role of regional politics in shaping telemedicine adoption.

Figure 3. State-by-state comparison of April 2020 telemedicine minutes per capita versus telemedicine per capita added during the onset of the COVID-19 pandemic (measured as April through November 2020). Linear regression analysis shows a strong, positive association (m=4.18; R2=0.96) between states that had more telemedicine capacity in April and greater overall telemedicine expansion during COVID-19. Note: Vermont was excluded from this graph as it was an outlier at (4.66, 16.34).

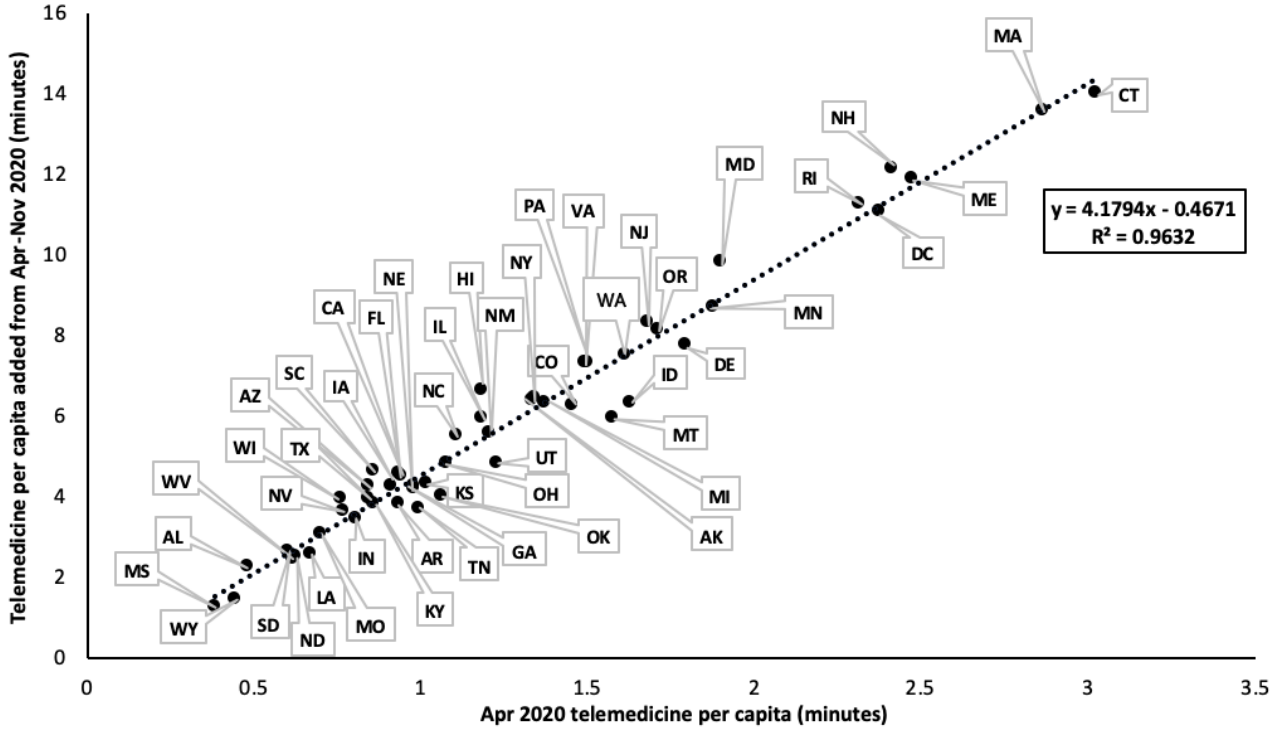
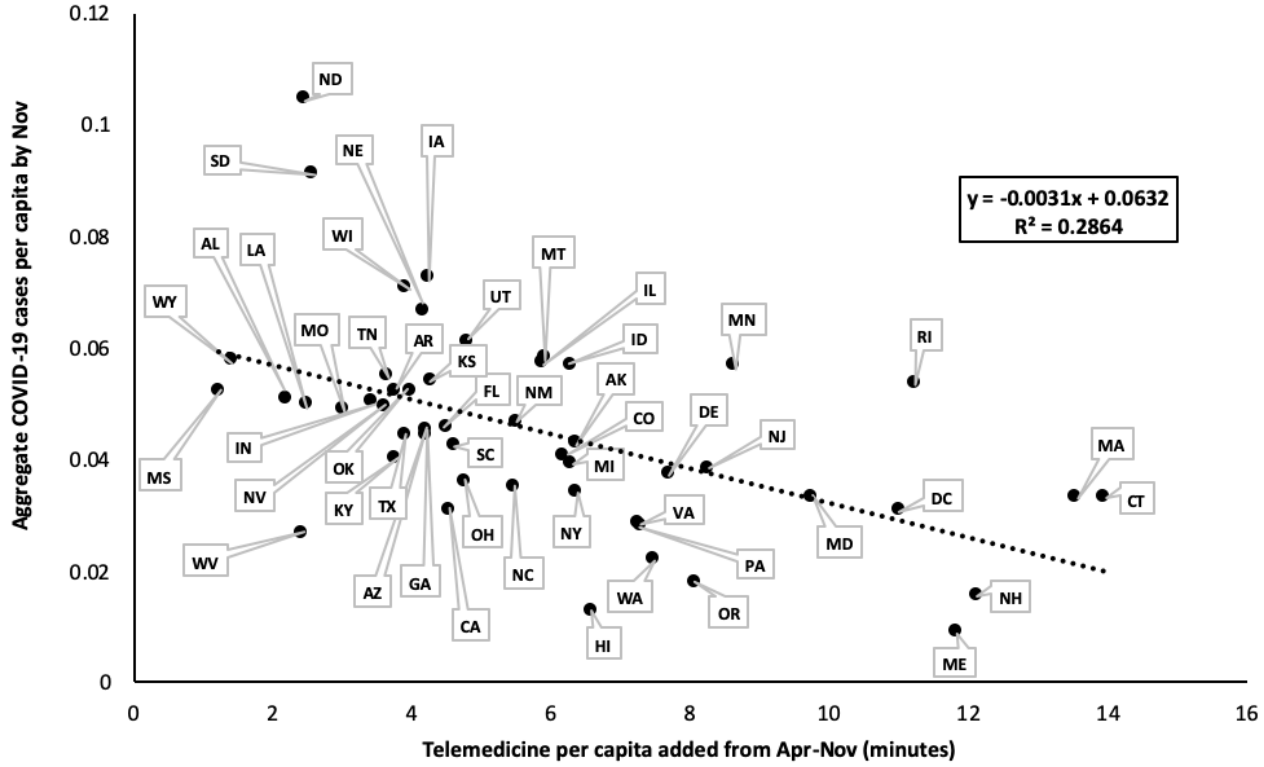


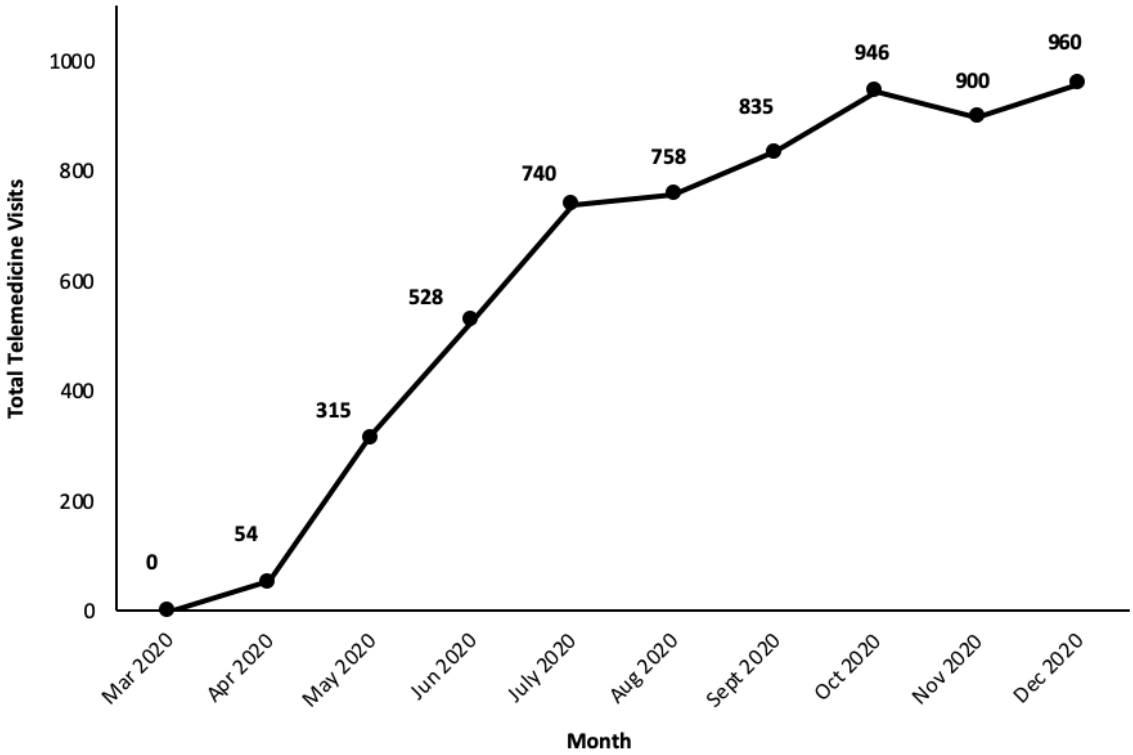
Figure 4. State-by-state comparison of telemedicine per capita added during the onset of the COVID-19 pandemic (measured as April through November 2020) versus aggregate COVID-19 cases per capita through November 2020. Linear regression analysis shows a moderately negative relationship (m=–0.0031, R2=0.29), indicating that greater telemedicine expansion was somewhat associated with fewer aggregate COVID-19 cases per capita. Note: Vermont was excluded from this graph as it was an outlier at (19.61, 0.0067).



NIH Clinical Center Program Outcomes

NIH Clinical Center data indicate that telemedicine rapidly became a useful tool for continuing clinical research efforts at a distance and has continued to accelerate in terms of utilization. Importantly, the NIH Clinical Center was closed to admissions except for COVID-19 cases and medical emergencies from April to June 2020, but the usage of telemedicine has continued to accelerate after the NIH Clinical Center was reopened to patients in most research studies. Although the NIH Clinical Center's telemedicine program has only existed since late April, it has already logged nearly 3000 visits across 15 different institutes (Figure 5). Most of these visits were between a remote patient and provider; however, some visits have been between patients at the NIH Clinical Center and specialists or other providers who are teleworking due to COVID-19 staffing restrictions. In addition, patient group meetings and staff meetings, including clinical rounds on inpatients, are conducted remotely or with a large remote component even though they are beyond the scope of the data presented here.

Figure 5. Increase in telemedicine visits at the NIH Clinical Center between its initiation on April 15 and December 2020. Trends show that there was a steadily increasing number of monthly visits since its initiation, with November as the only month that did not have more visits than the previous one. In just 7 months, the program tallied more than 6000 total visits.



Discussion

Principal Results

In summary, this study demonstrated that telemedicine expanded by more than 50-fold its previous level at the onset of the COVID-19 pandemic and has continued to maintain this high level of utilization across subsequent months, particularly within states in New England and the Mid-Atlantic. The rapid expansion of telemedicine during the COVID-19 has been widely appreciated and qualitatively described [20,31-34]. Here, we have added a quantitative analysis of trends in telemedicine demonstrating a sudden expansion (ie, over only 2 months) at a very large magnitude expansion, probably representing more than a 50-fold increase in telemedicine on a nationwide basis. In this expansion, an earlier adoption of telemedicine was associated with a greater magnitude of increase. There is strong evidence that, on a national level, the use of telemedicine has plateaued at a formerly inconceivable 200 million minutes/month on the doxy.me platform alone. Furthermore, there is evidence that the nature of the typical telemedicine visit has changed during the pandemic, now averaging only about 15 minutes instead of twice that, and with twice as many visits per provider per day. This change in pattern of use seems to reflect the fact that much of the use of telemedicine previously was for behavioral health purposes, rather than the broader clinical applications that have defined the COVID-19 era.

Doxy.me's telemedicine utilization indexes the private telemedicine market—a space that is likely to continue to grow and, through its growth, shift the landscape of health care in the years to come. Data from doxy.me's platform demonstrates that telemedicine suddenly became a critical tool for remote care during March and April 2020, when hospitals were overwhelmed with patients with COVID-19 and other facilities were largely shut down. Comparing March 2020 to March 2019, there was a 48-fold increase in the number of sessions on this platform, with it being one of the most widely used platforms before and after the beginning of the pandemic. Interestingly, the unprecedented surge in telemedicine usage has been relatively sustained since it peaked at 291 million monthly minutes in April, with the months of May to November 2020 still averaging 200 to 220 million monthly minutes each. This finding suggests that telemedicine utilization was not strictly determined by fluctuating COVID-19 case rates and stay-at-home orders; rather, telemedicine has sustained its role in serving both patients with and without COVID-19, a role that will likely continue in a postpandemic era.

Regional differences in telemedicine utilization identify geographically clustered states that either had greater (in the case of New England and the Mid-Atlantic) or lesser (in the case of southern and western states) degrees of telemedicine implementation and utilization during COVID-19 (Figures 3 and 4). One explanation for this finding is that northeastern states were forced to adopt telemedicine earlier (Figure 3). Additionally, the moderately negative association between telemedicine expansion and aggregate COVID-19 cases demonstrates that the earlier expansion of telemedicine capacity meant there were fewer in-person, health-related appointments in these states, which helped to contain the spread of COVID-19 and keep case rates lower than states without the same degree of telemedicine utilization. This association demonstrates the efficacy of telemedicine as a public health tool, especially in pandemic-related situations. However, it is also likely that the

earlier adoption of telemedicine in certain states was attributable to other factors, including differences in health care infrastructure, reimbursement policies, and licensing restrictions. More specifically, we would like to note that while many insurance companies had waived previous telemedicine-related reimbursement restrictions in light of the need for virtual care during the pandemic [35], these cost- and insurance-related barriers can significantly skew who is able to access telemedicine or other means of virtual care delivery [15]. While it is beyond the scope of this paper to explore these other factors, our results paint a picture of how telemedicine has evolved throughout the COVID-19 pandemic on a macroscopic level and provides a basis for speculating how it may continue to evolve, in tandem with necessary policy changes to make telemedicine more accessible to all patients and caregivers.

In contrast, the NIH Clinical Center's telemedicine program provides an alternative perspective on how a program that emerged in response to COVID-19 could still impact daily clinical research operations during both lockdown and return-to-work phases. The results from the NIH program indicate that there is an accelerating role for the use of telemedicine for clinical researchers to follow up with patients, schedule remote study visits, and provide virtual support to groups of patients. Given the sheer number of research protocols suspended or halted during COVID-19, this telemedicine program has evolved to support nearly 1000 patient visits each month, a number that represents most active clinical research visits. While the majority of telemedicine visits in the initial weeks of the program were within the National Cancer Institute, there has been a marked uptick in the proportion of visits for inpatient mental health units including those of the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute of Mental Health (NIMH) from June to August. This shift can largely be attributed to the fact the NIAAA and NIMH inpatient units were closed at the onset of the pandemic and did not reopen until the summer; however, it is also interesting to consider whether some of the increases in telemedicine utilization by these units may be indicative of broader increases in mental health needs following the trauma and anxiety caused by COVID-19.

Limitations

While both the doxy.me platform and the NIH Clinical Center program provide a broad snapshot of changes in telemedicine utilization and program development during the COVID-19 pandemic, they are by no means all-inclusive of the changes happening across all health systems and platforms. Additionally, while this analysis was focused on quantitative trends in national and regional differences in telemedicine utilization, we did not have access to meaningful data to analyze patient- and provider-level differences in access to telemedicine during COVID-19, which may be important for further shaping policy efforts. Future studies should aim to understand how these patient- and provider-level factors, including provider specialty, provider age, patient age, and patient socioeconomic status, may influence access to telemedicine-based care.

Conclusions

Overall, data from both these programs provide a quantitative lens for examining how trends in telemedicine have changed in response to COVID-19, with meaningful implications for local and national health care policy. Telemedicine utilization increased more than 48-fold in the first year of the pandemic, most notably in New England and the Mid-Atlantic. While telemedicine has provided significant bandwidth and has played an important role in covering remote care delivery needs, there are some apparent limitations of telemedicine as the sole option for care. Many specialties are limited in terms of the care they can provide virtually. Additionally, data from this analysis revealed that there were marked decreases in the amount of time providers spent per session in virtual appointments, whether attributable to telemedicine, the demands of the pandemic, or both. It is highly unlikely that telemedicine will displace in-person care efforts in any medical specialty. However, this study illustrates some important considerations as we evolve to a more hybrid model of virtual and in-person care, including privacy- and security-related nuances, regional differences, and clinical setting considerations. Even beyond the COVID-19 pandemic, telemedicine will continue to shape the evolving nature of health care delivery and hold critical importance for increasing access to health care.

Acknowledgments

This study was supported by grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) Division of Intramural Clinical and Biological Research Z1A AA000130, which is the broader grant covering research activities related to authors ND, DG, and ELV's work as part of the Office of the Clinical Director at the NIAAA, and a grant from the National Institute on Mental Health (MH118482), which supports BEB's ongoing research at the University of South Florida. Neither of the funding sources were explicitly for this retrospective study.

Authors' Contributions

ELV and DG developed the idea for the study and led both data analysis and manuscript development. MO, PC, BMW, and BEB provided data and reviewed the manuscript. ND, SRP, and JFB assisted in manuscript development and editing.

Conflicts of Interest

BMW is the founder of doxy.me. BEB, JFB, and SRP are all current employees of doxy.me. ELV, DG, ND, MO, and PC have no conflicts of interest to declare.

References

1. Wilson LS, Maeder AJ. Recent directions in telemedicine: review of trends in research and practice. Healthc Inform Res 2015 Oct;21(4):213-222 [FREE Full text] [doi: 10.4258/hir.2015.21.4.213] [Medline: 26618026]
2. Telehealth. National Institute of Biomedical Imaging and Bioengineering. 2020. URL: https://www.nibib.nih.gov/science-education/science-topics/telehealth [accessed 2021-04-23]
3. Quick Statistics About Hearing. National Institute on Deafness and Other Communication Disorders. 2021. URL: https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing [accessed 2021-04-23]
4. Eye Disease Statistics. NIH National Eye Institute. 2014 Mar. URL: https://www.nei.nih.gov/sites/default/files/2019-04/NEI_Eye_Disease_Statistics_Factsheet_2014_V10.pdf [accessed 2021-04-23]
5. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017 Oct 19;377(16):1585-1592. [doi: 10.1056/NEJMsr1503323] [Medline: 29045204]
6. Alvandi M. Telemedicine and its role in revolutionizing healthcare delivery. Amer J Managed Care 2017;5(1):e1-e5 [FREE Full text]
7. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. Telemed J E Health 2013 Jun;19(6):444-454 [FREE Full text] [doi: 10.1089/tmj.2013.0075] [Medline: 23697504]
8. Telemedicine or telehealth? Timeline results by year. PubMed.gov, NIH National Library of Medicine: National Center for Biotechnology Information. 2021. URL: https://tinyurl.com/v3raz48s [accessed 2021-04-23]
9. Chakrabarti S. Usefulness of telepsychiatry: A critical evaluation of videoconferencing-based approaches. World J Psychiatry 2015 Sep 22;5(3):286-304 [FREE Full text] [doi: 10.5498/wjp.v5.i3.286] [Medline: 26425443]
10. Molfenter T, Boyle M, Holloway D, Zwick J. Trends in telemedicine use in addiction treatment. Addict Sci Clin Pract 2015 May 28;10:14 [FREE Full text] [doi: 10.1186/s13722-015-0035-4] [Medline: 26016484]
11. Eibl JK, Gauthier G, Pellegrini D, Daiter J, Varenbut M, Hogenbirk JC, et al. The effectiveness of telemedicine-delivered opioid agonist therapy in a supervised clinical setting. Drug Alcohol Depend 2017 Jul 01;176:133-138 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.01.048] [Medline: 28535455]
12. Mahmoud H, Vogt EL, Dahdouh R, Raymond ML. Using continuous quality improvement to design and implement a telepsychiatry program in rural Illinois. Psychiatr Serv 2020 Aug 01;71(8):860-863. [doi: 10.1176/appi.ps.201900231] [Medline: 32290807]
13. Deslich S, Stec B, Tomblin S, Coustasse A. Telepsychiatry in the 21(st) century: transforming healthcare with technology. Perspect Health Inf Manag 2013;10:1f [FREE Full text] [Medline: 23861676]
14. Shore JH. Telepsychiatry: videoconferencing in the delivery of psychiatric care. Am J Psychiatry 2013 Mar;170(3):256-262. [doi: 10.1176/appi.ajp.2012.12081064] [Medline: 23450286]
15. Mahmoud H, Vogt EL, Sers M, Fattal O, Ballout S. Overcoming barriers to larger-scale adoption of telepsychiatry. Psychiatric Annals 2019 Feb;49(2):82-88. [doi: 10.3928/00485713-20181228-02]
16. Cowan KE, McKean AJ, Gentry MT, Hilty DM. Barriers to use of telepsychiatry: clinicians as gatekeepers. Mayo Clin Proc 2019 Dec;94(12):2510-2523. [doi: 10.1016/j.mayocp.2019.04.018] [Medline: 31806104]
17. Iancu AM, Kemp MT, Alam HB. Unmuting medical students' education: utilizing telemedicine during the COVID-19 pandemic and beyond. J Med Internet Res 2020 Jul 20;22(7):e19667 [FREE Full text] [doi: 10.2196/19667] [Medline: 32614776]
18. Jumreornvong O, Yang E, Race J, Appel J. Telemedicine and medical education in the age of COVID-19. Acad Med 2020 Dec;95(12):1838-1843 [FREE Full text] [doi: 10.1097/ACM.0000000000003711] [Medline: 32889946]
19. Dzuben M. 3 Telehealth Lessons Learned from COVID-19 Pandemic. HIT Consultant. 2020 Apr 10. URL: https://hitconsultant.net/2020/04/10/telehealth-lessons-learned-from-covid-19-pandemic/#.X-eSLs1Kg2y [accessed 2021-04-23]
20. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly B, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
21. Rosen CS, Morland LA, Glassman LH, Marx BP, Weaver K, Smith CA, et al. Virtual mental health care in the Veterans Health Administration's immediate response to coronavirus disease-19. Am Psychol 2021 Jan;76(1):26-38. [doi: 10.1037/amp0000751] [Medline: 33119331]
22. Miner H, Fatehi A, Ring D, Reichenberg JS. Clinician telemedicine perceptions during the COVID-19 pandemic. Telemed J E Health 2021 May 07;27(5):508-512. [doi: 10.1089/tmj.2020.0295] [Medline: 32946364]
23. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: Implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Mar 20;26(5):309-313. [doi: 10.1177/1357633x20916567]
24. Russel D, Hopkins J. Amid Coronavirus, calls and texts to mental-health hotlines are surging. The Center for Public Integrity. 2020 Apr 2. URL: https://tinyurl.com/4t7p653x [accessed 2021-11-11]
25. Cunningham PW. The Health 202: Texts to federal government mental health hotline up roughly 1,000 percent. The Washington Post. 2020 May 4. URL: https://tinyurl.com/yue2acmr [accessed 2021-11-11]
26. Singh J, Brown J, Heusser A. Telehealth Industry: Beyond the Pandemic. Healthcare Technology Industry Primer. Credit Suisse. 2020 Jun 01. URL: https://tinyurl.com/me48peku [accessed 2021-04-20]

27. Jercich K. Healthcare Organizations Turning to Wide Range of Vendors in Pandemic. Healthcare IT News. 2020 Jul 24. URL: https://www.healthcareitnews.com/news/healthcare-organizations-turning-wide-range-vendors-pandemic [accessed 2021-04-21]
28. Owings-Fonner N. Comparing the latest telehealth solutions. Let's Get Technical, American Psychological Association. 2019 Dec. URL: https://www.apaservices.org/practice/business/technology/tech-column/telehealth-solutions [accessed 2021-04-21]
29. NIH Clinical Center. URL: https://clinicalcenter.nih.gov/ [accessed 2021-12-06]
30. Surveillance Review and Response Group. United States COVID-19 Cases and Deaths by State Over Time. Centers for Disease Control and Prevention. 2020 Jun 11. URL: https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36 [accessed 2021-11-11]
31. Hollander JE, Carr BG. Virtually perfect? Telemedicine for Covid-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
32. Patel SY, Mehrotra A, Huskamp HA, Uscher-Pines L, Ganguli I, Barnett ML. Trends in outpatient care delivery and telemedicine during the COVID-19 pandemic in the US. JAMA Intern Med 2021 Mar 01;181(3):388-391 [FREE Full text] [doi: 10.1001/jamainternmed.2020.5928] [Medline: 33196765]
33. Breitinger S, Gentry MT, Hilty DM. Key opportunities for the COVID-19 response to create a path to sustainable telemedicine services. Mayo Clin Proc 2020 Dec;95(12):2602-2605 [FREE Full text] [doi: 10.1016/j.mayocp.2020.09.034] [Medline: 33276833]
34. Bashshur R, Doarn CR, Frenk JM, Kvedar JC, Woolliscroft JO. Telemedicine and the COVID-19 pandemic, lessons for the future. Telemed J E Health 2020 May;26(5):571-573. [doi: 10.1089/tmj.2020.29040.rb] [Medline: 32275485]
35. Weigel G, Ramaswamy A, Sobel L, Salganicoff A, Cubanski J, Freed M. Opportunities and Barriers for Telemedicine in the U.S. during the COVID-19 Emergency and Beyond. Kaiser Family Foundation. 2020 May 11. URL: https://tinyurl.com/ycktm79t [accessed 2021-08-05]

Abbreviations

CDC: Centers for Disease Control and Prevention
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
HITECH: Health Information Technology for Economic and Clinical Health
NIAAA: National Institute on Alcohol Abuse and Alcoholism
NIH: National Institutes of Health
NIMH: National Institute of Mental Health

Edited by G Eysenbach; submitted 26.04.21; peer-reviewed by R Krukowski, S Liu; comments to author 10.07.21; revised version received 06.08.21; accepted 06.09.21; published 28.01.22.

Please cite as:
Vogt EL, Welch BM, Bunnell BE, Barrera JF, Paige SR, Owens M, Coffey P, Diazgranados N, Goldman D
Quantifying the Impact of COVID-19 on Telemedicine Utilization: Retrospective Observational Study
Interact J Med Res 2022;11(1):e29880
URL: https://www.i-jmr.org/2022/1/e29880
doi:10.2196/29880
PMID:34751158

Original Paper

Assessment of Glycemic Control at St. Luke's Free Medical Clinic: Retrospective Chart Review

Wade Hopper[1*], OMS-III; Justin Fox[1*], OMS-III; JuliSu Dimucci-Ward[2*], MPH, PhD

[1]Department of Surgery, Edward Via College of Osteopathic Medicine - Carolinas, Spartanburg, SC, United States

[2]Department of Preventive Medicine and Public Health, Edward Via College of Osteopathic Medicine - Carolinas, Spartanburg, SC, United States

[*]all authors contributed equally

Corresponding Author:
Wade Hopper, OMS-III
Department of Surgery
Edward Via College of Osteopathic Medicine - Carolinas
350 Howard St
Spartanburg, SC, 29303
United States
Phone: 1 8643279800
Email: whopper@carolinas.vcom.edu

Abstract

Background: A free clinic is a health care delivery model that provides primary care and pharmaceutical services exclusively to uninsured patients. With a multidisciplinary volunteer clinical staff, which includes physicians, social workers, dieticians, and osteopathic medical students, St. Luke's Free Medical Clinic (SLFMC) cares for over 1700 patients annually in Spartanburg, South Carolina.

Objective: This study aims to measure the change, over time, in patient hemoglobin A_{1c} measurements at the SLFMC to quantify the success of the clinic's diabetes treatment program.

Methods: A prospective-retrospective chart review of patients (n=140) enrolled at the SLFMC between January 1, 2018, and January 1, 2021, was performed. Patients were stratified as having controlled (hemoglobin A_{1c}<7.0, n=53) or uncontrolled (hemoglobin A_{1c}≥7.0, n=87) diabetes relative to a therapeutic hemoglobin A_{1c} target of 7.0, which is recommended by the American Diabetes Association. For both controlled and uncontrolled groups, baseline hemoglobin A_{1c} values were compared to subsequent readings using a Wilcoxon matched-pairs signed rank test. Results from the SLFMC population were compared to the published literature on hemoglobin A_{1c} from other free clinics.

Results: Patients with uncontrolled diabetes experienced significant reductions in median hemoglobin A_{1c} at both 6 months (*P*=.006) and 1 year (*P*=.002) from baseline. Patients with controlled diabetes showed no significant changes. Black and Hispanic patients with uncontrolled diabetes experienced a 1.0% mean improvement in hemoglobin A_{1c} over the study window. The SLFMC's wholly uninsured patient population showed a population rate of controlled diabetes (42%), which was similar to recent nationwide averages for adults with diabetes (51% to 56%), as reported by the National Health and Nutrition Examination Survey. The clinic's Hispanic population (n=47) showed the greatest average improvement in hemoglobin A_{1c} of any ethnic group from baseline. Additionally, 61% of the SLFMC's Black population (n=33) achieved a hemoglobin A_{1c} of <7.0 by the end of the study window, which surpassed the nationwide averages for glycemic control.

Conclusions: We present free clinic hemoglobin A_{1c} outcomes obtained through a retrospective chart review. Uninsured patients treated for diabetes at the SLFMC show a reduction in hemoglobin A_{1c}, which is comparable to nationwide standards, although average hemoglobin A_{1c} levels in this study were higher than nationwide averages. Black and Hispanic patients with uncontrolled diabetes showed a mean 1% improvement in hemoglobin A_{1c} levels. These results represent some of the first in the literature emerging from a free clinic that is not affiliated with a major medical school.

(Interact J Med Res 2022;11(1):e31123) doi:10.2196/31123

KEYWORDS

free clinic; glycemic control; underserved; uninsured; diabetes; health care delivery; treatment program

Introduction

Free medical clinics are private nonprofit practices that provide a health care safety net for uninsured individuals in a community. Free clinics provide essential primary care services that include minimal paid staff and maximize the use of a volunteer clinical staff [1]. According to the Institute of Medicine, safety net practices are defined as "those providers that organize and deliver a significant level of health care and other needed services to uninsured, Medicaid, and other vulnerable patients" [2-4]. Free clinics seem to have arisen in response to specific populations' health care needs that are not satisfied by existing public or private programs [5]. Approximately 1400 free clinics provide care to over 2 million Americans each year [5]. Free clinics rely on charitable donations and grants to fund their daily practices, as most free clinics receive no federal funding [1,6]. In 2020, the National Association of Free Clinics (NAFC) formalized quality-of-care standards for all member associations [7]. Outcome reporting is becoming increasingly critical for individual free clinics to sustain growth through charitable donations.

An understanding of free clinic patients, practices, and outcomes can yield important insights into the efficacy of free clinics compared to that of other safety net institutions. Recent research by Julie Darnell has elucidated some commonalities across free clinics [1,5]. Namely, most of them are young: nearly 90% of existing free clinics were founded more recently than 2000 [5]. Furthermore, almost all free clinics rely upon volunteer health practitioners to provide the bulk of their care [1]. Many outcome studies have been published, which outlined free clinics' management of chronic disease [8-13]. Most publications in the free clinic sector represent student-ran, school-affiliated clinics. As entities, most free clinics have little to no incentive to publish outcomes data—after all, they do not receive government funding or participate in traditional insurance models. Moreover, reliance on volunteer staff places a human resources constraint on the publishing power of free clinics. This study represents one of the first of its kind to emerge from a free clinic that is not run by medical students.

The uninsured population is generally underrepresented in the medical literature. In 2019, over 26 million American adults—8.0% of the population—lacked health insurance [14]. An estimated 13%-16% of South Carolina adults aged 18-64 years lacked health insurance as of 2019 [15,16]. It is known that being uninsured places people at risk for incurring health care–associated debt, thereby delaying treatment [17]. Black and Hispanic people are overrepresented in the uninsured pool compared to the general population [18]. Despite the implementation of the Affordable Care Act, approximately 1 in 10 nonelderly Americans remained uninsured at the start of the 2020s [19]. Additionally, health concerns of the uninsured hold current relevance owing to the COVID-19–related economic shutdown in the United States. Health insurance in the United States is often tied to employment. Early during the COVID-19 pandemic, unemployment claims peaked in April 2020 at 36.5 million—14.7% of the population—and the largest figure observed since such records began [20,21]. Because free clinics exclusively serve the uninsured, research in this sector helps characterize the health behaviors and outcomes of uninsured Americans.

In 2020, St. Luke's Free Medical Clinic (SLFMC) enrolled 1700 patients and averaged about 140 visits per week (personal communication from Patricia Whitney, SLFMC clinic director, May 2021). Of these 1700 patients, 920 (54%) were non-White. The clinic annually receives over 7000 hours of help from over 50 clinical volunteers, including physicians, mid-level practitioners, nurses, pharmacists, social workers, and a dietician. Patients generally see the same providers on a recurring basis, and this facilitates long-term clinical relationships. There are two clusters of food deserts in Spartanburg which span 8 of the city's 20 census tracts [22]. The SLFMC provides a food bank and transportation assistance to patients with needs in those areas. The SLFMC's licensed dispensing pharmacy is stocked with a wide variety of generic medications, and in 2020, upward of US $6 million worth of medications were provided by the clinic at no or minimal cost to patients. Further, of note, the SLFMC provides in-house language interpretation and translation services. Some of the clinic's 334 Hispanic patients exclusively speak Spanish, and language assistance played a part in over 300 visits in 2020.

The SLFMC's patients apply for and are awarded contracts of care based on financial need. Patients may continue receiving care so long as they lack health insurance—if a patient becomes old enough to qualify for Medicare, for example, or if a patient finds insurance through an employer, then the SLFMC will rotate that patient out of their care to free up space for awaiting applicants. Type 2 diabetes mellitus (T2DM) is the most prevalent diagnosis at the SLFMC. SLFMC patients diagnosed with T2DM often have a few other diagnoses for which they are receiving appropriate care. Patients usually see the same physician or nurse practitioner on a recurring basis. Psychiatrists, dieticians, and social workers are also involved in diabetes care teams. Patients with diabetes receive counseling in accordance with the Association of Diabetes Care and Education Specialists' ADCES7 Self-Care Behaviors guidelines. Diabetes-specific pharmacotherapy was initiated as seen fit per practitioner. Finally, diabetes-specific follow-up visits typically occur every 3 to 6 months depending on the patient's glycemic control.

This study attempts to answer a handful of questions: what is the quality of diabetes care at the SLFMC as measured by glycemic control over time? With respect to hemoglobin A_{1c}, how does treatment at the SLFMC measurably compare to that provided by other free clinics? How do hemoglobin A_{1c} levels at the SLFMC compare to nationwide averages? Finally, do the clinic's hemoglobin A_{1c} outcomes differ depending on race or Hispanic ethnicity?

Methods

Methods Overview

Study methods were approved by the institutional review board at Edward Via College of Osteopathic Medicine. A retrospective chart review was used to gather data. This method is well-established in free clinic literature as being a cost- and time-efficient means of reporting outcomes and comparing them to existing professional standards [8-13]. The use of hemoglobin A_{1c} as a diabetes quality-of-care metric is endorsed by the US government and the American Diabetes Association (ADA) [23-25]. Hemoglobin A_{1c} is an indicator of 3-month average blood glucose as well as a predictor of diabetes-related morbidity and microvascular complications [26-28]. The ADA endorses a hemoglobin A_{1c} threshold of 7.0 as being appropriate for most adults [23].

