HEALTH TECHNOLOGIES AND INNOVATIONS TO EFFECTIVELY RESPOND TO THE COVID-19 PANDEMIC

EDITED BY: Björn Wolfgang Schuller, Phuong N. Pham, Synho Do, Constantinos S. Pattichis and Pradeep Nair
PUBLISHED IN: Frontiers in Digital Health

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-639-2
DOI 10.3389/978-2-88974-639-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

HEALTH TECHNOLOGIES AND INNOVATIONS TO EFFECTIVELY RESPOND TO THE COVID-19 PANDEMIC

Topic Editors:
Björn Wolfgang Schuller, Imperial College London, United Kingdom
Phuong N. Pham, Harvard Medical School, United States
Synho Do, Massachusetts General Hospital, Harvard Medical School, United States
Constantinos S. Pattichis, University of Cyprus, Cyprus
Pradeep Nair, Central University of Himachal Pradesh, India

Citation: Schuller, B. W., Pham, P. N., Do, S., Pattichis, C. S., Nair, P., eds. (2022). Health Technologies and Innovations to Effectively Respond to the COVID-19 Pandemic. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-639-2

Table of Contents

frontiers in Digital Health

EDITORIAL
published: 14 February 2022
doi: 10.3389/fdgth.2022.849652

Editorial: Health Technologies and Innovations to Effectively Respond to the Covid-19 Pandemic

*Pradeep Nair**

Department of New Media, School of Journalism, Mass Communication & New Media, Central University of Himachal Pradesh, Dharamsala, India

Keywords: health technology, innovation, healthcare, digital health, COVID-19

Editorial on the Research Topic

Health Technologies and Innovations to Effectively Respond to the Covid-19 Pandemic

Covid-pandemic has an unprecedented impact on global, regional, and national health systems across countries. Both public and private healthcare sectors had struggled and are still struggling to respond to the impact of the pandemic. The struggle is not only about adopting diverse healthcare responses in terms of cutting-edge technological tools and innovations in the areas of public health, medicine and wellness to take prompt decisions to address the pandemic by flattening the disease curve but also to revisit and reopen the realm of "digital health" in the policy and public discourse.

While keeping in view, the short, medium and long term response strategies, it is an appropriate time to take a tangible shift toward holistic technology and data-driven digital tools. This will help to engage both public and private healthcare systems across the country in facilitating policy dialogue, technical assistance and training on specialized policy and response interventions at the regional and national level. It is also a fact that the well-tested hardware are already in place and quite a few experiments have been conducted but has not been in good use, particularly in developing nations, for lack of national will, want of resources, and strategic planning to reach those who really need it most.

An alignment of expertise, leadership, and practices are required to synthesize knowledge and experience to assess capacities and avenues of emerging technologies such as artificial intelligence (AI), machine learning, Blockchain, health wearables, remote patient monitoring trackers, sensor-enabled hospital beds, medication-tracking systems and medical supplies and equipment inventory tracking systems. This will provide policy dialogue and responses to make the healthcare systems inclusive and accessible in terms of better patient experience and cost (1).

With the advent of newer technology, it will percolate down. The real concern is its application up to the last point. The first telemedicine experiments in developing nations were conducted way back in 1970s and 80s and a couple of more "resourceful" healthcare operators use it off and on since the early 1990s. The problem with the developing countries is to integrate the emerging healthcare technologies into the operational situation, that too for all. We should remember that any technology has the inherent characteristics of being used by those who can afford it. So, either it is a tool in the hand of private players (many of them have state of art technology) or the cost is subsumed by the State, which failed in the last so many decades to take to the last point. Educating and communicating people about the benefits of healthcare technology would come later when it is made available to them.

OPEN ACCESS

Edited and reviewed by:
Giovanni Ferrara, University of Alberta, Canada

***Correspondence:**
Pradeep Nair
nairdevcom@yahoo.co.in

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *06 January 2022*
Accepted: *24 January 2022*
Published: *14 February 2022*

Citation:
Nair P (2022) Editorial: Health Technologies and Innovations to Effectively Respond to the Covid-19 Pandemic. Front. Digit. Health 4:849652. doi: 10.3389/fdgth.2022.849652

The potential of telemedicine is well observed during the pandemic time as several consultations with doctors/health experts happened without going to the clinic or hospital. This was a positive sign and urges to look into new business models to help further telemedicine across the world. On this front, it is a time to encourage companies to come to the forefront with a wide range of tech-driven digital healthcare tools and products. It was predicted by the health industry experts that tele-health market across the world is likely to witness a massive spike both from the demand and supply sides. The service providers are assessing capacities and avenues, pursuing evidence-based innovations and technologies across the board, including diagnostic and telemedicine tools, cellphone apps for fitness, well-being, medical and healthcare, and data-driven software. But the argument is that maybe it is the pandemic which compelled to explore and extend these technologies to deal with the spatial and temporal gaps to access healthcare. The concern shall be to look into how these technologies and innovations will be used in a post-covid world to deal with public health issues (2). How in coming days, technologies and innovations will be integrated with public health response system for greater accessibility and affordability to health care services should be a priority area of research in public health studies. The pandemic has unraveled the myriad avenues and opportunities in boosting healthcare and life sciences in the world, it would be interesting to look into how health shall be observed as wealth in the post Covid world with technology and innovations as its greatest investment.

The recent studies on Covid-19 have shown that the rates of severity of the impact of the pandemic are more on people from the less privileged communities. Medical devices and applications that are connected to healthcare IT systems via the internet known as "The Internet of Medical Things (IoMT)" has the potential to provide regular access to health care services through cloud-connected health care professionals. The concept of "data governance" in healthcare can offer safe and secure digital ecosystem to manage the handling, storing and sharing of real-time patient data 24 x 7 across hospitals and healthcare institutions to the next level with clearly defined policies and procedures.

This Research Topic collection attempted to explore the new post-pandemic health realities with a focus on various policy and response strategies spearheaded by the countries across the world to make public healthcare services more accessible to citizens in terms of cost, security and data privacy issues. The diverse studies published under the Research Topic collection looked into the new structural and institutional shifts taking place to make healthcare services safe and convenient with the help of empowered frontline care leveraging technologies. To make the healthcare services more inclusive and accessible, healthcare organizations and institutions, both public and private need to orchestrate the myriad interconnected changes required to design, implement and sustain digitally-enabled healthcare delivery platforms.

The Research Topic collection addressed various policy and response strategies adopted to deal with restricted physical access to socio-economic infrastructure, facilities and services amid the pandemic with a focus on cutting-edge health-technologies at its core. The topic collection strongly advocates that health technologies and innovations are going to be one of the significant sectors for investment and innovations over the next 30 years. This will not only transform the global health care sector in terms of diagnosis, disease management, treatment and prevention but also help to better prepare for future emergencies.

The published topic collection had studies having a diverse and vast coverage. The mini review "Covid-19 and Computer Audition: An overview on what speech and sound analysis could contribute in the SARS-CoV-2 Corona crisis" by Schuller et al. advocates the use of Computer Audition (CA) for implementation of (pre-) diagnosis and monitoring tools, and more generally provides rich and significant contribution in the fight against Covid-19 spread. A brief research report titled "Early warning signs of a mental health tsunami: A coordinated response to gather initial data insights from multiple digital services providers" by Becky Inkster and Digital Mental Health Data Insights Group (DMHDIG) provides evidence based concept for researchers and private companies to work collaboratively on a diverse range of mental health concerns. The research report observed that there is an increasing demand for digital mental health support during Covid due to an increased presentation of anxiety and loneliness. The opinion on "Out-of-Hospital care of heart failure patients during and after Covid-19 pandemic: Time for Telemedicine" by Faragli et al. talks about how telemedicine has turned as an essential requirement during covid time. The opinion piece strongly advocates the use of telemedicine as a home monitoring solution to manage the patients' health during and after the pandemic. Another opinion piece titled "Digital Covid credentials: an implementation process" by Nehme et al. reflects on the acceptance of digital covid credentials to ensure the implementation of adequate safeguards. It urges that the digital aspects of published information within a trust framework can be very useful to certify an individual's most recent Covid related status.

A policy and practice review on "A crisis-responsive framework for medical device development applied to the Covid-19 pandemic" by Antonini et al. urges for a crisis-responsive design framework to assist with product development, under pandemic conditions. The review emphasizes on stakeholder engagement, needs assessment, rapid manufacturing, and modified product testing to enable accelerated development of healthcare products. The study highlighted the use of crisis-responsive framework in a case study of face shield design and production for a large US academic hospital. A mini-review on "Covid-19 prognostic models: A pro-con debate for machine learning vs. traditional statistics" by Hindawi et al. explored the possibilities of data science to access open datasets, tutorials, programming languages, and hardware to create mathematical models to address the Covid-19 pandemic. It advocates the use of data science models to trace the impact of the virus on population and individuals for further diagnostic, prognostic, and epidemiological observations and analysis. It also had a comparative analysis between the classical

statistics and machine learning techniques used for predicting covid outcomes.

The perspective by Fan et al. on "Factors to consider in the use of vital signs wearables to minimize contact with stable Covid-19 patients: experience of its implementation during the pandemic" cited reasons for choosing vital signs wearables for the purpose of reducing patient contact and preserving personal protective equipment. The study provides an overview of the factors needs to be considered while implementing vital signs wearable solutions in healthcare institutions. An original research study by Zhang et al. titled –"A real time portable IoT system for Telework tracking" validates the idea of using an Internet of Things (IoT) system to monitor the working status in real-time so as to record the working pattern and nudge the user to have a behavior change. The constructed shallow convolutional neural network (CNN) in the study helps to recognize the working status from a common working routine. The method adopted in the study is useful to the workers wellness during the Covid pandemic and also have a significant contribution in dealing with post-pandemic realities.

The perspective on "Integrated Care in the era of Covid-19: Turing vision into reality with digital health" by Kouroubali et al. looks into how digital health has been accepted as care platform across the globe to deal with the pandemic challenges. The perspective highlights the importance of digital health as an integrated care to support sharing and reusing of healthcare data for prevention, prediction and disease management. The study also addresses the political and social barriers and how to overcome them to achieve integrated care in practice. The perspective by Ramadi and Srinivasan on "Pre-emptive innovation infrastructure for medical emergencies: accelerating healthcare innovation in the wake of a global pandemic" proposes a pre-emptive innovation infrastructure incorporating in-house hospital innovation teams, consortia-based assembly of expertise, and novel funding mechanisms to combat health emergencies like Covid. The perspective talks about a framework to improve ongoing innovation and infrastructure for healthcare agencies by leveraging the strengths of academic, medical, government, and industrial institutions. The research study titled—"Covid-19 in Brazil—preliminary analysis of response supported by Artificial Intelligence in Municipalities" by Morales et al. emphasizes on the use of Artificial Intelligence to empower telehealth to increase coordinated patient access to health system during covid pandemic. The study describes a case report analyzing the use of Laura Digital Emergency Room as AI-powered telehealth platform in three different cities of Brazil. The study says that the implementation of an AI-powered telehealth will increase the access to healthcare services amid the unprecedented impact of Covid. The study urges for efforts to sustain affordable and scalable solutions to leverage value in health care systems in the context of middle and low income countries.

The research study by Kyriacou et al. on "Operating an eHealth System for pre-hospital and emergence health care support in light of Covid-19" talks about creating an electronic system (eEmergency System) in order to support, improve, and help the procedure of handling emergence calls. The study while quoting examples of case studies from Cyprus, focuses on developing an electronic system to support ambulance fleet handling, emergency call evaluation, triage procedure, and the improvement of communication between the call center and the ambulance vehicles. The system was further expanded during the Covid time in order to support the handling of patients infected with the new virus. The brief research report on "combinational analysis of phenotypic and clinical risk factors associated with hospitalized Covid-19 patients" by Das et al. is a follow-up study of using genomic data to identify a potential role of calcium and lipid homeostasis in severe Covid cases. The study attempts to identify similar combinations of features (disease signatures) associated with severe disease in a separate patient population with purely clinical and phenotypic data. The study used a Precision Life Combinational Analytics Platform to analyze features derived from de-identified health records in the United Health Group Covid-19 Data Suite. The study found several disease signatures where lower levels of lipids were found co-occurring with lower levels of serum calcium and leukocytes. The study looked into how these signatures are attributed to similar mechanisms linking calcium and lipid signaling where changes in cellular lipid levels during inflammation and infection affect calcium signaling in host cells. The study demonstrates that combinational analysis can identify disease signatures associated with the risk of developing severe Covid-19 separately from genomic or clinical data in different populations.

Another research study on "digital contact tracing against Covid-19 in Europe: current features and ongoing developments" by Blasimme et al. examines the evolution of digital contract tracing in eight European countries and highlights that privacy and data protection are at the core of contact tracing applications in Europe, even though the countries differ in their technical protocols, and their capacity to utilize collected data beyond proximity tracing alone. The study reflects a shift from a strict interpretation of data minimization and purpose limitation toward a more expansive approach to digital contact tracing in Europe. The study by Moura et al. on "explainable machine learning for Covid-19 Pneumonia classification with texture-based features extraction in Chest Radiography" provides evidential grounds for understanding the distinctive Covid-19 radiographic texture features using supervised ensemble machine learning models based on trees through the interpretable Shapley Additive Explanations (SHAP) approach. The study used 2611 Covid-19 chest X-ray images and 2611 non-Covid-19 chest X-rays. After segmenting the lung in three zones, histogram normalization is applied to extract radiomic features. SHAP Recursive Feature Elimination with Cross-Validation is used to select features and Hyperparameter optimization of XGBoost and Random Forest ensemble tree models were applied through random search. The study showed a predominance of radiomic feature selection in the right lung, leading to the upper lung zone.

The policy and practice review titled–"toward a common performance and effectiveness terminology for digital proximity tracing applications" by Lueks et al. explores digital proximity

tracing (DPT) for Sars-Cov-2 pandemic mitigation as a complex intervention to notify application users about possible risk exposures to infected persons. The review describes differences between performance and effectiveness measures and attempts to develop a terminology and classification system for DPT evaluation. It further discusses key aspects for critical assessments of integration of additional data measurements into DPT applications to facilitate understanding of performance and effectiveness of the applications. The research–"Retrospective analysis and forecasted economic impact of a virtual cardiac rehabilitation program in a third-party payer environment" by Harzand et al. forecasts the potential clinical and economic benefits of delivering a home-based virtual cardiac rehabilitation program based on a retrospective analysis of cardiac rehabilitation utilization and cost in a third party payer environment. The study performed a retrospective cohort study using insurance claims data from a large, third-party payer in the state of Pennsylvania.

Another study on "using machine learning to predict mortality for Covid-19 patients on day Zero in the ICT" by Jamshidi et al. urges for the use of machine learning based on typical laboratory results and clinical data registered on the day of ICU admission for the early prediction of severity of the disease in intensive care unit (ICU) patients to optimize treatment strategies. The study reveals that several machine learning algorithms, including Random Forest (RF), logistic regression, gradient boosting classifier, support vector machine classifier, and artificial neural network algorithms can be easily used to build classification models and the predictions can be studied by implementing the local interpretable model-agnostic explanation technique. The study by Gupta et al. on "Trends in Covid-19 publications: streamlining research using NLP and LDA" developed a comprehensive Latent Dirichlet Allocation (LDA) model with 25 topics using natural language processing (NLP) techniques on PubMed research articles about Covid. The study proposed a novel methodology to develop and visualize temporal trends to improvise existing online literature hubs. The last study in the topic collection as a review article by Kostkova et al. on "data and digital solutions to support surveillance strategies in the context of the Covid-19 pandemic" provides a comprehensive overview for the application of data and digital solutions to support surveillance strategies. It also talks about digital epidemiology, available data sources, and essential components of 21st Century digital surveillance, early warning and response, outbreak management, control and digital interventions in the context of Covid-19 pandemic and beyond.

The challenge in a post-covid world will be to validate the new technologies and innovations without assuming that they will work because they worked in an uncontrolled emergency situation during the pandemic. This requires new observations regarding the adoption and use of digital health tools and services to receive information and seek social and medical assistance and support in a post-covid world when again the people will have physical access to socio-economic infrastructure, facilities and services (3). The concern will be to look into how medium and low income countries will refocus on preventive healthcare in a post-covid world with diverse and massive population, with disproportionate health infrastructure that often impacts medical responses to health emergency situations (4). When the world will regain its feet in post-covid realties, the challenge will be to recalibrate public health strategies with the help of emerging digital healthcare technologies to strengthen preventive healthcare. Investigations are required to explore the role of government in incentivizing and encouraging healthcare industries to deal with the issues of post-covid safety and care. The integration of health technologies and innovations into health policies and response strategies could be one of the concerned areas of research in post-covid world. The understanding of comprehensive responses of countries that have been successful to deal with covid through the timely and effective deployment of health technologies to facilitate planning, surveillance, testing, contact tracing, quarantine, and clinical management will help other countries to improve their health strategies and response systems (5).

We hope that this Research Topic collection will give the audience some good readings to look into how healthcare tools and innovations facilitate and enable easy-to-use and affordable healthcare services by providing accessible interfaces connecting all the stakeholders of healthcare system in a holistic manner. If these healthcare innovations and technologies will be used sensibly and rationally, they have the potential to revolutionize healthcare access and delivery in a post-covid world especially for women, vulnerable and marginalized groups, and economically disadvantaged sections of the society. To effectively deal with the post-covid realities, integrating medical and healthcare services with health technologies and innovations should be the priority for healthcare providers and policy makers. To effectively deal with a Covid like situation in future, a mechanism is required to enable patient tracking, movement of citizens, identification of viral loads and disease clusters to strengthen monitoring and containment measures and these also need to discuss at various academic and research forums.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

The author acknowledge the suggestions made by the reviewer and Editor Prof. Giovanni Ferrara, University of Alberta, Canada, for facilitating my understanding of the validation of new technologies and innovations in a post-covid world.

REFERENCES

1. McCall B. Covid-19 and artificial intelligence: protecting healthcare workers and curbing the spread. *Lancet*. (2020) 2:e166–7. doi: 10.1016/S2589-7500(20)30054-6
2. Health Affairs. Ensuring the growth of telehealth during Covid-19 does not exacerbate disparities in care. Available online at: https://www.healthaffairs.org/do/10.1377/hblog20200505.591306/full/ (accessed January 23, 2022).
3. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol.* (2016) 69:245–7. doi: 10.1016/j.jclinepi.2015.04.005
4. Wallis L, Blessing P, Dalwai M, Shin SD. Integrating mHealth at point of care in low-and middle-income settings: the system perspective. *Glob Health Action.* (2017) 10:1327686. doi: 10.1080/16549716.2017.1327686
5. Naude W. Artificial intelligence versus Covid-19: limitations, constraints and pitfalls. *AI Soc*. (2020) 28:1–5. doi: 10.1007/s00146-020-00978-0

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

OPINION
published: 12 November 2020
doi: 10.3389/fdgth.2020.593885

Out-of-Hospital Care of Heart Failure Patients During and After COVID-19 Pandemic: Time for Telemedicine?

Alessandro Faragli[1,2,3,4]*, *Edoardo La Porta*[5,6,7], *Carlo Campana*[8], *Burkert Pieske*[1,2,3,4], *Sebastian Kelle*[1,2,4], *Friedrich Koehler*[9] *and Alessio Alogna*[2,3,4]

[1] Department of Internal Medicine and Cardiology, Deutsches Herzzentrum Berlin, Berlin, Germany, [2] Charité–Universitätsmedizin Berlin, Department of Internal Medicine and Cardiology, Campus Virchow-Klinikum, Berlin, Germany, [3] Berlin Institute of Health (BIH), Berlin, Germany, [4] DZHK (German Center for Cardiovascular Research), Partner Site Berlin, Berlin, Germany, [5] Department of Cardionephrology, Clinical Ligurian Institute of High Specialty, Villa Maria Group (GVM) Care and Research, Rapallo, Italy, [6] Department of Internal Medicine, University of Genoa and IRCSS Azienda Ospedaliera Universitaria San Martino-IST, Genoa, Italy, [7] Unit of Dialysis, IRCSS Istituto Giannina Gaslini, Ospedale Pediatrico, Genoa, Italy, [8] Department of Cardiology, Sant'Anna Hospital, ASST-Lariana, Como, Italy, [9] Center for Cardiovascular Telemedicine, Department of Cardiology and Angiology, Charité–Universitätsmedizin Berlin, Berlin, Germany

Keywords: heart failure, COVID-19, home monitoring, body fluids, telemedicine

OPEN ACCESS

Edited by:
Esteban J. Pino,
University of Concepcion, Chile

Reviewed by:
Xiaorong Ding,
University of Oxford, United Kingdom
Carolina Varon,
KU Leuven, Belgium

***Correspondence:**
Alessandro Faragli
faragli@dhzb.de

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *11 August 2020*
Accepted: *16 October 2020*
Published: *12 November 2020*

Citation:
Faragli A, La Porta E, Campana C, Pieske B, Kelle S, Koehler F and Alogna A (2020) Out-of-Hospital Care of Heart Failure Patients During and After COVID-19 Pandemic: Time for Telemedicine? Front. Digit. Health 2:593885. doi: 10.3389/fdgth.2020.593885

OPINION LETTER

The current letter has been driven by the clinical observation of the events that happened in the last months during the coronavirus disease 2019 (COVID-19) pandemic in the European countries, with specific reference to the situation of patients in the North of Italy and in Germany.

> *"Specialists are people who always repeat the same mistake."*—Walter Gropius, German architect and founder of the Bauhaus School

A 71-year-old male, Caucasian, is affected by chronic heart failure (CHF) New York Heart Association (NYHA) class III and chronic kidney disease stage III. The first diagnosis of CHF has been performed 4 years ago after hospitalization for acute coronary syndrome resulting in a percutaneous coronary intervention with primary stenting. Since then, he has been hospitalized at an average of 1.5 times per year. Two thirds of the patient's hospitalizations were caused by worsening of his chronic body fluid congestion with peripheral edema and impaired renal function, while for one third of the cases, the main cause was volume depletion. This has been manifesting with hypotension and hypokalemia as a result of challenges in managing the correct intake of diuretics and blood pressure-lowering medications.

When admitted to the cardiology ward, such a paradigmatic patient represents a challenge, especially with regard to the body fluid management. This requires a specialized heart failure (HF) team with extensive experience in the field. The clinical approach to such complex patients includes daily physical exam and control of body fluid balance through fluid intake and urine output. On top of this, biomarkers, chest X-ray, lung ultrasound, and, for cardiorenal patients, bioimpedance analysis are performed during the hospitalization to assess the patients' congestion status. Moreover, it requires a fine tuning of medications, diet, and liquid restrictions to achieve a proper balance between body volume and blood pressure.

After recompensation and discharge, the patient is left alone with a single method to monitor himself: a standard weight scale. He weighs himself every day, trying to keep contact with his physician on the phone. He relies on elective appointments in the outpatient clinic, three times a year.

It is February 2020, and the COVID-19 pandemic takes hold. The patient is not able to get a prompt appointment with his physician. In case of worsening of his clinical condition, he is being told to call the emergency number.

This situation could evolve into three different scenarios:

- No Hospitalization Needed and no SARS-CoV-19 Infection During the Pandemic

The patient independently manages his chronic fluid congestion based on the experience of the past years. He can avoid any contact with COVID-19^{+} patients. However, the lack of a proper medical assistance may increase the risk of experiencing a sudden decompensation event. Compared to the time pre-COVID-19 pandemic, his mortality risk may look the same in the short term, but it will probably drastically increase in the medium to long term.

- Clinical Deterioration of the Patient Condition

The patient constantly deteriorates, gains weight, and his quality of life is strongly affected. He does not get an appointment with his general practitioner, neither in the outpatient clinics. He manages to survive without an emergency hospitalization, but in a poor condition, for a few weeks or months until the pandemic situation has improved, and he receives medical attention.

- Hospitalization for Acute Decompensation

The patient gets admitted to the hospital because of an acute exacerbation of his condition. The hospitals are under great pressure because of the pandemic, and the intensive care units have limited capacity. His mortality risk may still be higher than before the pandemic.

- Hospitalization for COVID-19

The patient is infected with SARS-CoV-19, his condition quickly worsens, and he needs to be quickly hospitalized. The mortality risk in this scenario is possibly the highest.

The last two scenarios are unfortunate and, most importantly, avoidable. However, a drastic change in the management of chronic patients is today, more than ever, of paramount importance. In addition to the direct impact on public health, COVID-19 has been challenging the way of living, the habits, as well as many long-existing cultural and social structures on which societies are based. Maybe for the first time in modern medicine, the major strategy of healthcare policymakers has been to keep patients outside hospitals to avoid the spread of the infection. However, this is not enough, and remote monitoring strategies are necessary for the future of a sustainable healthcare system (1) for many reasons.

We are experiencing since many years a clear mismatch between specialized physicians and patients in need of care (2, 3). Aging in western countries has led to an increase in the number of patients with multiple comorbidities (4). Since the very first beginning of medicine and then throughout modern times, healthcare systems have been structured on a face-to-face patient–physician interaction. This kind of approach has contributed to a hospital influx of patients during the COVID-19 outbreak.

Several European healthcare systems seemed unprepared to fight the pandemic, while many hospitals even contributed to the initial spread of COVID-19. In this scenario, most of the scheduled medical and surgical procedures were rescheduled, while many chronic patients have been temporarily lost at follow-up.

For these reasons, telemedicine has turned from being a "nice to have" approach to an essential requirement (5) for a more efficient system. Chronic HF patients are facing an increased challenge regarding the management of body fluids.

While during a hospitalization, the volume status of the patient is generally addressed by the medical doctors with various methods and solutions such as physical examination, ultrasound, chest X-ray, or blood examinations, at home, the solutions available and specific for the prevention of decompensation events in such patients are relatively limited, as described in **Figure 1**. We now believe that a natural shift toward home monitoring solutions should be considered and encouraged to manage the patients' body volume during and after the pandemic.

While invasive solutions such as cardioMEMS have already demonstrated to decrease HF patients' hospitalization (6), their utilization has been limited mainly due to invasiveness and related adverse events (7) or lack of penetrance among medical doctors (8). The spectrum of non-invasive telemedicine is, instead, broad, and the recent positive results achieved by Köhler et al. (9) and confirmed by the meta-analysis of Zhu et al. (10) are encouraging.

Audio–video tools able to connect patients and physicians for real-time consultations are widely available and, even if still not extensively adopted, during the pandemic, and the relative lockdown, virtual visits (VVs) represented the first tangible action in favor of a home monitoring of chronic patients, obtaining positive results (11). A recent work published by Salzano et al. (12) was able to show in a cohort of 103 HF patients how a 24/7 audio and video management during the pandemic is able to decrease hospitalizations and mortality compared to a previously observed comparable population in which telemedical support was not present or available. While the feasibility and utility of such solutions for HF patients have been shown to be beneficial even before the pandemic (13), a lot of work is necessary to support a routine utilization. Even if advancements were made in terms of reimbursements, audio/video tools are not yet part of an organized widespread telemonitoring plan in all countries (11). Moreover, VVs require an important engagement by the medical doctors that many times does not match with the time available. Centralized hospitals dedicated only to telemedicine may solve such a problem. The Center for Cardiovascular Telemedicine in Charité Berlin is an example of how a centralized management of telemedicine information is able to act successfully on distant and rural territories (9).

Portable or wearable devices collecting vital parameters while involving the use of Web apps or smartphones are increasingly reliable, and they represent the next generation of solutions available for CHF patients (14). However, even if most research

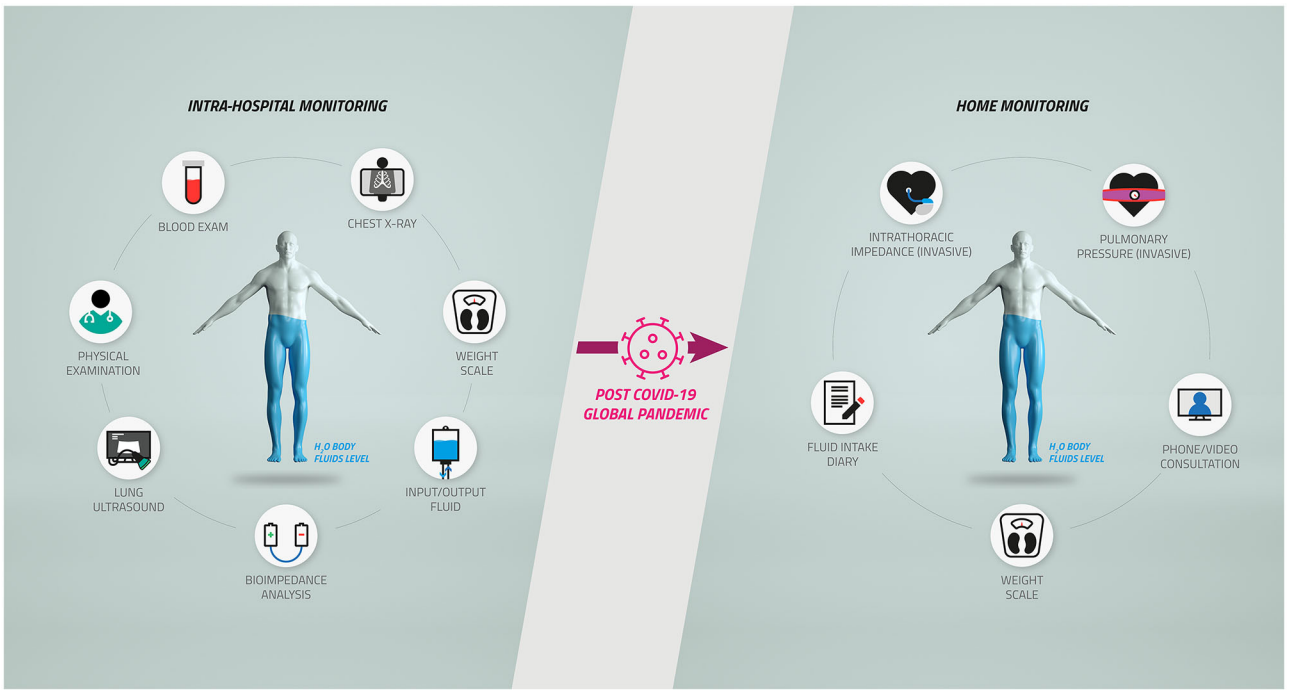


FIGURE 1 | Body fluid management in chronic heart failure and cardiorenal patients after the coronavirus disease 2019 (COVID-19) pandemic. The complex workflow for the diagnosis and monitoring of body fluids of chronic heart failure and cardiorenal patients inside the hospitals **(Left)** has been restricted by the COVID-19 pandemic, causing a potentially relevant shift toward home monitoring protocols **(Right)**.

has shown the cost-effectiveness of such devices, regulatory authorities have slowed a full penetration of wearables in the medical market until further clinical evidence is available (15).

Another important aspect is technological, since the correct technology should match with the correct clinical indication. For example, for HF patients, a further step should be taken to move beyond the utilization of weight scales, known since many years to be poorly accurate in detecting body fluid congestion and body volume imbalances (14). While remote monitoring through implantable cardioverter defibrillators (ICDs) works really well for the detection of arrhythmias, the same cannot be said for the management of body fluids through intrathoracic impedance mainly due to the high risk of false positives that slowed the initial enthusiasm (16). Even if a lot of research is undergoing in new non-invasive technologies for the assessment of body fluids, their clinical value still needs to be demonstrated (14).

The complex clinical scenario offered by the COVID-19 pandemic should finally be the springboard for telemedicine. Telemedicine still presents challenges, such as the identification of the correct patients' populations in need, a variable that should always be addressed first. This has been nicely demonstrated in a recent randomized, multicenter, open-label telemedicine study by Galinier et al. (17), where patients at higher risk and the ones more socially isolated presented better clinical outcomes than more stable patients, showing how telemedicine may be more useful in such patients. Usability of technologies and increased adherence to the monitoring plan are some of the topics that need to be addressed to finally make telemedicine affordable and efficient for the post-pandemic healthcare system. On top of that, we believe that optimization of the healthcare organization and automatization of the management processes, meaning data collection, data interpretation, and clinical action toward the patients, need to proceed in a highly structured and fast path to be completely effective. This could be potentially achieved by departments or hospitals dedicated to telemedicine in conjunction with general practitioners. We believe that the introduction of working telemedicine programs needs to enter a novel stage, assigning specific duties and responsibilities to trained personnel. A collaboration between general practitioners and specialized centers is necessary mainly for medically underserved and rural areas. However, the roles need to be precisely defined to avoid confusion.

Both the Heart Failure Society of America and European Society of Cardiology strongly encourage the use of telemedicine for HF management during the COVID-19 outbreak (18). However, a functioning widespread system that allows the reimbursement of home monitoring solutions is still lacking (19). Germany is moving in an innovative direction with the so-called Digital Care Act, entitling all individuals covered by statutory health insurance to reimbursement for certain digital health applications (20). The chance of having a digital solution reimbursed encourages the hospitals to adopt novel telemedical solutions and produce proactively a much faster tangible outcome.

In Italy, during the COVID-19 outbreak, several patients experienced a poor outcome because they did not access to health system (21). Telemedicine owns nowadays the potential of delivering a better healthcare by empowering patients and

by providing individualized healthcare, especially during a pandemic (22).

Yes, it is time for telemedicine. But, first, let us make telemedicine a matter of routine. Let us learn from our mistakes.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

We acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Fund of Charité–Universitätsmedizin Berlin. AF and AA received the grant Validation Fund: Track 1 from the Berlin Institute of Health (BIH) on the project cardioBIA: a new algorithm for the prediction and prevention of congestive events in cardiological patients. FK is supported by the grant TELEMED5000 https://www.telemed5000.de/ founded by the German Federal Ministry of Economics and Energy (BMWi), and the project is associated with the Charité-COVID-19 Research project.

ACKNOWLEDGMENTS

This short manuscript is dedicated to the loving memory of Giuseppe Faragli, father of AF, heart failure patient, who died during the COVID-19 pandemic. Furthermore, we acknowledge and thank Simone Proietti Timperi for performing the graphical illustration.

REFERENCES

1. Fineberg HV. Shattuck lecture. A successful and sustainable health system–how to get there from here. *N Engl J Med.* (2012) 366:1020–7. doi: 10.1056/NEJMsa1114777
2. Scheffler RM, Liu JX, Kinfu Y, Dal Poz MR. Forecasting the global shortage of physicians: an economic- and needs-based approach. *Bull World Health Organ*. (2008) 86:516–23B. doi: 10.2471/BLT.07.046474
3. Liu JX, Goryakin Y, Maeda A, Bruckner T, Scheffler R. Global health workforce labor market projections for 2030. *Hum Resour Health.* (2017) 15:11. doi: 10.1186/s12960-017-0187-2
4. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nat Rev Cardiol.* (2018) 15:505–22. doi: 10.1038/s41569-018-0064-2
5. Duffy S, Lee TH. In-person health care as option B. *N Engl J Med*. (2018) 378:104–6. doi: 10.1056/NEJMp1710735
6. Abraham WT, Perl L. Implantable hemodynamic monitoring for heart failure patients. *J Am Coll Cardiol.* (2017) 70:389–98. doi: 10.1016/j.jacc.2017.05.052
7. Vaduganathan M, DeFilippis EM, Fonarow GC, Butler J, Mehra MR. Postmarketing adverse events related to the cardioMEMS HF system. *JAMA Cardiol.* (2017) 2:1277–9. doi: 10.1001/jamacardio.2017.3791
8. Singh R, Varjabedian L, Kaspar G, Zughaib M. CardioMEMS in a busy cardiology practice: less than optimal implementation of a valuable tool to reduce heart failure readmissions. *Cardiol Res Pract.* (2018) 2018:4918757–4918757. doi: 10.1155/2018/4918757
9. Koehler F, Koehler K, Deckwart O, Prescher S, Wegscheider K, Kirwan B-A, et al. Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): a randomised, controlled, parallel-group, unmasked trial. *Lancet.* (2018) 392:1047–57. doi: 10.1016/S0140-6736(18)31880-4
10. Zhu Y, Gu X, Xu C. Effectiveness of telemedicine systems for adults with heart failure: a meta-analysis of randomized controlled trials. *Heart Fail Rev.* (2019) 25:231–43. doi: 10.1007/s10741-019-09801-5
11. Gorodeski EZ, Goyal P, Cox ZL, Thibodeau JT, Reay RE, Rasmusson K, et al. Virtual visits for care of patients with heart failure in the Era of COVID-19: a statement from the heart failure society of America. *J Card Fail.* (2020) 26:448–56. doi: 10.1016/j.cardfail.2020.04.008
12. Salzano A, D'Assante R, Stagnaro FM, Valente V, Crisci G, Giardino F, et al. Heart failure management during the COVID-19 outbreak in Italy: a telemedicine experience from a heart failure university tertiary referral centre. *Eur J Heart Fail.* (2020) 22:1048–50. doi: 10.1002/ejhf.1911
13. Inglis SC, Clark RA, Dierckx R, Prieto-Merino D, Cleland JGF. Structured telephone support or non-invasive telemonitoring for patients with heart failure. (2017) 103:255–7. doi: 10.1136/heartjnl-2015-309191
14. Faragli A, Abawi D, Quinn C, Cvetkovic M, Schlabs T, Tahirovic E, et al. The role of non-invasive devices for the telemonitoring of heart failure patients. *Heart Fail Rev.* (2020). doi: 10.1007/s10741-020-09963-7. [Epub ahead of print].
15. de la Torre-Díez I, López-Coronado M, Vaca C, Aguado JS, de Castro C. Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review. *Telemed J E Health.* (2015) 21:81–5. doi: 10.1089/tmj.2014.0053
16. Böhm M, Drexler H, Oswald H, Rybak K, Bosch R, Butter C, et al. Fluid status telemedicine alerts for heart failure: a randomized controlled trial. *Eur Heart J.* (2016) 37:3154–63. doi: 10.1093/eurheartj/ehw099
17. Galinier M, Roubille F, Berdague P, Brierre G, Cantie P, Dary P, et al. Telemonitoring versus standard care in heart failure: a randomised multicentre trial. *Eur J Heart Fail.* (2020) 22:985–94. doi: 10.1002/ejhf.1906
18. Seferovic PM, Ponikowski P, Anker SD, Bauersachs J, Chioncel O, Cleland JGF, et al. Clinical practice update on heart failure 2019: pharmacotherapy, procedures, devices and patient management. An expert consensus meeting report of the heart failure association of the European society of cardiology. *Eur J Heart Fail.* (2019) 21:1169–86. doi: 10.1002/ejhf.1531
19. Frederix I, Caiani EG, Dendale P, Anker S, Bax J, Böhm A, et al. ESC e-cardiology working group position paper: overcoming challenges in digital health implementation in cardiovascular medicine. *Eur J Prev Cardiol.* (2019) 26:1166–77. doi: 10.1177/2047487319832394
20. Gerke S, Stern AD, Minssen T. Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. *NPJ Dig Med.* (2020) 3:94. doi: 10.1038/s41746-020-0306-7
21. De Filippo O, D'Ascenzo F, Angelini F, Bocchino PP, Conrotto F, Saglietto A, et al. Reduced rate of hospital admissions for ACS during Covid-19 outbreak in Northern Italy. (2020) 383:88–9. doi: 10.1056/NEJMc2009166
22. Cleland JGF, Clark RA, Pellicori P, Inglis SC. Caring for people with heart failure and many other medical problems through and beyond the COVID-19 pandemic: the advantages of universal access to home telemonitoring. *Eur J Heart Fail.* (2020) 22:995–8. doi: 10.1002/ejhf.1864

Conflict of Interest: AF and ELP are shareholders of the company BOCAhealthcare GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

BRIEF RESEARCH REPORT
published: 10 February 2021
doi: 10.3389/fdgth.2020.578902

Early Warning Signs of a Mental Health Tsunami: A Coordinated Response to Gather Initial Data Insights From Multiple Digital Services Providers

Becky Inkster and Digital Mental Health Data Insights Group (DMHDIG)†*

Wolfson College, University of Cambridge, Cambridge, United Kingdom

OPEN ACCESS

Edited by:
Phuong N. Pham,
Harvard Medical School,
United States

Reviewed by:
Aikaterini Bourazeri,
University of Essex, United Kingdom
Milena B. Cukic,
Amsterdam Health and Technology Institute (AHTI), Netherlands

****Correspondence:***
Becky Inkster
becky@beckyinkster.com

†Members of the Digital Mental Health Data Insights Group are listed at the end of the article

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *01 July 2020*
Accepted: *22 December 2020*
Published: *10 February 2021*

Citation:
Inkster B and Digital Mental Health Data Insights Group (DMHDIG) (2021) Early Warning Signs of a Mental Health Tsunami: A Coordinated Response to Gather Initial Data Insights From Multiple Digital Services Providers. Front. Digit. Health 2:578902. doi: 10.3389/fdgth.2020.578902

Introduction: The immediate impact of coronavirus 2019 (COVID-19) on morbidity and mortality has raised the need for accurate and real-time data monitoring and communication. The aim of this study is to document the initial observations from multiple digital services providers during the COVID-19 crisis, especially those related to mental health and well-being.

Methods: We used email and social media to announce an urgent call for support. Digital mental health services providers ($N = 46$), financial services providers ($N = 4$), and other relevant digital data source providers ($N = 3$) responded with quantitative and/or qualitative data insights. People with lived experience of distress, as service users/consumers, and carers are included as co-authors.

Results: This study provides proof-of-concept of the viability for researchers and private companies to work collaboratively toward a common good. Digital services providers reported a diverse range of mental health concerns. A recurring observation is that demand for digital mental health support has risen, and that the nature of this demand has also changed since COVID-19, with an apparent increased presentation of anxiety and loneliness.

Conclusion: Following this study, we will continue to work with providers in more in-depth ways to capture follow-up insights at regular time points. We will also onboard new providers to address data representativeness. Looking ahead, we anticipate the need for a rigorous process to interpret insights from an even wider variety of sources in order to monitor and respond to mental health needs.

Keywords: COVID-19, financial stress, isolation, anxiety, data insights, digital mental health

INTRODUCTION

During the coronavirus 2019 (COVID-19) pandemic, traditional mental health services and related activities declined, in part, due to outpatient clinics being closed to adhere to social distancing requirements, mental health staff redeployment, and inpatient beds being converted into COVID-19 units. As governments attempt to contain the virus, we must mitigate the mental health impact of the pandemic and economic crisis, especially given that pre-COVID-19 predictions already indicated that by 2030, depression will be the leading cause of disease burden globally (1).

During the severe acute respiratory syndrome (SARS) (2002–2004) epidemic, social disengagement, mental stress, and anxiety were associated with increased suicide rates in the elderly population (2). Another study found that 30% of children and 25% of parents who were quarantined or isolated during pandemic diseases met the clinical criteria for post-traumatic stress disorder (3). Data from previous economic depressions and recessions suggest profound increases in substance use disorder, depression, and suicide (4, 5).

In the current pandemic, frontline healthcare workers face the possibility of anxiety and burnout (6, 7) and moral injury (8), alongside fears of becoming ill. This is more pronounced among ethnic minority frontline healthcare workers due to the apparent increased health risks associated with COVID-19 (9). For others living in highly conflicted households, social distancing has meant prolonged social contact and abuse. For example, in the UK, the number of suspected domestic homicide victims more than doubled during the first 3 weeks of the lockdown (10). In France, calls to the national violence against children helpline increased by 89% (11). From an economic perspective, a survey of UK households 3 weeks into the "lockdown" found that 49% of households feel anxious about their finances, rising to 95% among the households experiencing serious financial difficulties (12). A survey conducted in March 2020, just as the lockdown rules were coming into place in the USA, also highlighted higher levels of psychological distress among lower income households (13).

There is a need to obtain more granular and real-time information to help us understand the nature and scale of the mental health crisis. A possible source of this information is the large number of digital mental health services providers used by millions of people globally. These include patient to clinician communication tools, digitally enabled treatments, self-managed care solutions, mental health and well-being apps, online forums, support networks, and digital communities. In addition to this, given the established links between health, social, and economic factors (e.g., (14)), insights should also be obtained from financial services providers and other relevant digital data sources. The potential value of healthcare insights in financial data is already recognized (15, 16), and financial services firms not only are a source of uniquely constructive data on household economies (17) but can also offer possible mechanisms of direct and indirect mental health interventions.

To investigate the impact of COVID-19 on mental health, we set out to collect observations from multiple digital services providers (**Supplementary Table 1**). To our knowledge, this has never been done at scale before, and we did not know how many providers would respond or what the nature of their data insights might be. With rapid turnaround, a diverse range of providers came forward with collective information sourced from a user base of at least 10 million people, but possibly reaching upwards of 50 million globally.

MATERIALS AND METHODS

We used email and social media to announce an urgent call for support to investigate the scale and nature of the mental health impact of COVID-19 [1]. Starting 6 April 2020, BI sent emails to all speakers who had presented at previous "Digital Innovation in Mental Health" (DIMH) conferences [2], as well as to members of the FinHealthTech Consortium [3], and also to a much wider digital community *via* LinkedIn, Twitter, and Facebook. Additionally, we encouraged providers and co-authors to ensure that they sought the views of people in their own networks.

We reached out directly to 55 digital services providers. We received confirmation of support to contribute from 53 providers (i.e., a positive response rate of 96%), which consisted of 46 digital mental health services providers, four financial services providers, and three other digital data source providers ($N = 3$). Respondents were asked to provide qualitative and/or quantitative insights with no exchange of data or identifiable information. A list of digital services providers can be found in **Supplementary Table 1**. This study was purely exploratory. We deliberately did not provide a framework for insights or any analytic specifications (e.g., what specific hypotheses to test). Therefore, all insights should be considered illustrative examples, not primary research.

We asked providers to be compliant with the General Data Protection Regulation (GDPR) and the Data Protection Act 2018 if their users were within Europe. To set a good example of responsible innovation, this document only accepted data insights from providers with clear and accessible privacy policies. Other than these ethical grounds, there were no other exclusion criteria. There were no specific inclusion criteria, but many of the respondents had a pre-existing association with members of the study team, for example, through the annually run conference, DIMH, created by Dr. Becky Inkster[2].

We deliberately did not select a specific methodology for this study, and we did not test any specific hypotheses. Providers collected very different types of data and analyzed it in their own way using techniques that were appropriate for their data. Trying to develop some common methodologies is a future goal, which

[1] Available online at: https://www.beckyinkster.com/covid19 (accessed May 17, 2020).

[2] Available online at: https://www.beckyinkster.com/summer-2021-conference (accessed May 17, 2020).

[3] Available online at: https://www.beckyinkster.com/fhtc (accessed May 17, 2020).

will require more time, and increased collaboration between different providers and other stakeholders.

Data insights and draft versions of the paper were shared among all co-authors for feedback, including from people with a range of lived experiences of distress and service use.

RESULTS

Given the anecdotal nature of many of the insights and the non-systematic way in which providers were chosen, we are reluctant to draw conclusions from the content provided in **Supplementary Tables 2, 3**. Instead, we summarize some of the more frequent observations reported by providers.

Intentions

Insights suggest changes in the type of information individuals are seeking or presenting. From Google Trends data, searches for "anxiety symptoms" doubled between the weeks beginning 8 March and 22 March 2020. In a similar timeframe, Mental Health America (MHA) witnessed a 22% increase in the numbers of Generalized Anxiety Disorder 7-item (GAD7) anxiety screens ($N = 11{,}033$) taken in March 2020 compared with February 2020. Qualitative insights suggest that individuals are seeking practical resources and coping strategies. Participants in the It's Ok To Talk discussion raised questions about anxiety, strategies to manage work, studies, sleep, dealing with domestic violence, and difficult home relationships. Babylon reports that many patients are seeking advice on information about local council support services, seeking advice for activities to keep busy and how to remain healthy, and how to get support to access food and financial concerns. Ieso Digital Health reports up to a third of patients mentioning COVID-19 as a reason for presenting for mental health treatment and also reports a rise in patient worries about viruses, with up to 15% of in-session worries about COVID-19.

Affiliative Tendencies

Papa reported that 53% of users felt less lonely, and that virtual companions have performed a range of tasks with elderly users (e.g., obtaining medications, online grocery shopping). Peer support specialists are being rapidly trained. Digital Peer Support trained 750 peer support specialists between 10 March and mid-April 2020. Wisdo reported a 283% increase in the numbers of people replying to other people's messages and an increase of 115% in the numbers of people signing up for roles to provide support for others. Mentally Aware Nigeria Initiative (MANI) trained over 200 psychosocial support team specialists/counselors on mental health.

Support-Seeking

Many providers are experiencing increased support-seeking behaviors. For example, Ieso Digital Health reports an 84% increase in referrals. Vala Health reports a doubled volume of mental health-related consultations with general practitioners (GPs) during the period 10 March to 8 April 2020. By week 4 of the UK lockdown, general health enquiries had returned to almost pre-COVID levels, but mental health consultations continued to rise. National Alliance on Mental Illness (NAMI) reports a 41% increase in demand for HelpLine resources and information. Ieso Digital Health reported an 84% increase in referrals to their 1–1 online cognitive behavioral therapy (CBT) service in the weeks since the lockdown was announced in the UK, relative to the same period in 2019. Wysa witnessed a 77% increase in new users during February to March 2020, as compared with the same period in 2019. MANI recorded the highest number of emergency calls in the month of April. Qualitative insights from Orygen (Australia) revealed that young people report privacy concerns in having telehealth consults with family members in the background.

Outcomes

Many providers report observations suggesting increased anxiety, uncertainty, loneliness, and loss. MHA reports that 45% of people who took an anxiety screen in March ($N = 11{,}033$) scored for severe anxiety. In a self-reported questionnaire to members of The Mighty, 89% of members reported that their daily life has been at least somewhat impacted by increased anxiety; 43% say that it has been extremely impacted. This is consistent with reports from Kooth demonstrating increases in sadness (up 161%), health anxiety (up 155%), sleep difficulties (up 90%), concerns over body image (up 43%), eating difficulties (up 31%), loneliness (up 23%), and bereavement (up 20%). The Mental Health Foundation survey reported that respondents felt increasingly lonely, and that this was most pronounced for people aged 18–24 (44%) and 25–34 (35%). Multiple providers report users mentioning their loss of access to care and human support [The Mighty, MeeTwo, Wysa, consultant National Health Service (NHS) nurse].

Qntfy's observations suggest decreased well-being in the general public, and that at times, this has been greater among those who identify as healthcare providers. Unmind and Wysa reported higher anxiety levels in health staff than in the rest of their populations. Sangath reports that community health workers face "fears and insecurities among their patients, as well as added anxieties about the health and well-being of their own children and family members." CBTClinics report a rise of anxiety and depressive type disorders from people emotionally close to frontline health staff (e.g., parents, spouses, and children).

Other outcomes include increases in reporting of unsafe domestic settings (Babylon, Wysa, Teen Line, Kooth), suicidal risk/ideation (MeeTwo, Qntfy, Mental chat, Beyond Blue, Mumsnet), and sleep disturbances (It's OK to Talk, Kooth, Mumsnet, Qare, BioBeats, Wysa). There have also been increased prescription of anti-depressant medications (Jasvinder Kandola), increased requests for pain killers *via* telehealth (Vala Health), and increased activities on darknet markets mentioning psychiatric medications (The TellFinder Alliance).

Financial Concerns

Financial insights show an overarching theme of the interrelation between mental health and financial health worries. Three sub-themes emerged from the data: (1) uncertainty and a sense of loss of control particularly "at-risk" individuals and groups; (2) anger, anxiety, and concerns over access to financial support especially

those who feel that they are "falling through the gaps"; and (3) negative mental health and/or physical health with financial health outcomes.

The Money and Mental Health Policy Institute survey (N = 568) reported a range of concerns by respondents with lived experience of mental health problems about how changes, as a result of COVID-19, might affect their finances: 62% worried about having to access the benefits system, 57% worried about losing their job, and 56% worried about creditors chasing them for money. Tully and OpenWrks Group reported that 81% of self-employed customers (N = 650) have declared that they do not have any work coming in due to COVID-19. Furthermore, 50% of their wider sample (N = 1,822) have had income reduced, and 19% have lost their income. The Turn2us survey showed that 70% of respondents (N = 6,198) who have had employment affected are unable to afford rent or mortgages. An anonymous financial services provider also shared concerns that their on-site cashiers may be vulnerable and distressed by customer behaviors.

Qualitative insights also make it clear both how emotive and tangible the impacts of financial concerns and outcomes are on mental health worries and outcomes. For example, "we are… dead… no money no food… 4 weeks in isolation UC no answers… I have no other way to provide for my children and I don't care about the bills… I will have to go out and improvise something."; "what about the thousands who started new jobs to better ourselves after the Feb 28th cut off and before the #coronavirus hit the UK but now sit suicidal in the gap entitled to nothing despite being lifelong tax payers? #newstarterjustice #newstarterprotest #newstarterfurlough"; and "got my letter yesterday to tell me it's being taken away. The welfare system has kicked me when I'm down already, made me physically ill & caused a flare up of my health just when I don't need to go to a hospital mid pandemic."

DISCUSSION

To our knowledge, this study is the first of its kind to bring together a large number of private organizations, including financial services providers, to share digital data insights about the mental health concerns of millions of people online. Our study is novel and radical because this is the first time that something like this has been achieved in this field. Many people questioned the feasibility of being able to bring together a large group of digital services providers (some of which are in competition with each other) to share their insights. We believe that our study provides a proof-of-concept for the viability of using this approach.

The information that we have quickly compiled has been sourced from different geographies, demographics, and types of digital interactions and provides insights into the diversity of individual mental health needs. During our study, a paper (18) called for mental health monitoring to move beyond NHS linkage, in order to capture the real incidence in the community and embrace new technologies measuring moment-to-moment change. This initial snapshot of data that we collected could help inform future studies, for example, it could help the research community to understand what questions could be asked (especially those without expertise in mental health), to aid in the generation of specific hypotheses, or to help with the formulation of prior probabilities. Additionally, we hope that this study increases the research community's awareness of the digital mental health landscape and the services providers who are currently collecting data, as well as the types of data insights and metrics that they might be able to provide.

We recognize that this study is not rigorous in terms of data collection and methodology. We did not choose these providers in a systematic way. Using data from digital services providers limits our population to people who have access to these digital platforms and many "hidden" populations are not registering in digital spaces. Furthermore, we do not know whether our demographic is representative of any larger population or whether whole-population impacts can be inferred from digital service impacts. In addition, we did not verify the insights shared by providers. This avoided privacy issues, but has the potential to have introduced inaccuracies or biases in the reported information. This study is also unable to characterize mental health problems at a clinical level because most digital providers did not report clinically-validated measurements. The use of digital measurements to monitor mental states and distress is still a developing space.

Prior to this study, we did not know what the response rate would be or what types of insights we would be able to obtain. Developing new methodologies to combine such insights will be a substantial undertaking, which should involve many stakeholders. Developing such methodologies is a future goal, but it was not the purpose of this study.

It is important to note that this study was conducted in the midst of the initial pandemic, a time of significant uncertainty. Between the time of the data gathering for and the publication of this study, there have been countless responses across countries announced and enacted. The insights discussed here capture an important moment in time during the initial pandemic phase and also offer a useful reference for on-going data monitoring and subsequent study follow-ups.

Following this study, we will continue to work with providers to capture follow-up insights at later time points[2] and onboard new providers to address issues of data representativeness. We will continue to engage with and include people with a range of experiences of distress and service use, so that we are inclusively influenced by their insights and inputs. It will also be important to capture insights that relate to resilience and recovery. An important next step will be to develop rigorous means to bring together public and private sector data to monitor mental health needs in real-time (just as contact tracing is used to manage the viral epidemic). This can fuel research and understanding and help to inform high-quality responses, which can be delivered remotely to those in need on global and local scales.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: we only had access to the data insights provided by digital services providers, we did not access the

data. Requests to access these datasets should be directed to becky@beckyinkster.com.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements. Digital services providers followed their own in house ethical procedures, terms and conditions and consent procedures for their own data sets.

AUTHOR CONTRIBUTIONS

BI formulated the idea, operationalised and co-ordinated the response, inviting co-authors to join, all having different professional and/or lived experiences who have made important contributions in various ways, such as performing literature searches, writing, helping us to connect with digital providers, idea generation, editing, interpretation, etc.

ACKNOWLEDGMENTS

The Lifeos for additional insights and encouragement during the study.