Study Population

All patients' medical charts at the SLFMC between January 1, 2018, and January 1, 2021, were screened for a diagnosis of T2DM. Inclusion criteria were as follows: (1) a diagnosis of diabetes, prediabetes, hyperglycemia, or metabolic syndrome; and (2) two or more different hemoglobin A_{1c} levels documented, with at least 6 months between the first and last recorded hemoglobin A_{1c} levels. Patients with a documented pregnancy within the study timeframe were excluded owing to the potential confounding effect of gestational diabetes. Patients with diabetes were stratified into two groups: controlled (baseline hemoglobin A_{1c}<7.0) and uncontrolled (baseline hemoglobin A_{1c}≥7.0) [12,13]. Stratification was essential because the two groups have two different therapeutic goals: the goal when treating controlled diabetes is to maintain hemoglobin A_{1c} levels under 7.0, while the goal when treating uncontrolled diabetes is to reduce hemoglobin A_{1c} levels to <7.0 [23].

Data Collection

All available descriptive data and hemoglobin A_{1c} readings—including dates—for individual patients were transcribed from a hard copy medical record into Microsoft Excel. Descriptive data included age, gender, race/ethnicity, and BMI. Smoking status and the presence of concomitant diagnosis with essential hypertension (ICD-10-CM I10) were also recorded. Characteristics of the SLFMC sample were compared to data from Spartanburg County to explore what kinds of patients the SLFMC attracts [29,30]. Baseline hemoglobin A_{1c} values were defined as each patient's first recorded hemoglobin A_{1c} value within the study window. To facilitate comparison, subsequent hemoglobin A_{1c} values of each patient were rounded to either 6 months, 1 year, or 2 years from baseline. The baseline hemoglobin A_{1c} value was also used to sort individuals into 1 of 2 groups on the basis of whether they had controlled (hemoglobin A_{1c}<7.0) diabetes or uncontrolled (hemoglobin A_{1c}≥7.0) diabetes relative to a best-practice hemoglobin A_{1c} treatment goal of 7.0, as recommended by the ADA [23]. Finally, each patient's most recent hemoglobin A_{1c} values within the study timeframe were recorded. The most recent hemoglobin A_{1c} value on file was necessary for comparing treatment efficacy because not all patients received treatment for the same duration—some of the participants enrolled at the SLFMC in the middle of the study's timeframe, and some patients also left the clinic outright or aged out into Medicare.

Outcomes

The primary outcome measured per individual was the net change in hemoglobin A_{1c} ($\triangle A_{1c}$) from baseline to the most recent visit on record. At the group level, successful treatment was defined as maintenance of hemoglobin A_{1c} in patients with controlled diabetes and reduction in hemoglobin A_{1c} in those with uncontrolled diabetes. The primary outcome measured per group was median hemoglobin A_{1c}. Finally, the success of the SLFMC diabetes treatment program was assessed by stratifying the clinic's A_{1c} outcomes into three tiers of <7.0, 7.0-8.0, and >8.0. Proportions of patients within each range were compared to the hemoglobin A_{1c} data from various populations in the recent literature.

Statistical Analysis

A Wilcoxon matched-pairs signed rank test was performed for the median group's hemoglobin A_{1c} outcomes (controlled vs uncontrolled diabetes). The hemoglobin A_{1c} distribution was expected to show a skew, so the median value was chosen over the mean value to dampen the influence of outliers. A Bonferroni-Dunn post-test was applied to the resultant *P* values to standardize them to a single significance level (a) per group. Between-group medians, *P* values of <.05 were considered significant. Graphs were generated using GraphPad Prism (version 9; GraphPad Software, Inc).

Results

This study examined 140 individuals who had the clinical diagnosis of T2DM and were treated at the SLFMC. Of them, 38% (n=53) of patients had controlled diabetes and 62% (n=87) had uncontrolled diabetes at baseline. A demographic profile (n=140) is detailed in Table 1. The study group had a high proportion of Hispanic patients (34%) and a low proportion of Caucasian patients (41%) compared to county data (7% and 73%, respectively). High rates of comorbid obesity (61%) and hypertension (74%) were also observed at rates of nearly twice the local prevalence for these diagnoses. The mean hemoglobin A_{1c} value at baseline was 8.3 (SD 2.5). The distribution of hemoglobin A_{1c} values at baseline is presented in Figure 1.

As a group, patients with controlled diabetes were defined by a baseline hemoglobin A_{1c} value of <7.0. According to the most recent observations, 75% of the patients with controlled diabetes remained controlled hemoglobin A_{1c} values and 25% experienced increases in hemoglobin A_{1c} values, which transitioned them to the uncontrolled category (Table 2). Conversely, patients with uncontrolled diabetes were defined by baseline hemoglobin A_{1c} values of ≥7.0. At 1 year from baseline, 25% of patients with uncontrolled diabetes experienced

reductions in hemoglobin A_{1c} values, which transitioned them to the controlled category. This proportion declined over time to 21% according to most recent data.

Patients with uncontrolled diabetes experienced a significant reduction in median hemoglobin A_{1c} values at both 6 months (P=.006) and 1 year (P=.002) from baseline (Figure 2A). Within this group, both mean and median hemoglobin A_{1c} values decreased compared to baseline at every measurement interval. Patients with controlled diabetes showed no significant changes in median hemoglobin A_{1c} values (Figure 2B). Within this group, both mean and median hemoglobin A_{1c} values increased compared to baseline at every measurement interval. These data support the hypotheses that patients with uncontrolled diabetes would experience significant changes in median hemoglobin A_{1c} values, whereas those with controlled diabetes would not.

Table 1. Patient demographics.

Characteristics	St. Luke's Free Medical Clinic (N=140)	Spartanburg County [24,25]
Age (years), mean (SD)	52.4 (10.2)	—[a]
Sex, n (%)		
Male	62 (44)	— (48)
Female	78 (56)	— (52)
Race, n (%)		
White	57 (41)	— (73)
Hispanic	48 (34)	— (7)
Black	32 (23)	— (21)
Asian	3 (2)	— (2)
Baseline hemoglobin A_{1c} (%), mean (SD)	8.3 (2.5)	—
Hemoglobin A_{1c} readings per patient (%), mean (SD)	2.7 (0.9)	—
BMI (kg/m^2), mean (SD)	33.6 (8.6)	—
Obesity, n (%)	85 (61)	— (32)
Hypertension, n (%)	104 (74)	— (37)
Smoking tobacco use, n (%)	49 (35)	— (35)

[a]—: not available.

Figure 1. Baseline hemoglobin A_{1c} values for all 140 patients ranged from 4.4 to 15.8. Columns indicate ranges centered at the halfway point of each integer; column at 5, for example, contains all baseline hemoglobin A_{1c} values between 5 and 5.99. The number of patients within each hemoglobin A_{1c} range at baseline is indicated in red above each column.

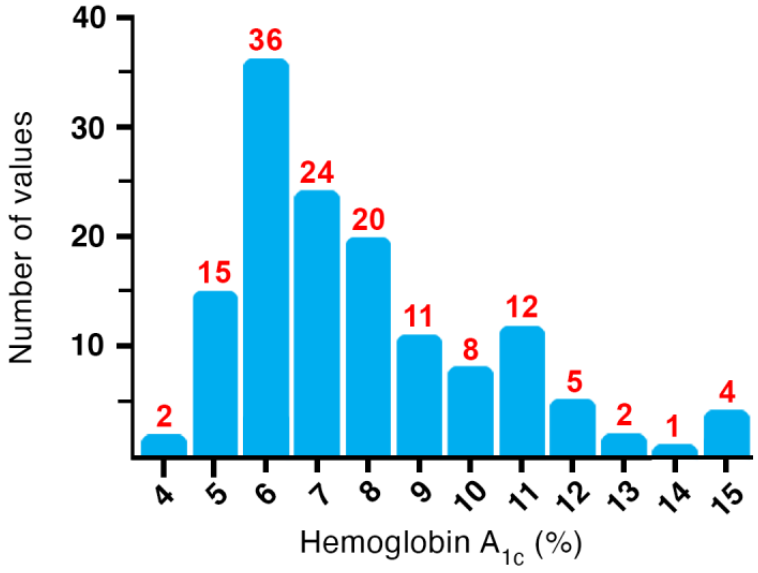
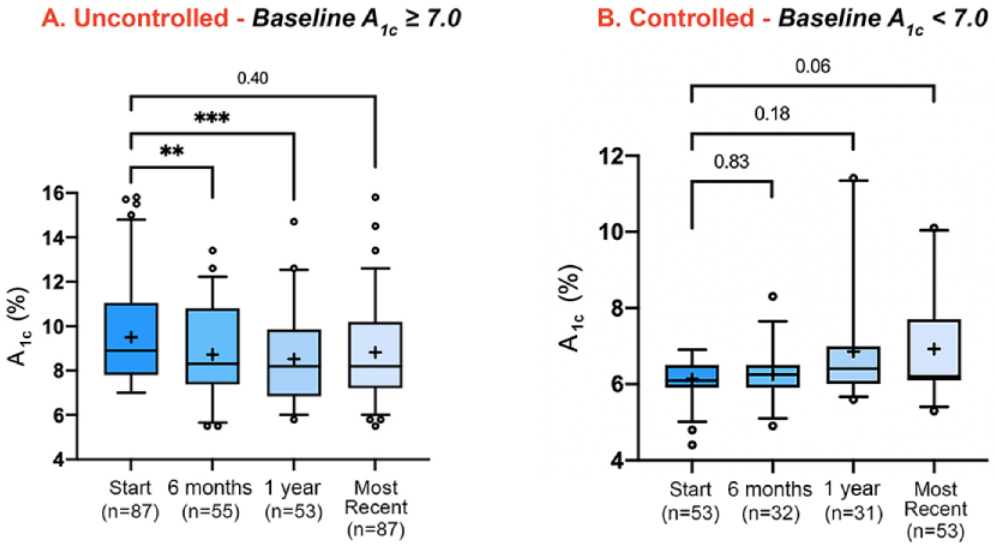


Table 2. Hemoglobin A_{1c} goal[a] achievement.

Time point	Controlled diabetes, n (%)	Uncontrolled diabetes, n (%)
Baseline	53 (100)	87 (0)
6 months	32 (85)	55 (20)
1 year	31 (77)	53 (25)
Most recent	53 (75)	87 (21)

[a]Proportion of patients who achieved a hemoglobin A_{1c} value of <7.0 at the aforementioned intervals relative to baseline. The sample size of each patient group is reported, followed by (in parentheses) the percentage of patients per sample achieving adequate glycemic control. For example, there were 55 patients with controlled diabetes who had hemoglobin A_{1c} data at 6 months from baseline. Of these 55 patients, 85% had a hemoglobin A_{1c} value of <7.0 at this time.

Figure 2. Boxes represent IQRs (25th-75th percentile) of hemoglobin A_{1c} values from patients with uncontrolled type 2 diabetes mellitus (A; left panel) and controlled type 2 diabetes mellitus (B; right panel) at baseline. Whiskers indicate 2 SDs from mean values (5%-95%). Lines within boxes represent median values. The "+" sign indicates mean values. Open dots represent outliers 2 SDs from the mean. ***P*=.006 ****P*=.002.



Glycemic control data were further stratified by race and Hispanic origin (Table 3). The clinic's Hispanic population initially had the highest mean hemoglobin A_{1c} values among all demographic groups (8.6%) and had the greatest reduction in mean ΔA_{1c} (–0.6) across groups. White patients experienced the poorest response to treatment with a mean ΔA_{1c} of +0.2. The clinic's Black population showed the highest rate of glycemic control (48% at baseline and 61% most recently).

Follow-up was measured as the percentage of patients in each group with a recorded hemoglobin A_{1c} value at 6 months and at 1 year. Patients with controlled diabetes had almost the same follow-up hemoglobin A_{1c} testing rates (63% at 6 months and 61% at 1 year) as patients with uncontrolled diabetes (61% at 6 months and 58% at 1 year). Follow-up rates did not differ significantly on the basis of gender, race, or Hispanic origin.

SLFMC data were compared to those of studies that similarly measured hemoglobin A_{1c} levels within target populations. The National Health and Nutrition Examination Survey (NHANES) generates samples with the intent of accurately representing the American adult population. Also included are 3 studies that reported posttreatment measures from student-run free clinics affiliated with 3 different medical schools. Rates of glycemic control at the SLFMC were superior to those observed at other free clinics and within 10% of 2 recently reported nationwide averages.

Finally, the rate at which the study group achieved controlled diabetes (hemoglobin A_{1c}<7.0) was calculated at baseline and from the most recent available data (Table 4). Initially, 38% of the SLFMC patients had controlled diabetes; this proportion rose to 42% by the end of the study. The outcomes observed at the SLFMC were compared to similar hemoglobin A_{1c} data sets gathered from a literature review. Rates of control at the SLFMC were within 10% of two recent NHANES samples reported by Ali et al and Carls et al [31,32]. A more recent NHANES sample reported by Fang [33] showed a 56% rate of glycemic control. Beyond the nationwide data sets, the SLFMC outcomes were also compared to those published by three different student-run free clinics in various parts of the country. All these free clinic studies used similar methods and reported stratified posttreatment glycemic control data [12,13]. The glycemic control rate of the SLFMC exceeded all posttreatment hemoglobin A_{1c} control rates in the free clinic literature, but the SLFMC population displayed a lower mean hemoglobin A_{1c} at baseline than did the included clinics [12,13]. The stratified data reported from other free clinics represents posttreatment (rather than baseline) hemoglobin A_{1c} levels.

Table 3. Hemoglobin A_{1c} data stratified by race and Hispanic origin among patients at St. Luke's Free Medical Clinic.

Metric	Non-Hispanic White (n=57)	Black (n=33)	Hispanic (n=47)
Hemoglobin A_{1c} value at baseline, mean	8.0	7.9	8.6
ΔA_{1c}[a], mean	+0.2	–0.3	–0.6
ΔA_{1c} in patients with controlled diabetes, mean (n)	+1.0 (22)	–0.3 (16)	+0.6 (14)
ΔA_{1c} in patients with uncontrolled diabetes, mean (n)	–0.4 (35)	–0.9 (17)	–1.1 (33)
Proportion of patients with a hemoglobin A_{1c} value of <7.0 at baseline, %	39	48	30
Proportion of patients with a hemoglobin A_{1c} value of <7.0 measured most recently, %	37	61	34

[a]ΔA_{1c}: change in hemoglobin A_{1c} values; "+" indicates an increase and "–" indicates a decrease.

Table 4. Cross-comparison of hemoglobin A_{1c} data.

Hemoglobin A_{1c} value, %	SLFMC[a] baseline measurement (N=140)	SLFMC most recent measurement (N=140)	NHANES[b] 2007-2010 data (N=1444)	NHANES 2011-2014 data (N=1326)	NHANES 2011-2016 data (N=5800)	UCSD[c] Free Clinic data (N=181)	Vanderbilt University Free Clinic data (N=45)	Icahn School of Medicine Free Clinic data (N=44)
<7.0, %	38	42	52	51	56	30	38	21
7.0-8.0, %	20	20	27	21	—[d]	29	24	18
>8.0, %	42	38	21	28	—	41	38	61
Hemoglobin A_{1c}, mean	8.3	8.3	7.2	—	7.2	9.2	9.6	10.1
ΔA_{1c}[e], mean	–0.3	–0.3	—	—	—	–1.0	–1.7	–1.3
Most recent hemoglobin A_{1c} value	8.0	8.0	—	—	—	8.2	7.9	8.8

[a]SLFMC: St. Luke's Free Medical Clinic.

[b]NHANES: National Health and Nutrition Examination Survey.

[c]UCSD: University of California San Diego.

[d]—: not applicable.

[e]ΔA_{1c}: change in hemoglobin A_{1c} values.

Discussion

Principal Findings

In analyzing the follow-up behaviors of the SLFMC's patients, we noted no significant difference in follow-up rates between patients with controlled and those with uncontrolled diabetes. However, there were some differences between the follow-up rates of different ethnic groups at the SLFMC. Prior research shows that Black and Hispanic patients tend to have higher baseline hemoglobin A_{1c} levels [34-37] and poorer self-management of diabetes, are less likely to achieve good glycemic control with treatment [38], and are more likely to experience diabetes-related complications resulting in worse diabetes outcomes and higher rates of mortality than non-Hispanic White patients [39]. Hispanic patients may specifically face obstacles associated with cultural differences between patients and providers, placing them at an increased risk for underuse of services, poor-quality care, and worse outcomes compared with non-Hispanic Whites [40]. The literature suggests that language barriers have a negative influence on health behaviors and outcomes among Latino patients [41]. To our knowledge, no unifying theory explains these known discrepancies in hemoglobin A_{1c} values among ethnic groups.

Black and Hispanic patients at the SLFMC defied the data cited above, with higher use of resources and better health outcomes than non-Hispanic White patients at the clinic (Table 3). Hispanic patients had a mean ΔA_{1c} of –0.6 while enrolled at the SLFMC. Black patients had a remarkable 61% glycemic control rate by the end of the study, which was the highest of any other race. SLFMC provides free care, including many free prescription medications, and we believe that ease of access to care is a primary driver of the trend-reversing outcomes achieved by the SLFMC's Black and Hispanic patients with diabetes. Generally, Black and Hispanic populations are subject to disproportionately low health care access [42,43]. Furthermore, patient trust in providers is a direct predictor of subjective treatment effect [44,45]. The authors offer the following original

thoughts on care for minorities within the free clinic setting: we suggest that patients are more likely to trust practitioners who provide their services free of charge. It is likely that patients feel more comfortable and less intimidated in the setting of charitable care compared to payment-based models. We also suggest that the ratio of immigrants in the Hispanic population is higher than that in other ethnicities [46]. We think that immigrants are more likely to use free health services out of appreciation for some of the things that American culture takes for granted. The aforementioned factors probably influenced the stronger improvements in hemoglobin A_{1c} values observed in Black and Hispanic patients than in Non-Hispanic White patients (Table 3).

Free clinic populations universally show hemoglobin A_{1c} averages at least a full point above nationwide averages for adults with diabetes (Table 4). The NHANES, which was chosen as the national comparison data set, generates annual samples with the intent of representing the American population as a whole. Most of the patients included in the NHANES sample have health insurance coverage; hence, NHANES suitably represents the glycemic control of the insured population. The SLFMC data are encouraging and suggest that free clinics may be able to approach the quality of care provided by traditional models of outpatient practices.

The glycemic control data gained in this study add to the growing body of knowledge characterizing the health of patients who utilize free clinic services. At baseline, The SLFMC population demonstrated the highest rate of controlled diabetes (38%) and the lowest mean hemoglobin A_{1c} value (8.3) observed in free clinic outcome literature, as well as the smallest mean improvement in hemoglobin A_{1c} values (–0.3). These findings are likely attributable to the relatively good health of the clinic's population at the start the study. The COVID-19 pandemic occurred within the study timeline, and it probably influenced the modest mean improvement in hemoglobin A_{1c} value of –0.3. Altogether, this study presents the current best-case scenario of free clinic glycemic control at baseline.

Quantifying free clinic quality of care is a matter of importance to individual free clinics for the purpose of fundraising, but this aim is complicated by many factors. One such factor is time. The SLFMC has used paper charts for decades and has not completely transitioned to electronic medical records (EMRs). Many other free clinics are likely in the same situation because EMRs are costly to adopt. Sifting through paper charts takes more time than searching through a well-kept EMR. Another factor is available labor. Volunteer providers may not always be able to spare the time it takes to write up and publish outcomes. Despite obstacles of time and cost, the NAFC recently approved quality reporting standards, which it expects each member clinic to abide by. This presents a valuable professional growth opportunity for medical students and other postgraduate medical professionals. We recommend that free clinics take further initiative and publish outcomes when possible.

Limitations

The retrospective cross-sectional design is limited in scope and cannot explain relationships between variables. Study data were extracted largely from paper charts and were subject to potential inaccuracies. The clinic is currently transitioning from a paper chart system to an EMR, and some charts were unavailable for evaluation. The SLFMC yielded a population of 140 patients with diabetes with a unique demographic profile that may not generalize well to other clinics. Furthermore, this study did not consider the specific diabetes treatment modalities provided to each patient (pharmaceuticals, nutrition counseling, home glucose monitoring, etc) which were excluded from the study because treatment varies widely across patients and providers. One final limitation is that not all patients had the same number of follow-up hemoglobin A_{1c} tests—such variations were possibly due to, for example, duration of enrollment (including loss to follow-up) and missed appointments.

Conclusions

This report highlights a free clinic diabetes treatment program that provides significant benefits to patients with uncontrolled T2DM. The SLFMC was successful in treating minority patients, with uncontrolled Black and Hispanic patients showing a mean 1% improvement in hemoglobin A_{1c} values. Rates of glycemic control at the SLFMC were superior to those observed at other free clinics and within 10% below two recently reported nationwide averages. This study presents one of the largest sample sizes yet observed in the free clinic hemoglobin A_{1c} literature. Limitations include loss to follow-up and incomplete patient records. Meta-analysis of comparable outcomes data is the intuitive next step in the journey toward a better understanding of the quality of care provided at free clinics.

Acknowledgments

We would like to thank Patsy Whitney (executive director of the SLFMC), Petra Warren, MD, Nancy Anderson (registered nurse), and Julie Darnell, PhD.

Conflicts of Interest

None declared.

References

1. Darnell JS. Free clinics in the United States: a nationwide survey. Arch Intern Med 2010 Jun 14;170(11):946-953. [doi: 10.1001/archinternmed.2010.107] [Medline: 20548006]

2. Lewin M, Altman S, Institute of Medicine (US) Committee on the Changing Market, Managed Care, and the Future Viability of Safety Net Providers. Americas's Health Care Safety Net: Intact but Endangered. Washington, DC: National Academies Press; 2000.
3. Moore SL, Fischer I, Havranek EP. Translating Health Services Research into Practice in the Safety Net. Health Serv Res 2016 Feb;51(1):16-31 [FREE Full text] [doi: 10.1111/1475-6773.12341] [Medline: 26646189]
4. Nguyen OK, Makam AN, Halm EA. National Use of Safety-Net Clinics for Primary Care among Adults with Non-Medicaid Insurance in the United States. PLoS One 2016;11(3):e0151610 [FREE Full text] [doi: 10.1371/journal.pone.0151610] [Medline: 27027617]
5. Darnell J. What is the role of free clinics in the safety net? Med Care 2011 Nov;49(11):978-984. [doi: 10.1097/MLR.0b013e3182358e6d] [Medline: 22005605]
6. National Association of Free and Charitable Clinics. Definition of a Free or Charitable Clinic. URL: https://www.nafcclinics.org/content/clinics [accessed 2021-01-19]
7. National Association of Free and Charitable Clinics Launches Membership Quality Standards Program. National Association of Free and Charitable Clinics. 2020 Jun 10. URL: https://www.nafcclinics.org/sites/default/files/NAFC%20Quality%20Standards_2.pdf [accessed 2021-02-21]
8. Atkinson B, Uichanco L, Kattih Z. Review of Hypertension Management at University of South Florida BRIDGE Clinic. J Student-Run Clin 2018;4(1).
9. Rojas S, Smith S, Rojas S, Vaida F. Longitudinal hyperlipidemia outcomes at three student-run free clinic sites. Fam Med 2015 Apr;47(4):309-314 [FREE Full text] [Medline: 25853603]
10. Smith SD, Marrone L, Gomez A, Johnson ML, Edland SD, Beck E. Clinical outcomes of diabetic patients at a student-run free clinic project. Fam Med 2014 Mar;46(3):198-203 [FREE Full text] [Medline: 24652638]
11. Zucker J, Gillen J, Ackrivo J, Schroeder R, Keller S. Hypertension management in a student-run free clinic: meeting national standards? Acad Med 2011 Feb;86(2):239-245. [doi: 10.1097/ACM.0b013e31820465e0] [Medline: 21169778]
12. Gorrindo P, Peltz A, Ladner TR, Reddy I, Miller BM, Miller RF, et al. Medical students as health educators at a student-run free clinic: improving the clinical outcomes of diabetic patients. Acad Med 2014 Apr;89(4):625-631 [FREE Full text] [doi: 10.1097/ACM.0000000000000164] [Medline: 24556762]
13. Laitman B, Mosley G, Thomas D, Meah Y. How Well Does a Student-Run Free Clinic Care for Diabetic Patients? A Retrospective Study of the East Harlem Health Outreach Partnership. J Student-Run Clin 2017;3(1):1-9.
14. Keisler-Starkey K, Bunch L. Health Insurance Coverage in the United States: 2020. United States Census Bureau. 2021 Sep 14. URL: https://www.census.gov/library/publications/2021/demo/p60-274.html [accessed 2022-01-13]
15. Health Insurance Coverage: Estimates from the National Health Interview Survey (2010-2020). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nhis/healthinsurancecoverage.htm [accessed 2021-04-18]
16. Health Insurance Coverage of Adults 19-64. Kaiser Family Foundation. URL: https://tinyurl.com/ye2xw5cz [accessed 2021-04-18]
17. Garfield R, Orgera K, Damico A. The Uninsured and the ACA: A Primer - Key Facts about Health Insurance and the Uninsured amidst Changes to the Affordable Care Act. Kaiser Family Foundation. 2019 Jan 25. URL: https://tinyurl.com/2p8ecnkb [accessed 2020-12-14]
18. Rhoades JA, Cohen SB. The Long-Term Uninsured in America, 2007-2010 (Selected Intervals): Estimates for the U.S. Civilian Noninstitutionalized Population under Age 65. In: Statistical Brief (Medical Expenditure Panel Survey. Rockville, MD: Agency for Healthcare Research and Quality; 2013.
19. Garfield R, Orgera K, Damico A. Key Facts about the Uninsured Population. Kaiser Family Foundation. URL: https://www.kff.org/uninsured/issue-brief/key-facts-about-the-uninsured-population/ [accessed 2020-12-14]
20. Transmission of material in this news release is embargoed until 8:30 a.m. (EDT) Friday, May 8, 2020. Bureau of Labor Statistics, US Department of Labor. URL: https://www.bls.gov/news.release/archives/empsit_05082020.pdf [accessed 2020-05-08]
21. Unemployment Insurance Weekly Claims: Week of May 9, 2020. US Department of Labor. URL: https://oui.doleta.gov/press/2020/051420.pdf [accessed 2020-05-21]
22. Carter CL. Race, Resources, and Income Profiles. Data Visualization of Census Tracts in Spartanburg. South Carolina Commission for Minority Affairs. 2019. URL: https://dc.statelibrary.sc.gov/bitstream/handle/10827/33168/CMA_Race_Resources_and_Income_Profiles_2019-12.pdf?sequence=1 [accessed 2021-04-18]
23. American Diabetes Association. 6. Glycemic Targets:. Diabetes Care 2019 Jan;42(Suppl 1):S61-S70. [doi: 10.2337/dc19-S006] [Medline: 30559232]
24. Health Insurance Exchange: 2020 Quality Rating System Measure Technical Specifications. Centers for Medicare & Medicaid Services. 2019. URL: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/Downloads/2020-QRS-Measure-Tech-Specs_20190925_508.pdf [accessed 2021-01-10]
25. Centers for Medicare & Medicaid Services. URL: https://cmit.cms.gov/CMIT_public/ViewMeasureProgramDetails?measureId=2837&measureProgramId=2530 [accessed 2021-01-10]

26. Diabetes Control Complications Trial Research Group, Nathan DM, Genuth S, Lachin J, Cleary P, Crofford O, et al. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 1993 Sep 30;329(14):977-986. [doi: 10.1056/NEJM199309303291401] [Medline: 8366922]
27. HbA1c and Estimated Average Glucose (eAG). National Institutes of Diabetes and Digestive and Kidney Diseases. 2010. URL: http://www.ngsp.org/A1ceAG.asp [accessed 2021-01-10]
28. Stratton I, Adler A, Neil H, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ 2000 Aug 12;321(7258):405-412 [FREE Full text] [doi: 10.1136/bmj.321.7258.405] [Medline: 10938048]
29. Spartanburg County Community Health Needs Assessment. University of South Carolina. 2019 Dec 13. URL: https://tinyurl.com/3mxsry8t [accessed 2021-02-28]
30. Brady K. The Status of Public Health in Spartanburg County, SC. Metropolitan Studies Institute at University of South Carolina Upstate. 2016. URL: https://tinyurl.com/ycyf4vfw [accessed 2021-02-28]
31. Carls G, Huynh J, Tuttle E, Yee J, Edelman SV. Achievement of Glycated Hemoglobin Goals in the US Remains Unchanged Through 2014. Diabetes Ther 2017 Aug;8(4):863-873 [FREE Full text] [doi: 10.1007/s13300-017-0280-5] [Medline: 28646411]
32. Ali MK, Bullard KM, Saaddine JB, Cowie CC, Imperatore G, Gregg EW. Achievement of goals in U.S. diabetes care, 1999-2010. N Engl J Med 2013 Apr 25;368(17):1613-1624. [doi: 10.1056/NEJMsa1213829] [Medline: 23614587]
33. Fang M. Trends in Diabetes Management Among US Adults: 1999-2016. J Gen Intern Med 2020 May;35(5):1427-1434 [FREE Full text] [doi: 10.1007/s11606-019-05587-2] [Medline: 31898135]
34. Herman WH, Ma Y, Uwaifo G, Haffner S, Kahn SE, Horton ES, Diabetes Prevention Program Research Group. Differences in A1C by race and ethnicity among patients with impaired glucose tolerance in the Diabetes Prevention Program. Diabetes Care 2007 Oct 29;30(10):2453-2457 [FREE Full text] [doi: 10.2337/dc06-2003] [Medline: 17536077]
35. Herman W, Dungan K, Wolffenbuttel B, Buse JB, Fahrbach JL, Jiang H, et al. Racial and ethnic differences in mean plasma glucose, hemoglobin A1c, and 1,5-anhydroglucitol in over 2000 patients with type 2 diabetes. J Clin Endocrinol Metab 2009 May;94(5):1689-1694. [doi: 10.1210/jc.2008-1940] [Medline: 19276235]
36. Ziemer D, Kolm P, Weintraub W, Vaccarino V, Rhee MK, Twombly JG, et al. Glucose-independent, black-white differences in hemoglobin A1c levels: a cross-sectional analysis of 2 studies. Ann Intern Med 2010 Jun 15;152(12):770-777. [doi: 10.7326/0003-4819-152-12-201006150-00004] [Medline: 20547905]
37. Kirk JK, Passmore LV, Bell RA, Narayan KV, D'Agostino RB, Arcury TA, et al. Disparities in A1C levels between Hispanic and non-Hispanic white adults with diabetes: a meta-analysis. Diabetes Care 2008 Feb;31(2):240-246. [doi: 10.2337/dc07-0382] [Medline: 17977939]
38. Campbell JA, Walker RJ, Smalls BL, Egede LE. Glucose control in diabetes: the impact of racial differences on monitoring and outcomes. Endocrine 2012 Dec;42(3):471-482 [FREE Full text] [doi: 10.1007/s12020-012-9744-6] [Medline: 22815042]
39. Diabetes Disparities Among Racial and Ethnic Minorities. US Department of Health and Human Services. URL: https://archive.ahrq.gov/research/findings/factsheets/diabetes/diabdisp/diabdisp.html [accessed 2021-03-28]
40. Brown AF, Gerzoff RB, Karter AJ, Gregg E, Safford M, Waitzfelder B, TRIAD Study Group. Health behaviors and quality of care among Latinos with diabetes in managed care. Am J Public Health 2003 Oct;93(10):1694-1698. [doi: 10.2105/ajph.93.10.1694] [Medline: 14534224]
41. Fernandez A, Schillinger D, Warton EM, Adler N, Moffet HH, Schenker Y, et al. Language barriers, physician-patient language concordance, and glycemic control among insured Latinos with diabetes: the Diabetes Study of Northern California (DISTANCE). J Gen Intern Med 2011 Feb;26(2):170-176 [FREE Full text] [doi: 10.1007/s11606-010-1507-6] [Medline: 20878497]
42. Canedo JR, Miller ST, Schlundt D, Fadden MK, Sanderson M. Racial/Ethnic Disparities in Diabetes Quality of Care: the Role of Healthcare Access and Socioeconomic Status. J Racial Ethn Health Disparities 2018 Feb;5(1):7-14 [FREE Full text] [doi: 10.1007/s40615-016-0335-8] [Medline: 28078660]
43. James CV, Moonesinghe R, Wilson-Frederick SM, Hall JE, Penman-Aguilar A, Bouye K. Racial/Ethnic Health Disparities Among Rural Adults - United States, 2012-2015. MMWR Surveill Summ 2017 Nov 17;66(23):1-9 [FREE Full text] [doi: 10.15585/mmwr.ss6623a1] [Medline: 29145359]
44. Du L, Wu R, Chen X, Xu J, Ji H, Zhou L. Role of Treatment Adherence, Doctor-Patient Trust, and Communication in Predicting Treatment Effects Among Tuberculosis Patients: Difference Between Urban and Rural Areas. Patient Prefer Adherence 2020;14:2327-2336 [FREE Full text] [doi: 10.2147/PPA.S277650] [Medline: 33262582]
45. Birkhäuer J, Gaab J, Kossowsky J, Hasler S, Krummenacher P, Werner C, et al. Trust in the health care professional and health outcome: A meta-analysis. PLoS One 2017;12(2):e0170988 [FREE Full text] [doi: 10.1371/journal.pone.0170988] [Medline: 28170443]
46. Key findings about U.S. immigrants. Pew Research Center. 2020 Aug 20. URL: https://www.pewresearch.org/fact-tank/2020/08/20/key-findings-about-u-s-immigrants/ [accessed 2021-09-24]

Abbreviations

△A_{1c}: net change in hemoglobin A_{1c}
ADA: American Diabetic Association
EMR: electronic medical record
NAFC: National Association of Free Clinics
NHANES: National Health and Nutrition Examination Survey
SLFMC: St. Luke's Free Medical Clinic
T2DM: type 2 diabetes mellitus

Edited by A Mavragani; submitted 09.06.21; peer-reviewed by A Ables, P Swamy; comments to author 11.09.21; revised version received 09.10.21; accepted 10.12.21; published 24.01.22.

Please cite as:
Hopper W, Fox J, Dimucci-Ward J
Assessment of Glycemic Control at St. Luke's Free Medical Clinic: Retrospective Chart Review
Interact J Med Res 2022;11(1):e31123
URL: https://www.i-jmr.org/2022/1/e31123
doi:10.2196/31123
PMID:35072636

Original Paper

Learning Agility of Learning and Development Professionals in the Life Sciences Field During the COVID-19 Pandemic: Empirical Study

XinYun Peng[1], MEd; Nicole Wang-Trexler[2], PhD; William Magagna[3], MEd; Susan Land[1], PhD; Kyle Peck[1], PhD

[1]Department of Learning and Performance Systems, Pennsylvania State University, University Park, PA, United States

[2]Vanguard Group, Inc, Valley Forge, PA, United States

[3]Siemens Healthineers, Newark, DE, United States

Corresponding Author:
XinYun Peng, MEd
Department of Learning and Performance Systems
Pennsylvania State University
301 Keller Building
University Park, PA, 16802
United States
Phone: 1 4154201314
Email: xypeng@psu.edu

Abstract

Background: The COVID-19 pandemic has impacted the life sciences field worldwide. Life sciences organizations (eg, pharmaceutical and med-tech companies) faced a rapidly increasing need for vital medical products, patient support, and vaccine development. Learning and development (L&D) departments play a crucial role in life sciences organizations as they apply learning initiatives to organizational strategy within a constantly evolving sector. During the COVID-19 pandemic, the work of L&D professionals in life sciences organizations changed profoundly during the abrupt shift to remote work, since learning and training normally occur in a face-to-face environment. Given the complex and dynamic situation of the pandemic, both individuals and organizations needed to learn quickly and apply what they learned to solve new, unprecedented problems. This situation presents an opportunity to study how characteristics of learning agility were evidenced by life sciences organizations and individual employees in the remote working mode.

Objective: In collaboration with Life Sciences Trainers & Educators Networks (LTEN), this study investigated the responses and learning agility of L&D professionals and their organizational leadership within the life sciences sector to the work changes due to the pandemic. The study answered the following questions: (1) How did L&D professionals in the life sciences sector respond to the changes in their work environment during the COVID-19 pandemic? (2) How did L&D professionals in the life sciences sector demonstrate learning agility during remote working?

Methods: We adopted a mixed methods approach that included a semistructured interview and a survey. Participants who were life sciences or health care L&D practitioners and in relevant positions were recruited via email through the LTEN and its partner pharmaceutical, biotech, or medical devices organizations. Interviews with 12 L&D professionals were conducted between June and August 2020 through phone or online conferencing, covering 22 open-ended questions to stimulate ideas that could be explored further in the survey. The semistructured interview questions were grounded in theory on learning agility. In total, 4 themes were developed from the interviews, which formed the basis for developing the survey items. The subsequent survey regarding 4 specific themes was conducted from August to October 2020 using Qualtrics. Both interview and survey data were analyzed based on a learning agility framework.

Results: Findings revealed generally positive organizational and individual responses toward the changes brought about by the pandemic. Results also indicated that a disruptive crisis, such as the shift from working in the office to working from home (WFH), required professionals' learning agility to both self-initiate their own learning and to support the learning agility of others in the organization.

Conclusions: This study was designed to better understand education and training in the life sciences field, particularly during the unique circumstances of the global COVID-19 pandemic. We put forward several directions for future research on the learning agility of L&D professionals in life sciences organizations.

(Interact J Med Res 2022;11(1):e33360) doi:10.2196/33360

KEYWORDS

COVID-19; learning agility; learning and development professionals; life sciences professionals; training and development; mixed methods

Introduction

During 2020 and 2021, almost every organization evolved and shifted to address the COVID-19 pandemic, including extensive numbers of employees working from home (WFH) due to lockdowns and shelter-in-place orders [1]. As the coronavirus continues to spread, many organizations (at the time of this writing) still have not set a date to return to their physical offices. Learning and development (L&D) professionals excel at understanding the organization's future capability needs and identifying priorities and learning solutions for the organization. Therefore, they are playing a pivotal role in transitioning and implementing changes within their organizations. Recent L&D studies have investigated learning in the health and life sciences with informational technologies, virtual programs, or online platforms, as well as their feasibility and effectiveness during the pandemic [2-7]. Much of the recent pandemic-related research on health and life sciences is concerned with digital mental health, especially stress and depression trends [8-11]. However, there is little research on the learning agility of L&D professionals in the health and life sciences, particularly how they dealt with the abrupt change to remote working. We argue that the learning agility of L&D professionals is essential to the survival and growth of organizations. This study describes the learning agility of L&D professionals in the life sciences sector, as they encountered drastic changes in their job requirements due to COVID-19.

The concept of learning agility was coined by Lombardo and Eichinger [12]. It is defined as the willingness and ability to learn from experience and subsequently apply that learning to perform successfully under new or first-time conditions. Learning agility is tied closely to developmental job experiences and reflects the complexity of challenging jobs. As such, it is considered an early indicator of one's potential and leadership effectiveness and therefore is used by organizations to identify and develop high-potential employees [13-15]. To elaborate on the Lombardo and Eichinger [12] definition, learning agility describes the following characteristics of a person: the willingness to adapt to new job requirements, the ability to continuously learn new things, to overcome difficulties, and to manage multiple, sometimes contradictory, tasks. Being mentally prepared for job requirements that are different and unfamiliar and being prepared to constantly learn new things for them are prerequisites of being agile. Effective leaders normally have a broad portfolio of leadership roles and can vary their performance of leadership skills, depending on the situation, showing high learning agility when encountering difficulties. In addition, existing and widely acknowledged theories of leadership tend to classify one's leadership into contrasting categories, for example, task oriented versus relation oriented [16], directive versus participative [17], autocratic versus consultative [18], and transactional versus transformational [19]. Lombardo and Eichinger [12] proposed that effective leaders should be able to accommodate multiple opposing categories to react to dilemmas under multiple circumstances.

Learning agility is frequently raised in corporate conversations and business reports when discussing how the working mode for people changed so abruptly due to the COVID-19 pandemic. Extending from the Lombardo and Eichinger [12] concept, other researchers [20] have studied or applied learning agility in different contexts. For instance, De Meuse [20] studied the development and validation of the TALENTx7 Assessment, which is a psychological measure of learning agility [20]. This work expanded the original model into a 7-factor model, and later Burke and Mitchinson [21] developed a 9-factor model based on the original learning agility concept [21]. Norton [22] studied leadership flexibility and included learning agility as 1 of the definitional perspectives. Likewise, DeRue and Myers [23] built a framework for leadership development called PREPARE using learning agility as a key element. In Kaiser and Craig's [24] view, learning agility is a meta-competency, meaning that it is the fundamental capacity that enables other technical competencies.

Other empirical studies show how learning agility can be applied. For example, Nesbit [25] used learning agility to assess leaders' knowledge and skill acquisition and therefore assist with their behavioral repertoire expansion in self-directed leadership development. White and Shullman [26] discussed learning agility, especially the ability to accept the ambiguity of the working environment, as an indicator of effective leadership. When it comes to leadership across different managerial levels within an organization, learning agility was a positive predictor of leaders' effectiveness in talent management [27]. De Meuse et al [27] noted that most people in leadership roles do not increase all-around competencies simultaneously when the job requirements change. Since it was reportedly rare in the management population to have high learning agility, it would be prudent for organizations to select individuals for key positions using the learning agility framework as a reference [27].

Studies on learning agility are often based on the premise that individuals actively seek professional development opportunities [13]. However, learning agility can also manifest within a specific circumstance where it is passively triggered (eg, the turbulent environment of the COVID-19 pandemic), which has not been previously discussed. In the case of COVID-19, both work content and format shifted. Individuals and organizations did not choose, but were forced, to develop and apply learning agility to survive. Learning agility can also exist on an organizational level. An organization's reactions to turbulent environments and its strategies to solve novel challenges that impact many employees are also critical indicators of the long-term success of the organization [28]. Therefore, this study

aims to describe the learning agility of L&D professionals in the life sciences sector when they encountered abrupt changes in their job requirements due to the emergence of the COVID-19 pandemic. Specifically, we answered the following research questions:

- How did L&D professionals in the life sciences sector respond to the changes in their work environment during the COVID-19 pandemic?
- How did L&D professionals in the life sciences sector demonstrate learning agility during remote working?

Methods

Study Design

This research study used a mixed methods approach to understanding how L&D professionals in the life sciences sector dealt with changes in their work due to the pandemic. Specifically, we were interested in their perceptions, solutions, and expectations for the future. Mixed methods research requires data triangulation from quantitative and qualitative data, which strengthens the construct validity of the study [29].

Participants and Recruitment

Participants were recruited through an email list of Life Sciences Trainers & Educators Networks (LTEN) and its partner organizations, which included pharmaceutical companies, medical device manufacturing companies, biotechnology companies, and training and consulting companies with core services in the life sciences sector. The invitation emails were sent to the L&D departments of these organizations. Additional personnel who work closely with L&D departments, for example, the sales department, were also invited to participate. Salespersons were an important data source, as they are served by L&D departments and they directly interact with health care workers. After receiving the invitation emails, anyone who was interested in participating in this study could contact the researchers to complete informed consent, schedule an interview, or access the questionnaire through a link in the email. In the first phase of this study, we recruited 12 L&D professionals, whose experience ranged from 10 years to more than 30 years and held director or c-suite L&D positions in pharmaceutical, biotechnology, and medical device organizations, to participate in the interview. We intentionally focused our sampling for the interviews of experienced L&D practitioners, as they worked in the life sciences and health care L&D longer and witnessed the evolution of this area. Additionally, they had more connections with stakeholders, allowing them to have a macrolevel perspective. In the second phase of the study, we collected survey responses from 74 different individuals who held a variety of leadership positions.

Qualitative Method

The interview was used to gather insights into overarching changes of professionals' perceptions and mindsets about working remotely through the lens of learning agility. The semistructured interview was designed based on the existing literature and our subject matter experts' understanding of the status quo of L&D in life sciences and health care. It contained 22 questions, with topics covering experiences and opinions, virtual solutions, digital literacy, and the future, making the conversations flow naturally. Interviewees' responses and researcher's notes served as data sources for the second phase of data collection. See Multimedia Appendix 1 for the interview questions.

Quantitative Method

A follow-up survey questionnaire was designed based on the preliminary data collected through the interviews and expanded to 37 questions in total, with 8 (22%) demographic questions and 29 (78%) questions regarding 4 specific themes: organizational actions, remote working, L&D, and the future. Respondents were asked about their perceptions and expectations on these themes. See Multimedia Appendix 2 for the survey questions.

Data Collection and Analysis

The interview data collection started during June 2020 and ended in August 2020. Interviews were conducted through videoconferencing or over the phone with 12 individuals that lasted approximately 30-60 minutes each. The survey data were collected from August to October 2020 using Qualtrics. Interviews were first transcribed and then coded and organized into groups of topics. These topics were expanded and specified into survey questions that were used in the second phase. The survey questions were designed for exploratory descriptive analysis with the intention to capture nuances of how professionals in a greater scale adapted their professional lives during the pandemic in contrast to exploring their psychological states or traits. Later, the survey responses and the interview data were coded and reorganized based on the learning agility framework adapted from Eichinger et al [30]. This framework of learning agility could be generalized as 4 key characteristics of learning agility: the willingness to adapt to new job requirements, the ability to handle jobs with increasing complexity, the ability to continuously learn new things, and the ability to overcome difficulties.

Ethics Approval

In compliance with our university's Institutional Review Board protocols (Study ID STUDY00009028), all participants signed an informed consent release prior to their data being collected. The research procedures were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Declaration of Helsinki of 1975, as revised in 2000.

Participants were told that they did not have to answer any question they did not want to answer and could stop their participation at any time. All identities and data were kept confidential and anonymous.

Results

Demographics of Respondents

The participants of this study were professionals (58 [88%] of 66) who held leadership roles in the L&D or equivalent departments of their life sciences organizations. As Figure 1 illustrates, 46 (70%) of 66 survey participants had 11+ years of experience, with 14 (30%) of these having 21 or more years of

experience. More than 34 (51%) of the 66 participants were in director-level positions, 13 (20%) were managers, 12 (18%) were executives, and 7 (11%) were developers or trainers (Figure 1).

Survey results also showed the respondents' organization information. Respondents worked for medical device manufacturers (16 [23%] of 71 responses), pharmaceutical companies (30 [42%] of 71 responses), biotech companies (10 [14%] of 71 responses), suppliers (6 [9%] of 71 responses), and other types of organizations (eg, training companies, consulting firms, and labs). There were 28 (42%) of 66 respondents who worked in organizations that have more than 10,000 employees. Among the organizations of all respondents, 41 (62%) of 66 are entirely US based and 28 (42%) of 66 are directly involved in COVID-19 diagnostics or treatment (Figure 2). This information about participants' leadership experience and type of organizations helped us interpret the survey and interview responses regarding their behaviors in response to the drastic shift of work, as well as their thoughts about the changes.

Figure 1. Information about respondents. L&D: learning and development.

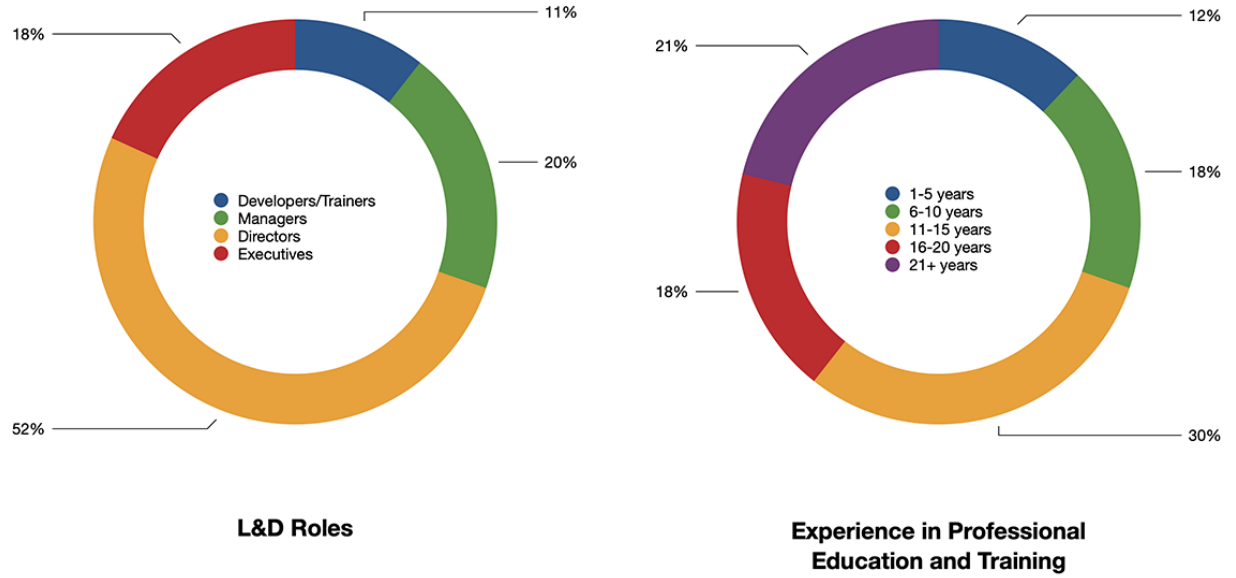
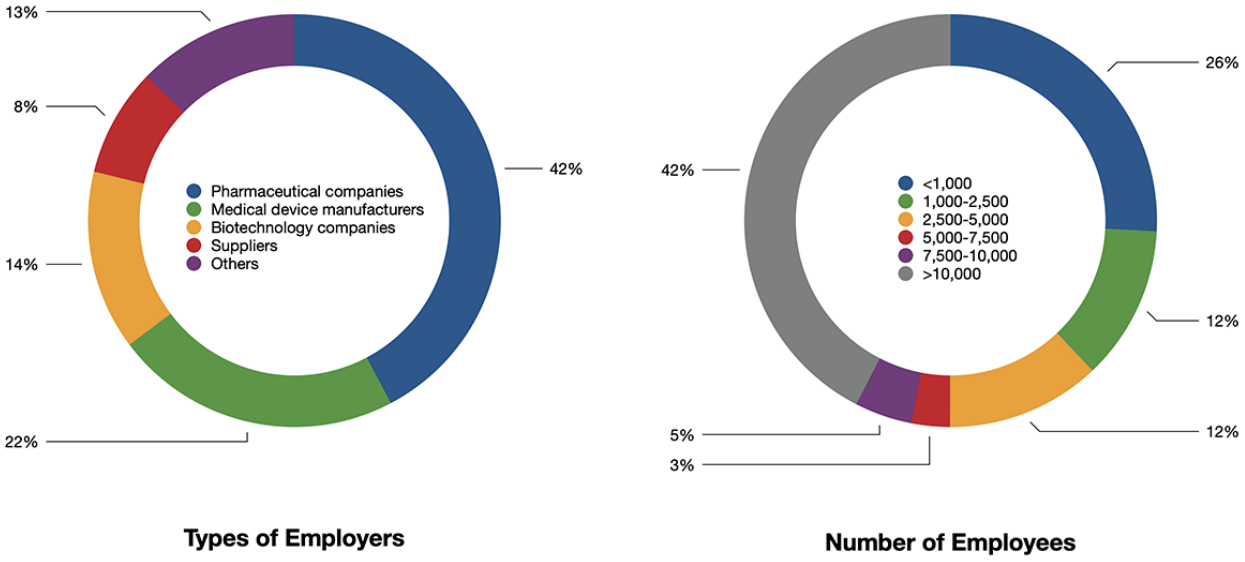


Figure 2. Information about respondents' organizations.



How Organizations Responded to the Impact of COVID-19

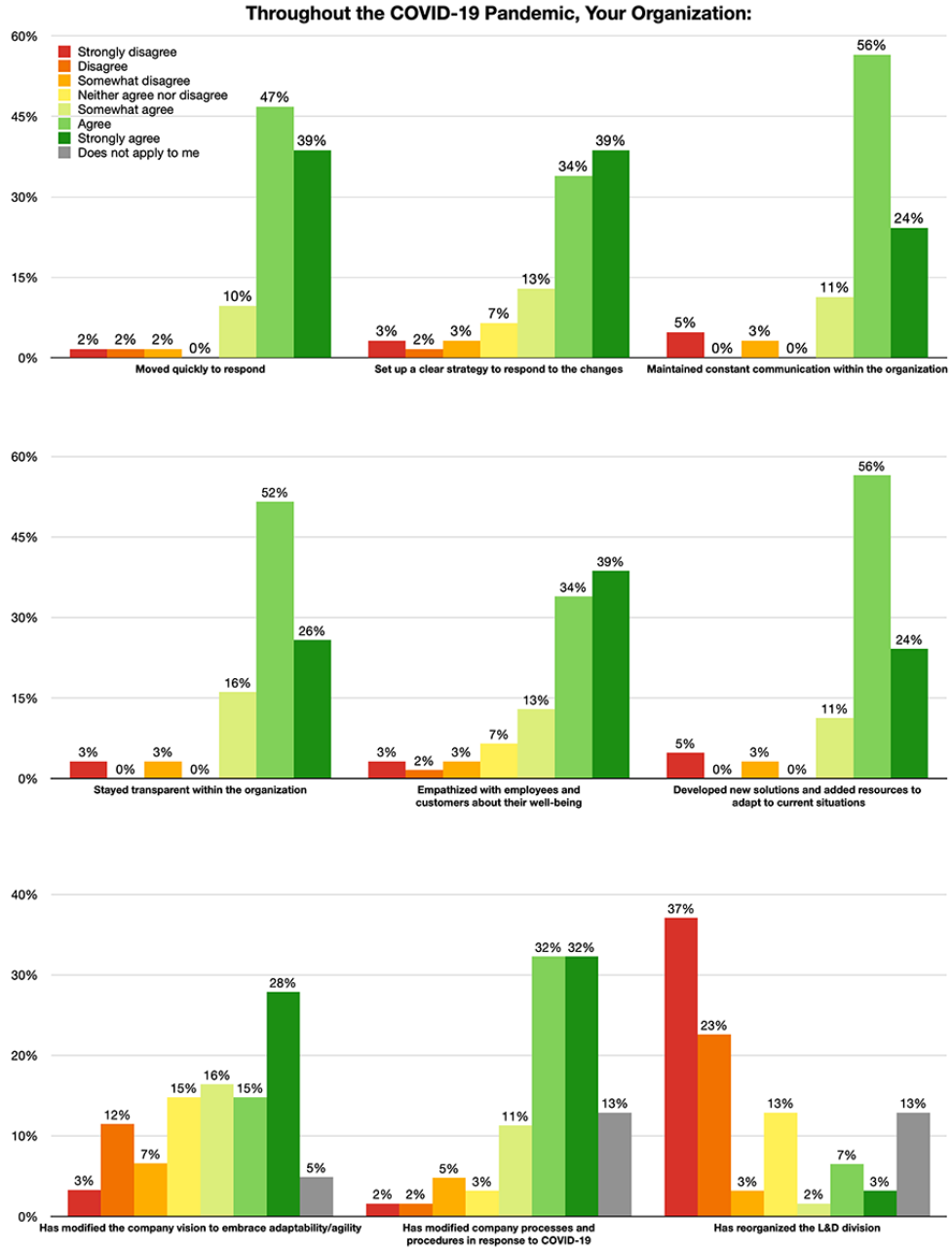
The respondents were asked about their perceptions of their organizations during the work environment change: How quickly the organization leadership reacted to the pandemic, how they responded to emerging problems, and what they did to keep employees doing and feeling well.

Overall Responses

According to the majority of the respondents and interviewees (53-59 [86%-97%] of 62), they were appreciative of the overall response of the leadership during the pandemic (Figure 3). They

agreed in varying degrees that the leadership moved quickly by setting up clear strategies, developing new solutions, modifying company processes and procedures, and adding resources to adapt to current situations. The leadership also remained in constant communication and were transparent within the organizations.

Figure 3. Levels of agreement to organizations' responses to the pandemic. L&D: learning and development.



In other aspects, employees were not as satisfied. In terms of the modification of the company vision and removing barriers to embracing adaptability, these organizations were perceived as less successful by their employees. About 35-36 (58%-59%) of 61 respondents agreed with these items, but only 4 (7%) respondents strongly agreed that their organizations had removed barriers for them and customers. Moreover, fewer respondents (27-30 [44%-49%] of 61) agreed rather than disagreed or were indifferent that their organizations had maintained the pre-COVID-19 work atmosphere, increased decision-making speed, or modified the support system.

In the L&D department of these organizations specifically, 30-47 (56%-87%) of 54 respondents agreed that training- and curriculum-related work had been completely or partially moved online since beginning to work virtually. Such work tasks included, but were not limited to, onboarding processes, knowledge-based training, sales-related skills training, leadership skills training, soft/power skills training, and compliance training.

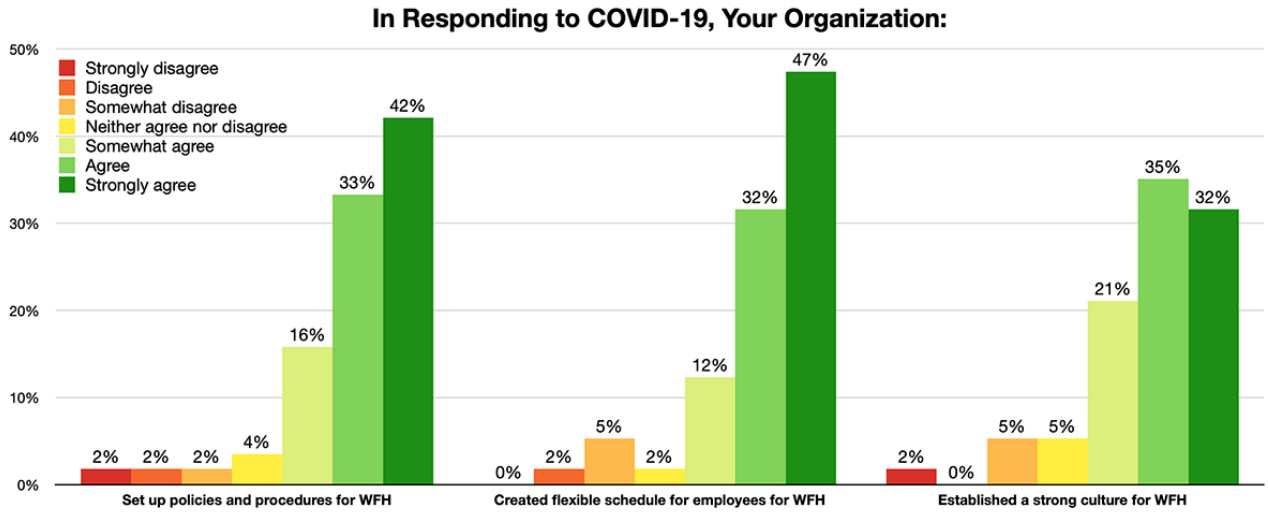
Responses to Employees' and Customers' Needs

It was acknowledged by 47-55 (82%-96%) of 57 respondents that their organizations had created solutions to meet customers' emerging needs by modifying or designing new products and services and leveraging technologies. Organizations had made efforts to allow employees to smoothly transition to the virtual working mode (Figure 4). According to the respondents' ratings,

50-52 (88%-91%) of 57 of them agreed that their organizations had set up policies, created flexible schedules, and established strong cultures for the WFH situation.

Organizations also made efforts to indirectly meet customers' needs by reskilling and upskilling their customer-facing and training-related employees. There were 46-47 (84%-86%) of 55 respondents who had participated in the reskilling and upskilling opportunities related to digital competencies either through virtual microlearning provided by their employers or by locating these upskilling resources on their own. The top 3 ranked uses of technologies perceived to be the most valuable were (1) tools for videoconferencing, (2) engaging customers, and (3) helping employees with information recall. Some interview participants reported that their organizations deployed classes for upskilling, such as virtual selling skills training and onboarding for just-in-time learning. However, a few others reported that they struggled as virtual sales and training limited customers' and trainees' engagement and the performance and training of soft skills for virtual sales were insufficient.

Figure 4. Levels of agreement to organizations' support to remote working. WFH: working from home.



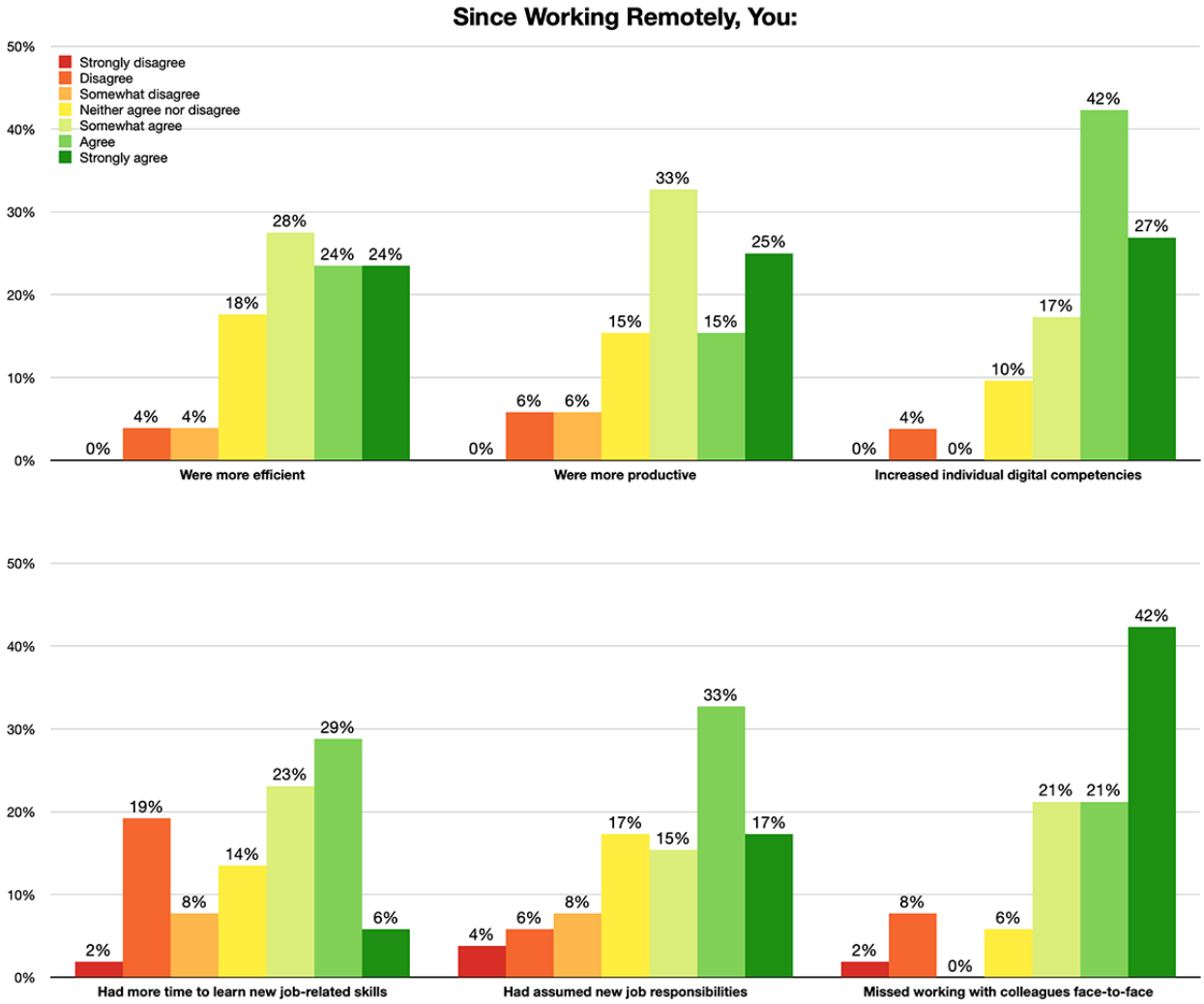
How Individuals Responded to the Changes Caused by the Pandemic

With respect to the individual perspective of professionals in L&D, the respondents were asked about their perceptions of WFH experiences, customer-facing colleagues, and the changes in job responsibilities.

Respondents Treated Remote Working Positively

Although 31 (58%) of 53 respondents perceived WFH as temporary, 40 (70%) of 57 respondents still thought their organizations would consider WFH as a long-term strategy after the pandemic. They had gradually adapted to the new working mode and were mentally prepared for working remotely in the foreseeable future. Working remotely stimulated productivity and promoted the development of digital competencies (Figure 5). Approximately 38-45 (73%-87%) of 52 respondents reflected that they had become more efficient and productive and that they gained digital competencies since WFH began, yet 44 (85%) of them still missed working with colleagues face-to-face. Comparatively fewer respondents (30 [58%] of 52) perceived that they had more time to learn new things related to their jobs. A few participants addressed that they had been working remotely since before COVID-19, so they did not experience a significant change in their jobs.

Figure 5. Levels of agreement to personal working experiences.



The Change in Customer-Facing Employees' Job Requirements

Among all the employees of these surveyed organizations, customer-facing employees experienced the most extreme changes in their day-to-day work. There were 45 (79%) of 57 respondents who perceived that customer-facing employees and field teams had more time to participate in trainings. More than 51 (90%) of the respondents perceived that these colleagues had increased digital selling competency. However, more respondents disagreed (26 [46%] of 57) than agreed (22 [39%] of 57) that digital sales allowed customer-facing employees to engage with more customers than traditional sales before COVID-19.

Moreover, 34 (65%) of 52 respondents realized they had assumed new and more job responsibilities. Several respondents stated in the comment area to explain their choices that demand services had far exceeded the ability to respond and that they could not keep up with the demand at the beginning. One respondent to a survey question (Multimedia Appendix 2, question 22) provided the following response that could explain why this discrepancy occurred:

> *The biggest gap is that clients thought virtual classes were easier to prepare, so the demand went up. However, the resources were limited…[and there was] high expectation from the clients [and their] insecure psychological state.* [Direct quote from a participant]

Perceptions of "The New Normal"

Only 9 (17%) of 53 respondents believed that the field would go back to pre-COVID-19 training modes. The majority (31 [59%] of 53) believed that more organizational goals would be accomplished online even after COVID-19, and 9 (17%) of them believed that the pandemic triggered a paradigm shift and an evolution of this field. In the short-term (3-6 months), 30-33 (56%-62%) of 54 respondents thought that there would be minor changes in company culture, organizational operations, L&D, and recruitment and onboarding processes. There were 32 (59%) of these 54 respondents who thought that major changes would take place in customer-facing interactions. In other words, more respondents chose the minor-changes options than the major-changes options for the 4 aspects of the working environment, although in the long-term (3-5 years), more respondents (22-32 [41%-59%] of 54) tended to think there

would be major changes versus minor changes in organizational operations, L&D, recruitment and onboarding, and customer-facing interactions (see Figure 6). It is worth noticing that although less than 27 (50%) of the respondents chose major changes for new talent recruitment and onboarding, it was still greater than the respondents who chose minor changes. See Multimedia Appendix 2 for the survey questions.

Figure 6. Prediction of the working environment in the short and the long term. L&D: learning and development.

	3-6 Months	3-5 Years
Company Culture	59% minor changes	41% minor changes
Organizational Operation	56% minor changes	59% major changes
L&D	62% minor changes	56% major changes
New Talent Recruitment and Onboarding	56% minor changes	41% major changes
Customer Facing Interactions	59% major changes	59% major changes

How the Defining Characteristics of Learning Agility Were Represented in the Data

The leadership of these organizations reacted to the pandemic by shifting to remote working as soon as possible to maximize the security of their organizations, instead of resisting and waiting for the impact. This reflected their willingness to adapt to the new working conditions following the nationwide lockdown policies. One interviewee noted,

> *The people who provide education to the customers on-site, they immediately have switched to a virtual solution…We did everything we can…We figured out ways to do it right, [even though] it was not the most effective and efficient.* [Direct quote from a participant]

This response is aligned with the generally positive survey responses (see Table 1). On the individual level, participants mentioned that they initially experienced a phase of accepting the situation and then they endeavored to find out solutions. They also mentioned that a supportive team could make people resilient and proactive when faced with challenges.

As the job requirements changed, the work included new tasks and the task complexity increased. To handle jobs with increasing complexity, L&D and sales professionals were required to learn new things continuously, especially at the beginning stage of the transition. More than 65% (34/52) of the survey respondents agreed that they had assumed new job responsibilities since the onset of the pandemic. Organizations put extra emphasis on virtual trainings of various skills, sales, leadership, and soft skills. For L&D professionals, the amount of new training to be designed caused an increase in job responsibilities. Sales professionals also spent more time to participate in trainings, which left them with less time to work on their primary job tasks. For example, 1 (8%) of the 12 interviewees talked about the increasing frequency of their new employee orientations to address emerging issues in a timely manner, so they had to re-create orientation materials to suit shorter sessions. Training professionals reported that they learned new ways to achieve their goals of creating new orientation training. This, therefore, required more time to train professionals to design and deliver new content and for employees from other departments to help with the training. During this process, everyone explored new territories and learned something new, for example, virtual platforms and learning resources. Respondents reported that the most prominent gain in skills was digital competencies. Survey responses show that 45 (87%) of the 52 respondents agreed that they had increased individual digital competencies (Figure 5). In addition, almost every interviewee had referred to the adaptation to a new way of communication, collaboration, and operation through virtual platforms. There were difficulties many of them needed to overcome, since a significant number of senior employees were not familiar with virtual ways of working prior to the pandemic onset. In sum, the data revealed how the characteristics of learning agility were present in the work life of the professionals in our study in response to the pandemic. In Table 1, we identify 4 categories from the data that align with characteristics of learning agility: the willingness to adapt to new job requirements, the ability to handle jobs with increasing complexity, the ability to continuously learn new things, and the ability to overcome difficulties.

Table 1. Characteristics of learning agility and supporting evidence from the survey.[a]

Facet of learning agility	Supporting evidence in the data
Willingness to adapt to job requirements	• Of 57 respondents, 52 (91%) agreed that their organizations set up policies and procedures for WFH[b]. • Of 61 respondents, 27 (44%) agreed that their organizations had increased the speed of decision making. • Of 61 respondents, 56 (91%) agreed that the leadership maintained constant communication within the organization. • Of 61 respondents, 57 (93%) agreed that the leadership stayed transparent within the organization. • Of 61 respondents, 52 (85%) agreed that the leadership empathized with employees and customers about their and their families' well-being.
Ability to learn new things continuously	• Of 62 respondents, 57 (92%) agreed that the leadership developed new solutions and added resources to adapt to current situations. • Of 52 respondents, 45 (87%) agreed that they increased individual digital competencies since working remotely. • Of 55 respondents, 46/47 (84%/85%) agreed that their organization provided reskilling/upskilling for the virtual training world. • Of 56 respondents, 46 (82%) agreed that their organizations designed new products and services to meet clients' current needs. • Of 52 respondents, 30 (58%) agreed that they had more time to learn new things related to their jobs since working remotely. • Of 61 respondents, 30 (49%) agreed that their organizations had modified the supporting system.
Ability to overcome difficulties	• Of 61 respondents, 36 (59%) agreed that their organizations had modified the company vision to embrace adaptability/agility. • Of 62 respondents, 47 (76%) agreed that their organizations had modified company processes and procedures in response to COVID-19. • Of 57 respondents, 52 (91%) agreed that their organization created flexible schedule for employees to WFH. • Of 56 respondents, 41 (73%) agreed that their organization had become more agile. • Of 62 respondents, 36 (58%) agreed that their organizations had removed barriers. • Of 62 respondents, 53 (85%) agreed that the leadership set up a clear strategy to respond to the changes.
Ability to handle jobs with increasing complexity	• Of 56 respondents, 46 (82%) agreed that their organizations designed new products and services to meet clients' current needs. • Of 57 respondents, 45 (79%) agreed that their customer facing employees and field teams had more time to participate in training sessions. • Of 51 respondents, 38 (75%) agreed that they were more efficient since working remotely. • Of 52 respondents, 34 (65%) agreed that they had assumed new job responsibilities since working remotely. • Of 57 respondents, 23 (40%) agreed that their customer facing employees and field teams had exhibited higher levels of productivity.