DIGITAL MENTAL HEALTH DATA INSIGHTS GROUP (DMHDIG)

Becky Inkster, Wolfson College, Cambridge University, Cambridge, UK & The Alan Turing Institute, London, UK; Ross O'Brien, Central and North West London NHS Foundation Trust, and Healthy London Partnership, NHS, UK; Kate Niederhoffer, Knowable Research, Texas, USA; Niranjan Bidargaddi, College of Medicine & Public Health, Flinders University, Adelaide, South Australia, Australia; Anne-Claire Camille Stona, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore; Glen Coppersmith, Qntfy, USA; Amanda Towler, The TellFinder Alliance; The TellFinder Alliance, USA; Philip Resnik, Department of Linguistics and Institute for Advanced Computer Studies, University of Maryland, Maryland, USA; Rebecca Resnik, Rebecca Resnik and Associates, Bethesda, Maryland, USA; Maria Liakata, Queen Mary University of London, UK; The Alan Turing Institute, UK; University of Warwick, UK; Helen Barker, London, UK; Abdullahi Abubakar Kawu, Ibrahim Badamasi Babangida University, Lapai, Nigeria, Africa; Karen Machin, School of Health & Social Work, University of Hertfordshire, UK; Survivor Researcher Network, UK; Pattie Pramila Gonsalves, Sangath, India; Sweta Pal, Sangath, India; Swetha Ranganathan, Sangath, India; John A. Naslund, Department of Global Health and Social Medicine, Harvard Medical School, Harvard University, Massachusetts, USA; Jo Robinson, Orygen, Parkville, Melbourne, Victoria, Australia, Centre for Youth Mental Health & The University of Melbourne, Parkville, Melbourne, Victoria, Australia; Munmun De Choudhury, School of Interactive Computing, Georgia Institute of Technology, Atlanta Georgia, USA; Glenn Melvin, School of Psychology, Faculty of Health, Deakin University, Melbourne, Australia; Terry Hanley, University of Manchester, UK; Matthew Jackman, Lived Experience Academic, Western Pacific Region, Global Mental Health Peer Network, Melbourne, Australia; Ed Humpherson, Director General for Regulation, United Kingdom Statistics Authority, UK; Bo Wang, Department of Psychiatry, University of Oxford; The Alan Turing Institute, UK; Bilal A. Mateen, Kings College Hospital, London, UK & The Alan Turing Institute, London, UK; Akeem Sule, Department of Psychiatry, University of Cambridge, UK; Wolfson College, University of Cambridge, UK; Essex Partnership University NHS Foundation Trust, UK; Ezinne Nwankwo, Cambridge University, UK & Harvard University, USA; Gabriela Pavarini, Department of Psychiatry, University of Oxford; Wellcome Centre for Ethics and Humanities, University of Oxford, UK; Josip Car, Centre for Population Health Sciences; WHO Collaborating Centre for Digital Health and Health Education; Health Services and Outcomes Research, LKCMedicine; Imperial College London, UK; David Crepaz-Keay, Head of Applied Learning, Mental Health Foundation, UK; Fellow, Royal Society for Public Health, UK; Jasvinder Kandola, Division of Medicine, Hammersmith Hospital Imperial College London, UK; Hannah Stewart, The University of Texas Health Science Center at Houston (UT Health) School of Public Health, Department of Health Promotion & Behavioral Sciences, Texas, USA; Eiman Kanjo, Nottingham Trent University, Smart Sensing Lab (MA220), Clifton Lane, Nottingham, UK; Sarah Ticho, Hatsumi, London, UK; April C. Foreman, American Association of Suicidology, Louisiana, USA; Emma Selby, Digital Mentality, London, UK; Stan Shepherd, Instant Access Medical, London, UK; Karen L Fortuna, Dartmouth College, Hanover, New Hampshire, USA; Emachi Eneje, Birmingham Mind, UK; Tamra Huesers, Harmony Center, Minot, North Dakota, USA; Stephen Jeffreys, Survivor Researcher Network, London, UK; Mat Rawsthorne, NIHR Biomedical Research Centre for Mental Health & Technology, University of Nottingham, UK; Gerry Craigen, Department of Psychiatry, Faculty of Medicine, University of Toronto & Associate Attending Staff Psychiatrist, Department of Psychiatry, University Health Network, Toronto General Hospital Toronto, Canada; Kristina Barger, Cogenta, UK; Neha Kumar, Georgia Institute of Technology, Atlanta, USA; Sachin Pendse, Georgia Institute of Technology, Atlanta, USA; Errin Riley, Sense About Science, London, UK; Elvira Perez Vallejos, Nottingham NIHR Biomedical Research Centre for Mental Health, UK & Institute of Mental Health, Nottingham University, UK; Mark Embrett, Dalhousie University, Canada; Ernest Okyere-Twum, Universite Paris Descartes, France & Centre for mental health research in Africa (CEMHRA); Kumar Jacob, MindWave Ventures, UK; Janak Gunatilleke, MindWave Ventures, UK; Mirantha Jayathilaka, MindWave Ventures, UK; Mariana Pinto Da Costa, Unit for Social and Community Psychiatry, WHO Collaborating Centre for Mental Health Services Development, Queen Mary University of London, London, UK; Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal;

Hospital de Magalhães Lemos, Porto, Portugal; Ana Catarino, Ieso Digital Health; Ronan Cummins, Ieso Digital Health; Tom Clelford, Ieso Digital Health; James de Bathe, Ieso Digital Health; Valentin Tablan, Ieso Digital Health; Sarah Bateup, Ieso Digital Health; Andrew D Blackwell, Ieso Digital Health; Tejal Patel, Babylon; Keith Grimes, Babylon; Ed Sykes, Huma Therapuetics, UK; Pete Trainor, Vala Health; Daf Rakphetmanee, Ooca; Kanpassorn Eix Suriyasangpetch, Ooca; Annie Meharg, Kooth; Aaron Sefi, Kooth; Derek Richards, SilverCloud Health and Trinity College Dublin; Angel Enrique, SilverCloud Health and Trinity College Dublin; Jorge Palacios, SilverCloud Health and Trinity College Dublin; Antony Brown, CBTClinics; Eva Papadopoulou, Minddistrict; Charlotte Lee, Big Health; Fanny Jacq, Qare; Loïc Tse, Qare; David Plans, Huma Therapuetics, UK; Senior Lecturer in Organisational Neuroscience, University of Exeter, UK; Anika Sierk, Unmind; Heather Bolton, Unmind; Knut Schroeder, Expert Self Care; Tarek R. Besold, Alpha Health, Telefonica Innovation Alpha; Institute of Cognitive Science, University of Osnabrueck, Germany; Aleksandar Matic, Alpha Health, Telefonica Innovation Alpha; Department of Psychological and Behavioural Science, London School of Economics, London, UK; Iñaki Estella Aguerr, Alpha Health, Telefonica Innovation Alpha; Department of Electrical and Electronic Engineering, Imperial College London, London, UK; Liz Ashall-Payne, ORCHA; Rob Daly, ORCHA; Simon Leigh, ORCHA; Jo Aggarwal, Wysa; Ramakant Vempati, Wysa; Smriti Joshi, Wysa; Vinod Subramanian, Wysa; Madhura Kadaba, Wysa; Clara Falala Sechet, Owlie; Geby Chyntia Irwan, Riliv; Audrey Maximillian Herli, Riliv; Karine Chevreul, StopBlues; Anais Le Jeannic, StopBlues; Kathleen Turmaine, StopBlues; Christopher Rainbow, BeyondBlue; Megan Chor Kwan Lam, Neurum Health; Christine Hiu Man Chiu, Neurum Health; Will Allen-Mersh, Spill; Justine Roberts, Mumsnet; Sara Ray, The Mighty; Angelica Catalano, The Mighty; Jennifer Russell, TalkLife; TalkCampus; Jamie Druitt, TalkLife; TalkCampus; Boaz Gaon. Wisdo; Suzi Godson, MeeTwo; Kerstyn Comley, MeeTwo; Satu Raappana, MIELI Mental Health Finland Mental-chat & Mental Gaming, Finland; Michelle Carlson, Teen Line, USA; Andrew Parker, Papa; Ken Duckworth, National Alliance on Mental Illness, USA; Dan Gillison, National Alliance on Mental Illness, USA; Theresa Nguyen, Mental Health America, USA; Madeline Reinert, Mental Health America, USA; Victor Ugo, Mentally Aware Nigeria Initiative (MANI), Nigeria; Ifedayo Ward, Mentally Aware Nigeria Initiative (MANI), Nigeria; Chantelle Booysen, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Ashley Foster-Estwick, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Grace Gatera, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; David Karorero, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Kumba Philip-Joe, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Damian Juma, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Claudia Sartor, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Chinwendu Ukachukwu, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Lian Zeitz, Young Leaders for the Lancet Commission on Global Mental Health and Sustainable Development; Alex Fine, Qntfy, USA; Merlyn Holkar, Money & Mental Health Policy Institute, UK; Conor D'Arcy, Money & Mental Health Policy Institute, UK; Katie Alpin, Money & Mental Health Policy Institute, UK; Jo Kerr, Turn2Us, UK; Lee Healey, IncomeMax, UK; Olly Betts, Tully and OpenWrks Group, the team behind Tully; Andrea Severino, Healthy Virtuoso, Italy; Will Van Der Hart, The Mind and Soul Foundation, UK; Danielle Smalls, The TellFinder Alliance, USA; Chris Dickson, The TellFinder Alliance, USA; Andrew Stroz, The TellFinder Alliance, USA; Sebastian Vollmer, Warwick University, Warwick, UK & The Alan Turing Institute, London, UK; Hoang D. Nguyen, School of Computing Science, University of Glasgow, Singapore; Daniel Albert Rosello, Nottingham Trent University, Nottingham, UK; Valentino Megale, Softcare Studios, Rome, Italy; Jan D. Smeddinck, Open Lab, Newcastle University, Newcastle upon Tyne, UK; Rosanna Bellini, Open Lab, Newcastle University, Newcastle upon Tyne, UK; Craig A. DeLarge, The Digital Mental Health Project, Wise Working, California, USA; Shivani Patel, South London and Maudsley NHS Trust, London, UK; Jerome Uriko-kang, Global Mental Health Peer Network, Ghana, Africa; Tunde Olatunji, Lyrical Kombat, London, UK; Vanessa Lalo, Liberal Clinical Psychologist, Paris, France; Robert Walker, Department of Mental Health Office of Recovery and Empowerment, Massachusetts Department of Mental Health, USA; Ann John, Population Data Science, National Centre for Mental Health, Swansea University, Swansea, Wales, UK; Diana Rayes, The Johns Hopkins Bloomberg School of Public Health, Baltimore, USA; Marwa Elnahass, Newcastle University, UK; Karen Elliott, Newcastle University, UK; Lil Tonmyr, Family Violence, Mental Health & Suicide Surveillance Team, Behaviours, Environments and Lifespan Division, Centre for Surveillance and Applied Research, Public Health Agency of Canada; Andrew MacKenzie, Centre for Surveillance and Applied Research, Public Health Agency of Canada; Michael L. Birnbaum, The Zucker Hillside Hospital, Psychiatry Research, Northwell Health, New York, USA; Eric D. Caine, University of Rochester Medical Center, New York, USA; John Pestian, Cincinnati Children's Hospital Medical Center, University of Cincinnati; Oak Ridge National Laboratory/VA-MPV Champion program, Cincinnati, USA; Dan Jacobson, Oak Ridge National Laboratory/VA-MPV Champion program, Cincinnati, USA; Mike Sorter, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, USA; Tracy Glauser, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, USA; Michael Meaney, Translational Neuroscience programme, Singapore Institute for Clinical Sciences, Singapore & McGill University, Quebec, Canada; Vincent M. B. Silenzio, Rutgers School of Public Health, Rutgers University, New Brunswick and Newark, New Jersey, USA; Jenny Edwards, London, UK; Ricardo Araya, Centre for Global Mental Health, King's College London, UK; Chris Fitch, Personal Finance Research Centre, University of Bristol, UK; Jamie Evans, Personal Finance Research Centre, University of Bristol,

UK; Kevin Telford, University of Edinburgh, Scotland, UK; Peggy Loo, Legal & General Group Plc, UK; Andrea Stevenson, Independent Consultant, London, UK; Tatyana Marsh, Open Banking Excellence, UK; Helen Child, Open Banking Excellence, UK; Roger S. McIntyre, Mood Disorders Psychopharmacology Unit, University Health Network, Toronto, Canada & Institute of Medical Science, University of Toronto, Toronto, Canada & Department of Pharmacology, University of Toronto, Toronto, Canada & Department of Psychiatry, University of Toronto, Toronto, Canada & Brain and Cognition Discovery Foundation, Toronto, Canada; Henrietta Bowden-Jones, National Centre for Behavioural Addictions, UK (National Problem Gambling Clinic + National Centre for Gaming Disorders) & Medical Women's Federation & Royal Society of Medicine & Royal College of Psychiatrists & University College London, UK; John Torous, Digital Psychiatry Division, Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Harvard University, Massachusetts, USA; Thomas R. Insel, Humanest Care, USA.

ORCID IDS

Becky Inkster: 0000-0003-1201-455X; Anne-Claire C. Stona: 0000-0003-0350-0345; Philip Resnik: 0000-0002-6130-8602; Terry Hanley: 0000-0001-5861-9170; David Crepaz-Keay: 0000-0003-3845-4721; Neha Kumar: 0000-0002-7014-5585; Ann John: 0000-0002-5657-6995; Mariana Pinto Da Costa: 0000-0002-5966-5723; Mark Embrett: 0000-0002-3969-0219; Munmun De Choudhury: 0000-0002-8939-264X; Mat Rawsthorne: 0000-0002-7481-693X; Vincent M. B. Silenzio: 0000-0003-1408-7955; Aleksandar Matic (Alpha Health): 0000-0002-8752-4098; Sebastian Vollmer: 0000-0002-9025-0753; Jo Robinson: 0000-0001-5652-918X; Pete Trainor (Vala Health): 0000-0002-7778-3934; Karen L Fortuna: 0000-0003-0343-2346; Pattie Pramila Gonsalves: 0000-0003-3780-4523; Smriti Joshi: 0000-0001-7446-2804; Hannah Stewart: 0000-0003-2536-9405; Tarek R. Besold (Alpha Health): 0000-0002-8002-0049; Mirantha Jayathilaka: 0000-0002-2462-4833; Gabriela Pavarini: 0000-0001-5574-4021; Sarah Bateup: 0000-0003-3926-0021; Iñaki Estella Aguerri: 0000-0001-5110-6858; Janak Gunatilleke: 0000-0003-1474-5735; Hoang D. Nguyen: 0000-0003-2541-3269; Ricardo Araya: 0000-0002-0420-5148; Karen Machin: 0000-0002-0374-4238; Glenn Melvin: 0000-0002-6958-3908; John A. Naslund: 0000-0001-6777-0104; Stephen Jeffreys: 0000-0002-5088-9309; John Torous: 0000-0002-5362-7937; Thomas R. Insel: 0000-0001-5031-0160; Bo Wang: 0000-0002-3412-3768; Abdullahi Abubakar Kawu: 0000-0003-2531-9539; Lil Tonmyr: 0000-0002-8722-7616; Clara Falala-Séchet (Owlie): 0000-0003-3331-5255; Maria Liakata: 0000-0001-5765-0416; Simon Leigh: 0000-0002-6843-6447; Liz Ashall-Payne: 0000-0001-7325-195X; Karen Elliott: 0000-0002-2455-0475; Sarah Ticho: 0000-0002-5585-0497; Bilal A. Mateen: 0000-0003-4423-6472; David Plans: 0000-0002-0476-3342; Rosanna Bellini: 0000-0002-2223-2801; Tejal Patel: 0000-0002-7356-7054; Derek Richards: 0000-0003-0871-4078; Jorge Palacios: 0000-0002-2103-5507; Angel Enrique: 0000-0003-0585-4008; Niranjan Bidargaddi: 0000-0003-2868-9260; Valentino Megale: 0000-0003-2150-696X; Elvira Perez Vallejos: 0000-0002-0258-9440; Aaron Sefi: 0000-0002-0776-3858.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2020.578902/full#supplementary-material

REFERENCES

1. World Health Organisation. *Global Burden of Mental Disorders and the Need for a Comprehensive, Coordinated Response From Health and Social Sectors at the Country Level*. (2011). Available online at: https://apps.who.int/gb/ebwha/pdf_files/EB130/B130_9-en.pdf (accessed May 17, 2020).
2. Yip PS, Cheung YT, Chau PH, Law YW. The impact of epidemic outbreak: the case of severe acute respiratory syndrome (SARS) and suicide among older adults in Hong Kong. *Crisis*. (2010) 31:86–92. doi: 10.1027/0227-5910/a000015
3. Sprang G, Silman M. Posttraumatic stress disorder in parents and youth after health-related disasters. *Disaster Med Public Health Prep*. (2013) 7:105–10. doi: 10.1017/dmp.2013.22
4. Granados JAT, Roux AVD. Life and death during the great depression. *Proc Natl Acad Sci USA*. (2009) 106:17290–5. doi: 10.1073/pnas.0904491106
5. Gunnell D, Chang SS. Economic recession, unemployment, and suicide. In: O'Connor RC, Pirkis J, editors. *The International Handbook of Suicide Prevention*. New Jersey; New York, NY: John Wiley & Sons, Ltd. (2016). p. 284–300. doi: 10.1002/9781118903223.ch16
6. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. *JAMA Netw Open*. (2020) 3:e203976. doi: 10.1001/jamanetworkopen.2020.3976
7. Tan BYQ, Chew NWS, Lee GKH, Jing M, Goh Y, Yeo LLL, et al. Psychological impact of the COVID-19 pandemic on health care workers in Singapore. *Ann Intern Med*. (2020) 17:M20-1083. doi: 10.7326/M20-1083
8. Greenberg N, Docherty M, Gnanapragasam S, Wessely S. Managing mental health challenges faced by healthcare workers during covid-19 pandemic. *BMJ*. (2020) 368:m1211. doi: 10.1136/bmj.m1211
9. *Impact of COVID-19 on Black, Asian and Minority Ethnic (BAME) Staff in Mental Healthcare Settings|Assessment and Management of Risk*. (2020). Available online at: https://www.rcpsych.ac.uk/news-and-features/latest-news/detail/2020/06/11/more-support-needed-for-bame-psychiatrists-during-the-pandemic-according-to-rcpsych-survey (accessed May 17, 2020).
10. *Domestic Abuse Killings 'More Than Double' Amid Covid-19 Lockdown*. (2020). Available online at: https://www.theguardian.com/society/2020/apr/15/domestic-abuse-killings-more-than-double-amid-covid-19-lockdown (accessed May 17, 2020).
11. *Child Abuse: L'express. Violences intrafamiliales: lors du confinement, les appels au 119 ont presque doublé*. (2020). Available online at: https://www.leparisien.fr/faits-divers/violences-intrafamiliales-les-appels-au-119-ont-presque-double-22-04-2020-8304171.php (accessed May 17, 2020).
12. Kempson E, Poppe C, Standard Life Foundation. *Coronavirus Financial Impact Tracker*. (2020). Available online at: https://www.standardlifefoundation.org.uk/__data/assets/pdf_file/0030/57486/COVID-19-Financial-Impact-Tracker-April-2020-FINAL.pdf (accessed May 17, 2020).
13. Keeter S. *People Financially Affected by COVID-19 Outbreak Are Experiencing More Psychological Distress Than Others*. Pew Research Center (2020). Available online at: https://www.pewresearch.org/fact-tank/2020/03/30/

people-financially-affected-by-covid-19-outbreak-are-experiencing-more-psychological-distress-than-others/ (accessed May 17, 2020).
14. Richardson T, Elliott P, Roberts R. The relationship between personal unsecured debt and mental and physical health: a systematic review and meta-analysis. *Clin Psychol Rev*. (2013) 33:1148–62. doi: 10.1016/j.cpr.2013.08.009
15. Inkster B, Loo P, Mateen B, Stevenson A. Improving insights into health care with data linkage to financial technology. *Lancet Digit Health*. (2019) 1:110–2. doi: 10.1016/S2589-7500(19)30061-5
16. Skatova A, Shiells K, Boyd A. Attitudes toward transactional data donation and linkage in a longitudinal population study: evidence from the Avon Longitudinal Study of Parents and Children. *Wellcome Open Res*. (2019) 4:192. doi: 10.12688/wellcomeopenres.15557.1
17. Frankham C, Richardson T, Maguire N. Psychological factors associated with financial hardship and mental health: a systematic review. *Clin Psychol Rev*. (2020) 77:101832. doi: 10.1016/j.cpr.2020.101832
18. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry*. (2020) 7:547–60. doi: 10.1016/S2215-0366(20)30168-1

Conflict of Interest: BI is an advisor to Wysa and TalkLife and has worked with the majority of these providers previously via either her Digital Innovation in Mental Health conferences and/or her FinHealthTech Consortium. KJ, JG, and MJ are employed by MindWave Ventures. AC, RC, TC, JB, VT, SB, and AB are employed by Ieso Digital Health. TP, KG, and ES are employed by Huma Therapuetics. PT is employed by Vala Health. DR and KS are employed by Ooca. AM and AS are employed by Kooth. DR, AE, and JP are employed by Silver Cloud Health and Trinity College Dublin. AB is employed by CBTClinics. EP is employed by Minddistrict. CL is employed by Big Health. FJ and LT are employed by Qare. DP is employed by Huma Therapeutics. AS and HB are employed by Unmind. KS is employed by Expert Self Care Ltd (distrACT app). TB, AM, and IA, Alpha Health, Telefonica Innovation Alpha. LA-P, RD, and SL work with ORCHA. JA RV, SJ, VS, and MK are employed by Wysa. CS is employed by Owlie. GI and AH are employed by Riliv. KC, AL, and KT are employed by Stop Blues. CR is employed by BeyondBlue. ML and CC are employed by Neurum Health. WA-M is employed by Spill. JR is employed by Mumsnet. SR and AC are employed by The Mighty. JR and JD are employed by TalkLife. BG is employed by Wisdo. SG and KC are employed by MeeTwo. SR is employed by MIELI Mental Health Finland Mental-chat & Mental Gaming, Finland. MC is employed by Teen Line, USA. AP is employed by Papa. KD and DG are employed by National Alliance on Mental Illness, USA. TN and MR are employed by Mental Health America, USA. VU and IW are employed by Mentally Aware Nigeria Initiative (MANI), Nigeria. CB, AF-E, GG, DK, KP-J, DJ, CS, CU, and LZ are part of the Young Leaders for the Lancet Commission on GlobalMental Health and Sustainable Development. GC and AF are employed by Qntfy, USA. MH, CD'A, and KA are employed by Money & Mental Health Policy Institute, UK. JK is employed by Turn2Us, UK. LH is employed by IncomeMax, UK. OB is employed by Tully and OpenWrks Group, the team behind Tully. AS is employed by Healthy Virtuoso, Italy. WV is employed by The Mind and Soul Foundation, UK. AT, DS, CD, and AS work for The TellFinder Alliance, USA. SV has received funding from iqvia for toolbox development. TI works for Humanest Care.

frontiers in Digital Health

POLICY AND PRACTICE REVIEWS
published: 22 March 2021
doi: 10.3389/fdgth.2021.617106

OPEN ACCESS

Edited by:
Björn Wolfgang Schuller, Imperial College London, United Kingdom

Reviewed by:
Niamh Lennox-Chhugani, TaoHealth Research, United Kingdom
Milena B. Cukic, Amsterdam Health and Technology Institute (AHTI), Netherlands

***Correspondence:**
Peter K. Sorger
peter_sorger@hms.harvard.edu; sorger_admin@hms.harvard.edu
Ben Linville-Engler
benle@mit.edu

†ORCID:
Marc-Joseph Antonini
orcid.org/0000-0002-9774-1483
Deborah Plana
orcid.org/0000-0002-4218-1693
Shriya Srinivasan
orcid.org/0000-0002-2508-1324
Lyla Atta
orcid.org/0000-0002-6113-0082
Aditya Achanta
orcid.org/0000-0002-7610-3538
Helen Yang
orcid.org/0000-0002-9455-5300
Avilash K. Cramer
orcid.org/0000-0003-0014-8921
Jacob Freake
orcid.org/0000-0002-5198-835X
Michael S. Sinha
orcid.org/0000-0002-9165-8611
Sherry H. Yu
orcid.org/0000-0002-1432-9128
Nicole R. LeBoeuf
orcid.org/0000-0002-8264-834X
Ben Linville-Engler
orcid.org/0000-0002-1251-8275
Peter Sorger
orcid.org/0000-0002-3364-1838

‡*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *14 October 2020*
Accepted: *17 February 2021*
Published: *22 March 2021*

A Crisis-Responsive Framework for Medical Device Development Applied to the COVID-19 Pandemic

Marc-Joseph Antonini [1,2,3,4†‡], ***Deborah Plana*** [1,4,5†‡], ***Shriya Srinivasan*** [1,6,7,8†‡], ***Lyla Atta*** [1,9†], ***Aditya Achanta*** [1,10†], ***Helen Yang*** [1,11†], ***Avilash K. Cramer*** [1,4†], ***Jacob Freake*** [1,12†], ***Michael S. Sinha*** [1,11†], ***Sherry H. Yu*** [1,13†], ***Nicole R. LeBoeuf*** [1,14†], ***Ben Linville-Engler*** [1,15,16*†] ***and Peter K. Sorger*** [1,5,11*†]

[1] *Greater Boston Pandemic Fabrication Team (PanFab) c/o Harvard-MIT Center for Regulatory Science, Harvard Medical School, Boston, MA, United States,* [2] *Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, United States,* [3] *McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, United States,* [4] *Harvard-MIT Division of Health Sciences and Technology Program, Cambridge, MA, United States,* [5] *Department of Systems Biology, Harvard Ludwig Cancer Research Center and Harvard Medical School, Boston, MA, United States,* [6] *Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, United States,* [7] *Division of Gastroenterology, Hepatology and Endoscopy, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States,* [8] *David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, United States,* [9] *Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD, United States,* [10] *Harvard Medical School, Boston, MA, United States,* [11] *Harvard-MIT Center for Regulatory Science, Harvard Medical School, Boston, MA, United States,* [12] *Fikst Product Development, Woburn, MA, United States,* [13] *Department of Dermatology, Yale University School of Medicine, New Haven, CT, United States,* [14] *Department of Dermatology, Center for Cutaneous Oncology, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Boston, MA, United States,* [15] *System Design and Management, Massachusetts Institute of Technology, Cambridge, MA, United States,* [16] *Massachusetts Manufacturing Emergency Response Team (MA M-ERT), Massachusetts Technology Collaborative, Westborough, MA, United States*

The disruption of conventional manufacturing, supply, and distribution channels during the COVID-19 pandemic caused widespread shortages in personal protective equipment (PPE) and other medical supplies. These shortages catalyzed local efforts to use nontraditional, rapid manufacturing to meet urgent healthcare needs. Here we present a crisis-responsive design framework designed to assist with product development under pandemic conditions. The framework emphasizes stakeholder engagement, comprehensive but efficient needs assessment, rapid manufacturing, and modified product testing to enable accelerated development of healthcare products. We contrast this framework with traditional medical device manufacturing that proceeds at a more deliberate pace, discuss strengths and weakness of pandemic-responsive fabrication, and consider relevant regulatory policies. We highlight the use of the crisis-responsive framework in a case study of face shield design and production for a large US academic hospital. Finally, we make recommendations aimed at improving future resilience to pandemics and healthcare emergencies. These include continued development of open source designs suitable for rapid manufacturing, education of maker communities and hospital administrators about rapidly-manufactured medical devices, and changes in regulatory policy that help strike a balance between quality and innovation.

Keywords: personal protective equipment (PPE), COVID-19, manufacturing, prototyping, biocompatibility, 3D printing, regulatory sciences, medical device design

INTRODUCTION

Rapid Product Development to Meet Emergent Shortages of Medical Supplies

In the face of a global COVID-19 pandemic, widespread disruption of international supply chains and local distribution networks has led to severe shortages in personal protective equipment (PPE) and other medical equipment such as ventilators (1). These shortages have spurred numerous local efforts to supply alternative products. Such efforts involve a diverse community of scientists, engineers, physicians, hobbyists (the "maker" community), community-based organizations, and industrial manufacturers not previously involved in supplying healthcare products. Numerous international collaborations have been formed, anchored in open-source designs, rapid dissemination of pre-prints (on medRxiv or bioRxiv) and repositories such as the National Institute of Health's 3D Print Exchange (2). Such non-traditional fabrication of medical equipment is made feasible by the widespread availability and low cost of manufacturing techniques including 3D printing and laser cutting. These approaches are ideal for low-volume production of face shields, masks, frames for N95 filtering facepiece respirators ("N95 masks"), swabs for diagnostic kits, and potentially more complex medical products such as ventilator parts (3, 4). Many of these devices are safety-critical items designed to control infection risk or sustain life. There is therefore reason for concern about medical products that are manufactured using non-traditional methods and supplied by relatively inexperienced fabricators. We nonetheless propose that the capacity for crisis-responsive local manufacturing be further developed so that it can contribute to resilience to pandemics and healthcare emergencies at local, national, and international levels. By analogy, local repair and rebuilding capacities have long been recognized as critical aspects of resilience to natural disasters (5).

A Crisis-Responsive Design Framework

One of the greatest challenges facing non-traditional producers of medical equipment is the complex and unfamiliar regulatory landscape in place for safety-critical products. Thus, in an emergency setting, design validation and testing—not initial design and final fabrication—are often the biggest barriers to the introduction of new or alternative products. As a consequence, there have been multiple instances in which maker communities or small manufacturers have created a needed product, only to find it turned away by healthcare providers and hospitals (6). The primary goal of this perspective is to prevent such situations by providing an overview of medical device development to makers, engineers, and manufacturers who are not traditionally involved in the medical industry. We also elaborate on the development of design and regulatory frameworks relevant to future emergencies, with a focus on PPE and similar "low-risk" medical devices. We end with some considerations for regulators that could be applied to future pandemics.

Like the traditional framework for medical device development, the crisis-responsive framework outlined here incorporates systems-level interactions among producers and stakeholders that impact product development, testing, and deployment. In a crisis however, it is necessary to reframe a traditionally deliberate, iterative, and highly controlled process for medical device development into a methodology that can be performed on an accelerated timescale with unfamiliar stakeholders and without compromising product safety. Use of a crisis-responsive framework ensures that hospital incident commands, healthcare leadership, institutional review boards (ethics committees), product designers, and fabricators can work efficiently together in pursuit of enhanced resiliency to medical emergencies.

Comparing Traditional and Crisis-Responsive Design Frameworks for Medical Device Development

A variety of models have been developed to describe the different stages of medical device development and their relationships to each other (7–12). Key steps include: (1) problem definition and needs assessment, (2) solution definition, verification, and validation and (3) regulatory approval and implementation (**Figure 1**). Here we highlight two development models: the traditional waterfall process, first developed in 1970 to describe software development (13) and historically used by most medical device manufacturers, and a crisis-responsive framework, better suited to tackle the rapidly-changing demands of the pandemic. The later framework borrows from "agile product development" (14) and emphasizes flexibility, rapid implementation of new features to respond to changing requirements, and fast delivery of a working product (**Figure 1**). While the waterfall model emphasizes feedback and iteration primarily at the product validation stages when a design has been fully implemented (15), the crisis-responsive model involves review and iteration at earlier stages in a design; this is essential because it is rarely possible to undertake formal market research or systematic needs assessment under pandemic conditions. From the perspective of time scales, agile development parallelizes steps to the extent possible to eliminate waiting periods. Use of agile product development and rapid manufacturing makes it possible to create finished prototypes on a time scale of days to weeks as opposed to months to years, as in the case of traditional waterfall-type development, facilitating iterative design and testing with end-users. In a healthcare setting this is likely to include senior physicians and hospital leaders who would not normally be involved in PPE selection. Crisis-responsive development relies on the willingness of hospital stakeholders to consider unfamiliar, innovative, and more costly designs based on an assessment of risks posed by the unavailability of traditional products.

In conventional needs assessment, the impact of selling price, access to retail and wholesale channels, and brands are carefully considered; channel access and branding allow commodity products lacking strong intellectual property protection (e.g., face shields, N95 masks, gowns) to sell at a premium price. Nonetheless, the pressure on price is high, and low price margins pose the primary limitation on innovation. In many cases, low margins make domestic production infeasible causing a small number of overseas manufacturers to provide the majority of PPE products. For proprietary products, margins

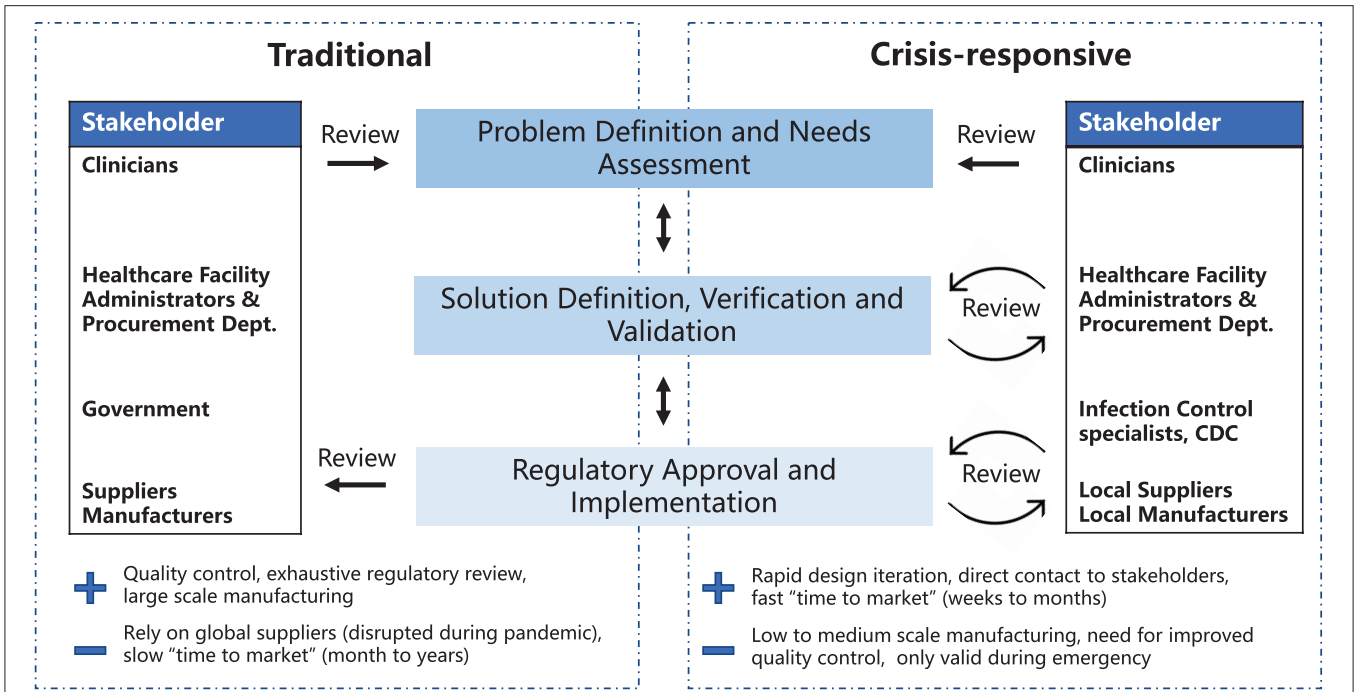


FIGURE 1 | Traditional and crisis-responsive design framework for medical device development. The crisis-responsive framework places emphasis on repeated design review using input from a broad range of stakeholders, all of whom may face unique constraints due to the pandemic.

are typically higher and innovation more important, but the willingness of third parties or national healthcare systems to reimburse for a product is a major consideration. Price sensitivity varies among private and public health systems, nursing homes, and independent living facilities, leading to a plethora of functionally similar products distinguished primarily in branding and distribution channels.

In a crisis, however, the prioritization of these concerns is shifted because the goal is typically to produce the best possible product in the shortest amount of time. Given limitations in fabrication facilities, materials, and the skill of the design team, price and branding are deprioritized, particularly for items such as PPE that suddenly become essential and require significant effort for procurement teams to obtain in volume. Similarly, access to retail and wholesale hospital supply channels, which is typically dominated by a small number of large companies, becomes a secondary consideration during a pandemic. This is particularly true because in the COVID-19 pandemic it is precisely the failure of traditional supply chains to meet urgent requirements that has created the need for nontraditional suppliers. These changes fundamentally alter the stakeholder landscape and the design process. Getting products from a fabricator to a customer is still essential, and involves procurement departments, but is made easier when production is local to the customer and the customer is directly engaged in specifying and testing prototypes.

A crisis-responsive framework has many potential limitations in terms of regulatory compliance and sustainability and is not suitable for use under non-emergency conditions; it is intended to provide stopgap solutions to meet immediate needs. Crisis-responsive design typically does not include the documentation needed for regulatory review and often relies on small-scale production. Designs are sensitive to unanticipated substitution of input materials due to supply shortages. Brand identity is rarely considered, and the analysis of intellectual property may be incomplete. Despite these limitations, even in a crisis it is essential that a rational and well-considered process be followed to ensure that products are functional, reliable, and as safe as possible. Only then can manufacturing by local communities help rather than hinder emergency response. Governments also have an important role to play in creating emergency authorizations and temporarily overriding some intellectual property protections.

STAGE 1: PROBLEM DEFINITION AND NEEDS ASSESSMENT

The Importance of Stakeholder Input

Traditionally, problem definition involves assessing the needs of end-users or healthcare systems through market research. Alternatively, in a crisis, it is common for designers and fabricators to work directly with end-users, such as healthcare workers, rather than with traditional procurement departments. Health care workers will be most concerned with the usability, reliability, and testing of a product and least concerned with branding and cost. Design, manufacturing, distribution, risk-mitigation, and lifecycle considerations (e.g., sterilization) remain the purview of the design and fabrication teams, but we have found that end-users are often willing to engage in issues of

design and fabrication. Notably, direct contact between designers and users provides a rare opportunity for innovation in areas such as PPE, a type of product for which new devices are slow to emerge despite long-known deficiencies in current products.

The process of defining requirements and engaging stakeholders will differ in private and public hospitals, private practices, nursing homes, and independent living facilities but in general, it is end-users who will drive the process. Designers may need to coordinate with individuals empowered by hospital administrators, hospital incident commands (in charge of emergency response) (16), purchasing and procurement departments, and hospital administrators in order to better understand current needs. Local and state government officials can also be a resource for regulatory, purchasing power, and supply chain information and should be consulted if possible. In many cases, non-traditional medical products use components that were manufactured for other purposes (e.g., vacuum cleaner filters for use in PAPRs). Local suppliers and distributors can be an invaluable source of information on the availability of such materials and equipment and their technical performance. We have found that, during the COVID-19 pandemic, many materials suppliers are willing to provide extra help to fabricators who are not part of their traditional customer base. A final and important aspect of needs assessment is soliciting requirements from a diversity of end-users who differ in gender, body size and shape (e.g., differing face dimensions in the context of respirators) and also in clinical roles (e.g., nursing staff, physicians in emergency rooms, outpatient consultants, orderly staff, custodial staff). Non-traditional fabricators must take care to reduce inequities in the workplace and in patient access to health care, not amplify them.

Coordinating With Multiple Stakeholders

At the outset of the COVID pandemic, many municipalities and even hospitals had their own design and fabrication teams working largely independently of each other, although often using shared designs and methods. Several months into the pandemic, particularly after the first wave of hospitalization passed, local fabrication teams started to work together to improve efficiency and share expertise. State programs such as Massachusetts Manufacturing Emergency Response Team [M-ERT (17, 18)] and national efforts such as America Makes (19) are playing an increasingly important role in matching end-users with suppliers and in providing access to tested designs, materials, and supply chains. Including such groups in the design and fabrication processes can bring substantial benefits in terms of the suitability of the design and feasibility of fabrication.

Traditional and Nontraditional Supply Procurement

The first stage in meeting urgent supply shortages is to look for alternative medical manufacturers of similar products. When such products are not available, an alternative solution is to find non-medical suppliers of functionally related products and components. For instance, the Greater Boston Pandemic Fabrication Team (PanFab) (20) PAPR design uses commercially available high-efficiency particulate air (HEPA) vacuum filters since supply shortages have made it challenging to procure filters traditionally used in healthcare settings (20). In addition to meeting device shortages, looking "sideways" in the supply chain through the creation of modified products can help meet PPE shortages in novel ways. For instance, many N95 respirators become unusable after several donning and doffing cycles due to poor fit of the nosepiece (the metal tabs often become distorted) and degradation or breakage of the elastic straps that hold masks in place. Because manufacturing N95-type respirators requires highly specialized fabrics and equipment, it is more feasible to fix the problems with existing masks than to make new ones. As a result, multiple groups have developed 3D printed mask frames that can fit over existing N95 masks and take the place of degraded or broken nosepieces and straps (21–27). Like many other innovative products developed to meet emergency needs, mask frames may also have a role in respiratory protection under non-crisis conditions.

Determining Raw Materials Needs

Needs assessment in a crisis must not only consider the requirements of end-users, but also the capabilities of manufacturers and suppliers. During a pandemic, acquiring raw materials is often challenging, as suppliers may be closed or an entire class of material may be out of stock (e.g., thin BoPET sheets commonly used in face shields). It is therefore important to consider equipment and supply constraints and assess alternatives for raw materials, fabricators, or suppliers early in the design process. In a crisis, multiple fabricators who might normally compete may be willing to collaborate to increase production volume and provide complementary capabilities.

Converting Needs Assessment Results to Technical Specifications

Requirements identified via needs assessment must be converted into functional and technical specifications that guide design. For example, if an end-user needs a face shield that is adaptable to different individuals, the functional requirement is for a product that fully covers faces of different sizes and has adjustable straps and attachment hardware. The technical specification might then be a shield of length of 22.5 to 30 cm and headband circumference of 50–60 cm. Analogously, an end-user requirement for reusability triggers a requirement for input from infection control experts and results in a functional specification for sterilizable designs and materials. The technical specification would then call for materials compatible with sanitizing wipes or hydrogen peroxide sterilization and an absence of crevices that can trap contaminants.

STAGE 2: SOLUTION DEFINITION, VERIFICATION AND VALIDATION

Solution definition, verification and validation is an iterative process in which functional and technical specifications defined in Stage 1 are transformed into actual designs. Designs are

then compared to specifications to verify that all requirements have been met. General considerations applicable to all medical products must also be included, such as the biocompatibility of materials in contact with humans. During a crisis, high demand for some types of equipment and raw materials may impose additional requirements on device components and processes.

Rapid Manufacturing Techniques

Rapid manufacturing methods have a well-established role in facilitating rapid cycles of design and testing to reduce uncertainty and shorten production timelines (28). The use of rapid manufacturing methods is increasing in healthcare, facilitated in part by a series of FDA workshops (29). In the case of prosthetics (30), orthotics (31), tools for surgical planning (32), and dental and surgical equipment (31, 33, 34), additive manufacturing has opened new possibilities for designing products with complex geometries and allowed manufacturers to move away from providing only standardized products in a few sizes toward custom, patient-matched products. During the COVID-19 pandemic, additive manufacturing has been widely used to make face shields (31, 35, 36), nasopharyngeal swabs (37–39), face mask brackets (22), components for portable-air purifying respirators (PAPRs) (35), and ventilator splitters (40). In response to this activity, the FDA has released relevant guidance (3). **Table 1** describes key manufacturing methods that are suitable for the production of substitutes for medical devices currently in short supply. The methods described use machinery that is available in both commercial (industrial) and consumer (maker) grades; however, industrial machinery is more precise, faster, and can typically process larger size products or materials. In many cases, designs prototyped on consumer-grade equipment can be successfully transitioned to higher-capability industrial machines.

Rapid manufacturing is used infrequently for PPE under normal circumstances primarily because the approach is typically more expensive (per unit) than conventional, large-scale production using methods such as injection molding. The limited production volumes of rapid manufacturing also pose a significant challenge to meeting the large demand created by the pandemic. Thus, they are most effective as a means for prototyping and short-term production or highly-distributed production, while conventional large-scale production methods ramp up.

Sterilization and Reuse

To address acute shortages in devices that are traditionally single-use, such as respirators or face shields, the CDC issued guidance allowing for extended-use, reprocessing, and reuse of PPE (42). According to FDA regulations, hospitals and third-party reprocessors are considered "manufacturers" of the reprocessed devices and must comply with the same regulatory requirements as the original equipment manufacturers (43). Given the difficulty of compliance, the FDA and CDC guidelines have been relaxed for single-use devices during the pandemic. However, designers must take the necessary steps to ensure the compatibility of their devices with anticipated sterilization, decontamination, and cleaning procedures. This usually requires consultation with hospitals' infection control experts as well as empirical testing. Additionally, fabricators should take the appropriate steps to ensure initial disinfection and sterilization of their products prior to delivery to end users. For devices such as faceshields, wiping down newly-fabricated units with approved disinfectants is most likely to be the appropriate procedure. The EPA provides a detailed record of products meeting criteria for use against SARS-CoV-2, along with corresponding directions for use (44).

In a pandemic, many products that are normally disposable end up being reused because they are in short supply. It should therefore be assumed that face shields, masks frames, and other items will be sterilized or decontaminated if at all possible. Sterilization methods of autoclaving, applying bleach-containing solutions, and alcohol-based wipes are widely available in healthcare and may be suitable for sterilization of products such as face shields. However, such methods often degrade key components, damage labels and safety warnings, and are not compatible with products such as N95 masks. Short wavelength ultraviolet (UV) light (45), vaporized or ionized hydrogen peroxide (46–48), and moist heat are more generally applicable but less widely available alternatives currently being developed for sterilization of masks and similar devices (45). Early into the design process it is important to determine which sterilization methods are available for testing and possible use with deployed products and then ensure their compatibility with a proposed product. The use environment should also be taken into consideration; for example, access to the sterilization equipment or decontamination solutions may be limited during a crisis and the process of getting products to and from centralized sterilization facilities must be considered. In many cases, the fundamental desirability of product reuse runs up against practical challenges with logistics. This is particularly true in the case of PPE that needs to be sterilized or decontaminated and then returned to the original users. We have found that many healthcare providers have been unable to put the necessary tracking procedures in place to make "return to original user" possible.

Biocompatibility

Biocompatibility is defined by the FDA as the "ability of a material to perform with an appropriate host response in a specific situation" (49) where response refers to a host immune or inflammatory reaction to the material. Evaluation of biocompatibility is one part of the FDA's overall determination of safety and effectiveness for new or modified devices that come into direct or indirect contact with the human body (50). In the US, two documents outline standard biocompatibility testing: the International Standard ISO 10993-1 (51) and the guidance related to ISO 10993-1 (49). A separate biocompatibility standard exists specifically for respiratory devices: ISO 18562-1:2017 (52). Among other factors, a biocompatibility assessment focuses on: (1) material chemistry and any changes caused by the manufacturing process, (2) material physical properties, (3) nature of the body contact (direct or indirect), (4) contact

TABLE 1 | Prototyping and manufacturing methods applicable to production of five medical devices, based on Open Source COVID-19 Medical Supply Guide (41).

Categories	Set-up cost and time	Prod. cost and time	Specific methods	Face shields	Nasapharyngeal swabs	Surgical face masks	N95 respirators	PAPRs	Ventilator splitters
Rapid Prototyping	Low	High	3D Printing (FDM)	Yes	Yes	No	Yes	Yes	Yes
			3D Printing (SLA)	Yes	Yes	No	Yes	Yes	Yes
			Machining	Yes	No	Yes	Yes	Yes	TBD
			Laser Cutting	Yes	No	Yes	No	Yes	No
High Volume Production Processes	High	Low	Die Cutting	Yes	No	Yes	Yes	Yes	No
			Injection Molding	Yes	Yes	No	Yes	TBD	Yes
			Compression Molding	Yes	Yes	No	No	TBD	TBD
			Thermoforming	Yes	Yes	No	No	TBD	TBD
Fabrication and Assembly	Variable	Variable	Sewing	No	No	Yes	Yes	Yes	No
			Gluing and Bonding	Yes	Yes	Yes	Yes	Yes	TBD
			Fastening	No	No	No	No	Yes	No
			Electronics	No	No	No	No	Yes	No
			Assembly	Yes	No	Yes	Yes	TBD	Yes

***FDM**: Fused Deposition Modeling. An additive manufacturing (3D printing) process that can be implemented using low-cost consumer-grade equipment. A printer heats up a plastic in filament form and uses a print head to create a 3D structure from the bottom up.*
***SLA:** Stereolithography. Another additive manufacturing process using low cost consumer-grade equipment in which a vat of liquid resin is hardened with a laser to create desired shapes.*
***Machining**: Conventional (subtractive) manufacturing in which raw material is cut into a desired shape by a controlled material-removal process such as milling and routing. Small, low-volume, computer numerical control (CNC) machines are increasingly widespread.*
***Laser cutting**: A technology in which a high-powered laser is used to cut sheet materials to size; consumer-grade laser cutters for plastics are increasingly inexpensive.*
***Die cutting**: A technology in which metal knives (dies) with sharp edges are pressed into sheet materials to cut them to size and shape. Production time per piece is significantly faster than laser cutting but requires more expensive and longer setup due to the custom-fabrication of the die.*
***Waterjet cutting.** A technology in which a narrow jet of high pressure water and abrasive is used to cut sheets of material; desk-top waterjet machinery is now available.*
***Injection Molding**: A high-volume manufacturing technology that produces parts by injecting molten material (polymers, glass, metal) into a mold. The upfront cost and setup time of injection molding is high but unit cost and production time are low.*
***Compression molding**: A system by which heat and pressure are applied to a material in order to shape it.*
***Thermoforming**: A technique by which plastic is heated, formed to a specific shape in a mold, and trimmed to create a desired shape.*

duration, and (5) prior history of safe use [as defined in ISO 10993-1 (49)].

Biocompatibility requirements for materials are application-specific and vary greatly based on the part of the body in contact with the device and the duration of contact; sustained internal contact is substantially more problematic than brief or external contact with the skin. Thus, approval of a material in one application does not constitute approval for another application. In addition to material considerations, the method by which a material is handled or processed during manufacturing may influence its biocompatibility. During biocompatibility evaluation, testing is performed on the "final finished form" of the device, which includes all manufacturing processes including packaging and sterilization. Rapid prototyping can be advantageous in conducting biocompatibility testing early in a product development cycle.

From a practical standpoint, most devices being subjected to rapid fabrication for pandemic response are for external use only and primarily contact the skin (or hair). In this setting, it is reasonable to use materials previously shown to be safe, such as silicones, parylene coatings, and many common fabrics. Particular attention should be paid to foam, elastic materials, and adhesives with respect to skin contact and latex should always be avoided; Monprene® (PR-23040) is an FDA-approved alternative elastic material that is widely used in phlebotomy and is readily available.

If materials previously documented to be biocompatible in a particular setting are unavailable or functionally unsuitable, then biological endpoints ranging from cytotoxicity and sensitization to material degradation and carcinogenicity must be considered. Attachment A of FDA's guidance related to ISO 10993-1 provides a list of the recommended biological endpoints to consider for the development of biocompatibility evaluation as well as the rationale for these endpoints (49). The flow chart of **Figure 2** illustrates how one might evaluate biocompatibility to determine if a newly-developed device requires supplementary testing. During the current crisis, safeguards have been relaxed by regulatory authorities to enable more rapid response, so long as a specific medical device's product code is explicitly mentioned in an FDA guidance or enforcement policy. These guidance documents are freely available on the FDA website (53).

Design Verification and Validation (V&V)

Design verification is an iterative and empirical process in which objective evidence is sought to assess whether a product satisfies specifications. Design validation is a summative exercise that assesses the integrity of the final product and ensures that it meets user needs in a real or simulated use environment via testing, measurement, and observation of user interaction [commonly to ISO 13485:2016 (54)]. It should be noted that a product's packaging, labeling, and instructions are considered to be essential parts of the product. Even in a crisis, it is important to provide inserts and labels with products describing their intended use, composition, and regulatory compliance (e.g., reference to an EUA).

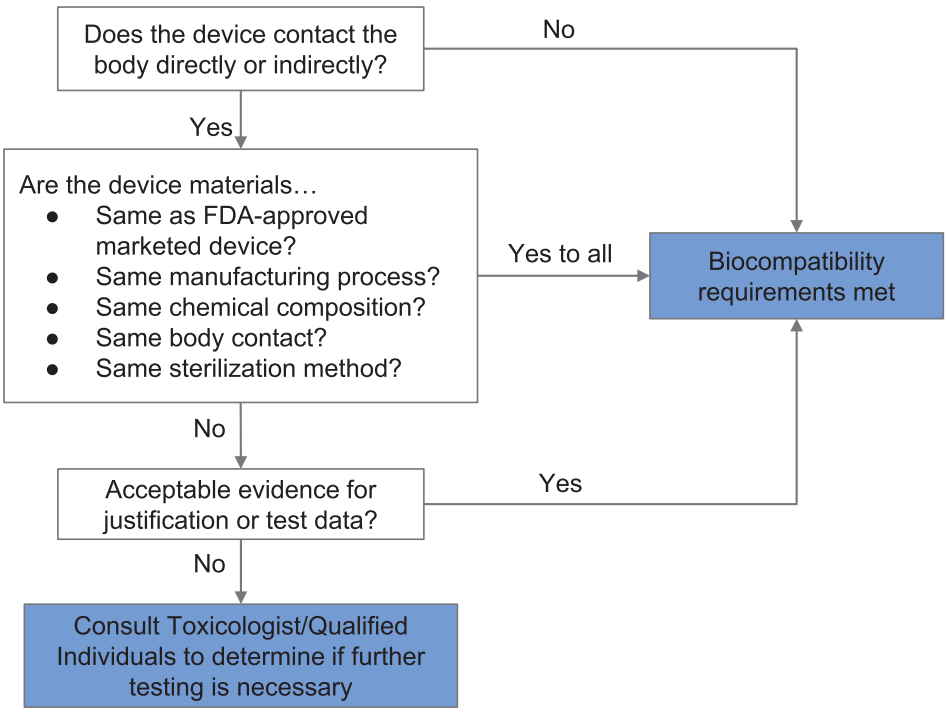


FIGURE 2 | Systematic approach to a biological evaluation of medical devices as part of a risk management process. The process is simplified from FDA Guidance document for ISO 10993-1 (49).

Scaling Up

While rapid manufacturing techniques such as 3D printing and laser cutting are efficient approaches for fabricating prototypes, low throughput and high unit costs do not make them feasible for large-scale production. In a traditional waterfall process, once a design converges on a final set of specifications, design transfer takes place. During the design transfer, the prototype design is adapted to the demands of large-scale manufacturing methods such as injection molding and die cutting. These methods are high-throughput and inexpensive per piece but are associated with significant up-front cost and set-up time (up to several months). Moreover, highly specific expertise is generally required to make design dies and molds compatible with a specific type of equipment. In a crisis, in which time is usually limited and production costs are less of a concern, parallel production using the prototyping facilities of many manufacturers, colleges, and makers can be a good way to meet demand. For example, in the case of the Panfab/BWH face shield (**Box 1**), 3,000 face shields were fabricated in a few weeks using 3D printing and laser cutting at multiple sites (36) and the Czech 3D producer PRUSA has described a highly parallel face shield printing process using inexpensive machines (35).

Conversations with suppliers and manufacturers will help guide prototyping and design processes and prepare for design transfer to a manufacturer, if relevant. A responsible entity may be required to register with the FDA as a manufacturer and list the product(s) they are distributing or selling, or the services they are providing to other manufacturers, depending on the product code, respective risk classification, and regulatory compliance pathway (55). Meeting these requirements will be especially important once the public health emergency has ended. Additional stakeholders such as payers and post market surveillance organizations may also come into play (**Supplementary Figure 1**). It is currently unknown whether traditional manufacturers will adopt some of the innovative designs developed by non-traditional suppliers during the current pandemic and shepherd them through the regulatory process. We hope that this is the case, but much depends on the creation of better market incentives (see below).

STAGE 3: REGULATORY APPROVAL AND IMPLEMENTATION

On January 31, 2020, the US HHS Secretary Alex Azar declared a public health emergency involving COVID-19, noting that the circumstances justified emergency use of *in vitro* diagnostics and other medical devices that aid in the detection or diagnosis of COVID-19 (56). Pursuant to this declaration, the FDA has issued Emergency Use Authorizations (EUAs) for a number of medical devices. As mentioned above, EUAs allow certain non-FDA approved medical products to be used in the absence of adequate FDA-approved alternatives (57). While some EUAs are manufacturer-specific, others are broader in scope. For example, EUAs for face shields and respirators waive certain FDA requirements for all prospective manufacturers, authorize

BOX 1 | Case study: nontraditional design and fabrication of face shields during a health care emergency.

A Case Study on Fabricating Face Shields in the Northeast U.S.

In response to the COVID-19 pandemic, which grew rapidly in Massachusetts during March 2020, multiple teams formed to address rapidly growing shortages in medical supplies. Some of these teams were established by institutional mandate and others arose spontaneously through the efforts of engaged individuals. An example of the latter is the Greater Boston Pandemic Fabrication Team [PanFab (20)], a student-faculty initiative organized by the Harvard-MIT Center for Regulatory Sciences. It consists of a group of volunteers with expertise in engineering, biomedicine, manufacturing, and regulatory review working closely with the physicians at Boston-area hospitals, the local maker community, and manufacturing experts from local companies contributing outside of normal work hours. A physician assigned to a hospital incident command was a particularly important member of the PanFab team because she could provide timely and accurate information on current and emerging needs. PanFab has been effective in designing and rapidly fabricating face shields, mask frames, PAPRs, and other types of PPE (36) and its activities are representative of local design in response to a healthcare crisis.

Stage 1: *Problem Definition and Needs Assessment*.

Defining the problem

Faceshields are a critical component of PPE per the CDC (77) and are used in conjunction with surgical masks or N95 FFRs to protect the face and neck, particularly mucous membranes in the eyes, nose, and mouth, from splatter by contaminated bodily fluids (78). However, during the COVID-19 pandemic, a severe shortage of face shields developed. Face shields are composed of a clear shield (commonly made from BoPET, PETG, acetate, or polycarbonate) and a headband frame that is in contact with a user's forehead and commonly made from a lightweight plastic or foam (78). In a US health care setting, face shields are traditionally single use devices, but because of shortages they are being worn for the full duration of a shift (6 to 12 h) and then sterilized for reuse.

Needs assessment results and associated technical specifications:

From interviews with healthcare providers, hospital administrators, and infection control specialists, the following needs were identified:

1. The shield must protect mucosal membranes from splashes of bodily fluid by extending from the forehead to the points of both ears and down the neck.
2. The shield should not fog or otherwise obscure the user's view, even during strenuous activity that may produce perspiration (78).
3. The shield should fit a range of facial lengths and heights while not interfering with range of head and neck motion.
   a. Associated technical specification: facial lengths 23-30 cm and range of motion 180 degrees in each direction.
4. The shield should remain firmly in place and remain comfortable when the user moves their head up, down, and laterally at varying speeds.
5. Face shields are typically worn for the duration of a shift (up to 12 h). Thus, the shield cannot be so heavy as to cause discomfort when worn for a shift.
6. Attachment mechanisms should remain firm while preventing skin sensitization, irritation, or imprints on the skin. The attachment must be adjustable.
   a. The FDA EUA allows specific materials to be re-used without thorough re-testing; these are materials of choice for the design.
7. Device cost in terms of materials should be <$5 per shield; design and fabrication time are to be donated.
8. Hospital demand requires a production rate of 2,000 shields per month.
9. The face shield should require minimal assembly, both for ease of use and to limit the presence of hard to reach surfaces, which could interfere with the decontamination process.
10. To be reusable, the face shield should be compatible with the hospital's commonly used sterilization techniques, which include ionized hydrogen peroxide, germicidal disposable wipes, or 70% isopropanol wipes.

Stage 2: *Solution Definition and Validation*

Prototyping and Manufacturability:

Given the low-risk nature of the device, various 3D printing approaches were considered acceptable. We began with the open-source Prusa design (35) and iterated it based on feedback from healthcare providers at Brigham and Women's Hospital (BWH) (**Figure 3**). The new design added a forehead fin with a drip guard that protected the otherwise-exposed forehead from body fluid exposure. This requirement was not identified during initial assessment but became obvious once prototypes were in the hands of emergency room users. The prototype shield was also too narrow and short and did not provide sufficient splash protection for the neck and sides of the face for all users; the length and width of the shield was therefore increased.

Sterilization and Reprocessing:

Given that safe sterilization of 3D printed face shields had not been extensively studied, we followed CDC's guidance (79) relating to sterilization and reprocessing compatibility for goggles, a related product. We checked for changes in material properties and visor transparency following several days of use and following regular cleaning with sanitizing wipes (Germicidal Disposable Wipes). We also ensured that the face shield could withstand ionized hydrogen peroxide sterilization (iHP; TOMI SteraMist), resulting in effective killing of test bacterial spores as measured by standard biological indicators (80).

Biocompatibility

A number of materials were assessed during fabrication of the face shield based on (i) resource availability, (ii) previous uses in marketed and FDA-approved medical devices, and (iii) compatibility with common sterilization techniques. In the case of the face shield, only the 3D printed headband, the foam pad, and the Velcro strap of the face shield were in contact with user's skin or hair. Applying the workflow shown in **Figure 2,** the requirement for safe interaction with intact skin was met based on limited duration body contact (< 24 h) and prior material biocompatibility information obtained from material Material Safety Data Sheets (MSDSs), reported experience from manufacturers, and published literature. PLA was selected as the material for 3D printing the headband because of its wide availability and because it is known to be safe in contact with the skin. The Velcro strap, which is made from polyethylene and nylon, is also known to be safe for skin contact. Similarly, we established that EVA foam was safe for skin contact. We selected closed-cell EVA over its open-cell counterpart—which is absorbent—and evaluated its compatibility with common sterilization and reprocessing techniques [70% isopropanol wipes, ionized hydrogen peroxide (80)].

Stage 3: ***Regulatory approval and Implementation***

Regulatory approval:

Face shields are an FDA regulated Class I 510(k) exempt device (78) making them appropriate to attempt to fabricate locally. Additional information on the FDA regulation of face shield products during the COVID-19 pandemic can be found in **Table 2**.

Clinical Testing:

To assess face shield usability and safety, a cohort of 97 physicians, physician assistants, emergency department technicians, environmental service staff, and other individuals with patient-facing roles were recruited to an IRB-approved study from the Emergency Department at BWH. Users wore the face shield during their shifts and completed a questionnaire on baseline experiences and attitudes. A majority of users indicated that they had a better experience with the PanFab face shield

Continued

BOX 1 | Continued

as compared to the hospital standard-issue, disposable model. This data was then presented to the appropriate stakeholders at BWH, and the product received approval for use as part of a clinical workflow.

Production and Implementation:

Following validation of the final design through an IRB-approved study at BWH over 3,000 face shields have been manufactured and deployed to meet local demand. These face shields were produced via distributed rapid manufacturing in collaboration with a local makerspace community (BoroBot), Salesforce, a small-scale prototyping company (SunPe Prototype), large-scale manufacturers of non-medical supplies (iRobot and Velcro), in addition to academic partners (Harvard Graduate School of Design, the Shin Laboratory at the Brigham and Women's Hospital's Stepping Strong Foundation for Trauma Innovation, and the Wentworth Institute of Technology). All design files were made openly available on the PanFab website and on the NIH 3DPrint exchange platform to allow other manufacturers and localities to produce the face shield.

Timelines and key dependencies

Once the need for an alternative source of face shields was established via consultation with hospital incident command, four individuals from the PanFAB group became engaged with the project. Problem definition and preparation of documents for the IRB occurred concurrently and required approximately two weeks. Following this, prototyping and soliciting user feedback on modified designs took another two weeks. User testing in a clinical setting was dependent on having IRB approval and a final design. Collecting and consolidating results took another week and was overseen by three physician-scientists. During this time, production materials were procured and four production sites were lined up. Production began when analysis of testing data was complete and first delivery of face shields occurred six weeks after the project started; production continued for six weeks at four sites, yielding a total of ∼3,000 face shields.

FIGURE 3 | Face shields fabricated during the COVID-19 pandemic. **(A)** Image of Prusa RC2 design and **(B)** final PanFab face shield prototype. See text for references and details.

the modification of approved products, and allow for extended use, widespread production, and distribution of devices so long as manufacturers adhere to requirements outlined in the EUA notice (58, 59). EUAs expire upon resolution of the public health emergency. The agency has also provided extensive and frequently-updated guidance documents for manufacturers seeking to produce rapid diagnostics, personal protective equipment (PPE), and other devices for front-line use in responding to COVID-19. EUAs have also extended into the realm of sterilizing PPE for re-use.

A lack of strict regulatory oversight does not absolve designers and fabricators from doing their best to ensure that products are safe and functional and do not put healthcare providers or patients at risk. Additionally, disclosure of risk to the end-user is important: the FDA EUAs and enforcement policies include specific requirements for labeling, information for use, and use environments to ensure that products do not create undue risk.

Overview of Medical Device Regulatory Review and Implementation

In the US, medical devices are typically made available via the FDA's 510(k) premarket notification process. Depending on the degree of risk associated with the use of a medical device it is classified by the FDA as either Class I (low risk), Class II (moderate risk), or Class III (high-risk) (60, 61). The 510(k) submission process requires that a manufacturer of a device new to the market demonstrate "substantial equivalence" to one or more legally marketed devices, thereby avoiding a requirement for extensive clinical testing. Substantial equivalence does not require that a device be identical, but instead mandates that it be as safe and effective as an existing marketed device (62). Some Class I products are exempt from the 510 k premarket submission process based on the evaluation that these products pose a particularly low risk. This includes devices such as face shields, most forms of PPE, and nasopharyngeal swabs. Class II devices include more complex life-critical products such as ventilators and Class III devices traditionally encompass products implanted into a person, such as pacemakers, defibrillators, or prosthetics. Almost all devices currently being supplied by non-traditional fabricators and maker communities fall into Class I, and the majority are Class I exempt. However, ventilator splitters, which have been much discussed as a way of increasing ventilator capacity (63), may be Class II devices (64) and considerable controversy has attended their development (65).

The premarket submission process only clears a manufacturer to market a product; additional requirements must be met prior to manufacturing, selling, or distributing it. Quality controls in manufacturing, referred to as good manufacturing practices (GMP), include a requirement of product traceability in case of manufacturing flaws or product recalls. Notably, no Class II devices, and only a small number of Class I 510 k exempt devices, are also exempt from GMP regulations (66). Meeting GMP standards constitutes a substantial barrier for formal certification of devices fabricated through local, nontraditional manufacturing practices.

After implementation, unique device identification, an important part of the FDA's post-market surveillance process, ensures that manufacturers vigilantly and proactively monitor the use of their product(s) for adverse events and patient injuries. The post-market surveillance process can also act as means to collect user feedback, which can lead to product improvements or innovative alternatives to existing designs. In the current crisis, US regulators have put in place emergency authorizations that waive some of the normal GMP and tracing requirements in favor of less stringent discretionary enforcement. In many cases, this forms the regulatory foundation for repurposing products from the non-medical supply chain and for hospitals to engage with informal networks of maker communities and similar organizations (67). The underlying logic is that a more permissive stance on low-risk Class I products better balances the hazards associated with a lack of supply against the hazards of using new designs and non-traditional manufacturing processes.

The use of research protocols under the purview of Institutional Review Boards (IRBs, similar to ethics committees outside of the US) provides one well-established route for testing new products in a clinical setting. Risk assessment, informed consent, and certification of human subjects training are standard components on an IRB submission, which should be familiar to many investigators in academic medical institutions. Collaborations between such individuals and maker communities or private fabrications are therefore encouraged. While, private "pre-certification" laboratories that perform testing of devices to prevailing regulatory standards in the US, Europe, and Japan can provide a substantial degree of confidence in a non-traditional product, they may still require GMP certification for use under the FDA guidelines.

Product Safety Validation for PPE

Product safety validation for PPE and similar low-risk devices is commonly performed by certified testing laboratories, many of which are commercial and provide fee-based testing for compliance with specific regulatory standards (e.g., NIOSH standards for PPE). Under normal circumstances, regulations require that quality management systems (QMS, e.g., to ISO 9001) be in place. A QMS specifies how design and development inputs and outputs should be documented, how records should be maintained on the skills, the experience and qualifications of key personnel, and how calibration records should be established and monitored. Statistical Process Control (SPC) is also widely used throughout medical device manufacturing (and manufacturing in general) to ensure product consistency and quality. In a crisis-responsive setting, QMS and SPC are likely to be infeasible, which makes formal certification by NIOSH or FDA standards impossible. In such instances, makers must do their best to maintain quality and test either independently or via the use of other local resources such as academic laboratories. Consortia such as the Mass General Brigham Center for COVID Innovation (68), N95Decon (69), and M-ERT (17) may be able to provide guidance in specific situations.