[a]All the items listed in the table are items whose "agrees" options were selected by more respondents than "disagrees" options, even though some "agrees" responses were lower than 50%. See Multimedia Appendix 2 for survey questions.

[b]WFH: working from home.

Discussion

Principal Results

The interview and survey results indicated that L&D professionals were overall positive in their perceptions of their organization's leadership, their colleagues, and themselves in terms of finding solutions and supporting one another. The majority of respondents reported increased productivity and opportunities to reskill and upskill during WFH. In addition, they upgraded their digital competencies, especially technologies for videoconferencing, engaging customers, and helping learners' information recall. In addition, this study revealed some insights that the framework adapted from Eichinger et al [30] did not discuss. Their framework was developed to evaluate an employee's leadership potential and effectiveness. Our findings suggest that elements of learning agility were demonstrated by most survey participants in response to a highly disruptive crisis, when the working mode abruptly shifted. This study provides insights into how the pandemic created a context that sparked a different entry point (ie, the drive for job survival) into the learning agility development cycle. Results also provide insights into how the COVID-19 crisis demanded professionals' learning agility to both self-initiate their own learning and to support the learning agility of others in the organization.

Limitations and Future Directions

In the quantitative component of the study, 74 participants responded to the survey, which represents ~5% of the sample pool of 1500 people. It is below the average of medium-length web-based surveys (12-25 questions)—less than 10% [31-33]. Given that our survey consisted of 37 questions, it was reasonable to have a lower response rate, as the length of a survey has a negative influence on the response rate [33,34]. One possible direction for future research is to increase the sample size of participants by expanding the survey to a larger membership body [35]. In addition, the participating population of this study was skewed to high management roles and training professionals. Sales professionals, however, are the trainees closely interacting with these training professionals and are as much impacted by the pandemic as training professionals. Therefore, future studies could shift the recruiting focus to

first-line sales professionals to investigate how specifically the overall life cycle of training and education for sale professionals changed.

Another possible direction for future studies is to dig deeper into the learning agility framework by addressing the mental and emotional aspects. To implement this idea, we need to ask more why and how questions about people's motivations and ways of predicting and solving problems, and to intentionally differentiate people's pandemic leadership behaviors from those of the prepandemic period.

Last but not least, the transition to the new way of working that we have observed was reactionary at the core. It would be worthwhile to see whether the L&D professionals would adopt a more proactive approach to learning agility as the pandemic subsides or whether they drift back to the old normal.

Conclusion

Prior research has explored current and future trends in life sciences training [36], noting trends for more remote L&D initiatives. This study examined a more specific circumstance where life sciences L&D professionals faced unprecedented challenges, requiring a sudden and unexpected shift to remote work. This study explored the organizational and individual reactions of life sciences organizations toward the pandemic and the shift in the working environment that it entailed. In retrospect, the first 3 months of the pandemic were a survival phase for most of the organizations in the life sciences sector. The leadership lacked experience and confidence in their solutions during such a turbulent time. The next several months presented an adoption phase, which was marked by a collective will to make change. This phase aligns with the period during which we collected our data and, as such, reflects how learning agility appeared at the organizational and individual levels during this time. Looking toward the future as vaccines are more prevalent, organizations are now more competent to shift into a proactive phase. They have more experience dealing with unpredictable and abrupt changes in the global environment. The changes in work life, due to the COVID-19 crisis, may have pushed the development of the industry 5-7 years forward into the future in a few months [37]. The forces that prompted the development include the organization's drives for financial success, customer demands, and, more prominently, the practitioners' adaptation and agility. They demonstrated a willingness to be agile that may not have emerged as quickly outside of the pandemic crisis. Future work should investigate how life sciences L&D organizations structure their remote work and remote training in a post-COVID-19 world. Both our interview and survey results predicted that a mix of live and virtual solutions for working, training, and learning would emerge to better suit the ever-changing market in life sciences organizations.

Disclaimer

The views and opinions expressed in this publication are those of the authors. They do not purport to reflect the opinions or views of their employers. This research was supported by Life Sciences Trainers & Educators Networks (LTEN). The LTEN is the only global 501(c)(3) nonprofit organization specializing in meeting the needs of life sciences learning professionals.

Authors' Contributions

The interview and survey protocols were set up by NWT, WM, and KP in 2020. The practical part of the study was conducted by NWT and XYP in 2020; WM, SL, and KP fulfilled an advisory role. The manuscript was designed and drafted by XYP, and SL, WM, and NWT contributed to revising the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview questions.
[DOCX File , 16 KB - ijmr_v11i1e33360_app1.docx]

Multimedia Appendix 2
Survey questions.
[DOCX File , 38 KB - ijmr_v11i1e33360_app2.docx]

References

1. Parker K, Horowitz J, Minkin R. How the Coronavirus Outbreak has – and Hasn't – Changed the Way Americans Work. URL: https://www.pewresearch.org/social-trends/2020/12/09/how-the-coronavirus-outbreak-has-and-hasnt-changed-the-way-americans-work [accessed 2021-09-01]
2. Jones SE, Campbell PK, Kimp AJ, Bennell K, Foster NE, Russell T, et al. Evaluation of a novel e-Learning program for physiotherapists to manage knee osteoarthritis via telehealth: qualitative study nested in the PEAK (Physiotherapy Exercise and Physical Activity for Knee Osteoarthritis) randomized controlled trial. J Med Internet Res 2021 Apr 30;23(4):e25872 [FREE Full text] [doi: 10.2196/25872] [Medline: 33929326]

3. Thakur A, Pereira C, Hardy J, Bobbette N, Sockalingam S, Lunsky Y. Virtual education program to support providers caring for people with intellectual and developmental disabilities during the COVID-19 pandemic: rapid development and evaluation study. JMIR Ment Health 2021 Oct 07;8(10):e28933 [FREE Full text] [doi: 10.2196/28933] [Medline: 34617917]
4. Jeyakumar T, Ambata-Villanueva S, McClure S, Henderson C, Wiljer D. Best practices for the implementation and sustainment of virtual health information system training: qualitative study. JMIR Med Educ 2021 Oct 22;7(4):e30613-e30616 [FREE Full text] [doi: 10.2196/30613] [Medline: 34449402]
5. Pilabré AH, Ngangue P, Barro A, Pafadnam Y. An imperative for the national public health school in Burkina Faso to promote the use of information and communication technologies in education during the COVID-19 pandemic: critical analysis. JMIR Med Educ 2021 May 18;7(2):e27169 [FREE Full text] [doi: 10.2196/27169] [Medline: 33970868]
6. Akour I, Alshurideh M, Al Kurdi B, Al Ali A, Salloum S. Using machine learning algorithms to predict people's intention to use mobile learning platforms during the COVID-19 pandemic: machine learning approach. JMIR Med Educ 2021 Feb 04;7(1):e24032 [FREE Full text] [doi: 10.2196/24032] [Medline: 33444154]
7. Kawakami N, Imamura K, Watanabe K, Sekiya Y, Sasaki N, Sato N, SMART-CBT Project Team. Effectiveness of an internet-based machine-guided stress management program based on cognitive behavioral therapy for improving depression among workers: protocol for a randomized controlled trial. JMIR Res Protoc 2021 Sep 29;10(9):e30305 [FREE Full text] [doi: 10.2196/30305] [Medline: 34460414]
8. Balcombe L, De Leo D. Psychological screening and tracking of athletes and digital mental health solutions in a hybrid model of care: mini review. JMIR Form Res 2020 Dec 14;4(12):e22755 [FREE Full text] [doi: 10.2196/22755] [Medline: 33271497]
9. Balcombe L, De Leo D. Digital mental health challenges and the horizon ahead for solutions. JMIR Ment Health 2021 Mar 29;8(3):e26811 [FREE Full text] [doi: 10.2196/26811] [Medline: 33779570]
10. Zhang Y, Lyu H, Liu Y, Zhang X, Wang Y, Luo J. Monitoring depression trends on twitter during the COVID-19 pandemic: observational study. JMIR Infodemiol 2021;1(1):e26769 [FREE Full text] [doi: 10.2196/26769] [Medline: 34458682]
11. Jha IP, Awasthi R, Kumar A, Kumar V, Sethi T. Learning the mental health impact of COVID-19 in the United States with explainable artificial intelligence: observational study. JMIR Ment Health 2021 Apr 20;8(4):e25097 [FREE Full text] [doi: 10.2196/25097] [Medline: 33877051]
12. Lombardo MM, Eichinger RW. High potentials as high learners. Hum Resour Manag 2001;39(4):321-329. [doi: 10.1002/1099-050X(200024)39:4<321::AID-HRM4>3.0.CO;2-1]
13. De Meuse KP, Dai G, Hallenbeck GS. Learning agility: a construct whose time has come. Consult Psychol J: Pract Res 2010 Jun;62(2):119-130. [doi: 10.1037/a0019988]
14. Dries N, Vantilborgh T, Pepermans R. The role of learning agility and career variety in the identification and development of high potential employees. Pers Rev 2012;41(3):340-358. [doi: 10.1108/00483481211212977]
15. Mitchinson AG, Gerard NM, Roloff KS, Burke W. Learning about learning agility. 2012 Jul Presented at: Academy of Management Annual Meeting Proceedings; July 1, 2012; Boston, MA p. 17742. [doi: 10.5465/ambpp.2012.288]
16. Fiedler F. A Theory of Leadership Effectiveness. New York, NY: McGraw-Hill; 1967.
17. House RJ. A path goal theory of leader effectiveness. Adm Sci Q 1971 Sep;16(3):321. [doi: 10.2307/2391905]
18. Vroom V, Yetton P. Leadership and Decision-Making (Vol. 110). Pittsburgh, PA: University of Pittsburgh Press; 2010.
19. Burns J. Leadership. New York, NY: Harper and Row; 1978.
20. De Meuse KP. Using Science to Identify Future Leaders: Part III - the TALENTx7 Assessment of Learning Agility. 2015. URL: https://www.researchgate.net/publication/281640087_Using_Science_to_Identify_Future_Leaders_Part_III_-_The_TALENTx7_Assessment_of_Learning_Agility [accessed 2022-04-14]
21. Burke W, Roloff KS, Mitchinson A. Learning Agility: A New Model and Measure. New York, NY: Teachers College, Columbia University; 2016.
22. Norton LW. Flexible leadership: an integrative perspective. Consult Psychol J: Pract Res 2010 Jun;62(2):143-150. [doi: 10.1037/a0019990]
23. DeRue DS, Myers CG. Leadership development: a review and agenda for future research. In: Day DV, editor. The Oxford Handbook of Leadership and Organizations. Oxford, UK: Oxford University Press; 2014:832-855.
24. Kaiser RB, Craig SB. Do the behaviors related to managerial effectiveness really change with organizational level? An empirical test. Psychol-Manag J 2011 Apr;14(2):92-119. [doi: 10.1080/10887156.2011.570140]
25. Nesbit PL. The role of self-reflection, emotional management of feedback, and self-regulation processes in self-directed leadership development. Hum Resour Dev Rev 2012 Apr 04;11(2):203-226. [doi: 10.1177/1534484312439196]
26. White RP, Shullman SL. Acceptance of uncertainty as an indicator of effective leadership. Consult Psychol J: Pract Res 2010 Jun;62(2):94-104. [doi: 10.1037/a0019991]
27. De Meuse KP, Dai G, Wu J. Leadership skills across organizational levels: a closer examination. Psychol-Manag J 2011 Apr;14(2):120-139. [doi: 10.1080/10887156.2011.570143]
28. Ganguly A, Nilchiani R, Farr JV. Evaluating agility in corporate enterprises. Int J Prod Econ 2009 Apr;118(2):410-423. [doi: 10.1016/j.ijpe.2008.12.009]

29. Creswell J, Clark V. Designing and Conducting Mixed Methods Research. 3rd edition. Thousand Oaks, CA: SAGE Publications; 2017.
30. Eichinger R, Lombardo M, Capretta C. FYI for Learning Agility. Minneapolis, MN: Lominger International: A Korn/Ferry Company; 2010.
31. Anseel F, Lievens F, Schollaert E, Choragwicka B. Response rates in organizational science, 1995–2008: a meta-analytic review and guidelines for survey researchers. J Bus Psychol 2010 Feb 9;25(3):335-349. [doi: 10.1007/s10869-010-9157-6]
32. Rogelberg SG, Stanton JM. Introduction: understanding and dealing with organizational survey nonresponse. Org Res Methods 2016 Nov 08;10(2):195-209. [doi: 10.1177/1094428106294693]
33. PeoplePulse. Response Rates: Tips on How to Increase Your Survey Response Rates. URL: https://peoplepulse.com/resources/useful-articles/survey-response-rates [accessed 2021-09-02]
34. Qualtrics. How to Increase Online Survey Response Rates. URL: https://www.qualtrics.com/experience-management/research/tools-increase-response-rate [accessed 2021-09-02]
35. Boddy CR. Sample size for qualitative research. QMR 2016 Sep 12;19(4):426-432. [doi: 10.1108/qmr-06-2016-0053]
36. Magagna W, Wang N, Peck K. Current and future trends in life sciences training: questionnaire study. JMIR Med Educ 2020 Apr 24;6(1):e15877 [FREE Full text] [doi: 10.2196/15877] [Medline: 32329739]
37. How COVID-19 Has Pushed Companies Over the Technology Tipping Point And Transformed Business Forever. URL: https://tinyurl.com/bdfxjt2u [accessed 2021-08-29]

Abbreviations

L&D: learning and development
LTEN: Life Sciences Trainers & Educators Networks
WFH: working from home

Edited by T Leung; submitted 09.09.21; peer-reviewed by A Joseph, M Reynolds; comments to author 07.11.21; revised version received 29.11.21; accepted 18.03.22; published 26.04.22.

Please cite as:
Peng X, Wang-Trexler N, Magagna W, Land S, Peck K
Learning Agility of Learning and Development Professionals in the Life Sciences Field During the COVID-19 Pandemic: Empirical Study
Interact J Med Res 2022;11(1):e33360
URL: https://www.i-jmr.org/2022/1/e33360
doi:10.2196/33360
PMID:35417403

Original Paper

Statistical Methods for Item Reduction in a Representative Lifestyle Questionnaire: Pilot Questionnaire Study

Alessio Staffini[1,2,3], MSc; Kento Fujita[4], MSc; Akiko Kishi Svensson[1,5,6], MD, MPH, PhD; Ung-Il Chung[1,7,8], MD, PhD; Thomas Svensson[1,5,7], MD, PhD

[1]Precision Health, Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

[2]Department of Economics and Finance, Catholic University of Milan, Milan, Italy

[3]Data Solutions Division, Project Promotion Department, Albert Inc, Tokyo, Japan

[4]Data Service Infrastructure Development Department, Service Infrastructure Division, IT-OT Innovation Division, Mobile Technology Unit, SoftBank Corp, Tokyo, Japan

[5]Department of Clinical Sciences, Lund University, Malmö, Sweden

[6]Department of Diabetes and Metabolic Diseases, The University of Tokyo, Tokyo, Japan

[7]School of Health Innovation, Kanagawa University of Human Services, Kawasaki, Japan

[8]Clinical Biotechnology, Center for Disease Biology and Integrative Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Akiko Kishi Svensson, MD, MPH, PhD
Precision Health, Department of Bioengineering
Graduate School of Engineering
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo, 113-8656
Japan
Phone: 81 358414737
Email: kishi@bioeng.t.u-tokyo.ac.jp

Abstract

Background: Reducing the number of items in a questionnaire while maintaining relevant information is important as it is associated with advantages such as higher respondent engagement and reduced response error. However, in health care, after the original design, an *a posteriori* check of the included items in a questionnaire is often overlooked or considered to be of minor importance. When conducted, this is often based on a single selected method. We argue that before finalizing any lifestyle questionnaire, *a posteriori* validation should always be conducted using multiple approaches to ensure the robustness of the results.

Objective: The objectives of this study are to compare the results of two statistical methods for item reduction (variance inflation factor [VIF] and factor analysis [FA]) in a lifestyle questionnaire constructed by combining items from different sources and analyze the different results obtained from the 2 methods and the conclusions that can be made about the original items.

Methods: Data were collected from 79 participants (heterogeneous in age and sex) with a high risk of metabolic syndrome working in a financial company based in Tokyo. The lifestyle questionnaire was constructed by combining items (asked with daily, weekly, and monthly frequency) from multiple validated questionnaires and other selected questions. Item reduction was conducted using VIF and exploratory FA. Adequacy tests were used to check the data distribution and sampling adequacy.

Results: Among the daily and weekly questions, both VIF and FA identified redundancies in sleep-related items. Among the monthly questions, both approaches identified redundancies in stress-related items. However, the number of items suggested for reduction often differed: VIF suggested larger reductions than FA for daily questions but fewer reductions for weekly questions. Adequacy tests always confirmed that the structural detection was adequate for the considered items.

Conclusions: As expected, our analyses showed that VIF and FA produced both similar and different findings, suggesting that questionnaire designers should consider using multiple methods for item reduction. Our findings using both methods indicate that many questions, especially those related to sleep, are redundant, indicating that the considered lifestyle questionnaire can be shortened.

(Interact J Med Res 2022;11(1):e28692) doi:10.2196/28692

KEYWORDS

item reduction; surveys and lifestyle questionnaires; feedback measures; questionnaire design; variance inflation factor; factor analysis; mobile phone

Introduction

Background

Short questionnaires have several advantages over long ones: they are less expensive to design [1]; are associated with less random or systematic error or noise in the reported results caused by lack of motivation, fatigue, or boredom [2-4]; and have higher respondent engagement, with empirical evidence indicating that long questionnaires are associated with a low response rate [5]. Therefore, reducing the number of survey items is important.

A number of approaches to reduce survey items have been studied. First, in psychometrics, the Cronbach α coefficient is often used in Classical Test Theory [6] to assess the internal consistency of a given questionnaire. Second, factor analysis (FA) and principal component analysis were used to extract a latent structure and explain most of the data variance with fewer factors or components than the original questionnaire items (eg, the studies by McHorney et al [7], Bai et al [8], and Brosnan et al [9]). Third, researchers have also focused on Item Response Theory (IRT) [10,11]; however, IRT methods are most effective when the original questionnaire or survey is developed using IRT or when there are theoretical reasons to expect it to fit an IRT model [12]. Recent studies have used the variance inflation factor (VIF) to better address collinearity problems among covariates when performing regression analysis of survey data [13]. However, despite the advantages and practicality of automatic variable selection, simulations suggest that the VIF does not necessarily identify a true underlying model [14]. In general, it is better not to rely on a single method for variable reduction and instead to compare the results of multiple approaches.

Although it is important to minimize errors in questionnaire responses in all fields, it is paramount in medicine and public health, where omissions and inaccuracies can lead to possible misdiagnoses and subsequent incorrect treatments or interventions. Despite this, emphasis in health care is often placed only on the initial design, where a questionnaire is typically evaluated by comparing its internal consistency with that of other similar questionnaires or with a different version of the same questionnaire [15-17]. The questionnaires also sometimes include open-ended questions, despite evidence showing that such questions often do not provide sufficiently solid insights [18]. Although an *a priori* carefully designed questionnaire is extremely important, we believe that including an *a posteriori* check of the questions before making any result-based inferences will help improve the overall quality of the instrument and should, therefore, become part of standard procedure.

Objective

Our study builds upon other research conducted in this direction: to name a few related works, Cappelleri et al [19] applied FA to develop a questionnaire to measure the satisfaction of patients with type 1 diabetes, highlighting that the two key factors were convenience and social comfort; Juniper et al [20] compared the results of the impact method, which preserves items according to their relative importance as perceived by the patients, with the results of FA in a quality of life questionnaire, suggesting that different approaches lead to different results; Arifin and Yusoff [21] applied confirmatory FA, as well as performed other statistical tests, for an emotional intelligence inventory to be used among medical course applicants. In the Classical Test Theory branch, researchers often compared Cronbach α with Rasch analysis; for example, Prieto et al [22] did so on a 38-item health questionnaire, concluding that the methods led to similar results; Erhart et al [23] analyzed the possibility of performing item reduction by comparing the results of Cronbach α with those of Rasch item fit for a health-related quality of life questionnaire administered to children and adolescents, concluding, on the other hand, that both methods should be accompanied by additional analyses.

We found that very often, researchers focused on a single technique or similar approaches without comparing the results from multiple different methods. This was the starting point of our study. Despite its desirable properties (such as automatic identification and remotion of multicollinear items) and statistical foundations, VIF is rarely directly used as an item reduction technique. FA indirectly deals with the same multicollinearity problem by grouping *similar* variables under the same factor; however, its inner logic is different, as multicollinearity is solved because of dimensionality reduction. To the best of our knowledge, this is the first study to directly compare how similar or how different the results obtained from these methods are when applied to the items of a lifestyle questionnaire.

The aim of this study is to compare the results of VIF and FA for item reduction in a lifestyle questionnaire constructed by combining items from different sources, with questions asked with daily, weekly, or monthly frequency.

Methods

Study Participants

The collected questionnaire data (described in detail in the following sections) form part of a 1-month pilot study for lifestyle interventions, with a more general aim than the one described in this paper; that is, in particular, to investigate the feasibility of using a smartphone lifestyle intervention app and assess the feasibility of collecting data from a wearable device. The recruited participants were individuals with a high risk of metabolic syndrome. The sample size was calculated based on

accepting an α risk of .05 and a β risk of .25 in a 1-tailed test, with an estimated 10% loss to follow-up. The estimated sample size was 120 participants. A total of 117 individuals consented to participate; however, 2 (1.7%) individuals were withdrawn, leaving a total of 115 (98.3%) participants. The 115 participants were randomly allocated into 2 groups: the intervention group (79/115, 68.7% individuals who followed a lifestyle education program over the study period) and the control group (36/115, 31.3% individuals). Allocation to the intervention and control groups occurred through randomization, with stratification by sex and age (<40 years or >40 years). Participants in the intervention group (the focus of this study), men and women aged between 29 and 58 years, were selected from a financial company based in Tokyo. To minimize sampling bias in the collected questionnaire data, we confirmed that the intervention group was sufficiently heterogeneous in terms of age, sex, and employment conditions. The initial study excluded those who had a history of serious medical conditions, had received any other lifestyle intervention, planned to take long vacations, had night shifts, and were pregnant (or those with suspected pregnancy). The study period of the research presented in this paper was from March 2 to March 30, 2018.

All research was performed in accordance with relevant ethical guidelines and regulations. All the participants received detailed information about the purpose of the study in writing and during explanatory face-to-face meetings. All participants provided written informed consent and understood that participation was completely voluntary and could be discontinued at any time without any disadvantage or penalty. Participants were given a wearable device as an incentive for participation.

Questionnaire

The intervention group participants were asked to respond to lifestyle-related questions through a smartphone-based mobile app for a month. For the purposes of this study, we included participants who answered any of the questions.

We constructed a lifestyle questionnaire comprising 51 questions, focusing in particular on the domains of sleep, stress, nutrition, and alcohol and tobacco consumption. For brevity, the questionnaire items are indicated by *item X*, where *X* represents the item number. The questionnaire is provided in Multimedia Appendix 1.

Daily questions on sleep quality were selected from the validated St Mary's Hospital Sleep Questionnaire (items 2-5) [24]. Daily questions to assess dietary habits were taken from the Dietary Guidelines provided by the Ministry of Agriculture, Forestry, and Fisheries of Japan (Food Safety and Consumer Affairs Bureau; items 6-16). Daily questions about caffeine intake (items 17 and 18) and stress levels (items 20 and 21) have been widely used in many population-based cohort studies. In addition, we added 2 other daily questions, not taken from any validated questionnaire, about opening the smartphone-based app in the morning (item 1) and alcohol consumption (item 19). Most of the weekly questions (21/26, 81%) were taken from the validated General Sleep Disturbance Scale (items 25-45) [25]. We added 5 other questions to assess the participants' commitment to reducing alcohol intake (item 47), determine the number of cigarettes smoked per day over the past week and interest in a smoking cessation program (items 23 and 46), and information about working conditions (items 22 and 24). Finally, monthly questions were taken from the validated Perceived Stress Scale–4 (items 48-51) [26]. As shown in Multimedia Appendix 1, items 48 and 51 are reverse-coded items; this is sometimes recommended to cross-check the validity of the responses.

When a participant did not complete a daily questionnaire item, the data were treated as missing and excluded from the analysis. However, when a participant did not complete a weekly or monthly questionnaire item, we assumed that their subsequent response to the item was applicable for the entire respective week or month and, thus, backfilled missing values.

Statistical Analyses

Overview

All statistical analyses were performed using Python (version 3.7.4) on the Anaconda platform (Anaconda Inc). To perform effective item reduction, it is important to verify that the variables we want to exclude, in fact, explain the same underlying variability as the other variables remaining in the questionnaire. We considered two statistical approaches: VIF and FA.

VIF Analysis

VIF is the quotient of the variance in a model with multiple variables and the variance in a model with only 1 variable. It indicates the strength of multicollinearity among a set of variables, assuming that they have a linear relationship. Each variable, in turn, is regressed on all other variables present in the set. Considering a set A of n variables, the VIF associated with variable $i \in A$ is defined as follows:

Here, R_i^2 is the coefficient of determination obtained by ordinary least squares regression and regressing variable i on all the other $j \neq i$ variables in the set. A higher VIF_i reportss greater collinearity between variable i and the other predictors. As $0 \leq R_i^2 \leq 1$, $VIF_i \in [1, \infty)$.

There is no general consensus on the ideal VIF threshold for indicating multicollinearity, with many different values having been used in the empirical literature. For example, Vittinghoff et al [27] suggested a threshold value of 10, whereas Johnston et al [28] were more conservative and used a threshold value of 2.5. Keeping in mind that multicollinearity is a bigger problem with a small sample size [29], we selected a threshold value of 5; this cutoff is often selected to establish a high *risk* of multicollinearity [30,31].

We applied VIF iteratively: first, we calculated VIF_i for each variable in set A; then, we removed the variable with the highest calculated VIF value (as it is already well-explained by the remaining variables) and recalculated VIF for each remaining variable. The process was stopped when all remaining variables had a calculated VIF ≤5 so that there was no concern of high collinearity among the variables.

FA Method

FA is a statistical method used to describe variability among observed and correlated variables in terms of a fewer number of unobserved and underlying variables called factors. This method aims to identify the independent factors that explain the different sources of variability in the original variables. It assumes that any observed variable is directly associated with an underlying factor. Informally, FA shares many similarities with principal component analysis: they are both data analysis techniques, their goal being to reduce a large number of variables into a fewer number of more treatable and interpretable new variables, trying to minimize the information loss at the same time. This dimensionality reduction is performed by projecting the original data onto a lower-dimensional space, preserving as much variability as possible. More formally, consider a set A of n observable variables for which we assume that each variable x_i can be expressed as a linear combination of $k<n$ factors and intercept β_i:

$$x_i = \beta_i + l_{i1}F_1 + l_{i2}F_2 + ... + l_{ik}F_k + \varepsilon_i$$

where $\varepsilon_i \sim N(0, \sigma^2)$. We can rewrite the above in matrix notation:

$$X = \beta + LF + \varepsilon,$$

where $X=(x_1,\ldots,x_n)^T$, $\beta=(\beta_1,\ldots,\beta_n)^T$, and denote the factor-loading matrix, $F=(F_1,\ldots,F_k)^T$ is the vector of common factors, and $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_n)^T$ is the vector of unobserved error terms. The factor-loading matrix L expresses the relationship of each observed variable with the unobserved factors, showing the variance explained by each observed variable for each factor. If we consider $X\in R^n$ to be the vector representing all questionnaire items, we can replace it with the vector $F\in R^k$ of unobserved factors, which lies in a lower-dimensional space. We can think of each factor as being capable of explaining a certain variance in the original items, and we can further exclude factors with the lowest amount of explained variance (the criterion for selecting the number of factors is explained in the following sections).

We looked for joint variations in the observed variables in response to unobserved latent factors. Three assumptions should be satisfied: there should be no outliers in the data, the sample size should be *big enough*, and there should be no perfect multicollinearity between the observed variables [32]. Our data set satisfied all of these assumptions.

We further applied two adequacy tests before proceeding with FA: the Bartlett test of sphericity [33] and the Kaiser–Meyer–Olkin (KMO) test [34]. The Bartlett test of sphericity checks whether the observed variables are effectively correlated with each other by comparing their correlation matrix *corr(X)* against the identity matrix *I*. If the null hypothesis (*corr(X)=I*) cannot be rejected, this indicates that the original variables are orthogonal (hence, unsuitable for structure detection), and any data reduction technique (such as FA) would not produce any meaningful result. The KMO test is another way of measuring the suitability of data for FA; it estimates the proportion of variance that may be common (ie, caused by the same underlying factor) among all observed variables, with a lower proportion being a more suitable condition for FA. KMO values range between 0 and 1 (according to the original paper by Kaiser [34], values of a statistic <0.5 mean that performing FA is not adequate). Together, the 2 adequacy tests check the data distribution and sampling adequacy.

Finally, we used the Kaiser criterion [35] to select an adequate number of factors. The eigenvalues λ_i of the correlation matrix *corr(X)* can be used to measure the degree to which the factors explain the variance in the observed variables. Thus, any factor with an associated eigenvalue $\lambda_i>1$ explains more of the variance than a single (observed) variable. As a selection criterion, the number of factors was chosen such that it was equal to the number of eigenvalues >1. However, the Kaiser criterion has been criticized for being an arbitrary approach [36]. Thus, we further examined the scree plot of the eigenvalues and checked whether our stopping criterion matched the possible points of inflection. Finally, we used ordinary least squares to identify the minimum residual solution, which is also the default method for exploratory FA, when estimating the factor loadings.

Ethical Approval

This study was approved by the research ethics committee of the Faculty of Medicine, University of Tokyo (application number 11781).

Results

For meaningful comparisons, we grouped the questions according to the frequency at which they were asked (daily, weekly, and monthly) and analyzed each group separately.

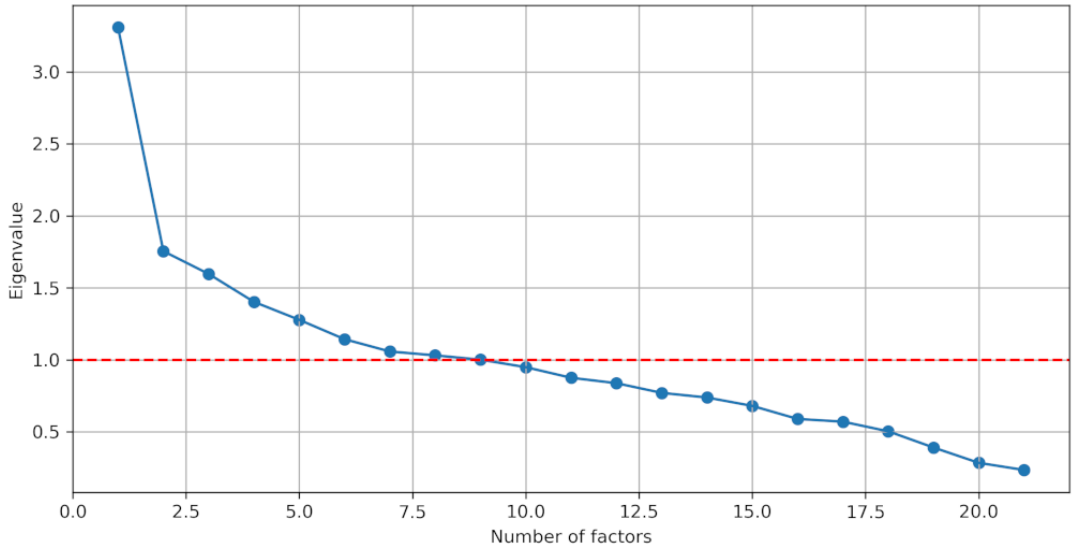
Daily Questions

A total of 21 questions were asked daily to 79 users for a total of 1746 collected data points. After excluding missing values, of the 1746 collected data points, 1658 (95%) data points were left for analysis.

As described in the *Methods* section, we iteratively excluded all the variables for which $VIF_i>5$. By doing so, the 21 initial variables were reduced to 3 (items 4, 20, and 21).

When FA was applied to the daily questions set, we obtained a *P* value of 0 using the Bartlett test; thus, we rejected the null hypothesis *corr(X)=I* at all significance levels, confirming that the considered variables were correlated with each other. The KMO test yielded a value of 0.73, further confirming that the structure detection was adequate for the considered variables. The Kaiser criterion led to the selection of 9 factors (with the first factor having a significantly higher eigenvalue), as shown in the eigenvalue scree plot in Figure 1.

Figure 1. Eigenvalues scree plot for daily questions.



By analyzing the factor-loading matrix *L* of the 9 factors for daily questions, we observed the following: (1) factor 1 had high factor loadings for items 2, 3, 4, and 5 (sleep-related questions); (2) factor 2 had high factor loadings for item 20 (stress level); (3) factor 3 had high factor loadings for item 6 (whether eating 3 times a day); (4) factor 4 had high factor loadings for items 16, 17, and 18 (related to caffeinated drinks); and (5) factor 5 had high factor loadings for items 9 and 12 (having lunch in a restaurant or having a takeout lunch). The other factors were less interesting (eg, as they were each associated with 1 or 2 variables only and displayed factor loadings smaller in magnitude than those of the 5 abovementioned factors) or were more difficult to interpret, as the amount of variance they explained was relatively low and spread across multiple variables. The factor-loading matrix for daily questions is provided in Multimedia Appendix 2.

When the sample size is small (typically <300, as in our analysis), it is also worth looking at the average communality of the retained items [37]. Using 9 underlying factors, we obtained an average communality of 0.492, which is an acceptable value when using Promax rotation [38], as we did for our analysis.

For the main 5 abovementioned factors, we also conducted a reliability analysis using Cronbach α to measure the internal consistency of our underrepresentation. The α coefficient for factor 1 was high (.895), indicating high internal consistency for the sleep-related items, confirming the reliability of the factor. Factors 2 and 3 had high factor loadings for a single variable; thus, we could not check any interitem internal consistency. Factors 4 and 5 displayed lower α coefficients (.470 and .536, respectively), which might be a consequence of the low number of represented original items (3 and 2, respectively) than factor 1 (4), as well as a reduced scale with respect to the sleep-related questions.

The 9 underlying factors (instead of the 21 observed variables) explained 49.25% of the total variance.

Weekly Questions

A total of 26 questions were asked weekly to 79 users. The VIF approach reduced the initial 26 questions to 21 (excluding items 28, 30, 35, 46, and 47).

When applying FA to the weekly question set, we again obtained a *P* value of 0 using the Bartlett test, and the KMO test produced a value of 0.69. The Kaiser criterion led to the selection of 9 factors. Similar to the eigenvalues for the daily questions, the first factor exhibited a significantly higher eigenvalue (Figure 2).

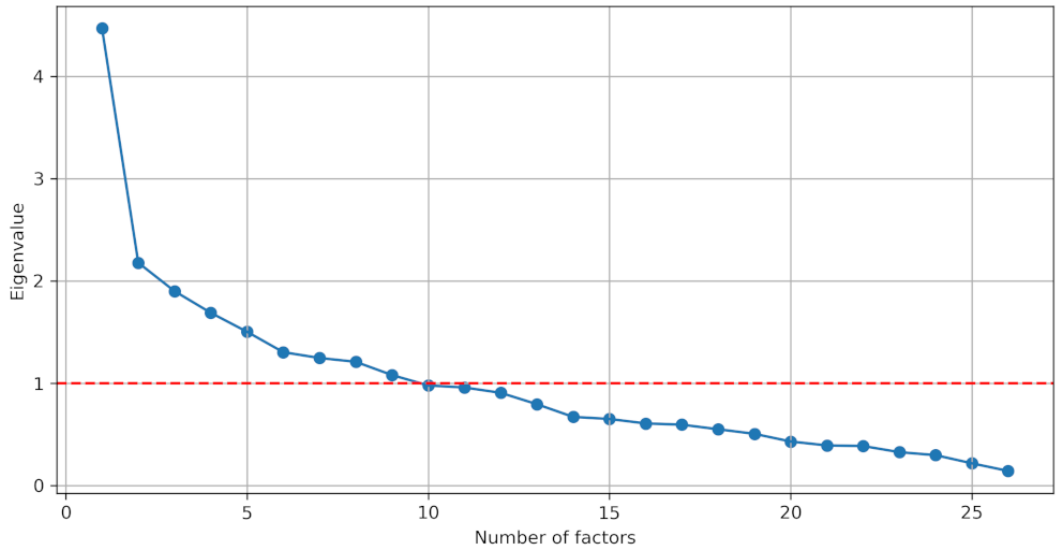
By analyzing the factor-loading matrix *L* of the 9 factors for weekly questions, we observed the following: (1) factor 1 had high factor loadings for items 26, 30, 31, and 39 (sleep-related questions); (2) factor 2 had high factor loadings for items 28, 34, and 35 (questions related to satisfaction and feeling well); (3) factor 3 had high factor loadings for items 32 and 33 (feeling annoyed and feeling tired during the day); (4) factor 4 had high factor loadings for items 22 and 24 (weekly number of days off and weekly hours of work); and (5) factor 5 had high factor loadings for items 23 and 46 (amount of tobacco consumption and degree of interest in a smoking cessation program). The other factors were less interesting (eg, as they were each associated with 1 or 2 variables only and displayed factor loadings smaller in magnitude than those of the 5 abovementioned factors) or were more difficult to interpret, as the amount of variance they explained was relatively low and spread across multiple variables. The factor-loading matrix for weekly questions is provided in Multimedia Appendix 3.

Using 9 underlying factors, we obtained an average communality of 0.493, which is acceptable.

For the main 5 abovementioned factors, the results of the reliability analysis indicated high or acceptable Cronbach α coefficients for all considered factors (.757, .821, .721, .679, and .609), further confirming the internal consistency of the considered underrepresentation.

The 9 underlying factors (instead of the 26 observed variables) explained 49.29% of the total variance.

Figure 2. Eigenvalues scree plot for weekly questions.



Monthly Questions

A total of 4 questions were asked monthly to the 79 users. The VIF approach reduced the initial 4 questions to 2 (leaving items 50 and 51).

When applying FA to the monthly question set, we again obtained a *P* value of 0 using the Bartlett test, and the KMO test produced a value of 0.64. The Kaiser criterion led to the selection of a single factor (Figure 3).

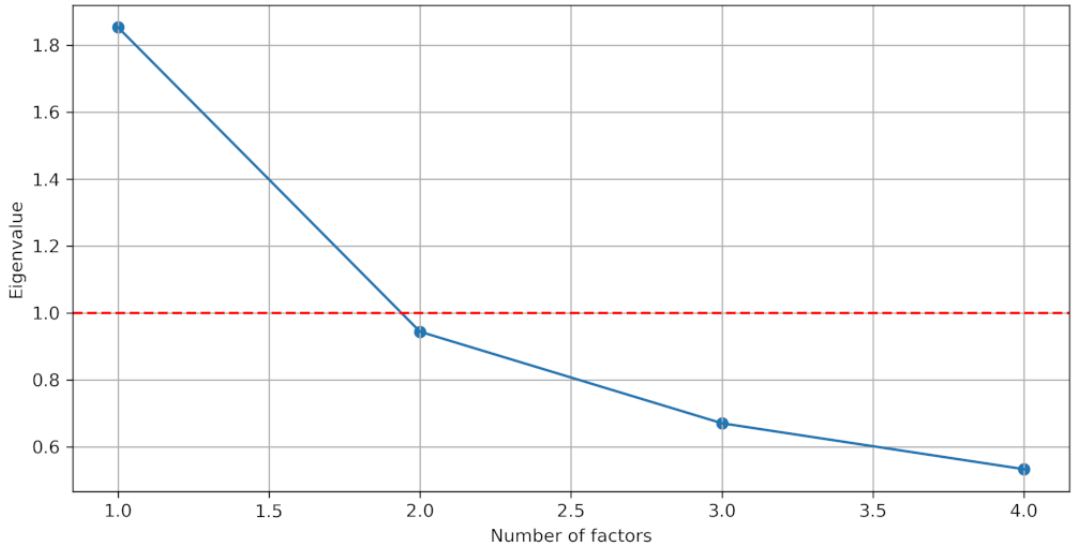
By analyzing the factor-loading vector *L* of the factor for monthly questions, we observed that factor 1 had relatively high factor loadings for all 4 monthly questions, in particular items 50 and 51 (*degree of feeling on top of things* and *degree of not being able to cope with all the things that needed to be done*, respectively). The factor-loading vector for monthly questions is provided in Multimedia Appendix 4.

Using a single underlying factor, we obtained an average communality of 0.308, which is too low. However, as communalities are the proportion of each variable's variance that can be explained by the factors (in this case, a single factor), it makes sense that their average matches the cumulative variance explained by the single factor itself.

For the single considered factor, the results of the reliability analysis indicated an acceptable Cronbach α coefficient (.607), further confirming the internal consistency of the considered underrepresentation.

The single underlying factor (instead of the 4 observed variables) explained 30.76% of the total variance.

Figure 3. Eigenvalues scree plot for monthly questions.



Discussion

Principal Findings

The objective of this study was to compare 2 different statistical methods, which are often used as reduction techniques, and their results in a lifestyle questionnaire, which was constructed using a variety of questions (asked with daily, weekly, and monthly frequency) aimed at evaluating the general well-being of an individual. Our main findings suggest not only that existing validated lifestyle questionnaires might benefit from further

item reduction (in the questions about sleep quality and satisfaction, in particular) but also that different algorithms lead to different results for what concerns other groups of items.

Indeed, the results obtained using the two considered methods (VIF and FA) shared some similarities but also exhibited substantial differences.

Among daily questions, VIF led to the exclusion of many more variables than FA, retaining only 1 question about sleep and 2 questions aimed at evaluating stress levels at work and at home, respectively. Although we expected the sleep items to be highly correlated, the iterative approach of deleting correlated variables also led to the exclusion of all items on eating habits, suggesting that, at least in our data sample, these variables shared some correlation with sleep satisfaction and stress level components. The two types of stress do not seem to be well correlated, suggesting that stress at home does not depend on the amount of stress felt at work and vice versa. FA suggested the inclusion of only 9 factors instead of the original 21 variables; however, in line with the VIF, the factor with the highest eigenvalue captured the variance of the sleep-related items.

Among the weekly questions, the VIF reduced the initial 26 weekly questions to 21 questions. Among the 5 deleted questions, 3 (60%) were sleep-related, and it made sense that they were correlated with the other weekly sleep items. However, interestingly, there was also a correlation between desiring a healthier lifestyle (identified as an interest in quitting smoking and consideration for reducing alcohol intake) and work habits or sleep satisfaction level. This is in line with findings from Hidaka et al [39], who observed a positive correlation between low sleep satisfaction and unhealthy lifestyle patterns in the Japanese population. FA produced similar results, identifying 1 factor in particular for sleep-related questions and another factor for general well-being.

Finally, the monthly questions selected from the Perceived Stress Scale–4 aimed to assess different aspects of the consequences of stress. For example, stress is known to negatively affect self-confidence [40] and, in the long-run, also mental and physical health [41]. Therefore, we expected that the analyses would lead to the exclusion of a significant number of these items. Indeed, the VIF reduced the initial 4 questions to 2, and the FA identified a single significant factor.

We observed that despite the fact that both methods deal with the same underlying problem of multicollinearity, VIF led to greater item reduction in some instances, whereas FA did so in others. Thus, we suggest that questionnaire designers use both methods and, in the event of a discrepancy in results, adopt other additional measures such as comparing both results with the consistency of the internal questionnaire obtained using Cronbach α for the final selection.

As previously mentioned, our questionnaire included 2 reverse-coded questions. In our methods, we analyzed the absolute value of the correlation coefficients to eliminate the need for any additional operations to identify the direction of the correlation (positive or negative).

The principal result of our study shows that even in validated lifestyle questionnaires, many items (particularly sleep-related ones) are indeed redundant. Therefore, when aiming for short questionnaires, we suggest that questionnaire designers should always consider the application of item reduction instruments after a trial phase, as certain items could, in principle, be deleted without incurring significant information loss.

Our study had some limitations that warrant mention. First, because of factor loadings, FA can also be used to analyze the amount of information that is lost when switching from the original variables to the underlying factors. In general, it should be noted that the exclusive use of statistical methods to shorten questionnaires can lead to the loss of valuable information [42]. Indeed, this was observed in our results; the total variance explained exclusively by the identified factors never exceeded 50%. Therefore, rather than providing a recipe that is indiscriminately valid, our approach was to focus on identifying areas in which there is a high possibility of reducing questionnaire items with as little information loss as possible. Our empirical results show that sleep-related questions are, by far, the area where such a reduction seems the most feasible. Second, as this was a pilot study, the number of participants was relatively small. This is the major limitation of our study, as a bigger data set would provide more statistically sound results. However, despite its small size, the size of our data set was considered acceptable for the analyses we conducted [43,44]. Furthermore, despite differences in age and sex, the participants were all selected from the same company, which could have introduced some selection bias. Thus, the results obtained in this study should be validated in further studies.

Conclusions

We constructed a lifestyle questionnaire by combining items from various authoritative sources. We then applied two different statistical methods for item reduction (VIF and FA) to check whether the existing items in the three groups of questions (asked with daily, weekly, and monthly frequency) were redundant. The results of the applied methods did not always match but nevertheless provided evidence that many items related to sleep, in particular, were indeed redundant. Two reduced questionnaires (according to VIF and FA) are proposed in Multimedia Appendix 5. We also conducted reliability analyses for each group of questions using Cronbach α to measure the consistency of the obtained underrepresentation, obtaining satisfactory results. Our results suggest that questionnaire designers should always conduct a *trial* phase on a sample of participants, and examine the correlation between the items, before finalizing any lifestyle questionnaire.

Acknowledgments

This research was supported by the Center of Innovation Program of the Japan Science and Technology Agency (grant number JPMJCE1304). The funders had no role in the design of the study; collection, analysis, and interpretation of data; writing of the report; or decision to submit the paper for publication. The authors would like to thank all staff members at the Center of Innovation, University of Tokyo, for their extensive efforts and help in conducting the study. The authors would also like to thank all members of Precision Health, The University of Tokyo, for their invaluable assistance and, in particular, Dr Masahiro Nakamura for his help with the preparation of the data set. The authors wish to thank the Nomura Securities Health Insurance Association for recruiting participants and collecting data. The authors also wish to thank Shinya Uezono (SoftBank Corp), Shunya Arita (SoftBank Corp), Toshiki Murata (SoftBank Corp), Saeru Yamamuro (SoftBank Corp), Tomoki Kinoue (SoftBank Corp), Tomihisa Ikeura (The University of Tokyo, Center of Innovation Program), and Yoshifumi Matsui (Albert Inc) for their support in this project. The authors would like to thank Tatsunori Seki (SoftBank Corp) and Yoshiya Shigeeda (SoftBank Corp) for their insightful comments and suggestions.

Authors' Contributions

TS, AKS, and AS designed the main concepts of this study. AS and KF analyzed and interpreted the data. AS drafted the paper. TS, KF, AKS, and UC critically revised the manuscript for important intellectual content. All the authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Lifestyle questionnaire.
[XLSX File (Microsoft Excel File), 13 KB - ijmr_v11i1e28692_app1.xlsx]

Multimedia Appendix 2
Factor loading matrix for daily questions.
[PNG File , 69 KB - ijmr_v11i1e28692_app2.png]

Multimedia Appendix 3
Factor loading matrix for weekly questions.
[PNG File , 85 KB - ijmr_v11i1e28692_app3.png]

Multimedia Appendix 4
Factor loading vector for monthly questions.
[PNG File , 3 KB - ijmr_v11i1e28692_app4.png]

Multimedia Appendix 5
Proposed reduced lifestyle questionnaires.
[XLSX File (Microsoft Excel File), 12 KB - ijmr_v11i1e28692_app5.xlsx]

References

1. Lavrakas PJ. Encyclopedia of survey research methods. Thousand Oaks, CA: Sage Publications; 2008.
2. Herzog AR, Bachman JG. Effects of questionnaire length on response quality. Public Opin Q 1981;45(4):549-559. [doi: 10.1086/268687]
3. Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation checks: detecting satisficing to increase statistical power. J Exp Soc Psychol 2009;45(4):867-872. [doi: 10.1016/j.jesp.2009.03.009]
4. Maloney P, Grawitch MJ, Barber LK. Strategic item selection to reduce survey length: reduction in validity? Consult Psychol J Pract Res 2011;63(3):162-175. [doi: 10.1037/a0025604]
5. Stanton JM, Sinar EF, Balzer WK, Smith PC. Issues and strategies for reducing the length of self-report scales. Pers Psychol 2002;55(1):167-194. [doi: 10.1111/j.1744-6570.2002.tb00108.x]
6. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297-334. [doi: 10.1007/BF02310555]
7. McHorney CA, Ware Jr JE, Raczek AE. The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31(3):247-263. [doi: 10.1097/00005650-199303000-00006] [Medline: 8450681]

8. Bai Y, Li J, Bai Y, Ma W, Yang X, Ma F. Development and validation of a questionnaire to evaluate the factors influencing training transfer among nursing professionals. BMC Health Serv Res 2018;18(1):107 [FREE Full text] [doi: 10.1186/s12913-018-2910-7] [Medline: 29433487]
9. Brosnan K, Grün B, Dolnicar S. Identifying superfluous survey items. J Retail Consum Serv 2018;43:39-45. [doi: 10.1016/j.jretconser.2018.02.007]
10. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Qual Life Res 2007;16 Suppl 1:5-18. [doi: 10.1007/s11136-007-9198-0] [Medline: 17375372]
11. Ledbetter MK, Diawara N, Porter BE. The use of item response theory in survey methodology: application in seat belt data. Am J Oper Res 2018;8:17-32. [doi: 10.4236/ajor.2018.81002]
12. Allen DS. The impact of shortening a long survey on response rate and response quality. Brigham Young University. 2016. URL: https://scholarsarchive.byu.edu/etd/5968 [accessed 2022-03-03]
13. Liao D, Valliant R. Variance inflation factors in the analysis of complex survey data. Surv Methodol 2012;38(1):53-62.
14. Liao D. Collinearity diagnostics for complex survey data. University of Maryland. 2010. URL: https://drum.lib.umd.edu/handle/1903/10881 [accessed 2022-03-03]
15. Roberts JG, Tugwell P. Comparison of questionnaires determining patient satisfaction with medical care. Health Serv Res 1987;22(5):637-654 [FREE Full text] [Medline: 3692863]
16. Wilson A, McDonald P. Comparison of patient questionnaire, medical record, and audio tape in assessment of health promotion in general practice consultations. BMJ 1994;309(6967):1483-1485 [FREE Full text] [doi: 10.1136/bmj.309.6967.1483] [Medline: 7804055]
17. Dell-Kuster S, Sanjuan E, Todorov A, Weber H, Heberer M, Rosenthal R. Designing questionnaires: healthcare survey to compare two different response scales. BMC Med Res Methodol 2014;14:96 [FREE Full text] [doi: 10.1186/1471-2288-14-96] [Medline: 25086869]
18. LaDonna KA, Taylor T, Lingard L. Why open-ended survey questions are unlikely to support rigorous qualitative insights. Acad Med 2018;93(3):347-349. [doi: 10.1097/ACM.0000000000002088] [Medline: 29215376]
19. Cappelleri JC, Gerber RA, Kourides IA, Gelfand RA. Development and factor analysis of a questionnaire to measure patient satisfaction with injected and inhaled insulin for type 1 diabetes. Diabetes Care 2000;23(12):1799-1803. [doi: 10.2337/diacare.23.12.1799] [Medline: 11128356]
20. Juniper EF, Guyatt GH, Streiner DL, King DR. Clinical impact versus factor analysis for quality of life questionnaire construction. J Clin Epidemiol 1997;50(3):233-238. [doi: 10.1016/s0895-4356(96)00377-0] [Medline: 9120521]
21. Arifin WN, Yusoff MS. Confirmatory factor analysis of the Universiti Sains Malaysia emotional quotient inventory among medical students in Malaysia. SAGE Open 2016;6(2):215824401665024. [doi: 10.1177/2158244016650240]
22. Prieto L, Alonso J, Lamarca R. Classical test theory versus Rasch analysis for quality of life questionnaire reduction. Health Qual Life Outcomes 2003;1:27 [FREE Full text] [doi: 10.1186/1477-7525-1-27] [Medline: 12952544]
23. Erhart M, Hagquist C, Auquier P, Rajmil L, Power M, Ravens-Sieberer U, European KIDSCREEN Group. A comparison of Rasch item-fit and Cronbach's alpha item reduction analysis for the development of a quality of life scale for children and adolescents. Child Care Health Dev 2010;36(4):473-484. [doi: 10.1111/j.1365-2214.2009.00998.x] [Medline: 19702637]
24. Ellis BW, Johns MW, Lancaster R, Raptopoulos P, Angelopoulos N, Priest RG. The St. Mary's hospital sleep questionnaire: a study of reliability. Sleep 1981;4(1):93-97. [doi: 10.1093/sleep/4.1.93] [Medline: 7232974]
25. Lee KA. Self-reported sleep disturbances in employed women. Sleep 1992;15(6):493-498. [doi: 10.1093/sleep/15.6.493] [Medline: 1475563]
26. Cohen S. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The social psychology of health: the Claremont symposium on applied social psychology. Thousand Oaks, CA: Sage Publications; 1988:31-67.
27. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. 2nd edition. Boston, MA: Springer; 2012.
28. Johnston R, Jones K, Manley D. Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. Qual Quant 2018;52(4):1957-1976 [FREE Full text] [doi: 10.1007/s11135-017-0584-6] [Medline: 29937587]
29. De Jongh PJ, De Jongh E, Pienaar M, Gordon-Grant H, Oberholzer M, Santana L. The impact of pre-selected variance inflation factor thresholds on the stability and predictive power of logistic regression models in credit scoring. Orion 2015;31(1):17-37. [doi: 10.5784/31-1-162]
30. Sheather S. A modern approach to regression with R. New York, NY: Springer; 2009.
31. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: with applications in R. 2nd edition. New York, NY: Springer; 2013.
32. Fabrigar LR, Wegener DT. Exploratory factor analysis: understanding statistics. Oxford, UK: Oxford University Press; 2011.
33. Bartlett MS. Tests of significance in factor analysis. Br J Stat Psychol 1950;3(2):77-85. [doi: 10.1111/j.2044-8317.1950.tb00285.x]
34. Kaiser HF. An index of factorial simplicity. Psychometrika 1974;39(1):31-36. [doi: 10.1007/BF02291575]

35. Kaiser HF. The varimax criterion for analytic rotation in factor analysis. Psychometrika 1958;23(3):187-200. [doi: 10.1007/BF02289233]
36. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods 1999;4(3):272-299. [doi: 10.1037/1082-989X.4.3.272]
37. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. Psychol Methods 1999;4(1):84-99. [doi: 10.1037/1082-989X.4.1.84]
38. Osborne JW, Costello AB, Kellow JT. Best practices in exploratory factor analysis. In: Osborne JW, editor. Best practices in quantitative methods. Thousand Oaks, CA: Sage Publications; 2008:205-213.
39. Hidaka T, Kasuga H, Endo S, Masuishi Y, Kakamu T, Takeda A, et al. Are lifestyle pattern changes associated to poor subjective sleep quality?: a cross-sectional study by gender among the general Japanese population underwent specified medical check-ups in 2014 and 2015. BMJ Open 2020;10(12):e037613 [FREE Full text] [doi: 10.1136/bmjopen-2020-037613] [Medline: 33328256]
40. Goette L, Bendahan S, Thoresen J, Hollis F, Sandi C. Stress pulls us apart: anxiety leads to differences in competitive confidence under stress. Psychoneuroendocrinology 2015;54:115-123 [FREE Full text] [doi: 10.1016/j.psyneuen.2015.01.019] [Medline: 25705012]
41. Toussaint L, Shields GS, Dorn G, Slavich GM. Effects of lifetime stress exposure on mental and physical health in young adulthood: how stress degrades and forgiveness protects health. J Health Psychol 2016;21(6):1004-1014 [FREE Full text] [doi: 10.1177/1359105314544132] [Medline: 25139892]
42. Jensen J, Burlingame G. An item reduction analysis of the group questionnaire. Psychotherapy (Chic) 2018;55(2):144-150. [doi: 10.1037/pst0000145] [Medline: 29863394]
43. de Winter JC, Dodou D, Wieringa PA. Exploratory factor analysis with small sample sizes. Multivariate Behav Res 2009;44(2):147-181. [doi: 10.1080/00273170902794206] [Medline: 26754265]
44. Jung S, Lee S. Exploratory factor analysis for small samples. Behav Res Methods 2011;43(3):701-709. [doi: 10.3758/s13428-011-0077-9] [Medline: 21431996]

Abbreviations

FA: factor analysis
IRT: Item Response Theory
KMO: Kaiser–Meyer–Olkin
VIF: variance inflation factor

Edited by T Leung; submitted 05.04.21; peer-reviewed by A Chang, R Watson, M Nieroda; comments to author 26.04.21; revised version received 18.06.21; accepted 09.02.22; published 18.03.22.

Please cite as:
Staffini A, Fujita K, Svensson AK, Chung UI, Svensson T
Statistical Methods for Item Reduction in a Representative Lifestyle Questionnaire: Pilot Questionnaire Study
Interact J Med Res 2022;11(1):e28692
URL: https://www.i-jmr.org/2022/1/e28692
doi:10.2196/28692
PMID:35302507

Original Paper

Pilot Project for a Web-Based Dynamic Nomogram to Predict Survival 1 Year After Hip Fracture Surgery: Retrospective Observational Study

Graeme McLeod[1,2*], MD; Iain Kennedy[1*], MBChB; Eilidh Simpson[3], MBChB; Judith Joss[1*], MBChB; Katriona Goldmann[4], MSc

[1]Department of Anaesthesia, Ninewells Hospital, National Health Service Tayside, Dundee, United Kingdom

[2]School of Medicine, University of Dundee, Ninewells Hospital, Dundee, United Kingdom

[3]Crosshouse Hospital, National Health Service Ayrshire and Arran, Kilmarnock, United Kingdom

[4]William Harvey Research Institute, Barts and the London School of Medicine & Dentistry, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Graeme McLeod, MD
Department of Anaesthesia
Ninewells Hospital
National Health Service Tayside
James Arnott Drive
Dundee, DD1 9SY
United Kingdom
Phone: 44 1382 632175
Email: g.a.mcleod@dundee.ac.uk

Abstract

Background: Hip fracture is associated with high mortality. Identification of individual risk informs anesthetic and surgical decision-making and can reduce the risk of death. However, interpreting mathematical models and applying them in clinical practice can be difficult. There is a need to simplify risk indices for clinicians and laypeople alike.

Objective: Our primary objective was to develop a web-based nomogram for prediction of survival up to 365 days after hip fracture surgery.

Methods: We collected data from 329 patients. Our variables included sex; age; BMI; white cell count; levels of lactate, creatinine, hemoglobin, and C-reactive protein; physical status according to the American Society of Anesthesiologists Physical Status Classification System; socioeconomic status; duration of surgery; total time in the operating room; side of surgery; and procedure urgency. Thereafter, we internally calibrated and validated a Cox proportional hazards model of survival 365 days after hip fracture surgery; logistic regression models of survival 30, 120, and 365 days after surgery; and a binomial model. To present the models on a laptop, tablet, or mobile phone in a user-friendly way, we built an app using Shiny (RStudio). The app showed a drop-down box for model selection and horizontal sliders for data entry, model summaries, and prediction and survival plots. A slider represented patient follow-up over 365 days.

Results: Of the 329 patients, 24 (7.3%) died within 30 days of surgery, 65 (19.8%) within 120 days, and 94 (28.6%) within 365 days. In all models, the independent predictors of mortality were age, BMI, creatinine level, and lactate level. The logistic model also incorporated white cell count as a predictor. The Cox proportional hazards model showed that mortality differed as follows: age 80 vs 60 years had a hazard ratio (HR) of 0.6 (95% CI 0.3-1.1), a plasma lactate level of 2 vs 1 mmol/L had an HR of 2.4 (95% CI 1.5-3.9), and a plasma creatinine level of 60 vs 90 mol/L had an HR of 2.3 (95% CI 1.3-3.9).

Conclusions: In conclusion, we provide an easy-to-read web-based nomogram that predicts survival up to 365 days after hip fracture. The Cox proportional hazards model and logistic models showed good discrimination, with concordance index values of 0.732 and 0.781, respectively.

(Interact J Med Res 2022;11(1):e34096) doi:10.2196/34096

KEYWORDS

hip fracture; survival; prediction; nomogram; web; surgery; postoperative; machine learning; model; survival; mortality; hip; fracture

Introduction

As many as 7 of 100 patients die in the first 30 days after hip fracture [1-4]. Mortality in the first 365 days after hip fracture surgery varies between 14% and 23% of patients [5,6]. Identification of individual risk can inform anesthetic and surgical decision-making and potentially improve outcomes. However, mathematical models can be complex and difficult to interpret, and the effect of changes in continuous or categorical variables may not be obvious. Graphical presentation of data is a pivotal technique in science and key to better communication [7]. Nomograms present covariables in a relatively easy-to-understand way and are commonly used to inform clinicians and patients of the risk of mortality in prostate cancer [8]. However, interpreting predictive indices and applying them to individual patients is difficult [9]. Apps based on the R package (RStudio v 1.3.1093; R Foundation for Statistical Computing), such as Shiny, can be used to translate statistical models into easy-to-understand, web-based interactive nomograms that readily demonstrate differences between low-risk and high-risk patients. One example is the DynNom package [7] in R that predisplays the results of statistical models as a dynamic nomogram and readily allows individual prediction with 95% CI.

Anesthetic guidelines and protocols increasingly drive standardization of practice [10]. However, we believe that individual identification of risk is more likely to improve outcomes [11]. Several risk-specific and generic surgical risk indices are available that predict mortality after hip fracture surgery, but most are limited to prediction of mortality 30 days after operation. These indices discriminate well but lack adequate calibration [12]. The Nottingham Hip Fracture Score has been validated nationally [2] and internationally [13] and is commonly used, but its 365-day score discriminates only between low- and high-risk patients. Moreover, the more complex a statistical model is due to nonlinearity and interactions, the more difficult it is to comprehend and apply. As such, survival models based on time-to-death data [14] are uncommon. Therefore, there is a need to develop an easy-to-interpret app for time-to-event as well as binary-outcome data. Proving such an app's utility locally would provide a platform for prospective development of a large multicenter database that could inform a statistical model with high calibration and discrimination. This model could easily be used at the bedside with a laptop, tablet, or mobile phone in order to inform staff and laypeople of outcomes after hip fracture surgery. Therefore, our primary objective was to develop a web-based nomogram from clinical data collected over an 8-month period from patients undergoing hip fracture surgery at a single tertiary center.

Methods

We conducted a retrospective study of patients undergoing hip fracture surgery. Our study included data collection, statistical modeling, and app development.

Data Collection

We collected preoperative and operative data from all patients presenting for hip fracture surgery at Ninewells Hospital, Dundee, Scotland, over an 8-month period between May 1, 2016, and December 31, 2016. The patients' case notes, anesthetic charts, and operative notes for the first year after surgery were reviewed as part of a fourth-year medical student project.

The data included patient characteristics, comorbidities, and health status. Patient characteristics recorded on admission included age, sex, BMI, fracture side (left or right), type of fracture (intracapsular or extracapsular), their type of residence before the fracture, and a social deprivation score based on the Scottish Index of Multiple Deprivation 2016 (SIMD16) database, which measures deprivation in 6976 residential areas in Scotland [15]. We used the SIMD16 vigintile database, which ranks deprivation from 1 (the most deprived residential areas) to 20 (the least deprived residential areas). Blood tests were taken on hospital admission and included white cell count and levels of hemoglobin, creatinine, lactate, and C-reactive protein. With regard to surgery, we noted the physical status of the patient according to the American Society of Anesthesiologists (ASA) Physical Status Classification System, the type of anesthesia (general or spinal), the type of surgical implant, and the time the operation took place. Operations performed between 9 AM and 5 PM were classified as daytime operations, those performed between 5 PM and 10 PM as evening operations, and those performed between 5 PM and 9 AM as nighttime operations. Postoperatively, we noted the need for transfusion, presence of acute kidney injury, cardiovascular complications such as pulmonary embolus or myocardial infarction, and infection from any source (wound, urinary, or respiratory). We recorded the date of hospital discharge and the destination of the patient. The type of residence of the patient before the fracture and upon discharge from the ward were classified as the following: home (either the patient's own home or sheltered housing), care home, acute-care hospital, rehabilitation hospital, or long-term-care hospital. Our primary outcome was time to death by any cause within 365 days of hip fracture surgery.

Model Development

We developed 4 statistical models: a global Cox proportional hazards model using all available covariates, a final Cox proportional hazards model, a generalized linear model, and a logistic regression model. Models and nomograms were developed using the R packages "shiny," "ggplot2," "ggpub," "stargazer," "rms," "shinythemes," and "plotly."

Our modeling strategy was based on that recommended by Harrell and Steyerberg [9]. We selected variables based on our clinical experience and evidence from published studies. We collected as much pertinent data as possible, with wide distributions for predictor values. We hypothesized that continuous variables were nonlinear. We used imputation to replace missing covariables with the median value. We restricted the number of events per variable in the model according to the following equation: events per variable = events or outcomes/15. We prespecified the complexity of the model and initially allotted 3 cubic splines (knots) to continuous variables in order to detect any nonlinear relationships between variables and outcomes and allotted 1 *df* to categorical data.

We first created a global model using all variables and tested the association of each predictor with outcomes adjusted for all other predictors and the *df* used. We reduced the model by calculating the *df* that could be spent and deciding how they should be spent. We ranked the apparent importance of predictors of death by plotting the Akaike information criterion, defined as $\chi^2 - 2$ *df*. Initial estimation of shrinkage (γ) needed used the formula $\gamma = (\chi^2 - p)/\chi^2$). We also interpreted the model graphically and decided which parameters should be retained for bootstrap validation of calibration and discrimination. Continuous variables that showed a linear relationship with outcome were restricted to 1 *df*.

Overfitting and effects of shrinkage were assessed using the corrected calibration slope. This was obtained using bootstrapping bias–corrected (overfitting minus corrected) estimates of predicted vs observed values. In order to check proportional hazards assumptions, we examined scaled Schoenfeld residuals.

Model Validation

Prediction errors were assessed using the log-likelihood ratio (χ^2) for continuous data and the Brier score for binary data. The ability to discriminate between low-risk and high-risk patients was measured with R^2, the Gini index from 0 to 1, a robust measure of variation, and measures of rank discrimination, such as the *C* index and Somers Dxy. The *C* index represents the probability of concordance, *C*, between predicted and observed survival, and is equivalent to the area under the receiver operator characteristics curve (AUROC). Concordance is defined as the proportion of all pairs of subjects whose survival time can be ordered such that the subject with the higher predicted survival is the one who survived longer. Dxy is the difference between concordance and discordance probabilities and relates to the *C* index by the equation Dxy = 2(*C* − 0.5). Internal calibration and validation used the bootstrap.

Model Comparisons

Our secondary objectives were to develop a 365-day logistic regression model and a 365-day generalized linear model for binomial response data for sensitivity analysis, and to develop additional 30-day and 120-day logistic models in order to compare accuracy against the routinely used Nottingham Hip Fracture Score.

App Development

A data scientist (KG) developed an app using Shiny, a package from RStudio that builds interactive web applications with R. We created 3 files: ui.R to define the user interface A; server.R to interrogate data from the user interface and define the app logic; and functions.R to combine these 2 files and create the Shiny application.

The user interface (ui.R) consisted of a title, side panel, and main panel. The side panel contained a drop-down box with 4 models: the global Cox proportional hazards model, the final Cox proportional hazards model, the generalized linear model, and the logistic regression model. The side panel also had sliders for input of continuous variables over their range of values and follow-up time (0 to 365 days). The main panel consisted of 3 tabs: a prediction plot, a survival plot, and a model summary.

Prediction plots were displayed on a graph with probability on the x-axis. The mean was displayed as a colored square with horizontal lines representing the 95% CI for the outcome. Survival models showed a Kaplan-Meier plot of estimated survival probability over time. The app can be viewed at our page on the shinyapps website [16].

Statistical Analysis

Continuous variables are presented as the mean (SD) and were analyzed using the Aspin-Welch unequal variance test. Nonparametric data were presented as the median (IQR, full range) and analyzed using the Mann-Whitney *U* test. Cross-tabulation of categorical data count (n) was analyzed using the χ^2 test. Calculation of the AUROC for Nottingham Hip Fracture Score used GraphPad Prism 9 (GraphPad).

Ethics Approval

Caldicott guardian approval was obtained from the University of Dundee on October 16, 2016. In the United Kingdom, Caldicott guardians provide ethical approval for interrogation of anonymous clinical databases.

Results

Data Collection

We recorded data from 329 patients, of whom 224 (68%) were female and 85 (32%) were male. We found that 4% of biochemical data were missing and replaced them with the median value. Over two-thirds of patients (224/329, 68%) were classified as ASA category III or IV. These categories indicate severe systemic disease and disease that is a constant threat to life, respectively. We found that 24 (7.3%) patients died within 30 days, 65 (19.8%) within 120 days, and 94 (28.6%) within 365 days of surgery. Patient characteristics, categorized according to survival or death within 365 days, are shown in Table 1.

Table 1. Characteristics of surviving and deceased patients 365 days after hip fracture surgery.