Use of Research Protocols

An effective and well-established way to test non-traditional medical products in a healthcare setting is to use research protocols overseen by an Institutional Review Board (ethical review board; IRB); these are part of the normal operation of virtually all teaching hospitals. The use of a research protocol makes clear to participants, via the process of informed consent, that a product is being tested and that it has not necessarily been through the usual regulatory review. IRB approval is also almost always required for conducting user surveys and receiving direct feedback from end-users. Devices such as face shields (described in **Table 2**) are FDA regulated Class I products with the lowest risk to human health. They are therefore suitable for testing in an IRB-supervised research protocol under even the suboptimal conditions of a healthcare emergency. We have found that the use of research protocols for product testing increases buy-in from healthcare leadership, in part because it is a familiar process. In the case of a non-traditional face shield as described in the case study below (**Box 1**), it was possible to perform IRB protocol review and product specification, validation, and testing in a clinical setting in a period of roughly three weeks (36). IRB review was performed in less than one week as a result of specific steps put in place by area hospitals to accelerate the introduction of COVID-19 related protocols. More commonly, IRB review for non-invasive, low-risk protocols takes several months.

GENERAL RECOMMENDATIONS

A crisis-responsive design framework aims to couple user needs, rapid manufacturing technologies, and local fabrication to fulfill unmet demand for simple medical devices in times of crisis. Below we provide three sets of recommendations intended to serve as a checklist for (1) designers, maker communities and fabricators, (2) healthcare providers and administrators, and (3) regulatory bodies overseeing pandemic response. Note that detailed regulatory documents from NIOSH, ANSI, ISO, and DIN typically cost several hundred USD but many are being made freely available for the duration of the current healthcare emergency.

Recommendations for Designers, Maker Communities, and Fabricators

1. **Conduct a thorough needs assessment** to understand the key features and requirements for the proposed product and its labeling, assess likely demand from different types of users, establish requirements for sterilization or decontamination,

TABLE 2 | Summary of regulatory and testing standard for each device.

Medical supply	Device FDA product code	Traditional regulatory process	Revised process during COVID-19 pandemic	Standards for testing (NIOSH/ANSI)
Face shield	LYU	FDA regulated Class I 510(k) exempt	FDA Emergency Use Authorization (EUA) (58) FDA Enforcement Policy	ANSI/ASSE Z88.2-2015
Nasopharyngeal (NP) swabs	KXF, KXG	FDA regulated Class I 510(k) exempt	No EUA	Not applicable. See text for description of materials, PCR compatibility, transport media compatibility, mechanical performance, and length considerations.
Surgical face masks	FXX	FDA regulated Class II 510(k) (70)	FDA Enforcement Policy (71)	6-254 ASTM F2100-11; 6-335 ASTM F2101-14; 6-406 ASTM F1862; 6-425 ASTM F2100-19; 6-427 ASTM F2101-19
N95 respirator	ONT, ORW, NZJ	FDA regulated Class II 510(k) clearance	FDA Emergency Use Authorization (EUA) (59)	NIOSH (72, 73)
PAPR	N/A	NIOSH regulated, FDA approved (74)	FDA Emergency Use Authorization (EUA) (59) as a subset of filtering facepiece respirators (FFRs) (75)	NIOSH (76)

**Only true for the duration of the crisis. These EUA guidelines are up to date as of the time of submission.*

secure appropriate materials and mitigate any supply chain issues (e.g., by using alternative raw materials), and establish tolerances for different styles and types of products. Perform this analysis by speaking to a broad range of stakeholders early in the process and make sure to consider diverse body types and clinical roles. Secure examples of existing products and assess their strengths and weaknesses. Download and review relevant NIOSH/ANSI or regulatory documents from other agencies.

2. **Solicit end-user, expert, and clinical feedback early and often** throughout the design process to ensure the product satisfies anticipated and unanticipated needs. Frequent assessment increases the likelihood of end-user buy-in and enables rapid adoption of new solutions in response to changing demands. Consulting with senior medical staff and division heads can be helpful in this setting. Establish a testing process and determine whether an IRB-approved clinical protocol may be required; it is necessary if a formal survey is to be conducted.
3. **Conduct a rigorous assessment of biocompatibility, sterilization, and risk**. For PPE-type devices, research issues related to biocompatibility, manufacturing, intended duration/frequency of use, and anatomic location. Use materials previously established to be safe for the intended use if at all possible. This assessment is usually based on the scientific literature, medical device standards, or data on devices previously reviewed by the FDA. Be aware of likely sterilization requirements and feasibility since not every method of sterilization is compatible with every device. Review ISO 14971 (55), ISO 10993−1 (51) and FDA's associated guidance (49), and ISO 18562-1 (52) for additional information. Determine if proposed sterilization methods will in fact be available for testing and use when products are deployed.
4. **Search design repositories for suitable designs** that can be used as-is or modified to meet local requirements and fabrication capabilities. Many online forums have emerged to assist in the dissemination of best practices. Publicly funded designs should be available under nonrestrictive licenses from resources such as the National Institutes of Health 3D Print Exchange (4).
5. **Consider manufacturing methods**, such as 3D printing and laser cutting, which enable rapid prototyping, and low-volume manufacturing, while keeping in mind a possible transition to other approaches (e.g., injection molding) for large scale manufacturing. Consider which types of equipment will be available for the duration of the project and whether local shops can be contacted for access to higher-end equipment.
6. **Avoid action bias**, which results in premature action and over-rapid development of potentially suboptimal or undesired solutions. Even in a crisis, it is important to proceed deliberately to ensure that products are usable, safe, and durable.
7. **Develop a written process for device performance validation** through measurement of fit-for-purpose criteria and alignment with end-user needs and regulatory standards. Check with target users regarding the testing requirements and IRB-based testing capabilities. Consider life cycle issues including whether products will be withdrawn from service at the end of a medical emergency.
8. **Provide documentation.** Make sure to include accurate and complete labels and product information as inserts with the final products; consider complementing this with QR codes and online resources to provide users with the most up to date information. Consider making new designs and any improvements on existing designs publically available.
9. **Perform multiple activities in parallel.** Given the need to provide products as rapidly as possible in a pandemic,

processes that are traditionally performed in a sequential matter should be parallelized to the greatest extent possible. Needs assessment, literature review, and IRB submission can occur in parallel. Collecting information on biocompatibility, sterilization, and risk can be started based on some assumptions about materials. Once a prototype is available, input from potential users can be solicited. Note however, than any data collection via surveys or field studies requires IRB approval so securing this approval is a critical step and potential bottleneck. At this time potential fabricators and sources of material can be lined up. However, before proceeding to actual manufacturing, test data must be analyzed and a final validation assessment performed. This is a second critical step and the information should be reviewed by the design, clinical, and fabrication team members. Labels and product documentation must be completed before devices can be delivered to the end user and all products must be decontaminated or sterilized; this is the third critical step in the path to providing useful products into a healthcare setting.

Recommendations for Healthcare Providers and Administrators

1. **Assemble a diverse crisis response team.** Healthcare providers should add one or more individuals with experience in manufacturing, product design, or with maker communities to incident command and crises response teams. These individuals should be charged with outreach to non-traditional suppliers and emerging resources (e.g., the 3D Print Exchange) prior to a crises. In many cases it will be possible to identify individuals with valuable expertise in clinical teams. The engineering organizations in many hospitals can also be helpful. A diversity of perspectives and providing key personnel with the time and resources to understand regulatory documents associated with non-traditional fabrication is essential for sourcing non-traditional medical supplies.
2. **Develop more robust supply chains.** Health care providers should consider the resilience of supply chains in addition to cost. The likelihood that all suppliers will simultaneously be unable to supply key products must be considered, since the number of original equipment manufacturers is often much smaller than the large number of branded products would suggest.
3. **Use IRB Review Process to test non-traditional medical supplies.** Using expedited IRB review is an effective way to test non-traditional medical products in a healthcare setting. However, it is necessary that IRBs work efficiently under emergency circumstances when low-risk Class I and exempt products are being tested.

Recommendations for Regulatory Bodies

1. **Maintain current regulatory standards under normal conditions.** We do not propose that regulatory standards for PPE and simple medical products be relaxed, but consideration must be given to foster innovation. Federal reports have repeatedly identified deficiencies in existing PPE but few if any improvements have been made in designs or supply chains. This situation should be rectified before the next emergency and may require government funding.
2. **Develop policies that facilitate fabrication of high quality non-traditional products** for use specifically in medical emergencies. These policies might reasonably include some of features of a crisis-responsive framework outlined above as well as pre-tested public domain designs and prescriptive fabrication approaches.
3. **Increase the transparency of Emergency Use Authorizations to improve submission and approval.** The minimum standards needed for submission of a request for an FDA EUA for new technologies must be improved. As it currently stands, it is difficult to ascertain key features of products in the non-traditional supply chain from commercial manufacturers. For example, the FDA authorized distribution of foreign manufactured N95-style masks and did not require that companies provide basic operational data including name and place of business, proprietary or brand name, model number, marketing authorization, and a copy of the product labeling (81). At the very least, data submitted to the FDA in support of an EUA should be made public to the greatest extent possible.
4. **Provide an accelerated pathway to transition products authorized under an FDA EUA to either traditional or emergency-only approval.** It is highly desirable that innovative designs and approaches developed during the pandemic be further developed, tested, and integrated into normal supply chains so that they are ready for future emergencies. In the US such a task falls outside of the remit of regulators such as the FDA and NIOSH but might be tackled by the Dept. of Health and Human Services or even the Department of Defense. Since PPE for pandemic response is intended to be a public good, international cooperation would be highly desirable.

CONCLUSIONS AND FUTURE PROSPECTS

The COVID-19 pandemic has made clear the fragility of medical supply chains whose breakdown has resulted in rapid and severe shortages for many essential medical supplies. The communities of designers, fabricators, and healthcare professionals who have come together to supply locally made substitutes using rapid manufacturing methods have revealed a hitherto untapped capacity to make the provision of medical supplies more resilient. The great strength of community efforts is that they avoid extended product development cycles and bypass international lowest-cost production in favor of rapid innovation and efficient execution. However, even in a crisis, the risks of using alternative, locally-manufactured medical devices must be carefully evaluated and mitigated. Hastily designed products made without appropriate stakeholder input are unlikely to be used clinically, representing a loss of time and resources by well-intentioned makers. If poorly designed devices make their way into clinical use, they can pose a substantial hazard. This perspective provides a crisis-responsive framework for

medical product design that is intended to avoid these risks. The framework is informed by traditional FDA guidance on product approval, implementation, and validation, but adapt that guidance to accommodate constraints in time, material resources, and human capital.

What does the current state of the COVID-19 response portend for the future? From the perspective of nontraditional fabricators of low-risk medical devices and PPE, a transition from purely local efforts to national and international collaborations is already underway. These initiatives include several efforts focused on PPE including Get Us PPE (82), Covid19 Masks (83), Mask Match (84), Project N95 (85), and PPELink (86) and larger efforts such as Open Source Medical Supplies (OSMS) (87), America Makes (19), and the NIH 3-D Print Exchange (4). State governments and companies such as Gillette (a P&G Company) (88), and Lovepop (89) are coming together through initiatives like the Manufacturing Emergency Response Team (M-ERT) (17) to design, develop, manufacture, and donate or sell thousands of devices. These consortia facilitate stakeholder interactions and device testing, operations that are hard to perform on individual bases. Continuing these efforts will build greater resilience for future pandemics.

A key unanswered question is what will happen to non-traditional medical devices and innovative designs when the COVID-19 pandemic recedes. It is possible that even the best innovations will be abandoned in favor of a return to standardized lowest-cost alternatives. The design deficiencies of these existing products and the fragility of supply chains have long been known and analyzed in a comprehensive series of government reports spanning two decades, several of which were undertaken in response to MERS, SARS, influenza, and other zoonotic transfers. A wide variety of potentially effective solutions were proposed by government and private entities (67) but few were actually implemented. A key problem is the high pressure on costs, which for commodity products inevitably results in the use of lowest cost suppliers and products that barely meet operational requirements. The current pandemic has revealed the weaknesses of this approach, but we have witnessed little discussion about the necessity of paying more to maintain supply chain resilience.

To avoid forgetting the lessons of COVID-19 pandemic, it is essential that innovative designs and approaches be further developed, tested, and integrated into normal regulatory chains so that they can be used in standard products or readied for future emergencies. Changes in regulatory policies are needed to balance the rigidity of the pre-crisis approach with counterfeiting (81) unintentionally enabled by EUAs. We envision the development of open-source repositories, such as NIH 3D Print Exchange, to facilitate independent testing of products that meet key NIOSH, FDA and other requirements for functionality and safety. These should be accompanied by simple prescriptive approaches and best practices geared to the capabilities of small-scale fabrications. Such guidance could mimic the practice in building codes [e.g., those from the International Code Council (76)] of providing both a limited number of pre-engineered solutions that are ready for application in the field while also allowing for a wider range of solutions when engineering resources are available. In a crisis, guidance of this type would take the place of QMS and SPC and improve products even when the requirements for GMP cannot be met. Training materials for designers, fabricators, and healthcare institutions should be also developed to help break down barriers to communication and should be placed in the public domain rather than behind paywalls. Institutional changes should also include adding makers and engineers to incident command and procurement teams so that non-traditional designs can be more effectively vetted; model IRB protocols should be prepared for products that need clinical testing. Finally, patented designs should be placed in repositories for public use during health emergencies (67) to avoid delays from patent disputes. Such improvements to our infrastructure will dramatically improve our ability to respond to future medical pandemics and medical emergencies in the US and across the globe.

AUTHOR CONTRIBUTIONS

M-JA, DP, SS, NL, BL-E, and PS: article conception. M-JA, DP, SS, LA, AA, HY, JF, MS, BL-E, and PS: writing. M-JA, DP, SS, LA, AA, HY, AC, JF, MS, SY, NL, BL-E, and PS: editorial feedback. DP, HY, and PS: Greater Boston Pandemic Fabrication Team (PanFab) Consortium Coordination. All authors contributed to the article and approved the submitted version.

FUNDING

Local citizens and engineers have generously donated their time and resources to PanFab and they are essential to program success. This work was supported by the Harvard-MIT Center for Regulatory Science and by NIH/NCI grants P30-CA006516, U54-CA225088 (to PS, NL, and DP) and by T32-GM007753 (to DP) and by the Harvard Ludwig Center. AC is supported by the Hugh Hampton Young Fellowship of MIT. M-JA is a recipient of the Friends of McGovern Graduate Fellowship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.617106/full#supplementary-material

REFERENCES

1. *Shortage of Personal Protective Equipment Endangering Health Workers Worldwide*. WHO. Available online at: https://www.who.int/news-room/detail/03-03-2020-shortage-of-personal-protective-equipment-endangering-health-workers-worldwide.
2. *3D Printing Medical Equipment in Response to the COVID-19 Pandemic*. NIH Library (2020). Available online at: https://www.nihlibrary.nih.gov/services/3d-printing-service/3d-printing-medical-equipment-response-covid-19-pandemic.
3. *FAQs on 3D Printing of Medical Devices, Accessories, Components, Parts During the COVID-19 Pandemic*. U.S. Food and Drug Administration. (2020). Available online at: https://www.fda.gov/medical-devices/3d-printing-medical-devices/faqs-3d-printing-medical-devices-accessories-components-and-parts-during-covid-19-pandemic
4. NIH 3D Print Exchange. *A Collection of Biomedical 3D Printable Files and 3D Printing Resources Supported by the National Institutes of Health (NIH)*. Available online at: https://3dprint.nih.gov/.
5. Hanefeld J, Mayhew S, Legido-Quigley H, Martineau F, Karanikolos M, Blanchet K, et al. Towards an understanding of resilience: responding to health systems shocks. *Health Policy Plan*. (2018) 33:355–67. doi: 10.1093/heapol/czx183
6. Stevens R. *Why DIY 3D-Printed Face Masks and Shields Are So Risky*. Slate Magazine (2020). Available online at: https://slate.com/technology/2020/04/diy-3d-printed-face-masks-shields-coronavirus.html
7. Aitchison GA, Hukins DWL, Parry JJ, Shepherd DET, Trotman SG. A review of the design process for implantable orthopedic medical devices. *Open Biomed Eng J*. (2009) 3:21–27. doi: 10.2174/1874120700903010021
8. Alexander K, Clarkson PJ. A validation model for the medical devices industry. *J Eng Des*. (2002) 13:197–204. doi: 10.1080/09544820110108890
9. Pietzsch JB, Shluzas LA, Paté-Cornell ME, Yock PG, Linehan JH. Stage-gate process for the development of medical devices. *J Med Devices*. (2009) 3:021004.
10. Ogot MM, Kremer G. *Engineering Design: A Practical Guide*. Pittsburgh, PA: Trafford (2004).
11. Medina LA, KremerGEO, Wysk RA. Supporting medical device development: a standard product design process model. *J Eng Des*. (2013) 24:83–119. doi: 10.1080/09544828.2012.676635
12. Pietzsch JB, Aquino LM, Yock PG, Paté-Cornell ME, Linehan JH. Review of U.S. medical device regulation. *J Med Devices*. (2007) 1:283–92. doi: 10.1115/1.2812429
13. Winston W, Royce. Managing the development of large software systems. *Tech Pap West Electron Show Conv*. (1970) 8:1–9.
14. *Agile Project Management in Product Development Projects - ScienceDirect*. Available online at: https://www.sciencedirect.com/science/article/pii/S1877042814021259.
15. *Design Control Guidance for Medical Device Manufacturers*. U.S. Food and Drug Administration. (1997). Available online at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/design-control-guidance-medical-device-manufacturers
16. Fletcher B, Knight A, Pockrus B, Wain, MJ, Lehman-Huskamp K. Hospital incident command: First responders or receiving centers? *Am J Disaster Med*. (2016) 11:125–30. doi: 10.5055/ajdm.2016.0231
17. The Mass. *Manufacturing Community Responds to COVID-19*. MassTech. Available online at: https://masstech.org/M-ERT
18. Zeidel ML, Kirk C, Linville-Engler B. Opening up new supply chains. *N Engl J Med*. (2020) 32:e72. doi: 10.1056/NEJMc2009432
19. *America Makes - National Additive Manufacturing Innovation Institute*. America Makes. Available online at: https://www.americamakes.us/
20. *PanFab News - Latest Results*. Available online at: https://www.panfab.org/
21. *Hack the Pandemic – Copper 3D*. Antibacterial 3D Printing. Available online at:. Available online at: https://copper3d.com/hackthepandemic/
22. *Thingiverse.com. Mask Frame 'CEG Extreme' and Small n95 fitter (High Filtration with Halyard H600) by ctwiles*. Available online at: https://www.thingiverse.com/thing:4262131
23. *3D Printed Face Mask—No Worries on Mask Shortage and Coronavirus Infection*. Available online at: https://creality.com/info/makers-guide-3d-printed-face-mask-no-worries-on-mask-shortage-and-virus-infection-i00248i1.html
24. *COVID-19 Response*. Lowell Makes. (2020). Available online at: https://lowellmakes.com/covid-19-response/
25. Sher, D. *WASP Shares Open Source Processes for Production of Personalized PPE Masks and Helmets*. 3D Printing Media Network. (2020). Available online at: https://www.3dprintingmedia.network/personalized-ppe-mask/
26. *How to Make Bellus3D's Face Mask Fitter. Bellus3D: High-quality 3D Face Scanning*. Available online at: http://www.bellus3d.com/solutions/facemask
27. McAvoy M, Bui A-TN, Hansen C, Plana D, Said JT, Yu Z, et al. 3D printed frames to enable reuse and improve the fit of N95 and KN95 respirators. *medRxiv*. (2020). doi: 10.1101/2020.07.20.20151019
28. Attaran, M. The rise of 3-D printing: The advantages of additive manufacturing over traditional manufacturing. *Bus Horiz*. (2017) 60:677–88. doi: 10.1016/j.bushor.2017.05.011
29. *Technical Considerations for Additive Manufactured Medical Devices Guidance for Industry Food Drug Administration Staff*. U.S. Food and Drug Administration. (2017). Available online at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/technical-considerations-additive-manufactured-medical-devices
30. Dimitroulis G, Austin S, Sin Lee PV, Ackland D. A new three-dimensional, print-on-demand temporomandibular prosthetic total joint replacement system: preliminary outcomes. *J Cranio-Maxillo-fac Surg Off Publ Eur Assoc Cranio-Maxillo-fac Surg*. (2018) 46:1192–8. doi: 10.1016/j.jcms.2018.05.028
31. Popescu D, Laptoiu D. Rapid prototyping for patient-specific surgical orthopaedics guides: A systematic literature review. *Proc Inst Mech Eng*. (2016) 230:495–515. doi: 10.1177/0954411916636919
32. Ganguli A, Pagan-Diaz GJ, Grant L, Cvetkovic C, Bramlet M, Vozenilek J, et al. 3D printing for preoperative planning and surgical training: a review. *Biomed. Microdevices*. (2018) 20:65. doi: 10.1007/s10544-018-0301-9
33. George M, Aroom KR, Hawes HG, Gill BS, Love J. 3D printed surgical instruments: the design and fabrication process. *World J. Surg*. (2017) 41:314–9. doi: 10.1007/s00268-016-3814-5
34. Oliveira TT, Reis AC. Fabrication of dental implants by the additive manufacturing method: a systematic review. *J Prosthet Dent*. (2019) 122:270–4. doi: 10.1016/j.prosdent.2019.01.018
35. *Prusa Face Shield*. PrusaPrinters. Available online at: https://www.prusaprinters.org/prints/25857-prusa-face-shield
36. Mostaghimi A, Antonini M-J, Plana D, Anderson PD, Beller B, Boyer EW, et al. Regulatory and safety considerations in deploying a locally fabricated, reusable face shield in a hospital responding to the COVID-19 pandemic. *Med*. (2020) 1:139–51.e4. doi: 10.1016/j.medj.2020.06.003
37. *BIDMC-Led Clinical Trial Identifies Four Novel 3D-Printed Swabs for Use in COVID-19 Testing*. Available online at: https://www.bidmc.org/about-bidmc/news/2020/04/3d-printed-swabs
38. Callahan CJ, Lee R, Zulauf KE, Tamburello L, Smith KP, Previtera J, et al. Open development and clinical validation of multiple 3D-Printed sample-collection swabs: Rapid resolution of a critical COVID-19 testing bottleneck. *medRxiv*. (2020). doi: 10.1101/2020.04.14.20065094
39. *Home. COVID Swabs*. Available online at: https://printedswabs.org/
40. *Prisma Health introduces VESper*. Prisma Health. Available online at: https://www.prismahealth.org/vesper/
41. *Open Source COVID19 Medical Supply Guide*. Google Docs. Available online at: https://docs.google.com/document/d/1-71FJTmI1Q1kjSDLP0EegMERjg_0kk_7UfaRE4r66Mg/edit?usp=sharing&usp=embed_facebook
42. *Single-Use Devices. Disinfection & Sterilization Guidelines. Guidelines Library. Infection Control*. CDC. (2019). Available online at: https://www.cdc.gov/infectioncontrol/guidelines/disinfection/reuse-of-devices.html
43. *Enforcement Priorities for Single-Use Devices Reprocessed by Third Parties Hospitals*. U.S. Food and Drug Administration. (2000). Available online at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/enforcement-priorities-single-use-devices-reprocessed-third-parties-and-hospitals
44. *List N: Disinfectants for Use Against SARS-CoV-2*. US EPA. (2020). Available online at: https://www.epa.gov/pesticide-registration/list-n-disinfectants-use-against-sars-cov-2

45. Coronavirus Disease 2019 (COVID-19). *Decontamination and Reuse of Filtering Facepiece Respirators*. Centers for Disease Control and Prevention. (2020). Available online at: https://www.cdc.gov/coronavirus/2019-ncov/hcp/ppe-strategy/face-masks.html
46. US Food and Drug Administration. *Emergency Use Authorization*. Washington, DC: ASP STERRAD Sterilization Systems. (2020).
47. US Food and Drug Administration. *Emergency Use Authorization*. Washington, DC: Battelle Critical Care Decontamination System. (2020).
48. US Food and Drug Administration. *Emergency Use Authorization*. Washington, DC: Steris Corporation. (2020).
49. Use of International Standard ISO 10993-1. *Biological Evaluation of Medical Devices - Part 1: Evaluation Testing Within A Risk Management Process*. U.S. Food and Drug Administration. (2019). Available online at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-international-standard-iso-10993-1-biological-evaluation-medical-devices-part-1-evaluation-and
50. *Biocompatibility Assessment*. U.S. Food and Drug Administration. (2018). Available online at: https://www.fda.gov/medical-devices/cdrh-research-programs/biocompatibility-assessment
51. *14:00-17:00. ISO 10993-1:2018. ISO*. Available online at: https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/89/68936.html
52. *ISO 18562-1:2017(en), Biocompatibility Evaluation of Breathing Gas Pathways in Healthcare Applications — Part 1: Evaluation and Testing Within a Risk Management Process*. Available online at: https://www.iso.org/obp/ui#iso:std:iso:18562:-1:ed-1:v1:en
53. *COVID-19-Related Guidance Documents for Industry, FDA Staff, and Other Stakeholders*. FDA. (2020). Available online at: https://www.fda.gov/emergency-preparedness-and-response/coronavirus-disease-2019-covid-19/covid-19-related-guidance-documents-industry-fda-staff-and-other-stakeholders
54. *14:00-17:00. ISO 13485:2016*. ISO. Available online at: https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/97/59752.html
55. Speer J, Rish T. *ISO 14971 Risk Management For Medical Devices: The Definitive Guide*. Indianapolis: Greenlight. Guru (2020).
56. U.S. Department of Health & Human Services. *Secretary Azar Declares Public Health Emergency for United States for 2019 Novel Coronavirus*. HHS.gov. (2020). Available online at: https://www.hhs.gov/about/news/2020/01/31/secretary-azar-declares-public-health-emergency-us-2019-novel-coronavirus.html
57. *Emergency Use Authorizations*. U.S. Food and Drug Administration. (2020). Available online at: https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations
58. US Food and Drug Administration. *Emergency Use Authorization*. Washington, DC: Manufacturers of Face Shields. (2020).
59. US Food and Drug Administration. *Emergency Use Authorization*. Washington, DC: Respirators. (2020).
60. *Classify Your Medical Device*. U.S. Food and Drug Administration. (2020). Available online at: https://www.fda.gov/medical-devices/overview-device-regulation/classify-your-medical-device
61. *Learn if a Medical Device Has Been Cleared by FDA for Marketing*. U.S. Food and Drug Administration. (2018). Available online at: https://www.fda.gov/medical-devices/consumers-medical-devices/learn-if-medical-device-has-been-cleared-fda-marketing
62. *Safety Performance Based Pathway Guidance for Industry Food Drug Administration*. U.S. Food and Drug Administration. (2019). Available online at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/safety-and-performance-based-pathway
63. Neyman G, Irvin CB. A single ventilator for multiple simulated patients to meet disaster surge. *Acad Emerg Med*. (2006) 13:1246–9. doi: 10.1197/j.aem.2006.05.009
64. Denise MH. *Food and Drug Administration Emergency Use Authorization Issued in Response to Concerns Relating to Insufficient Supply and Availability of FDA-Cleared Ventilators for Use in Healthcare Settings to Treat Patients During the COVID-19 Pandemic*. Washington, DC: FDA (2020).
65. *Joint Statement on Multiple Patients Per Ventilator*. (2020). Available online at: https://www.asahq.org/about-asa/newsroom/news-releases/2020/03/joint-statement-on-multiple-patients-per-ventilator
66. *Medical Device Exemptions 510(k) GMP Requirements*. U.S. Food and Drug Administration. Available online at: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpcd/315.cfm
67. Sinha MS, Bourgeois FT, Sorger PK. Personal protective equipment for COVID-19: distributed fabrication and additive manufacturing. *Am J Public Health*. (2020) 110:1162–4. doi: 10.2105/AJPH.2020.305753
68. *MGB Center for COVID Innovation*. Available online at: https://covidinnovation.partners.org/.
69. N95DECON. Available online at: https://www.n95decon.org
70. *Product Classification*. U.S. Food and Drug Administration. Available online at: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPCD/classification.cfm?ID=FXX
71. *Enforcement Policy for Face Masks Respirators During the Coronavirus Disease (COVID-19) Public Health Emergency (Revised)*. U.S. Food and Drug Administration. (2020). Available online at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/enforcement-policy-face-masks-and-respirators-during-coronavirus-disease-covid-19-public-health
72. *Part 84—Approval of Respiratory Protective Devices. Electronic Code of Federal Regulations vol. §84.174 Filter Efficiency Level Determination Test—Non-Powered Series N, R, And P Filtration*. Washington, DC: CDC.
73. *Approval Tests and Standards for Air-Purifying Particulate Respirators*. Federal Register. (2020). Available online at: https://www.federalregister.gov/documents/2020/04/14/2020-07804/approval-tests-and-standards-for-air-purifying-particulate-respirators
74. *Medical Devices; Exemption From Premarket Notification: Class II Devices; Surgical Apparel*. Federal Register. (2018). Available online at: https://www.federalregister.gov/documents/2018/05/17/2018-10563/medical-devices-exemption-from-premarket-notification-class-ii-devices-surgical-apparel?fbclid=IwAR0p6t1z5BttFbyGIK6Krch565d0Y96UKauIqrXcDtqJbnntRBg0F7HtfeQ
75. Ouellette LL. *Written Description: Regulatory Responses to N95 Respirator Shortages. Written Description*. (2020). Available online at: https://writtendescription.blogspot.com/2020/04/regulatory-responses-to-n95-respirator.html
76. *ICC - International Code Council*. ICC. Available online at: https://www.iccsafe.org/.
77. Coronavirus Disease 2019 (COVID-19). *Interim Infection Prevention and Control Recommendations for Patients with Suspected or Confirmed Coronavirus Disease 2019 (COVID-19) in Healthcare Settings*. Centers for Disease Control and Prevention. (2020). Available online at: https://www.cdc.gov/coronavirus/2019-ncov/hcp/infection-control-recommendations.html
78. Roberge RJ. Face shields for infection control: a review. *J Occup Environ Hyg*. (2016) 13:235–242. doi: 10.1080/15459624.2015.1095302
79. Coronavirus Disease 2019 (COVID-19). *Strategies for Optimizing the Supply of Eye Protection*. Centers for Disease Control and Prevention. (2020). Available online at: https://www.cdc.gov/coronavirus/2019-ncov/hcp/ppe-strategy/eye-protection.html
80. Cramer AK, Plana D, Yang H, Carmack MM, Tian E, Sinha MS, et al. Analysis of SteraMist ionized hydrogen peroxide technology as a method for sterilizing N95 respirators and other personal protective equipment. *medRxiv*. (2020). doi: 10.1101/2020.04.19.20069997
81. Plana D, Tian E, Cramer AK, Yang H, Carmack MM, Sinha MS. Assessing the quality of nontraditional N95 filtering face-piece respirators available during the COVID-19 pandemic. *medRxiv [Preprint]*. (2020). doi: 10.1101/2020.07.25.20161968
82. *#GetUsPPE - Getting Protective Equipment to our Healthcare Heroes. Get Us PPE*. Available online at: https://getusppe.org/
83. *COVID19Masks: A PPE exchange. Helping Those Who Help Others*. Available online at: https://covid19masks.info/webapps/f?p=2222:1
84. *Mask Match*. Available online at: https://www.mask-match.com
85. *The National COVID-19 Critical Equipment Clearinghouse*. Project N95. Available online at: https://www.projectn95.org/
86. *PPE Link*. Available online at: https://ppelink.wordpress.com/

87. *Open Source Medical Supplies*. Available online at: https://opensourcemedicalsupplies.org/
88. Chesto J. *P&G Is Making Tens of Thousands Of Face Shields at Gillette plant in Southie - The Boston Globe*. Available online at: https://www.bostonglobe.com/2020/04/13/business/pg-is-making-tens-thousands-face-shields-gillette-plant-southie/
89. Walrath R. *Lovepop Begins Manufacturing Personal Protective Equipment*. AmericanInno. (2020). Available online at: https://www.americaninno.com/boston/inno-news-boston/3d-card-startup-lovepop-begins-manufacturing-protective-equipment/

Conflict of Interest: PS is a member of the SAB or Board of Directors of Applied Biomath, Glencoe Software and RareCyte Inc., and has equity in these companies. In the last five years the Sorger lab has received research funding from Novartis and Merck. PS declares that none of these relationships are directly or indirectly related to the content of this manuscript. NL is a consultant for or has received honoraria from the following companies: Seattle Genetics, Sanofi, and Bayer. JF was employed by company Fikst Product Development.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Citation: Antonini M-J, Plana D, Srinivasan S, Atta L, Achanta A, Yang H, Cramer AK, Freake J, Sinha MS, Yu SH, LeBoeuf NR, Linville-Engler B and Sorger PK (2021) A Crisis-Responsive Framework for Medical Device Development Applied to the COVID-19 Pandemic. Front. Digit. Health 3:617106. doi: 10.3389/fdgth.2021.617106

frontiers in Digital Health

MINI REVIEW
published: 29 March 2021
doi: 10.3389/fdgth.2021.564906

COVID-19 and Computer Audition: An Overview on What Speech & Sound Analysis Could Contribute in the SARS-CoV-2 Corona Crisis

Björn W. Schuller[1,2,3], Dagmar M. Schuller[3], Kun Qian[4], Juan Liu[5], Huaiyuan Zheng[6] and Xiao Li[7]*

[1] GLAM – Group on Language, Audio & Music, Imperial College London, London, United Kingdom, [2] EIHW – Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Augsburg, Germany, [3] audEERING GmbH, Gilching, Germany, [4] Educational Physiology Laboratory, The University of Tokyo, Tokyo, Japan, [5] Department of Plastic Surgery, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [6] Department of Hand Surgery, Wuhan Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [7] Department of Neurology, Children's Hospital of Chongqing Medical University, Chongqing Medical University, Chongqing, China

OPEN ACCESS

Edited by:
Matthew Crowson, Massachusetts Eye and Ear Infirmary and Harvard Medical School, United States

Reviewed by:
Chi-Chun Lee, National Tsing Hua University, Taiwan
Pablo Arias, Lund University, Sweden

***Correspondence:**
Björn W. Schuller
schuller@ieee.org

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *22 May 2020*
Accepted: *03 February 2021*
Published: *29 March 2021*

Citation:
Schuller BW, Schuller DM, Qian K, Liu J, Zheng H and Li X (2021) COVID-19 and Computer Audition: An Overview on What Speech & Sound Analysis Could Contribute in the SARS-CoV-2 Corona Crisis. Front. Digit. Health 3:564906. doi: 10.3389/fdgth.2021.564906

At the time of writing this article, the world population is suffering from more than 2 million registered COVID-19 disease epidemic-induced deaths since the outbreak of the corona virus, which is now officially known as SARS-CoV-2. However, tremendous efforts have been made worldwide to counter-steer and control the epidemic by now labelled as pandemic. In this contribution, we provide an overview on the potential for computer audition (CA), i.e., the usage of speech and sound analysis by artificial intelligence to help in this scenario. We first survey which types of related or contextually significant phenomena can be automatically assessed from speech or sound. These include the automatic recognition and monitoring of COVID-19 directly or its symptoms such as breathing, dry, and wet coughing or sneezing sounds, speech under cold, eating behaviour, sleepiness, or pain to name but a few. Then, we consider potential use-cases for exploitation. These include risk assessment and diagnosis based on symptom histograms and their development over time, as well as monitoring of spread, social distancing and its effects, treatment and recovery, and patient well-being. We quickly guide further through challenges that need to be faced for real-life usage and limitations also in comparison with non-audio solutions. We come to the conclusion that CA appears ready for implementation of (pre-)diagnosis and monitoring tools, and more generally provides rich and significant, yet so far untapped potential in the fight against COVID-19 spread.

Keywords: corona virus, SARS-CoV-2, COVID-19, computer audition, machine listening, computational paralinguistics

1. INTRODUCTION

The World Health Organisation's (WHO) office in China was first made aware of the previously unknown SARS-CoV-2 "Corona" virus on the last day(s) of the last year. On March 11, 2020, the WHO declared the disease triggered by the virus—COVID-19—as pandemic. The spread of the disease induced by the SARS-CoV-2 or "Corona" virus is assumed to underlie an exponential growth. However, whether there are long-term effects after recovery is yet to be fully researched. In the light of this dramatic spread, one is currently internationally witnessing drastic countermeasures that have not been seen in this form over decades in many countries. These include significant public "shut-down" measures to foster "social distancing" in order to slow down and control further spread.

As research globally is making massive efforts to contribute to better understand and fight the phenomenon from a medical and interdisciplinary point of view, also computer science and engineering in terms of "Digital Health" solutions aim at maximum exploitation of available and realisable means. In particular, in combination with artificial intelligence (AI), one can exploit a powerful tool, which so far has largely been tapped for prediction of COVID-19 spread [cf., e.g., (1)], and computer vision (CV) approaches in the corona context such as for automatic screening for COVID-19 on CT images (2, 3). There is, however, broader belief that also other signals including such from sensors on a smartphone could help even in the diagnosis of COVID-19 (4), e.g., the heart rate sensor.

In the following, we aim to provide an overview on what computer audition (CA), i.e., the application of computing for audio processing including "machine listening," "computational paralinguistics," and more general speech and sound analysis, but also synthesis, could contribute in this situation. To the best of the authors' knowledge, this resource is so far not used in practise despite offering a plethora of opportunities in this context.

The remainder of this overview is structured as follows: We first summarise phenomena more and less closely related to the case of COVID-19 that have already been targeted by CA and would be readily available. Examples include automatic recognition of speakers suffering from a cold or wearing a mask, breathing, coughing and sneezing sounds, or recognition of audio in spatial proximity. We then shift to the introduction of concrete use-cases how CA could benefit the ongoing global fight against the corona crisis. Subsequently, we introduce challenges and entry barriers from a technical as well as ethical and societal point of view, and discuss limitations before concluding this overview.

2. COMPUTER AUDITION: RELATED PHENOMENA

In the following, we set out by show-casing what CA has already successfully targeted as audio use-cases for recognition, and appears related to the task of interest in this contribution—fighting the ongoing COVID-19 spread.

2.1. Speech Analysis

Speech analysis by computational means is highly related to the field of computational paralinguistics (5). The field has several related recognition tasks on offer. These are often well-documented in the framework of competitive challenge events such as the Interspeech Computational Paralinguistics Challenge (ComParE). The latter has—perhaps closest related to the COVID-19 case—in its 2017 edition featured the automatic recognition of speech under cold (6), i.e., automatically recognising speakers affected by a cold from the acoustics of their voice. In the challenge of last year, the continuous assessment of breathing patterns from the speech signal appears relevant (7), e.g., as basis to recognise often witnessed symptoms of short-breathiness and breathing difficulties related to COVID-19. The last ComParE challenge further targets the recognition of speech under mask, i.e., the automatic recognition whether a speaker is wearing a facial protective mask, and the recognition of emotion of elderly, which may become interesting in monitoring the aftermath of social isolation of elderly, as was discussed, e.g., in the U.K. for 3 months. On the age scale's opposite end, toddlers' crying sounds seem to be the only indicator to understand if they are suffering from COVID-19 symptoms. In the ComParE challenge series, infant crying was investigated in 2018 (8), and the valence, i.e., positivity of baby sounds in 2019 (9). As symptoms of COVID-19 can also include lack of appetite, it seems further interesting to reference to the EAT challenge (10): In this event, it was demonstrated that one can infer from audio whether speech under eating and eating sounds indicate eating difficulty and "likability" related to whether one enjoys eating. The assessment of sleepiness—a further symptom of COVID-19—was first featured in ComParE in 2011 (11) as binary task, and as continuous sleepiness assessment on the Karolinska sleepiness scale in 2019 (9). Also pain such as headache or bodily pain can accompany COVID-19; speech under pain has also been shown to be automatically accessible (12, 13). When it comes to individual risk assessment and monitoring, speaker traits may be of interest. High mortality risk groups include the elderly, and a (slightly) higher mortality rate was so far seen in male individuals (14). Age and gender were also shown in the context of ComParE, and can be automatically determined reliably given sufficient speech material (15). A history of health issue can further indicate high risk. A number of health-related speaker state information relevant in this context has been shown feasible such as individuals suffering from asthma (16), head-and-neck cancer (17), or smoking habits (18, 19).

Speaker diarization, i.e., determining who is speaking when, and speaker counting (20) can become of interest in the ongoing social distancing. When it comes to counter measures such as quarantine, or risk assessment of individuals, one could also consider the usage of automatic recognition of deceptive speech when people are questioned about their recent contacts or whereabouts, as their personal work and life interests may interfere with the perspective of being sent to quarantine. Deception and sincerity were targeted in ComParE in 2016 (21). Monitoring well-being of individuals during social distancing and quarantine can further find interest in depression and fear recognition. Both were shown feasible to be assessed from speech

in the Audio/Visual Emotion Challenge (AVEC) event series (22) including from speech only at reasonable deviation on a continuous scale.

Generally speaking, speech audio also includes textual cues. Broadening up to Spoken Language Processing (SLP), this can also be of help to gather and analyse information from spoken conversations available in individual communications, news, or social media. For textual cues, this has already been considered (23). From a speech analysis perspective, this includes automatic speech recognition (ASR) and natural language processing (NLP).

2.2. Sound Analysis

From a sound analysis perspective, one may first consider such interest for COVID-19 use-cases that are produced by the human body. In the context of COVID-19, this includes mostly the automatic recognition of coughs (24–26) including dry vs. wet coughing (27) and dry vs. productive coughing (28) and sneeze (26), swallowing, and throat clearing (25) sounds—all showcased at high recognition rates. As severe COVID-19 symptoms are mostly linked to developing a pneumonia, which is the cause of most deaths of COVID-19 as suggested by post-mortem biopsies (29, 30), it further appears of interest that different breathing patterns, respiratory sounds, and lung sounds of patients with pneumonia can be observed through CA (31), even with mass devices such as smart-phones (32). Of potential relevance could also be the already possible monitoring of different types of snoring sounds (33), including their excitation pattern in the vocal tract and their potential change over time to gain insight on symptoms also during sleep. Further, highest risk of mortality from COVID-19 has been seen for such suffering from cardiovascular disease followed by chronic respiratory disease. In ComParE 2018, heart beats were successfully targeted from audio for three types of heart status, namely, normal, mild, and moderate/severe abnormality. Hearing local proximity from ambient audio further appears possible (34), and could be used to monitor individuals potentially too close to each other in the "social distancing" protective countermeasure scenarios. 3D audio localisation (35) and diarization further allows for locating previously recognised sounds and attributing them to sources. This could further help in the monitoring of public spaces or providing warnings to users as related to individuals potentially being locally too close with directional pointers. Audio source separation and denoising (36) of stethoscope sounds and audio (37) for clinicians and further processing can additionally serve as tool.

3. POTENTIAL USE-CASES

Let us next elaborate on use-cases we envision as promising for CA in the context of COVID-19. A coarse visual overview on the dependence of CA tasks and these use-cases is provided in **Table 1**. Check-marks indicate that the already available automatic audio analysis tasks listed in the left column appear of interest in the three major use-case groups listed in the right-most three columns. Note that these are indicative in nature. Further, to provide an impression of the "readiness," performance indications are given. For a strict comparability of these, they are only provided for tasks that have been featured in the Interspeech Computational Paralinguistics Challenge (ComParE) series.[1] Shown are the best results after the challenge including by fusion of best participant systems. Likewise, it is assured that test-set labels were unknown to participants and a strict subject independence and challenging conditions including no ability for "cherry picking" alike preselection of test examples are assured. The results overall show that under realistic conditions, the tasks are handled highly above chance level, yet, clearly below "perfect" recognition.

3.1. Risk Assessment

A first use-case targets the prevention of COVID-19 spread by individual risk assessment. As shown above, speaker traits such as age, gender, or health state can be assessed automatically from the voice to provide an estimate on the individual mortality risk level. In addition, one can monitor if oneself or others around are wearing a mask when speaking, count speakers around oneself, and locate these and their distance to provide a real-time ambient risk assessment and informative warning.

3.2. Diagnosis

While the standard for diagnosis of COVID-19 is currently a nasopharyngeal swab, several other possibilities exist including chest CT-based analysis as very reliable resource as outlined above. Here, we consider whether an audio-based diagnosis could be possible. While it seems clear that such an analysis will not be suited to compete with the state-of-the-art in professional testing previously named, its non-invasive and ubiquitously available nature would allow for individual pre-screening "anywhere," "anytime," in real-time, and available more or less to "anyone." To the best of the authors' knowledge, no study has yet systematically investigated audio from COVID-19 patients vs. highly varied control group data including such suffering from influenza or cold and healthy individuals. Unfortunately, coughing and sneezing of COVID-19 patients does not differ significantly to human perception from "normal" patients. This includes lung and breathing sounds. However, (38) assume that abnormal respiratory patterns can be a clue for diagnosis. Overall, by that, it seems unclear if diagnosis from short audio samples of patients could be directly possible, given that most speech or body sounds are likely not to show significant differences for closely related phenomena such as influenza or cold, but a number of encouraging results show that breathing, coughing, and speech sounds could be suited (39). The current Interspeech 2021 ComParE event therefore features COVID-19 recognition from forced cough and speech.

Rather, we believe that a histogram of symptoms over time in combination with their onset appears highly promising. **Table 2** visualises this concept in a qualitative manner by coarse ternary quantification of each symptom or "feature" from a machine

[1] http://www.compare.openaudio.eu

TABLE 1 | Interdependence of computer audition (CA) tasks and potential use-cases in the context of the corona crisis.

CA task	COVID-19 relation	Performance	Risk assessment	Diagnosis	Monitoring
SPEECH ANALYSIS					
Age & Gender	L	53.6/85.7% UAR (4 age, 3 gender classes)	√		
Breathing	H	0.778 CC_r (with breathing sensor in [−1,1])		√	√
Cold	H	71.0% UAR (2 classes: yes/no)		√	
COVID-19	H		√	√	√
Crying (infants)	L	78.6% UAR (3 valence classes)		√	
Deception and sincerity	L	72.1%/.654 CC_p (2 classes: yes/no/in [0,1])			√
Depression	L				√
Emotion (incl. of elderly)	L	63.8% UAR (mean 2 dimensions × 3 levels)			√
Health state	H		√		√
Lung sounds	H			√	√
Mask	H	82.6% UAR (2 classes: yes/no)	√		√
Pain	L		√	√	
Personality	L	70.4% UAR (mean 5D × 2-classes: ±)	√	√	
Sleepiness	L	72.5% UAR (2 classes)		√	√
SLP	L				√
Speaker count	L		√		√
SOUND ANALYSIS					
Coughing (dry, wet, productive)	H			√	√
Cardiovascular disease	L	56.2% UAR (3 classes)	√		√
Diarization	L		√		√
Localisation	L		√		√
Proximity	L		√		√
Sneezing	H			√	√
Snoring	L	58.5% UAR (4 classes)		√	√
Source separation	L		√		√
Swallowing	H			√	√
Throat clearing	H			√	√

The first column indicates the degree of immediate relation to COVID-19 as high ("H") or low ("L"). The column performance gives an impression of the reliability; for comparability, this is only shown for tasks featured under the same subject-independent challenging conditions in the Interspeech Computational Paralinguistics Challenge (ComParE)—the number of classes or scale and measure of performance are given together with the best result at the end of the event including by fusion of participating systems. UAR, unweighted average recall, i.e., sum of recall divided by number of classes, CC_r/CC_p: Pearon's/Spearman's correlation coefficient. SLP, spoken language processing. Health state encompasses a wider range of health factors such as asthma, head and neck cancer, or smoking habits that can be inferred from speech audio.

learning perspective.[2] Each of the symptoms in the table can—as outlined above—(already) be assessed automatically from an intelligent audio sensor. In a suited personal application such as on a smartphone or smartwatch, smart home device with audio abilities, or via a telephone service, etc., one could collect frequency of symptoms over time and from the resulting histogram differentiate with presumably high success rate between COVID-19, influenza, and cold. By suited means of AI, a probability could be given to users how likely their symptoms speak for COVID-19. Of particular interest thereby is also the "Onset Gradient" feature in **Table 2**. It alludes to whether the onset of symptoms over time is gradual (i.e., over the span of up to 2 weeks or more) or rather abrupt (i.e., within hours or a few days only), which can be well-observed by AI analysis in a histogram sequence updated over time. Collecting such information from many users, this estimate for histogram-based diagnosis of COVID-19 can be improved in precision over time if users "donate their data." In addition, clinicians could be given access to the histogram or be pointed to typical audio examples in a targeted manner remotely that have been collected over longer time to speed up the decision whether the users should go for other more reliable forms of testing. This could help to highly efficiently pre-select individuals for screening.

3.3. Monitoring of Spread

Beyond the idea of using smartphone-based surveys and AI methods to monitor the spread of the virus (40), one could use CA for audio analysis via telephone or other spoken conversation. An AI could monitor the spoken conversations and screen for speech under cold or other symptoms as shown in **Table 2**. Together with GPS coordinates from smart phones or knowledge of the origin of the call from the cell, one could establish real-time spread maps.

[2]based on https://www.qld.gov.au, https://www.medicinenet.com/, https://www.medicalnewstoday.com, all assessed on March 20, 2020.

TABLE 2 | Qualitative behaviour of symptoms of COVID-19 vs. cold and influenza (flu): Tentative histogram by symptom ("feature"/"variable") in ternary quantification [from no/low ("+") to frequent/high ("+++")].

Symptom	COVID-19	Influenza	Cold
Breathing: Dypnea (Shortness)	+++	++	+
Breathing: Difficulty	+++	++	+
Rhinorrhea (Running nose)	+	++	+++
Nasal congestion	+	+	+++
Coughing	dry ++	dry ++	+
Sneezing	+	+	+++
Sore throat	+	++	+++
Pain: Body	+	+++	++
Pain: Head (Headache)	++	+++	+
Fatigue, Tiredness	mild ++	+++	+
Appetite loss	+	+++	+
Onset gradient	+	+++	+

Shown is also the symptom gradient onset behaviour. Further frequently related variables include behaviour and personality changes, diarrhoea, fever, sore tongue, or watery eyes, which could partially be assessed also by audio—the latter two rather by physiological and visual sensors, respectively. Assembled from diverse references, the table is indicative in nature on purpose, and more fine-grained quantification could apply.

3.4. Monitoring of Social Distancing and Effects

Social distancing—in already diagnosed cases of COVID-19 or direct contact isolation of individuals—might lead to different negative side effects. People who have less social connexion might suffer from even a weaker immune system, responding less well to pathogens (41). Especially, the high-risk target group of elderly could even encounter suicidal thoughts and develop depression or other clinical conditions in isolation (42). CA might provide indications about social interaction, exemplary speaking time during the day via phone or other devices, as well as measure emotions of the patient throughout the day or detecting symptoms of depression and suicidal risk (43).

In addition, the public obedience and discipline in social distancing could be monitored with the aid of CA. AI allows to count speakers, locate them and their potential symptoms as reflecting in the audio signal (cf. **Table 2**), and "diarize" the audio sources, i.e., attribute which symptoms came from which (human) individual. Likewise, public spaces could be empowered by AI that detects potentially risky settings, which are overcrowded, under-spaced in terms of distance between individuals, and spot potentially COVID-19 affected subjects among a crowd, and whether these and others are wearing a protective mask while speaking.

3.5. Monitoring of Treatment and Recovery

During hospitalisation or other forms of treatment and recovery, CA can monitor the progress, e.g., by updating histograms of symptoms. In addition, the well-being of patients could be monitored similarly to the case of individual monitoring in social distancing situations as described above. This could include listening to their emotions, eating habits, fatigue, or pain, etc.

3.6. Generation of Speech and Sound

While we have focused entirely on the analysis of audio up to this point, it remains to state that there may be also use-cases for the generation of audio by AI in a COVID-19 scenario. Speech conversion and synthesis could help those suffering from COVID-19 symptoms to ease their conversation with others. In such a setting, an AI algorithm can fill in the gaps arising from coughing sounds, enhance a voice suffering from pain or fatigue and further more by generative adversarial networks (44). In addition, alarm signals could be rendered which are mnemonic and re-recognisable, but adapt to the ambient sound to be particularly audible.

4. CHALLENGES

4.1. Time

The fight against COVID-19 has been marked by a race to prevent too rapid spread that could lead to peak infection rates that overburden the national health systems and availability of beds in the intensive care units leading to high mortality rates. Further, at presence, it cannot be clearly stated whether or not COVID-19 will persistently stay as disease. However, recent research and findings (45) as well as model calculations indicate that COVID-19 will continue to heavily spread over the next months in different areas of the world. Enhanced social distancing might delay the spreading. Additionally, at the moment there is no solid research available to prove persistent immunity against the virus after an infection with COVID-19. Therefore, the need to apply measures of enhanced risk assessment, diagnosis, monitoring, and treatment is urgently necessary to support the current medical system as well as to get COVID-19 under control.

4.2. Collecting COVID-19 Patient Data

Machine learning essentially needs data to learn. Accordingly, for any kind of CA application targeting speech or sounds from patients suffering from COVID-19 infection, we will need collected and annotated data. At present, such data are hardly publicly available for research purposes, but urgently needed. Hence, a crucial step in the first place will be to collect audio data including highly validated such from diagnosed patients and ideally control subjects under equal conditions and demographic characteristics including control data with a rich representation of further respiratory and other diseases.

4.3. Model Sharing

In order to accelerate the adaptation of machine learning models of CA for COVID-19, exchange of data will be crucial. As such data are usually highly private and sensitive in nature, the recent advances in federated machine learning (46) can benefit the exchange of personal model parameters rather than audio to everyone's benefit. Likewise, users of according services can "donate their data" in a safe and private manner.

4.4. Real-World Audio Processing

Most of the tasks and use-cases listed above require processing of audio under more or less constrained "in-the-wild" conditions

such as audio recording over telephone, VoIP, or audio takes at home, in public spaces, or in hospitals. These are usually marked by noise, reverberation, varying distance to microphone(s), transmissions with potential loss, and further disturbances. In addition, given the pandemic character of the SARS-CoV-2 corona crisis, one will ideally need to be robust against multilingual, multicultural, or local speech and sound variability.

4.5. Green Processing

Green processing summarises here the idea of efficiency in computing. This will be a crucial factor for mobile applications, but also for big(ger) data speech analysis (47) such as in the case of telephone audio data analysis. It includes conservative consumption of energy such as on mobile devices, efficient transmission of data such as in the above named federated machine learning in order not to burden network transmission, memory efficiency, model update efficiency, and many further factors.

4.6. Trustability of Results

Machine learning and pattern recognition methods as used in CA are usually statistical learning paradigms and hence prone to error. The probability of error needs to be (a) estimated, known as confidence measure estimation, and (b) communicated to users of CA services in the COVID-19 context to assure trustability of these methods. One step further is that results should ideally be explainable. However, eXplainable AI (XAI) itself is at this time a young discipline, but provides an increasing method inventory allowing for interpretation of results (48).

4.7. Ethics

Many of the above suggested use-cases come at massive ethical responsibility and burden, which can often only be justified in times of global crisis as the current one. This includes mostly many of the above sketched applications of CA for monitoring. Assuring privacy at all times will be crucial to benefit only the goal of fighting COVID-19 spread without opening doors for massive miss-use. At the same time, balancing between individual interests and the beneficence of groups and societies will need to be carefully considered.

In addition, apart from responsible research, it needs to be assured that the data are representative of all users in all use-cases avoiding potential algorithmic bias. Indeed, the suggested CA algorithms could function better for some parts of the population, because algorithms were trained with data from only one subculture due to different access to resources/technologies. As an example, this could create an asymmetry in the detection of symptoms of subparts of a population (inter- and intra-countries). Deploying the same solution at scale would favour certain social groups and disfavour others (49).

Next, one must assure that common points of reference for comparison across studies are given, the aim of an audio task is well-decided upon, results are interpretable, and communicated to all, including in particular communication of potential limitations (50). Further concerns in this context will discuss legal and societal implications. All of these cannot be discussed here—rather, we can provide pointers for the interested reader as starting points (50–55).

While the technology seems ripe for application, one may ask if we should use it? Or, are the ethical questions that rise from these technologies enough to pause the development of the suggested CA techniques at scale and think on the ethical solutions first? And, do we have enough ethical knowledge today to put enough constraints on the suggested CA applications to make them secure/ethical? These questions touch upon many actions that have been taken during the ongoing pandemic, but of course, this will not justify risking massive personal data leakage or restrictions of personal rights due to missclassifications by AI or more specifically CA. Decisions will need to be made individually per use-case and potential CA solution carefully weighing benefit against risk.

5. LIMITATIONS

Following the described advantages and the potential uses of these technologies in the case of COVID-19, we now provide a critical thinking about their limitations and discussion about their usability in the described use-cases.

First, as to the tasks described in **Table 1**, in a non-negligible number of cases the data used for the experiments are still simulated. Hence, the extrapolation from this to real-world scenarios is far from being trivial. Also, in all use-cases, we assumed that there is access to ideal sound recordings so one can track a person at home and in public spaces. Today, cities and homes are not equipped with microphone networks, so smartphones are the preferred choice. On the one hand, these recordings are potentially extremely noisy, reverberated, and marked by package loss, which limits the applicability of the previous research; on the other hand, most use-cases assume that the smartphone is constantly listening, and that AIs are able to detect, for example, if a person is eating (even before wondering if they are eating with appetite or not). Furthermore, it may appear difficult to see how we can envision the application for locating and detecting sound sources from recordings made by smartphones.

Few or none of the technologies mentioned are fully operational today, so we can use them effectively for the proposed objectives (as opposed to computer vision which is already commonly used). And if time is indeed a challenge in this period, it seems that the time necessary to exploit CA efficiently for the COVID-19 task could be the biggest challenge, as software development, deployment, maintenance, testing including medical such, and alike are usually very demanding in time.

Further, it has been noticed that only some specific elements are directly related to COVID-19 (such as pre-diagnosis of COVID-19 from breathing, coughing, or speech). In particular, for the various paralinguistis recognition tasks, these would otherwise further include lung sounds as compared to others diseases such as common cold/influenza. An indication on the degree of immediate relation to COVID-19 is provided in **Table 1**. On the other hand, many of the introduced aspects

TABLE 3 | Nine key aspects: Promising complimentarity of CA with other methods, tentative advantages ("+") vs. disadvantages ("-") or equality ("0") of using audio as the modality as compared to other more established methods used at scale in the medical and related setting.

Task	Method	Complementary	Reliability	Robustness	Speed	Effort	Area covered	Accessibility	Privacy	Cooperation requirement	Information richness
Risk assessment	5G/BT contact tracing	√	−	−	+	0	+	+	−	+	+
	Temperature measurement	√	+	+	0	0	+	0	−	−	+
Diagnosis	Blood sample, chest x-ray, heart-rate sensing, upper respiratory sample	√	−	−	+	+	+	+	−	0	+
Monitoring: spread,	Wastewater tracking		0	+	+	0	+	−	−	+	
social distancing,	Video-based	√	−	+	0	+	+	+	0	−	0
	Chip-based		−	−	0	0	+	+	−	+	+
distancing effects,	Analysis by physiology	√	+	+	0	+	+	+	−	+	+
	Analysis by text		0	−	0	−	−	0	−	0	+
	Analysis by video	√	0	0	0	+	+	+	0	−	0
Treatment/recovery	Medical testing		−	−	+	+	+	+	−	0	0
	Human assessment		0	−	0	+	0	+	+	0	−

Obviously, in general, audio-based assessment requires presence of audio in the first place, which can be positive in terms of privacy, but negative for continuous assessment or such without cooperation requirement. BT, bluetooth.

bear interest even from monitoring of cold/influenza and other respiratory or even related viral diseases perspective.

Also, in many of the cases described, another simpler means can certainly be used instead to arrive at the same information. For example, bio-signals allow a more direct and more efficient measurement of a person's state of health and its evolution; smartphones and GPS tracking are very effective in locating individuals. Hence, a multi-modal combination of audio and other modalities appears very promising depending on the individual requirements and settings.

In **Table 3**, we hence investigate whether the suggested CA applications would work better than those that are already implemented at scale or in high technology-readiness state. On a similar line, we indicate where the mentioned CA techniques could complement the current monitoring methods particularly well. We provide the most common alternative methods for the three major use-cases risk-assessment, diagnosis, and monitoring as introduced in section 3. Other alternatives are recently developed and need to be related in a similar fashion to CA (56), once being ready for usage at scale. Also note that the table merely presents a coarse indication. It will depend on the detailed usage which approach is to be preferred. For example, the indicated equality in effort for contact tracing by bluetooth or alike vs. CA is a coarse estimate as, on the one hand, a high cost for development, advertisement, and distribution of such solutions is required. On the other hand, a centralised service based on CA will also come at a high effort: CA population tracking could be extremely expensive in terms of resources and development. Indeed, application development, server infrastructure, data analysis centers, data encryption, storage, and anonymisation as well as all the costs of maintaining these services could easily add up transforming CA solutions in over-priced servers difficult to maintain and without the certainty that they will detect large numbers of COVID-19 cases.

As to the alternatives to CA by use-case, for risk assessment, this is currently mainly achieved by the named contact tracing apps run on one's smart phone by bluetooth or 5G methods (57). When active on the phones of two individuals in sufficiently close proximity such as less than 2 m for a minimum set time of more than 15 min, the contacts are stored (usually only locally in anonymous ways). Users that report COVID-19 positive diagnosis are informed back to the service in anonymous forms. Such an approach has been used already for other infectious diseases (e.g., for HIV or tuberculosis). Another increasingly used method is thermal camera based body temperature measurement (58) often in the context of access, which has also been vastly used before, e.g., at airports. A single thermal camera can be used together with subject tracking to assess many individuals by a single device.

For diagnosis, the common present alternatives used at scale are upper respiratory samples such as regarding reverse transcription-polymerase chain reaction (RT-PCR), and blood samples (58), or, chest X-ray. A mobile health alternative is found by intelligent heart rate analysis such as from wrist-worn devices (59).

Monitoring of spread can alternatively, for example, be fulfilled by analysing wastewater (60).

Monitoring of social distancing is realisable also by video-based tracking (61) of individuals or chip-based such as via 5G, bluetooth, or NFC and related technologies, as used, e. g., in factories (57) usually requiring each participant to be equipped, accordingly. Monitoring of social distancing effects is largely related to affective, behavioural, and social computing in more general, for which there exist a range of other modalities—mainly physiology, text, or video (51). For monitoring of

social distancing treatment and recovery, mainly the usage of medical testing and mere human assessment form major alternative options.