Variable	Surviving (n=235)	Deceased (n=94)	Difference (95% CI), odds ratio (95% CI)	*P* value
Age in years, mean (SD)	82.5 (10.0)	80.9 (9.6)	1.5 (0.8 to 3.9)	.21
Sex, n (%)			1.0 (0.6 to 1.7)	.94
Male	61 (26)	24 (25.5)		
Female	174 (74)	70 (74.5)		
BMI in kg/m^2, mean (SD)	24.2 (5.7)	21.8 (4.2)	2.4 (0.9 to 3.8)	.002
Status according to American Society of Anesthesiologists Physical Status Classification System, n (%)			N/A[a]	<.001
I	7 (3)	0 (0)		
II	62 (26.4)	5 (5.3)		
III	113 (48.1)	58 (61.7)		
IV	27 (11.5)	26 (27.7)		
Type of residence, n (%)			N/A	<.001
Home	189 (80.4)	46 (48.9)		
Care home	41 (17.4)	45 (47.9)		
Rehabilitation hospital	3 (1.3)	3 (3.2)		
Acute-care hospital	1 (0.4)	0 (0)		
Long-term–care hospital	1 (0.4)	0 (0)		
Scottish Index of Multiple Deprivation 2016 score, median (IQR, full range)	11 (6 to 16, 1 to 20)	12 (8 to 16, 1 to 20)	1.0 (–1.0 to 2.0)	.42
Stay in days, mean (SD)	12.6 (10.2)	12.1 (8.2)	0.5 (–1.7 to 2.6)	.67
Side, n (%)			1.0 (0.6 to 0.6)	.89
Left	123 (52.3)	50 (53.2)		
Right	112 (47.7)	44 (46.8)		
Implant type, n (%)			N/A	.70
Bipolar	18 (7.7)	3 (3.2)		
Compression hip screw	80 (34)	35 (37.2)		
Collarless, polished, tapered	26 (11.1)	1 (1.1)		
Thompson	88 (37.4)	45 (47.9)		
Femoral nail	23 (9.8)	10 (10.6)		
Hemoglobin in g/L, mean (SD)	120.3 (18.0)	115.6 (15.0)	4.7 (0.8 to 8.6)	.01
White cell count in 10^9/L, mean (SD)	11.8 (6.1)	11.1 (3.2)	0.8 (–0.2 to 1.8)	.14
C-reactive protein in mg/L, median (IQR, full range)	6 (3 to 25, 2 to 299.0)	13 (3 to 46, 3 to 273)	1.0 (0.0 to 3.0)	.046
Lactate in mmol/L, mean (SD)	1.47 (0.74)	1.70 (0.92)	0.24 (0.0 to 0.48)	.04
Creatinine in µmol/L, mean (SD)	71.1 (27.5)	89.2 (42.0)	18.2 (8.4 to 27.8)	<.001
Operation time (%)			N/A	.02
Daytime (9 AM to 5 PM)	196 (83.4)	78 (83)		
Evening (5 PM to 10 PM)	37 (15.7)	14 (14.9)		
Night (10 PM to 9 AM)	2 (0.9)	2 (2.1)		
Type of residence after discharge (%)			N/A	<.001
Home	100 (42.6)	13 (13.8)		
Care home	61 (26)	42 (44.7)		
Rehabilitation setting	54 (23)	18 (19.1)		

Variable	Surviving (n=235)	Deceased (n=94)	Difference (95% CI), odds ratio (95% CI)	*P* value
Acute-care hospital	15 (6.4)	6 (6.4)		
Long-term-care hospital	7 (3)	3 (3.2)		
Died in hospital	2 (0.9)	8 (8.5)		

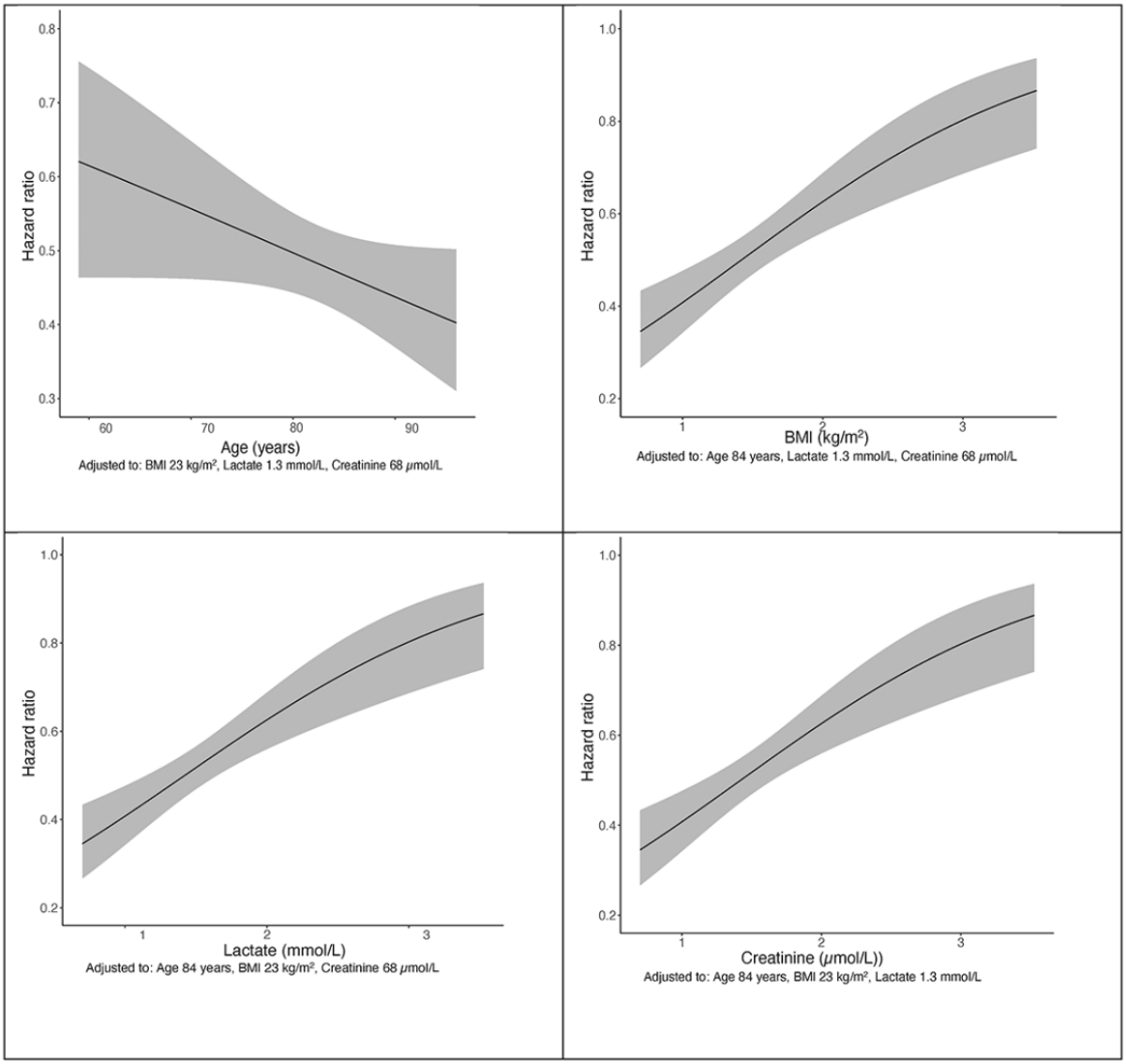
[a]N/A: not applicable.

Model Development

A global Cox proportional hazards model using all covariates, a final Cox proportional hazards model, logistic regression models, and a generalized linear model were constructed from the data. The global Cox proportional hazards model took all covariates into account, whereas the final validated model was built within the statistical constraints discussed in the methods. The graphs in Figure 1 show continuous relationships between covariates and the probability of death using the final Cox proportional hazards model. The nonlinear relationship between creatinine and the risk of death (ie, HR) was calculated using cubic splines (Figure 1).

Independent predictors of mortality in the final Cox proportional hazards model included increased age, BMI, creatinine, lactate, and the combination of these factors. Examples of differences in mortality on admission included the following: age 80 vs 60 years had an HR of 0.6 (95% CI 0.3-1.1), a plasma lactate level of 2 vs 1 mmol/L had an HR of 2.4 (95% CI 1.5-3.9), and a plasma creatinine level 60 vs 90 µmol/L had an HR of 2.3 (95% CI 1.3-3.9).

Figure 1. Final Cox proportional hazards model 365 days after hip fracture surgery. Hazard ratios show a reduced risk of death with increasing age and lower BMI. Risk of death rose with increased creatinine and lactate levels. Note the nonlinear increase in risk with creatinine level, and the increase in risk from values immediately above the physiological range.



Model Validation

Validation results for the global and final Cox proportional hazards models are shown in Table 2.

Table 2. Model validation. Global and final Cox proportional hazards models. The final model was developed after iterative data reduction and calibration using bootstrap and showed good validation in 329 patients.

Model	R^{2a}	LR (χ^2)[b]	*P* value	Dxy[c]	C index[d]	g[e]
Global Cox proportional hazards model 365 days after surgery	0.364	χ^2_{22}=45.328	.002	0.623	0.812	1.897
Final Cox proportional hazards model 365 days after surgery	0.231	χ^2_{7}=43.113	<.01	0.474	0.732	1.360

[a]R^2 coefficient of determination.

[b]Likelihood ratio chi-square test.

[c]Somers Dxy test.

[d]Concordance index.

[e]Gini index.

Model Comparisons

The predictive variables identified using the final Cox proportional hazards model were similar to the predictive variables identified using the 365-day logistic regression and 365-day binomial models (Table 3).

Validation results for our secondary outcomes and the 30, 120, and 365-day logistic regression models are presented in Table 4.

Using our data, we calculated the AUROC for Nottingham Hip Fracture Score to be <0.61 (95% CI) at all time points (Table 5).

An example of an easy-to-interpret dynamic nomogram is presented in Figure 2. The variables of this nomogram can be altered using sliders. The digital nomogram is available online at our website [16].

Table 3. Independent variables predicting mortality in the final Cox proportional hazards model, a logistic model, and a binomial model. All models are 365 days after hip fracture surgery. Variables common to all models included age, BMI, lactate, and creatinine. Apostrophes indicate nonlinear restricted cubic splines.

Dependent variable	Final Cox proportional hazards model, regression coefficient (95% CI)	Logistic regression model, regression coefficient (95% CI)	Binomial model, regression coefficient (95% CI)
Age	0.976 (0.947 to 1.007)	–0.023 (–0.062 to 0.016)	–0.018 (–0.056 to 0.020)
BMI	0.913 (0.862 to 0.967)	–0.115 (–0.199 to –0.032)	–0.126 (–0.205 to –0.047)
White cell count	N/A[a]	0.138 (–0.109 to 0.385)	–0.028 (–0.105 to 0.048)
White cell count’	N/A	–0.196 (–0.453 to 0.062)	N/A
Lactate	0.003 (<0.001 to 0.199)	–5.519 (–10.812 to –0.226)	–0.899 (–0.095 to 1.893)
Creatinine	0.906 (0.817 to 1.005)	–0.072 (–0.198 to 0.055)	–0.031 (0.008 to 0.055)
Creatinine’	1.185 (1.030 to 1.364)	0.133 (–0.042 to 0.308)	N/A
Lactate*Creatinine	1.110 (1.037 to 1.189)	0.098 (0.013 to 0.183)	-0.007 (–0.018 to 0.004)
Lactate*Creatinine’	0.865 (0.788 to 0.951)	–0.134 (–0.250 to –0.018)	N/A
Constant	N/A	5.471 (–3.329 to 14.270)	0.491 (–3.379 to 4.360)

[a]N/A: not applicable.

Table 4. Logistic regression validation results.

Model	R^{2a}	LR (χ^2_9)[b]	*P* value	Brier	Dxy[c]	C index[d]	g[e]
Logistic model (30 days)	0.714	17.390	.004	0.069	0.541	0.770	1.348
Logistic model (120 days)	0.396	21.280	.002	0.114	0.706	0.853	2.051
Logistic model (365 days)	0.277	37.252	<.001	0.147	0.562	0.781	1.619

[a]R^2 coefficient of determination.

[b]Likelihood ratio chi-square test.

[c]Somers Dxy test.

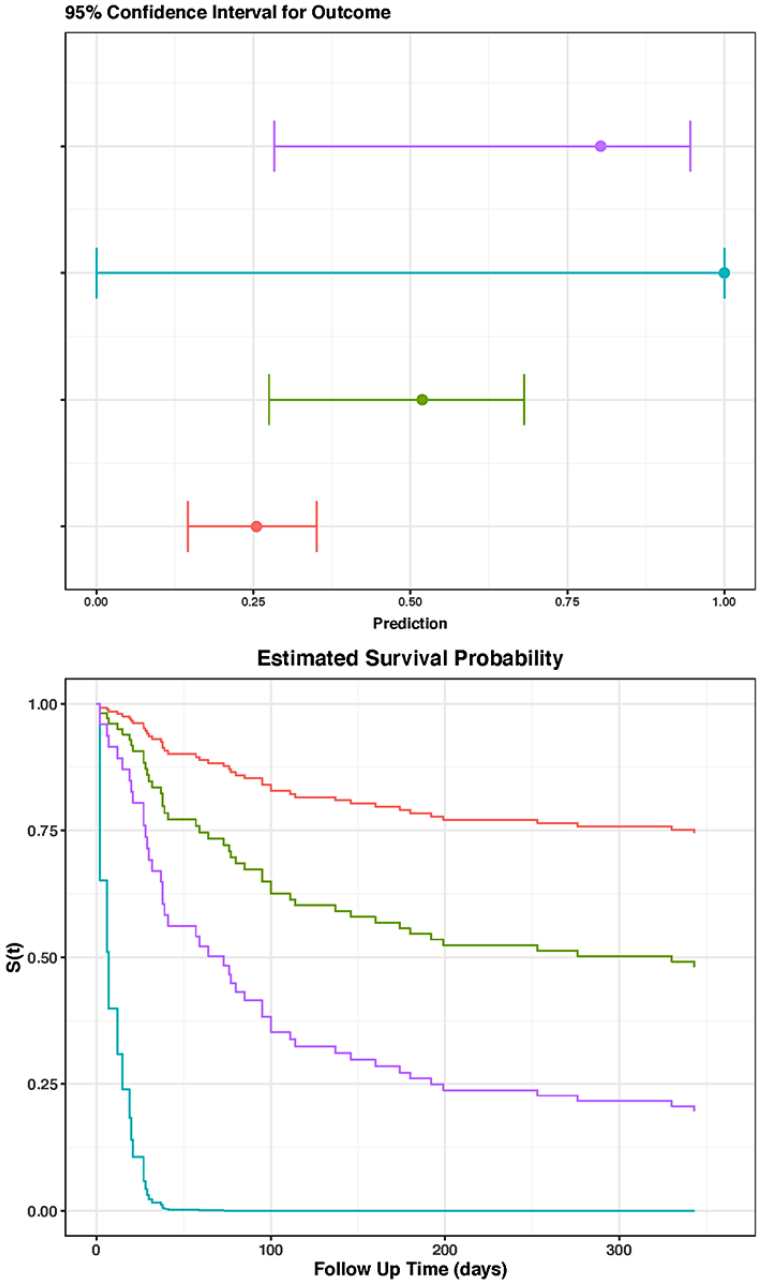
[d]Concordance index.

[e]Gini index.

Table 5. Diagnostic results for Nottingham Hip Fracture Score.

Time	Area under the receiver operating characteristics curve (95% CI)	*P* value
30 days	0.576 (0.454-0.698)	.22
120 days	0.606 (0.538-0.674)	.003
365 days	0.602 (0.526-0.678)	.01

Figure 2. Dynamic nomogram. Top: sliders that are used to enter data for age (standardized to 80 years) and white cell count (standardized to 10 x 109/L). Bottom: 4 imaginary scenarios, differing in BMI, creatinine, and lactate. The red, green, blue, and purple lines represent the following values for BMI, creatinine, and lactate, respectively: 25 kg/m^2, 80 µmol/L, 1.5; 15 kg/m^2, 80 µmol/L, 1.5 mmol/L; 15 kg/m^2, 80 µmol/L, 4 mmol/L; and 15 kg/m^2, 140 µmol/L, 4 mmol/L. The dynamic nomogram is available on our website [16].



Discussion

Principal Findings

We provide proof of concept of a simple, dynamic digital nomogram created in R and Shiny that shows individual survival with the 95% CI after hip fracture surgery. The nomogram offers an easy, intuitive means of interpreting complicated models. Our models showed good discrimination and calibration. Lactate, creatinine, age, and BMI emerged as important predictors of mortality in all models.

Comparison to Prior Work

Our data are consistent with previous studies demonstrating an association between higher serum lactate and mortality following hip fracture [18-20]. For example, we found that a rise in plasma lactate from 2 to 3 mmol/L increased the hazard ratio by >25%. Unlike previous studies, we did not arbitrarily define raised lactate as a level >2.5 mmol/L [21] or 3.0 mmol/L [18,20]. In fact, our nonlinear modeling of continuous lactate data showed an early, steep rise in the risk of mortality from 1 mmol/L. This has implications for clinical practice. It suggests that a lower-than-anticipated lactate level has an impact on short-term and long-term mortality and the need for early resuscitation. However, we are not aware of any randomized controlled trials that have examined fluid resuscitation in patients presenting with hip fracture. An association between prolonged lactate clearance and mortality may occur in the surgical intensive care

unit population [22], but this cannot be extrapolated to the management of elderly patients with hip fracture.

Nonlinear modeling of our creatinine data also showed an early, steep rise in the risk of mortality. For example, a rise in plasma creatinine from 60 to 90 μmol/L more than doubled the risk of death. Once more, this demonstrates that changes just outside the normal physiological range may profoundly impact outcomes; clinicians should take note of such changes, rather than wait for grossly deranged blood results.

Our models also revealed that there was an inverse association of outcome with BMI [23,24] and that frailty and muscle mass had a significant long-term negative impact on survival after hip fracture surgery. For example, a reduction in BMI from 25 to 20 kg/m^2 increased the hazard ratio by approximately one-third. Unlike other studies, we failed to show a significant effect of anemia. This probably reflects changes in patient blood management strategies since initial studies into this association were published [2,3]. Surprisingly, we showed an inverse relationship between age and outcome, in contrast to many other models [12,25,26]. This reflects increased comorbidities in our younger population, limiting the applicability of our model to other populations. Nevertheless, for comparison, we applied the Nottingham Hip Fracture Score to our data. Surprisingly, the Nottingham score showed much poorer discrimination compared to our dataset, with an AUROC <0.61 and a lack of statistical significance for 30-day predictions of mortality.

Strengths and Limitations

Our study had 3 key strengths. First, rather than just focus on 30-day mortality, we observed our patients for 12 months in order to obtain a detailed temporal overview of outcomes after hip fracture surgery. Most models, in contrast, focus on measurement of 30-day mortality [4,7,9-12] and tend to reflect events during hospital stay. By contrast, the Nottingham Hip Fracture Score predicts 1-year mortality [13], but it divides patients according to a binary low risk/high risk classification based on a cut-off score.

Second, we used modeling techniques available in R. The nonlinearity of creatinine and the interaction with lactate justified our application of restricted cubic splines to continuous data. Although this allocated 3 degrees of freedom to continuous variables, this technique improved the accuracy of the model. We also used bootstrapping to validate our model. The advantage of bootstrapping is that the entire dataset can be used, unlike data splitting, which reduces the sample size for both model development and testing. Variable selection or stopping rules were not used, because these methods provide regression coefficients that are too high and confidence intervals that are too small. Neural networks, such as support vector machines, naive Bayes classifiers, and random forest classifiers, have been applied to hip fracture data sets, but were no better than logistic regression in predicting outcomes after surgery [27].

Third, our mortality was in line with national data. Mortality increased from 7.3% at 30 days to 28.6% at 365 days and allowed us to incorporate 5 variables with good calibration and validation.

A limitation of this study was insufficient data; we could not generate a model that incorporated all potential confounders. We suggest investigators capture data from the dimensions of risk recommended by Iezzoni [28], as these are most likely to explain variations in mortality. Such variables should not only include patient characteristics, recent health status, mental acuity, and quality of life, but also markers of acute clinical stability.

Future Directions

We present an example of our dynamic nomogram online [16] but emphasize that, based on our global model, predictions can be improved by recruiting more patients. While this study identifies important risk factors for mortality in hip fracture patients and robustly demonstrates a proof of concept for an app-based dynamic nomogram of individualized mortality risk, medical apps in the United Kingdom must be registered with the Medicines and Healthcare Products Regulatory Agency as Class 1 medical devices prior to any clinical use, which requires prospective registration of data gathering. Our app is not registered and should not be used to guide specific patient treatment; the prototype app provided online is for educational purposes and to inform future research. As such, validation against a larger patient population is needed to validate the model and support a future application for Medicines and Healthcare Products Regulatory Agency device registration.

Conclusion

We developed a dynamic nomogram for prediction of survival using Shiny that presents a Cox proportional hazards model and logistic and binomial models in an easy, intuitive, and interpretable format. All models identified lactate and creatinine levels at admission as independent predictors of mortality. Although our relatively small numbers limit external application at this time, our findings nevertheless show that acute hemodynamic changes drive mortality not just in the first 30 days, but also up to 1 year after operation.

Conflicts of Interest

GM is a member of the B Braun-Philips scientific advisory panel and has received funding for presentation of research at international meetings.

References

1. Royal College of Physicians. National Hip Fracture Database annual report 2018. London: RCP; 2018.

2. Marufu TC, White SM, Griffiths R, Moonesinghe SR, Moppett IK. Prediction of 30-day mortality after hip fracture surgery by the Nottingham Hip Fracture Score and the Surgical Outcome Risk Tool. Anaesthesia 2016 May;71(5):515-521 [FREE Full text] [doi: 10.1111/anae.13418] [Medline: 26940757]
3. Pincus D, Ravi B, Wasserstein D, Huang A, Paterson JM, Nathens AB, et al. Association Between Wait Time and 30-Day Mortality in Adults Undergoing Hip Fracture Surgery. JAMA 2017 Nov 28;318(20):1994-2003 [FREE Full text] [doi: 10.1001/jama.2017.17606] [Medline: 29183076]
4. White SM, Moppett IK, Griffiths R, Johansen A, Wakeman R, Boulton C, et al. Secondary analysis of outcomes after 11,085 hip fracture operations from the prospective UK Anaesthesia Sprint Audit of Practice (ASAP-2). Anaesthesia 2016 May;71(5):506-514 [FREE Full text] [doi: 10.1111/anae.13415] [Medline: 26940645]
5. Menéndez-Colino R, Alarcon T, Gotor P, Queipo R, Ramírez-Martín R, Otero A, et al. Baseline and pre-operative 1-year mortality risk factors in a cohort of 509 hip fracture patients consecutively admitted to a co-managed orthogeriatric unit (FONDA Cohort). Injury 2018 Mar;49(3):656-661. [doi: 10.1016/j.injury.2018.01.003] [Medline: 29329713]
6. Kim B, Lee Y, Park H, Sohn H, Oh A, Jeon Y, et al. C-reactive protein is an independent predictor for 1-year mortality in elderly patients undergoing hip fracture surgery: A retrospective analysis. Medicine (Baltimore) 2016 Oct;95(43):e5152 [FREE Full text] [doi: 10.1097/MD.0000000000005152] [Medline: 27787371]
7. Jalali A, Alvarez-Iglesias A, Roshan D, Newell J. Visualising statistical models using dynamic nomograms. PLoS ONE 2019 Nov 15;14(11):e0225253. [doi: 10.1371/journal.pone.0225253] [Medline: 31730633]
8. Zhou X, Ning Q, Jin K, Zhang T, Ma X. Development and validation of a preoperative nomogram for predicting survival of patients with locally advanced prostate cancer after radical prostatectomy. BMC Cancer 2020 Feb 04;20(1):97 [FREE Full text] [doi: 10.1186/s12885-020-6565-5] [Medline: 32019501]
9. Harrell F. Regression Modeling Strategies. 2nd ed. Heidelburg: Springer; 2015.
10. Griffiths R, Babu S, Dixon P, Freeman N, Hurford D, Kelleher E, et al. Guideline for the management of hip fractures 2020: Guideline by the Association of Anaesthetists. Anaesthesia 2021 Feb;76(2):225-237. [doi: 10.1111/anae.15291] [Medline: 33289066]
11. Munirama S, McLeod G. 'Stratified' approach to individualized anaesthetic care. Br J Anaesth 2015 Apr;114(4):543-545 [FREE Full text] [doi: 10.1093/bja/aeu345] [Medline: 25296913]
12. Karres J, Heesakkers NA, Ultee JM, Vrouenraets BC. Predicting 30-day mortality following hip fracture surgery: evaluation of six risk prediction models. Injury 2015 Feb;46(2):371-377. [doi: 10.1016/j.injury.2014.11.004] [Medline: 25464983]
13. Tilkeridis K, Ververidis A, Kiziridis G, Kotzamitelos D, Galiatsatos D, Mavropoulos R, et al. Validity of Nottingham Hip Fracture Score in Different Health Systems and a New Modified Version Validated to the Greek Population. Med Sci Monit 2018 Oct 27;24:7665-7672 [FREE Full text] [doi: 10.12659/MSM.909943] [Medline: 30367027]
14. Guzon-Illescas O, Perez Fernandez E, Crespí Villarias N, Quirós Donate FJ, Peña M, Alonso-Blas C, et al. Mortality after osteoporotic hip fracture: incidence, trends, and associated factors. J Orthop Surg Res 2019 Jul 04;14(1):203 [FREE Full text] [doi: 10.1186/s13018-019-1226-6] [Medline: 31272470]
15. McAlpine. Scottish Index of Multiple Deprivation. URL: https://www.gov.scot/publications/scottish-index-multiple-deprivation-2016/pages/5/ [accessed 2021-01-28]
16. Goldmann K. Hip fracture modelling. URL: https://hip-fracture-modelling.shinyapps.io/hip-fracture-modelling/ [accessed 2021-02-28]
17. Scottish Index of Multiple Deprivation. URL: https://www.gov.scot/publications/scottish-index-multiple-deprivation-2016/pages/5/ [accessed 2021-01-28]
18. Venkatesan M, Smith RP, Balasubramanian S, Khan A, Uzoigwe CE, Coats TJ, et al. Serum lactate as a marker of mortality in patients with hip fracture: A prospective study. Injury 2015 Nov;46(11):2201-2205. [doi: 10.1016/j.injury.2015.06.038] [Medline: 26337798]
19. Chen X, Zhang J, Zhang Y, Lin Y, Liu Z, Wang X, et al. Serum lactate concentration on admission to hospital predicts the postoperative mortality of elderly patients with hip fractures 30 days after surgery. Am J Transl Res 2021;13(9):10363-10371 [FREE Full text] [Medline: 34650704]
20. Uzoigwe CE, Venkatesan M, Smith R, Burnand HGF, Young PS, Cheesman CL, et al. Serum lactate is a prognostic indicator in patients with hip fracture. Hip Int 2012;22(5):580-584. [doi: 10.5301/HIP.2012.9762] [Medline: 23112077]
21. Salottolo KM, Mains CW, Offner PJ, Bourg PW, Bar-Or D. A retrospective analysis of geriatric trauma patients: venous lactate is a better predictor of mortality than traditional vital signs. Scand J Trauma Resusc Emerg Med 2013 Feb 14;21:7 [FREE Full text] [doi: 10.1186/1757-7241-21-7] [Medline: 23410202]
22. Haas SA, Lange T, Saugel B, Petzoldt M, Fuhrmann V, Metschke M, et al. Severe hyperlactatemia, lactate clearance and mortality in unselected critically ill patients. Intensive Care Med 2016 Feb;42(2):202-210. [doi: 10.1007/s00134-015-4127-0] [Medline: 26556617]
23. Akinleye SD, Garofolo G, Culbertson MD, Homel P, Erez O. The Role of BMI in Hip Fracture Surgery. Geriatr Orthop Surg Rehabil 2018;9:2151458517747414 [FREE Full text] [doi: 10.1177/2151458517747414] [Medline: 29468090]
24. Hjelholt TJ, Johnsen SP, Brynningsen PK, Knudsen JS, Prieto-Alhambra D, Pedersen AB. Development and validation of a model for predicting mortality in patients with hip fracture. Age Ageing 2022 Jan 06;51(1):233-233. [doi: 10.1093/ageing/afab233] [Medline: 34923589]

25. Chen C, Chen Y, Chen H, Hung C, Shi H. Artificial Neural Network and Cox Regression Models for Predicting Mortality after Hip Fracture Surgery: A Population-Based Comparison. Medicina (Kaunas) 2020 May 19;56(5):243-243 [FREE Full text] [doi: 10.3390/medicina56050243] [Medline: 32438724]
26. Moppett I. Accuracy of POSSUM and Nottingham Hip Fracture Score. Acta Anaesthesiol Scand 2018 Oct;62(9):1333-1334. [doi: 10.1111/aas.13201] [Medline: 29963689]
27. Forssten MP, Bass GA, Ismail AM, Mohseni S, Cao Y. Predicting 1-Year Mortality after Hip Fracture Surgery: An Evaluation of Multiple Machine Learning Approaches. J Pers Med 2021 Jul 27;11(8):727-727 [FREE Full text] [doi: 10.3390/jpm11080727] [Medline: 34442370]
28. Iezzoni LI. The risks of risk adjustment. JAMA 1997 Nov 19;278(19):1600-1607. [doi: 10.1001/jama.278.19.1600] [Medline: 9370507]

Abbreviations

ASA: American Society of Anesthesiologists
AUROC: area under the receiver operator characteristics curve
HR: hazard ratio
SIMD16: Scottish Index of Multiple Deprivation 2016

Edited by T Leung; submitted 13.10.21; peer-reviewed by A Jalali, H Ayatollahi; comments to author 01.12.21; revised version received 18.01.22; accepted 13.02.22; published 30.03.22.

Please cite as:
McLeod G, Kennedy I, Simpson E, Joss J, Goldmann K
Pilot Project for a Web-Based Dynamic Nomogram to Predict Survival 1 Year After Hip Fracture Surgery: Retrospective Observational Study
Interact J Med Res 2022;11(1):e34096
URL: https://www.i-jmr.org/2022/1/e34096
doi:10.2196/34096
PMID:35238320

Original Paper

Pulmonary Screening Practices of Otolaryngology–Head and Neck Surgeons Across Saudi Arabia in the Posttreatment Surveillance of Squamous Cell Carcinoma: Cross-sectional Survey Study

Majed Alnefaie[1*], MBBS; Abdullah Alamri[2*], MBBS; Asalh Saeedi[2*], MBBS; Awwadh Althobaiti[2*], MBBS; Shahad Alosaimi[2], MBBS; Yousuf Alqurashi[2*], MBBS; Hani Marzouki[2*], MD; Mazin Merdad[2*], MD

[1]King Fahad Armed Forces Hospital, Medical Services of The Armed Forces, Jeddah, Saudi Arabia

[2]Department of Otolaryngology Head and Neck Surgery, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

[*]these authors contributed equally

Corresponding Author:
Majed Alnefaie, MBBS
King Fahad Armed Forces Hospital
Medical Services of The Armed Forces
Al Kurnaysh Rd
Al Andalus
Jeddah, 23311
Saudi Arabia
Phone: 966 500900450
Fax: 966 126823614
Email: Majed.n.md@gmail.com

Abstract

Background: With respect to patients with head and neck squamous cell carcinoma (HNSCC), posttreatment surveillance for distant disease has mostly focused on the lungs, as HNSCC distant metastasis occurs in this organ in 90% of HNSCC cases. Additionally, the incidence rate of primary tumors in the lungs is high due to the field cancerization of the entire upper aerodigestive tract.

Objective: Our cross-sectional survey study aims to evaluate the current beliefs and pulmonary screening practices of otolaryngology–head and neck surgeons across Saudi Arabia with respect to the posttreatment surveillance of HNSCC.

Methods: This nationwide cross-sectional survey was conducted among head and neck surgeon members of the Saudi Society of Otolaryngology from June 1 to June 30, 2020. A predesigned questionnaire was used for data collection, and a descriptive analysis was carried out.

Results: This study included 22 participants and had a 78% (22/28) response rate. This study found that the majority of participants (9/22, 41%) used lung radiography for routine lung screening during posttreatment follow-ups, whereas 32% (7/22) used low-dose computed tomography (CT; 7/22, 32%). With regard to the number of years for which participants perform lung screening during follow-ups, the majority of participants (17/22, 77%) reported 5 years, and only 9% (2/22) have performed lifelong lung screening. With regard to the frequency of lung screening, 77% (17/22) of participants conduct screening annually, 18% (4/22) conduct screening half-yearly, and 5% (1/22) conduct screening biennially. With regard to beliefs about the effectiveness of screening procedures in reducing lung cancer mortality rates during follow-ups, 36% (8/22) of participants believed them to be very effective or somewhat effective, 18% (4/22) did not know, and only 9% (2/22) believed that they were not effective.

Conclusions: The participants mainly used lung radiography (9/22, 41%), low-dose CT (7/22, 32%), or positron emission tomography/CT (6/22, 27%) as a routine lung screening method during the posttreatment follow-up of patients with head and neck cancer for 5 years (17/22, 77%) or 10 years (3/22, 14%), and only a small percentage of participants have performed lifelong lung screening (2/22, 9%). Lung screening was mostly conducted annually or half-yearly. Such screening was believed to be very effective or somewhat effective.

(Interact J Med Res 2022;11(1):e24592) doi:10.2196/24592

KEYWORDS

squamous cell carcinoma of head and neck; lung neoplasms; radiography; otolaryngology; surgeons; survey

Introduction

Distant head and neck squamous cell carcinoma (HNSCC) metastases are discovered at various sites, but they most often occur in the lungs (66%-83% of HNSCC cases) and commonly occur in bones (22%-31% of HNSCC cases) and in the liver (6%-10% of HNSCC cases) [1]. The reported prevalence rates of HNSCC metastases in different clinical studies vary from 4% to 26%, and those reported in postmortem research vary from 37% to 57% [2].

There are many variables that impact the growth of distant metastases, such as the primary site, histological differentiation, patients' immunological abilities, advanced tumor stages, locoregional primary tumor control, and the extracapsular metastasis of the lymph nodes [3].

Metastasis is the natural evolution of primary tumors in patients with advanced HNSCCs who are not undergoing locoregional primary tumor control. Both distant metastases and secondary primary carcinomas may grow over time in curatively treated patients. This patient group may benefit from undergoing follow-ups after receiving therapy for the main HNSCC if the secondary cancer has been cured [4].

Screening for distant metastases and secondary primary tumors in the lungs is helpful, as it allows clinicians to make prognoses and allows for adapted patient counseling. Further, such screening has a beneficial effect on the prognoses of patients when it results in the early detection of distant metastases and secondary primary tumors [5].

Pulmonary follow-ups for secondary lesion identification can be performed in many ways, including via chest x-rays, computed tomography (CT) scans, positron emission tomography (PET) scans, bronchoscopy, brushes, and cytology [6].

A relatively recent field of research is the design of follow-up programs. Posttreatment follow-ups have been acquiring importance in clinical settings. The optimal form of surveillance is not clear, and there is a lack of data on the cost-effectiveness of rigorous monitoring. Despite the different timing protocols and the different modalities that are used among clinicians, the common denominator is the objective of promptly detecting and treating recurrent diseases as well as secondary primaries. Nowadays, multi-professional teams with skills in treatment toxicity management and prevention conduct follow-ups. The current follow-up methods include full head and neck clinical examination and structural examination [7-9].

In a previous study that was conducted by otolaryngology–head and neck surgeons, 26 out of 32 participants performed routine pulmonary screening, and of these 26 participants, 23 (88%) believed that chest radiography should be the preferred screening method. Most participants thought that mortality could be beneficially affected by lung screening. The most preferred modality for screening symptomatic patients was low-dose spiral CT (48%), followed by PET/CT (14%) and sputum cytology (14%). Additionally, 31% of respondents performed a chest x-ray for high-risk asymptomatic patients (current smokers, patients exposed to radiation, patients with a family history of cancer, and patients with advanced HNSCC). The same percentage of respondents conducted low-dose CT, while 19% relied on PET scans. Further, 19% of respondents did not screen any high-risk patients. Most respondents (77%) had more than 10 years of medical practice since graduating from medical school in the provinces of Quebec, Ontario, and Alberta [10].

Our cross-sectional survey study aimed to evaluate the current beliefs and pulmonary screening practices of otolaryngology–head and neck surgeons across Saudi Arabia with respect to the posttreatment surveillance of HNSCC. In this study, the findings of our survey were compared to the most recent data from the literature.

Methods

Study Design, Duration, and Participants

We used an analytical cross-sectional study design and collected data during the period from June 1 to June 30, 2020. Head and neck surgeon members of the Saudi Society of Otolaryngology in Saudi Arabia were surveyed in this study.

Inclusion and Exclusion Criteria

All head and neck surgeons of the Saudi Society of Otolaryngology who worked in Saudi Arabia hospitals and consented to participate in this study were included. There were no exclusion criteria.

Sample Size

The total sample consisted of all head and neck surgeons of the Saudi Society of Otolaryngology in Saudi Arabia. A total of 22 participants were included in this study.