How CA or CA combinations would compare (e.g., in terms of false positives/false negative rates or detection time) to the already implemented medical and alternatives systems in society will need to be broken down in detail. For instance, questions such as do we have any clues to think that CA will be more robust in terms of diagnostic than the current medical monitoring system will need careful further investigation.

Besides such more technical questions, practical questions on acceptance will also largely dominate the usefulness of CA methods for COVID-19. The tracing application experience has in some countries shown that the population was not fully ready to give away their data. Similar of even bigger societal challenges and limitations of the deployment of CA applications in the context of COVID-19 at scale need to be expected.

6. DISCUSSION

In this short overview, we provided pointers toward what CA could potentially contribute to the ongoing fight against the world-wide spread of the SARS-CoV-2 virus known as "corona crisis" and the COVID-19 infection triggered by it. We have summarised a number of potentially useful audio analysis tasks by means of AI that have already been demonstrated feasible. We further elaborated use-cases how these could benefit in this battle, and shown challenges arising from real-life usage. The envisioned use-cases included automated audio-based individual risk assessment, audio-only-based diagnosis directly from speech or (forced) cough-sounds and by symptom frequency and symptom development histograms over time in combination with machine learning, and several contributions to monitoring of COVID-19 and its effects including spread, social distancing, and treatment and recovery besides use-cases for audio generation. At the time of writing, it seems that what matters most is a rapid exploitation of this largely untapped potential. Obviously, in this short overview, not all possibilities could be included, and many further potential use-cases may exist. We also showed key limitations, but others will exist. Further, the authorship is formed by experts on CA, digital health, and clinicians having worked with COVID-19 infected patients over the last months—further insights from other disciplines will be highly valuable to add. The corona crisis demands for common efforts on all ends—we truly hope computer audition can add a significant share to an accelerated success of the crisis' defeat.

AUTHOR CONTRIBUTIONS

BS wrote the manuscript. All authors contributed ideas and input and read over the manuscript.

FUNDING

This work was further partially supported by the Zhejiang Lab's International Talent Fund for Young Professionals (Project HANAMI), P. R. China, the JSPS Postdoctoral Fellowship for Research in Japan (ID No. P19081) from the Japan Society for the Promotion of Science (JSPS), Japan, and the Grants-in-Aid for Scientific Research (No. 19F19081 and No. 17H00878) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. We acknowledge funding from the EU's HORIZON 2020 Grant No. 115902 (RADAR CNS).

ACKNOWLEDGMENTS

This manuscript has been released as a pre-print in a shorter version at arXiv.org (46). We express our deepest sorrow for those who left us due to COVID-19; they are not numbers, they are lives. We further express our highest gratitude and respect to the clinicians and scientists, and anyone else these days helping to fight against COVID-19, and at the same time help us maintain our daily lives.

REFERENCES

1. Hu Z, Ge Q, Jin L, Xiong M. Artificial intelligence forecasting of covid-19 in China. *arXiv preprint arXiv:200207112*. (2020). doi: 10.18562/IJEE.054
2. Gozes O, Frid-Adar M, Greenspan H, Browning PD, Zhang H, Ji W, et al. Rapid AI development cycle for the coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis. *arXiv preprint arXiv:2003.05037*. (2020).
3. Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *medRxiv*. (2020) 27. doi: 10.1101/2020.02.14.20023028v5
4. Maghdid HS, Ghafoor KZ, Sadiq AS, Curran K, Rabie K. A Novel AI-enabled framework to diagnose coronavirus COVID 19 using smartphone embedded sensors: design study. *arXiv preprint arXiv:200307434*. (2020). doi: 10.1109/IRI49571.2020.00033
5. Schuller B, Batliner A. *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing*. Chichester: Wiley (2013). doi: 10.1002/9781118706664
6. Schuller B, Steidl S, Batliner A, Bergelson E, Krajewski J, Janott C, et al. Computational paralinguistics challenge: addressee, cold & snoring. In: *Proceedings of Interspeech*. Stockholm: ISCA (2017). p. 3442–6. doi: 10.21437/Interspeech.2017-43
7. Schuller BW, Batliner A, Bergler C, Messner EM, Hamilton A, Amiriparian S, et al. Computational paralinguistics challenge: elderly emotion, breathing & masks. In: *Proceedings of Interspeech*. Shanghai: ISCA (2020). p. 2042–6. doi: 10.21437/Interspeech.2020-0032
8. Schuller BW, Steidl S, Batliner A, Marschik PB, Baumeister H, Dong F, et al. Computational paralinguistics challenge: atypical & self-assessed affect, crying & heart beats. In: *Proceedings of Interspeech*. Hyderabad: ISCA (2018). p. 122–6. doi: 10.21437/Interspeech.2018-51
9. Schuller BW, Batliner A, Bergler C, Pokorny F, Krajewski J, Cychosz M, et al. Computational paralinguistics challenge: styrian dialects, continuous sleepiness, baby sounds & orca activity. In: *Proceedings of Interspeech*. Graz: ISCA (2019). p. 2378–82. doi: 10.21437/Interspeech.2019-1122
10. Schuller D, Schuller B. The challenge of automatic eating behaviour analysis and tracking. In: Costin HN, Schuller BW, Florea AM, editors. *Recent Advances in Intelligent Assistive Technologies: Paradigms and Applications.*

Intelligent Systems Reference Library. Berline; Heidelberg: Springer (2020). p. 187–204. doi: 10.1007/978-3-030-30817-9_8
11. Schuller B, Steidl S, Batliner A, Schiel F, Krajewski J, Weninger F, et al. Medium-term speaker states-a review on intoxication, sleepiness and the first challenge. *Comput Speech Lang*. (2014) 28:346–74. doi: 10.1016/j.csl.2012.12.002
12. Oshrat Y, Bloch A, Lerner A, Cohen A, Avigal M, Zeilig G. Speech prosody as a biosignal for physical pain detection. In: *Proceedings 8th Speech Prosody*. Boston, MA (2016). p. 420–4. doi: 10.21437/SpeechProsody.2016-86
13. Ren Z, Cummins N, Han J, Schnieder S, Krajewski J, Schuller B. Evaluation of the pain level from speech: introducing a novel pain database and benchmarks. In: *Proceedings 13th ITG Conference on Speech Communication*. Oldenburg: ITG/VDE (2018). p. 56–60.
14. Caramelo F, Ferreira N, Oliveiros B. Estimation of risk factors for COVID-19 mortality-preliminary results. *medRxiv*. (2020) 12. doi: 10.1101/2020.02.24.20027268
15. Weninger F, Marchi E, Schuller B. Improving recognition of speaker states and traits by cumulative evidence: intoxication, sleepiness, age and gender. In: *Proceedings of Interspeech*. Portland, OR: ISCA (2012). p. 1159–62.
16. Mazić I, Bonković M, Džaja B. Two-level coarse-to-fine classification algorithm for asthma wheezing recognition in children's respiratory sounds. *Biomed Signal Process Control*. (2015) 21:105–18. doi: 10.1016/j.bspc.2015.05.002
17. Maier A, Haderlein T, Stelzle F, Nöth E, Nkenke E, Rosanowski F, et al. Automatic speech recognition systems for the evaluation of voice and speech disorders in head and neck cancer. *EURASIP J Audio Speech Music Process*. (2009) 2010:926951. doi: 10.1186/1687-4722-2010-926951
18. Poorjam AH, Bahari MH, Van hamme H. Multitask speaker profiling for estimating age, height, weight and smoking habits from spontaneous telephone speech signals. In: *Proceedings 4th International Conference on Computer and Knowledge Engineering (ICCKE)*. Masshad: IEEE (2014). p. 7-12. doi: 10.1109/ICCKE.2014.6993339
19. Satori H, Zealouk O, Satori K, Elhaoussi F. Voice comparison between smokers and non-smokers using HMM speech recognition system. *Int J Speech Technol*. (2017) 20:771–7. doi: 10.1007/s10772-017-9442-0
20. Xu C, Li S, Liu G, Zhang Y, Miluzzo E, Chen YF, et al. Crowd++ unsupervised speaker count with smartphones. In: *Proceedings ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*. Zurich (2013). p. 43–52. doi: 10.1145/2493432.2493435
21. Schuller B, Steidl S, Batliner A, Hirschberg J, Burgoon JK, Baird A, et al. Computational paralinguistics challenge: deception, sincerity & native language. In: *Proceedings of Interspeech*. San Francisco, CA: ISCA (2016). p. 2001–5. doi: 10.21437/Interspeech.2016-129
22. Valstar M, Gratch J, Schuller B, Ringeval F, Cowie R, Pantic M. Summary for AVEC 2016: depression, mood, and emotion recognition workshop and challenge. In: *Proceedings 24th ACM International Conference on Multimedia (MM)*. Amsterdam: ACM (2016). p. 1483–4. doi: 10.1145/2964284.2980532
23. Pandey R, Gautam V, Bhagat K, Sethi T. A Machine learning application for raising WASH awareness in the times of covid-19 pandemic. *arXiv preprint arXiv:200307074*. (2020). doi: 10.2196/preprints.25320
24. Matos S, Birring SS, Pavord ID, Evans H. Detection of cough signals in continuous audio recordings using hidden Markov models. *IEEE Trans Biomed Eng*. (2006) 53:1078–83. doi: 10.1109/TBME.2006.873548
25. Olubanjo T, Ghovanloo M. Tracheal activity recognition based on acoustic signals. In: *Proceedings 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Chicago, IL: IEEE (2014). p. 1436–9. doi: 10.1109/EMBC.2014.6943870
26. Amiriparian S, Pugachevskiy S, Cummins N, Hantke S, Pohjalainen J, Keren G, et al. CAST a database: rapid targeted large-scale big data acquisition via small-world modelling of social media platforms. In: *Proceedings 7th biannual Conference on Affective Computing and Intelligent Interaction (ACII)*. San Antionio, CA: IEEE (2017). p. 340–5. doi: 10.1109/ACII.2017.8273622
27. Moradshahi P, Chatrzarrin H, Goubran R. Improving the performance of cough sound discriminator in reverberant environments using microphone array. In: *Proceedings International Instrumentation and Measurement Technology Conference (I2MTC)*. Graz: IEEE (2012). p. 20–3. doi: 10.1109/I2MTC.2013.6555454
28. Schröder J, Anemiiller J, Goetze S. Classification of human cough signals using spectro-temporal Gabor filterbank features. In: *Proceedings International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Shanghai: IEEE (2016). p. 6455–9. doi: 10.1109/ICASSP.2016.7472920
29. Beigmohammadi MT, Jahanbin B, Safaei M, Amoozadeh L, Khoshavi M, Mehrtash V, et al. Pathological findings of postmortem biopsies from lung, heart, and liver of 7 deceased COVID-19 patients. *Int J Surg Pathol*. (2020). doi: 10.1177/1066896920935195
30. Tian S, Xiong Y, Liu H, Niu L, Guo J, Liao M, et al. Pathological study of the 2019 novel coronavirus disease (COVID-19) through postmortem core biopsies. *Modern Pathol*. (2020) 33:1007–14. doi: 10.1038/s41379-020-0536-x
31. Murphy RL, Vyshedskiy A, Power-Charnitsky VA, Bana DS, Marinelli PM, Wong-Tse A, et al. Automated lung sound analysis in patients with pneumonia. *Respirat Care*. (2004) 49:1490–7.
32. Song I. Diagnosis of pneumonia from sounds collected using low cost cell phones. In: *Proceedings International Joint Conference on Neural Networks (IJCNN)*. Killarney: IEEE (2015). p. 1–8. doi: 10.1109/IJCNN.2015.7280317
33. Janott C, Schmitt M, Zhang Y, Qian K, Pandit V, Zhang Z, et al. Snoring classified: the munich passau snore sound corpus. *Comput Biol Med*. (2018) 94:106–18. doi: 10.1016/j.compbiomed.2018.01.007
34. Pokorny FB, Fiser M, Graf F, Marschik PB, Schuller BW. Sound and the city: Current perspectives on acoustic geo-sensing in urban environment. *Acta Acust United Acust*. (2019) 105:766–78. doi: 10.3813/AAA.919357
35. Delikaris-Manias S, Pavlidi D, Pulkki V, Mouchtaris A. 3D localization of multiple audio sources utilizing 2D DOA histograms. In: *Proceedings 24th European Signal Processing Conference (EUSIPCO)*. Budapest: IEEE (2016). p. 1473–7. doi: 10.1109/EUSIPCO.2016.7760493
36. Liu S, Keren G, Schuller BW. N-HANS: introducing the Augsburg Neuro-Holistic Audio-eNhancement System. *arXiv preprint arXiv:1911.07062*. (2019).
37. Yang K, He Z, Yang W, Tang Q, Li D, Wang Z, et al. Heart sound denoising using computational auditory scene analysis for a wearable stethoscope. In: *Proceedings 56th International Midwest Symposium on Circuits and Systems (MWSCAS)*. Columbus, OH: IEEE (2013). p. 1220–3. doi: 10.1109/MWSCAS.2013.6674874
38. Wang Y, Hu M, Li Q, Zhang XP, Zhai G, Yao N. Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with COVID-19 in an accurate and unobtrusive manner. *arXiv preprint arXiv:2002.05534*. (2020).
39. Qian K, Schuller BW, Yamamoto Y. Recent advances in computer audition for diagnosing COVID-19: an overview. *arXiv preprint arXiv:2012.04650*. (2020).
40. Rao ASS, Vazquez JA. Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey in the populations when cities/towns are under quarantine. *Infect Control Hospital Epidemiol*. (2020) 41:826–30. doi: 10.1017/ice.2020.61
41. Cole SW, Levine ME, Arevalo JM, Ma J, Weir DR, Crimmins EM. Loneliness, eudaimonia, and the human conserved transcriptional response to adversity. *Psychoneuroendocrinology*. (2015) 62:11–7. doi: 10.1016/j.psyneuen.2015.07.001
42. Luo Y, Hawkley LC, Waite LJ, Cacioppo JT. Loneliness, health, and mortality in old age: a national longitudinal study. *Soc Sci Med*. (2012) 74:907–14. doi: 10.1016/j.socscimed.2011.11.028
43. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. *Speech Commun*. (2015) 71:10–49. doi: 10.1016/j.specom.2015.03.004
44. Pascual S, Bonafonte A, Serra J. SEGAN: Speech enhancement generative adversarial network. *arXiv preprint arXiv:170309452*. (2017). doi: 10.21437/Interspeech.2017-1428
45. Wu Z, McGoogan JM. Characteristics of and important lessons from the Coronavirus disease (2019). (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*. (2020) 323:1239–42. doi: 10.1001/jama.2020.2648
46. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: concept and applications. *ACM Trans Intell Syst Technol*. (2019) 10:1–19. doi: 10.1145/3298981

47. Verma JP, Agrawal S, Patel B, Patel A. Big data analytics: Challenges and applications for text, audio, video, and social media data. *Int J Soft Comput Artif Intell Appl*. (2016) 5:41–51. doi: 10.5121/ijscai.2016.5105
48. Adadi A, Berrada M. Peeking inside the black-box: a survey on Explainable Artificial Intelligence (XAI). *IEEE Access*. (2018) 6:52138–60. doi: 10.1109/ACCESS.2018.2870052
49. Ahmad MA, Patel A, Eckert C, Kumar V, Teredesai A. Fairness in machine learning for healthcare. In: *Proceedings of the KDD*. New York, NY: ACM (2020). p. 3529–30. doi: 10.1145/3394486.3406461
50. Batliner A, Hantke S, Schuller BW. Ethics and good practice in computational paralinguistics. *IEEE Trans Affect Comput*. (2020). doi: 10.1109/TAFFC.2020.3021015
51. Reynolds C, Picard R. Affective sensors, privacy, and ethical contracts. In: *Proceedings CHI'04 Extended Abstracts on Human Factors in Computing Systems*. (2004). p. 1103–06. doi: 10.1145/985921.985999
52. Kummer AW, Turner J. Ethics in the practice of speech-language pathology in health care settings. *Semin Speech Lang*. (2011) 32:330–7. doi: 10.1055/s-0031-1292758
53. Batliner A, Schuller B. More than fifty years of speech processing-the rise of computational paralinguistics and ethical demands. In: *Proceedings of ETHICOMP*. Paris: CERNA (2014).
54. Greene D, Hoffmann AL, Stark L. Better, nicer, clearer, fairer: a critical assessment of the movement for ethical artificial intelligence and machine learning. In: *Proceedings 52nd Hawaii International Conference on System Sciences (HICSS)*. Kauai, HI (2019). p. 2122–31. doi: 10.24251/HICSS.2019.258
55. Nebeker C, Torous J, Ellis RJB. Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Med*. (2019) 17:137. doi: 10.1186/s12916-019-1377-7
56. Sheikhzadeh E, Eissa S, Ismail A, Zourob M. Diagnostic techniques for COVID-19 and new developments. *Talanta*. (2020) 220:121392. doi: 10.1016/j.talanta.2020.121392
57. Siriwardhana Y, De Alwis C, Gür G, Ylianttila M, Liyanage M. The fight against the COVID-19 pandemic with 5G technologies. *IEEE Eng Manage Rev*. (2020) 48:72–84. doi: 10.1109/EMR.2020.3017451
58. Dzien C, Halder W, Winner H, Lechleitner M. Covid-19 screening: are forehead temperature measurements during cold outdoor temperatures really helpful? *Wiener klinische Wochenschrift*. (2020) 132:1–5. doi: 10.1007/s00508-020-01754-2
59. Quer G, Radin JM, Gadaleta M, Baca-Motes K, Ariniello L, Ramos E, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat Med*. (2020) 27:73–7. doi: 10.1038/s41591-020-1123-x
60. Larsen DA, Wigginton KR. Tracking COVID-19 with wastewater. *Nat Biotechnol*. (2020) 38:1151–3. doi: 10.1038/s41587-020-0690-1
61. Tachibana Y, Segawa N. Physical distance monitoring system for COVID-19 using raspberry Pi and a monocular camera. In: *Proceedings of the 18th Conference on Embedded Networked Sensor Systems (SenSys)*. New York, NY: ACM (2020). p. 772–3. doi: 10.1145/3384419.3430591

Conflict of Interest: BS and DS were employed by the company audEERING GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

PERSPECTIVE
published: 15 April 2021
doi: 10.3389/fdgth.2021.648520

Pre-emptive Innovation Infrastructure for Medical Emergencies: Accelerating Healthcare Innovation in the Wake of a Global Pandemic

Khalil B. Ramadi [1,2,3,4,5]* *and Shriya S. Srinivasan* [3,4,5,6]

[1] Division of Engineering, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, [2] Tandon School of Engineering, New York University, New York, NY, United States, [3] Hacking Medicine, Massachusetts Institute of Technology, Cambridge, MA, United States, [4] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, United States, [5] Division of Gastroenterology, Hepatology and Endoscopy, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States, [6] Society of Fellows, Harvard University, Cambridge, MA, United States

OPEN ACCESS

Edited by:
Phuong N. Pham, Harvard Medical School, United States

Reviewed by:
Wouter A. Keijser, University of Twente, Netherlands
Niamh Lennox-Chhugani, International Foundation for Integrated Care (IFIC), United Kingdom

***Correspondence:**
Khalil B. Ramadi
kramadi@nyu.edu

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *31 December 2020*
Accepted: *17 March 2021*
Published: *15 April 2021*

Citation:
Ramadi KB and Srinivasan SS (2021) Pre-emptive Innovation Infrastructure for Medical Emergencies: Accelerating Healthcare Innovation in the Wake of a Global Pandemic. Front. Digit. Health 3:648520. doi: 10.3389/fdgth.2021.648520

Healthcare innovation is impeded by high costs, the need for diverse skillsets, and complex regulatory processes. The COVID-19 pandemic exposed critical gaps in the current framework, especially those lying at the boundary between cutting-edge academic research and industry-scale manufacturing and production. While many resource-rich geographies were equipped with the required expertise to solve challenges posed by the pandemic, mechanisms to unite the appropriate institutions and scale up, fund, and mobilize solutions at a time-scale relevant to the emergency were lacking. We characterize the orthogonal spatial and temporal axes that dictate innovation. Improving on their limitations, we propose a "pre-emptive innovation infrastructure" incorporating in-house hospital innovation teams, consortia-based assembly of expertise, and novel funding mechanisms to combat future emergencies. By leveraging the strengths of academic, medical, government, and industrial institutions, this framework could improve ongoing innovation and supercharge the infrastructure for healthcare emergencies.

Keywords: innovation infrastructure, translational medical research, health innovation system, pre-emptive innovation, hackathon

MAIN TEXT

The COVID-19 pandemic has emphasized that ongoing innovation in healthcare is critical to meeting arising needs. Healthcare innovation, however, can be slow and costly (1) because of the complexity and risks involved. "Rapid" dissemination of health innovations are defined as those that disseminate in under 5 years (2). Drug development pipelines frequently take longer than a decade (3). While this pace pales in contrast to consumer technologies, it prevents spurious and poorly validated innovations from potentially harming patients. However, during emergency

periods, such as a pandemic, the pace of innovation must be accelerated to decrease mortality and minimize economic impact.

Successful innovation in health care requires three distinct steps: (1) inception, (2) implementation/testing, and (3) dissemination (4). Efforts to spur the invention and inception of innovations, such as hackathons and workshops, have sprouted across industries and achieved significant success (5–8). Testing and implementation have been facilitated through the creation of incubators and accelerators that provide financial, logistical, and legal stability to early-stage companies. Diffusion, the practice of disseminating technology for widespread uptake (2), has been arguably the most underdeveloped step and presents a major bottleneck in the innovation process (2, 9–11), even fostering designation as an "art" (2). By definition, diffusion is a passive process, where adoption varies significantly on individual perception (12). While this represents the natural pathway in a free market, it makes healthcare innovation incredibly costly and challenging, ultimately at the cost of human life. How can we enable active, widespread adoption of innovations, in a more convective fashion? We have glimpsed what such a process could look like in recent approvals of various COVID-19 vaccines and devices, where innovations backed by multiple stakeholders were able to penetrate the market quickly, having rapid, and widespread impact.

Here we conceptualize a spatial-temporal framework (**Figure 1**) to describe the components and processes of innovation. By analyzing points of hand-off and inefficiencies in this framework, we propose a new mechanism to streamline innovation, specifically geared toward emergency preparedness, co-localizing critical components on both axes. Healthcare innovations can be products (e.g. drugs, medical devices), processes, or services. While the complexity and requirements for each differ, their progression can be mapped onto this framework.

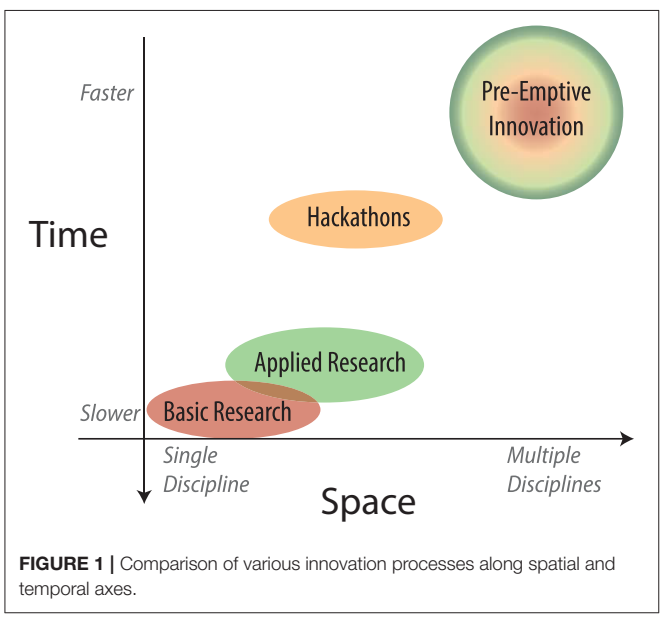


FIGURE 1 | Comparison of various innovation processes along spatial and temporal axes.

SPATIAL AND TEMPORAL AXES OF INNOVATION

The healthcare innovation pathway can be examined from spatial and temporal perspectives. Prior temporal frameworks have highlighted the linear paths of innovation, from idea generation through adoption (4) while spatial frameworks have focused on the role of the adoption system, health system, and broader context (13). We seek to bridge these two frameworks to describe a holistic landscape (**Figure 2A**), incorporating (1) the distinct set of disciplines required for each step, (2) the ease with which different fields can interact, and (3) their respective incentive systems.

Spatial Domains

Academia

Academic institutions have spearheaded multi-disciplinary and translational research through specific funding opportunities, creation of cross-departmental initiatives, and collaborative research centers. Basic science, engineering, and medical research are increasingly bridging fields that have been historically siloed. Multi-disciplinary academic and training programs play a central role in educating individuals to work at multi-disciplinary fields. An example is the Harvard-MIT Health Sciences and Technology (HST) program which integrates both engineering and medical training. At the interface of various disciplines, transformational devices and therapeutics are conceived, addressing challenges that conventionally-isolated disciplines were unable to tackle alone. Academic institutions provide depth in knowledge and skillsets, and generate substantially new ideas by students and faculty that come from a diversity of thought. However, innovation in academia is usually relatively slow. Academics are incentivized by publications, funding, patents, and overall impact. Research grants typically do not provide the funding required to translate and deploy innovations, although certain federal grant programs [e.g. National Institutes of Health (NIH)] have begun to offer this level of support.

Hospitals

Hospitals, and academic medical centers (AMCs) in particular, have incorporated a number of initiatives to enhance the integration of new health innovations (14–16). AMCs help determine the optimal method of integrating new innovations into existing healthcare pipelines and processes (17). AMCs have also developed innovation hubs that combine translational and clinical expertise to implement ideas. AMCs share incentives with traditional academic institutions, in addition to traditional hospital goals of improving patient care and patient experience while lowering overall costs.

Industry

Commercial entities offer expertise in scaling solutions beyond initial proof-of-concept testing. This expertise is vital for the dissemination of innovations. Industry can leverage deep expertise in these fields and significantly accelerate the pace of innovation. However, larger companies can struggle to switch focus between domains. How to manage such changes is the focus

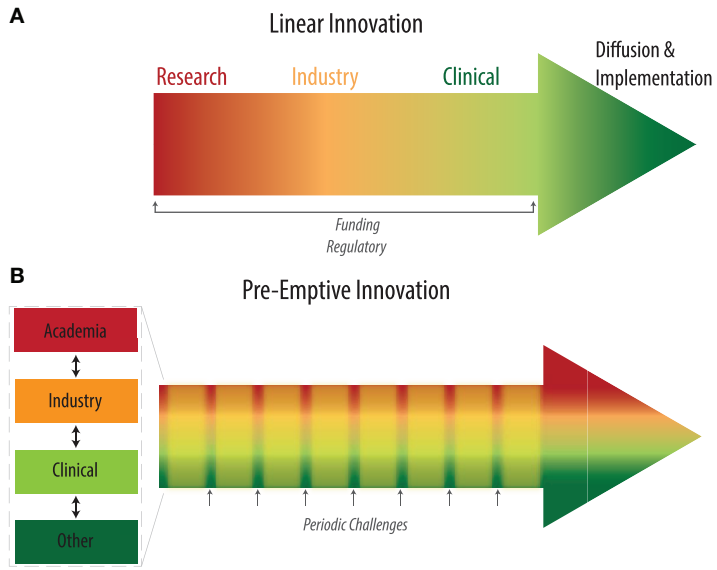


FIGURE 2 | (A) Existing linear pipeline for health innovation. **(B)** Pre-emptive innovation pathway incorporating regular periodic challenges that engage all stakeholders throughout.

of extensive research (18, 19). Involvement of industry early in the innovation process can select for ideas and designs that can be marketed and sustainably manufactured at scale in the long-term. Innovations need to have a clear path to long-term financial sustainability to enable buy-in from for-profit corporations.

Payers

Payers play an important role motivating the development of new innovations, particularly those that improve long-term patient health while simultaneously reducing costs. Payers can be private insurance companies or government entities. Depending on the nature of the innovation, payers may not be directly responsible for covering costs, however the broad dissemination of innovations are likely to be enhanced if there exist insurance reimbursement codes, particularly for digital health technologies. Payer motivations and reimbursement structure largely determine the success of innovations.

Government

The role of government in health innovation can vary significantly depending on the national context. In countries with a centralized, public health system, government acts as the primary customer whose buy-in is crucial for innovation uptake. Where healthcare is privatized, governments traditionally only regulate approval of innovations. In both cases, government can often be a node that bridges different spatial players when specific needs arise. During the COVID-19 pandemic, e.g., the US government actively facilitated collaborations, provided funding and resources through Operation Warp Speed. Chinese and Russian governments actively aided the development of the COVID-19 vaccines.

Temporal Domains

Most medical devices and therapies are conceived and preclinically validated in the academic realm over a period of years. Following the licensing or sale of intellectual property to a commercial entity, development of the product or service may take additional years before final testing via a clinical trial and regulatory approval. Large scale manufacturing and penetration into the market can further take a few years. Even for the most pressing of diseases, a timeline of 5 or more years is considered a reasonable time-to-market.

Innovation Challenges and Hackathons

Innovation challenges seek to expedite innovation by crowdsourcing solutions to problems (5–8, 20). Proposed problems can be open-ended (e.g. diabetes treatment) or have a specific focus of interest (e.g. multivitamin delivery in rural towns). Innovation challenges can occur over the course of a few days to a few months. Events generally incentivize the creation of teams that dedicate time and effort to developing ideas that may address existing problems in the sector. Hackathons are a specific type of innovation challenge where participants and teams are assembled and work together in a dedicated space for a prescribed length of time (21). This creates an environment where teams can enjoy protected time for focused, collaborative work. Innovation challenges also provide mentoring from experts across sectors, in order to answer questions teams may have and guide them as they refine and develop their ideas.

Incubators and Accelerators

Incubators and accelerators accept teams with ideas in various stages of development and guide them toward testing and implementation of their ideas, including pilot studies or clinical trials. These programs can provide funding, expertise, office space, and access to partners to facilitate testing. Incubators

can exist as stand-alone entities or be housed within academic institutions. AMCs specifically have increasingly developed in-house incubator-like programs in an effort to cultivate internal innovation capacity (22, 23). Teams emerging from these programs often spin-out as commercial for-profit entities, raising funding from angel and venture capital sources, and successfully marketing and selling their products or services.

Each phase in the temporal domain is characterized by specific expertise. Basic science and engineering are required for invention and inception. Translational and clinical expertise are necessary to integrate and test innovations in specific clinical contexts. Business and manufacturing knowhow are crucial for scaling and sustainability. Temporally, the timeframes for each step can vary and depend, in part, on how rapidly different disciplines can be onboarded and convalesced into a multi-disciplinary team.

Barriers to Diffusion of Health Innovations

Even the most successful health innovations face a steep battle in their diffusion and uptake. Previous studies have identified a number of barriers to health innovation diffusion including organizational structure, partnerships, lack of dedicated resources, and inadequate context (2, 9, 11, 12, 24–26). Proof of efficacy of an innovation does not guarantee its uptake. Rather, incentives for each active stakeholder need to be present to promote use. Rapidly-scaled innovations align these incentives (1) in an efficient manner.

A related challenge for organizations to incorporate innovations is to create an environment conducive to change management (10). Organizations need to create a climate for change and engage the whole organization in order to implement and sustain changes (18, 19). In healthcare, this challenge is amplified by the number of stakeholders involved.

The current paradigm of health innovations is much like a relay race, where each stakeholder takes the lead for one component of the translation process before handing it off to the next stakeholder. Each hand-off must contend with communication across disciplines that may each speak a different language. There also exists substantial inertia in each segment, given that innovation is often no one person's primary professional obligation, but rather a peripheral facet to their job (27).

In the face of a healthcare emergency such as the COVID-19 pandemic, the challenges enumerated above and the breadth of the spatial and temporal axes hinder innovation from being realized at a time-scale relevant to the situation.

Pre-emptive Innovation Infrastructure for Medical Emergencies

We propose the concept of a pre-emptive innovation infrastructure for medical emergencies (PRIME) which proactively builds a sturdy cross-disciplinary ecosystem that is primed to spring to action in times of crisis. The PRIME would consist of teams in each geography that co-localize the various spatial domains and parallelize the temporal domains of innovation (**Figure 2B**). For each instance, state governments could identify key university, private-industry, and hospital labs in major academic cities to serve as the R&D hubs. Large companies with local manufacturing and/or prototyping spaces would be partnered with these labs for rapid prototyping, design, and development. Regulatory officials would be regularly involved to familiarize themselves to the community and articulate necessary testing requirements proactively. Within the team, technical, business, legal, clinical and regulatory leads would be appointed. As a pre-selected entity, the team would have the impetus to work together and move swiftly when called upon in times of crisis.

The PRIME is motivated by observations of how various players were able to ramp up innovation efforts during the COVID-19 pandemic. A shared sense of urgency and purpose during an emergency enabled businesses, institutions, and governments to deploy emergency protocols that accelerated the pace of work, regulatory and committee approvals, institutional review boards (IRBs) and other bureaucratic steps. It is also modeled after the multidisciplinary teams bridging academia, industry, and regulatory bodies assembled regularly by the Defense Advanced Projects Research Agency (DARPA).

The PRIME could be assigned predetermined pathways for testing, implementation, regulatory approval, scale-up, marketing, and funding to facilitate rapid implementation. Ownership, protection, and sharing of intellectual property would follow pre-templated models, based on the type of device or therapy, to prevent delays in innovation. During a medical emergency, directors of the PRIME would designate the major challenges, unlock government funding for R&D and facilitate advance purchase orders (APCs) to incentivize the manufacturing and deployment.

Challenge-Based Preparation

In non-crisis periods, PRIME teams would engage in regular innovation drills that challenge the networks to solve existing healthcare problems with transformational, as opposed to incremental innovations, similar to the Gates Foundation challenges. These could be funded through government grants and APCs, creating a market for qualifying solutions.

Similar to regular exercises carried out by the armed forces to ensure optimal preparation, these challenges would prepare all stakeholders for emergency scenarios. In addition to bolstering emergency preparedness, this infrastructure would establish the capacity for regular, constant cross-disciplinary collaboration and translation in healthcare, solving many of the large healthcare problems that we currently face. The PRIME allows professionals within each vertical to operate in such a way that meets their conventional incentive structures while simultaneously contributing to a broader, impact-driven schema.

A Climate of Change

PRIME tackles barriers identified by previous studies that impede health innovation (28, 29). Firstly, it fosters a culture of mutual understanding and co-creation across industries that may not be accustomed to regular collaboration, facilitating innovation flow (28, 29). This framework would also actively involve workers from each industry across levels and departments, creating a climate of change (30).

Each pathway of innovation espouses certain advantages and limitations. Ironically, a pathway that is too rigid can actively impede innovations that arise from nonconventional sources (22). Thus, we intentionally avoid prescribing a predetermined method of collaboration in the pre-emptive innovation infrastructure and encourage the creation of "slack," whereby innovations from a variety of sources are nourished. Such slack can be created, e.g., by developing an open pathway whereby employees across industries can highlight problems and propose possible solutions. Employees could then be afforded some protected time from their primary responsibilities to pursue this idea, with appropriate guidance. This would also work toward changing organization structure to one of change for continuous quality improvement (31, 32).

The consolidated framework for implementation research (CFIR) has previously outlined various characteristics to facilitate implementation science. The weakest determinant was identified as the outer setting, external relationships across institutional and broader context in which innovations need to diffuse. PRIME would address this weakness directly (33).

Previous studies have identified that successful innovations require internal champions at every step of the innovation timeline to move them forward (34). A PRIME would take this one step further and obviate the need for individuals to move against the grain. Institutional frameworks would facilitate rather than impede innovative projects, while government and payer stakeholders would provide advocacy for emerging solutions. A sustainable health innovation ecosystem is a mechanism by which institutional entrepreneurship can be created across multiple players simultaneously (35). An established collaborative infrastructure would also facilitate data sharing and implementation, making diffusion and uptake of technologies less reliant on individuals and more intentionally integral to continuous healthcare improvement.

The race to develop devices and vaccines to address the COVID-19 pandemic has exemplified the potential of establishing sustained collaboration across stakeholders in healthcare for rapid innovation. The framework proposed here would be conducive to various types of innovations including digital innovations that have proven important in battling the ongoing COVID-19 pandemic (36). A PRIME would not only bolster emergency preparedness, but also catalyze ongoing innovation. Looking to the future, this infrastructure could also be applied to other impending crises including climate change, water scarcity, energy crises, and food security.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

KR and SS jointly conceived the ideas presented here and wrote the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Herzlinger RE. *Harvard Business Review*. Boston, MA: Harvard Business Publishing (2006).
2. Parston G, McQueen J, Patel H, Keown OP, Fontana G, Al Kuwari H, et al. The science and art of delivery: accelerating the diffusion of health care innovation. *Health Affairs.* (2015) 34:2160–6. doi: 10.1377/hlthaff.2015.0406
3. Cummings J, Lee G, Ritter A, Sabbagh M, Zhong. Alzheimer's disease drug development pipeline: 2020. *Alzheimer's Demen.* (2020) 6:e12050. doi: 10.1002/trc2.12050
4. Varkey P, Horne, A, Bennet KE. Innovation in health care: a primer. *Am J Med Qual.* (2008) 23:382–388. doi: 10.1177/1062860608317695
5. Ramadi KB, Srinivasan S, Atun R. Health diplomacy through health entrepreneurship: using hackathons to address Palestinian-Israeli health concerns. *BMJ Global Health.* (2019) 4:e001548. doi: 10.1136/bmjgh-2019-001548
6. Srinivasan S, Ramadi KB, Ippolito, A, Atun, R. Democratizing innovation through grass-roots entrepreneurship: lessons from efforts to address the opioid epidemic in the United States. *BMJ Global Health.* (2019) 4:e002079. doi: 10.1136/bmjgh-2019-002079
7. Olson KR, Walsh M, Garg P, Steel A, Mehta S, Data S, et al. Health hackathons: theatre or substance? A survey assessment of outcomes from healthcare-focused hackathons in three countries. *BMJ Innovations.* (2017) 3:37–44. doi: 10.1136/bmjinnov-2016-000147
8. Celi LA, Ippolito A, Montgomery RA, Moses, C, Stone. Crowdsourcing knowledge discovery and innovations in medicine. *J Med Internet Res.* (2014) 16:e216. doi: 10.2196/jmir.3761
9. Barnett J, Vasileiou K, Djemil F, Brooks, L, Young. Understanding innovators' experiences of barriers and facilitators in implementation and diffusion of healthcare service innovations: a qualitative study. *BMC Health Services Research.* (2011) 11:342. doi: 10.1186/1472-6963-11-342
10. Macfarlane F, Barton-Sweeney C, Woodard, F, Greenhalgh, T. Achieving and sustaining profound institutional change in healthcare: case study using neo-institutional theory. *Soc Sci Med.* (2013) 80:10–18. doi: 10.1016/j.socscimed.2013.01.005
11. Denis JL, Hébert Y, Langley A, Lozeau, D, Trottier. H. *Explaining diffusion patterns for complex health care innovations*. *Health Care Manage Rev.* (2002) 27:60–73. doi: 10.1097/00004010-200207000-00007
12. Safi S, Thiessen, T, Schmailzl KJ. Acceptance and resistance of new digital technologies in medicine: qualitative study. *JMIR Res Protoc.* (2018) 7:e11072–2. doi: 10.2196/11072
13. Atun R, de Jongh T, Secci F, Ohiri, K, Adeyi. Integration of targeted health interventions into health systems: a conceptual framework for analysis. *Health Policy Plan.* (2010) 25:104–11. doi: 10.1093/heapol/czp055
14. Ellner AL, Stout S, Sullivan EE, Griffiths EP, Mountjoy A, Phillips RS. Health systems innovation at academic health centers: leading in a new era of health care delivery. *Acad Med.* (2015) 90:872–80. doi: 10.1097/ACM.0000000000000679
15. Dzau VJ, Yoediono Z, Ellaissi WF, Cho AH. Fostering innovation in medicine and health care: what must academic health centers do? *Acad Med.* (2013) 88:1424–9. doi: 10.1097/ACM.0b013e3182a32fc2
16. DePasse JW, Chen CE, Sawyer A, Jethwani, K, Sim. Academic Medical Centers as digital health catalysts. *Healthc (Amst).* (2014) 2:173–6. doi: 10.1016/j.hjdsi.2014.05.006
17. Desveaux L, Kelley LT, Bhatia RS, Jamieson T. Catalyzing digital health innovation in ontario: the role of an academic medical centre. *Healthc Policy.* (2020) 16:55–68. doi: 10.12927/hcpol.2020.26353

18. Campbell RJ. Change management in health care. *Health Care Manag (Frederick).* (2008) 27:23–39. doi: 10.1097/01.HCM.0000285028.79762.a1
19. Auguste J. Applying kotters 8-step process for leading change to the digitaltransformation of an orthopedic surgical practice group in Toronto,Canada. *J Health Med Inform.* (2013) 3:1–4. doi: 10.4172/2157-7420.1000129
20. Ramadi KB, Nguyen FT. Rapid crowdsourced innovation for COVID-19 response and economic growth. *NPJ Digit Med.* (2021) 4:18. doi: 10.1038/s41746-021-00397-5
21. Gubin TA, Iyer HP, Liew SN, Sarma A, Revelos A, Ribas J, et al. A systems approach to healthcare innovation using the mit hacking medicine model. *Cell Syst.* (2017) 5:6–10. doi: 10.1016/j.cels.2017.02.012
22. Srimathveeravalli G, Balesh E, Cheng C P, Chen D. If you build it, they will come: how to establish an academic innovation enterprise. *Tech Vasc Interv Radiol.* (2017) 20:121–6. doi: 10.1053/j.tvir.2017.04.005
23. Mann DM, Kuppin Chokshi S, Lebwohl R, Mainiero M, Dinh-Le C, Driscoll K, et al. Building digital innovation capacity at a large academic medical center. *NPJ Digit Med.* (2019) 2:13. doi: 10.1038/s41746-019-0088-y
24. Keown OP, Parston G, Patel H, Rennie F, Saoud F, Al Kuwari H, et al. Lessons from eight countries on diffusing innovation in health care. *Health Aff (Millwood).* (2014) 33:1516–22. doi: 10.1377/hlthaff.2014.0382
25. Dixon-Woods M, McNicol, S, Martin, G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Quality Safety.* (2012) 21:876–84. doi: 10.1136/bmjqs-2011-000760
26. Berwick DM. Disseminating innovations in health care. *JAMA.* (2003) 289:1969–75. doi: 10.1001/jama.289.15.1969
27. Ostrovsky, A, Barnett, M. Accelerating change: fostering innovation in healthcare delivery at academic medical centers. *Healthc (Amst).* (2014) 2:9–13. doi: 10.1016/j.hjdsi.2013.12.001
28. Dhainaut JF, Blin O, Herry F, Bilbault P, Cauterman M, Favrel-Feuillade F, et al. Health research and innovation: can we optimize the interface between startups/pharmaceutical companies and academic health care institutions or not? *Therapie.* (2020) 75:113–23. doi: 10.1016/j.therap.2019.11.010
29. Ii SS, Fitzgerald L, Morys-Carter MM, Davie N L, Barker R. Knowledge translation in tri-sectoral collaborations: an exploration of perceptions of academia, industry and healthcare collaborations in innovation adoption. *Health Policy.* (2018) 122:175–83. doi: 10.1016/j.healthpol.2017.11.010
30. Kottler JP. *Leading Change*. Harvard Business School Press (1996).
31. McAlearney AS, Terris D, Hardacre J, Spurgeon P, Brown C, Baumgart A, et al. Organizational coherence in health care organizations: conceptual guidance to facilitate quality improvement and organizational change. *Qual Manag Health Care.* (2013) 22:86–99. doi: 10.1097/QMH.0b013e31828bc37d
32. Gagliardi, P. The creation and change of organizational cultures: a conceptual framework. *Organization Studies.* (1986) 7:117–34. doi: 10.1177/017084068600700203
33. Leeman J, Baquero B, Bender M, Choy-Brown M, Ko LK, Nilsen P, et al. Advancing the use of organization theory in implementation science. *Prev Med.* (2019) 129s:105832. doi: 10.1016/j.ypmed.2019.105832
34. McGloughlin EK, Anglim P, Keogh, I, Sharif, F. Innovation for the future of Irish MedTech industry: retrospective qualitative review of impact of BioInnovate Ireland's clinical fellows. *BMJ Innov.* (2018) 4:32–38. doi: 10.1136/bmjinnov-2016-000184
35. Melder A, Burns P, Mcloughlin, I, Teede, H. Examining 'institutional entrepreneurship' in healthcare redesign and improvement through comparative case study research: a study protocol. *BMJ Open.* (2018) 8:e020807. doi: 10.1136/bmjopen-2017-020807
36. Budd J, Miller BS, Manning EM, Lampos E, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. *Nat Med.* (2020) 26:1183–92. doi: 10.1038/s41591-020-1011-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 10 June 2021
doi: 10.3389/fdgth.2021.643042

A Real-Time Portable IoT System for Telework Tracking

*Yongxin Zhang[1], Zheng Chen[1], Haoyu Tian[1], Koshiro Kido[1], Naoaki Ono[1,2], Wei Chen[3], Toshiyo Tamura[4], M. D. Altaf-Ul-Amin[1], Shigehiko Kanaya[1,2] and Ming Huang[1]**

[1] Computational Systems Biology, Division of Information Science, Nara Institute of Science and Technology, Nara, Japan, [2] Data Center, Nara Institute of Science and Technology, Nara, Japan, [3] Department of Electronic Engineering, Center for Intelligent Medical Electronics, School of Information Science and Technology, Fudan University, Shanghai, China, [4] Institute for Healthcare Robotics, Waseda University, Tokyo, Japan

Telework has become a universal working style under the background of COVID-19. With the increased time of working at home, problems, such as lack of physical activities and prolonged sedentary behavior become more prominent. In this situation, a self-managing working pattern regulation may be the most practical way to maintain worker's well-being. To this end, this paper validated the idea of using an Internet of Things (IoT) system (a smartphone and the accompanying smartwatch) to monitor the working status in real-time so as to record the working pattern and nudge the user to have a behavior change. By using the accelerometer and gyroscope enclosed in the smartwatch worn on the right wrist, nine-channel data streams of the two sensors were sent to the paired smartphone for data preprocessing, and action recognition in real time. By considering the cooperativity and orthogonality of the data streams, a shallow convolutional neural network (CNN) model was constructed to recognize the working status from a common working routine. As preliminary research, the results of the CNN model show accurate performance [5-fold cross-validation: 0.97 recall and 0.98 precision; leave-one-out validation: 0.95 recall and 0.94 precision; (support vector machine (SVM): 0.89 recall and 0.90 precision; random forest: 0.95 recall and 0.93 precision)] for the recognition of working status, suggesting the feasibility of this fully online method. Although further validation in a more realistic working scenario should be conducted for this method, this proof-of-concept study clarifies the prospect of a user-friendly online working tracking system. With a tailored working pattern guidance, this method is expected to contribute to the workers' wellness not only during the COVID-19 pandemic but also take effect in the post-COVID-19 era.

Keywords: telework, nudge, wearable, realtime tracking, convolutional neural network

OPEN ACCESS

Edited by:
Susanna Spinsante,
Marche Polytechnic University, Italy

Reviewed by:
Alpha Agape Gopalai,
Monash University Malaysia, Malaysia
Ian Cleland,
Ulster University, United Kingdom

***Correspondence:**
Ming Huang
alex-mhuang@is.naist.jp

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *17 December 2020*
Accepted: *16 April 2021*
Published: *10 June 2021*

Citation:
Zhang Y, Chen Z, Tian H, Kido K, Ono N, Chen W, Tamura T, Altaf-Ul-Amin MD, Kanaya S and Huang M (2021) A Real-Time Portable IoT System for Telework Tracking. Front. Digit. Health 3:643042. doi: 10.3389/fdgth.2021.643042

1. INTRODUCTION

A great leap of digital healthcare is expected under the background of COVID-19, which has changed the style of socioeconomic organization. Now, telework has been being adopted pervasively to decrease the transmission risk so as to contain the pandemic. On the other side of the strengthened working efficiency by saving the commuting time, the side effects caused by prolonged working hours become more visible.

Given that the adverse effect of prolonged working hours and the consequent stress have been identified (1, 2), attempts that chopped up the working time with breaks have been tried and

validated (3, 4). Additionally, the sit-stand working pattern has been proposed recently as a new working fashion to reduce the sitting time and has shown the potential effectiveness (5). However, the change of posture may not be able to eliminate the adverse effect of prolonged working hours. Given the difficulty in keeping up to an intervention schedule by the worker himself, a fully automatic intervention system, which can recognize the working status in real-time and then generate the behavior-change notification automatically, is valuable. To meet the requirements in simultaneousness and interactiveness, the system should (1) be able to recognize the motion by extracting the information from video or time-series sensor signal in real time, and (2) be able to feedback to the subject in a convenient way.

The first requirement is widely studied in the topic of human activity recognition (HAR) by using the camera, ambient sensor, and wearable sensors (6, 7). The first two modalities are not suitable for the recognition of working status due to the higher computation cost and the rigid requirement in the system setup. Numerous previous research utilizing wearable sensor fusion in HAR show the clear prospect of accurate HAR recognition using portable sensors (8–10). Nevertheless, the aforementioned two requirements make it more difficult for most of the methods to be applied.

The smartphones and the accompanying smartwatch ecosystem, which is one of the major modalities of IoT devices, provides an excellent platform to implement the HAR in terms of performance and availability. The research of using a smartphone as the hub of data-stream acquiring, processing, and modeling have emerged in recent years. Cao et al. have tried to conceptualize the smartphone-based implementation by optimizing the number and types of features and validated the algorithm by using an open dataset (11). Cvetkovic et al. tried to fuse the accelerometer signal from a smartphone and a wristband to recognize the daily activities, which are majorly locomotions and achieved an 87% average accuracy (12). Bianchi et al. proposed a system consists of sensing and data transmitting via a wearable gadget and the accompanying HAR recognition in the form of cloud service based on the deep learning model (13). In the IoT context, the user's loyalty is an unavoidable factor in system design, which suggests that an established ecosystem, such as the iOS and Android system, with a massive user base, would be ideal. Moreover, it is natural for these commercial systems to provide a direct interaction via smartphone or smartwatch devices. With the new dedicated neural network accelerator being added in, the recent smartphones are becoming the appropriate hub for edge computing as research concerning using the signals of IMU sensor of a smartwatch can be seen lately (14, 15).

Although the monitoring of working status did not become a practical need until last year when remote working become an elementary style and few researchers contributed to the topic, human activity recognition based on wearable devices, which may be applicable, have been developed. Mannini et al. have tried to use the wrist-worn and ankle-worn accelerometer to identify the locomotions (walking, cycling, and resting) (16). Specially, as it can be imaged for a real working situation, a few relevant elementary actions, such as reading and typing, as well as irrelevant interruptions are mixed in, a coarse information, such as the record of screen time and sedentary time may cause overestimate or underestimate of working time. Therefore, Kwon and Choi have tried to construct a pipeline based on smartwatch and artificial neural network model to recognize the working relevant activities from other daily living activities (17). In considering the requirement on the accuracy and the restriction of privacy protection, the activity recognition based on wearable sensors is considered to be an appropriate solution, in comparison with the video-based, which may cause concern of the privacy disclosure and the self-report-based methods, which is prone to be coarse and data missing.

However, in the previous studies, the influence of individual difference is unclear and the overall method including signal source and the classification model can be optimized. In view of the results of the previous researches above, in this research, we extend the signal source by adding the signal of gyroscope and comparing the convolutional neural network (CNN) classification model with random forest (RF) and support vector machine (SVM) in more rigorous off-line [5-fold and leave-one-out (LOO)] and real-time experiments to evaluate the feasibility of working recognition with standalone pair of smartphone and smartwatch.

2. MATERIALS AND METHODS

2.1. Data Acquisition

The data used in this research were generated by the accelerometer and gyroscopes enclosed in an Apple Watch (series 5). The coreMotion of Apple provides the application programing interface (APIs) for the developer to access the data generated by the sensors in Apple watch and iPhone, which can be extracted for off-line use or direct use in a dedicated app. Both the raw data and the processed values can be extracted by the APIs provided by coreMotion in a user-defined interval. The sampling rate is 50 Hz in this research, which is higher than 20 Hz, and it is often used in human activity recognition, because most of the target activities requires a minimal of 30 Hz to prevent aliasing (18, 19). In this study, the data streams being used in this study are tabulated in **Table 1**.

The data streams will then be used by the coreML, which is a framework for machine learning provided by Apple. In recent models (iPhone11 and later), the computation performance is greatly boosted by the dedicated hardware—the neural engine. The machine learning model can be trained separately by python and then converted back by the coremltools to the coreML format to run on the iPhone. In this study, we extracted the data streams to train the model first, and then put the converted model back into the iPhone.

2.2. Experiments

In this study, we simulated the activity pattern at-home by performing the combinations of locomotion and limbs movements, which are described in **Table 2**. Two experiments were carried out subsequently. The first offline experiment, which also served the purpose of initial data collection, was conducted

on 12 subjects (college students, 10 males, two females; ages: [23–34] years old), whose willingness in participating in the experiments was confirmed by written informed consent, with an Apple watch worn on their right wrists. The subjects were guided to conduct the prescribed locomotions and limb movements at their own paces. The timing for each action was recorded by an assistant to ensure the correctness of class labeling. In every 2 s, statistical features which is described in section 2.3 of the nine-channel signals sensed by the watch simultaneously were calculated on site and sent to the paired iPhone11 via Bluetooth. In the experiment protocol, the working-while-standing style was added to the conventional sitting style given that more and more people use at least partially the standing position in working time. The actions of reading, keyboard typing, and writing were attributed as working status. Therefore, given the similarity of the nine-channel signals, we combined limb movements of reading, typing, and writing in both standing and sitting status, which results in three classes of combinations for working status. Consequently, there are seven classes of locomotion-limbs-movements combinations to be recognized in this research as shown in the [Label] row of **Table 2**.

TABLE 1 | The nine channels of data streams used in this research, which can be accessed by the APIs of CoreMotion.

Data	Description
Motionuseraccelerationx (G) Motionuseraccelerationy (G) Motionuseraccelerationz (G)	The acceleration that the user is giving to the device along the corresponding axis. The total acceleration of the device is equal to gravity plus the acceleration the user impacts to the device
Motionyaw (rad)	Angular rotation around an axis that runs vertically through the device.
Motionroll (rad)	Angular rotation around a longitudinal axis that passes through the the device from its top to Bottom
Motionpitch (rad)	Angular rotation around a lateral axis that passes through the device from side to side
Motionrotationratex (rad/s) Motionrotationratey (rad/s) Motionrotationratez (rad/s)	Rotation rate along the corresponding axis

Finally, a real-time validation, which used the the pre-trained model to predict the simultaneous activity, was conducted. Two subjects, who have participated in the off-line experiment, were invited to repeat the experiment in order to test the repeatability of the model; another three subjects who did not participate in the offline experiment (males, 24–28 years old) were invited to perform the actions of working status for 30 min, during which time they could stop working or drinking water freely.

2.3. Features Extraction

Regarding the offline experiment, the nine-channel signals were stored in the iPhone during the experiment. Whereafter, signals were extracted from iPhone as CSV files with time stamps and used to train the classification model offline. Although the model can be trained directly in iPhone via Xcode, Python was used in the training period given the higher flexibility and the abundance of packages.

The preprocessing is closely related to the construction of the CNN models. Although the deep ANN is adequate in automatic feature generation, given the relatively small dataset and the effectiveness of the engineered feature extraction in reducing the input dimension and model complexity (19, 20), it is assumed that noise in a short interval (2 s) is normally distributed so that the lower order statistics features (mean and standard deviation) are sufficient in describing the short segment without filtering and the low order statistics. On the other hand, the mean and standard deviation have been proved to be informative features in activity recognition problems (19, 21). Therefore, the mean and standard deviation of the nine-channel signals were extracted from the right wrist and used as the input to the models.

During the experiment, it was noticed that the inter-subject difference in performing the writing, reading, and typing mainly resides in the pace. Therefore, a time length, which is set as 2 s, that can probably include one bout of the action is desired. This size of the window yields highest accuracy in human activity recognition, even though it includes fewer cycles of an action (18). In this logic, the means and standard deviation of a short interval (2 s) are the unit input to the CNN models.

Moreover, although it is used by default that the data are fragmented into short intervals and input into machine learning model for classification, the automatic segmentation may occur in the same semantic event (the same action), which would then worsen both the training and predicting period, especially for a

TABLE 2 | The protocol of the experiment (upper) and the samples number for each class (lower).

Episode	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Locomotion	Walking	Standing	Standing	Standing	Standing	Sitting	Sitting	Sitting	Sitting	Lying
Limb movements	Neutral	Neutral	Reading	Typing	Writing	Neutral	Reading	Typing	Writing	Neutral
Duration (min)	10	5	3	3	3	5	3	3	3	10
Label	W_N	St_N	Re	Ty	Wr	Si_N	Re	Ty	Wr	L_N
Class	W_N	St_N	Re	Ty	Wr	Si_N	L_N			
#Sample	892	429	541	539	546	444	900			
Duration (min)	119	57	72	72	73	60	120			

relatively short 2-s window. In view of this concern, we tried to mitigate this problem by using a longer time interval of the data while keeping the basic interval as 2 s. The feature vector of each basic interval is then arranged into a row and stacked up to a feature matrix $\boldsymbol{M} \in \Re^{m \times n}$, where m represents $m \times 2$ s data and n represents n features.

In addition to the mean and standard deviation, we assumed that the cooperativity of the nine channels is informative in describing the differences between the seven classes. Eventually, the features of the nine channels can be extracted in the following way:

- first, the pair-wise Pearson correlation of the nine-channel signals were calculated, which resulted in a symmetrical 9×9 matrix $\boldsymbol{R}^i$, where the superscript represents the ith 2s;
- second, the maximum eigenvector $\boldsymbol{\varepsilon}^i (\in \Re^{1\times 9})$ of the $\boldsymbol{R}^i$ is then calculated and extracted as the features of cooperativity;
- third, the means $\boldsymbol{\mu}^i = \mu_1^i, \mu_2^i, \cdots, \mu_9^i$ and standard deviations $\boldsymbol{\sigma}^i = \sigma_1^i, \sigma_2^i, \cdots, \sigma_9^i$ are calculated for each channel and concatenated to $\boldsymbol{\varepsilon}^i$ to form the feature vector $\boldsymbol{v}^i \in \Re^{1\times 27}$;
- finally, the feature matrix for a longer interval is formed by stacking the feature vectors row by row to form the matrix M as the input of the CNN model.

The m and n are the hyperparameters, and the m is set as 8 s, an interval that includes most of the underlying hand movements. Regarding the n, the ($\boldsymbol{\varepsilon}$, $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) combination and the ($\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) combination were tried in the model building stage. The shape of the input matrix is illustrated in **Figure 1**.

With regard to the real-time test, the nine-channels signals were sampled at 50 Hz via the **CoreMotion** framework and then were segmented into short episodes of 2-s length for further features generation by using the vDSP module in the **Accelerate** framework. On the other side, the iPhone used the Passive Response Delegate Mechanism to receive and reorganize the consecutive four 2-s feature vectors, $\boldsymbol{f} \in \Re^n$, into an 8-s feature matrix $\boldsymbol{M}$, which were then fed into the pre-trained model.

2.4. Classification Model

The classification model was built with a 2D CNN model with one to three hidden layers (**Figure 2**) in light of the finding of Münzner et al., which shows that a sallow CNN model with 2–3 CNN layers is better than the random forest model (22). The simplicity of the model also benefits the edge computation by using the iPhone at one's disposal, where the computation latency and the power consumption should be taken into account seriously. To minimize the negative influence of the deep learning model on the general performance of the smartphone, developers tend to restrict the complexity of the model (23, 24). A simpler model may also be beneficial to its generalizability and robustness when being applied to a new dataset. **Figure 2** shows the structure of the CNN model, where the feature matrix $\boldsymbol{M}$ was provided as input.

As it is shown in **Figure 2**, in the first CNN layer, we paid special attention to the kernel size and strides, because we thought that a double length in unit interval (2 s) will be helpful in extracting the latent features of each action. This consideration

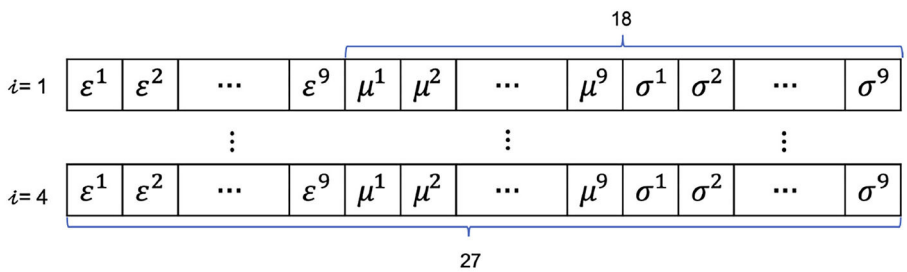


FIGURE 1 | The features and the shape of model input, i represents the ith 2 s, and a sample of input consists of 8 s. The ($\boldsymbol{\varepsilon}$, $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) and ($\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) features combinations are represented by the structure of 27 and 18, respectively.

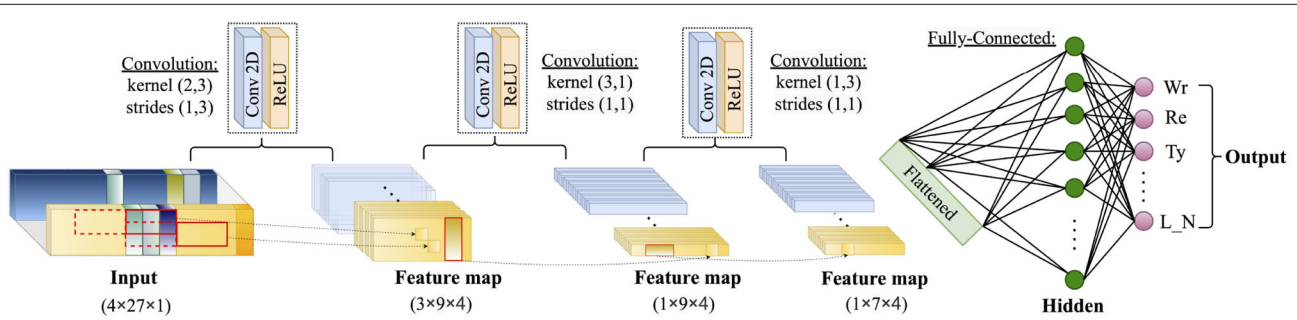


FIGURE 2 | The structure of the convolutional neural network (CNN) models. The 8-s nine-channel data streams were transformed into a 4×27 (or 4×18, not shown in the figure) 2D matrix to input into the network. There are three variants, where CNN(1) uses the first convolutional layer, CNN(2) uses the first two convolutional layers, and CNN(3) uses all the three convolutional layers.

also applied in the tuning of the second CNN layer, where the whole time length of the input to the second CNN layer was covered by the filters. The resultant feature map of the second CNN layer is further abstracted by the third CNN layer, which combines the features of three channels. Based on this structure, three variants were tested in the 5-fold and LOO validations, where CNN(1) corresponds to the model with the first CNN layer, CNN(2) corresponds to the model with the first two CNN layers, and CNN(3) corresponds to the model with all the three CNN layers.