Study Procedures

We adapted a questionnaire that consisted of 6 questions regarding actual practices and was previously designed and reviewed by Madana et al [10]. The questionnaire was used after obtaining permission from the main author. The questions inquired about the characteristics of routine lung screening during the posttreatment follow-up of patients with head and neck cancer [10]. The questionnaire was distributed to all head and neck surgeons of the Saudi Society of Otolaryngology in Saudi Arabia. No translation was needed, as the distributed form was written in the English language.

Data Management and Statistical Analysis

We used SPSS version 26 (IBM Corporation) to analyze the study data. Descriptive statistics were used to present the frequencies and percentages of the categorical variables.

Ethical Considerations

We prepared an informed consent form for the participants and gave them a brief description of this study's rationale and objectives. Afterward, we asked them to sign the consent form.

The anonymity and confidentiality of data were maintained throughout the study. Records were retained in a password-protected computer, and they will be retained for at least 7 years. There were no conflicts of interest. The study was approved by the Unit of Biomedical Ethics Research Committee at King Abdulaziz University.

Results

With regard to the methods of routine lung screening that were used during the posttreatment follow-up of patients with head and neck cancer, our study found that the majority of participants (9/22, 41%) used lung radiography, whereas 32% (7/22) used low-dose CT and 27% (6/22) used PET/CT. With regard to the number of years for which physicians perform lung screening for head and neck cancer during follow-ups, the majority of participants (17/22, 77%) reported 5 years and 14% (3/22) reported 10 years; only 9% (2/22) have performed lifelong lung screening. With regard to the frequency of lung screening, 77% (17/22) of participants conduct screening annually, 18% (4/22) conduct screening half-yearly, and 5% (1/22) conduct screening biennially. With regard to the believed effectiveness of the screening procedures (ie, those listed in question 1) in reducing lung cancer mortality rates during the follow-up of patients with head and neck cancer, 36% (8/22) of participants believed them to be very effective or somewhat effective, 18% (4/22) did not know, and only 9% (2/22) believed that they were not effective (Table 1).

Table 1. Characteristics of routine lung screening during the posttreatment follow-up of patients with head and neck cancer (respondents: N=22).

Parameters	Respondents, n (%)
Methods of routine lung screening during the posttreatment follow-up of patients with head and neck cancer	
Low-dose computed tomography	7 (32)
Lung radiography	9 (41)
Positron emission tomography/computed tomography	6 (27)
Types of patients with head and neck cancer who underwent routine lung screening during posttreatment follow-ups	
All patients	9 (41)
Only high-risk patients (smokers, patients exposed to radiation, patients with a family history of cancer, and patients with advanced HNSCC[a])	9 (41)
Only symptomatic patients	4 (18)
Number of years for which physicians perform lung screening for head and neck cancer during follow-ups	
10 years	3 (14)
5 years	17 (77)
Lifelong	2 (9)
Physicians' frequency of conducting lung screening for head and neck cancer during follow-ups	
Annually	17 (77)
Biennially	1 (5)
Half-yearly	4 (18)
Believed effectiveness of the screening procedures (ie, those listed in question 1) in reducing lung cancer mortality rates during the follow-up of patients with head and neck cancer	
Did not know	4 (18)
Not effective	2 (9)
Somewhat effective	8 (36)
Very effective	8 (36)
Have any of the patients during the past 12 months inquired about lung screening?	
No	11 (50)
Yes	11 (50)
Number of years of clinical head and neck practice and number of years since graduation from medical school	
0-5	1 (5)
11-20	9 (41)
6-10	6 (27)
>20	6 (27)
Practicing census region	
Asir	1 (5)
Dammam	2 (9)
Jeddah	8 (36)
Jazan	1 (5)
Mecca	2 (9)
Riyadh	7 (32)
Ta'if	1 (5)
Patient volume during a typical week of head and neck practice (number of patients/week)	
20-50	11 (50)
50-75	3 (14)

Parameters	Respondents, n (%)
75-100	1 (5)
<20	7 (32)

[a]HNSCC: head and neck squamous cell carcinoma.

Discussion

Principal Findings

A high response rate among otolaryngology–head and neck surgeons across Saudi Arabia was achieved in our study (22/28, 78%). This shows a high level of interest in postoperative screening practices.

Head and neck cancer refers to a group of malignant neoplastic lesions that have similar biological behaviors and are found in the upper aerodigestive tract. Head and neck cancer is the sixth most common cancer in the world; each year, over 500,000 new cases are diagnosed and 200,000 related deaths occur [11,12]. The most common sites of distant metastases are the lungs, the skeletal system, and the liver [13]. Due to the high incidence rate of metastasis (90% of cases) in patients with HNSCC, the posttreatment examination of the pulmonary region is critical [14,15]. Patients with HNSCC need posttreatment care that does not end with the completion of definitive treatment.

Our nationwide survey was conducted among head and neck surgeon members of the Saudi Society of Otolaryngology. The purpose of this survey study was to assess otolaryngology–head and neck surgeons' current beliefs and pulmonary screening practices with respect to the posttreatment surveillance of HNSCC in Saudi Arabia.

Regrettably, there is no consensus in the literature on the frequency and mode of posttreatment follow-up. Different investigational modalities each have their own set of advantages and disadvantages [16].

Similar to our results, another study, which was conducted by the Canadian Society of Otolaryngology to evaluate head and neck surgeons, reported that the majority of respondents performed routine lung screening and preferred chest radiography over low-dose CT or PET [10]. There is evidence however that PET/CT may be the most sensitive of these modalities, but further research is needed to show improvements in patient outcomes [16]. According to the Centers for Disease Control and Prevention, the most recommended method for lung cancer screening is low-dose CT [17]. Additionally, the present guidelines of the US Preventive Services Task Force suggest using the same method [18]. The reason why physicians prefer to avoid low-dose CT in follow-ups that are conducted after the treatment of head and neck cancer may be the modality's low specificity. The overdiagnosis of lung cancer was reported in more than 18% of cancer cases during the screening process of the National Lung Screening Trial [19]. Depending on radiography however cannot be the correct decision to make since, in another study, radiography was proven to be a poor method for diagnosing lung tumors in more than 65% of patients with cancer, and these patients were later diagnosed with pulmonary cancer [20]. Thus, the most recent findings among physicians must be disseminated more frequently.

With regard to the number of years for which participants perform lung screening during follow-ups, the majority of participants (17/22, 77%) reported 5 years, and only 9% (2/22) have performed lifelong lung screening. The study conducted in Canada found that 60% of their respondents conduct lung screening for 5 years, some of their respondents conduct lifelong lung screening, and the fewest number of participants conduct lung screening for 10 years [10]. This difference in the number of years for which physicians perform lung screening can be attributed to the variations in the current evidence concerning the posttreatment follow-up of patients with head and neck cancer; some studies have suggested that physicians should continue to conduct follow-ups once per year after 5 posttreatment years [21]. However, there is little evidence that supports the effectiveness of conducting follow-ups for more than 5 years [22].

The British Association of Head and Neck Oncologists [23] recommends 4- to 6-week follow-up visits for the first 2 years after treatment, 3-month follow-up visits for the third posttreatment year, 6-month follow-up visits for the fourth and fifth posttreatment years, and annual visits after that. With regard to the frequency of lung screening, in our study, the majority of participants (17/22, 77%) conduct screening annually, 18% (4/22) conduct screening half-yearly, and 5% (1/22) conduct screening biennially.

The Canadian study reported that most respondents were screening their patients annually, while less than 15% screened patients biennially or half-yearly [10]. Variations in the number of follow-up visits have also been evident in the guidelines present in the literature and can account for the differences between Canada and Saudi Arabia, since clinicians from different countries can follow different guidelines [9]. The Saudi head and neck surgeons in our study stated that conducting scheduled visits is also the best way to provide adequate follow-ups to patients with HNSCC, since these follow-ups address many concerns and not just the early detection of recurrence or secondary primaries.

Strengths and Limitations

To date, no study has been conducted in Saudi Arabia to analyze the current practices of head and neck surgeons with respect to detecting post-HNSCC pulmonary metastasis. Our study provides a highly comprehensive view of current practices, given that all certified head and neck surgeons of the Saudi Society of Otolaryngology in Saudi Arabia participated in this study. Despite its exploratory nature, this study offers some insight into the lack of evidence-based practices for the posttreatment pulmonary surveillance of HNSCC. Conducting well-controlled trials to evaluate different modalities of

surveillance for different subtypes and stages of HNSCC would shed light on the survival rates associated with and the cost-effectiveness of these modalities. However, we are limited by the cross-sectional nature of this study. We are also limited by the lack of literature on this topic and, hence, our inability to obtain enough confidence in our results.

Conclusion and Recommendations

The participants mainly used lung radiography (9/22, 41%), low-dose CT (7/22, 32%), or PET/CT (6/22, 27%) as a routine lung screening method during the posttreatment follow-up of patients with head and neck cancer for 5 years (17/22, 77%) or 10 years (3/22, 14%), and only a small percentage of participants have performed lifelong lung screening (2/22, 9%). Lung screening was mostly conducted annually or half-yearly. Such screening was believed to be very effective or somewhat effective. However, controversy still exists due to the lack of evidence-based protocols worldwide. Therefore, future research should explore the importance of this subject by using a more comprehensive methodology and enrolling patients with HNSCC in comparative studies. We also recommend conducting further follow-up studies to obtain more knowledge on the effects of the positions that physicians hold.

Conflicts of Interest

None declared.

References

1. Ferlito A, Shaha AR, Silver CE, Rinaldo A, Mondin V. Incidence and sites of distant metastases from head and neck cancer. ORL J Otorhinolaryngol Relat Spec 2001;63(4):202-207. [doi: 10.1159/000055740] [Medline: 11408812]
2. León X, Quer M, Orús C, del Prado Venegas M, López M. Distant metastases in head and neck cancer patients who achieved loco-regional control. Head Neck 2000 Oct;22(7):680-686. [doi: 10.1002/1097-0347(200010)22:7<680::aid-hed7>3.0.co;2-j] [Medline: 11002323]
3. Shaha AR, Hoover EL, Mitrani M, Marti JR, Krespi YP. Synchronicity, multicentricity, and metachronicity of head and neck cancer. Head Neck Surg 1988;10(4):225-228. [doi: 10.1002/j.1930-2398.1988.tb00003.x] [Medline: 3235352]
4. Leong SCL, Javed F, Elliot S, Mortimore S. Effectiveness of x-ray and computed tomography screening for assessing pulmonary involvement in patients with head and neck squamous cell carcinoma. J Laryngol Otol 2008 Sep;122(9):961-966. [doi: 10.1017/S0022215107001296] [Medline: 18177532]
5. Roland NJ, Bradley PJ. The role of surgery in the palliation of head and neck cancer. Curr Opin Otolaryngol Head Neck Surg 2014 Apr;22(2):101-108. [doi: 10.1097/MOO.0000000000000031] [Medline: 24419405]
6. Gourin CG, Watts TL, Williams HT, Patel VS, Bilodeau PA, Coleman TA. Identification of distant metastases with positron-emission tomography-computed tomography in patients with previously untreated head and neck cancer. Laryngoscope 2008 Apr;118(4):671-675. [doi: 10.1097/MLG.0b013e3181625737] [Medline: 18197136]
7. Pignon JP, le Maître A, Maillard E, Bourhis J, MACH-NC Collaborative Group. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. Radiother Oncol 2009 Jul;92(1):4-14. [doi: 10.1016/j.radonc.2009.04.014] [Medline: 19446902]
8. Kazi R, Manikanthan K, Pathak KA, Dwivedi RC. Head and neck squamous cell cancers: need for an organised time-bound surveillance plan. Eur Arch Otorhinolaryngol 2010 Dec;267(12):1969-1971. [doi: 10.1007/s00405-010-1377-x] [Medline: 20835832]
9. De Felice F, Musio D, Tombolini V. Follow-up in head and neck cancer: A management dilemma. Adv Otolaryngol 2015 Jan 14;2015:1-4 [FREE Full text] [doi: 10.1155/2015/703450]
10. Madana J, Morand GB, Barona-Lleo L, Black MJ, Mlynarek AM, Hier MP. A survey on pulmonary screening practices among otolaryngology-head & neck surgeons across Canada in the post treatment surveillance of head and neck squamous cell carcinoma. J Otolaryngol Head Neck Surg 2015 Feb 04;44(1):5 [FREE Full text] [doi: 10.1186/s40463-015-0057-7] [Medline: 25649793]
11. Watkinson J, Gaze M, Wilson J. Stell and Maran's Head and Neck Surgery, 4Ed. Boca Raton, Florida: CRC Press; Jun 05, 2000.
12. Robinson KL, Macfarlane GJ. Oropharyngeal cancer incidence and mortality in Scotland: are rates still increasing? Oral Oncol 2003 Jan;39(1):31-36. [doi: 10.1016/s1368-8375(02)00014-3] [Medline: 12457719]
13. Dennington ML, Carter DR, Meyers AD. Distant metastases in head and neck epidermoid carcinoma. Laryngoscope 1980 Feb;90(2):196-201. [doi: 10.1288/00005537-198002000-00002] [Medline: 7354687]
14. de Bree R, Deurloo EE, Snow GB, Leemans CR. Screening for distant metastases in patients with head and neck cancer. Laryngoscope 2000 Mar;110(3 Pt 1):397-401. [doi: 10.1097/00005537-200003000-00012] [Medline: 10718426]
15. Zbären P, Lehmann W. Frequency and sites of distant metastases in head and neck squamous cell carcinoma. An analysis of 101 cases at autopsy. Arch Otolaryngol Head Neck Surg 1987 Jul;113(7):762-764. [doi: 10.1001/archotol.1987.01860070076020] [Medline: 3580159]
16. Zhao X, Rao S. Surveillance imaging following treatment of head and neck cancer. Semin Oncol 2017 Oct;44(5):323-329. [doi: 10.1053/j.seminoncol.2018.01.010] [Medline: 29580434]

17. Who should be screened for lung cancer? Centers for Disease Control and Prevention. URL: https://www.cdc.gov/cancer/lung/basic_info/screening.htm [accessed 2021-11-17]
18. US Preventive Services Task Force, Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, et al. Screening for lung cancer: US Preventive Services Task Force recommendation statement. JAMA 2021 Mar 09;325(10):962-970. [doi: 10.1001/jama.2021.1117] [Medline: 33687470]
19. Patz Jr EF, Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemägi MC, NLST Overdiagnosis Manuscript Writing Team. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med 2014 Feb 01;174(2):269-274 [FREE Full text] [doi: 10.1001/jamainternmed.2013.12738] [Medline: 24322569]
20. Shah SI, Applebaum EL. Lung cancer after head and neck cancer: role of chest radiography. Laryngoscope 2000 Dec;110(12):2033-2036. [doi: 10.1097/00005537-200012000-00010] [Medline: 11129015]
21. National Comprehensive Cancer Network, Forastiere AA, Ang KK, Brizel D, Brockstein BE, Burtness BA, et al. Head and neck cancers. J Natl Compr Canc Netw 2008 Aug;6(7):646-695. [doi: 10.6004/jnccn.2008.0051] [Medline: 18691457]
22. Merkx MAW, van Gulick JJM, Marres HAM, Kaanders JHAM, Bruaset I, Verbeek A, et al. Effectiveness of routine follow-up of patients treated for T1-2N0 oral squamous cell carcinomas of the floor of mouth and tongue. Head Neck 2006 Jan;28(1):1-7. [doi: 10.1002/hed.20296] [Medline: 16155911]
23. British Association of Head and Neck Oncologists. Practice care guidance for clinicians participating in the management of head and neck cancer patients in the UK. Drawn up by a Consensus Group of Practising Clinicians. Eur J Surg Oncol 2001 Jun;27 Suppl A:S1-S17. [doi: 10.1053/ejso.2000.1090] [Medline: 11583107]

Abbreviations

CT: computed tomography
HNSCC: head and neck squamous cell carcinoma
PET: positron emission tomography

Edited by G Eysenbach; submitted 25.09.20; peer-reviewed by F Gomez, C Phellas; comments to author 10.01.21; revised version received 07.05.21; accepted 31.05.21; published 18.03.22.

Please cite as:
Alnefaie M, Alamri A, Saeedi A, Althobaiti A, Alosaimi S, Alqurashi Y, Marzouki H, Merdad M
Pulmonary Screening Practices of Otolaryngology–Head and Neck Surgeons Across Saudi Arabia in the Posttreatment Surveillance of Squamous Cell Carcinoma: Cross-sectional Survey Study
Interact J Med Res 2022;11(1):e24592
URL: https://www.i-jmr.org/2022/1/e24592
doi:10.2196/24592
PMID:35302511

Original Paper

A Machine Learning Approach to Predict the Outcome of Urinary Calculi Treatment Using Shock Wave Lithotripsy: Model Development and Validation Study

Reihaneh Moghisi[1*], MSci; Christo El Morr[2*], PhD; Kenneth T Pace[3,4], MD, PhD; Mohammad Hajiha[5], MD; Jimmy Huang[1], PhD

[1]School of Information Technology, York University, Toronto, ON, Canada
[2]School of Health Policy and Management, York University, Toronto, ON, Canada
[3]Division of Urology, St Michael's Hospital, Toronto, ON, Canada
[4]Department of Surgery, University of Toronto, Toronto, ON, Canada
[5]Department of Urology, Loma Linda University Health, Loma Linda, CA, United States
[*]these authors contributed equally

Corresponding Author:
Reihaneh Moghisi, MSci
School of Information Technology
York University
4700 Keele St
Toronto, ON, M3J 1P3
Canada
Phone: 1 416 736 2100
Email: rmoghisi88@gmail.com

Abstract

Background: Shock wave lithotripsy (SWL), ureteroscopy, and percutaneous nephrolithotomy are established treatments for renal stones. Historically, SWL has been a predominant and commonly used procedure for treating upper tract renal stones smaller than 20 mm in diameter due to its noninvasive nature. However, the reported failure rate of SWL after one treatment session ranges from 30% to 89%. The failure rate can be reduced by identifying candidates likely to benefit from SWL and manage patients who are likely to fail SWL with other treatment modalities. This would enhance and optimize treatment results for SWL candidates.

Objective: We proposed to develop a machine learning model that can predict SWL outcomes to assist practitioners in the decision-making process when considering patients for stone treatment.

Methods: A data set including 58,349 SWL procedures performed during 31,569 patient visits for SWL to a single hospital between 1990 and 2016 was used to construct and validate the predictive model. The AdaBoost algorithm was applied to a data set with 17 predictive attributes related to patient demographics and stone characteristics, with success or failure as an outcome. The AdaBoost algorithm was also applied to a training data set. The generated model's performance was compared to that of 5 other machine learning algorithms, namely C4.5 decision tree, naïve Bayes, Bayesian network, K-nearest neighbors, and multilayer perceptron.

Results: The developed model was validated with a testing data set and performed significantly better than the models generated by the other 5 predictive algorithms. The sensitivity and specificity of the model were 0.875 and 0.653, respectively, while its positive predictive value was 0.7159 and negative predictive value was 0.839. The C-statistics of the receiver operating characteristic (ROC) analysis was 0.843, which reflects an excellent test.

Conclusions: We have developed a rigorous machine learning model to assist physicians and decision-makers to choose patients with renal stones who are most likely to have successful SWL treatment based on their demographics and stone characteristics. The proposed machine learning model can assist physicians and decision-makers in planning for SWL treatment and allow for more effective use of limited health care resources and improve patient prognoses.

(Interact J Med Res 2022;11(1):e33357) doi:10.2196/33357

KEYWORDS

lithotripsy; urolithiasis; machine learning; treatment outcome; ensemble learning; AdaBoost; renal stones; kidney disease

Introduction

Urinary stone disease, also known as urolithiasis, is a disease that occurs when a solid particle of minerals and salts is formed inside the urinary tract. A recent systematic review suggests an increasing prevalence of urolithiasis in North America over the past 3 decades [1]. In Canada, urinary stone disease is prevalent with a lifetime risk of 10% among both men and women, whereas there is a 75% chance of recurrence in 20 years after initial treatment [2].

Historically, shock wave lithotripsy (SWL) has been the most used procedure for treating upper tract urolithiasis and stones smaller than 20 mm in diameter due to its noninvasive nature, lower cost, fewer side effects, and faster recovery [3,4].

In Ontario, Canada, SWL is a regionalized and limited resource. St. Michael's Hospital in downtown Toronto is one of the only 3 centers in the province offering this service. Wait time to access SWL treatment in Canada ranges from 1 day to 1 year, with a mean wait time of 8.4 weeks in Ottawa and 8 weeks in Toronto [5]. Considering the intolerability of the pain associated with stone disease and long wait times, some patients opt for more invasive therapies such as ureteroscopy to gain access to faster treatment.

While SWL is the predominant treatment, the reported failure rate of SWL after the first session ranges from 30% to 89% [6-8]. The failure rate can be reduced significantly by identifying the candidates who are most likely to benefit from SWL, which would optimize treatment results for SWL candidates and allow for the most effective use of limited medical resources.

To identify the predictive factors of SWL outcome, several studies have focused on statistical analyses of patient characteristics using bivariate and/or multivariate analysis [4,9-11]. The advantage and strength of machine learning is its ability to synthesize complex combinations of various attributes [12,13]. Our objective for this study was to construct a robust machine learning model that can predict SWL results to assist practitioners in their decision-making.

Methods

Ethics Approval

This study received ethics approval from the Office of Research Ethics at York University (certificate number STU 2019-139) and St. Michael Research Ethics Board (approval number 16-167).

Data Set

We assessed a data set of patients aged ≥18 years receiving SWL treatment at St. Michael's Hospital between 1998 and 2016. The data set comprised the records of 37,013 patients.

We excluded the data of patients with special conditions (eg, staghorn calculi, horseshoe kidney, caliceal diverticula, duplex collecting systems, solitary kidneys, musculoskeletal deformities) and stones larger than 25 mm in diameter. The remaining data set consisted of 57,485 SWL procedures that were performed on 31,569 patients during this period, which were used as a training data set to build the model. Several factors can impact SWL treatment outcome, including stone location and age; the choice of the attributes was guided by input from clinical experts and a literature review [4,10,11,14]. We retained 17 attributes that were most relevant to SWL success and were available in our database (Table 1).

Table 1. Training set attributes and corresponding values.

Attribute	Value
Kidney side	Left or right
Electrode used	Integer (1 to 3)
Stone treatment number	Integer
Number of shocks	Integer
Stone locations	Lower calyx, lower ureter, middle calyx, middle ureter, pelvis, upper calyx, upper ureter, ureterovesical junction, renal pelvis
Area of stone	Integer (mm^2)
Gender	Female or male
BMI	Real number (kg/m^2)
Age	18-95
Number of stones	Integer
Family history	True or false
Asymptomatic	True or false
Stent insertion	True or false
Shock frequency	120, 90, 60
Antibiotic	True or false
Shock maximum voltage	Integer
Lithotripter models	Dornier MFL 5000, Philips LithoTron, Storz Modulith SLX-F2
Outcome	Success or failure

Defining Success and Failure of SWL on the Training Data Set

The failure or success of SWL in the training data set was based on whether there was a retreatment plan for the same patient and same stone within 90 days after initial treatment or not. The effectiveness of the lithotripter machine was measured by success rates on the training set.

Ensemble Learning Technique

To predict the treatment outcome for SWL candidates, we used the AdaBoost algorithm based on the ensemble learning method, a machine learning technique that combines several base classifiers in various formats to produce a more robust and optimal classification model. Compared to other conventional machine learning algorithms, ensemble learning techniques are more stable, faster, simpler, and easier to program [15-19].

AdaBoost combines multiple weak classifiers that are sequentially applied to the data set. In each iteration, after the weak classifier is called, misclassified item sets are detected and given higher weight to increase the emphasis of the weak classifier on them in the next round. The final classification model is then generated as a linear combination of these weak classifiers with their assigned weights as their coefficient [19]. We used 10-fold cross-validation for AdaBoost.

Performance Evaluation

To compare AdaBoost's performance to that of other classifiers, we used 5 classification algorithms to predict SWL failure (retreatment required <3 months), namely C4.5, naïve Bayes, Bayesian network, K-nearest neighbors, and multilayer perceptron, and used *t* tests to perform pairwise comparisons of the performance of the AdaBoost algorithm against that of the other 5 classification models. The measurements used to determine the models' performance were sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) [20], accuracy, F1 score [14,21], and Matthews correlation coefficient [22]. Machine learning was performed using WEKA (version 3.9; University of Waikato) [23]. We used 10-fold cross-validation for performance evaluation.

Generalizability of the Model

Classifiers were assessed for generalizability using the testing data set of 864 patients who had their preoperative and postoperative follow-ups conducted at the same center, and whose SWL procedure success and failure was determined by computed tomography (CT) scan of patients 3 months after the initial therapy. The testing data set was not included in the training set used to build the model. We employed the undersampling technique to resolve the imbalance in data by removing random examples from the majority class. *SpreadSubsample* was the Java class implemented for subsampling the original training set. We matched the ratio of success to failure in the training set to the ratio observed in our testing set, which was 40% to 60%.

Results

The AdaBoost Model

A total of 30 iterations were used for the AdaBoost model. Although increasing the number of iterations usually increases the accuracy of the model, we ceased adding more iterations to the model to avoid overfitting.

Research has shown that applying the boosting method to any weak classifier can drastically enhance the accuracy of the classification model [24]. Indeed, the accuracy of applying the base learner (Decision Stump) alone on our data set was 67.8%. However, with the ensemble method, we could boost this accuracy by 9% to 76.38%, which demonstrates the superiority of the boosting method.

Model Performance

Table 2 shows the comparison of the AdaBoost model against the other 5 classification techniques in terms of 4 different performance measurements. AdaBoost performed significantly better than all 5 other classifiers on all performance measures.

Table 2. Performance comparison of AdaBoost against 5 other classifiers.

Measurement	AdaBoostM1	C4.5	Naïve Bayes	Multilayer perceptron	Bayesian network	KNN[a]
Accuracy	77.59	75.26[b]	75.82[b]	69.11[b]	76.49[b]	57.52[b]
MCC[c]	0.53	0.46[b]	0.47[b]	0.34[b]	0.49[b]	0.09[b]
F1 score	0.84	0.82[b]	0.83[b]	0.76[b]	0.83[b]	0.66[b]
Area under ROC[d]	0.80	0.74[b]	0.75[b]	0.74[b]	0.78[b]	0.54[b]

[a]KNN: K-nearest neighbors.

[b]Statistically significant.

[c]MCC: Matthews correlation coefficient.

[d]ROC: receiver operating characteristic.

The sensitivity of the model was 0.875 (ie, 87.5% of all patients with successful SWL treatment were correctly identified by our model). On the other hand, the specificity was 0.6528 (ie, 65.3% of all patients with failed SWL treatment were correctly identified by our model).

Furthermore, the PPV (ie, the probability that subjects with a success prediction truly succeeded in the treatment) was 0.7159. Meanwhile, the NPV (ie, the probability that subjects with a failure prediction have truly failed the treatment) was 0.839.

Finally, we measured the correlation between the attributes and the class; the top 5 contributors detected were the number of stones, the area of the stone, the stone treatment number, the lithotripter machine, and the patient's age.

Discussion

Principal Findings

Our goal was to evaluate the ability of machine learning techniques to assist in effective decision-making for the treatment of urolithiasis with SWL by accurately predicting the SWL results. We have shown that AdaBoost provided superior prediction ability compared to 5 other classification techniques.

The AUC (area under the ROC [receiver operating characteristic] curve or C-statistic) of the ROC analysis for our prediction model was 0.843, which reflects an excellent test (a C-statistic value of 0.8-0.89 indicates an excellent test, 0.7-0.79 indicates a good test, and 0.51-0.69 indicates a poor test) [25].

The model had high sensitivity and medium specificity. Given that we are interested in identifying the patients for whom SWL has a low chance of success to plan for alternative procedures, the NPV of 0.839 demonstrated that the model can predict with high probability if a subject will fail the treatment. Considering how scarce and expensive health care resources are, it is important to allocate those limited resources appropriately [26,27]; our model allows for appropriate allocation by informing physicians about patients who are not likely to benefit from SWL.

Recently, Choo et al [28] developed a decision tree algorithm C 5.0 for the same purpose of predicting treatment outcomes for SWL, including 15 predictive attributes on only 791 patients. Although their model had high accuracy (92.3%), some of its branches included fewer than 10 patients each. Considering that our AdaBoost-based model outperformed the decision tree algorithm in all performance measurements, we can expect it to yield better accuracy if other predictive attributes (ie, skin-to-stone distance, stone Hounsfield unit, creatinine level, stone composition, etc [3,4]) were included in the data set in a future study.

Our results show that the 3 different models of lithotripters did not significantly change the SWL treatment success rate (P=.81). This finding suggests that frequently upgrading the technology of SWL machines does not necessarily result in a better outcome, whereas optimizing patient and stone selection is a more important factor in predicting the outcome of the SWL.

Limitations

A limitation of this study was the lack of follow-up data for some of the patients enrolled. As a result, a treatment's failure was defined only based on having retreatment of a stone in the same center (St. Michael's Hospital) within 3 months of the initial SWL. However, to overcome this limitation and test the robustness of our model, we used 864 records that included only patients who had their complete preoperative and

postoperative follow-ups conducted at St. Michael's Hospital. This subset of the data set was not used for training the model. The follow-up data, the stone-free rate, and the success of treatment for these patients were assessed based on the follow-up CT scan administered at St. Michael's Hospital 3 months after the initial SWL.

Another limitation is that some attributes that have been shown to be predictive of SWL outcome in recent studies, such as stone density, skin-to-stone distance, and stone composition [10], were not available in our database since these data points were not known or collected 20 years ago.

Conclusion

We built a machine learning model to assist physicians and decision-makers to choose the best treatment option for SWL candidates based on their demographics and stone characteristics, which can result in improved prognoses. The model was generated based on the AdaBoost algorithm.

A pairwise comparison was performed between the AdaBoost classifier and 5 other classification techniques in terms of their accuracy, Matthews correlation coefficient, area under the ROC curve, and root mean squared error. The findings of these comparisons suggest the superiority of AdaBoost compared to those algorithms.

We aim to explore several meaningful research directions in the future. First, we will develop new models and architectures that are more robust and efficient by utilizing deep learning techniques. Second, our proposed ensemble learning approach can be applied to more comprehensive databases for more applications to ascertain the applicability of the model [29-32].

Acknowledgments

We would like to thank Daniela Ghiculete for her continued support in establishing the research protocol, accessing the database, and assisting in getting the Research Ethics Board approval.

This research is supported by the Natural Sciences and Engineering Research Council of Canada, the York Research Chairs program, and the Faculty of Health at York University.

Data Availability

The data sets generated and/or analyzed in this study are not publicly available due to patient data privacy reasons but are available from the corresponding author upon reasonable request.

Authors' Contributions

RM preprocessed and compiled raw data, built the model, assessed the performance, and put in writing the findings. CE redacted, edited, and constructed the structure and content of the paper and study design, and assisted in the evaluation of the analysis. He gave continuous counsel and advised about the research objective. KTP provided support and training on the medical aspects of the research objective and contributed to the development of the research protocol and design. MH assisted and contributed to editing the draft, study design, and literature review. JH provided support and training for the study design, editing of the draft, and literature review.

Conflicts of Interest

None declared.

References

1. Chen Z, Prosperi M, Bird VY. Prevalence of kidney stones in the USA: The National Health and Nutrition Evaluation Survey. J Clin Urol 2018 Nov 26;12(4):296-302. [doi: 10.1177/2051415818813820]
2. Stamatelou KK, Francis ME, Jones CA, Nyberg LM, Curhan GC. Time trends in reported prevalence of kidney stones in the United States: 1976-1994. Kidney Int 2003 May;63(5):1817-1823 [FREE Full text] [doi: 10.1046/j.1523-1755.2003.00917.x] [Medline: 12675858]
3. El-Assmy A, El-Nahas AR, Abo-Elghar ME, Eraky I, El-Kenawy MR, Sheir KZ. Predictors of Success after Extracorporeal Shock Wave Lithotripsy (ESWL) for Renal Calculi Between 20–30 mm: A Multivariate Analysis Model. TSW Urology 2006 Mar 23;1:93-100. [doi: 10.1100/tswurol.2006.74]
4. Takahara K, Ibuki N, Inamoto T, Nomi H, Ubai T, Azuma H. Predictors of success for stone fragmentation and stone-free rate after extracorporeal shockwave lithotripsy in the treatment of upper urinary tract stones. Urol J 2012;9(3):549-552. [doi: 10.22037/uj.v9i3.1590]
5. Piggott K, Bell C. Looking for Lithotripsy: Accessibility and Portability of Canadian Healthcare. Hcpol 2013 Nov 27;9(2):65-75. [doi: 10.12927/hcpol.2013.23605]
6. Abe T, Akakura K, Kawaguchi M, Ueda T, Ichikawa T, Ito H, et al. Outcomes of shockwave lithotripsy for upper urinary-tract stones: a large-scale study at a single institution. J Endourol 2005 Sep;19(7):768-773. [doi: 10.1089/end.2005.19.768] [Medline: 16190825]

7. Park H, Park M, Park T. Two-year experience with ureteral stones: extracorporeal shockwave lithotripsy v ureteroscopic manipulation. J Endourol 1998 Dec;12(6):501-504. [doi: 10.1089/end.1998.12.501] [Medline: 9895251]
8. Srisubat A, Potisat S, Lojanapiwat B, Setthawong V, Laopaiboon M. Extracorporeal shock wave lithotripsy (ESWL) versus percutaneous nephrolithotomy (PCNL) or retrograde intrarenal surgery (RIRS) for kidney stones. Cochrane Database Syst Rev 2014 Nov 24(11):CD007044. [doi: 10.1002/14651858.CD007044.pub3] [Medline: 25418417]
9. El-Assmy A, El-Nahas AR, Abou-El-Ghar ME, Awad BA, Sheir KZ. Kidney stone size and hounsfield units predict successful shockwave lithotripsy in children. Urology 2013 Apr;81(4):880-884. [doi: 10.1016/j.urology.2012.12.012] [Medline: 23395121]
10. Gücük A, Uyetürk U. Usefulness of hounsfield unit and density in the assessment and treatment of urinary stones. World J Nephrol 2014 Nov 06;3(4):282-286 [FREE Full text] [doi: 10.5527/wjn.v3.i4.282] [Medline: 25374823]
11. Nakasato T, Morita J, Ogawa Y. Evaluation of Hounsfield Units as a predictive factor for the outcome of extracorporeal shock wave lithotripsy and stone composition. Urolithiasis 2015 Feb 20;43(1):69-75. [doi: 10.1007/s00240-014-0712-x] [Medline: 25139151]
12. El Morr C, Ali-Hassan H. Analytics in Healthcare: A Practical Introduction. Cham, Switzerland: Springer International Publishing; 2019.
13. Imbus JR, Randle RW, Pitt SC, Sippel RS, Schneider DF. Machine learning to identify multigland disease in primary hyperparathyroidism. J Surg Res 2017 Nov;219:173-179 [FREE Full text] [doi: 10.1016/j.jss.2017.05.117] [Medline: 29078878]
14. Bozzini G, Verze P, Arcaniolo D, Dal Piaz O, Buffi NM, Guazzoni G, et al. A prospective randomized comparison among SWL, PCNL and RIRS for lower calyceal stones less than 2 cm: a multicenter experience: A better understanding on the treatment options for lower pole stones. World J Urol 2017 Dec 5;35(12):1967-1975. [doi: 10.1007/s00345-017-2084-7] [Medline: 28875295]
15. Fern A, Givan R. Online Ensemble Learning: An Empirical Study. Machine Learning 2003;53:71-109. [doi: 10.1023/A:1025619426553]
16. Freund Y, Schapire R. Experiments with a new boosting algorithm. In: Proceedings of the Thirteenth International Conference on International Conference on Machine Learning. 1996 Presented at: ICML'96; July 3-6; Bari, Italy p. 148-156. [doi: 10.5555/3091696.3091715]
17. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). Ann Statist 2000 Apr 1;28(2). [doi: 10.1214/aos/1016218223]
18. Huang X, Huang Y, Wen M, An A, Liu Y, Poon J. Applying data mining to pseudo-relevance feedback for high performance text retrieval. 2006 Presented at: Sixth International Conference on Data Mining (ICDM'06); Dec 18-22; Hong Kong, China. [doi: 10.1109/icdm.2006.22]
19. Schapire R. Explaining AdaBoost. In: Schölkopf B, Luo Z, Vovk V, editors. Empirical Inference. Cham, Switzerland: Springer Berlin Heidelberg; 2013:37-52.
20. Qiu P. The Statistical Evaluation of Medical Tests for Classification and Prediction. J Am Stat Assoc 2005 Jun;100(470):705. [doi: 10.1198/jasa.2005.s19]
21. Thornley C. Advances in Information Retrieval. J Doc 2012 Aug 31;68(5). [doi: 10.1108/jd.2012.27868eaa.002]
22. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 2020 Jan 02;21(1):6 [FREE Full text] [doi: 10.1186/s12864-019-6413-7] [Medline: 31898477]
23. Frank E, Hall M, Holmes G, Kirkby R, Pfahringer B, Witten I. Weka-A Machine Learning Workbench for Data Mining. In: Data Mining and Knowledge Discovery Handbook. Cham, Switzerland: Springer; 2009:1269-1277.
24. Drucker H, Cortes C. Boosting Decision Trees. Advances in neural information processing systems. In: Proceedings of the 8th International Conference on Neural Information Processing Systems. 1996 Presented at: NIPS'95; Nov 27-Dec 2; Denver, Colorado p. 479-485.
25. Scott AJ, Hosmer DW, Lemeshow S. Applied Logistic Regression. Biometrics 1991 Dec;47(4):1632. [doi: 10.2307/2532419]
26. El Morr C, editor. Research Perspectives on the Role of Informatics in Health Policy and Management. Hershey, PA: IGI Global; 2014.
27. El Morr C. Introduction to Health Informatics: A Canadian Perspective. Toronto, ON: Canadian Scholars' Press; 2018.
28. Choo MS, Uhmn S, Kim JK, Han JH, Kim D, Kim J, et al. A Prediction Model Using Machine Learning Algorithm for Assessing Stone-Free Status after Single Session Shock Wave Lithotripsy to Treat Ureteral Stones. J Urol 2018 Dec;200(6):1371-1377. [doi: 10.1016/j.juro.2018.06.077] [Medline: 30036513]
29. Huang X, Zhong M, Si L. York University at TREC 2005: Genomics Track. 2005. URL: https://trec.nist.gov/pubs/trec14/papers/yorku-huang2.geo.pdf [accessed 2022-03-11]
30. Liu Y, An A, Huang X, editors. Boosting Prediction Accuracy on Imbalanced Datasets with SVM Ensembles. In: Advances in Knowledge Discovery and Data Mining. Berlin, Heidelberg: Springer; 2006.
31. Liu Y, Huang X, An A, Yu X. ARSA: a sentiment-aware model for predicting sales performance using blogs. In: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. 2007 Presented at: SIGIR '07; July 23-27; Amsterdam, The Netherlands p. 607-614. [doi: 10.1145/1277741.1277845]

32. Yu X, Liu Y, Huang X, An A. Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain. IEEE Trans Knowl Data Eng 2012 Apr;24(4):720-734. [doi: 10.1109/tkde.2010.269]

Abbreviations

AUC: area under the receiver operating characteristic curve
CT: computed tomography
NPV: negative predictive value
PPV: positive predictive value
ROC: receiver operating characteristic
SWL: shock wave lithotripsy

Edited by T Leung; submitted 03.09.21; peer-reviewed by CY Yang, J Kim; comments to author 13.12.21; revised version received 24.12.21; accepted 11.02.22; published 16.03.22.