With the development in machine learning, a number of classification models have been used for HAR, among which the SVM and random forest were proven to be suitable in this field (22, 25, 26). To confirm the performance of the CNN models, the SVM classifier with radial basis function (RBF) kernel (regularization parameter = 1.5) and the random forest classifier (number of trees = 200, maximum tree depth = 10), which are supported by CoreML, have been compared with the CNN models in the 5-fold and LOO validations.

2.5. Evaluation

Three kinds of evaluations were conducted to examine the performance of the overall pipeline. First, to examine the performance of the overall method, 5-fold cross-validation was used. The overall accuracy, F1 score and the Matthews correlation coefficient (MCC) is used to evaluate the overall performance of the model. The MCC metric is essentially a correlation coefficient, which is more suitable in evaluating the imbalanced dataset by taking the proportion of each class into consideration. The definition of MCC for the binary situation is:

$$\mathrm{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}, \quad (1)$$

and the MCC calculation can be easily extended to the multiclass situation. Moreover, the results for working status classification are specifically analyzed with class-wise metrics (recall, precision).

Second, because a portion of the samples from the same subject would be used in the training stage in the 5-fold validation, a validation that can evaluate the influence of the individual difference on the model is beneficial. Since the leave-one-out validation, which is more rigorous for the model generalizability by excluding the whole dataset of a subject from the training dataset and using it as the test dataset (18), it was implemented for all the 12 subjects. From this validation, we expected to determine the best model in terms of generalizability.

Finally, the real-time test, which is described in section 2.2, was finally conducted to validate the reproducibility and generalizability of the model by installing the best model into the prototype application. For the two subjects who have participated in the off-line experiment, a confusion matrix is used to show the results, whereas for the three subjects who were new to the experiment, individual compositional bar chart is used to show the detail of classification results.

3. RESULTS

By using the coreMotion API, the signals of the nine channels can be extracted at a constant frequency (50 Hz). Although slight fluctuation can be seen in the sampling interval from time to time, the fluctuations were all <10 ms. Measurements were successfully extracted from all the 12 subjects, and by separating the samples for 8 s, the sample numbers of the seven classes can be found in **Table 2** (lower part).

By plotting the averaged 8-s matrices (**Figure 3**), the differences of the seven classes can be visualized. From **Figure 3**, the inter-classes difference can be confirmed. Specifically, the three classes that belong to the working status are similar to each other, while the walking is a distinct one being different from the others. Moreover, it can be seen that the *St_N* is similar to the *Ty*, which causes the occasional misclassification between these two classes in the 5-fold cross-validation and the leave-one-out validation.

3.1. Results of 5-Fold Cross-Validation

Models that used 27 features ($\boldsymbol{\varepsilon}$, $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) combination and 18 features ($\boldsymbol{\mu}$, $\boldsymbol{\sigma}$) combination are summarized in **Figure 4**, from which it can be confirmed that the model input with 27 features combination with three CNN layers has the best performance for the working status recognition (recall: 0.965, precision: 0.967), whose confusion matrix can be seen in the right matrix of **Figure 5**, while the counterpart with 18 features combination (left matrix of **Figure 5**) has similar results (recall: 0.964, precision: 0.965).

Looking closer into the confusion matrices of the CNN(3) models with two combination of features (**Figure 5**), the fluctuation in the predicting accuracy can be seen. **Figure 5** shows the averaged confusion matrices of the models, from which it can be confirmed that the three actions of working status (*Re*, *Ty*, and *Wr*) can be recognized accurately, whereas the *St_N* is somewhat difficult to be separated from the *L_N*, which may be caused by the similarity in free movements of torso and limbs in these two actions. Within the three actions of working status, the Ty may be recognized as *St_N* occasionally, which seems plausible because the vibration of the torso while standing may also be reflected by the sensors on the right wrist.

3.2. Results of Leave-One-Out Validation

Similar to the 5-fold validation, models with two feature combinations were compared here. The results are summarized in **Figure 6**. In this validation, the RF model is the best for the overall classification, but not for the working status. The model with two CNN layer inputs with 18 features combination has the best performance (recall: 0.951, precision: 0.944), while the RF model has a similar performance (recall: 0.947, precision: 0.926). Please note that the recall and precision values in this paragraph are the averaged values for the three working activities.

The effect of different feature combinations is compared by examining 2-layers CNN model with 18 and 27 features (**Figure 7**). The 27-feature model outperforms the counterpart in identifying the *L_N*, but is inferior in identifying the *Re* and *Ty*. For both models, the *St_N* cannot be classified accurately,

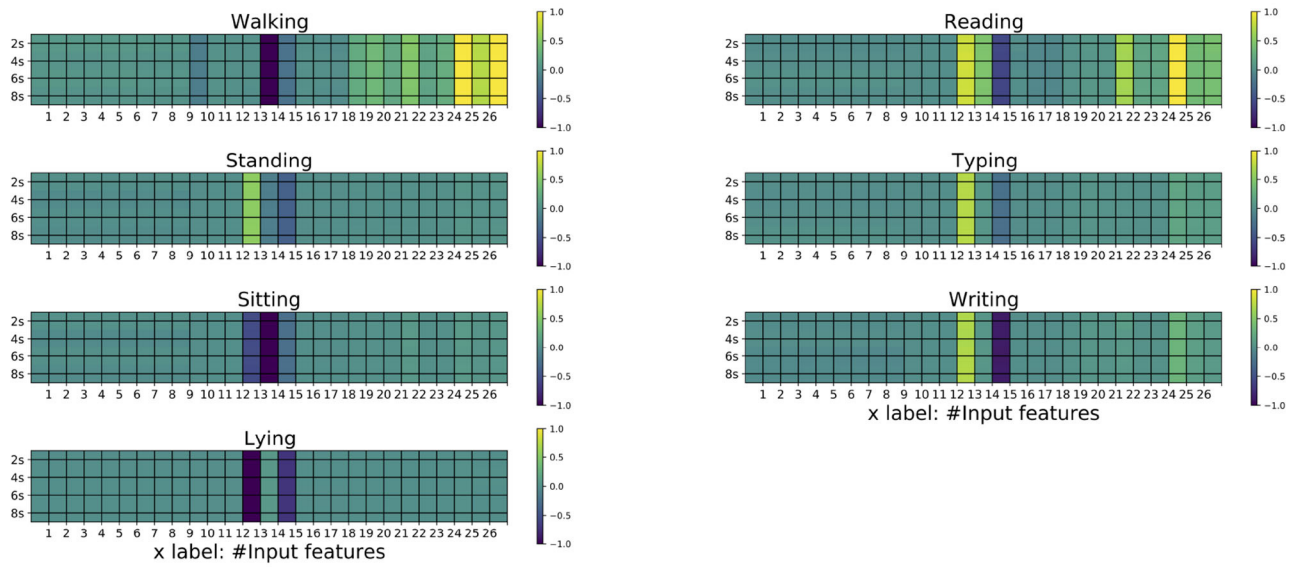


FIGURE 3 | The averaged 8-s matrices for the seven classes. The matrices on the left column corresponds to *W_N*, *St_N*, *Si_N*, *L_N*, respectively, whereas those on the right column corresponds to *Re*, *Ty*, and *Wr*, respectively. The horizontal axis corresponds the maximum eigenvector of the correlation matrix, the means, standard deviation, and the of the nine channels.

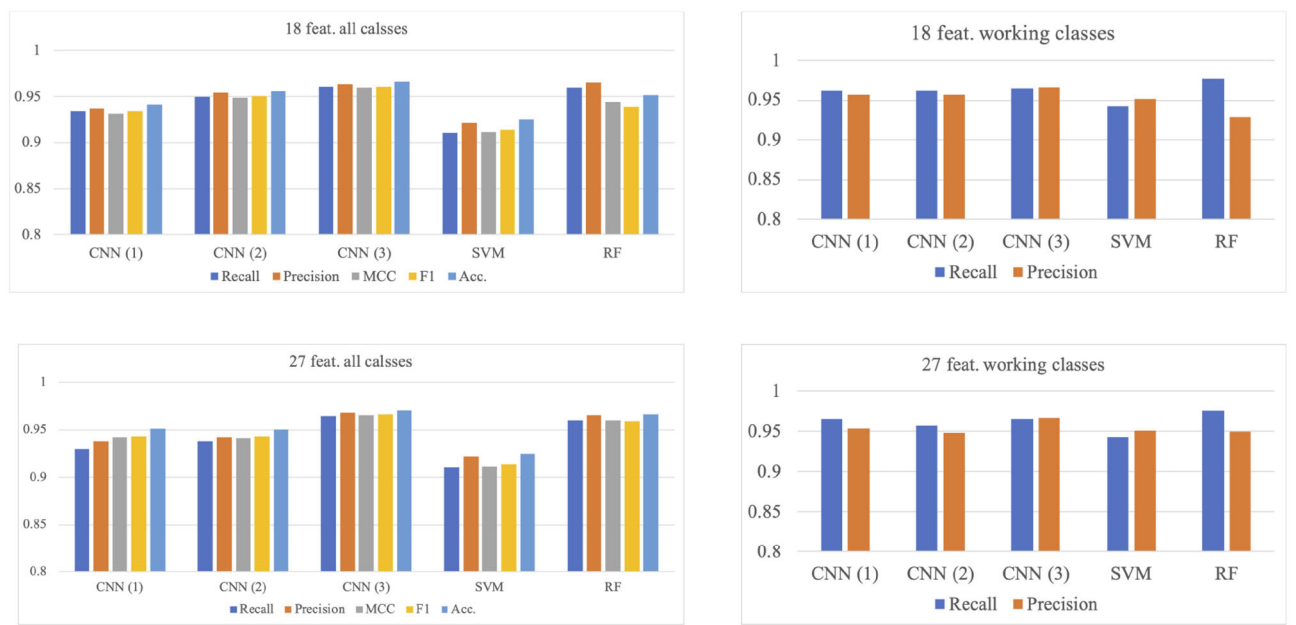


FIGURE 4 | Models comparison of 5-fold cross-validation. CNN(1), CNN(2), and CNN(3) correspond to the model with 1, 2, and 3 convolutional layer(s), respectively.

for being misclassified as *L_N* and *Si_N*. This inaccuracy is accountable; given that the subjects are free to move their hands occasionally, there is no consistent pattern for these three classes.

Figure 8 further shows the distribution of the recall and precision over the 12 validations. The long-tail distribution can be seen in the 27-feature model for *Re* and the 18-feature model for *Ty*. However, they are caused by only one case (subject) with extremely low result, respectively. Therefore, no severe overfitting is observed.

3.3. Results of Real-Time Validation

According to the results of the offline LOO validations, the 18-feature two-layers CNN model achieved the best results for

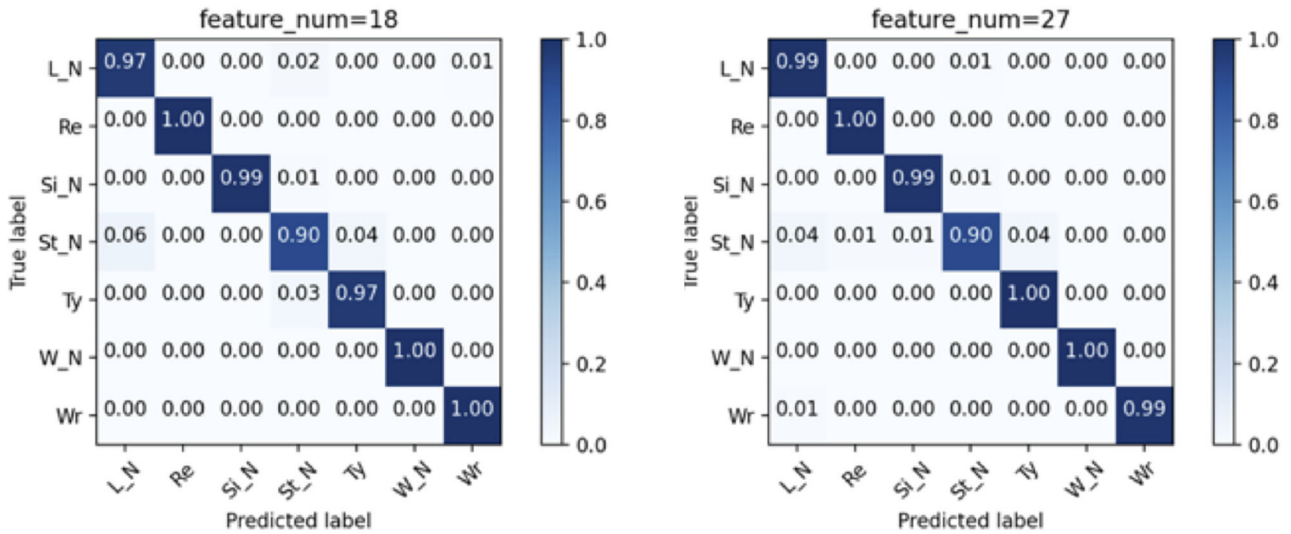


FIGURE 5 | Confusion matrices of the CNN(3) model by using the two combination of features in the 5-fold cross-validation. The correspondence is indicated by the title of the matrices.

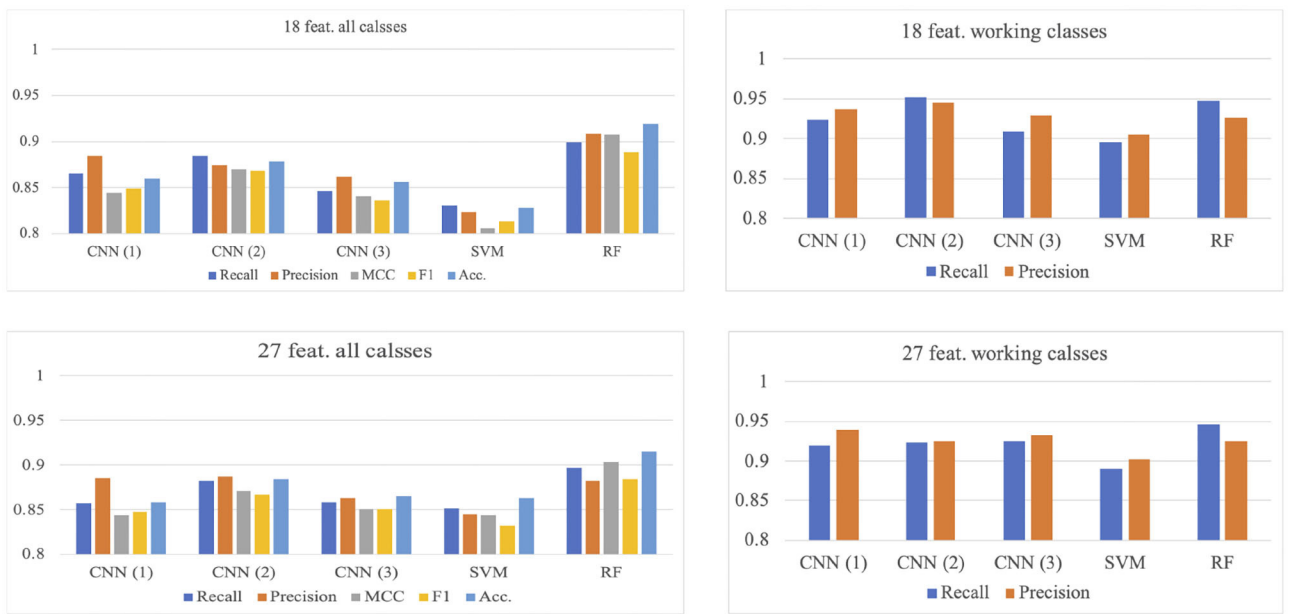


FIGURE 6 | Models comparison of LOO validation. CNN(1), CNN(2), and CNN(3) corresponds to the model with 1, 2, and 3 convolutional layer(s), respectively.

working status recognition. Therefore, it is implemented to the prototype application in iPhone and tested by real-time experiments. The averaged confusion matrix of the two subjects who have participated in the offline experiment is shown in **Figure 9A**, whereas the Matrix in **Figures 9B–D** shows the compositional bar chart of the three new subjects.

According to **Figure 8** , the results of the two subjects who have participated in the offline experiment are consistent with the 5-fold validation, having a generally accurate classification.

The results for the three new subjects vary mainly in the recognition of writing, while the reading and typing activities can be recognized accurately. The writing was recognized as typing in **Figure 9B** and as lying in **Figure 9C**. This variance may be caused by the unconstrained action of the hand, whereas the reading and typing require the hand to interplay with the book and keyboard. Noteworthily, the new subjects were free to interrupt their activities and perform irrelevant actions, such as drinking water, which made it more similar to the real working

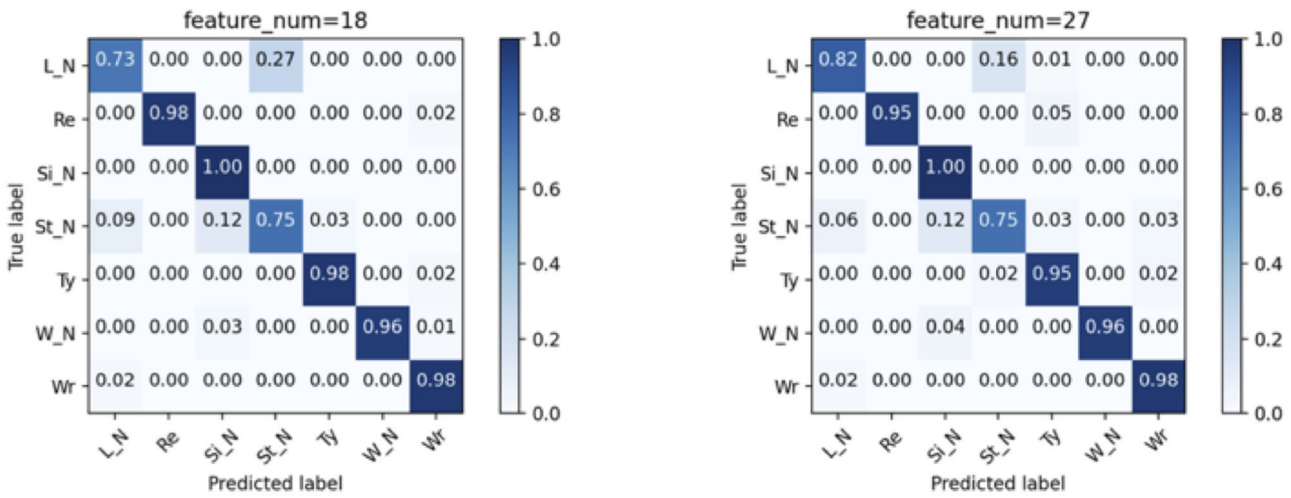


FIGURE 7 | Confusion matrices of the CNN(2) model by using the two combination of features in the LOO validation. The correspondence is indicated by the title of the matrices.

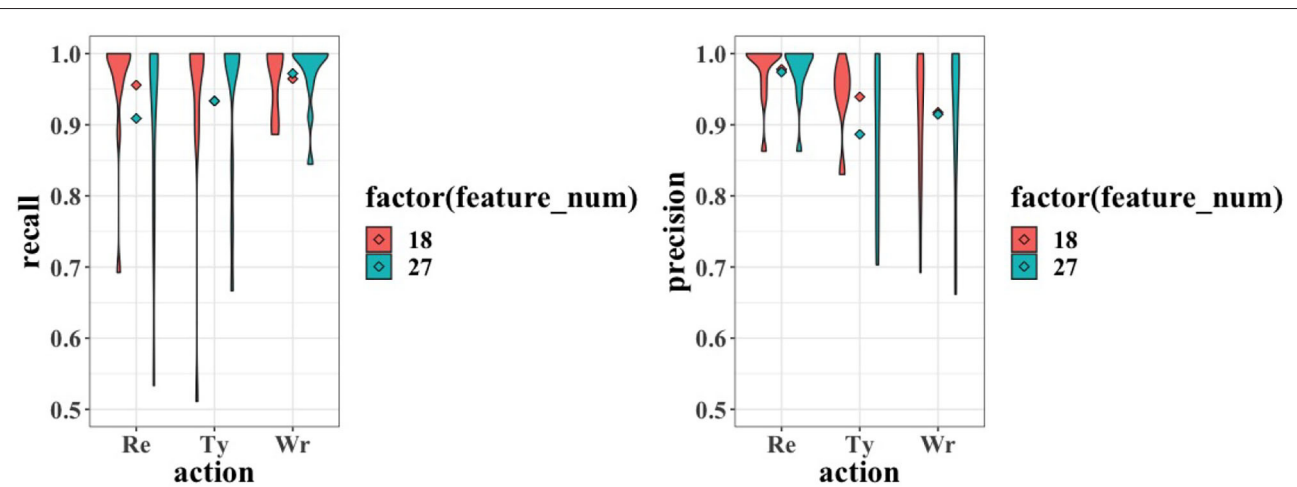


FIGURE 8 | Violin plot for the recall **(left)** and precision **(right)** values of the working status. Each violin shows the distribution of the corresponding metric over the LOO validation, whereas the dots in the middle of each pair of violins show their means.

situation. While the sitting can be recognized, the drinking is mistakenly classified as reading, probably due to the similarity in hand movement. Albeit the misclassification of these random actions irrelevant to the working status, the majorities of the three working activities have been recognized correctly. A simple statistical threshold can output a correct judge whether the user is working or not.

4. DISCUSSION

This proof-of-concept study focuses on the availability of the overall method of combining wearable sensing and edge computing based on the iOS ecosystem. Therefore, the usefulness of the wrist-worn sensors, the selection of the preprocessing scheme, and the performance of a complexity-restricted machine-learning model are the three major factors. First, the performance of the CNN model confirms the feasibility of working status recognition by using the wrist-worn sensors in the Apple Watch alone, where all the three actions that belong to the working status can be recognized accurately. Although more atomic actions, i.e., drinking water, can be added to the model, judging from the results of the classification, it is plausible to expect a similar outcome for the working status recognition.

4.1. Pre-processing

The selection of preprocessing can be reflected in the generation of input vectors. From the comparison of the two combinations, it can be seen that the cooperativity between channels expressed by the channel-wise linear correlation is not as important as the low-order statistical features (the mean and standard deviation). This piece of insight is beneficial for the edge computing because

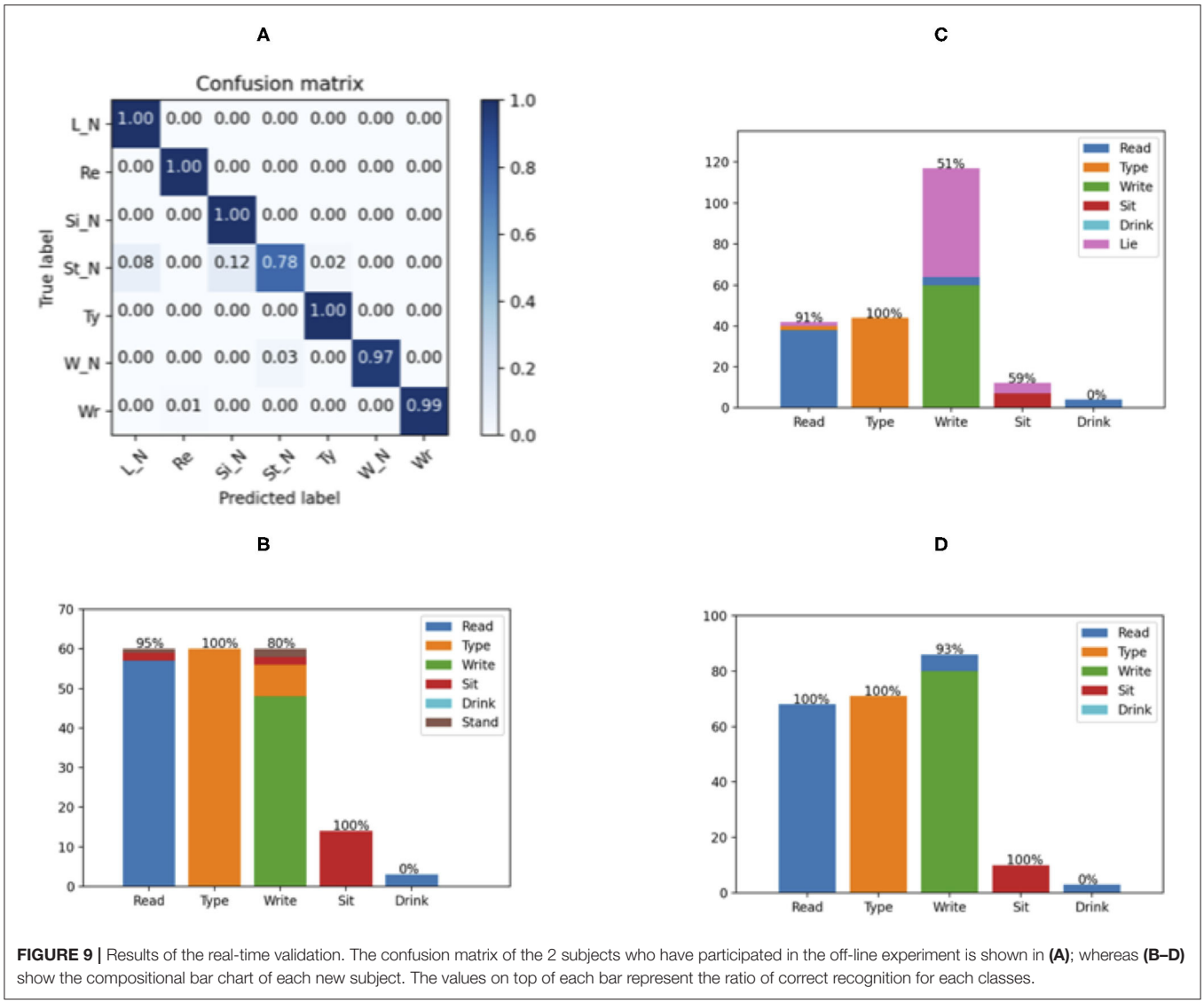


FIGURE 9 | Results of the real-time validation. The confusion matrix of the 2 subjects who have participated in the off-line experiment is shown in **(A)**; whereas **(B–D)** show the compositional bar chart of each new subject. The values on top of each bar represent the ratio of correct recognition for each classes.

a major part of computation in the preprocessing is for the calculation of cooperativity.

4.2. The CNN Model

Consistent with the results of Münzner et al. a shallow neural network is sufficient in activities recognition, this research attains the best results for LOO validation with a two-layer CNN model, a similar result as that of Münzner et al.'s research (22). By comparing the CNN models with different layers (1–3) and adding the spontaneous gyroscope signal, further improves the performance of a classification model in working status recognition without using the GPS information, which may cause misclassification of the activities that are irrelevant to the working status (17).

Along with the model explanation by using SHAP values, the importance of the features and their special patterns for each class can be inspected closer. **Figure 10** plots the SHAP values of six randomly selected samples of different classes. The subfigures from the second to the last columns represent the seven classes. First, for these six samples, the cooperativity features in the first nine columns of each subfigure have generally small SHAP values compared with the statistical features, which goes along well with the comparing results of the two different inputs. Second, although the samples are of different classes, the patterns are relatively constant for each class. For example, the subfigures in the third column correspond to the *Re* action. Although the classes of the samples vary, the patterns of the feature importance (a large absolute SHAP value implies an important feature) are very similar concentrating on the latter part of the feature vectors.

4.3. Limitations

We have to admit that the experiment setting is not as flexible as the real scenario, where the continuity of each action cannot be assured; by further summarizing these atomic actions into semantic features, a more realistic working pattern could be recognized using the current hardware setup. Furthermore, as

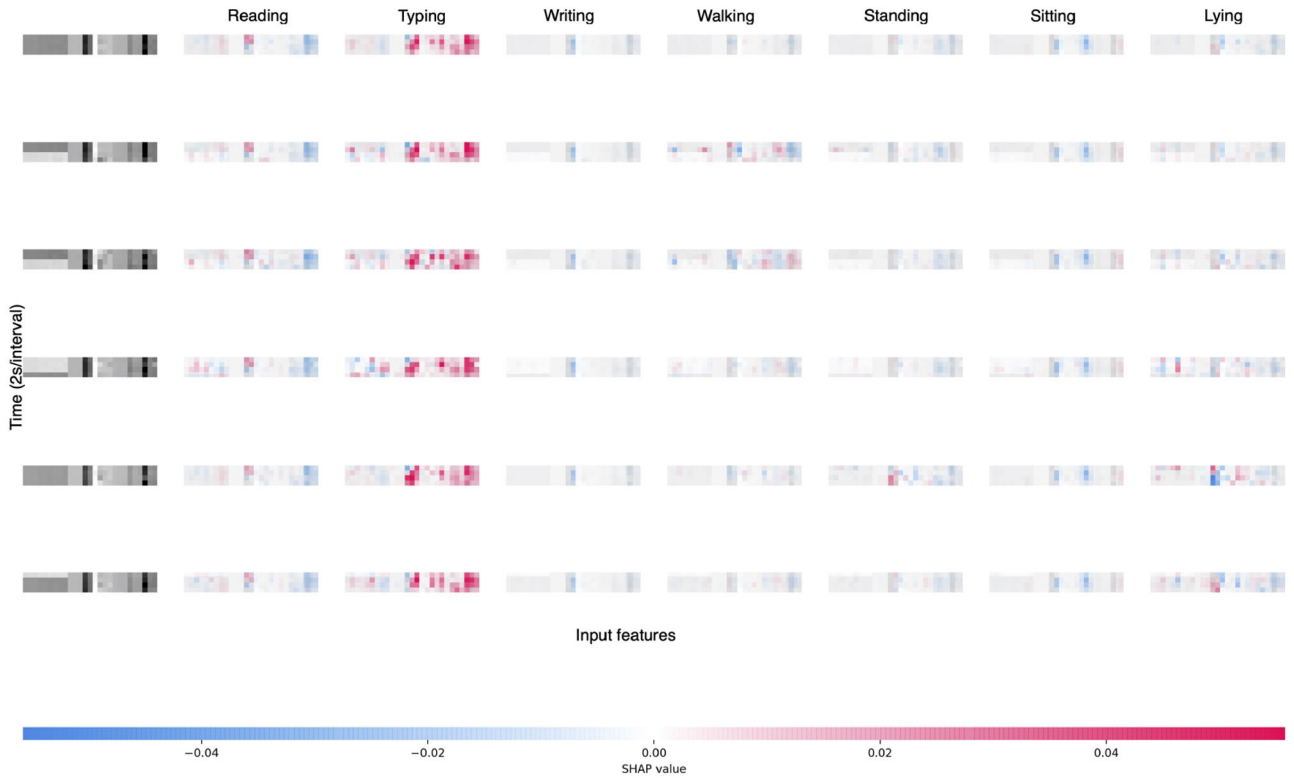


FIGURE 10 | SHAP values of six samples of different classes. The subfigures in the first column are the inputs of the samples, and the subfigures from the second to the last columns represent the seven classes in the order of L_N, Re, Si_N, St_N, Ty, W_N, Wr.

it can be seen in the real-time validation, new subjects may deteriorate the model accuracy, further extension of the dataset, and further training of the model with transfer learning could be considered in future works.

4.4. Prospect

Coming back to the ultimate purpose of this research, the usefulness and availability of the system are equally important. This premise drives us to point at the established ecosystem, where the APIs for data acquisition and modeling are expected to be further improved. Not only this working status recognition but detailed behavior pattern modeling for all-day routine and lifestyle-physiological outcome association can also be expected in the future research.

5. CONCLUSION

Aiming at providing a timely nudging to mitigate the minus effect of long-time telework without an additional device, this research examines the idea of using wrist-worn sensors in a commercial smartwatch and a smartphone to capture the real-time signals from sensors and conduct the recognition using a pre-trained CNN model. In this manner, the whole workflow can be implemented in real time with a ready hardware setup. On the other hand, by taking the power consumption of smartphone computing into account, shallow CNN structure with special consideration on the properties of the signal is validated. By rearranging the statistical features of an 8 s signal into a feature matrix and input it into the classification model, the CNN model show accurate performance [5-fold cross-validation: 0.97 recall and 0.98 precision; leave-one-out validation: 0.95 recall and 0.94 precision (SVM: 0.89 recall and 0.90 precision; random forest: 0.95 recall and 0.93 precision)] for the recognition of working status. This proof-of-concept study clarifies the prospect of a user-friendly online working tracking system, which recognizes the working status with standalone pair of smartphone and smartwatch and will nudge the user to take a break after a long working time. It is expected to contribute to the workers' wellness not only during the COVID-19 pandemic but also take effect in the post-COVID-19 era.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Review Committee concerning Research involving Human Subjects of Nara Institute of Science and

Technology. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This research and development work was supported by a Grant-in-aid for Young Scientists of the Japan Society for the Promotion of Science (JSPS) #20k19923 and Grant-in-Aid for JSPS Fellows #20J13779.

REFERENCES

1. Chandola T, Brunner E, Marmot M. Chronic stress at work and the metabolic syndrome: prospective study. *BMJ*. (2006) 332:521–5. doi: 10.1136/bmj.38693.435301.80
2. Borchini R, Veronesi G, Bonzini M, Gianfagna F, Dashi O, Ferrario MM. Heart rate variability frequency domain alterations among healthy nurses exposed to prolonged work stress. *Int J Environ Res Public Health*. (2018) 15:113. doi: 10.3390/ijerph15010113
3. Henning RA, Jacques P, Kissel GV, Sullivan AB, Alteras-Webb SM. Frequent short rest breaks from computer work: effects on productivity and well-being at two field sites. *Ergonomics*. (1997) 40:78–91.
4. Chaikumarn M, Nakphet N, Janwantanakul P. Impact of restbreak interventions on the neck and shoulder posture of symptomatic VDU operators during prolonged computer work. *Int J Occup Safety Ergon*. (2018). 24:251–9. doi: 10.1080/10803548.2016.1267469
5. Zhu W, Gutierrez M, Toledo MJ, Mullane S, Stella A, Diemar R, et al. Long-term effects of sit-stand workstations on workplace sitting: a natural experiment. *J Sci Med Sport*. (2017) 21:811–6. doi: 10.1016/j.jsams.2017.12.005
6. Murtaza F, Yousaf MH, Velastin S. PMHI: proposals from motion history images for temporal segmentation of long uncut videos. *IEEE Signal Process Lett*. (2017) 25:179–83. doi: 10.1109/LSP.2017.2778190
7. Wang Y, Cang S, Yu H. A survey on wearable sensor modality centred human activity recognition in health care. *Expert Syst Appl*. (2019) 137:167–90. doi: 10.1016/j.eswa.2019.04.057
8. Ordonez F, Roggen D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors*. (2016) 16:115. doi: 10.3390/s16010115
9. Cao J, Li W, Ma C, Tao Z. Optimizing multi-sensor deployment via ensemble pruning for wearable activity recognition. *Inform Fusion*. (2017) 41:68–79. doi: 10.1016/j.inffus.2017.08.002
10. Garcia Ceja E, Galvan Tejada C, Brena R. Multi-view stacking for activity recognition with sound and accelerometer data. *Inform Fusion*. (2018) 40:45–56. doi: 10.1016/j.inffus.2017.06.004
11. Cao L, Wang Y, Zhang B, Jin Q, Vasilakos A. GCHAR: an efficient group-based context–aware human activity recognition on smartphone. *J Parallel Distrib Comput*. (2017) 118:67–80. doi: 10.1016/j.jpdc.2017.05.007
12. Cvetkovic B, Szeklicki R, Janko V, Lutomski P, Lustrek M. Real-time activity monitoring with a wristband and a smartphone. *Inform Fusion*. (2017) 43:77–93. doi: 10.1016/j.inffus.2017.05.004
13. Bianchi V, Bassoli M, Lombardo G, Fornacciari P, Mordonini M, De Munari I. IoT wearable sensor and deep learning: an integrated approach for personalized human activity recognition in a smart home environment. *IEEE Intern Things J*. (2019) 6:8552–62. doi: 10.1109/JIOT.2019.2920283
14. Abdu-Aguye M, Gomaa W. Robust human activity recognition based on deep metric learning. In: *Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2019)*. Prague (2019). p. 656–63. doi: 10.5220/0007916806560663
15. Mohammed S, Ogawa T, Gomaa W. CHARM-deep: continuous human activity recognition model based on deep neural network using IMU sensors of smartwatch. *IEEE Sens J*. (2020) 20:8757–70. doi: 10.1109/JSEN.2020.2985374
16. Mannini A, Intille SS, Rosenberger M, Sabatini AM, Haskell W. Activity recognition using a single accelerometer placed at the wrist or ankle. *Med Sci Sports Exerc*. (2013) 45:2193. doi: 10.1249/MSS.0b013e31829736d6
17. Kwon M-C, Choi S. Recognition of daily human activity using an artificial neural network and smartwatch. *Wireless Commun Mobile Comput*. (2018) 2018:2618045. doi: 10.1155/2018/2618045
18. Morales J, Akopian D. Physical activity recognition by smartphones, a survey. *Biocyber Biomed Eng*. (2017) 37:388–400. doi: 10.1016/j.bbe.2017.04.004
19. Cruciani F, Cleland I, Nugent C, McCullagh P, Synnes K, Hallberg J. Automatic annotation for human activity recognition in free living using a smartphone. *Sensors*. (2018) 18:2203. doi: 10.3390/s18072203
20. Nweke HF, The YW, Mujtaba G, Al-garadi MA. Data fusion and multiple classifier systems for human activity detection and health monitoring: review and open research directions. *Inf Fusion*. (2019) 46:147–70. doi: 10.1016/j.inffus.2018.06.002
21. Bulling A, Blanke U, Schiele B. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput Surv*. (2014) 46:1-33. doi: 10.1145/2499621
22. Münzner S, Schmidt P, Reiss A, Hanselmann M, Stiefelhagen R, Dürichen R. CNN-based sensor fusion techniques for multimodal human activity recognition. In: *Proceedings of the 2017 ACM International Symposium on Wearable Computers*. (2017). p. 158–65. doi: 10.1145/3123021.3123046
23. Redmon J, Farhadi A. Yolov3: an incremental improvement. *arXiv* [Preprint]. (2018) arXiv:1804.02767.
24. Wang H, Kim B, Xie J, Han Z. How is energy consumed in smartphone deep learning apps? Executing locally vs. remotely. In: *2019 IEEE Global Communications Conference (GLOBECOM)*. Waikoloa (2019). p. 1–6. doi: 10.1109/GLOBECOM38437.2019.9013647
25. Chen Z, Zhu Q, Soh YC, Zhang L. Robust human activity recognition using smartphone sensors via CT-PCA and online SVM. *IEEE Trans Ind Inform*. (2017) 13:3070–80. doi: 10.1109/TII.2017.2712746
26. Khan A, Hammerla N, Mellor S, Plötz T. Optimising sampling rates for accelerometer-based human activity recognition. *Patt Recog Lett*. (2016) 73:33–40. doi: 10.1016/j.patrec.2016.01.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 17 June 2021
doi: 10.3389/fdgth.2021.660823

Digital Contact Tracing Against COVID-19 in Europe: Current Features and Ongoing Developments

Alessandro Blasimme[*†], *Agata Ferretti*[†] *and Effy Vayena*

Health Ethics and Policy Lab, Department of Health Sciences and Technology, ETH Zürich, Zürich, Switzerland

The SARS-CoV-2 pandemic is a public health challenge of unprecedented scale. In the midst of the first wave of the pandemic, governments worldwide introduced digital contact tracing systems as part of a strategy to contain the spread of the virus. In Europe, after intense discussion about privacy-related risks involving policymakers, technology experts, information technology companies, and—albeit to a limited extent—the public at large, technical protocols were created to support the development of privacy-compatible proximity tracing apps. However, as the second wave of SARS-CoV-2 sweeps the continent, digital contact tracing in Europe is evolving in terms of both technological and governance features. To enable policymakers to harness the full potential of digital health tools against SARS-CoV-2, this paper examines the evolution of digital contact tracing in eight European countries. Our study highlights that while privacy and data protection are at the core of contact tracing apps in Europe, countries differ in their technical protocols, and in their capacity to utilize collected data beyond proximity tracing alone. In particular, the most recently released apps tend to offer users more granular information about risk in specific locations, and to collect data about user whereabouts, in order to enhance retrospective contact tracing capacity. These developments signal a shift from a strict interpretation of data minimization and purpose limitation toward a more expansive approach to digital contact tracing in Europe, calling for careful scrutiny and appropriate oversight.

Keywords: APP, digital contact tracing, COVID-19, governance, privacy, epidemiology, Europe

OPEN ACCESS

Edited by:
Pradeep Nair, Central University of Himachal Pradesh, India

Reviewed by:
Monika Semwal, Nanyang Technological University, Singapore
Baobao Zhang, Cornell University, United States

***Correspondence:**
Alessandro Blasimme
alessandro.blasimme@hest.ethz.ch

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *29 January 2021*
Accepted: *20 May 2021*
Published: *17 June 2021*

Citation:
Blasimme A, Ferretti A and Vayena E (2021) Digital Contact Tracing Against COVID-19 in Europe: Current Features and Ongoing Developments. Front. Digit. Health 3:660823. doi: 10.3389/fdgth.2021.660823

INTRODUCTION

The SARS-CoV-2 pandemic is a public health challenge of unprecedented scale. Worldwide, 163 million people have tested positive for SARS-CoV-2 and 3.38 million have lost their life to the coronavirus disease (COVID-19) (1). As of May 2021, Europe alone has had more than 31 million cases and 700 thousand deaths according to the most recent estimates of the European Center for Disease Control and Prevention (2). Since late summer 2020, Europe has faced a resurgence of new cases as a second wave of SARS-CoV-2 spread across the continent, placing health systems under severe pressure and forcing governments to reinstate restrictions similar to those adopted in the first quarter of the year.

Alongside restrictions to population movement during the first wave of the pandemic, governments throughout the world introduced digital contact tracing (DCT) systems, in the hope that this new digital health technology would help contain the spread of the virus (3). DCT systems

mostly come in the form of smartphone apps that, using technologies commonly present in such devices–such as the Bluetooth data exchange standard or the global positioning system (GPS)–can keep track of the proximity between devices that have the app installed. Proximity data can be used to infer the risk that two users might have been close enough and for a sufficient amount of time to infect each other with the SARS-Cov-2 virus. Once a user tests positive for the virus, DCT apps can send an alert to other users who have been in close contact with her according to the proximity data recorded by the system. Alerted users can then test and isolate thus reducing the circulation of the virus in a given population. Asian countries were among the first to adopt DCT. Recognizing the public health potential of DCT, many European countries followed suit during the spring, developing national DCT systems in an attempt to expand their contact tracing capability.

Despite the promising potential of DCT, its introduction gave rise to intense debate over ethical, legal, and societal implications (ELSI). In particular, some characteristics of the Asian approach (mandatory use, centralized protocols, GPS- or cell tower-based geolocation) are seen by many as incompatible with European legal provisions and ethical views about the value of individual privacy.

For this reason, European policymakers, in close collaboration with technology experts and IT companies, started developing DCT standards based on the exchange of anonymized Bluetooth data. The European approach to DCT is defined in specific guidelines issued by the European Commission (EC) on April 17, 2020. This guidance is centered around the principle of data minimization, including precisely defined limits for data disclosure, use, and storage (4).

Meanwhile, in mid-April, the eHealth Network (comprising representatives of authorities responsible for digital health in the 27 EU Member States plus Norway) published a common toolbox, specifying essential requirements for European DCT apps. This toolbox emphasized a preference for decentralized protocols which store anonymized proximity data exclusively on users' mobile phones, over protocols storing data on centralized servers that are run by national health authorities. In particular, echoing the opinion of the European General Data Protection Board, this guidance underscored decentralized approaches as better suited to "keep personal data processing to the absolute minimum," enhance citizens' willingness to download and use DCT apps, and prevent "risks of data breaches and cyberattacks" (5).

At this time, many European technology experts were still collaborating on a centralized protocol called the Pan-European Privacy-Preserving Proximity Tracing protocol (PEPP-PT). Ultimately, though, some members of the PEPP-PT project resigned from this consortium in order to form a new protocol (6). The privacy-preserving decentralized protocol (Decentralized Privacy-Preserving Proximity Tracing, or DP-3T for short) was developed by a number of European academic institutions, in conjunction with the Swiss Federal Institutes of Technology (ETH Zurich and the EPFL of Lausanne).

In the meantime, Google and Apple released an application programming interface (API) to implement this protocol on Apple and Android mobile operating systems (Google/Apple Exposure Notification system, or GAEN for short). Most decentralized DCT systems in Europe, including the Swiss model, run on this protocol. Countries such as Germany and the UK used the centralized model initially, but adopted the decentralized scheme powered by Google and Apple for the final version of their national DCT apps, introduced on June 16 and September 24, 2020, respectively (7, 8).

At the time of writing, 19 of the 27 EU Member States plus Switzerland have created a national DCT app (9). Of these, only France and Hungary have opted for a centralized solution (10).

In this comparative study of national proximity tracing apps, we seek to characterize the European approach to DCT, and to examine its evolution between the first and second waves of SARS-CoV-2. Our analysis shows that European DCT systems, to some extent, are evolving to incorporate new features extending their capabilities beyond mere proximity tracing—a development that calls for careful scrutiny and adequate oversight.

METHODS

In order to examine the evolution of the European DCT landscape, we collected information from primary sources about national DCT apps in the following countries: France, Germany, Ireland, Italy, the Netherlands, Switzerland, and the UK (including England, Wales, and Scotland).

We included DCT systems released between March and October 2020. All the systems we included in our analysis revolve around a smartphone app as their key implementation technology. For inclusion in our study, the language of the app had to be English, French, Italian, or German (languages spoken by the authors). For each app, we collected *Privacy Policy* and the *Terms of Use* documentation from the app itself or its associated website. When available, we also analyzed the "FAQ section" and "press release" documentation, which usually contain a series of questions and answers, as well as concise information about the app's functionality and data processing. A list of the primary sources analyzed is available as **Supplementary Material**. Each source was archived (on archive.org) as it appeared at the time of review.

From this documentation, we first extracted and recorded general information and technical features for each DCT app (via MS Excel). Next, we imported the retrieved documents into Nvivo for qualitative content analysis. Two researchers (AB and AF) inductively created analytic codes from the text until thematic saturation was achieved (11). Semantically similar codes were further grouped into themes and subthemes. Two researchers (AB and AF) coded the text independently and resolved any coding discrepancies through discussion.

For comparative purposes, we collected information from secondary sources about national DCT systems in Asia. A list of these sources is available as **Supplementary Material**.

RESULTS

Common Characteristics

Table 1 provides a summary of select descriptive features for each included DCT app. A certain degree of similarity is evident across the analyzed DCT systems. For example, all of them

TABLE 1 | Characteristics of DCT systems in selected European countries.

	Switzerland	Italy	Germany	Ireland	Netherlands	United Kingdom		France
						Scotland	England & Wales	
								
Total COVID-19 cases[§]	426,199	2,038,759	1,640,858	85,394	754,171	2,256,009		2,507,532
Cumulative prevalence per 1 million population[§]	49,245,25	33,719,77	19,584,4	17,293,99	44,013,81	33,232,31		38,415,77
Total deaths[§]	6,508	71,620	29,778	2,200	10,974	70,405		62,197
App name	SwissCovid	Immuni	Corona-Warn-App	COVIDTracker	CoronaMelder	ProtectScotland	NHS Covid-19	TousAntiCovid
Release date	25.05.20	15.06.20	16.06.20	07.07.20	17.08.20	14.09.20	24.09.20	22.10.20
N. of downloads	2,863,858[a]	10,072,742[b]	24,200,000[c]	2,700,000[d]	4,330,264[e]	1,739,806[f]	20,739,925[g]	11,897,809[h]
Developed in public-private partnership	✓	✓	✓	✓	✓	✓	✓	✓
Voluntariness	✓	✓	✓	✓	✓	✓	✓	✓
De-centralized protocol	✓	✓	✓	✓	✓	✓	✓	✗
Exposure parameters	1.5 meters for 15 min	<8 meters 10 min	2 meters for 15 min	2 meters for 15 min	"near" for 15 min	2 meters for 15 min	2 meters for 15 min	2 meters for 15 min
Data retention: Random ID Exposure code	14 days 14 days	14 days 14 days	14 days 21 days	14 days 14 days	14 days 21 days	14 days 14 days	14 days 14 days	14 days 14 days

[§]*at 28.12.2020 source: https://covid19.who.int.*
[a]*at 21.12.2020 source: https://www.experimental.bfs.admin.ch/expstat/en/home/innovative-methods/swisscovid-app-monitoring.html.*
[b]*at 28.12.2020 source: https://www.immuni.italia.it/dashboard.html.*
[c]*at 17.12.2020 source: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/WarnApp/Archiv_Kennzahlen/Kennzahlen_18122020.pdf?__blob=publicationFile.*
[d]*at 28.12.2020 source: CovidTracker App.*
[e]*at 23.12.2020 source: https://github.com/minvws/nl-covid19-notification-app-statistics/blob/main/statistics/appstore_statistics.csv.*
[f]*at 28.12.2020 source: ProtectScotland App.*
[g]*at 16.12.2020 source: https://www.gov.uk/government/publications/nhs-test-and-trace-england-statistics-10-december-to-16-december.*
[h]*at 28.12.2020 source: TousAntiCovid App.*

were developed in public-private partnerships between the state (or national health authority), software development companies, and, at times, research institutions. Furthermore, all of the apps function on a voluntary basis, in order to safeguard individual freedom. Moreover, a strong focus on privacy preservation and data protection is a common feature of the European approach to DCT. However, not all countries use the same architecture to achieve this aim.

The majority of DCT apps rely on decentralized protocols. These apps operate with the privacy-preserving technology framework released by Google and Apple, which allows matching codes to be kept on the user's phone, and in the case of a positive test, fetches only an anonymized ID from a centralized database, in order to check for high risk contacts. Among the apps we analyzed, only the French *TousAntiCovid* adopts a centralized approach to data storage. To justify this decision, the French government argued that the Google/Apple system contradicts the digital sovereignty of the state and does not provide sufficient privacy safeguards, as sensitive data about positive cases, albeit encrypted, are accessed by users' apps (12–14). Moreover, as the FAQ section of the French app specifies, "*the Government considers that protecting the health of the French people is a mission that is the exclusive responsibility of the State and not of private international actors*"[1] (15).

From a technical perspective, European DCT apps employ similar exposure parameters (two meters apart for 15 min) to notify app users of a potentially dangerous contact. Taking a precautionary approach, the German *Corona-Warn-App* uses the most stringent exposure parameters, alerting a user who is within eight meters and for at least 10 min from an individual with a confirmed SARS-CoV-2 infection. The French *TousAntiCovid* employs the least stringent criteria of one meter apart for 15 min.

Despite differences in data storage locations across countries, we noted that data retention periods are consistent, both for randomly generated ID codes as well as temporary exposure codes. Randomly generated ID codes are generally stored for 14 days, while positive exposure codes are kept for 14 (Ireland, Italy, France, Netherlands, Scotland, and the UK) or 21 (Germany and Switzerland) days.

All of the reviewed systems collect statistical data concerning the number of users who downloaded the app, the number of apps actually in use, the positive cases uploaded to the system, the number of alerts sent to users, and the functioning of the app (e.g., Bluetooth signal strength, success of the data exchange, and the time at which the data must be destroyed). Some apps such as *SwissCovid* (Switzerland), *Immuni* (Italy), and *Corona-Warn-App* (Germany) have dedicated web pages offering aggregate information on how the respective DCT systems are being used (16–18). The apps also collect metrics data for public health surveillance, such as the day, time, and duration of a contact; whether the infected user is asymptomatic; the 1st day of illness; and the date of testing. Countries may retain such anonymous data for epidemiological surveillance or research purposes, however retention periods vary across countries. In Italy, metric (i.e., aggregated statistical) data is kept until the end of the emergency, but no later than 31 December 2021 (a limit previously set to the end of 2020.) In Ireland, England, and Scotland, metric data are retained, respectively for at least 7 years, 20 years, and indefinitely.

The seamless functioning of national DCT apps across borders motivated the European Commission to create an EU-wide system called *getaway*, to enable interoperability and help break the chain of COVID-19 infection across borders. The *getaway* would allow users who have installed one DCT app to travel to another participating European country and still receive contact tracing alerts (18). So far, however countries with interoperable apps include only Croatia, Denmark, Italy, Ireland, Germany, Latvia, the Netherlands, Poland, and Spain.

Country-Specific Features

European DCT apps differ in three respects: what happens upon notification of a contact with a positive case; how positive test results are handled; and additional features beyond proximity tracing.

In all cases analyzed, DCT apps advise users on what to do upon notification of close contact with someone who has tested positive for SARS-CoV-2. Most apps give users instructions for how to self-isolate, register for testing, and contact health authorities if symptoms emerge. The Irish *COVID Tracker* app allows users to voluntarily add a phone number, which is shared with health authorities. In case of close contact, the user not only is alerted by the app, but also phoned by the health authority that provides information about next steps and eventually arrange a COVID-19 test.

Each country follows its own procedure for uploading a positive test result into the DCT system. In Scotland for instance, health authorities send an exposure code via SMS to users who tests positive. Users enter the code, active for 72 h, into the *ProtectScotland* app. In France, the code is sent to users in a link via email, and via post as a QR code. Users must thus enter personal information (mobile number, email, address) in order to communicate the outcome of a positive test and trigger notification to other users. Other countries have chosen methods which avoid this requirement. Users of the German *Corona-Warn-App* can scan a QR code linked to test results, automatically activating the exposure code. In Switzerland, Italy, and the Netherlands, users must phone the health authority upon notification of a positive test result, in order to activate the exposure code.

Our qualitative assessment explored the evolution of DCT apps as one component of broader policy efforts intended to curb the economic and public health effects of the pandemic. The most recently released DCT apps were introduced after the summer, when a second wave of SARS-CoV-2 was already apparent in most European countries. At this time, some national apps released features that went beyond simple proximity tracing (see **Figure 1**). For instance, the French *TousAntiCovid* [the successor to a previous app called *StopCovid*, which was downloaded by a mere 2.6 million people and therefore replace by *TousAntiCovi*d (19)] expanded its functionality, allowing users to enter their postal code to receive more granular information about the local epidemiological situation. Moreover, users of

[1]Translated by the authors.

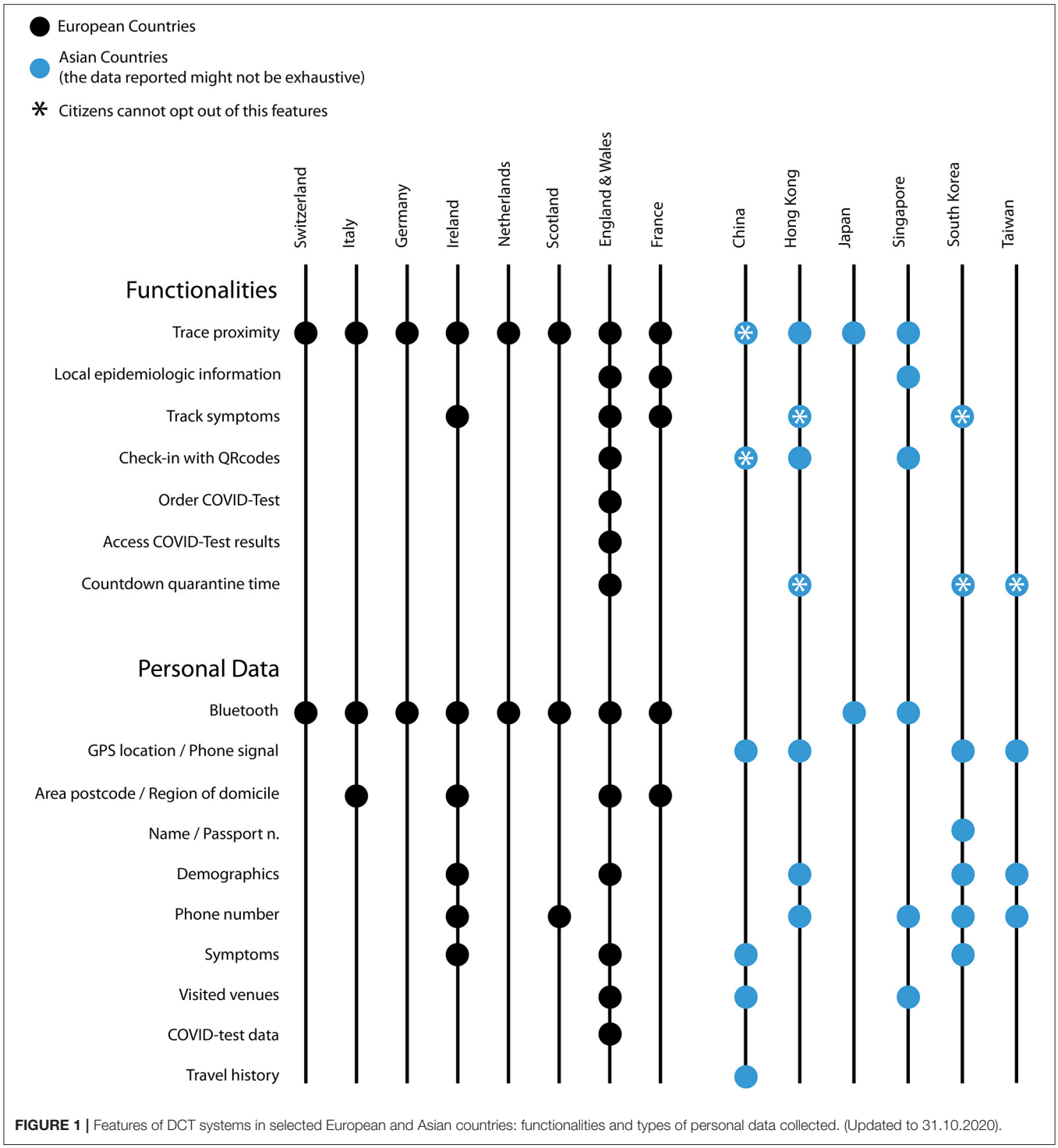


FIGURE 1 | Features of DCT systems in selected European and Asian countries: functionalities and types of personal data collected. (Updated to 31.10.2020).

the French app can access a government website (*Depistage COVID-19*) with a map of open testing centers and their current waiting times (20). The *NHS Covid-19* app, available in England and Wales, offers COVID-19 risk estimates as well. When users enter their postcode, they receive a notification of risk-level (low, medium, high) based on aggregate COVID-19 case information available to local authorities in a given area (21).

A daily symptom checker is integrated into the Irish *COVID Tracker* app, alongside the contact tracing function. This feature

enables users to receive personalized recommendations (e.g., self-quarantine, call their physician, request a COVID-19 test) in relation to any symptoms and their severity, and to demographic data voluntarily entered into the system. The French app allows users to connect to a similar symptom checker, hosted on a separate government website (22). The *NHS Covid-19* app, one of the latest to be released in October 2020, integrates a symptom checker tool alongside the option to order a COVID-19 test, via a link to the NHS Test and Trace website. Users can then receive results directly through the app. These new features qualify the app as a medical device, as they enable collection of health data and provide personalized health recommendations to users. The *SwissCovid* and *NHS Covid-19* apps are the only European DCT apps registered as Class I medical devices (23).

The *NHS Covid-19* app also functions as a countdown tool for self-isolation, and a check-in instrument when visiting public venues. The first function, which calculates the length of time a user should self-isolate, activates automatically when a user is notified of contact with a positive case. Based on the encounter date, the app recommends the user self-isolate for 10 days, beginning with the last encounter with the infected person. The countdown tool is also activated when a user enters COVID-19 symptoms or a positive COVID-19 test result into the app. The countdown is initiated, respectively on the day on which symptoms first appear or on the day of the test.

The check-in function allows users to scan a QR code when entering public spaces such as restaurants, bars, shops, cinemas, or religious centers (24, 25). Location data is stored on the user's phone for 21 days. Authorities cannot access this information unless users decide to make it available. NHS documentation explains that the check-in function enables users to record locations visited. App users can then decide to voluntarily disclose this information to contact tracers in case they receive a positive test result. Contact tracers routinely collect information from individuals who test positive (whether they use the NHS app or not) about places visited in the days prior to the test. While individuals have the right not to disclose recent locations that they visited, this information allows contact tracers to alert others who visited the same location. UK health authorities use this information also to assess the level of risk based on the aggregate number of coronavirus cases reported at a particular venue in a certain time period, together with the type of venue (e.g., its architecture). This activity enables health authorities to update the list of places considered to be risky. When public health officials identify a venue as "at risk," they add to a national reference list that is synchronized with the *NHS Covid-19* app. The app can thus issue an alert to users who have checked in at a risky venue. The tone of the alert message is calibrated according to the level of risk identified by the local health protection team. If risk is high, the user may be urged to call the health authority immediately. The alert does include information about the venue itself.

Governance and Oversight

Privacy Policies and Terms of Use documents provide information concerning the ethical conditions and legal bases for treatment of personal data by the respective national DCT systems. This information is meant to lend legitimacy to DCT activities, and to reassure the public about legal compliance.

All privacy policy documents of EU member states also make reference to the General Data Protection Regulation (GDPR), particularly concerning the protection of the rights of data subjects ensured by this Europe-wide legislation. For example, the Irish *COVID Tracker* privacy policy declares that "*The app is voluntary to use and the legal basis for the processing of the data is consent—namely Article 6(1)(a) of the GDPR for the processing of personal data and Article 9(2)(a) of the GDPR for the processing of special categories of personal data, in this case health related data*." (26).

In some cases, data governance principles are also reported. The privacy notice of the Italian DCT app *Immuni* declares compliance with Articles 13-14 of EU GDPR and respect for the principles of privacy ("*Under no circumstances will the users' movements be tracked, thus excluding any form of geolocation.*"), purpose limitation, and data minimization ("*Only the data necessary to alert the users that they have been exposed to a risk of infection, as well as to enable the adoption of any prevention and healthcare measures, are collected*") (27).

In Switzerland as well, DCT documentation provides the legal basis for the processing of collected data, referencing both existing and new, *ad hoc* provisions: *"The federal legislation on data protection is applicable to the data processing. In addition, the Data Protection Statement is in line with the Epidemics Act of 28 September 2012 (EpG; SR 818.101) and the Ordinance of 24 June 2020 on the Proximity Tracing System for the Coronavirus SARS-CoV-2 (VPTS; SR 818.101.25)"* (28).

These documents frequently mention the role national data protection authorities and various expert bodies played in the early assessment of DCT apps. The FAQ section of *TousAntiCovid*, for example, explains that before the launch of the app, a number of national advisory bodies was consulted on the question of digital tools and privacy protection. The *Conseil Scientifique COVID-19* came out in favor of the app, affirming the usefulness of digital tools in light of the updated "*Test, Alert, Protect*" strategy. Furthermore, the documentation notes the approval of the CNIL (*Commission Nationale Informatique et Libertés*, the French data protection authority), which was responsible for assessing whether adequate data protection measures were in place, both before and after the launch of the app (15).

Documentation from various countries describes the effort to engage a broader array of societal actors in the development of the DCT system. For example, the documentation of the *ProtectScotland* app states that "*The Scottish Government and the NHS Scotland have rigorous information governance process in place. From the early stages of the design of the app, a thorough consultation with relevant Scottish groups of interests and advocacy has taken place, including: The Health and Social Care (Scotland) Public Benet and Privacy Panel; The Scottish Privacy Forum; The Open Rights Group; The COVID-19 Data and Intelligence Network—Data ethics and public engagement subgroup; and representatives of the general public*" (29). However, no details are provided as to public engagement initiatives for the rest of the DCT systems in our sample.

In all cases, the analyzed documentation offers information concerning accountability for the lawful and responsible handling of personal data. For example the German *Corona-Warn-App*'s privacy policy reports that the app *"is provided by the Robert Koch Institute [. . .]. The RKI is also what is called the controller under data protection law, meaning it is responsible for the processing of App users' data. You [the user] can contact the RKI's data protection officer at the above address"*(30). Likewise, the Dutch *CoronaMelder* privacy policy cites the Minister of Health, Welfare and Sport, and the Regional Health Service, as controllers and accountable bodies for the protection of user data against potential abuse, loss, unauthorized access, unwanted disclosures, and unauthorized changes (31). Our study indicates that national governments and departments of health are the authorities responsible for the good functioning of DCT apps, as well as for communication with users and/or intervention when issues arise.

Despite efforts toward the transparent governance of DCT apps, limited information is available about oversight bodies and mechanisms charged with regularly assessing the functioning of DCT systems.

Two exceptions are the commitment by the NHS in England and Wales to review the privacy impact assessment in the event of software updates. As mentioned previously, this app "*is CE marked as Class I medical device in the United Kingdom and developed in compliance with European Commission Directive 93/42/EEC for Class I devices*" (25). As such, the app is subject to stricter oversight regulation (32). The Scottish DCT app also provides some details about the oversight mechanism in place; its documentation states that "*any future changes [to the app] will follow rigorous scrutiny; the decision will be balanced against public health benefit and cost (balanced against other health priorities) and this privacy notice will be updated accordingly for transparency*" (29).

Apart from these two cases, DCT documents do not relay how the responsible institutions intend to monitor an app's activity and the addition of new features over time. Notably, the Dutch documentation stresses that it is the responsibility of the user to check for data information notice updates (which may be introduced with future app developments). These changes will be in immediate effect in the app following publication of the updated privacy policy. Similarly, all of the *Terms of Use* that we analyzed encourage users themselves to inspect the app's source code (via online platforms such as GitHub/GitLab), as well as to report back about their experience of using the app (including any potential issues).

DISCUSSION

The European approach to DCT has been characterized by marked attention to privacy preservation and data protection. The General Data Protection Regulation (GDPR), in force since May 2018 in European Member States, played a central role in shaping this approach. The GDPR considers the protection of natural persons in relation to personal data processing as a fundamental right (rec. 1 GDPR). Moreover, it recognizes the challenges that new technological developments, together with the global reach of big technology corporations, pose to the protection of personal data (rec. 6 GDPR). Article 25 of the GPDR espouses the principles of data protection by design and by default, making them a legal requirement. These requirements arguably played a key role in shaping the European approach to DCT.

In general, data protection by design asserts that data processing activities should adopt state-of-the-art data protection safeguards across all technical components and processes. Data protection by default refers to the principle that data processing options should automatically be set to the most privacy preserving mode. From a practical point of view, these principles translate into a series of requirements, including data minimization and individual control of personal data. Data minimization contends that only data strictly necessary for a specific purpose should be collected and used, and there must be fixed limits on the extent of processing and the duration of storage and accessibility (art 25.2). Individual control refers to the principle that personal data should be made accessible only upon authorization of data subjects.

These provisions ensure the voluntary nature of European DCT systems, and the selection of privacy-preserving technological solutions for DCT. In particular, the use of GPS-based DCT was never given consideration in Europe, as all countries surveyed recognize Bluetooth-based models as the only legally viable option. In some countries, such as Italy for example, technology experts did not rule out *a priori* the possibility of collecting limited amounts of geolocation data for DCT purposes, but this option never gained support in policy circles. The rationale, based on data protection by design, is that geolocation data is considered redundant to the aim of proximity tracing, since it contains more information than is necessary to notify users about contact with positive cases. However, this argument depends upon a specific view of DCT as a personal warning system, rather than a public health surveillance tool.

European policymakers and advisors however showed a lesser degree of consensus as to the best IT architecture for DCT systems. In the view of some stakeholders, the GDPR did not appear to pose a concrete constraint on specific technological options for DCT. Germany and the UK initially favored a centralized model, to later change to a decentralized one. France and Hungary (not reviewed) ultimately implemented centralized DCT, while remaining fully compliant with GDPR rules. Switzerland, while not a member of the European Union, is revising its Federal Act on Data Protection (FADP) in a way that will also ensure general alignment with the provisions of the GDPR, especially regarding the rights of data subjects. The newly approved law (expected to come into effect in 2022) endorses privacy by design and also by default.

The European model differs in meaningful ways from the DCT approaches adopted by Asian countries in the earliest phases of the pandemic. While it is not possible to speak of an "Asian model" due to the great diversity among DCT systems in Asian countries, it is evident that a more expansive approach characterizes DCT in countries such as China, Hong Kong, Singapore, South Korea, and Taiwan (see **Figure 1**).

DCT apps developed in China at the beginning of the pandemic became mandatory immediately (33). Hong Kong, Taiwan, and South Korea also deployed mandatory apps and wearable trackers for those living under quarantine, either due to testing positive for COVID-19 or returning from foreign travel (34–36). These apps record GPS geolocation data or use cell tower data to ensure that individuals remain in their homes while in quarantine, and ask the user to enter symptoms, in order to monitor the course of the disease. Taiwan for example used the quarantine DCT feature in combination with rigorous manual contact tracing, which successfully helped contain the spread of the disease (37). South Korea relied on more intrusive surveillance measures, including a number of system tracking citizens' movement, and interactive maps displaying locations visited by COVID-19 positive individuals (38). Singapore was one of the first countries worldwide to introduce a voluntary centralized digital contact tracing app called *TraceTogether*, which was later integrated with a check-in system (called *Safe Entry*) for entry into public spaces (mandatory from the beginning of January 2021) (39). A similar feature was adopted in October 2020 in Hong Kong, where the government is still debating whether the app will be made mandatory.

During the period examined in this study (March to October 2020) European DCT systems showed stability in their overall technical architecture. In all of the reviewed countries, participation in DCT was originally designed to be, and remained, entirely voluntary. Data collection remained limited to randomly generated and periodically deleted Bluetooth IDs. While organizational and technical improvements were implemented to streamline the uploading of positive test results, this process remained fully voluntary, with disclosure of test results possible only with explicit authorization by a DCT user. One partial exception is presented by the England & Wales app, which automatically uploads test results when a test is booked directly through the app.

We have observed an expansion of DCT app features beyond basic proximity tracing in European apps released or updated during the second wave of SARS-CoV-2 that swept through Europe during late summer 2020. Novel features include the capability to track symptoms (Ireland, France, England, and Wales), acquire more detailed epidemiological information about a given area (France, England, and Wales), check in at venues (England and Wales), order COVID-19 tests and access results (England and Wales), and count down the quarantine time (England and Wales).

Our study indicates that privacy preservation through state-of-the-art technological solutions and alignment with data protection laws is the key defining feature of the European approach to DCT. However, we also demonstrated how such an approach is evolving to incorporate novel technological capabilities beyond mere proximity tracing. These developments signal a shift from a strict interpretation of data minimization and purpose limitation, toward a more expansive approach to digital contact tracing in Europe. This evolutionary trajectory seems to reflect technological capacities already seen in Asian countries.

In Europe, the incorporation of novel capacities seems a response to two aims. On one hand, adding features can be viewed as a way to encourage users to download and use DCT apps by offering additional functionalities that users may find useful or interesting. Considering the relatively low level of uptake of DCT apps in European countries compared with the adoption rates needed to ensure effectiveness (40), adding new features may be seen as one way to deliver more personal utility to app users, thus incentivizing participation. On the other hand, novel features such as digital check-ins may increase the aggregate data available to public health authorities, expanding their capacity to monitor how the epidemic is evolving and how the population responds to containment measures. Furthermore, this feature is an ingenious way to integrate manual and digital contact tracing. Both manually and digitally collected information about the whereabouts of positive cases can contribute to map out risky locations. In turn, this information can be used to alert people about potential contacts with positive cases irrespective of whether the use the DCT app or not, thus extending the utility of the DCT app beyond the section of the population that is actually using it.

The panorama of European DCT systems is evolving also in other respects. In December 2020, the privacy policy of the *Corona-Warn-App* was updated, allowing users to record symptoms and retrieve test results (41). In France, the government is considering adding a check-in function to the *TousAntiCovid* app when reopening restaurants (42). These updates may prelude to further expansion of DCT app capabilities in the near future. In Italy for example, the possibility of using the *Immuni* app as a tool in the imminent vaccination campaign is being discussed. The app could evolve into a digital booking system for vaccination appointments, and could then be licensed to store a digital copy of the vaccination certificate for display to health authorities, for entry to designated places or activities (43).

The possible evolution of European DCT systems calls for careful scrutiny and appropriate oversight, especially with respect to GDPR provisions. It must be noted that novel features do not necessarily contravene the principle of data minimization, as they can still be based on minimum necessary data collection for data processing purposes. However, such new features expand the scope of DCT apps beyond the purpose of proximity tracing and warnings to individual users. The legally mandated safeguards regarding data collection and storage may therefore be insufficient to capture additional privacy risks linked to novel functionalities. In other words, data protection by design and by default may be inadequate to address the evolution of DCT systems. To be sure, DCT innovation does not necessarily create greater privacy risks. Such technological evolution should not be prevented, and both public health and ethical rationale support changes aimed at improving the effectiveness of DCT systems against the spread of the virus. Yet as the purpose of DCT apps expands to incorporate new capacities, privacy risks should be regularly reassessed. An adaptive governance approach to DCT seems best suited to regularly fine tuning governance structures and oversight mechanisms over time (44), thus capturing the

technical evolution of such systems and their ethical, legal and societal implications.

CONCLUSION

As they face subsequent epidemic waves, European countries are tasked with deploying all possible means to mitigate the spread of the virus and minimize the health-related, personal, economic, and social damage it has caused since early 2020. Digital methods offer valuable aid to contain this disaster. In the context of harsh measures and restrictions to individual freedom made necessary by the emergency, DCT is relatively more tolerable, especially in its European incarnation, which offers a comprehensive set of technical and legal safeguards against potential abuse of personal data. Nevertheless, the vast majority of European citizens have not downloaded national DCT apps, despite their ethical and technical robustness. Lack of trust regarding the privacy-preserving features of DCT systems as well as about governments' oversight capacity, may be have contributed to such relatively low figures. Other explanatory hypotheses include insufficient public education campaigns, lack of familiarity on the part of the public with the use of digital health in the context of a public health crisis, and lack of consensus on best practices to implement DCT (14).

In this study we reviewed DCT systems in a number of European countries. We highlighted the strong emphasis that all such systems place on privacy and data protection, their fully voluntary character, and their adoption of the same Bluetooth-based standards for proximity tracing. We noted that such ethical and technological commitment is enshrined in both centralized and decentralized DCT systems. Furthermore, we reported an emerging evolutionary trajectory resulting in the incorporation of novel technological features beyond mere contact tracing that are, to some extent, reminiscent of those already seen in Asia. However, additional policy efforts seem necessary to account for such developments, to gain public trust and to foster more widespread adoption of DCT as a valuable means for containing the effects of the SARS-CoV-2 pandemic.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

AB and AF contributed to data collection, analysis, manuscript drafting, and editing. EV contributed to writing and editing the manuscript. All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This manuscript was supported by the Swiss National Science Foundation under awards NRP75 (#407540_167223) and NRP77 (#407740_187356).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.660823/full#supplementary-material

REFERENCES

1. Johns Hopkins. *COVID-19 Map - Johns Hopkins Coronavirus Resource Center*. (2021). Available online at: https://coronavirus.jhu.edu/map.html (accessed January 27, 2021).
2. European Centre for Disease Prevention and Control. *COVID-19 Situation Update for the EU/EEA, as of Week 2 2021*. (2021). Available online at: https://www.ecdc.europa.eu/en/cases-2019-ncov-eueea (accessed May 14, 2021).
3. Han E, Tan MMJ, Turk E, Sridhar D, Leung GM, Shibuya K, et al. Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. *Lancet*. (2020) 396:1525–34. doi: 10.1016/S0140-6736(20)32007-9
4. EUR-Lex. *Guidance on Apps Supporting the Fight Against COVID 19 Pandemic in Relation to Data Protection (2020/C 124 I/01)*. (2020). Available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020XC0417%2808%29 (accessed January 27, 2021).
5. eHealth Network. *Interoperability Guidelines for Approved Contact Tracing Mobile Applications in the EU*. (2020). Available from: https://ec.europa.eu/health/sites/health/files/ehealth/docs/contacttracing_mobileapps_guidelines_en.pdf (accessed January 27, 2021).
6. NS TECH. *PEPP-PT vs DP-3T: The Coronavirus Contact Tracing Privacy Debate Kicks Up Another Gear*. (2020). Available online at: https://tech.newstatesman.com/security/pepp-pt-vs-dp-3t-the-coronavirus-contact-tracing-privacy-debate-kicks-up-another-gear (accessed January 27, 2021).
7. Reuters. *Germany Flips to Apple-Google Approach on Smartphone Contact Tracing*. (2020). Available online at: https://www.reuters.com/article/us-health-coronavirus-europe-tech-idUSKCN22807J (accessed January 27, 2021).
8. Medical Device Network. *The UK's contact-tracing apps: why the long wait?* (2020). Available from: https://www.medicaldevice-network.com/features/uk-contact-tracing-app/ (accessed January 27, 2021).
9. European Commission. *Mobile contact tracing apps in EU Member States*. (2020). Available online at: https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/travel-during-coronavirus-pandemic/mobile-contact-tracing-apps-eu-member-states_en (accessed January 27, 21)
10. Henrich Böll Stiftung. *Pan-European contact-tracing – a chance for more normality and freer movement | Heinrich Böll Stiftung | Brussels office - European Union*. (2020). Available online at: https://eu.boell.org/en/2020/11/16/pan-european-contact-tracing-chance-more-normality-and-freer-movement (accessed January 27, 21)
11. Ando H, Cousins R, Young C. Achieving saturation in thematic analysis: development and refinement of a codebook. *Compr Psychol*. (2014) 3:03.CP.3.4. doi: 10.2466/03.CP.3.4
12. Politico. *How Google and Apple Outflanked Governments in the Race to Build Coronavirus Apps*. (2020). Available online at: https://www.politico.eu/article/google-apple-coronavirus-app-privacy-uk-france-germany/ (accessed January 27, 2021)
13. Sharon T. Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. *Ethics Inf Technol*. (2020) 1–13. doi: 10.1007/s10676-020-09547-x
14. Fahey RA, Hino A. COVID-19, digital privacy, and the social limits on data-focused public health responses. *Int J*

Inf Manage. (2020) 55:102181. doi: 10.1016/j.ijinfomgt.2020.102181
15. Ministère des solidarités et de la santé. *TousAntiCovid: Réponses à vos Questions*. (2020). Available online at: https://web.archive.org/web/20201031233234/https://solidarites-sante.gouv.fr/soins-et-maladies/maladies/maladies-infectieuses/coronavirus/tousanticovid (accessed October 31, 2020).
16. Federal Statistical Office. *SwissCovid App Monitoring*. (2021). Available online at: https://www.experimental.bfs.admin.ch/expstat/en/home/innovative-methoden/swisscovid-app-monitoring.html (accessed January 27, 2021).
17. Immuni Italia. *The Numbers of Immuni*. (2021). Available online at: http://www.immuni.italia.it (accessed January 27, 2021).
18. Robert Koch Institut. *Kennzahalen zur Corona-Warn-App*. (2021). Available online at: https://www.coronawarn.app/assets/documents/2021-01-22-cwa-daten-fakten.pdf (accessed January 27, 2021).
19. The Connexion France. *France to Release 'More Interactive' StopCovid App*. (2020). Available online at: https://www.connexionfrance.com/French-news/France-is-to-release-more-interactive-StopCovid-app-on-October-22 (accessed January 27, 2021).
20. Santé.fr. *DepistageCovid*. (2020). Available online at: http://web.archive.org/web/20201031215938/https://sante.fr/recherche/trouver/DepistageCovid (accessed October 31, 2020).
21. GOV.UK. *DPIA Annex 2: High-Risk Postcode District User Journey Flow*. (2020). Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/920835/HighRiskPostcode_20200923.pdf (accessed January 27, 2021).
22. Ministère des solidarités et de la santé. *Mes Conseils Covid. Des Conseils Personnalisés Pour Agir Contre le Virus*. (2020). Available from: https://mesconseilscovid.sante.gouv.fr/ (accessed January 27, 2021).
23. Decomplix. *Decomplix Supports FOPH With SwissCovid App for Contact Tracing*. (2020). Available online at: https://decomplix.com/swisscovid-app-contact-tracing/ (accessed January 27, 2021).
24. NHS COVID-19 app support. *FAQ NHS Covid-App: Which Venues in England Should Display the Official NHS QR Code Poster?* (2020). Available online at: https://faq.covid19.nhs.uk/article/KA-01183/en-us?parentid=CAT-01043&rootid=CAT-01027 (accessed January 27, 2021).
25. NHS COVID-19 app support. *NHS COVID-19 App: Terms of Use*. (2020). Available online at: https://web.archive.org/web/20201030072744/https://covid19.nhs.uk/our-policies.html (accessed October 31, 2020).
26. Health Service Executive. *COVID Tracker App: Data Protection Information Notice*. (2020). Available online at: https://web.archive.org/web/20201004220235mp_/https://covidtracker.gov.ie/privacy-and-data/data-protection/#6 (accessed October 31, 2020)
27. Immuni Italia. *Immuni: Privacy Notice*. (2020). Available online at: https://web.archive.org/web/20200805213934/https://www.immuni.italia.it/app-pn.html (accessed October 31, 2020).
28. Swiss Federal Office of Public Health. *SwissCovid App: Data Protection Statement*. (2020). Available online at: https://web.archive.org/web/20201026031819/https://www.bag.admin.ch/dam/bag/en/dokumente/cc/kom/swisscovid-app-datenschutz.pdf.download.pdf/FOPH_SwissCovid_Data_Protection_Statement_24_June2020.pdf (accessed October 31, 2020).
29. NHS Scotland. *ProtectScotland: How We Use Your Data*. (2020). Available online at: https://web.archive.org/web/20201030192549if_/https://www.protect.scot/how-we-use-your-data (accessed October 31, 2020)
30. Robert Koch Institute. *Corona-Warn-App: Privacy Notice*. (2020). Available online at: https://web.archive.org/web/20201130083701/https://www.coronawarn.app/assets/documents/cwa-privacy-notice-1.3-en.pdf (accessed October 31, 2020).
31. CoronaMelder. *CoronaMelder: Privacy Statement*. (2020). Available online at: https://web.archive.org/web/20201015200129/https://coronamelder.nl/en/privacy (accessed October 31, 2020).
32. Pane J, Francisca RD, Verhamme KM, Orozco M, Viroux H, Rebollo I, et al. EU postmarket surveillance plans for medical devices. *Pharmacoepidemiol Drug Saf*. (2019) 28:1155–65. doi: 10.1002/pds.4859
33. CNN Business. *China is Fighting the Coronavirus With a Digital QR Code. Here's How It Works*. (2020). Available online at: https://www.cnn.com/2020/04/15/asia/china-coronavirus-qr-code-intl-hnk/index.html (accessed January 27, 2021).
34. Quartz. *Hong Kong is Using Tracker Wristbands to Geofence People Under Coronavirus Quarantine*. (2020). Available online at: https://qz.com/1822215/hong-kong-uses-tracking-wristbands-for-coronavirus-quarantine/ (accessed January 27, 2021).
35. MIT Technology Review. *South Korea is Watching Quarantined Citizens With a Smartphone App*. (2020). Available online at: https://www.technologyreview.com/2020/03/06/905459/coronavirus-south-korea-smartphone-app-quarantine/ (accessed January 27, 2021).
36. Quartz. *How Taiwan is Tracking 55,000 People Under Home Quarantine in Real Time*. (2020). Available online at: https://qz.com/1825997/taiwan-phone-tracking-system-monitors-55000-under-coronavirus-quarantine/ (accessed January 27, 2021).
37. Cheng H-Y, Jian S-W, Liu D-P, Ng T-C, Huang W-T, Lin H-H. Contact tracing assessment of COVID-19 transmission dynamics in Taiwan and risk at different exposure periods before and after symptom onset. *JAMA Intern Med*. (2020) 180:1156–63. doi: 10.1001/jamainternmed.2020.2020
38. CNN Business. *Coronavirus Mobile Apps are Surging in Popularity in South Korea*. (2020). Available online at: https://www.cnn.com/2020/02/28/tech/korea-coronavirus-tracking-apps/index.html (accessed January 27, 2021).
39. SafeEntry. *Places Where SafeEntry Must be Deployed*. (2021). Available online at: https://safeentry.gov.sg/ (accessed January 27, 2021).
40. Toussaert S. Upping uptake of COVID contact tracing apps. *Nat Hum Behav*. (2021) 5:183–4. doi: 10.1038/s41562-021-01048-1
41. German Federal Government. *Corona- Warn-App Privacy Notice*. (2020). Available online at: https://www.coronawarn.app/assets/documents/cwa-privacy-notice-en.pdf (accessed January 27, 2021).
42. La Depeche. *TousAntiCovid: de Nouvelles Fonctionnalités Annoncées Pour la Future "Réouverture des Restaurants"*. (2020). Available online at: https://www.ladepeche.fr/2020/11/23/tousanticovid-de-nouvelles-fonctionnalites-annoncees-pour-la-future-reouverture-des-restaurants-9216559.php (accessed January 27, 2021).
43. La Repubblica. *Immuni, la Seconda Vita: "Così la App Può Dialogare Con il Piano Vaccini"*. (2020). Available online at: https://rep.repubblica.it/pwa/generale/2020/12/18/news/immuni_la_seconda_vita_cosi_la_app_puo_dialogare_con_il_piano_vaccini_-278976650/ (accessed January 27, 2021).
44. Blasimme A, Vayena E. What's next for COVID-19 apps? Governance and oversight. *Science*. (2020) 370:760–2. doi: 10.1126/science.abd9006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 17 June 2021
doi: 10.3389/fdgth.2021.648585

COVID-19 in Brazil—Preliminary Analysis of Response Supported by Artificial Intelligence in Municipalities

Hugo M. P. Morales[1]*, *Murilo Guedes*[1,2], *Jennifer S. Silva*[3] and *Adriano Massuda*[4]

[1] Department of Research, Instituto Laura Fressatto, Curitiba, Brazil, [2] School of Medicine, Pontifícia Universidade Católica Do Paraná, Curitiba, Brazil, [3] Department of Customer Success, Instituto Laura Fressatto, Curitiba, Brazil, [4] Department of Administration, São Paulo School of Business Administration, Fundação Getulio Vargas, São Paulo, Brazil

The novel coronavirus disease (COVID-19) forced rapid adaptations in the way healthcare is delivered and coordinated by health systems. Brazil has a universal public health system (Sistema Unico de Saúde—SUS), being the main source of care for 75% of the population. Therefore, a saturation of the system was foreseen with the continuous increase of cases. The use of Artificial Intelligence (AI) to empower telehealth could help to tackle this by increasing a coordinated patient access to the health system. In the present study we describe a descriptive case report analyzing the use of Laura Digital Emergency Room—an AI-powered telehealth platform—in three different cities. It was computed around 130,000 interactions made by the chatbot and 24,162 patients completed the digital triage. Almost half (44.8%) of the patients were classified as having mild symptoms, 33.6% were classified as moderate and only 14.2% were classified as severe. The implementation of an AI-powered telehealth to increase accessibility while maintaining safety and leveraging value amid the unprecedent impact of the COVID-19 pandemic was feasible in Brazil and may reduce healthcare overload. New efforts to yield sustainability of affordable and scalable solutions are needed to truly leverage value in health care systems, particularly in the context of middle-low-income countries.

Keywords: COVID-19, artificial intelligence, chatbot, public health, access, care coordination

OPEN ACCESS

Edited by:
Björn Wolfgang Schuller, Imperial College London, United Kingdom

Reviewed by:
Niamh Lennox-Chhugani, International Foundation for Integrated Care (IFIC), United Kingdom
Sonu Bhaskar, Liverpool Hospital & South West Sydney Local Health District (SWSLHD), Australia

***Correspondence:**
Hugo M. P. Morales
hugo@laura-br.com

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *31 December 2020*
Accepted: *22 April 2021*
Published: *17 June 2021*

Citation:
Morales HMP, Guedes M, Silva JS and Massuda A (2021) COVID-19 in Brazil—Preliminary Analysis of Response Supported by Artificial Intelligence in Municipalities. Front. Digit. Health 3:648585. doi: 10.3389/fdgth.2021.648585

INTRODUCTION

The novel coronavirus disease (COVID-19) seriously challenged health systems in the world. Emerging at the end of 2019 in Wuhan (China), the outbreak caused by the virus SARS-CoV-2 was declared as a pandemic by the World Health Organization (WHO) on March 11, 2020 (1). As 14 April 2021, the WHO COVID-19 situation report informed there have been over 137 million cases and over 3 million deaths since the start of the pandemic. With 354,617 deaths by the same date, Brazil ranks as the second most affected country in the world (1).

To face the Pandemic, countries developed different strategies. Lessons from previous coronavirus epidemics—SARS in 2003 and MERS in 2015—enabled governments in east and southeast Asia to take rapid and efficient health-systems responses to control SARS-CoV-2 transmission. In contrast, European and American countries which have the capacity to deal with an influenza epidemic were not prepared to face a coronavirus one. The late and disorganized response to the COVID-19 collapsed well-structured health systems and cost thousands of lives.

In this challenging context, innovative strategies have been described in order to support the COVID-19 response and the high pressure over the health system. Technological strategies played a pivotal role, especially in augmenting and organizing access to healthcare systems. Symptom checkers, which are self-directed risk assessments tools, has been shown to improve live quality and survival in oncology patients (2). Its use for COVID-19 is currently increasing and has been used to identify symptomatic patients or potential severe cases with the main objective to triage those individuals who would benefit from clinical assessment and/or management into further care (3). A recent French study showed that a national self-triage web application contributed to the alleviation of calls to the emergency calls centers and may be used for predicting increasing burden of hospital (4).

Another broadly adopted technology against COVID-19 are chatbots, which are Artificially intelligent (AI) based conversational agents. These applications can be accessed via a website or social media messaging platforms, such as WhatsApp and Facebook. Besides self-triage and personal risk-assessments, chatbot has been used to disseminate health information and knowledge, tracking COVID-19 symptoms and health aspects and monitor exposure and notifications (5). These novel technologies are excellent benefactors for facilitating progress in healthcare as they enable better accessibility and personalization for consumers and efficiency for healthcare providers and public health officials (6). Despite the increasing use, a recent review found only 9 studies, the majority of the studies coming from North America (5).

Brazil has a universal public health system (Sistema Unico de Saude—SUS), which is the main source of care for 75% of the population. Decentralized to the municipal level, health departments in the 5,570 municipalities largely handle the management of SUS, including co-financing, coordination of health programs, and delivery of health services. Despite the achievements in improving health outcomes since the implementation of SUS, deficiencies in its structure and governance persist, maintaining wide inequalities in distribution of health services across the country (7). These deficiencies have been exacerbated since fiscal austerity and health policy changes were implemented in 2016.

In the response to COVID-19, the absence of the federal government to assume its role of authority and leadership in combating the pandemic led municipal and state managers to build their own strategies to deal with the effects of the pandemic on their populations (8). Local responses varied in form, intensity, duration, and start and end times, to some extent associated with political alignments (9).

As a strategy to support health system management capacity at a municipal level, we developed a platform powered by Artificial Intelligence, called Laura Digital ER. By providing an innovative way for patients to have easy access to medical information via telehealth, the platform offers a solution to triage patients, to monitor the evolution of symptoms and to provide a credible and current source of information for patients. As a response to the pandemic, public-private partnerships, such as Inter-American Development Bank (IADB), BID Lab and Instituto Votorantim sponsored the implementation of the platform in municipal health systems.

Few studies have assessed the impact of deploying digital solutions at the municipal level. To fill this gap we describe three Brazilian municipality-level case reports setting out the early implementation of a digital triage and monitoring service which included the use of a chatbot utilizing algorithmic decision-making.

MATERIALS AND METHODS

Study Design

This is an observational, descriptive study that aims to summarize the early experience of the deployment of a platform consisting on a decision-tree algorithm to perform COVID-19 risk assessment in patients seeking emergency room evaluation for acute symptoms amid the pandemic combined with a chatbot to both disseminate health informations and track patient's symptoms evolution when needed. The algorithm's primary goal was to ensure safety while increasing accessibility and decreasing healthcare usage of resources during the COVID-19 pandemic. The means by which the tool was assumed to deliver its value were through a Kanban framework, in which healthcare facilities could dynamically allocate resources according to demand. This descriptive study reports on data collected from July to October 2020 from three large municipalities in Brazil, namely Curitiba (almost 2 million inhabitants), São Bernardo do Campo (almost 900,000 inhabitants), and Catanduva (roughly 100,000 inhabitants).

Algorithm Description

Employing a Natural Language Processing (NLP) framework, the Laura Digital ER platform was developed in April 2020 in response to the challenges created by the COVID-19 pandemic. Machine Learning algorithm was applied in order to ameliorate user experience in the process of answering basic questions and doubts about the pandemic (i.e., social distancing and preventive measures, such as how to wash hands, how and when to use masks). The chatbot was powered by RASA's NLP technology and the content was updated weekly by Laura's Team of health professionals, according to an analysis of the most commons question asked by the users. A powered chatbot was integrated into municipalities' websites to allow patient's interaction. Additionally, extensions to social media platforms and popular communication apps were ensured. The major objective of this chatbot is to provide the population with current and accurate information about COVID-19, including social distancing and preventive measures following CDC guidelines.

If a symptom was detected by the algorithm, the patient was redirected to a Triage Web Page in which a consent form was applied and, if signed, the individual was directed to a symptom checker. Once symptoms were verified, patients were classified into three main subgroups, following criteria defined by local authorities: mild, moderate or severe. Patients classified as mild disease were followed by a virtual non-human clinical pathway, in which every 72 h the algorithm interacted with individual through a messaging platform. Patients falling into

moderate or severe disease categories were prompted to phone calls or teleconsultations with a healthcare provider. All patients, stratified by their severity, could be monitored by a healthcare provider in a platform allowed by the interface. In such a way, the process deployed through the algorithm ensured prioritization of individuals with more severe COVID-19 presentations while a continuous oversight by a healthcare provider was afforded, potentially allowing dynamic and optimal healthcare utilization using a Kanban framework.

Ethical Aspects

This study follows the principles of the Declaration of Helsinki. All patients signed consent forms before data were extracted. Confidentiality and data privacy were ensured for all patients. The algorithm described here follows Brazilian local regulations for data protection n (Lei Geral de Proteção de Dados). Data were anonymized before any analyses were performed.

Data Presentation

In this descriptive study, we report data on the algorithm utilization during a period during the COVID-19 pandemic in Brazil. Demographic data from the three cities in which the algorithm was deployed was extracted from the Brazilian Institute of Geography and Statistics (IBGE).

We calculated the number of patients who used the platform, the proportion of individuals stratified by COVID-19 severity as assessed by the algorithm, the absolute number of patient interactions with the algorithm and number of patients exclusively followed by non-human interactions from July to October 2020.

RESULTS

The three cities in which the algorithm was deployed were Curitiba, São Bernardo do Campo, and Catanduva (**Table 1**). All these regions already had a call center organized to allow triage of patients with COVID-19 symptoms, in which healthcare providers who were not able to work in the frontline (e.g., elder of with comorbidities) were relocated to work in such a system. These healthcare providers were responsible to assess and determine workflow for patients with COVID-19 symptoms, such as evaluating potential severity and recommending in-person visits in the emergency room. Over the course of the COVID-19 pandemic, the total amount of calls gradually increased, with up to 1,000 calls per day in some of these regions. This rapid increase raised concerns regarding potential burden and lack of adequate follow-up for individuals with varying risks of time, as reflected in COVID-19 natural history in which more severe features ensue around 7–10 days. In this framework, healthcare works called individuals 3–5 times per day over a 14-day period of follow-up.

With the deployment of the AI algorithm in these cities, the workflow was changed to ensure greater patient interaction with the healthcare system by reducing the total amount of time spent by healthcare professionals in low-risk patients, such as those with non-severe COVID-19 presentations. During the period described in this study (July to October 2020), the NLP-powered chatbot interacted with 22,418 persons, performing more than 133,700 interactions (i.e., answered questions). Most common questions answered by the chatbot were related to COVID-19 test costs and locations to perform it, issues relating to transmission after close contact with a SARS-CoV2 infected individual and the average time to recover from COVID-19.

Risk assessment was categorized in four domains: gray (asymptomatic patient), blue (mild symptoms), yellow (moderate symptoms), and red (severe symptoms) based on interactions with health departments in each city. From the 24,162 interactions resulting in symptom evaluating by the NLP-algorithm, up to 45% were classified as mild disease and 14% as severe (**Figure 1**). Complete distribution of severity assessment is described in **Table 1**.

In terms of patient population, most individuals were older than 60 years old. The most commonly reported symptoms were cough (55%), sore throat (42.8%), and headache (39.8%). Symptom distribution is summarized in **Table 2**, stratified per location.

TABLE 1 | Stratification of alert severity per municipality.

Metric	Curitiba	Catanduva	Sao Bernardo do Campo	Total
City population	1,948,626	122,497	838,936	2,910,059
Chatbot interactions	96,626	31,732	9,420	133,778
Persons who interacted with the platform	14,646	6,262	1,510	22,418
Total alerts	17,498	5,903	761	24,162
Gray (asymptomatic)	1,406	258	92	1,756
Blue (mild)	7,077	3,185	345	10,607
Yellow (moderate)	5,637	2,092	219	7,948
Red (severe)	2,880	369	105	3,353

DISCUSSION

In this descriptive study, we aimed to summarize the initial phase of the first initiative in Brazil to provide a scalable tool to ensure better healthcare access to individuals while ensuring safety and contributing to better operations amid the COVID-19 pandemic. As far as we know, this is the first description of a telemedicine platform powered by an AI-chatbot to tackle COVID-19 issues as a decentralized and adjustable tool in Brazilian municipalities.

In the present study, almost half of patients who answered the symptoms checker were classified as having mild symptoms, therefore they could be safely followed without an in-person visit to the emergency room, as defined in COVID-19 management guidelines. Importantly, these patients classified as mildly symptomatic remained on the non-human clinical pathway, having their clinical status assessed by direct conversation with the chatbot. Since it is established a policy of making 3–4 phone calls to check the

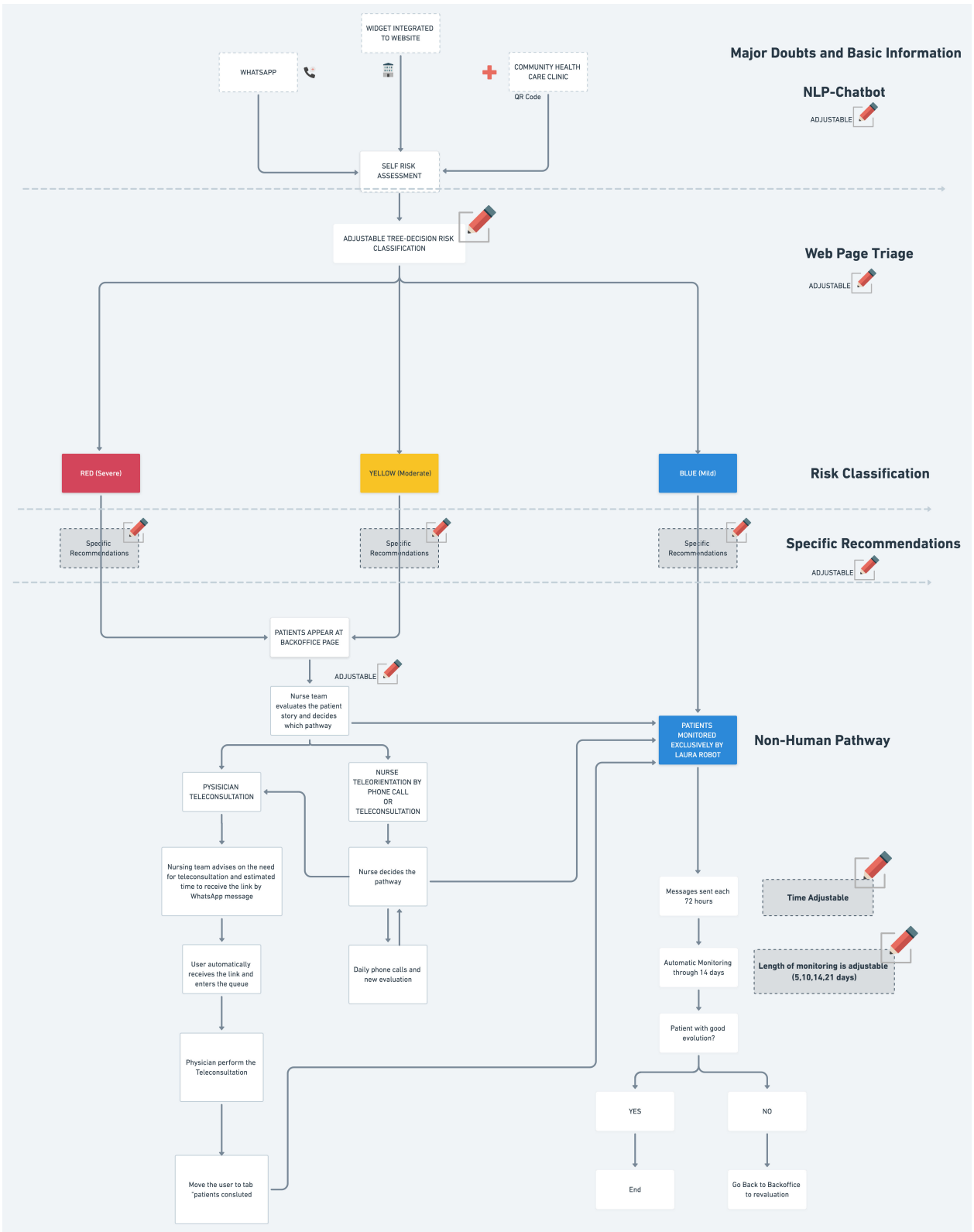


FIGURE 1 | AI-chatbot for risk COVID-19 severity risk assessment. The flow-chart describes all steps in the interaction between patients and the algorithm, as well as the indicated interventions according to the COVID-19 severity based on symptoms assessment evaluation by the technology.

TABLE 2 | Distribution of the most reported symptoms.

Symptom	Curitiba	Catanduva	Sao Bernardo do Campo	Total	
Total	**19,390**	**9,090**	**907**	**29,387**	
Cough *n* (%)	4,381 (22.6)	3,216 (35.4)	263 (29)	7,860	26.7%
Sore throat *n* (%)	4,304 (22.2)	2,452 (27)	238 (26.2)	6,994	23.8%
Headache *n* (%)	4,593 (23.7)	2,276 (25)	177 (19.5)	7,046	24.0%
Myalgia *n* (%)	3,122 (16.1)	1,913 (21)	196 (21.6)	5,231	17.8%
Fever *n* (%)	3,471 (17.9)	2,048 (22.5)	156 (17.2)	5,675	19.3%
Anosmia *n* (%)	2,244 (11.6)	1,009 (11.1)	95 (10.5)	3,348	11.4%
Diarrhea *n* (%)	1,569 (8.1)	627 (6.9)	49 (5.4)	2,245	7.6%
Dyspneia *n* (%)	1,412 (7.3)	1,497 (16.5)	88 (9.7)	2,997	10.2%

One patient can have more than one symptom. n = 24,162.

clinical evolution of these mild cases, we can assume that there was a saving of more than 60,000 phone calls after the algorithm's implementation. Therefore, even in this initial description, the technology is thought to have optimized human resources to more demanding areas in a severely overloaded system.

The sustainability of implementing such innovative tools in a country, such as Brazil requires several structural and cultural adaptations, as highlighted in this study (5). The algorithm was conceived to be used in the public health system in Brazil, leveraging available resources in specific regions. The three regions we describe in this study belong to a group of a relatively higher economic status in Brazil, in which technology use is broader and thereby patients can have better accessibility to technological solutions. The complexity of a continental country, such as Brazil implies the widespread adoption of innovative solutions in healthcare will need to address many challenges (10). Other than technological resources, sustainability in terms of financial structures for such projects are key limiting points to unleash innovative solutions (10). For the initiative presented here, a public and private collaborative effort was leveraged to implement a potential solution for better resource allocation amid the COVID-19 pandemic, as described. Continuing efforts to allow affordability in the processes of implementation of new, innovative tools in public-private partnerships are necessary in order to guarantee a stable environment in which continuous quality improvement can be developed. In a country in which the majority of individuals are covered by a public health system, sustainable innovation to allow wider access and to lever equity in healthcare implies coordinated efforts between multiple stakeholders. The coordination between these stakeholders needs to ensure a well-designed implementation plan, as well as rigorous metrics for performance and safety evaluation over time, which can be challenging in a system without a unified medical record (10–12). Indeed, in the public system in Brazil, the lack of a centralized medical record to allow follow-up of individuals who interacted with the algorithm has been a main limitation so far in terms of establishing outcomes related to technological interventions, such as the one described in this paper.

The total number of symptom checker evaluations was considered relatively low taking into account the total population assessed and it varied widely among the cities. The municipality of Catanduva (5,111 people who used the tool/100,000 inhabitants) presented higher usage rates in comparison to Curitiba (751/100,000 inhabitants) and São Bernardo do Campo (179/100,00 inhabitants). A possible explanation is the municipality's strong commitment to the platform in organizational and, mainly, marketing terms. Sao Bernardo do Campo, which have the lowest usage in the studied period, implemented a strong advertising campaign to use the tool with great results. During the months of October, November, and December (up to the 30th), there were 41,406 persons who interacted with the chatbot and more than 2,500 patients completed the self-risk assessment. The lack of national guidelines for the adoption of digital solutions in healthcare could be one of the main factors for this disparity because the project prioritization was totally dependent on each city's commitment. The Ministry of Health have implemented a centralized chatbot solution for COVID-19 but unfortunately it was not accompanied by national guidelines and there was no possibility of customization or adapting to the existing local workflow of the cities. We believe that a national plan for adopting digital healthcare is essential to transformation in healthcare toward leveraging optimal value to all stakeholders, but essentially and more importantly, to the end users, whose outcomes are at the core of all endeavors.

This study has several limitations. We provided a descriptive study on the early experience of the implementation of an AI-based solution to improve health care utilization amid the COVID-19 pandemic in the public health system in Brazil. We were unable to report health outcomes associated to the platform, due to the early stage of the report and because of the lack of uniform and centralized medical records data in the country. Additionally, although our symptom check was intuitive and validated in interactions with medical directors and healthcare leaders in each municipality in Brazil, our questionnaires were not previously validated or standardized.

Notably, this early report has several strengths. In spite of all limitations, we were able to describe a successful implementation of an AI-powered chatbot to assess COVID-19 related risks and to provide optimal healthcare utilization through flexible, on-demand and accessible structures. These findings corroborate with others from Austria, Switzerland and the USA, suggesting that the use of this technology may decrease the healthcare system overload and increase access to healthcare system (13–15). All efforts summarized here were validated along with health authorities in the country, therefore safety was ensured based on rigorous evaluation by state departments in each city where the algorithm was implemented. Therefore, we provide the first description of a successful algorithm to improve healthcare utilization amid the COVID-19 pandemic, leveraging value while minimizing burden during unprecedent times in a heterogeneous and widely affected country as Brazil. This report highlights that well-structured public-private initiatives, coupled with innovative

endeavors and sustainable incentives, yields value in healthcare even among low-middle income countries. Technology and telehealth can leverage existing structures toward greater healthcare management, improving accessibility through safe and optimal solutions.

CONCLUSIONS

The development and deployment of an AI powered telehealth aiming for greater access and optimizing care coordination is feasible, advisable, and may reduce healthcare overload. In order to succeed, the platform should be adaptable to local needs, including being capable of changing the decision-tree algorithm. A national policy for digital healthcare transformation could guide and enhance adoption of innovative technologies at municipalities-levels.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

HM and AM conceived the paper. JSS collected data. All authors made substantial contributions in writing the paper and contributed to the article and approved the submitted version.

REFERENCES

1. World Health Organisation. *WHO Announces COVID-19 Outbreak a Pandemic*. Geneva: World Health Organisation (2020).
2. Denis F, Basch E, Septans A-L, Bennouna J, Urban T, Dueck AC, et al. Two-year survival comparing web-based symptom monitoring vs routine surveillance following treatment for lung cancer. *JAMA.* (2019) 321:306–7. doi: 10.1001/jama.2018.18085
3. Mansab F, Bhatti S, Goyal D. Performance of national COVID-19 'symptom checkers': a comparative case simulation study. *BMJ Health Care Inform.* (2021) 28:187. doi: 10.1136/bmjhci-2020-100187
4. Galmiche S, Rahbe E, Fontanet A, Dinh A, Bénézit F, Lescure FX, et al. Implementation of a self-triage web application for suspected COVID-19 and its impact on emergency call centers: observational study. *J Med Internet Res.* (2020) 22:e22924. doi: 10.2196/22924
5. Almalki M, Azeez F. Health chatbots for fighting COVID-19: a scoping review. *Acta Inform Med.* (2020) 28:241–7. doi: 10.5455/aim.2020.28.241-247
6. Tudor Car L, Dhinagaran DA, Kyaw BM, Kowatsch T, Joty S, Theng YL, et al. Conversational agents in health care: scoping review and conceptual analysis. *J Med Internet Res.* (2020) 22:e17158. doi: 10.2196/17158
7. Brazil, Presidency of the Republic, General Secretariat, Law N. 13.709.
8. Brazil MoH.
9. Haase CB, Bearman M, Brodersen J, Hoeyer K, Risor T. 'You should see a doctor', said the robot: reflections on a digital diagnostic device in a pandemic age. *Scand J Public Health.* (2021) 49:33–6. doi: 10.1177/14034948209 80268
10. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health.* (2020) 20:1193. doi: 10.1186/s12889-020-09301-4
11. Lai L, Wittbold KA, Dadabhoy FZ, Sato R, Landman AB, Schwamm LH, et al. Digital triage: novel strategies for population health management in response to the COVID-19 pandemic. *Healthc (Amst).* (2020) 8:100493. doi: 10.1016/j.hjdsi.2020.100493
12. Major P, Stefura T, Wysocki M, Malczak P, Rzepa A, Proczko-Stepaniak M, et al. Impact of SARS-CoV-2 pandemic on bariatric care in Poland: results of national survey. *BMC Surg.* (2020) 20:314. doi: 10.1186/s12893-020-00990-7
13. Morse KE, Ostberg NP, Jones VG, Chan AS. Use characteristics and triage acuity of a digital symptom checker in a large integrated health system: population-based descriptive study. *J Med Internet Res.* (2020) 22:e20549. doi: 10.2196/20549
14. Munsch N, Martin A, Gruarin S, Nateqi J, Abdarahmane I, Weingartner-Ortner R, et al. Diagnostic accuracy of web-based COVID-19 symptom checkers: comparison study. *J Med Internet Res.* (2020) 22:e21299. doi: 10.2196/21299
15. Hautz WE, Exadaktylos A, Sauter TC. Online forward triage during the COVID-19 outbreak. *Emerg Med J.* (2021) 38:106. doi: 10.1136/emermed-2020-209792

Conflict of Interest: HM, JS, and MG are employees of Laura, the company that developed Laura Digital ER.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

OPINION
published: 25 June 2021
doi: 10.3389/fdgth.2021.594124

Digital COVID Credentials: An Implementation Process

Mayssam Nehme[1]*, *Laurent Kaiser*[2,3], *Philippe Gillet*[4], *Philippe Thevoz*[4], *Silvia Stringhini*[1,2,5] and *Idris Guessous*[1,2]

[1] Division of Primary Care Medicine, Geneva University Hospitals, Geneva, Switzerland, [2] Faculty of Medicine, University of Geneva, Geneva, Switzerland, [3] Division of Infectious Diseases, Geneva University Hospitals, Geneva, Switzerland, [4] SICPA, Prilly, Switzerland, [5] University Centre for General Medicine and Public Health, University of Lausanne, Lausanne, Switzerland

Keywords: digital, blockchain, COVID-19, decentralized governance, free movement, immunity, certificate, vaccination

Initial public health responses to the COVID-19 pandemic have focused on non-pharmaceutical interventions including stringent physical distancing measures, lockdowns, and restriction to free movement. This comes at significant costs however, both economically and socially (1, 2). As authorities begin to ease existing measures, governments are looking into specific alternatives to lockdown, such as phased mobilization of the economy (3), less stringent physical distancing measures, or immunity passports that would determine individual access or restrictions (4). Immunity passports vs. certificates differ in the rights related to their use and their issuing authority. Immunity passports have been cautioned against by the WHO and at international levels (5, 6) citing a lack of reliable interpretability of the presence or absence of COVID-19 antibodies, as well as ethical risks (7). With the advent of vaccines, these risks are potentially mitigated while other risks arise such as universal access to vaccination, and the debate around immunity passports is once again justifiably revived (8). COVID credentials could be an answer to facilitate some of the currently difficult scenarios in society and everyday life (travel, large gatherings, etc.). The need for a non-falsifiable solution is of utmost importance, especially with reports of fraud increasingly emerging (9).

Reflecting on the digital aspects of such a solution is important to ensure the implementation of adequate safeguards, display the right amount of information and use digital health systems to society's advantage. The European Union has recently published open source material detailing a potential trust framework and technical specificities that would be used in establishing a European Union Digital COVID Certificate that would be uniform and interoperable (10). COVID credentials taking into account vaccination, serology, PCR testing, and self-reported symptoms can employ algorithms to certify an individual's most recent COVID-related status. Certification would take into account results from pre-certified laboratories and pre-certified vaccination centers only, thus decreasing the prospect of false positive results and individuals inadvertently foregoing protective measures, putting themselves and others at risk (11). In addition, information could further assist individuals in making the right decisions and can also provide reminders to get tested or retested, vaccinated or re-vaccinated; which would also accommodate continually evolving aspects of the current COVID-19 pandemic and virus response. An example is setting reminders for individuals who received a vaccination to receive a booster shot, depending on the duration of the immune response (once defined), but also for individuals who received a specific vaccine to follow specific measures if a new variant turned out resistant to that vaccine. The presence of symptoms should also be part of the algorithm and could determine the need for fast-track testing or the implementation of isolation measures.

Here, we propose a very practical decentralized secured digital solution (**Figure 1**). The solution is securing the original data provided from a certified vaccination center, a certified laboratory or testing center. A digital security seal protects and guarantees the integrity of the data to be secured, through an unforgeable mathematical link between the hash of the data and the seal. To

OPEN ACCESS

Edited by:
Constantinos S. Pattichis, University of Cyprus, Cyprus

Reviewed by:
Gary Matkin, University of California, Irvine, United States
Joshua Coyne, University of Memphis, United States

***Correspondence:**
Mayssam Nehme
mayssam.nehme@hcuge.ch

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *25 August 2020*
Accepted: *04 June 2021*
Published: *25 June 2021*

Citation:
Nehme M, Kaiser L, Gillet P, Thevoz P, Stringhini S and Guessous I (2021) Digital COVID Credentials: An Implementation Process. Front. Digit. Health 3:594124. doi: 10.3389/fdgth.2021.594124

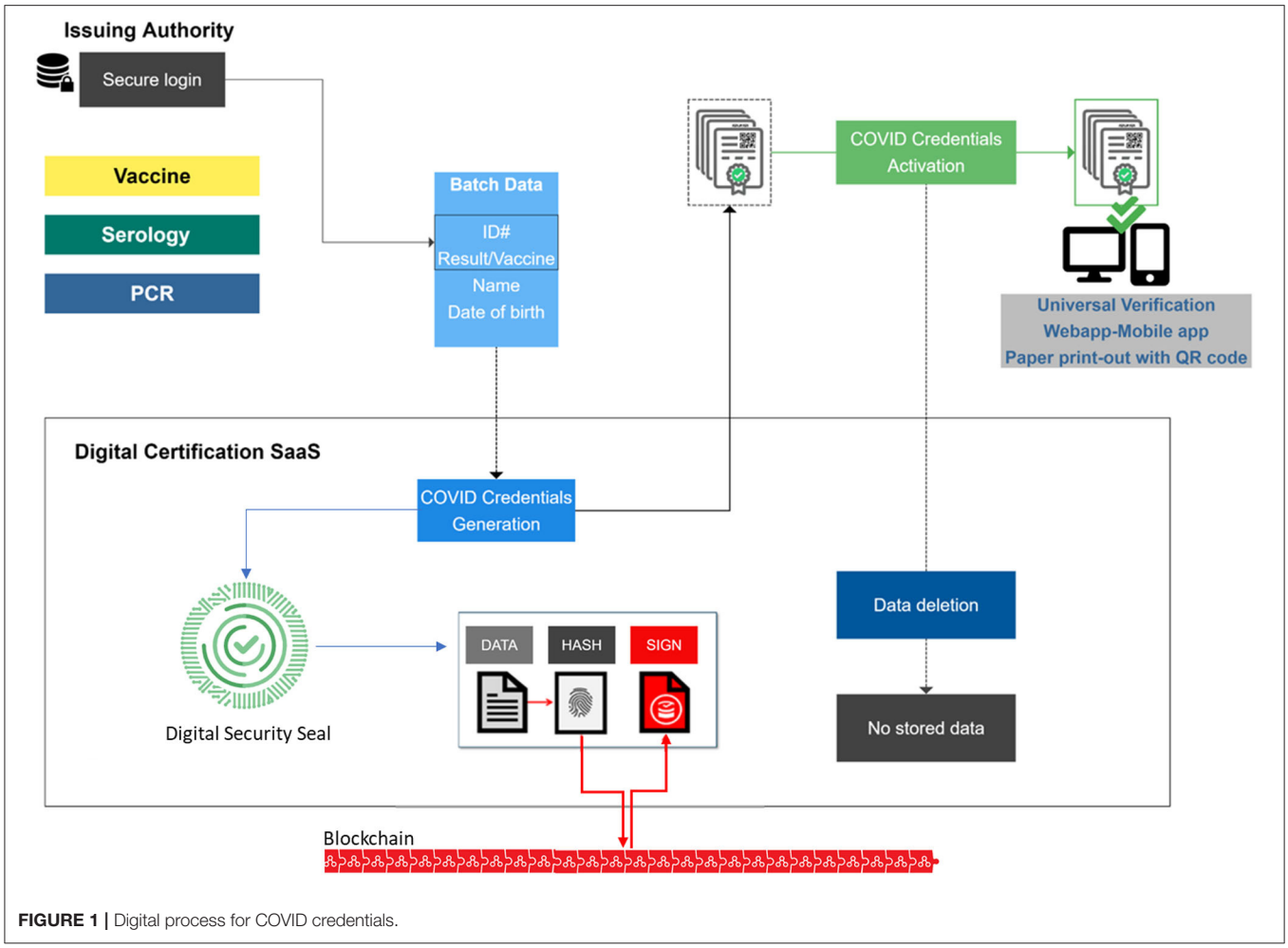


FIGURE 1 | Digital process for COVID credentials.

ensure the immutability, the digital security seal is timestamped on a blockchain. As the digital security seal contains only metadata, it guarantees privacy protection of the holder with personal and medical data only on the credential (QR code) itself. The blockchain is acting only as a secure "Trust Anchor," in the form of an undisputable timestamp. Thus, no data are ever exposed or stored on the blockchain. Unlike the European Union Digital COVID Certificate, this solution does not need to handle the complex management of cryptographic keys, thus avoiding the risk of having some of these keys being compromised or stolen.

The individual presents him or herself to the certified vaccination or testing center. His or her identity is verified (using an official ID) prior to testing, vaccination or determination of recovery. The information on vaccination status, or the test result or the recovery status is secured as COVID credentials. The COVID credentials consist of a certificate, secured by its QR code, containing the name of the person (previously verified), the medical information (vaccine, test result, recovery etc.) as well as the name and identification of the issuing authority. The COVID credentials are issued in batches (in the form of secured QR-codes) by the issuing authority (certified vaccination or testing center) using a Digital Certification SaaS. This Digital Certification SaaS is accessible online by the issuing authority only, with a secure login. Once the QR codes are generated, they are activated by the issuing authority and all information used to issue the credentials is deleted from the Digital Certification SaaS. This process reinforces the decentralized approach by removing the need for a central database that could be easily targeted, and safeguards are important to ensure only certified testing and vaccination centers are capable of issuing such credentials while respecting data protection and privacy regulations. The data remains in the issuing authority medical records (like any other laboratory or vaccination result and for a defined period of time if needed), enabling individuals to have their credentials re-issued when necessary (lost QR code for example). The secure QR code can be stored on an individual's phone or delivered as a print-out to reduce the digital divide. The secure QR code reduces the risk of forgery or tampering, and can be universally verifiable via a web-based portal or a mobile app, without the need to access a database containing personal or medical information. The individual has access to the web-based portal to verify his or her own credentials. The individual can choose to disclose information in specific contexts (airport control, access to a

venue, nursing home, etc.) and interpretation of the result ensues, based on the context-related requirements (for example negative PCR within the last 72 h to enter a specific country vs. 24 h etc.). Individuals can selectively decide who to show this information to and how many identifying details to reveal depending on the context. Selective disclosure and decentralized information can further assist in preserving privacy and confidentiality. A digitally secured solution can also reduce the risk of loss, identity theft and forgery while ensuring accessibility, bidirectional information and the possibility to revoke the credentials or update the expiration information when needed. In order to ensure more universal access, a paper version of the digital certificate and QR code is also available. This paper version provides the same level of security as the digital one, as its content is certified via the QR code which can be universally verified with the same security as the digital credential. QR code verification acting as a digital unforgeable stamp remains a cornerstone of certification in order to avoid any fraud or falsification. The QR code verification can also be performed offline as the verification keys (digital security seals) can be periodically replicated locally on the verification device when connected.

CONCLUSION

Immunity passports, certificates or COVID credentials will be increasingly at the forefront of medical and public policy discussions in the months and years to come. The adequate safeguards around a digital COVID credential should be discussed, and a non-falsifiable solution should be implemented especially if rights are linked to such credentials. The solution presented here provides a decentralized approach to databases as well as a secure certification process in line with the European Commission's recommendations (10). This solution also provides a secure approach, ensuring the integrity and validity of the information and respecting data protection regulations on privacy and confidentiality. The question of COVID credentials, now at the forefront, should be also be addressed at a policy level involving discussions between medical and public health actors, technology experts, ethicists and governing bodies. It is also of utmost importance to actively engage the public on the options and opinions connected with this issue in order to assess their trust and needs when proposing a digital health solution.

AUTHOR CONTRIBUTIONS

MN, PG, PT, LK, SS, and IG contributed to the writing of the manuscript. MN, PT, and IG contributed to the figures. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Edmond J. SAFRA Foundation for clinical research in internal medicine.