Please cite as:
Moghisi R, El Morr C, Pace KT, Hajiha M, Huang J
A Machine Learning Approach to Predict the Outcome of Urinary Calculi Treatment Using Shock Wave Lithotripsy: Model Development and Validation Study
Interact J Med Res 2022;11(1):e33357
URL: https://www.i-jmr.org/2022/1/e33357
doi:10.2196/33357
PMID:35293872

Original Paper

Microbial Ecosystem Therapeutic-2 Intervention in People With Major Depressive Disorder and Generalized Anxiety Disorder: Phase 1, Open-Label Study

Arthi Chinna Meyyappan[1,2,3], BScH, MSc; Evan Forth[1,2,3], BScH; Roumen Milev[1,2,3], MD, PhD

[1]Department of Psychiatry, Queen's University, Kingston, ON, Canada

[2]Centre for Neuroscience Studies, Queen's University, Kingston, ON, Canada

[3]Providence Care Hospital, Kingston, ON, Canada

Corresponding Author:
Arthi Chinna Meyyappan, BScH, MSc
Department of Psychiatry
Queen's University
752 King St. West
Kingston, ON, K7L 4X3
Canada
Phone: 1 6135444900
Email: 14acm@queensu.ca

Abstract

Background: Recent studies have investigated the potential of treatments that modify the gut microbiome, such as fecal microbiota transplantation and probiotics, in individuals with psychiatric illnesses.

Objective: The aim of this study was to investigate the safety, tolerability, and efficacy of a novel gut microbiome therapeutic, Microbial Ecosystem Therapuetic-2 (MET-2), in people with depression and anxiety.

Methods: In this phase 1, open-label trial, 12 adults diagnosed with major depressive disorder, generalized anxiety disorder, or both were recruited. Over 8 weeks, participants consumed three capsules per day, orally, of an encapsulated microbial therapeutic (MET-2), which contained 40 strains of bacteria that were purified and lab-grown from the stool of a single healthy donor. Participants were assessed biweekly using clinical scales and questionnaires in order to evaluate the safety, efficacy, and tolerability of the therapeutic.

Results: The therapeutic was found to be generally safe and tolerable, with limited adverse events and side effects and no serious adverse events. Of the 12 individuals included in this study, 9 (75%) responded to treatment (50% improvement in Montgomery-Asberg Depression Rating Scale [MADRS] scores, 7-item Generalized Anxiety Disorder scale [GAD-7] scores, or both, from baseline to the week-8 visit). Over the course of 10 weeks, MET-2 significantly decreased mean MADRS and GAD-7 scores (MADRS: $F_{2.731,\ 30.05}$=8.784, P<.001; GAD-7: $F_{2.778,\ 30.55}$= 9.638, P<.001). Multiple comparisons with Bonferroni adjustments showed a significant reduction in MADRS scores from baseline (mean 19.00, SD 4.843) to week 6 (mean 11.25, SD 8.001; P=.009), week 8 (mean 8.667, SD 8.732; P=.002), and week 10 (mean 8.250, SD 9.304; P=.006). Multiple comparisons showed a significant reduction in GAD-7 scores from baseline (mean 13.58, SD 4.010) to week 4 (mean 9.167, SD 5.096; P=.03), week 6 (mean 7.667, SD 4.539; P=.004), week 8 (mean 7.333, SD 6.583; P=.03), and week 10 (mean 7.500, SD 6.448; P=.03).

Conclusions: The findings from this study are the first to provide evidence for the role of microbial ecosystem therapy in treating depression and anxiety. However, a double-blind, randomized controlled trial with a larger sample size is needed for more conclusive results.

Trial Registration: ClinicalTrials.gov NCT04052451; https://www.clinicaltrials.gov/ct2/show/NCT04052451

International Registered Report Identifier (IRRID): RR2-10.2196/17223

(Interact J Med Res 2022;11(1):e32234) doi:10.2196/32234

KEYWORDS

gut-brain axis; microbiome; Microbial Ecosystem Therapeutic; depression; anxiety

Introduction

Major depressive disorder (MDD) is highly prevalent, affecting over 264 million people of all ages globally [1], and associated with high societal and personal burden. MDD is characterized by persistent depressed mood, loss of interest or pleasure and symptoms that cause clinically significant distress or impairment, or both [2]. MDD is often comorbid with other mental and physical illnesses, such as generalized anxiety disorder (GAD) [3]. GAD is characterized by excessive anxiety and worry about life circumstances, such as work, school, and relationships, among others [4], and has a lifetime prevalence of 5% [3,5]. The psychological symptoms of these illnesses are often accompanied by physical symptoms, such as abdominal issues, pain, and poor sleep quality [6,7]. Though there are a variety of gold standard and novel treatment methods that target symptoms of depression and anxiety [8], the heterogeneity of these disorders has led to difficulty in research in the field of mood and anxiety disorders [9].

Recent research has been exploring the connections between mood and anxiety disorders and the gut microbiome. As such, the gut-brain axis (GBA), which consists of bidirectional signaling between the gastrointestinal (GI) tract and the brain [10], has become a novel target for treatment of mood and anxiety symptoms. Studies suggest that this improvement of depressive and anxiety symptoms and severity may be related to the recolonization of the GI tract with healthy bacteria [11].

The purpose of this study was to evaluate the safety, efficacy, and tolerability of a GBA treatment method known as Microbial Ecosystem Therapeutic-2 (MET-2). MET-2 comes in an encapsulated form and is composed of 40 purified strains of lyophilized bacteria from a healthy 25-year-old donor; this was chosen for its favorable safety profile. MET-2 was developed in response to the growing body of literature supporting the ameliorative effects of fecal microbiota transplantation (FMT) on symptoms of depression. FMT is a procedure used to recolonize a patient's gut microbiota through the transplantation of feces from a donor to the recipient. A recent systematic review details the literature to date surrounding the safety, efficacy, and mechanisms of action for FMT [11]. Though FMT has been found to be effective in many cases, it is still an arduous, expensive, and invasive procedure. MET-2 provides an exciting alternative to FMT, with the possibility of conferring the same ameliorative properties with an easier and more tolerable mode of delivery. The objectives of this study were to assess the safety, tolerability, and efficacy of MET-2.

Methods

Study Design and Ethics Approval

This study was a 10-week, open-label, phase 1 clinical trial conducted out of Providence Care Hospital in Kingston, Ontario, Canada. This study was approved by the Health Science Regulatory Ethics Board of Queen's University, Ontario, Canada, and all methods were performed in accordance with the Declaration of Helsinki. The protocol has been previously published [11,12]. This study was registered with ClinicalTrials.gov on August 9, 2019 (NCT04052451).

Participants

Inclusion criteria for participants were as follows: (1) aged 18 to 65 years; (2) a diagnosis of MDD, GAD, or both, using the Mini International Neuropsychiatric Interview (MINI); (3) and no current use of any antidepressant medications. Mood, anxiety, sleep, GI symptoms, and severity of illness were assessed at screening using the Montgomery-Asberg Depression Rating Scale (MADRS) and the 7-item Generalized Anxiety Disorder scale (GAD-7). A minimum score of 15 on the MADRS or 8 on the GAD-7 were also required for inclusion in the study. For the full set of inclusion, exclusion, and discontinuation criteria, as well as the detailed study design, see the previously published protocol [12]. Participants were recruited from the local community using posters and online advertisements. Signed written informed consent was obtained from all participants in this study for the collection of all forms of data.

Intervention

The investigational product in this study was MET-2: capsules composed of 40 purified strains of lyophilized bacteria from a healthy 25-year-old donor. MET-2 was developed by NuBiyota in Guelph, Ontario, Canada. During the 8 weeks of treatment, all participants consumed three MET-2 capsules per day orally; each 0.5-g MET-2 capsule contains 3.2×10^5 to 3.2×10^{11} colony-forming units. This was known as the maintenance dose. Additionally, a loading dose of 5 g of MET-2 was taken for 2 days immediately following baseline and week-2 visits for all participants and following the week-4 visit for nonresponders (ie, those lacking a reduction in MADRS or GAD-7 scores by 50% by this time point) [12]. Mood, anxiety, GI symptoms, and sleep quality were assessed at all biweekly treatment visits. At the week-10 follow-up, only mood and anxiety were assessed.

Outcome Measures

The clinical measures included the following: (1) the GAD-7 [13] to assess anxiety symptoms and severity, with scores ranging from 0 (no anxiety) to 21 (severe anxiety); (2) the MADRS [14] to assess depressive symptoms and severity, with scores ranging from 0 (no depression) to 60 (severe depression); (3) the Snaith-Hamilton Pleasure Scale (SHAPS) [15] to assess anhedonia, with scores ranging from 0 (no anhedonia) to 14 (severe anhedonia); (4) the 16-item Quick Inventory of Depressive Symptomatology–Self-Report (QIDS-SR16) [16] to assess depressive symptoms, with scores ranging from 0 (no depression) to 27 (severe depression); (5) the Gastrointestinal Symptom Rating Scale (GSRS) [17] to assess GI symptoms, composed of five subscales—reflux, diarrhea, constipation, abdominal pain, and indigestion syndrome—with scores ranging from 1 (no discomfort) to 7 (severe discomfort); (6) the Pittsburgh Sleep Quality Index (PSQI) [18] to assess subjective sleep quality, with scores ranging from 0 (poor sleep quality) to 21 (good sleep quality); and (7) the Clinical Global Impressions-Severity scale (CGI-S) [19] to assess illness severity, with scores ranging from 0 (no illness) to 7 (severe illness).

Participants used a personal mood and symptom log to track any new symptoms that they have been experiencing since the beginning of treatment [20], assess the tolerability of treatment,

and keep track of their mood and sleep. Adverse events were assessed and recorded at all visits; they were categorized by frequency, severity, and causality. Only adverse events rated as a grade 2 or above were included in the analysis. Investigational product safety was assessed via recorded symptoms on the Toronto Side Effects Scale (TSES) [21-23], which is a 31-symptom scale with each symptom having a frequency and severity score ranging from 1 (never, no trouble) to 5 (every day, extreme trouble), respectively; the personal logs; and adverse events [12].

Statistical Analysis

Prism (version 8; GraphPad Software) was used to analyze all data from clinical measures obtained throughout the study and to create plots. A repeated-measures analysis of variance (ANOVA) was used to analyze changes in clinical measures from baseline to week 10. Paired *t* tests were used to compare the means of clinical measures at each time point to baseline. If a participant returned after a first course of treatment and later withdrew, their final clinical scores were projected to week 10.

Availability of Data and Materials

The data sets generated or analyzed during this study are not publicly available as of yet, but they are available from the corresponding author on reasonable request.

Results

Study Population

The final study cohort consisted of 12 participants, 8 (67%) of whom were female, with a total mean age of 28.8 (SD 12.8) years. The participants were recruited from May 16 to November 7, 2019. The trial profile can be found in Figure S1 in Multimedia Appendix 1. A total of 21 participants were originally screened for the study; 7 (33%) were ineligible due to a lack of MDD or GAD diagnosis or presence of mania, as per the MINI. Out of the remaining 14 participants, 2 (14%) withdrew prior to the week-2 visit for personal reasons and were not included in the analysis. The study population was diverse, with ages ranging from 19 to 59 years and representation from four different ethnicities. Further demographic information can be found in Table 1. Out of 12 participants, 10 (83%) were diagnosed with both MDD and GAD; 6 of these 10 (60%) were currently experiencing a major depressive episode, with the remainder having experienced at least one major depressive episode in the past. The 2 remaining participants out of 12 (17%) had sole diagnoses of MDD or GAD, respectively. All participants were combined into one group for analysis due to the high comorbidity of the two psychiatric illnesses. Mean baseline MADRS and GAD-7 scores were 19.0 (SD 4.8) and 13.6 (SD 4.0), respectively.

Table 1. Participant demographics.

Characteristic	Participants (N=12), n (%)
Gender	
Male	4 (33)
Female	8 (67)
Diagnosis	
Major depressive disorder (MDD) only	1 (8)
Generalized anxiety disorder (GAD) only	1 (8)
MDD and GAD	10 (83)
Ethnicity	
Caucasian	8 (67)
Chinese	2 (17)
Latin American	1 (8)
South Asian	1 (8)
Education level	
High school graduate or some college	3 (25)
College or university degree	9 (75)
Employment status	
Student	7 (58)
Working	2 (17)
On leave or disability	3 (25)

Efficacy Measures

The principal efficacy measures used were the MADRS, GAD-7, and CGI-S. A one-way repeated-measures ANOVA showed significant reductions in mean MADRS scores between visits ($F_{2.731,\ 30.05}$=8.784, P<.001). Multiple comparisons with Bonferroni adjustments showed a significant reduction in MADRS scores from baseline (mean 19.00, SD 4.843) to week

6 (mean 11.25, SD 8.001; *P*=.009), week 8 (mean 8.667, SD 8.732; *P*=.002), and week 10 (mean 8.250, SD 9.304; *P*=.006). There was a slight reduction in MADRS scores from baseline to week 2 (mean 14.67, SD 6.946; *P*=.08) and week 4 (mean 13.42, SD 9.443; *P*=.20); however, neither was significant. Additionally, 8 out of 12 (67%) participants were responders and improved by at least 50% in MADRS scores by week 8, of whom 7 (88%) remained responders by week 10. Of those who remained responders (n=7), 1 (14%) still worsened in mood symptoms, but not below 50% from baseline.

Similarly, results from the one-way repeated-measures ANOVA showed a significant reduction in mean GAD-7 scores ($F_{2.778, 30.55}$=9.638, *P*<.001). Multiple comparisons showed a significant reduction in GAD-7 scores from baseline (mean 13.58, SD 4.010) to week 4 (mean 9.167, SD 5.096; *P*=.03), week 6 (mean 7.667, SD 4.539; *P*=.004), week 8 (mean 7.333, SD 6.583; *P*=.03), and week 10 (mean 7.500, SD 6.448; *P*=.03). There was a slight reduction in GAD-7 scores from baseline to week 2 (mean 10.92, SD 4.542), which was not significant (*P*=.51). Additionally, 7 out of 12 (58%) participants were responders, and 6 (50%) improved by at least 50% in GAD-7 scores by week 8, of whom 5 (83%) remained responders by week 10. Of those who remained responders (n=5), 1 (20%) still worsened in anxiety symptoms, but not below 50% from baseline. A significant reduction in mean CGI-S scores was seen ($F_{2.833, 31.17}$=8.709, *P*<.001) from baseline (mean 3.667, SD 0.7785) to week 6 (mean 2.667, SD 0.9847; *P*=.006) and week 8 (mean 2.333, SD 1.073; *P*=.001). Although a slight reduction was seen in CGI-S scores from baseline to week 2 (mean 3.250, SD 0.7538; *P*=.26) and week 4 (mean 3.250, SD 0.9653; *P*=.56), neither was significant.

Additional efficacy measures included the QIDS-SR16, SHAPS, and PSQI, three self-report measures evaluating depressive symptomology, anhedonia, and subjective sleep quality, respectively. QIDS-SR16 scores were found to be significantly reduced ($F_{2.402, 26.42}$=7.111, *P*=.002) from baseline (mean 12.42, SD 3.147) to week 6 (mean 7.333, SD 5.836; *P*=.007) and week 8 (mean 6.750, SD 6.166; *P*=.005). Reductions from baseline to week 2 (mean 11.17, SD 4.877; *P*=.25) and week 4 (mean 8.667, SD 6.315; *P*=.10) were not significant.

No significant reduction in mean SHAPS scores was found ($F_{2.166, 23.82}$=0.9579, *P*=.40) from baseline (mean 3.917, SD 3.988) to week 2 (mean 3.333, SD 3.651), week 4 (mean 2.917, SD 4.641), week 6 (mean 3.417, SD 4.358), or week 8 (mean 2.583, SD 3.919).

Results from tests of between-subject contrasts in a one-way repeated-measures ANOVA showed a significant reduction in mean PSQI scores ($F_{2.547, 28.02}$= 3.100, *P*=.05). Multiple comparisons between PSQI scores at baseline (mean 9.333, SD 2.839) and week 2 (mean 8.500, SD 3.000), week 4 (mean 7.083, SD 3 .988), and week 6 (mean 7.917, SD 3.450) were not significant, but there was significance between PSQI scores at baseline (mean 9.333, SD 2.839) and week 8 (mean 7.000, SD 3.838). Graphs showing the change in efficacy outcome measures over the course of the study can be found in Figures S2-S7 in Multimedia Appendix 1.

Safety and Tolerability Measures

A total of 11 adverse events and zero serious adverse events were reported by participants during the course of the study. The majority of these adverse events were declared unrelated to the investigational product by the participants' family physicians, the principal investigator, or both. The most common adverse event was a stomachache, but this was reported multiple times by the same participant. Only one reported adverse event was considered to have a possible relationship to the study product: an instance of a stomachache rated level 2 with moderate pain. The full list of reported adverse events can be found in Table 2. As MET-2 is a therapeutic targeting the gut, its effect on GI symptoms was measured with the GSRS at five time points during the course of the study. One-way repeated-measures ANOVA showed no significant change in mean GSRS scores ($F_{1.998, 21.98}$=1.451, *P*=.26) from baseline (mean 1.871, SD 0.9754) to week 2 (mean 1.663, SD 0.5375), week 4 (mean 1.514, SD 0.5338), week 6 (mean 1.539, SD 0.6699), or week 8 (mean 1.514, SD 0.6273). A graph depicting the change in GSRS scores over the course of the study can be found in Figure S8 in Multimedia Appendix 1.

Side effect frequency and severity was measured using the TSES. A total of 31 side effects with an intensity greater than 14 were reported. These side effects can be found in Table 3.

Table 2. Adverse events reported by participants.

Adverse event	Incidence, n
Stomachache	3
Rash or black dots	1
Eye pain	1
Bloating	2
Diarrhea	1
Lower abdomen cramps	1
Itchy throat	1
Vomiting	1
Anxiety attack	1
Panicky feeling or jitters	1
Heart palpitations	2
Nightmares	1
Abdominal pain	1

Table 3. Side effects with an intensity greater than 14, as measured by the Toronto Side Effects Scale (TSES).

Side effect[a]	Incidence, n
Nervousness	5
Agitation	3
Tremor	2
Dyspepsia	1
Nausea	1
Diarrhea	1
Weakness or fatigue	3
Drowsiness	2
Increased sleep	1
Decreased sleep	1
Flushing	3
Headache	1
Dry mouth	1
Anorgasmia	1
Decreased libido	1
Delayed ejaculation	1
Impotence	1
Bloating	1
Heart palpitations	1

[a]Each symptom measured by the TSES had a frequency and severity score ranging from 1 (never, no trouble) to 5 (every day, extreme trouble), respectively.

Discussion

Principal Findings

In this study, MET-2, the novel gut microbiota–targeting treatment for symptoms of MDD and GAD, was found to be safe, generally tolerable, and efficacious. These findings are in line with what was expected, according to the literature to date regarding FMT, and suggest that gut microbiome manipulation can result in the alleviation of symptoms of a variety of psychiatric illnesses, including MDD and GAD. Preclinical research has found both the conferment of psychiatric symptoms through FMT from animals displaying behaviors related to

psychiatric illnesses to antibiotic-treated animals [24-27] and the amelioration of psychiatric symptoms after FMT from healthy animals to those displaying psychiatric symptoms [28,29]. Clinical research has found that the transfer of microbiota from a healthy donor to an ill recipient often results in the alleviation of psychiatric symptoms [11]. As MET-2 is a collection of bacteria cultured from the gut of a healthy human donor, it was suspected, and subsequently supported by this study, that treatment with MET-2 would result in symptom improvement similar to that seen in clinical FMT studies.

When comparing studies, MET-2 was found to be as safe as FMT, since both the FMT studies and this study reported relatively few adverse events that were deemed to be related to the treatment. When comparing the tolerability and burden to the patient, it is expected that MET-2 would be less of a burden, as FMT can be a rather arduous process, but further research would be needed to determine which treatment is more feasible.

Limitations

This study addresses some of the limitations in the literature around GBA treatments, such as fecal transplants, including safety, stigma, and labor costs. It also evaluates the use of a GBA treatment without other treatment interference, such as antidepressants or structured psychotherapy. The main limitation of this study, as with the reviewed clinical studies, was the small sample size and open-label design. The lack of large-scale, double-blind randomized controlled trials makes it difficult to determine efficacy and safety. The small sample size, in addition to the missing data, prevents large-scale analyses between parameters and may be the reason for the limited significance in the data, given the trends that were seen. Additionally, the lack of a placebo arm, in conjunction with the limited power in the sample size, suggests that the results may be merely due to the placebo effect or chance. That said, typically, the placebo effect is around 30% to 40% in psychiatric indications, and our study response was 75%, suggesting that it was unlikely to be a placebo effect. It also seems that the effects of MET-2 may begin to wear off after stopping treatment; however, given our study was only 10 weeks long and the follow-up period was only 2 weeks long, we cannot be certain at what point and rate the effects diminish. The follow-up period in future studies will need to be longer to determine if there is a need for maintenance therapy.

Further, given the nature of the study, participants were asked to come in every 2 weeks for an in-person visit, where, inevitably, they had an opportunity to open up about their mood and related symptoms, which could be therapeutic in and of itself. This could have contributed to the quick improvement in mood and anxiety symptoms.

Finally, though we looked to see whether participants may have had a diagnosis for an eating disorder or alcohol dependence, we did not ask how often these individuals were drinking or what their eating habits were. These components may have had an effect on the response to, or the transiency, of treatment.

Conclusions

In summary, our study has found MET-2 to be efficacious, as it significantly improved mood and anxiety symptoms in as early as 4 and 2 weeks of treatment, respectively. We found that 9 out of 12 (75%) participants had improved by at least 50% in their MADRS or GAD-7 scores from the start to the end of treatment. This improvement was seen in conjunction with limited side effects and a lack of serious adverse events.

With high individual variability in symptomatology and prognosis, high concentrations of comorbidity with other disorders, and genetic and environmental influences, progress in research in the treatment of psychiatric disorders has been challenging. Given the adaptable nature of the gut microbiome, it may be a good representation of the individual's history and could explain differences in risk of illness, disease course, and response to treatment.

With the complicated heterogeneity of psychiatric disorders, finding one treatment that works for all patients is not achievable, especially given the range of factors that influence the disorder and treatment response. Our study has shown MET-2 to considerably improve mood and anxiety symptoms, with limited side effects. While the research in this field is far from complete, the potential of targeting the GBA using GBA treatments, such as FMT and MET-2, to alleviate symptoms of psychiatric illness is promising. That said, further large-scale research in exclusively psychiatric indications is needed to strengthen the evidence that gut repopulation treatments, specifically MET-2, can be effective treatment methods.

Acknowledgments

This study was sponsored by NuBiyota LLC. Additional funding was received from Mitacs. The funder, NuBiyota, provided the investigational product; though they had a role in other data analysis in the trial, that data are not reported in this paper. Dr Casimiro Cabrera and Dr Gustavo Vazquez served as clinician-consultants during this trial.

Authors' Contributions

ACM is the corresponding author and was involved in designing and conducting the clinical trial, collecting and analyzing data, and writing, editing, and submitting this manuscript. EF was involved in writing and editing this manuscript. RM was the principal investigator of the study and was involved in writing and editing of this manuscript. ACM had full access to all the data in the study and had final responsibility for the decision to submit the paper for publication.

Conflicts of Interest

RM has received consulting and speaking honoraria from Allergan, Janssen, KYE, Lundbeck, Otsuka, Pfizer, and Sunovion, and has received research grants from CAN-BIND (Canadian Biomarker Integration Network in Depression), the Canadian Institutes of Health Research, Janssen, Lallemand Health Solutions, Lundbeck, NuBiyota, Ontario Brain Institute, the Ontario Mental Health Foundation, and Pfizer. The other authors have no conflicts to declare.

Multimedia Appendix 1
Trial profile and graphs of outcome measures.
[DOCX File , 316 KB - ijmr_v11i1e32234_app1.docx]

References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. Institute of Health Metrics and Evaluation. Depression. Global Health Data Exchange (GHDx). 2019. URL: http://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019-permalink/d780dffbe8a381b25e1416884959e88b [accessed 2022-01-07]
3. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):617-627 [FREE Full text] [doi: 10.1001/archpsyc.62.6.617] [Medline: 15939839]
4. Parekh R. What are anxiety disorders? American Psychiatric Association. 2017. URL: https://www.psychiatry.org/patients-families/anxiety-disorders/what-are-anxiety-disorders [accessed 2020-09-22]
5. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
6. Otheman Y, Fakir A, Kadiri M, Bichra MZ. Frequency of physical symptoms in anxiety-depressive disorders: cross-sectional study in a population of 202 psychiatric consultants [Article in French]. Pan Afr Med J 2018;31:149 [FREE Full text] [doi: 10.11604/pamj.2018.31.149.15738] [Medline: 31037209]
7. Johnson EO, Roth T, Breslau N. The association of insomnia with anxiety disorders and depression: Exploration of the direction of risk. J Psychiatr Res 2006 Dec;40(8):700-708. [doi: 10.1016/j.jpsychires.2006.07.008] [Medline: 16978649]
8. Lam RW, Kennedy SH, Parikh SV, MacQueen GM, Milev RV, Ravindran AV, CANMAT Depression Work Group. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: Introduction and methods. Can J Psychiatry 2016 Sep;61(9):506-509 [FREE Full text] [doi: 10.1177/0706743716659061] [Medline: 27486152]
9. Goldberg D. The heterogeneity of "major depression". World Psychiatry 2011 Oct;10(3):226-228 [FREE Full text] [doi: 10.1002/j.2051-5545.2011.tb00061.x] [Medline: 21991283]
10. Carabotti M, Scirocco A, Maselli MA, Severi C. The gut-brain axis: Interactions between enteric microbiota, central and enteric nervous systems. Ann Gastroenterol 2015;28(2):203-209 [FREE Full text] [Medline: 25830558]
11. Chinna Meyyappan A, Forth E, Wallace CJK, Milev R. Effect of fecal microbiota transplant on symptoms of psychiatric disorders: A systematic review. BMC Psychiatry 2020 Jun 15;20(1):299 [FREE Full text] [doi: 10.1186/s12888-020-02654-5] [Medline: 32539741]
12. Chinna Meyyappan A, Milev R. The safety, efficacy, and tolerability of Microbial Ecosystem Therapeutic-2 in people with major depression and/or generalized anxiety disorder: Protocol for a phase 1, open-label study. JMIR Res Protoc 2020 Jun 04;9(6):e17223 [FREE Full text] [doi: 10.2196/17223] [Medline: 32495743]
13. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
14. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr;134:382-389. [doi: 10.1192/bjp.134.4.382] [Medline: 444788]
15. Snaith RP, Hamilton M, Morley S, Humayan A, Hargreaves D, Trigwell P. A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. Br J Psychiatry 1995 Jul;167(1):99-103. [doi: 10.1192/bjp.167.1.99] [Medline: 7551619]
16. Rush A, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep;54(5):573-583. [doi: 10.1016/S0006-3223(02)01866-8]
17. Svedlund J, Sjödin I, Dotevall G. GSRS--A clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 1988 Feb;33(2):129-134. [doi: 10.1007/BF01535722] [Medline: 3123181]
18. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4]

19. Busner J, Targum SD. The clinical global impressions scale: Applying a research tool in clinical practice. Psychiatry (Edgmont) 2007 Jul;4(7):28-37 [FREE Full text] [Medline: 20526405]
20. Daily mood chart. Black Dog Institute. 2020. URL: https://www.blackdoginstitute.org.au/wp-content/uploads/2020/04/19-dailymoodchart.pdf [accessed 2020-09-28]
21. Vanderkooy JD, Kennedy SH, Bagby RM. Antidepressant side effects in depression patients treated in a naturalistic setting: A study of bupropion, moclobemide, paroxetine, sertraline, and venlafaxine. Can J Psychiatry 2002 Mar;47(2):174-180. [doi: 10.1177/070674370204700208] [Medline: 11926080]
22. Yin OQP, Wing YK, Cheung Y, Wang ZJ, Lam SL, Chiu HFK, et al. Phenotype-genotype relationship and clinical effects of citalopram in Chinese patients. J Clin Psychopharmacol 2006 Aug;26(4):367-372. [doi: 10.1097/01.jcp.0000227355.54074.14] [Medline: 16855453]
23. Crawford AA, Lewis S, Nutt D, Peters TJ, Cowen P, O'Donovan MC, et al. Adverse effects from antidepressant treatment: Randomised controlled trial of 601 depressed individuals. Psychopharmacology (Berl) 2014 Aug;231(15):2921-2931 [FREE Full text] [doi: 10.1007/s00213-014-3467-8] [Medline: 24525810]
24. Li N, Wang Q, Wang Y, Sun A, Lin Y, Jin Y, et al. Fecal microbiota transplantation from chronic unpredictable mild stress mice donors affects anxiety-like and depression-like behavior in recipient mice via the gut microbiota-inflammation-brain axis. Stress 2019 Sep;22(5):592-602. [doi: 10.1080/10253890.2019.1617267] [Medline: 31124390]
25. Lv W, Wu X, Chen W, Li Y, Zhang G, Chao L, et al. The gut microbiome modulates the changes in liver metabolism and in inflammatory processes in the brain of chronic unpredictable mild stress rats. Oxid Med Cell Longev 2019;2019:7902874 [FREE Full text] [doi: 10.1155/2019/7902874] [Medline: 31772709]
26. Xiao H, Ge C, Feng G, Li Y, Luo D, Dong J, et al. Gut microbiota modulates alcohol withdrawal-induced anxiety in mice. Toxicol Lett 2018 May 01;287:23-30 [FREE Full text] [doi: 10.1016/j.toxlet.2018.01.021] [Medline: 29391279]
27. Yang C, Fang X, Zhan G, Huang N, Li S, Bi J, et al. Key role of gut microbiota in anhedonia-like phenotype in rodents with neuropathic pain. Transl Psychiatry 2019 Jan 31;9(1):57 [FREE Full text] [doi: 10.1038/s41398-019-0379-8] [Medline: 30705252]
28. Langgartner D, Vaihinger CA, Haffner-Luntzer M, Kunze JF, Weiss AJ, Foertsch S, et al. The role of the intestinal microbiome in chronic psychosocial stress-induced pathologies in male mice. Front Behav Neurosci 2018 Oct 26;12:252 [FREE Full text] [doi: 10.3389/fnbeh.2018.00252] [Medline: 30464743]
29. Tillmann S, Abildgaard A, Winther G, Wegener G. Altered fecal microbiota composition in the Flinders sensitive line rat model of depression. Psychopharmacology (Berl) 2019 May;236(5):1445-1457 [FREE Full text] [doi: 10.1007/s00213-018-5094-2] [Medline: 30470860]

Abbreviations

ANOVA: analysis of variance
CGI-S: Clinical Global Impressions-Severity scale
FMT: fecal microbiota transplantation
GAD: generalized anxiety disorder
GAD-7: 7-item Generalized Anxiety Disorder scale
GBA: gut-brain axis
GI: gastrointestinal
GSRS: Gastrointestinal Symptom Rating Scale
MADRS: Montgomery-Asberg Depression Rating Scale
MDD: major depressive disorder
MET-2: Microbial Ecosystem Therapeutic-2
MINI: Mini International Neuropsychiatric Interview
PSQI: Pittsburgh Sleep Quality Index
QIDS-SR16: 16-item Quick Inventory of Depressive Symptomatology–Self-Report
SHAPS: Snaith-Hamilton Pleasure Scale
TSES: Toronto Side Effects Scale

Edited by G Eysenbach; submitted 19.07.21; peer-reviewed by V Taylor, C Kaylor-Hughes; comments to author 06.10.21; revised version received 22.10.21; accepted 22.11.21; published 21.01.22.

Please cite as:
Chinna Meyyappan A, Forth E, Milev R
Microbial Ecosystem Therapeutic-2 Intervention in People With Major Depressive Disorder and Generalized Anxiety Disorder: Phase 1, Open-Label Study
Interact J Med Res 2022;11(1):e32234
URL: https://www.i-jmr.org/2022/1/e32234
doi:10.2196/32234
PMID:35060914

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org