REFERENCES

1. OECD. *Evaluating the Initial Impact of COVID-19 Containment Measures on Economic Activity*. OECD (2020). Available online at: http://www.oecd.org/coronavirus/policy-responses/evaluating-the-initial-impact-of-covid-19-containment-measures-on-economic-activity-b1f6b68b/ (accessed July 18, 2020).
2. Correia S, Luck S, Verner E. *Pandemics Depress the Economy, Public Health Interventions do Not: Evidence From the 1918 Flu*. (2020). Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3561560 (accessed July 27, 2020).
3. Edmond Safra Center for Ethics, Harvard University. *Roadmap to Pandemic Resilience*. (2020). Available online at: https://ethics.harvard.edu/covid-roadmap (accessed July 18, 2020).
4. Persad G, Emanuel EJ. The ethics of COVID-19 immunity-based licenses ("immunity passports"). *JAMA*. (2020) 323:2241–2. doi: 10.1001/jama.2020.8102
5. WHO. *"Immunity Passports" in the Context of COVID-19 Scientific Brief*. WHO (2020). Available online at: https://www.who.int/publications-detail/immunitypassports-in-the-context-of-covid-19 / (accessed April 30, 2020).
6. National COVID-19 Science Task Force (NCS-TF) ELSI report. *Ethical, Legal, and Social Issues Associated With "Serological Passports"*. (2020). Available online at: https://ncs-tf.ch/en/policy-briefs (accessed June 15, 2020).
7. Olivarius K. Immunity, capital, and power in antebellum New Orleans. *Am Hist Rev*. (2019) 124:425–55. doi: 10.1093/ahr/rhz176
8. Hall MA, Studdert DM. "Vaccine passport" certification - policy and ethical considerations. *N Engl J Med*. (2021). doi: 10.1056/NEJMp2104289. [Epub ahead of print].
9. Europol. *Europol Warning on the Illicit Sale of False Negative COVID-19 Test Certificates*. Available online at: https://www.europol.europa.eu/newsroom/news/europol-warning-illicit-sale-of-false-negative-covid-19-test-certificates
10. European Commission. *EU Digital COVID Certificate*. Available online at: https://ec.europa.eu/health/ehealth/covid-19_en (accessed April 21, 2021).
11. Phelan AL. COVID-19 immunity passports and vaccination certificates: scientific, equitable, and legal challenges. *Lancet Lond Engl*. (2020) 395:1595–8. doi: 10.1016/S0140-6736(20)31034-5

Conflict of Interest: PG and PT has a patent WO2020011447 pending, and a patent WO2020030382 pending. SICPA has developed the CERTUS digital solution certificates with digital seal technology.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 06 July 2021
doi: 10.3389/fdgth.2021.686720

Trends in COVID-19 Publications: Streamlining Research Using NLP and LDA

Akash Gupta[1], *Shrey Aeron*[2], *Anjali Agrawal*[3] *and Himanshu Gupta*[4*]

[1] Department of Engineering, University of Cambridge, Cambridge, United Kingdom, [2] Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA, United States, [3] Harmony School of Innovation – Sugar Land (High School), Sugar Land, TX, United States, [4] Valley Health System, Ridgewood, NJ, United States

Background: Research publications related to the novel coronavirus disease COVID-19 are rapidly increasing. However, current online literature hubs, even with artificial intelligence, are limited in identifying the complexity of COVID-19 research topics. We developed a comprehensive Latent Dirichlet Allocation (LDA) model with 25 topics using natural language processing (NLP) techniques on PubMed® research articles about "COVID." We propose a novel methodology to develop and visualise temporal trends, and improve existing online literature hubs.

Our results for temporal evolution demonstrate interesting trends, for example, the prominence of "Mental Health" and "Socioeconomic Impact" increased, "Genome Sequence" decreased, and "Epidemiology" remained relatively constant. Applying our methodology to LitCovid, a literature hub from the National Center for Biotechnology Information, we improved the breadth and depth of research topics by subdividing their pre-existing categories. Our topic model demonstrates that research on "masks" and "Personal Protective Equipment (PPE)" is skewed toward clinical applications with a lack of population-based epidemiological research.

Keywords: natural language processing, latent dirichlet allocation, COVID-19, trends, LitCovid, topic model, Pubmed

OPEN ACCESS

Edited by:
Pradeep Nair, Central University of Himachal Pradesh, India

Reviewed by:
Viktoriya Semeshenko, Instituto Interdisciplinario de Economía Política de Buenos Aires (IIEP), Argentina
Monika Semwal, Nanyang Technological University, Singapore

***Correspondence:**
Himanshu Gupta
gupthi@valleyhealth.com

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *27 March 2021*
Accepted: *26 May 2021*
Published: *06 July 2021*

Citation:
Gupta A, Aeron S, Agrawal A and Gupta H (2021) Trends in COVID-19 Publications: Streamlining Research Using NLP and LDA. Front. Digit. Health 3:686720. doi: 10.3389/fdgth.2021.686720

INTRODUCTION

The COVID-19 outbreak was officially declared a pandemic by the World Health Organization in March 2020 (1). As the number of COVID-19 cases and deaths has increased, so has the research. Searching "COVID" in PubMed®'s database gives a list of over 32,000 unfiltered publications (as of July 2020). Due to the overwhelming stream of papers, there is now an urgent need for tools to automate the categorical organisation of research. More importantly, to sufficiently address COVID-19 and future pandemics, it is necessary to streamline the research and development process by allowing for quick identification of research areas that are either gaining popularity or lacking adequate research.

Latent Dirichlet Allocation (LDA) is an unsupervised topic modelling technique used to learn hidden topics within a corpus (2). It assumes topics are a soft clustering of words and outputs two probability distributions: a distribution of topics in the corpus, and distributions of words across each topic. Currently, LDA, with the aid of natural language processing (NLP) methodologies, has been used to investigate the response to government policies (3), analyse public sentiment on social media (4) and the news (5), and understand general

research hotspots in publications (6) as well as global trends (7). Most of these topic models either use a small number of topics or group a large number of topics into overarching themes, which eases comprehension at first glance. However, for a more in-depth analysis, a better understanding of the complexity of topics within each theme is required, which cannot be captured by searching for that topic in a literature repository using a simple query.

The National Center for Biotechnology Information (NCBI) has developed LitCovid, a central repository for curated COVID-19 research (8). The hub uses a combination of human and machine-learning methods to provide

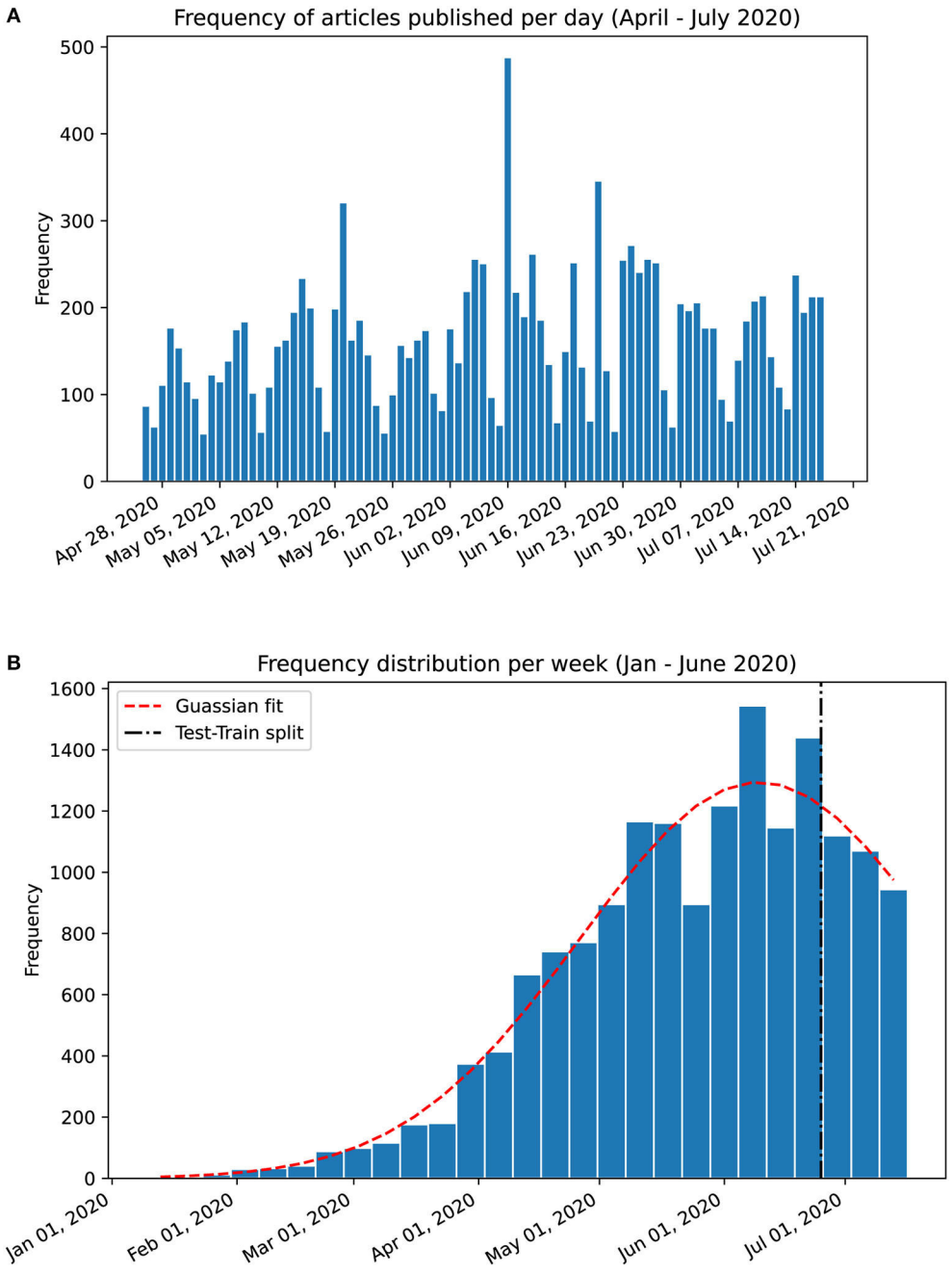


FIGURE 1 | Frequency charts of abstracts published in PubMed®. **(A)** Per day frequency shows a weekly pattern. **(B)** Per week frequency from January 17th to July 17th 2020 with a gaussian curve fitting (shown in red dashed line). The training set was split from the testing set at the 81% mark on June 22nd 2020. The weekly release pattern informed our decision to analyse trends on a per-week basis.

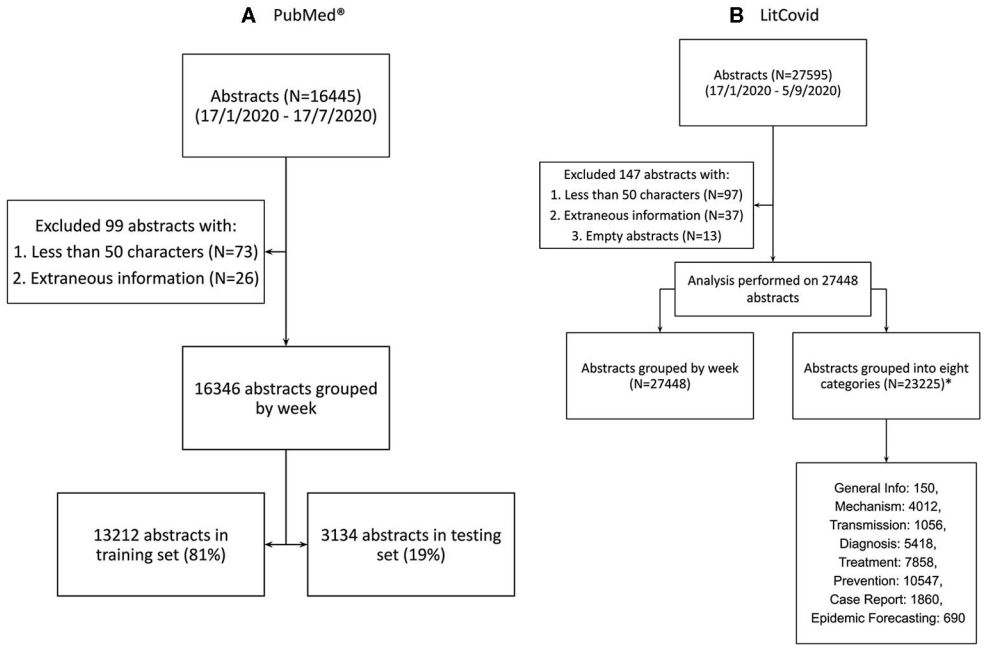


FIGURE 2 | Data sourcing and splitting. **(A)** Abstracts from PubMed® using the search term "COVID" **(B)** Abstracts from LitCovid Dataset using the search term "coronavirus," "ncov", "cov," "2019-nCoV," "COVID-19," and "SARS-CoV-2" (14). *Note that 4,223 LitCovid abstracts were not associated with any LitCovid category. Note that the total number of abstracts in all listed LitCovid categories is greater than the number of unique abstracts because an abstract can be placed in more than one category.

publications categorised by eight topics alongside temporal and geographic distributions. Although very useful, the defined categories are overly broad, limiting rapid assimilation of the ongoing research.

Here, we describe and implement an NLP methodology using LDA to identify topics in COVID-19 research. We then visualise and evaluate temporal trends to recognise the dominant and under-represented areas of research. We also apply our methodology to the LitCovid dataset and further subdivide their categories into topics. Although our results focus on COVID-19, we believe our method is generalisable and sustainable for evaluating rapidly evolving research fields.

METHODS

Data Sourcing

Our dataset is extracted using Entrez Programming Utilities (9) from the PubMed® Application Programming Interface (API) through the BioPython package. The API returns a list of metadata on all documents retrieved in a search query, which was the term "COVID." For each abstract, the PubMed® Document ID (PMID), date of publication on PubMed®, and abstract are stored.

A similar process is repeated for LitCovid abstracts. LitCovid (8), aided by machine-learning methods, manually assigns the articles into the eight categories of "Mechanism," "Transmission," "Diagnosis," "Treatment," "Prevention," "Case Report," "Forecasting," and "General." We use the NCBI Coronavirus API to extract the PMIDs corresponding to abstracts in each LitCovid category. We perform the same data cleansing process on the LitCovid abstracts as with the PubMed® abstracts.

For significantly large corpus sizes with a variety of journals, abstracts produce very similar topics compared to those produced by using the full text (10). Hence, we decided to use abstracts (3) due to the further benefits of much-reduced computation and free accessibility, aiding in reproducibility. A small number of abstracts are excluded from the corpus ($N = 99$) for the following reasons: those with <50 characters ($N = 73$), and others with a sentence including "This corrects the article DOI [...]" ($N = 26$).

Overview of the LDA Model

LDA is a standard topic modelling algorithm that learns the hidden topic structure within a corpus (2). Each topic has a weight within the corpus and is represented by a set of related words. We used Gensim's Python implementation of LDA to develop a topic model using a bag-of-words representation of the pre-processed training set. Further details about the NLP methodology and optimisation process are explained in the **Supplementary Material**. In brief, we set an upper limit on the number of topics to 50 and optimised hyperparameters based on the coherence value. Additionally, we selected a random seed

TABLE 1 | Topic description and summary about COVID-19 produced by LDA model.

Topic	Top 10 Most Relevant Words	Topic summary
1	Care, pandemic, service, telemedicine, health_care, need, resource, challenge, telehealth, healthcare	Health care, Telemedicine
2	Number, case, model, country, estimate, epidemic, estimated, data, daily, rate	Epidemiology
3	Case, day, chest, pneumonia, lesion, symptom, consolidation, fever, group, showed	Pulmonary
4	China, outbreak, epidemic, world_health, prevention_control, case, disease, novel_coronavirus, public_health, january	Disease Outbreak
5	Review, article, pandemic, research, vaccine, current, paper, scientific, literature, evidence	Scientific Development
6	Surgery, surgical, procedure, surgeon, hospital, dental, ppe, pandemic, emergency, staff	Surgery
7	Mortality, study, compared, included, outcome, risk_factor, higher, group, hospitalized, severe	Clinical Outcomes
8	Participant, survey, anxiety, questionnaire, mental_health, respondent, psychological, stress, fear, perceived	Mental Health
9	Disease, cardiovascular, acute_respiratory, cardiac, distress_syndrome, severe, ards, cytokine_storm, syndrome, cardiovascular_disease	Cardiovascular
10	Treatment, trial, clinical_trial, hydroxychloroquine, drug, study, remdesivir, hcq, tocilizumab, therapy	Clinical Trial
11	Cell, ace2, expression, receptor, sarscov2, tissue, human, virus, lung, pathway	Pathogenic Mechanism
12	Test, assay, testing, detection, sample, sarscov2, positive, specimen, sensitivity, antibody	Diagnosis
13	Child, symptom, pediatric, neurological, report, infection, infant, case, sarscov2, reported	Paediatrics
14	Social, crisis, health, economic, pandemic, impact, policy, consequence, psycinfo_database, right_reserved	Socio-economic Impact
15	Cancer, icu, intensive_care, treatment, ventilation, mechanical_ventilation, requiring, therapy, lung_cancer, ecmo	Oncology
16	Recommendation, risk, guideline, consensus, healthcare_worker, guidance, management, ibd, transplant, expert	Guidelines
17	Protein, sarscov2, compound, drug, target, vaccine, spike_protein, binding, inhibitor, epitope	Virus Structure
18	Virus, sequence, genome, human, bat, sarscov2, mutation, coronaviruses, genetic, animal	Genomic Sequence
19	Resident, person, county, older_adult, household, state, united_state, black, population, among	Demographics
20	Level, coagulation, thrombosis, ddimer, severe, elevated, crp, platelet, coagulopathy, aki	Haematology
21	Social_medium, information, public, video, news, medium, tweet, hand, misinformation, India	Communication
22	Use, vitamin, medication, drug, arb, angiotensin, angiotensin_receptor, ace_inhibitor, blocker, reninangiotensin_system	Existing Treatments
23	Liver, respiratory, skin, coinfection, ultrasound, gastrointestinal_symptom, diarrhea, imaging, symptom, fecal	Gastroenterology
24	Mask, aerosol, device, droplet, blood, particle, app, airborne, filter, mobile	Airborne transmission protection
25	Pregnant_woman, pregnancy, woman, maternal, pregnant, birth, neonatal, delivery, suspected, radiology_department	Pregnancy

After optimising the LDA model, 25 topics were produced. The topics are listed in descending order of prevalence in the corpus. The top 10 words (middle column) are chosen based on relevance to the topic, with lambda = 0.6 (15). The topic summary (last column) is based on the consensus of the authors after reviewing the top 20 words and the top 10 abstracts.

to use in the training process to ensure replicable results. We implemented the methodology in Python.

Temporal Evolution of Topics

To evaluate the temporal trends, we propose a novel method, which is applied to both PubMed® and LitCovid abstracts to produce an intuitive visualisation of the weekly temporal evolution of topic proportions. Analysis on a day-by-day basis can result in too much noise; instead, the documents are grouped by week, which is further justified by the apparent weekly release pattern of papers as shown in **Figure 1A**. However, since the weeks in January 2020 contain a substantially low number of papers, they are grouped into a single month. To evaluate how well the temporal evolution can predict future releases, we assign the weeks into training and testing sets with an approximate ratio of 8:2 (which is shown by the test-train split in **Figure 1B**).

For each week, in both the training and testing sets, all abstracts are combined into a single document. Applying the LDA model to these documents returns an estimate of the proportions of each topic that week. These results are graphed to produce a temporal topic evolution that is compared between PubMed® and LitCovid. We also create a heatmap to visualise all topics in the corpus in a concise chart.

Subdivision of LitCovid Categories

For each of the eight LitCovid categories, all abstracts are combined into a single document. Applying our model to the documents returns the proportions of each LDA topic in each LitCovid category. Statistical analysis (as described in the next section) is used to evaluate the categories and subdivide them into corresponding LDA topics.

Statistical Analysis

We incorporate tools from the Gensim Package (11) to perform statistical analyses on the topic model. We use perplexity (how surprised a model is to a sample) and topic difference to track the model convergence. The final performance of the model is evaluated using a coherence score (12).

Since LDA is a probabilistic model (2), the methodology used to capture temporal trends, as well as the subdivision of LitCovid categories is statistical. Applying the LDA topic model to a document returns a proportion for each topic, which is then used for further analysis.

Due to the non-linear nature of the temporal trends, a qualitative account of the trend fitting between the testing and training sets is discussed. Temporal trends of PubMed® and LitCovid publications are compared using Normalized Euclidean Distance, and topics are compared to each other

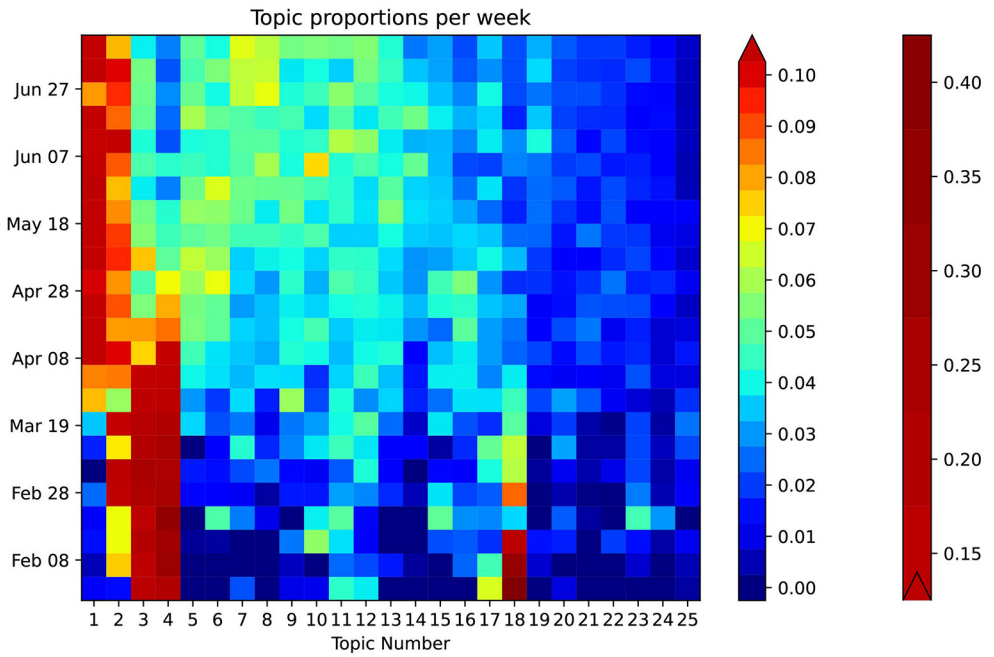


FIGURE 3 | Heatmap showing weekly temporal topic trends in PubMed®. The heatmap is enhanced using a bilinear scale for an intuitive visual representation of the temporal trends of COVID-19 topics. The topics are sorted in descending order of proportion on the horizontal axis. Exemplary trends can be observed: Topic 1 increases around April, while, simultaneously, Topic 4 decreases. Certain topics such as 24 and 25 show a consistently low topic proportion throughout.

using Jaccard Distance, a comparison between disjoint terms in two models (13).

LitCovid categories are compared against each other as well as the PubMed® corpus using Hellinger Distance (H):

$$\mathbf{H(P,Q)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2}}}\sqrt{\sum_{\mathbf{i=1}}^{\mathbf{k}}\left(\sqrt{\mathbf{p_i}}-\sqrt{\mathbf{q_i}}\right)^{\mathbf{2}}}$$

Where k is the number of topics; P and Q are the topic proportions returned by applying our LDA model to each LitCovid category.

RESULTS

The PubMed® API returned 16,445 abstracts with the search query "COVID" (as of July 17th, 2020). We exclude 73 abstracts with <50 characters and 26 abstracts with extraneous information, resulting in 16,346 abstracts in the corpus from January 17th to July 17th, 2020. The distribution of these follows a gaussian curve with mean on June 9th, 2020 (**Figure 1B**). The corpus is split (81:19) as per section Temporal Evolution of Topics: 13,212 abstracts for training and 3,134 abstracts for testing. The cutoff date for the training set is June 22nd (inclusive).

LitCovid generates 27,595 abstracts (as of September 8th, 2020) from the entirety of their dataset. We exclude 13 empty abstracts, 97 abstracts with <50 characters, and 37 abstracts with extraneous information. The final number of abstracts is 27,448, ranging from January 17th to September 5th 2020. 4,223 LitCovid abstracts could not be associated with any category and 23,225 are associated with at least one of the eight categories. The distribution of these categories is in **Figure 2**.

The LDA model is trained on 13,212 abstracts to produce a topic model with 25 topics. The final perplexity is 185.6, and the coherence score is 0.526. The average normalised Jaccard distance between the topics is 0.923 (STD 0.048) with a minimum of 0.701 between Topic 23 ("Gastroenterology") and Topic 13 ("Paediatrics"). A list of all the topics is displayed in **Table 1** alongside the top 10 relevant words and our summary interpretation based on the top words and abstracts associated with each topic. The topic of "Health care, telemedicine" has the highest proportion of 9.4%, whereas "pregnancy" has the lowest proportion of 1.0%.

Figure 3 provides a general breakdown of how topic proportion changes over the timeframe. The weekly topic proportions for the PubMed® test data were all within the bounds of the training data and are plotted as the top three rows in the heatmap, which show hardly any visible discrepancies. Most topics start at some high/low proportion, then trend in the opposite direction and begin to level off in April 2020. Others follow a relatively steady change with time. The highest topic proportion of 41% is during January 2020 in Topic 18, "genomic sequence." Topic 4, "disease outbreak," has the highest mean weekly topic proportion of 13.5% (STD 10.5). Topic 24, "airborne transmission protection," has the lowest mean of 0.98% (STD 0.66). This relatively lower topic proportion, as also illustrated in **Figure 4A**, was striking, because the usage of masks is a

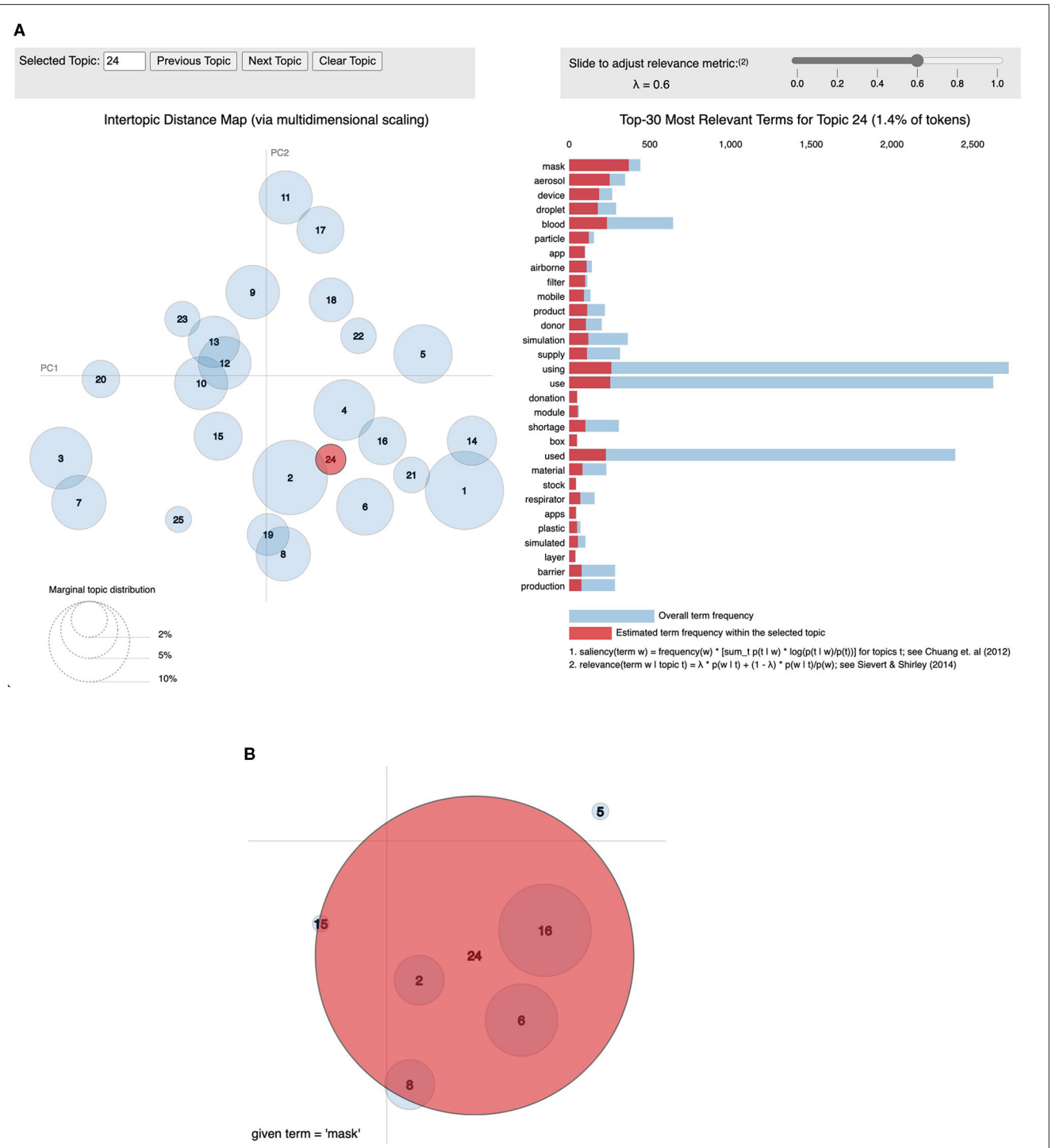


FIGURE 4 | Visualization of topics using pyLDAvis. This is a visualisation produced using the pyLDAvis module (15). **(A)** The left shows an inter-topic distance map created using classical multidimensional scaling. The right shows the words in Topic 24 ordered by relevance set to 0.6 (15). **(B)** The word "mask" has the highest weight in Topic 24 (2.3%) when compared to others: Topic 16 (0.078%), Topic 6 (0.033%), Topic 2 (.0090%), Topic 8 (0.018%).

highly emphasized guideline (16). We performed a *post hoc* analysis by analysing the proportion of the term "mask" (as shown in **Figure 4B**) in different topics: Topic 24 ("airborne transmission protection") 2.3%, Topic 16 ("guidelines") 0.078%, Topic 6 ("surgery") 0.033%, Topic 2 ("epidemiology") 0.0090%, Topic 8 ("mental Health") 0.018%.

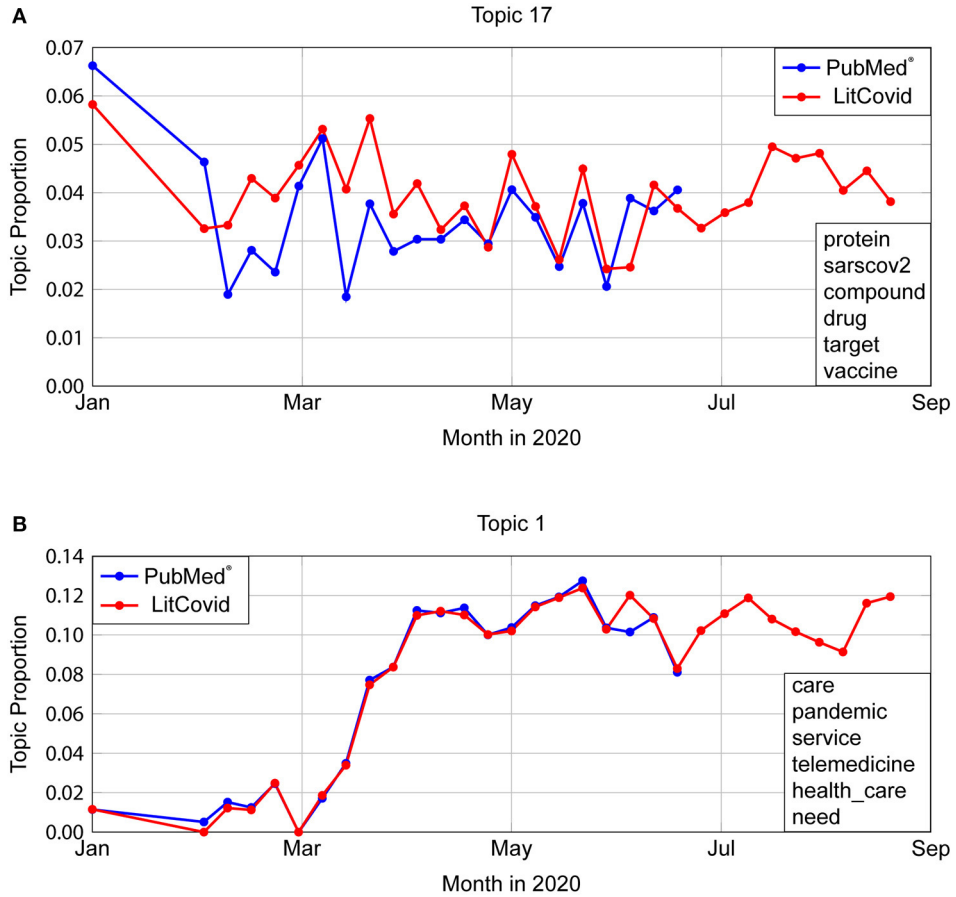


FIGURE 5 | Comparison of temporal trends in PubMed® and LitCovid. **(A)** Topic 17 depicts the most dissimilar topic trends, with a Normalised Euclidean Distance (NED) of 0.0604. Even though LitCovid's dataset was also taken from PubMed®, the mild variation can be attributed to differences in search queries (described in **Figure 2**). **(B)** Topic 1 represents the most similar trends (NED = 0.0303).

Applying our LDA model to LitCovid abstracts shows that temporal topic proportions have a very strong agreement with an average Normalized Euclidean Distance (NED) of 0.0303 (STD 0.0128). The most dissimilar looking graph is shown in **Figure 5A**, which is Topic 17, "Virus Structure" with NED 0.0604. Even though LitCovid's dataset was also taken from PubMed®, the mild variation can be attributed to differences in search queries (as described in **Figure 2**). The most similar trends occur for Topic 1 ("Health care, telemedicine") with NED = 0.0303, as show in **Figure 5B**.

In **Figure 6**, applying our model to the LitCovid categories shows that the LitCovid category "Epidemic forecasting" is most similar to LDA Topic 2, "Epidemiology" as it has a topic proportion of 74%. The most intuitive category to subdivide is "General Info"; our topic model split it into "Scientific Development" (21%), "Health Care, Telemedicine" (16%), "Socio-economic Impact" (15%), "Disease Outbreak" (14%), "Genomic Sequence" (9%) and "Communication" (6%). Another notable subdivision is of the "Case Report" category into the relevant medical fields of "Paediatrics" (20%), "Pulmonary" (18%), "Oncology" (13%), and "Cardiovascular" (12%), as well as identifying the under-represented topics "Haematology" (5%) and "Gastroenterology" (6%).

As shown in **Figure 7**, the LitCovid categories with the most overlap are "Treatment" and "Mechanism" with H = 0.160, whereas the least overlap is for "Epidemic forecasting" and "Case Report" with H = 0.572. When compared to the PubMed® corpus, "Prevention" has the lowest Hellinger distance of H = 0.116, suggesting that the "Prevention" category may be too broad. Conversely, "Epidemic Forecasting" has the greatest Hellinger distance (H = 0.476), suggesting that it is a well-defined category.

DISCUSSION

We provide a generalisable NLP methodology to extract abstracts from PubMed®, create an optimised LDA topic model, and visualise temporal trends. Applying our model to LitCovid abstracts helps to identify trending and under-represented research areas as well as subdivide its categories into relevant

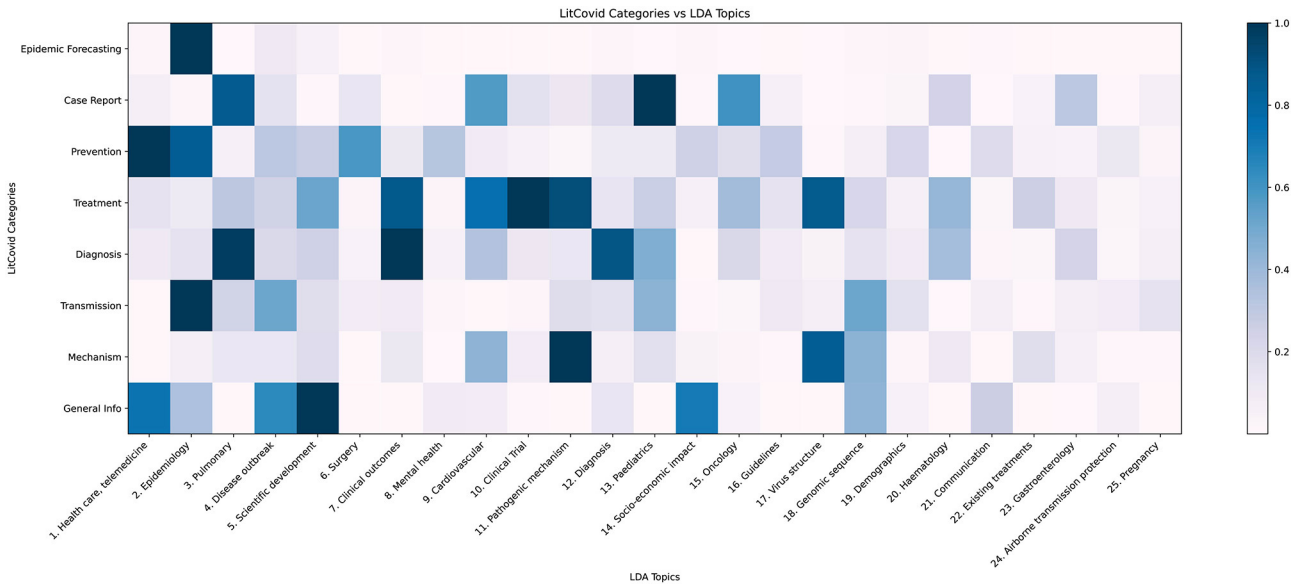


FIGURE 6 | LDA topic distribution across LitCovid categories. LitCovid divides COVID-19 research into eight categories. Our LDA topic model, when applied to each category, returns a topic distribution, which is normalised (dividing by the largest proportion in the category) and plotted as the heatmap.

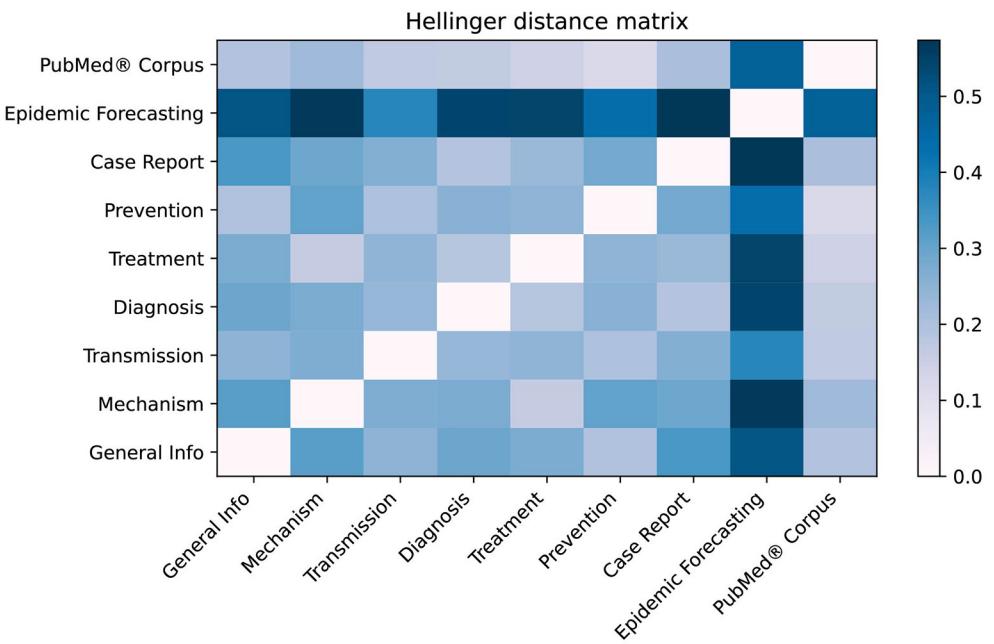


FIGURE 7 | Hellinger distance matrix comparing LitCovid categories and PubMed®. The Hellinger distance, which is used to quantify the dissimilarity between two probability distributions, was used to compare LitCovid categories against each other as well as the PubMed® corpus. This distance was calculated using the distribution of the 25 topics shown in **Figure 6**. For example, the most dissimilar categories are "Epidemic forecasting" and "Case Report" (H = 0.572).

topics. We find that an important topic related to masks and PPE is under-represented in population-based research.

The optimised LDA model developed in this study identifies 25 topics with no significant overlap, and all relate to the COVID-19 pandemic. Our model is trained and tested on a much larger number of documents compared to previous studies, and we believe our topic interpretation is improved using NLP techniques such as the use of bi-grams and ordering terms by relevance rather than frequency (further details are in **Supplementary Material**). Therefore, our model is comprehensive, and our analysis is in-depth. Topics not previously identified include "Pathogenic Mechanism,"

"Cardiovascular," and "Demographics" (Topics 9, 11, and 19). Previous topic models developed using research publications provide broad categories and do not sufficiently characterise the social science aspect of COVID-19 (6, 7, 17, 18). In contrast, our model covers this field in the topics "Mental Health," "Socio-economic Impact," and "Communication" (Topics 8, 14, and 21, respectively), representing 10.2% of the PubMed® test corpus.

We developed a customised heatmap shown in **Figure 3** representing topic trends on a holistic scale. This visualisation allows for intuitive understanding compared to the use of classic line graphs, which can be cumbersome to review, especially for a larger number of topics. From January to March 2020, Topic 18, "genomic sequence" has an unusually high topic proportion given its lower proportion in the overall corpus. This trend correlates well with our expectation as it was initially a hot topic in social media (19). The heatmap also clearly shows that Topic 24 ("Airborne Transmission Protection") has a lack of adequate representation. As a direct application of our topic model, we zoomed into the word "mask" (**Figure 4B**), demonstrating its representation in very few topics, including Topic 24, 16 ("guidelines"), and 6 ("surgery"). Moreover, the relative weight of "mask" within Topic 16 is two orders of magnitude less than Topic 24. Considering the existing relationship between masks and the spread of COVID-19 (20), one would expect "mask" to have a greater representation in "Epidemiology." The lack of research in this area suggests that a critical topic for future research is the usage of masks in a public context (21).

Analysis of temporal trends in **Figure 3** includes both the training and test data from PubMed®, which does not show many discrepancies between the two, suggesting that both our model and methodology are applicable on other relevant datasets. Existing research papers (4–6) limit their analysis by focusing only on the corpus they trained on. However, we perform further testing and analysis on related articles outside of our training set. Applying our methodology to LitCovid shows that the temporal trends are largely the same, further validating our hypothesis. When they differ, such as in **Figure 5A**, the shape remains similar, with disagreeing values, which can be attributed to the difference in search queries when curating the datasets we analyse.

PubMed® is a central repository for biomedical research literature, containing search tools to identify publications based on search queries. It lacks comprehensive tools to analyse publications for topic identification. LitCovid uses a combination of human curation and machine learning to provide eight broad categories for COVID-19 research. By applying our NLP methodology, we find that many of these general categories can be subdivided into multiple relevant topics, which provides more comprehensive insights for future specialised research. For example, the inherently broad category of "General Info" is split into more specific topics including socio-economic impact, communication, and telemedicine. Another subdivision we believe to be useful is for the category of "case report"; our model split it into six medical fields, including paediatrics, cardiovascular, and gastroenterology.

Figure 7 shows there is measurable overlap between LitCovid's categories, some of which is to be expected because they all contain the overarching theme of COVID-19. Interestingly, our model shows that the most heterogeneous category, "Prevention," comprises the topics "Health Care," "Telemedicine," "Epidemiology," "Surgical Procedures," and "Mental Health." One would not expect some of these, but rather it would be more reasonable to have "Guidelines" as a core topic. The closest categories identified in **Figure 7** are "Treatment" and "Mechanism," which is reasonable because effective preventative treatment should be inhibiting a mechanism. Closer analysis using **Figure 6** shows that "Mechanism" is almost a subset of "Treatment," with the key difference being that "genomic sequence" is a strong topic in "Mechanism," but not in "Treatment." We believe that these categories can be ambiguous to a researcher, and our methodology would significantly improve online literature hubs through more specific subdivisions.

As we performed searches in LitCovid, we noticed that abstracts were sometimes not relevant to the query even though a relevance option is provided. Improved comparative relevance of a topic within a research paper can add value for researchers as it distils a large mass of research papers by the strength of their major topics. Future research can utilise our methodology to enhance online literature hubs by ordering documents based on the proportion of topics.

Although the LDA topic model can be updated with new research, further investigation on hyperparameter adjustment will be required to identify new topics. However, our methodology is simple to re-run and provide up-to-date trends. To evaluate the temporal trends, we use the date of publication, because we found that not all articles had the research date. Although the publication date does not accurately reflect the time when the research was performed, for our purpose, our temporal visualisation disseminates the trending topics from the under-represented ones, which we believe is more crucial to the researchers. Another potential limitation in our dataset is the exclusive use of English abstracts. Since LDA is essentially a clustering algorithm, if more than one language is used, the topic model will likely return duplicate topics in different languages, which is redundant for our purpose. Instead, since many research papers written in other languages provide abstracts translated into English (22), this further justifies our use of abstracts instead of papers. Even though a significant number of publications are not analysed due their lack of abstracts, we believe this would not affect our identification of topics, given that it is reasonable to assume that this is random and not biased against any topic.

NLP techniques applied with LDA topic modelling results in a comprehensive topic listing and identification of important temporal trends. Our methodology has the potential to complement existing literature hubs. We identify topics for further research, such as studies on masks that may be of significant public interest.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

SA, AA, and AG: data generation. AA, SA, AG, and HG: data analysis and manuscript. AG, SA, AA, and HG: statistical analysis. HG: final manuscript approval. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank Aditya Gupta, MEng, for his valuable suggestions related to machine learning, the Entrez API, and topic classification. Our research paper appears on a preprint server from Social Science Research Network (SSRN) (23).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.686720/full#supplementary-material

REFERENCES

1. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. *Acta Biomed.* (2020) 91:157–60. doi: 10.23750/abm.v91i1.9397
2. Campbell JC, Hindle A, Stroulia E. Latent dirichlet allocation. *The Art and Science of Analyzing Software Data*. Elsevier (2015). p. 139–59.
3. Debnath R, Bardhan R. India nudges to contain COVID-19 pandemic: a reactive public policy analysis using machine-learning based topic modelling. *PLoS ONE*. (2020) 5:e0238972. doi: 10.1371/journal.pone.0238972
4. Ordun C, Purushotham S, Raff E. *Exploratory Analysis of Covid-19 Tweets using Topic Modeling, UMAP, and DiGraphs*. (2020). Available online at: http://arxiv.org/abs/2005.03082 (accessed September 15, 2020).
5. Liu Q, https://pubmed.ncbi.nlm.nih.gov/?term=Zheng$+$Z&cauthor_id=32302966Zheng Z, Zheng J, https://pubmed.ncbi.nlm.nih.gov/?term=Chen$+$Q&cauthor_id=32302966Chen Q, Liu G, https://pubmed.ncbi.nlm.nih.gov/?term=Chen$+$S&cauthor_id=32302966Chen S, et al. Health communication through news media during the early stage of the COVID-19 outbreak in China: digital topic modeling approach. *J Med Internet Res.* (2020) 22:e19118. doi: 10.2196/19118
6. Dong M, Cao X, Liang M, Li L, Liu G, Liang H. Understand research hotspots surrounding COVID-19 and other coronavirus infections using topic modeling. *medRxiv*. (2020). doi: 10.1101/2020.03.26.20044164
7. Tran BX, https://www.ncbi.nlm.nih.gov/pubmed/?term=Ha%20GH%5BAuthor%5D&cauthor=true&cauthor_uid=32521776Ha GH, Nguyen LH, https://www.ncbi.nlm.nih.gov/pubmed/?term=Vu%20GT%5BAuthor%5D&cauthor=true&cauthor_uid=32521776Vu GT, Hoang MT, https://www.ncbi.nlm.nih.gov/pubmed/?term=Le%20HT%5BAuthor%5D&cauthor=true&cauthor_uid=32521776Le HT, et al. Studies of novel coronavirus disease 19 (COVID-19) pandemic: a global analysis of literature. *Int J Environ Res Public Health*. (2020) 17:4095. doi: 10.3390/ijerph17114095
8. Chen Q, Allot A, Lu Z. Keep up with the latest coronavirus research. *Nature*. (2020) 579:193. doi: 10.1038/d41586-020-00694-1
9. Entrez Programming Utilities (E-Utilities). *Encyclopedia of Genetics, Genomics, Proteomics and Informatics*. Dordrecht: Springer (2008). p. 612.
10. Syed S, Spruit M. Full-Text or Abstract? Examining topic coherence scores using latent dirichlet allocation. In: *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. Tokyo (2017).
11. Rehurek R, Sojka P. *Software Framework for Topic Modelling with Large Corpora*. (2010). Available online at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.695.4595 (accessed October 6, 2020)
12. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. In: *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining - WSDM '15*. Shanghai (2015).
13. Huang A. Similarity measures for text document clustering. In *Proceedings of the Sixth New Zealand Computer Science Research Student Conference (NZCSRSC2008)*. Christchurch (2008). pp. 9–56.
14. *LitCovid*. Available online at: https://www.ncbi.nlm.nih.gov/research/coronavirus/ (accessed October 29, 2020).
15. Sievert C, Shirley K. LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*. Baltimore, MD (2014).
16. CDC. *Coronavirus Disease (2019). (COVID-19)*. (2020). Available online at: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/diy-cloth-face-coverings.html (accessed September 27, 2020).
17. Ebadi A, Xi P, Tremblay S, Spencer B, Pall R, Wong A. Understanding the temporal evolution of COVID-19 research through machine learning and natural language processing. *Scientometrics*. (2020) 126:725–39. doi: 10.1007/s11192-020-03744-7
18. Doanvo A, Qian X, Ramjee D, Piontkivska H, Desai A, Majumder M. Machine learning maps research needs in COVID-19 literature. *Patterns (NY)*. (2020) 1:100123. doi: 10.1101/2020.06.11.145425
19. Zhu B, Zheng X, Liu H, Li J, Wang P. Analysis of spatiotemporal characteristics of big data on social media sentiment with COVID-19 epidemic topics. *Chaos Solitons Fractals*. (2020) 140:110123. doi: 10.1016/j.chaos.2020.110123
20. Ma Q-X, Shan H, Zhang H-L, Li G-M, Yang R-M, Chen J-M. Potential utilities of mask-wearing and instant hand hygiene for fighting SARS-CoV-2. *J Med Virol.* (2020) 92:1567–71. doi: 10.1002/jmv.25805
21. Peeples L, Face masks: what the data say. *Nature*. (2020) 586:186–9. doi: 10.1038/d41586-020-02801-8
22. Amano T, González-Varo JP, Sutherland WJ. Languages Are Still a Major Barrier to Global Science. *PLoS Biol.* (2016) 14:e2000933. doi: 10.1371/journal.pbio.2000933
23. Gupta A, Aeron S, Agrawal A, Gupta H. *Trends in COVID-19 Publications: Streamlining Research Using NLP and LDA*. Available online at: https://ssrn.com/abstract=3708327 or http://dx.doi.org/10.2139/ssrn.3708327 (accessed October 15, 2020).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

BRIEF RESEARCH REPORT
published: 08 July 2021
doi: 10.3389/fdgth.2021.660809

Combinatorial Analysis of Phenotypic and Clinical Risk Factors Associated With Hospitalized COVID-19 Patients

Sayoni Das[1†], *Matthew Pearson*[1†], *Krystyna Taylor*[1], *Veronique Bouchet*[1], *Gert Lykke Møller*[1], *Taryn O. Hall*[2], *Mark Strivens*[1], *Kathy T. H. Tzeng*[2] *and Steve Gardner*[1*]

[1] PrecisionLife Ltd., Oxford, United Kingdom, [2] OptumLabs at UnitedHealth Group, Minnetonka, MN, United States

Characterization of the risk factors associated with variability in the clinical outcomes of COVID-19 is important. Our previous study using genomic data identified a potential role of calcium and lipid homeostasis in severe COVID-19. This study aimed to identify similar combinations of features (disease signatures) associated with severe disease in a separate patient population with purely clinical and phenotypic data. The PrecisionLife combinatorial analytics platform was used to analyze features derived from de-identified health records in the UnitedHealth Group COVID-19 Data Suite. The platform identified and analyzed 836 disease signatures in two cohorts associated with an increased risk of COVID-19 hospitalization. Cohort 1 was formed of cases hospitalized with COVID-19 and a set of controls who developed mild symptoms. Cohort 2 included Cohort 1 individuals for whom additional laboratory test data was available. We found several disease signatures where lower levels of lipids were found co-occurring with lower levels of serum calcium and leukocytes. Many of the low lipid signatures were independent of statin use and 50% of cases with hypocalcemia signatures were reported with vitamin D deficiency. These signatures may be attributed to similar mechanisms linking calcium and lipid signaling where changes in cellular lipid levels during inflammation and infection affect calcium signaling in host cells. This study and our previous genomics analysis demonstrate that combinatorial analysis can identify disease signatures associated with the risk of developing severe COVID-19 separately from genomic or clinical data in different populations. Both studies suggest associations between calcium and lipid signaling in severe COVID-19.

Keywords: COVID-19, SARS-CoV-2, severe COVID-19, disease risk, patient stratification, combinatorial analysis, real world data analysis

OPEN ACCESS

Edited by:
Pradeep Nair, Central University of Himachal Pradesh, India

Reviewed by:
Xiangjun Du, Sun Yat-sen University, China
Lin Song, University of California, San Francisco, United States

***Correspondence:**
Steve Gardner
steve@precisionlife.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Health Informatics, a section of the journal Frontiers in Digital Health

Received: *29 January 2021*
Accepted: *11 June 2021*
Published: *08 July 2021*

Citation:
Das S, Pearson M, Taylor K, Bouchet V, Møller GL, Hall TO, Strivens M, Tzeng KTH and Gardner S (2021) Combinatorial Analysis of Phenotypic and Clinical Risk Factors Associated With Hospitalized COVID-19 Patients. Front. Digit. Health 3:660809. doi: 10.3389/fdgth.2021.660809

INTRODUCTION

The Coronavirus disease 2019 (COVID-19) outbreak caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been declared a pandemic that has resulted in significant mortality, major social and economic disruption worldwide (1). The uncertainty surrounding the progression, management, and outcomes of COVID-19 has made it particularly challenging for healthcare systems. Studies have suggested that ∼80% of COVID-19 positive patients present with mild symptoms or are asymptomatic and that around 20% of the patients develop a more severe

response that may lead to hospitalization and, in some cases (2.3%), death (2–5).

The risk of developing severe COVID-19 is known to be higher in people who are older, male and have underlying health conditions such as hypertension, cardiovascular disease, diabetes, obesity, chronic respiratory diseases, and cancer (4, 5). Approximately 22% of the global population have at least one co-morbidity that puts them at increased risk of severe COVID-19 if exposed to the virus (6). Ethnicity and socio-economic deprivation have also been associated with severe illness (7).

SARS-CoV-2 binds to the host cell receptor through angiotensin-converting enzyme-2 (ACE2) (8) and starts replicating rapidly inside the host cells, which can trigger a hyperimmune response in some patients (9). This may be due to the generation of pro-inflammatory cytokines and chemokines called a cytokine storm that can cause acute respiratory distress syndrome (ARDS) in the lung and multi-organ failure (10, 11). Other studies have suggested that binding of SARS-CoV-2 increases the levels of ACE2 in lung cells that results in elevated levels of bradykinin (12) (bradykinin storm) leading to vascular leakage, hypotension, and pulmonary edema (13). These are manifested in COVID-19 patients with pneumonia and respiratory failure. Bradykinin's role in the regulation of clotting may be one mechanism for the extra-pulmonary manifestations such as thromboembolic complications, cardiac events, acute renal and hepatic injury (14, 15). Other symptoms such as neurological complications and gastrointestinal and endocrine symptoms have also been reported (14, 16). Recent evidence suggests that some patients with COVID-19 can also develop long-term complications or experience prolonged symptoms (17, 18).

Early identification and characterization of the risk factors associated with varying clinical outcomes of severely ill COVID-19 patients are crucial for accurate clinical stratification and the development of effective management and targeted therapeutic strategies. A previous case-control study using genomic data (19) identified 68 severe COVID-19 risk-associated genes in a population of hospitalized COVID-19 patients in the UK Biobank (20, 21). Nine of these were previously linked to differential response to SARS-CoV-2 infection. Several of these genes are related to key biological pathways associated with the development of severe COVID-19 and associated symptoms, including cytokine production cascades, endothelial cell dysfunction, lipid droplets, calcium signaling, and viral susceptibility factors (19).

In this study, we identified and assessed the phenotypic and clinical risk factors associated with hospitalized COVID-19 patients in the UnitedHealth Group (UHG) COVID-19 Data Suite using a similar combinatorial analysis approach. Using laboratory test data available for the UHG cohort, we investigated potential correlations with the genomic analysis findings and hypotheses from our previous UK Biobank COVID-19 study (19), including the potential association of calcium signaling and lipid dysregulation with severe clinical outcomes in COVID-19 patients.

METHOD

Cohort Generation

We used de-identified records of Medicare Advantage and commercially insured members with COVID-19 test results in the UHG COVID-19 Data Suite accessed through the UHG Clinical Discovery Portal for this study. The UHG COVID-19 Data Suite contains longitudinal health information on individuals representing diverse ethnicities, age groups, and geographical regions across the United States. The information includes data on COVID-19 test results, in-patient admission data for hospitalized individuals, medical and pharmacy claims, general diagnostic information, demographic data, and information on healthcare insurance plans.

We performed case-control studies on two cohorts to identify combinatorial disease signatures associated with the risk of hospitalization for COVID-19 positive patients. Cohort 1, consisting of 9,493 individuals (3,183 cases, 6,310 controls), was generated from the UHG COVID-19 Data Suite (dated August 2020). This contained 3,183 cases who had been hospitalized as a result of developing severe COVID-19 (based on primary diagnosis records) and 6,310 mild controls who had tested positive for COVID-19 but not been hospitalized (**Supplementary Table 1**). Patients who were enrolled in the Medicare Special Needs Plan were excluded to reduce any confounding factors associated with these patients, who are often above 65 years old and diagnosed with severe/disabling chronic conditions that increase their risk of hospitalization. Patients without linked clinical data since 2019 were also excluded.

To investigate the potential role of calcium and lipid homeostasis in COVID-19 patients with severe clinical outcomes, we selected five laboratory analytes that were relevant for this hypothesis and had good coverage in Cohort 1. These included serum calcium, low-density cholesterol (LDL), high-density cholesterol (HDL), triglycerides, and leukocyte count. A subcohort, Cohort 2, consisting of 1,581 patients (581 cases and 1,000 controls) was generated for the individuals with laboratory test results for these five analytes.

Feature Generation

The clinical, claims and pharmacy data were converted to categorical features for the study (**Supplementary Section** Feature Generation). The clinical and phenotypic data available for all individuals in Cohort 1 generated 1,339 binary features per patient (**Supplementary Table 2**). An additional, five laboratory analyte features were added for Cohort 2.

Combinatorial Analysis

The PrecisionLife platform uses a proprietary data analytics framework that enables efficient combinatorial analysis of large, n-dimensional, multi-modal patient datasets. Navigating this data space allows for the identification of combinations of features that are significantly associated with groups of cases in a case-control dataset.

Traditional analysis methods typically identify single features in a dataset that are important for a relatively large number of cases associated with a disease diagnosis. They may seek to

combine these single feature effects using a variety of methods. However, most large disease populations are heterogenous with multiple features coming together to exert non-linear influences on disease biology that lead to patient sub-populations having different symptoms, progression, and/or outcomes. These non-linear effects can only be observed in combination, i.e., they are a product of the interaction and so have to be observed and modeled at that level. The combinatorial approach used in this analysis enables us to capture the non-linear effects of these interactions on a disease (e.g., the effects of feedback loops in metabolic or genetic networks), which can only be seen in combinations found to be significant in such patient subgroups. This approach has been validated in multiple disease populations (19, 22, 23).

PrecisionLife's combinatorial analysis algorithm comprises two main phases: mining and processing (**Supplementary Figure 1**). In the mining phase, the algorithm identifies and validates combinations of feature states (for example SNP and associated genotype state) that are over-represented in cases. Multiple feature states are combined iteratively (using a Z-score statistic) until no additional single feature state is added. Combinations of feature states that have high odds ratios and high penetrance are prioritized. The mining process is repeated for 2,500 cycles of fully randomized permutation of all individuals in the dataset, keeping the same parameters and case:control ratio.

All combinations associated with each feature state are identified to form 'simple networks' for the original dataset and for each iteration of random permutation of the dataset. The simple networks are then validated using network properties such as minimum penetrance (number of cases in the simple network) as the null hypothesis when compared with the networks of the random permutations. Simple networks that appear in the random permutations above a preset FDR threshold are considered to be random and eliminated. All disease signatures from the validated simple networks are reported as validated disease signatures.

In the last phase, the validated disease signatures are processed. The features that connect all disease signatures in validated simple networks (known as critical features) are identified. These critical features are scored using a Random Forest (RF) algorithm based inside a 5-fold cross-validation framework to evaluate the accuracy with which the feature predicts the observed case:control split (minimizing Gini impurity) in a dataset. We use the resulting score to rank disease signatures.

Finally, a merged network architecture is generated by clustering all validated disease signatures based on their co-occurrence in patients in the dataset.

The PrecisionLife platform generated statistically significant disease signatures containing up to five features for each cohort. Each analysis took less than an hour to complete, running on a 32 CPU, 4 GPU cloud compute server. These were mapped to the cases in which they were found, and in-patient clinical data were used to generate a patient profile for each combinatorial disease signature.

RESULTS

Cohort Characteristics

Cohort 1 patients (3,183 cases) had a 19.1% (607 cases) mortality rate, while 51.3% (1,548 cases) were released from care and 29.6% (915 cases) were transferred to other healthcare facilities. Within Cohort 1, 51.3% were female, and 66.7% were Caucasian with a median age of 75 (**Table 1**, **Supplementary Figure 2**).

Around 54% of the hospitalized patients had at least one of the comorbidities previously linked with higher risk for COVID-19 severe response. Hypertension (52.1%) was the most common co-morbidity, followed by cardiovascular disease (38%), diabetes (31.5%), chronic lung disease (25.9%) and dementia (13.9%) (**Table 1**). The most common COVID-19 related diagnoses reported in hospital admissions data for cases were pneumonia (43%), followed by respiratory failure (18.3%) and septicemia (7.3%) (**Supplementary Figure 3**).

Combinatorial Disease Signatures Capture Phenotypic and Clinical Risk Factors for Severe COVID-19

The combinatorial analysis identified 1,147 combinations of clinical and phenotypic features (disease signatures) that were highly associated with hospitalized patients in Cohort 1 and 32,242 combinations in Cohort 2 (**Supplementary Table 3**, **Supplementary Figure 4**). A higher number of disease signatures was reported for Cohort 2. This is likely due to the relatively higher prevalence of the same clinical features among Cohort 2 individuals as compared to Cohort 1.

The disease signatures were filtered to exclude those that had any features indicating an absence of a disease diagnosis, symptom, or medication use, as these are likely to be generated as a result of incompleteness of the claims and pharmacy data rather than as a true disease association. Additionally, disease signatures that were found in fewer than 20 cases were also excluded. After filtering, 255 disease signatures in Cohort 1 and 531 disease signatures in Cohort 2 were used for further analysis.

All features in the disease signatures identified for each study were scored using a Random Forest (RF) algorithm based inside a 5-fold cross-validation framework to evaluate the accuracy with which a feature (e.g., a laboratory analyte value) predicts the observed case:control split (minimizing Gini impurity). One hundred sixty-six features in Cohort 1 and 41 features in Cohort 2 were identified as critical features as shown in **Supplementary Figure 5**. Many of these included diagnoses and symptoms associated with severe COVID-19 such as respiratory failure, pneumonia, acute renal failure, and septicemia because of their low incidence in controls.

We found that the combinatorial disease signatures capture clinical features associated with response to severe COVID-19 illness (**Figures 1**, **2**) These features include pneumonia and respiratory failure, which are frequently reported among hospitalized patients, and risk factors that increase the probability of developing severe response such as diabetes, hypertension and cardiovascular disease. Phenotypes related to the risk-associated comorbidities such as elevated glucose levels or blood pressure and common medications prescribed for them (e.g.,

TABLE 1 | Cohort characteristics for the hospitalization risk studies.

	Cohort 1 (n = 9,493)			Cohort 2 (n = 1,581)		
	Hospitalized patients	**Non-hospitalized patients**	**Two-sided *p-value***	**Hospitalized patients**	**Non-hospitalized patients**	**Two-sided *p*-value**
COVID-19 positive	3,183	6,310	N/A	581	1,000	N/A
Males (%)	1,549 (48.7%)	2,758 (43.7%)	5.0e-06	295 (50.1%)	348 (34.8%)	6.5e-10
Median Age (Range)	75 (29–89)	72 (15-89)	<2.2e-16	74 (31–89)	71.5 (33–89)	2.2e-14
Caucasian* (%)	1,632 (66.7%)	3,693 (66.8%)	0.938	298 (57.5%)	528 (59.5%)	0.465
Mortality (%)	607 (19.10%)	N/A	N/A	100 (17.2%)	N/A	N/A
Hypertension (%)	1,657 (52.1%)	3,864 (61.2%)	<2.2e-16	431 (58.7%)	672 (67.2%)	3.7e-03
Diabetes (%)	1,109 (34.8%)	2,114 (33.5%)	0.198	277 (47.7%)	427 (42.7%)	5.8e-02
Cardiovascular (%)	1,210 (38.0%)	1,468 (23.3%)	<2.2e-16	204 (35.1%)	267 (26.7%)	2.2e-03
Dementia (%)	443 (13.9%)	236 (3.7%)	<2.2e-16	27 (4.6%)	10 (1.0%)	7.2e-06
Chronic lung disease (%)	832 (25.9%)	1,198 (20.0%)	2.4e-15	135 (23.2%)	188 (18.9%)	3.8e-02

**The percentage of Caucasians was calculated as a fraction of those individuals for whom ethnicity information was available (∼85% of all records).*
p-values were calculated to assess the association of each feature in the two COVID-19 hospitalized risk cohorts using two-sided Fisher's exact tests for categorical data and Mann-Whitney U tests for continuous data.

insulin, statins, and dihydropyridines) were also commonly found. Many low-frequency features (<10% among hospitalized patients) such as ARDS (10), pneumothorax (24), hematuria (25), encephalopathy (16), pericarditis (26), and thrombosis (14) were frequently found in disease signatures in combination with other features. Some disease signatures also captured clinical features related to increased frailty such as senility or high risk of hospital readmission, whilst other features reflect conditions that are associated with prolonged hospital stay, such as pressure ulcers and secondary bacterial infections.

Networks generated by clustering disease signatures in the two cohorts highlighted the heterogeneity of clinical features observed in severe COVID-19. Such clustering enables the identification of disease signatures that co-occur in patient subgroups who are likely to have similar symptoms, underlying conditions, or clinical outcomes. For example, hospitalized patients who developed ARDS were likely to be influenced by the features nearest to ARDS in the network such as older age, development of pneumonia, pulmonary hemorrhage, sepsis, and high mortality (**Figure 3**, **Supplementary Figure 6**).

Disease Signatures Associated With Lower Levels of Serum Calcium and Lipids

In Cohort 2, features from five blood analytes (calcium, LDL, HDL, triglycerides, and leukocyte count) were available for patients. Hospitalized patients with severe COVID-19 were observed to be more likely to have lower serum calcium levels (<9.26 mg/dl), lower LDL levels (<78.23 mg/dl), lower HDL levels (<44.35 mg/dl), and higher levels of triglycerides (>206.20 mg/dl) when compared against the patients with mild disease (**Supplementary Table 4**). Both low and high levels of blood leukocyte count were observed in patients with severe COVID-19.

In Cohort 2 the PrecisionLife platform identified 18 disease signatures in 80 hospitalized patients with serum calcium values lower than 9.26 mg/dl (**Supplementary Figure 7**). Out of these, only four signatures were co-associated with the use of the dihydropyridines (calcium channel blockers) and proton-pump inhibitors which may have an effect on calcium homeostasis (27, 28). The hypocalcemia disease signatures were associated with COVID-19 symptoms such as pneumonia and respiratory failure, and comorbidities including diabetes, hypertension, and anemia. Two calcium disease signatures were found in 34 patients (42.5%), co-occurring with high mortality and hospital re-admission risk scores, which suggests that these patients had multiple underlying conditions. Another calcium disease signature in 33 (41.3%) patients was associated with low serum levels of HDL and pneumonia.

We also identified 45 disease signatures in 188 (32.4%) severe COVID-19 patients that were associated with comparatively low serum lipid (LDL, HDL, or triglyceride) levels (**Supplementary Figures 8–10**). Comorbidities such as hypertension, obesity, and cerebrovascular disease were found in these hypolipidemia signatures, which are not commonly co-associated in patients. We investigated whether the reduced lipid levels observed in these patients were caused by the use of statins. None of the disease signatures were associated with the feature indicating statin use by all associated cases. We found 12 hypolipidemia signatures where <10% of the patients were associated with any prescription records for statins within 90 days of the laboratory test result date, suggesting that these signatures were independent of statin use. Thus, dyslipidemia observed in many severe COVID-19 patients in Cohort 2 is not likely to represent an artifact of other comorbidities or medication use, but a consequential host response to SARS-CoV-2 infection which has been reported in many recent studies (29–31).

Mortality in the patients with either calcium or lipid disease signatures was not found to be significantly different. We were able to identify 15 disease signatures with lower levels of calcium and one signature with lower levels of cholesterol in this subcohort that were associated with at least

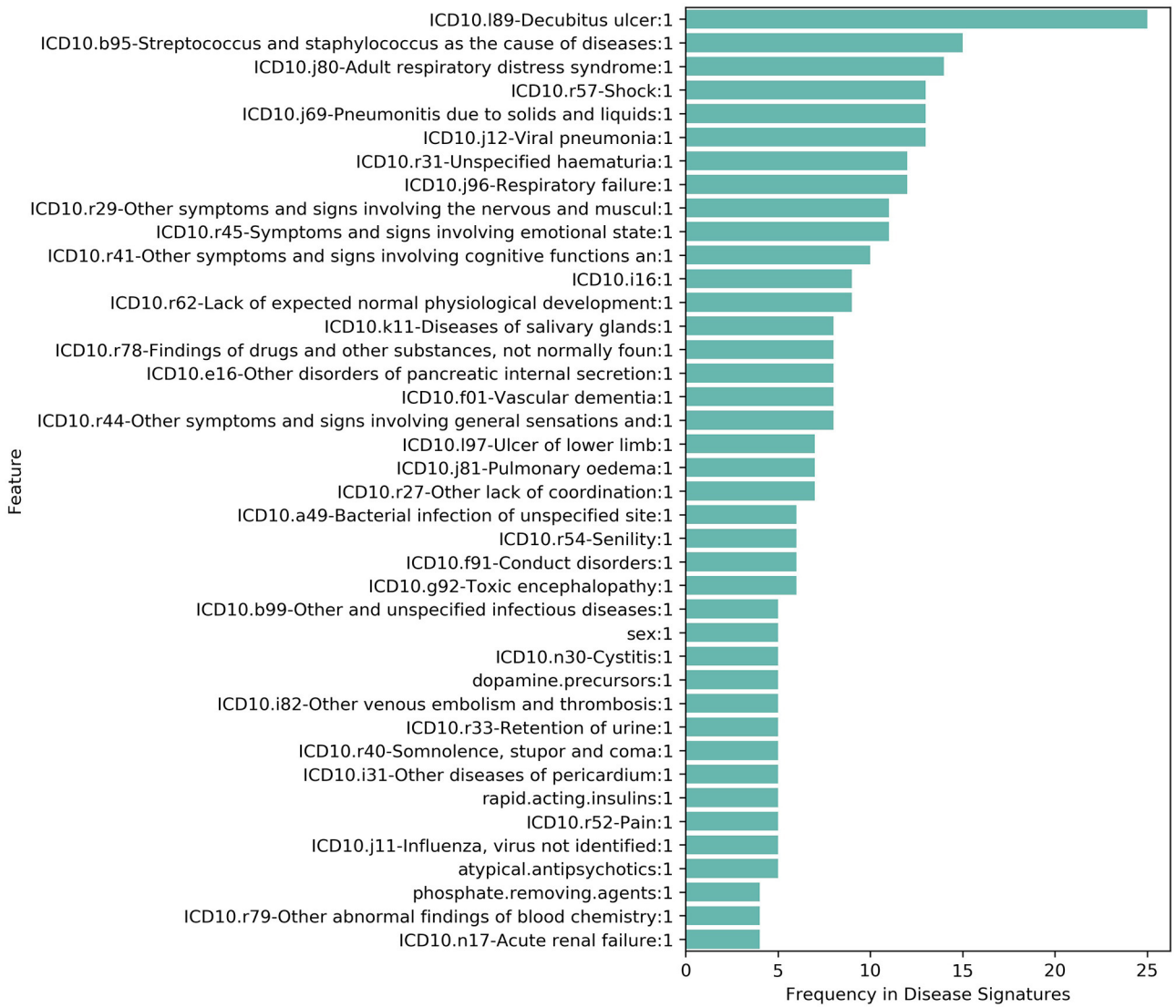


FIGURE 1 | Phenotypic and clinical features that were most frequently reported (top 40) in 255 filtered disease signatures in Cohort 1 were associated with an increased risk of hospitalization with severe COVID-19.

10 patients. The identification of calcium and lipid disease signatures in this subcohort strongly suggests that they reflect biochemical characteristics of patients with severe host response to COVID-19.

DISCUSSION

Pulmonary manifestations of COVID-19 such as respiratory failure and pneumonia were the most common symptoms in the two cohorts that were also prevalent in the combinatorial disease signatures identified by the PrecisionLife platform (**Supplementary Figures 3, 5**). Comorbidities such as hypertension, cardiovascular disease, chronic respiratory disease, and diabetes are known to be associated with COVID-19 risk from other studies (2–4), including our previous genetic study (18) in UK Biobank, were observed in hospitalized patients. These comorbidities co-occur with different COVID-19 symptoms, complications, medication use, and laboratory analyte values. This analysis enables us to gain useful insights into the likely associations between these clinical and phenotypic features that can improve the clinical management of patients.

A wide variety of severe COVID-19 manifestations, such as ARDS, sepsis, pericarditis, and thrombosis, were observed in the disease signatures representing patient sub-groups (2, 14, 24–26). This correlates with our previous genomic analysis on the UK Biobank COVID-19 cohort, which identified genes associated with some of these complications, including host pathogenic responses, inflammatory cytokine production, modulation of cardiac function, and endothelial cell function (19).

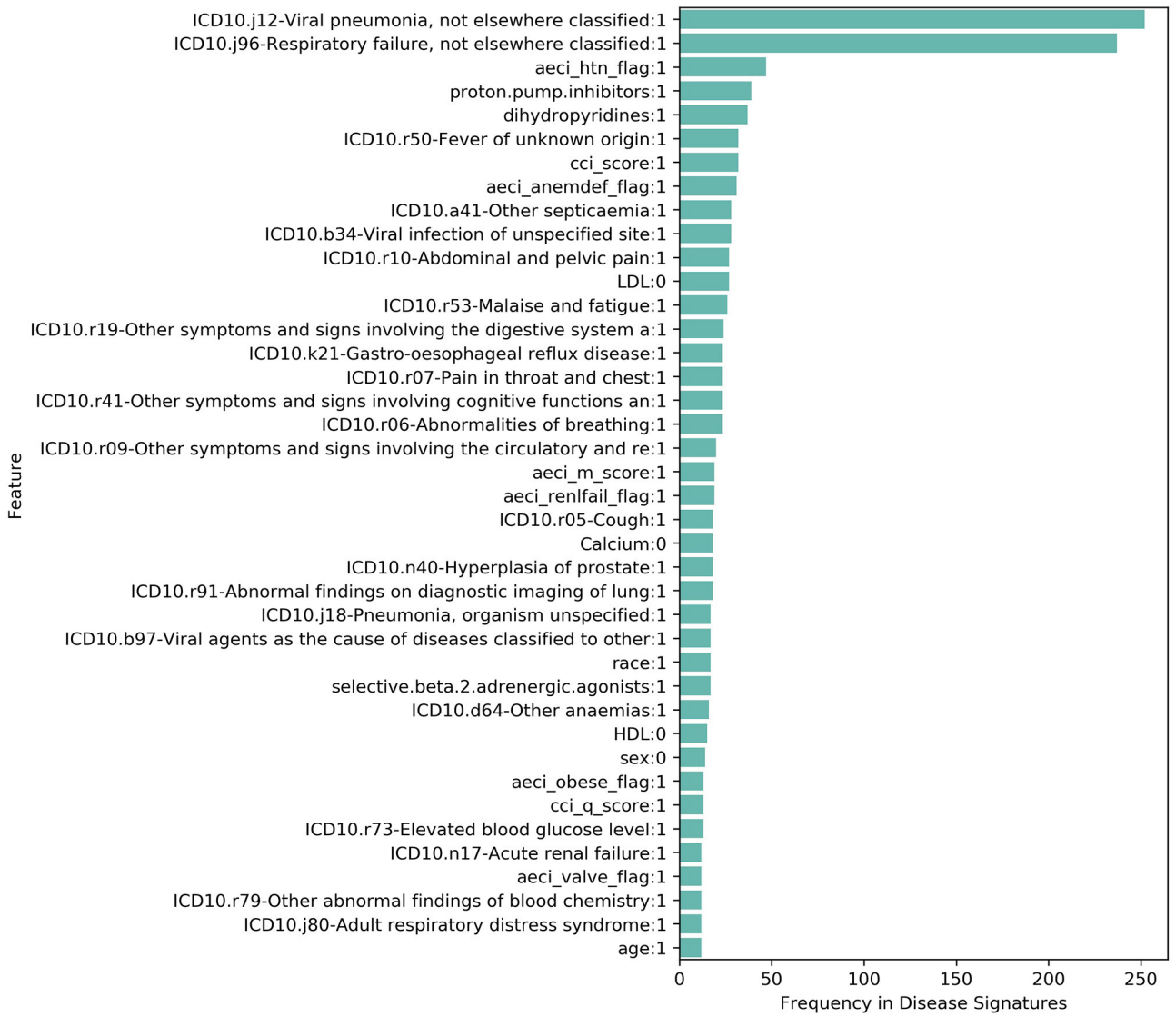


FIGURE 2 | Phenotypic and clinical features that were most frequently reported (top 40) in 581 filtered disease signatures in Cohort 2 (subcohort of Cohort 1 with additional laboratory test results) associated with increased risk of hospitalization with severe COVID-19. Features associated with hypocalcemia (Calcium:0) and hypolipidemia (LDL:0, HDL:0) were reported in multiple disease signatures.

The use of medications such as proton pump inhibitors, dihydropyridines, and beta-adrenergic blockers was observed in seven disease signatures in Cohort 1 and 80 signatures in Cohort 2. Dihydropyridines (32, 33)and beta-adrenergic blockers (34, 35) have been associated with improved outcomes for COVID-19 patients and suggested as potential treatments, while proton pump inhibitors have been associated with adverse outcomes in several studies (36, 37). The incidence of the medications in the disease signatures could be either due to adverse effects caused by the medication resulting in a more severe COVID-19 response or it could reflect the comorbidities in patients for which they are generally prescribed. Using the available data, it was not possible for us to ascertain the specific association of these medications in our study with certainty.

In Cohort 2, all hypocalcemia ($n = 18$) disease signatures and hypolipidemia ($n = 45$) signatures were found to be associated with severe pulmonary manifestations of COVID-19 (**Supplementary Figures 7–10**). There is increasing evidence that calcium and lipid homeostasis plays an important role in the viral replication cycle and they have been suggested as biomarkers for increased COVID-19 severity (29–31, 38). It has been demonstrated that the calcium signaling pathway or calcium-dependent processes in host cells are often perturbed by viral proteins that can bind calcium and/or calcium-binding protein domains, allowing them to modulate the host cellular machinery for viral replication, assembly, and release (39, 40). The mechanism of calcium regulation is not fully understood, as some viruses are known to increase

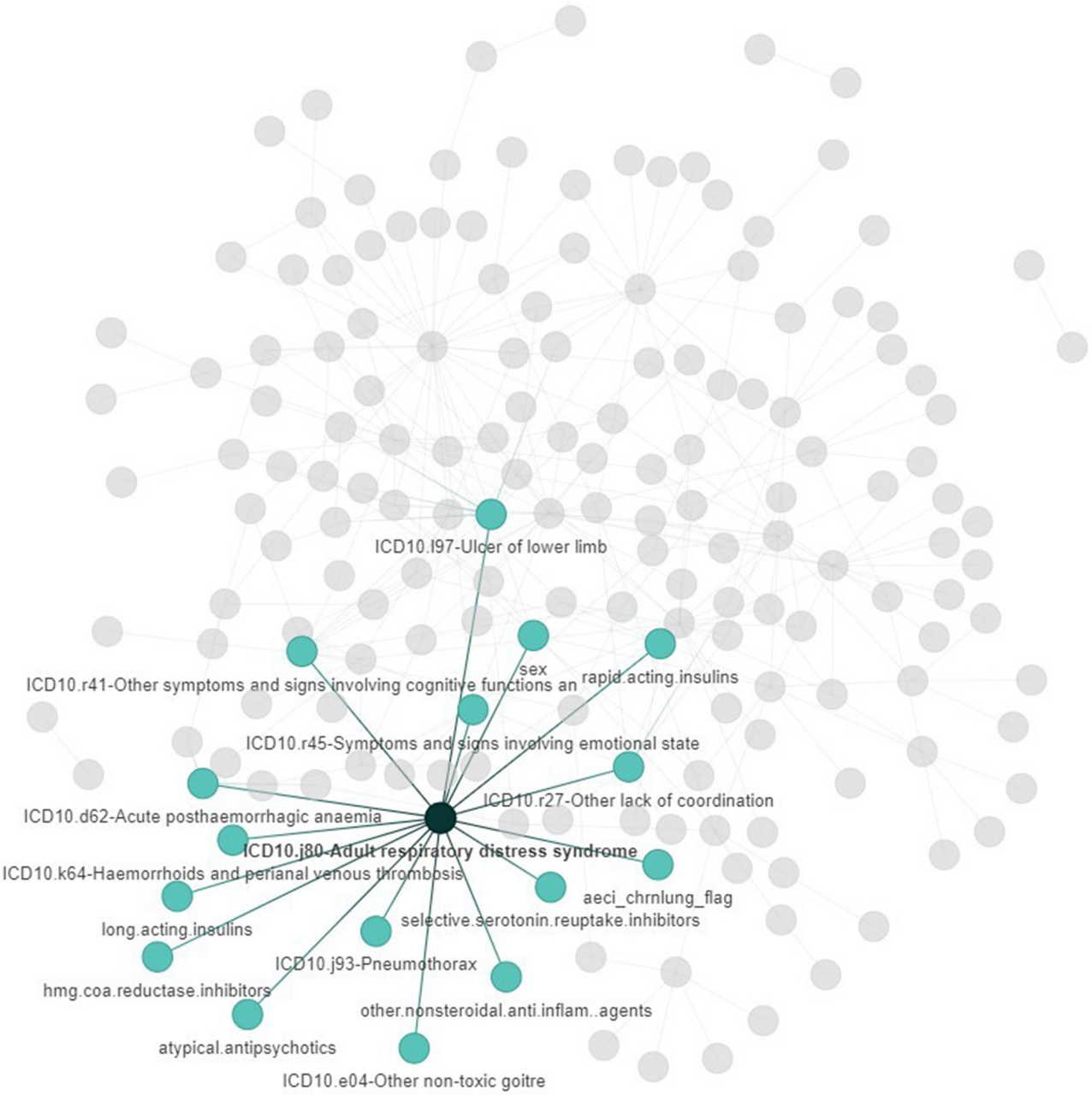


FIGURE 3 | The network architecture of filtered ($n = 255$) disease signatures associated with hospitalized COVID-19 patients in Cohort 1 generated by the PrecisionLife platform where each circle represents a feature and edges represent co-association in patients. The colored nodes and edges represent the disease signatures of patients who developed ARDS (shown in a darker shade) in Cohort 1. The co-associated features are shown in a lighter shade.

intracellular calcium levels while others are known to have a dynamic control based on the phase of infection (41). However, the SARS-CoV E protein has been shown to form protein-lipid channels that transport calcium ions, activating the NLRP3 inflammasome and increasing systemic inflammation via IL-1β (42).

Lower lipid levels have been reported in severe COVID-19 patients in many studies with a correlation observed between reduced lipid levels and disease severity (43–45). Many viruses, including SARS-CoV and MERS-CoV, can modulate lipid synthesis and signaling in host cells to divert cellular lipids to viral replication and exocytosis, facilitating the invasion of other host cells (46, 47). It has been suggested that the decrease in cellular cholesterol levels following SARS-CoV-2 infection leads to disruption of the signaling hub for inflammation and cholesterol metabolism, resulting in the dysregulation of cholesterol biosynthesis, inflammatory cytokine release, and vascular homeostasis (48, 49).

Regulation of cholesterol biosynthesis has been shown to be associated with six genes identified by a genome-scale CRISPR knockout screen that reduced SARS-Cov-2 infection in human alveolar basal epithelial carcinoma cells (50). The study also demonstrated that the use of dihydropyridines results in increased resistance to SARS-Cov-2 infection (50). Another study hypothesized that elevated unsaturated fatty acids in SARS-CoV-2 infected host cells bind calcium, resulting in hypocalcemia and triggering the production of pro-inflammatory mediators and cytokine storm induction (51, 52).

We found seven disease signatures in this study where lower levels of LDL were found co-occurring with lower levels of serum calcium, leukocyte count, or HDL. These signatures may be attributed to similar mechanisms linking calcium and lipid signaling where changes in cellular lipid levels during inflammation and infection (53) affects calcium signaling in host cells (54–56).

Retrospective analysis of the clinical histories of the hospitalized patients with lower calcium and lipid signatures was performed to identify whether the laboratory analyte values may be affected by other medical conditions. We found that 50% of cases represented by disease signatures featuring lower levels of calcium were reported to have vitamin D deficiency which is important for calcium homeostasis in both physiological and disease states (57). More than 25% of people above the age of 65 were vitamin D deficient. This suggests that the changes in calcium levels in some patients may be linked to vitamin D deficiency in severe COVID-19 (57, 58), which has also been associated with severe illness and which was found in eight disease signatures in Cohort 2. A recent study reported that lower serum calcium levels have been found to be associated with COVID-19 patients with pneumonia independent of vitamin D deficiency (59). This finding is consistent with our findings that a sub-group (50%) of patients with low serum calcium values were not reported with vitamin D deficiency. It is likely that in these patients, the changes in lipid levels following COVID-19 infection (53) affects the serum calcium levels (54–56), similar to the patients who had disease signatures that were combinations of lower serum calcium, leukocyte, and HDL levels.

Our previous analysis on the UK Biobank COVID-19 cohort (19) identified 16 calcium-binding/signaling genes and six genes relating to lipid droplet biology and correlated with serum lipid levels and coronary artery disease. In conjunction with the findings of this study, this adds further support to the role of calcium and lipid signaling in relation to viral pathogenesis and severe host response to COVID-19. To fully understand the role of calcium and lipid homeostasis in COVID-19, analysis of patient datasets that combine genetic, clinical, and hospital laboratory test data will be necessary.

Limitations of the Study

This study was limited by the completeness of data for features relevant to analyzing differential host response to COVID-19. Information on the onset of disease or symptoms, the clinical phase of disease, viral load, oxygen saturation, breathing rate, body mass index, and physiological measurements or biomarker levels during hospitalization was not consistently available. We used hospitalization status associated with a primary diagnosis of COVID-19 as a surrogate for severe COVID-19 patients. Mortality and diagnoses linked to clinical progression of COVID-19 were used to estimate the relative severity of disease among hospitalized patients.

The comorbidities, diagnoses, medications, and laboratory test results were derived from medical claims, pharmacy claims, and in-patient admission records. Since claims data are generated for reimbursement and administrative purposes rather than scientific research, the records may be missing information and there is potential for variability in their collection. Data sparsity of the available patient records was reflected in the low penetrance of many disease signatures. As more patient data becomes available, the disease signatures will become more predictive, enabling higher resolution patient stratification.

CONCLUSION

The PrecisionLife platform identified and analyzed 836 combinatorial disease signatures in two COVID-19 cohorts (Cohort 1 = 255, Cohort 2 = 531) associated with increased risk of hospitalization from COVID-19. These disease signatures were found to capture different symptomatic presentations of COVID-19, complications arising from the clinical progression of the disease, and underlying disease conditions that could be either associated with severe host response to COVID-19 or were indicative of conditions associated with older age or frailty.

In Cohort 2, we found 45 disease signatures that were associated with lower levels of serum calcium, LDL, HDL, and triglycerides in 188 (32.35%) hospitalized patients. This suggests that lower levels of calcium and cholesterol are biochemical characteristics associated with severe COVID-19 patients, which may also add further support to the role of calcium signaling and lipid dysregulation in SARS-CoV-2 pathogenesis.

These findings are consistent with the insights generated by multiple studies in different COVID-19 patient populations. This also validates our findings from our previous genomics study (19) on severe COVID-19 patients in UK Biobank (20) where we identified 16 risk-associated genes that had calcium-binding domains or were involved in calcium signaling and six genes linked to lipid droplet biology associated with serum lipid levels.

This study along with our previous genomic study (19) demonstrates that a combinatorial analysis approach is able to identify related groups of clinical and phenotypic features from both genomic and phenotypic data that are associated with the risk of developing severe forms of COVID-19. This enables us to gain unique insights into the non-linear combinatorial feature associations to a clinical phenotype in patient sub-groups, that is not detected by standard data analysis approaches. With the availability of more data, the combinatorial output of the analytical platform would be greatly enhanced and the insights derived from them would allow for the identification of targeted approaches to patient care.

This analysis also validates the association of calcium and lipid homeostasis with severe COVID-19 reported by our previous study, using real-world data in an independent cohort. We will extend these analyses in future to larger patient datasets that have both genetic and phenotypic data to fully ascertain the differences between mild and severe host responses to COVID-19 and the mechanism of calcium and lipid signaling in SARS-Cov-2 pathogenesis.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: the data analyzed in this study was obtained from UnitedHealth Group Clinical Discovery Portal. The data are proprietary and are not available for public use but, under certain conditions, may be made available to editors and their approved auditors under a data-use agreement to confirm the findings of the current study. Requests to access these datasets should be directed to Scott Schneweis, schneweis@uhg.com.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

SG conceived and supervised the study. MP and SD performed the studies and analyzed the data. SD wrote the manuscript. KT contributed to the study design, analysis of disease signatures, and manuscript. VB and MS contributed to the study design and manuscript. GM developed the core technology in PrecisionLife's platform. TH and KTHT contributed to the study design and coordinated access to the COVID-19 Data Suite through the UHG Clinical Discovery Portal. All authors contributed to the study and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We would like to acknowledge the UnitedHealth Group for providing us access to the COVID-19 Data Suite through the UHG Clinical Discovery Portal and the patients who provided their data. Special thanks to Megan Jarvis, Kae Tanudtanud, Yinglong Guo, Elena Fultz, Aditya Yellepeddi, and Teodi Enrik Racho from the UnitedHealth Group and the rest of the PrecisionLife team for their technical assistance and helpful discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.660809/full#supplementary-material

REFERENCES

1. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. *Acta Bio Med.* (2020) 91:157. doi: 10.23750/abm.v91i1.9397
2. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis.* (2020) 20:669–77. doi: 10.1016/S1473-3099(20)30243-7
3. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA.* (2020) 323:1239–42. doi: 10.1001/jama.2020.2648
4. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3
5. Hu B, Guo H, Zhou P, Shi ZL. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol.* (2020) 19:141–54. doi: 10.1038/s41579-020-00459-7
6. Clark A, Jit M, Warren-Gash C, Guthrie B, Wang HH, Mercer SW, et al. Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *Lancet Global Health.* (2020) 8:e1003–17. doi: 10.1016/S2214-109X(20)30264-3
7. Niedzwiedz CL, O'Donnell CA, Jani BD, Demou E, Ho FK, Celis-Morales C, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. *BMC Med.* (2020) 18:1–4. doi: 10.1186/s12916-020-01640-8
8. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell.* (2020) 181:271–80.e8. doi: 10.1016/j.cell.2020.02.052
9. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* (2020) 579:270–3. doi: 10.1038/s41586-020-2951-z
10. Ragab D, Salah Eldin H, Taeimah M, Khattab R, Salem R. The COVID-19 cytokine storm; what we know so far. *Front Immunol.* (2020) 11:1446. doi: 10.3389/fimmu.2020.01446
11. Coperchini F, Chiovato L, Croce L, Magri F, Rotondi M. The cytokine storm in COVID-19: an overview of the involvement of the chemokine/chemokine-receptor system. *Cytokine Growth Factor Rev.* (2020) 53:25–32. doi: 10.1016/j.cytogfr.2020.05.003
12. Garvin MR, Alvarez C, Miller JI, Prates ET, Walker AM, Amos BK, et al. A mechanistic model and therapeutic interventions for COVID-19 involving a RAS-mediated bradykinin storm. *Elife.* (2020) 9:e59177. doi: 10.7554/eLife.59177.sa2
13. Zwaveling S, van Wijk RG, Karim F. Pulmonary edema in COVID-19: explained by bradykinin? *J Allergy Clin Immunol.* (2020) 146:1454–5. doi: 10.1016/j.jaci.2020.08.038
14. Gupta A, Madhavan MV, Sehgal K, Nair N, Mahajan S, Sehrawat TS, et al. Extrapulmonary manifestations of COVID-19. *Nat Med.* (2020) 26:1017–32. doi: 10.1038/s41591-020-0968-3
15. Gavriatopoulou M, Korompoki E, Fotiou D, Ntanasis-Stathopoulos I, Psaltopoulou T, Kastritis E, et al. Organ-specific manifestations of COVID-19 infection. *Clin Exp Med.* (2020) 20:493–506. doi: 10.1007/s10238-020-00648-x
16. Garg RK, Paliwal VK, Gupta A. Encephalopathy in patients with COVID-19: a review. *J Med Virol.* (2020) 93:206–22. doi: 10.1002/jmv.26207
17. Iacobucci G. Long covid: damage to multiple organs presents in young, low risk patients. *BMJ.* (2020) 371:m4470. doi: 10.1136/bmj.m4470
18. Dennis A, Wamil M, Kapur S, Alberts J, Badley A, Decker GA, et al. Multi-organ impairment in low-risk individuals with long COVID. *medrxiv.* (2020). doi: 10.1101/2020.10.14.20212555

19. Taylor K, Das S, Pearson M, Kozubek J, Pawlowski M, Jensen CE, et al. Analysis of genetic host response risk factors in severe COVID-19 patients. *medRxiv*. (2020). doi: 10.1101/2020.06.17.20134015
20. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* (2018) 562:203–9. doi: 10.1038/s41586-018-0579-z
21. Armstrong J, Rudkin JK, Allen N, Crook DW, Wilson DJ, Wyllie DH, et al. Dynamic linkage of covid-19 test results between public health england's second generation surveillance system and uk biobank. *Microb Genom.* (2020) 6:mgen000397. doi: 10.1099/mgen.0.000397
22. Koefoed P, Andreassen OA, Bennike B, Dam H, Djurovic S, Hansen T, et al. Combinations of SNPs related to signal transduction in bipolar disorder. *PLoS ONE.* (2011) 6:e23812. doi: 10.1371/journal.pone.0023812
23. Taylor K, Das S, Pearson M, Kozubek J, Strivens M, Gardner S. Systematic drug repurposing to enable precision medicine: a case study in breast cancer. *Digit Med.* (2019) 5:180. doi: 10.4103/digm.digm_28_19
24. Zantah M, Castillo ED, Townsend R, Dikengil F, Criner GJ. Pneumothorax in COVID-19 disease-incidence and clinical characteristics. *Respir Res.* (2020) 21:1–9. doi: 10.1186/s12931-020-01504-y
25. Liu X, Zhang R, He G. Hematological findings in coronavirus disease 2019: indications of progression of disease. *Ann Hematol.* (2020) 99:1421–8. doi: 10.1007/s00277-020-04103-5
26. Tung-Chen Y. Acute pericarditis due to COVID-19 infection: an underdiagnosed disease? *Med Clin.* (2020) 155:44. doi: 10.1016/j.medcli.2020.04.007
27. Price D, Radke J, Albertson T. Hypocalcaemia after an occult calcium channel blocker overdose: a case report and literature review. *Basic Clin Pharmacol Toxicol.* (2014) 114:217–21. doi: 10.1111/bcpt.12121
28. Sivakumar J. Proton pump inhibitor-induced hypomagnesaemia and hypocalcaemia: case review. *Int J Physiol Pathophysiol Pharmacol.* (2016) 8:169.
29. Sun JK, Zhang WH, Zou L, Liu Y, Li JJ, Kan XH, et al. Serum calcium as a biomarker of clinical severity and prognosis in patients with coronavirus disease 2019. *Aging.* (2020) 12:11287. doi: 10.18632/aging.103526
30. di Filippo L, Formenti AM, Doga M, Frara S, Rovere-Querini P, Bosi E, et al. Hypocalcemia is a distinctive biochemical feature of hospitalized COVID-19 patients. *Endocrine.* (2020) 71:9–13. doi: 10.1007/s12020-020-02541-9
31. Yang C, Ma X, Wu J, Han J, Zheng Z, Duan H, et al. Low serum calcium and phosphorus and their clinical performance in detecting COVID-19 patients. *J Med Virol.* (2020) 93:1639–51. doi: 10.1002/jmv.26515
32. Solaimanzadeh I. Nifedipine and amlodipine are associated with improved mortality and decreased risk for intubation and mechanical ventilation in elderly patients hospitalized for COVID-19. *Cureus.* (2020) 12:e8069. doi: 10.7759/cureus.8069
33. Zhang L, Sun Y, Zeng HL, Peng Y, Jiang X, Shang WJ, et al. Calcium channel blocker amlodipine besylate is associated with reduced case fatality rate of COVID-19 patients with hypertension. *medRxiv.* (2020). doi: 10.1101/2020.04.08.20047134
34. Vasanthakumar N. Beta-adrenergic blockers as a potential treatment for COVID-19 patients. *BioEssays.* (2020) 42:2000094. doi: 10.1002/bies.202000094
35. Vasanthakumar N. Can beta-adrenergic blockers be used in the treatment of COVID-19? *Med Hypotheses.* (2020) 142:109809. doi: 10.1016/j.mehy.2020.109809
36. Lee SW, Ha EK, Yeniova AÖ, Moon SY, Kim SY, Koh HY, et al. Severe clinical outcomes of COVID-19 associated with proton pump inhibitors: a nationwide cohort study with propensity score matching. *Gut.* (2020) 70:76–84. doi: 10.1136/gutjnl-2020-323672
37. Almario CV, Chey WD, Spiegel BM. Increased risk of COVID-19 among users of proton pump inhibitors. *Am J Gastroenterol.* (2020) 115:1707–15. doi: 10.14309/ajg.0000000000000798
38. Wei X, Zeng W, Su J, Wan H, Yu X, Cao X, et al. Hypolipidemia is associated with the severity of COVID-19. *J Clin Lipidol.* (2020) 14:297–304. doi: 10.1016/j.jacl.2020.04.008
39. Zhou Y, Frey TK, Yang JJ. Viral calciomics: interplays between Ca2+ and virus. *Cell Calcium.* (2009) 46:1–7. doi: 10.1016/j.ceca.2009.05.005
40. Chen X, Cao R, Zhong W. Host calcium channels and pumps in viral infections. *Cells.* (2020) 9:94. doi: 10.3390/cells9010094
41. Moreno-Altamirano MM, Kolstoe SE, Sánchez-García FJ. Virus control of cell metabolism for replication and evasion of host immune responses. *Front Cell Infect Microbiol.* (2019) 9:95. doi: 10.3389/fcimb.2019.00095
42. Nieto-Torres JL, Verdiá-Báguena C, Jimenez-Guardeño JM, Regla-Nava JA, Castaño-Rodriguez C, Fernandez-Delgado R, et al. Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. *Virology.* (2015) 485:330–9. doi: 10.1016/j.virol.2015.08.010
43. Wang G, Zhang Q, Zhao X, Dong H, Wu C, Wu F, et al. Low high-density lipoprotein level is correlated with the severity of COVID-19 patients: an observational study. *Lipids Health Dis.* (2020) 19:1–7. doi: 10.1186/s12944-020-01382-9
44. Fan J, Wang H, Ye G, Cao X, Xu X, Tan W, et al. Low-density lipoprotein is a potential predictor of poor prognosis in patients with coronavirus disease 2019. *Metabolism.* (2020) 107:154243. doi: 10.1016/j.metabol.2020.154243
45. Hu X, Chen D, Wu L, He G, Ye W. Declined serum high density lipoprotein cholesterol is associated with the severity of COVID-19 infection. *Clin Chim Acta.* (2020) 510:105–10. doi: 10.1016/j.cca.2020.07.015
46. Abu-Farha M, Thanaraj TA, Qaddoumi MG, Hashem A, Abubaker J, Al-Mulla F. The role of lipid metabolism in COVID-19 virus infection and as a drug target. *Int J Mol Sci.* (2020) 21:3544. doi: 10.3390/ijms21103544
47. Yan B, Chu H, Yang D, Sze KH, Lai PM, Yuan S, et al. Characterization of the lipidomic profile of human coronavirus-infected cells: implications for lipid metabolism remodeling upon coronavirus replication. *Viruses.* (2019) 11:73. doi: 10.3390/v11010073
48. Guo C, Chi Z, Jiang D, Xu T, Yu W, Wang Z, et al. Cholesterol homeostatic regulator SCAP-SREBP2 integrates NLRP3 inflammasome activation and cholesterol biosynthetic signaling in macrophages. *Immunity.* (2018) 49:842–56. doi: 10.1016/j.immuni.2018.08.021
49. Lee W, Ahn JH, Park HH, Kim HN, Kim H, Yoo Y, et al. COVID-19-activated SREBP2 disturbs cholesterol biosynthesis and leads to cytokine storm. *Signal Transduction Target Ther.* (2020) 5:1. doi: 10.1038/s41392-020-00292-7
50. Daniloski Z, Jordan TX, Wessels HH, Hoagland DA, Kasela S, Legut M, et al. Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell.* (2020) 184:92–105.e16. doi: 10.1016/j.cell.2020.10.030
51. Thomas T, Stefanoni D, Reisz JA, Nemkov T, Bertolone L, Francis RO, et al. COVID-19 infection results in alterations of the kynurenine pathway and fatty acid metabolism that correlate with IL-6 levels and renal status. *medRxiv.* (2020). doi: 10.1101/2020.05.14.20102491
52. Singh VP, Khatua B, El-Kurdi B, Rood C. Mechanistic basis and therapeutic relevance of hypocalcemia during severe COVID-19 infection. *Endocrine.* (2020) 70:461–2. doi: 10.1007/s12020-020-02530-y
53. Feingold KR, Grunfeld C. The Effect of Inflammation and Infection on Lipids and Lipoproteins. In: Feingold KR, Anawalt B, Boyce A, et al., editors. *Endotext [Internet].* South Dartmouth, MA: MDText.com, Inc. (2000). Available online at: https://www.ncbi.nlm.nih.gov/sites/books/NBK326741/
54. Greineisen WE, Speck M, Shimoda LM, Sung C, Phan N, Maaetoft-Udsen K, et al. Lipid body accumulation alters calcium signaling dynamics in immune cells. *Cell Calcium.* (2014) 56:169–80. doi: 10.1016/j.ceca.2014.06.004
55. Mori S, Ito H, Yamamoto K. Effects of calcium antagonists on low density lipoprotein metabolism in human arterial smooth muscle cells. *Tohoku J Exp Med.* (1988) 154:329–33. doi: 10.1620/tjem.154.329
56. Ranganathan S, Harmony JA, Jackson RL. Effect of Ca2+ blocking agents on the metabolism of low density lipoproteins in human skin fibroblasts. *Biochem Biophys Res Commun.* (1982) 107:217–24. doi: 10.1016/0006-291X(82)91691-6
57. Jain A, Chaurasia R, Sengar NS, Singh M, Mahor S, Narain S. Analysis of vitamin D level among asymptomatic and critically ill COVID-19 patients

and its correlation with inflammatory markers. *Sci Rep*. (2020) 10:1–8. doi: 10.1038/s41598-020-77093-z

58. Vyas N, Kurian SJ, Bagchi D, Manu MK, Saravu K, Unnikrishnan MK, et al. Vitamin D in prevention and treatment of COVID-19: current perspective and future prospects. *J Am College Nutr*. (2020) 1–14. doi: 10.1080/07315724.2020.1806758
59. Mazziotti G, Lavezzi E, Brunetti A, Mirani M, Favacchio G, Pizzocaro A, et al. Vitamin D deficiency, secondary hyperparathyroidism and respiratory insufficiency in hospitalized patients with COVID-19. *J Endocrinol Investig*. (2021) doi: 10.1007/s40618-021-01535-2. [Epub ahead of print].

Conflict of Interest: SD, MP, KT, VB, GM, MS, and SG were employed by company PrecisionLife Ltd. TH and KTHT were employed by company OptumLabs at UnitedHealth Group.

frontiers in Digital Health

PERSPECTIVE
published: 03 August 2021
doi: 10.3389/fdgth.2021.647938

Integrated Care in the Era of COVID-19: Turning Vision Into Reality With Digital Health

Angelina Kouroubali†, Haridimos Kondylakis† and Dimitrios G. Katehakis†*

Institute of Computer Science, Foundation for Research and Technology-Hellas (FORTH-ICS), Heraklion, Greece

The lives of millions of people have been affected during the coronavirus pandemic that spread throughout the world in 2020. Society is changing establishing new norms for healthcare education, social life, and business. Digital health has seen an accelerated implementation throughout the world in response to the pandemic challenges. In this perspective paper, the authors highlight the features that digital platforms are important to have in order to support integrated care during a pandemic. The features of the digital platform *Safe in COVID-19* are used as an example. Integrated care can only be supported when healthcare data is available and can be sharable and reusable. Healthcare data is essential to support effective prevention, prediction, and disease management. Data available in personal health apps can be sharable and reusable when apps follow interoperability guidelines for semantics and data management. The authors also highlight that not only technical but also political and social barriers need to be addressed in order to achieve integrated care in practice.

Keywords: integrated care, interoperability, digital health, COVID-19, FHIR, eHealth, EHR

OPEN ACCESS

Edited by:
Pradeep Nair, Central University of Himachal Pradesh, India

Reviewed by:
Theodoros N. Arvanitis, University of Warwick, United Kingdom
Isaac Cano Franco, Institut de Recerca Biomèdica August Pi i Sunyer (IDIBAPS), Spain

***Correspondence:**
Angelina Kouroubali
kouroub@ics.forth.gr

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *30 December 2020*
Accepted: *28 June 2021*
Published: *03 August 2021*

Citation:
Kouroubali A, Kondylakis H and Katehakis DG (2021) Integrated Care in the Era of COVID-19: Turning Vision Into Reality With Digital Health. Front. Digit. Health 3:647938. doi: 10.3389/fdgth.2021.647938

INTRODUCTION

The coronavirus disease 2019 (COVID-19) has caused a worldwide pandemic (1), that has disrupted the lives of millions of people across the globe (2). Confinement measures have focused on the reduction of the spread of the epidemic and the minimization of the load of morbidity and mortality so that healthcare systems remain functional (3). Healthcare measures have focused on treatment, case isolation and contact tracing. Many digital tools have been developed to assist with contact tracing and case identification (4). Despite the fact that it is mobile solutions that enable the automatic detection of contacts, saving precious hours of work of public health staff, still several concerns have been raised from the very beginning in regards to legal and technical safeguards (5). In addition, several solutions have been built in isolation without taking into consideration interoperability specifications for rendering the collected data sharable and reusable. Interoperable systems can make data sharable and reusable introducing many opportunities for growth and improvement. Eliminating data silos and automating data integration supports the recognition of unseen patterns, offers opportunities to apply new intelligence to service patients and care-givers, creating value across the care continuum (6).

Due to the burden of multi-morbidity, which is growing in the European countries, because of the aging population, an increasing number of people find themselves in having to deal with more than one chronic or long-term conditions (7–10). This brings the need for the provision of integrated care (11, 12), in order for people to receive appropriate treatment across the continuum

of healthcare delivery, including rapid response in cases of a pandemic break that may affect disproportionally heavy patients with certain chronic conditions.

This paper presents the authors' perspective on how digital technologies can help make integrated care a reality addressing the needs of governments to efficiently respond to pandemics such as COVID-19. It elaborates on the requirements necessary so that digital platforms can effectually act as outbreak response tools. These include appropriate and effective data management and semantics as well as interoperable personal health apps that address the needs of citizens to effectively support contact tracing while staying safe in the pandemic. In addition, the authors highlight the existing technological, political and social barriers that need to be overcome for the effective digital management of the pandemic.

In order to illustrate this perspective, the authors take as an example *Safe in COVID-19* (13), a digital platform designed to address the need for sharing information between patients and physicians as well as providing data for public health authorities. The architecture of the platform has already been published in Kouroubali et al. (13), and is based upon existing work on personal health record systems (14, 15) and the development of integrated care solutions to effectively support personal health management and public health (16). We present the modules that such an app should employee for the effective management of COVID-19 and we elaborate on its effective promotion through the available political and social channels, presenting also the many potential benefits.

DIGITAL HEALTH IN COVID-19

Mobile technologies have been widely developed as a response to the pandemic with a variety of features (17). There are technologies that facilitate self-assessment at home, training, information sharing, contact tracing, and decision making, swiftly offering effective and usable tools against the COVID-19 pandemic. Some examples include platforms, like COVIDSafe in Australia (https://www.health.gov.au/resources/apps-and-tools/covidsafe-app) that offer the ability to document registered isolation and to allow public health to conduct appropriate analysis and research; COVID Symptom Tracker (https://covid.joinzoe.com/us) in the US, that tracks the onset and progression of symptoms to pinpoint virus hot spots and to help slow the spread of the disease; MaladieCoronavirus (https://maladiecoronavirus.fr/) in France for symptom checking and self management; Corona-Warn-App (https://www.coronawarn.app/en/) for tracing the spread and providing information; Covid19.es (https://covid19.es/) in Spain for symptom checking and tracing the spread; Helsenorge (https://www.helsenorge.no/) for symptom checking, tracing and information providing; and the software application Go.Data (https://www.who.int/godata), developed by the Global Outbreak Alert and Response Network (GOARN) to manage case-contact relationships and the follow-up of contacts. Go.data has been deployed to over 35 countries in support of the COVID-19 Pandemic response. In addition, several countries, such as France, Croatia, Germany, and Norway have already developed e-government apps for their citizens, under the direction or with the endorsement of public organizations (18).

A systematic survey (17) has shown that mobile apps can benefit citizens, health professionals and decision makers in facing the COVID-19 pandemic. The challenges that these apps have addressed, include increasing the reach of reliable information to both citizens and health professionals as well as reducing misinformation and confusion. App features include symptom tracking, mental health support, and home-monitoring. Additionally, the information collected in the mobile apps can help discover new predictors about the course of the disease, optimize healthcare resource allocation and reduce the burden to hospitals (17).

Although massive numbers of data are collected in digital health systems, they exist in isolated data islands that are not widely available and hence cannot be used to enable further statistical analysis and prediction making. Patient data, even when available, they cannot be readily used by mobile apps. Data structure and form are often not appropriate for data exchange. The COVID-19 pandemic has created an even greater need for healthcare professionals to gain knowledge about how their patients' health is evolving in real time. Also, decision makers need to be able to monitor citizens who report COVID-19 related symptoms but have not been tested with a diagnostic test for COVID-19.

Implementing remote care monitoring is likely to generate long-term benefits and help with the healthcare challenges as part of the national healthcare system. Despite the potential benefits of digital solutions, their adoption has been slow. Digital solution developers posit technical and privacy issues as the main reasons for the slow adoption of digital solutions (17). Non-functional specifications essential for the delivery of trustworthy apps are relevant to compliance with GDPR provisions, role-based access to patient depending on end-user roles, ensuring accuracy and security of the data, the ability to communicate with other applications and registries using international standards (e.g., COVID-19 registries, nationally approved terminologies, citizen and healthcare professional identification services, etc.), as well as compliance with approved medical protocols.

In addition, a full range of features is not included in most solutions (19). An important need includes the development and utilization of decision support tools to assist policy optimization for minimizing negative socio-economic impact, while contributing to the containment of the pandemic. Furthermore, the pandemic has revealed preparedness shortcomings in public health (3). Social, organizational, and technological factors need to be addressed in order to allow the adoption of these eHealth tools. Further steps will include the involvement of patients to further elaborate requirements for usable and meaningful solutions. Uptake of the mobile eHealth resources will require a significant change in management efforts and the redesign of existing models of care (20–22).

It is important to note that even though digital solutions for remote care can be a valuable alternative to the isolation conditions imposed during a pandemic crisis, they cannot substitute face to face physical examinations which include

human contact and communications (23). In addition, financial and legal constrains in technology readiness of healthcare systems can also act as a barrier to significant and long-term digital transformation, despite the rare opportunity of the current crisis (24). Integration with third-party applications would enable the use of the platform as an extension of the already existing hospital information systems enhancing the digital management of COVID-19 cases (25–27).

Safe in COVID-19 digital platform intends to respond to the challenges leveraging on existing experience and robust implemented technologies. It is built on an extensible architecture focusing on the needs of all involved stakeholders including citizens and their families, healthcare professionals and public authorities. The platform addresses user information and communication needs (13). Data is collected and stored in such a way so that semantic and big data technologies can utilize them to enable, decision support and appropriate proximity tracking when needed. Further details are provided in the next section whereas **Table 1** summarizes the differences of some of the available applications in the area, demonstrating the advantages of the Safe in COVID-19.

SAFE IN COVID-19 DIGITAL PLATFORM

Safe in COVID-19 is an expandable digital platform that was developed in an effort to support national authorities in Greece. The platform provides tools for public authorities, healthcare professionals, citizens and their families offering an integrated care perspective to the ongoing pandemic.

The platform architecture is shown in **Figure 1** and consists of the Application Tier providing the front-end applications to the end-users, and the Business Logic Tier offering the intelligent functionality (13). All these layers are supported by security and integrity services as described below.

Application Tier

The *Safe in COVID-19* solution includes the following applications: (i) a web application for public health authorities to give a complete picture of the status regarding the spread of the disease at a national level and the measures taken by healthcare services; (ii) a web application for healthcare professionals that supports online communication with registered patients in order to provide personalized information and coaching and instant access to patient reported symptoms related to COVID-19 disease for monitoring their health status; and (iii) a mobile application (Android/iOS) for citizens that allows recording on a daily basis their health status and communicate synchronously

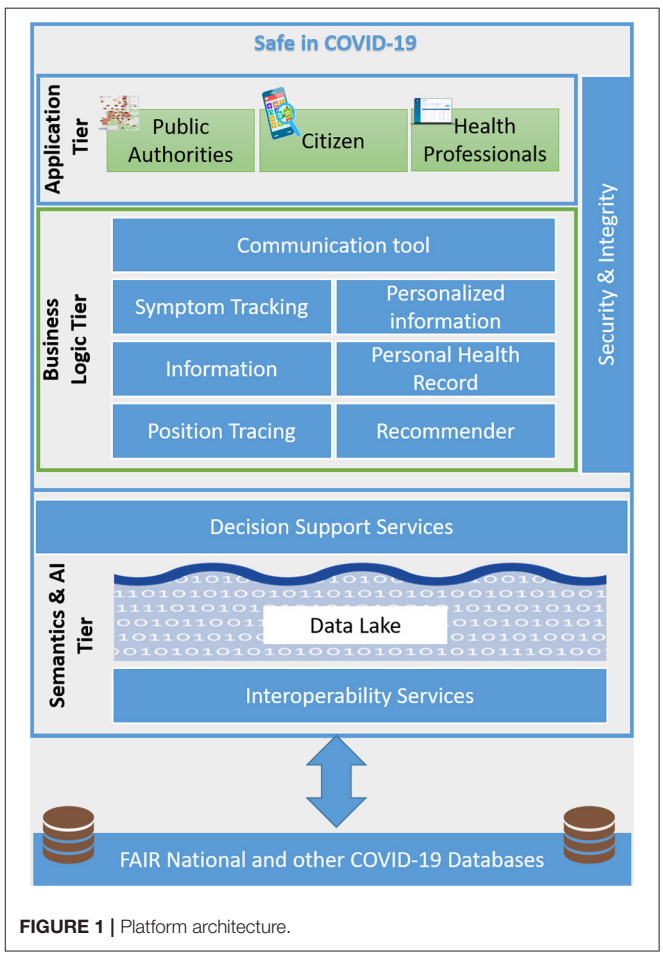


FIGURE 1 | Platform architecture.

TABLE 1 | A comparison of the features of key platforms for COVID-19.

	Country	Symptom checker and self diagnosis	Tracing the spread	Providing information	Home support for self-management	Communication with medical staff	AI enabled	Interoperable
Safe in COVID-19	Greece	x		x	x	x	x	x
COVIDsafe	Australia		x	x				
COVID symptom tracker	US	x	x					
Maladie coronavirus	France	x			x			
Corona-Warn-App	Germany		x	x				
Helsenorge	Norway	x	x	x				
COVID-19.es	Spain	x	x		x			
Go. Data	More than 35 countries		x					

or asynchronously with healthcare professionals in order to receive personalized instructions for managing their health. Certainly, web apps could be used instead of using smart phones for the citizens. However, we believe that packing the app as a mobile app makes it easier for users even without appropriate digital skills to navigate and enjoys additional integration with their mobile phone (push messages etc.).

Business Logic Tier

This tier includes the necessary services for handling the logic behind the graphical user interface. More specifically as the *Safe in COVID-19* framework is based on and significantly extends a fully-fledged, state of the art PHR system, it includes the personal health records and user profiling out of the box. However, specific modules have been customized for COVID-19. The symptom module has been extended to include specific symptoms of COVID-19, and the communication module has been extended to enable continuous communication with the healthcare team. Other modules such as the position tracing and the recommendations have been implemented specifically for addressing the COVID-19 outbreak. All collected data are stored in the data lake, enabling visual exploration through the application tier. Different levels of care and health services are brought together through online interaction with national digital health portals that provide secure application programming interfaces for patient, healthcare professional and insurance organizations applications. Typical services include access to national EHR services, ePrescription and national registries with COVID-19 patients.

Security and Integrity

Handling personal health data requires specific attention to the security and the integrity of the collected data. As such, the *Safe in COVID-19* framework does not store specific patient identification data, but only unique codes generated randomly by the system and shared with the health authorities. In addition, those codes can also be generated by the health authorities for the confirmed COVID-19 cases. Data processing has been minimized only for the persons who need it whereas, all data is encrypted in order to further increase security and privacy. Further, every data access is recorded, whereas the necessary consent is also granted by the citizens with full information about the intended processing of data. In order to verify legitimate users (i.e., entering on purpose false positives, undermining this way faith in the system) users may need to receive appropriate codes when entering data into the app to be confirmed by the server. Position tracking is legitimized by relying on a voluntary adoption by the users for each of the intended purposes.

Semantics and AI Tier

Further, the *Safe in COVID-19* framework goes beyond simple data dashboards, by providing intelligent dashboards that enable the multidimensional exploration and filtering of the available data. Internally we have (a) external data directly ingested in the data lake (b) the database of the adapted PHR-C (the schema is based on the Open EHR). All data are mapped to an ontology, which enables the semantic uplifting of the available data and their homogenization and integration. The ontology currently used is an extension of the MHA semantic core ontology (28), a high-level ontology integrating several standard taxonomies and terminologies used for modeling health data. We have extended it with just minor additions to completely cover our needs for COVID-19. Exploiting the semantic meaning of the data machine learning algorithms and epidemiological models can be incorporated in order to make predictions on virus spread on a specific geographic region, combining the available data with external regional, country and health-system data that might be available. Being able to predict a few days in advance the number of beds needed is essential for health organizations to have time to organize resources to meet the health needs of their reference population (outsourcing of services, staffing, hotel sanitation, purchase of respirators, etc.). We have to note that all modules in the business logic tier access a homogeneized, semantically uplifted and cleaned version of the available data staged in the data lake. However, we also maintain the raw staged data as we wouldn't like to lose any important information through errors in data cleaning and we would like to be able to repeat the cleaning steps if needed.

DIGITAL PLATFORMS AS OUTBREAK RESPONSE TOOLS

Establishing a close cooperation with public authorities is imperative in order to put the necessary applications into operation, since public health falls under their jurisdiction. Specific details about the development and deployment of COVID-19 relevant solution at a national level can only be established through compliance with approved medical protocols, interoperability with national registries for citizen identification and COVID-19, quality assurance, and the existence of the appropriate legal framework. Digital platforms can be effective when several interrelated factors are in place such as: (i) a public health system with a comprehensive national epidemiologic strategy (ii) a technology and architecture model that promotes interoperability for data sharing, and data re-use; (iii) widespread connection with mobile devices, and (iv) the appropriate regulatory and legislative conditions for the use of integrated digital solutions for all stakeholders safeguarding safety and privacy for all.

To leverage the individual benefits of mobile apps and other digital technologies it is important to have an interoperability layer to enable the various such solutions to exchange data. A unifying layer will facilitate the exchange of meaningful data in the appropriate format and needs to be implemented on top of the individual proposed solutions. All collected data should be made immediately findable and accessible through fully anonymization and publishing on the web using digital object identifiers (DOIs) and other permanent identifiers. This process is called data FAIRification.

Relevant ontologies, such as the MHA semantic core ontology (28), and other and approaches can be used for semantically uplifting available data through mapping and/ or annotating using ontology terms. This allows the rapid re-use of the

available data, enabling a common understanding and offering a rich set of terms for documenting and adding meta-data to the data provided by the platform. Appropriately homogenizing and integrating both app-specific data with external data from regional, country and health-system data sources can leverage and combine available knowledge for artificial intelligence models, enabling better and more useful predictions. Multiple epidemiological models are currently being developed for COVID-19 and could exploit the available data to make predictions on the disease spread further guiding healthcare services and personnel allocation. On the other hand, early patient identification through machine learning and the incorporation of risk models in such apps has a great potential for better helping those in higher risks.

The use of fast healthcare interoperability resources (FHIR) can enable and facilitate interoperability with existing hospital health systems. FHIR should be used for the representation of the medical data related to COVID-19 (29). COVID-19 Patient Reported Outcome Observations (30) have already been established including symptoms such as cough, fatigue, pain in throat, dyspnea, headache, diarrhea, nausea, loss of sense of smell, and temperature. FHIR resources such as Problem and Condition, are also appropriate for representation of the underlying diseases related to COVID-19. These resources can be automatically published and data properly anonymized through a FHIR Server. Any hospital information system may act as a data source for the *Safe in COVID-19* FHIR architecture, since they can publish patient laboratory results for COVID-19 tests into the FHIR Server. Available applications then will be able retrieve the test results from the FHIR server and present them to the end users.

Privacy and trust are key concerns for individuals considering the adoption of personal health systems. As data are expected to be directly shared upon their creation appropriate anonymization techniques should be enforced for sharing information among patients, and their relatives, doctors, researchers, policy makers and governments. Digital solutions need to provide appropriate explanations to citizens, highlighting the importance of sharing health information for "flattening the curve" and explaining all measures taken to ensure anonymity and confidentiality.

Extending existing digital services with novel solutions can increase public value. Efficiency and effectiveness are facilitated through the use of wider range of available tools such as the provision of legitimate user orientation, high level of participation, legality, and equity. Responses to COVID-19 have largely not been integrated, leading to adverse outcomes. Innovative digital solutions can exhibit their highest impact when they integrate and interoperate with existing healthcare services and infrastructures.

DISCUSSION AND CONCLUSION

Digital platforms can act as disease outbreak support tools when they are able to integrate within existing digital health services. This paper presented the overall requirements and conditions that are essential in order to develop digital platforms that can be useful to all healthcare stakeholders including public health authorities, healthcare professionals and citizens (31). The *Safe in COVID-19* platform was used to illustrate in practice what the healthcare ecosystem can gain form an integrated innovative digital service, demonstrating the components that should be available and the challenges for adopting them in practice.

Providing the means to securely link reliable data has the potential to support care anywhere at any time. When coupled with advanced analysis, then digital systems can be transformed into smart systems to better support integrated care. Of course, relevant challenges linked to organizational, financial, legal and interoperability issues will always need to be properly addressed at different configuration settings.

Integrating novel digital solutions can provide opportunities to support and strengthen health systems in the diagnosis, treatment, monitoring, citizen empowerment through information, public health surveillance and epidemiology (4, 12, 32).

Foreseen benefits for public health authorities include decongestion of healthcare units, provision of real-time information on the evolution of suspected, candidate and confirmed cases, online monitoring of the spread of the virus, and effective decision-making regarding required measures. Benefits to healthcare professionals include support in managing patients, reduced time for direct contact with patients, more efficient case management, and improved working conditions. Benefits for citizens include systematic recording of symptoms, self-assessment, access to personalized information, and instructions and reminders based on their overall health status.

Safe in COVID-19 offers an example of an integrated platform offering multiple benefits for the users in alignment with relevant national eHealth initiatives. It can operate efficiently under the appropriate regional or national infrastructures with links to medical and diagnostic centers around the country and the size of the population under consideration. Such platforms can support the return to the "new normal" with less stress and more security for individuals, more direct and safer management of patients by physicians, and better possibilities for monitoring the epidemic by public health authorities.

The healthcare systems can act as drivers of change and facilitate the adoption of similar eHealth technologies based on need. As technology adoption barriers have decreased due to the current pandemic (33), digital transformation in healthcare can accelerate the use of novel technologies. eHealth solutions can be implemented rapidly and can offer essential tools in supporting the COVID-19 pandemic for all stakeholders including citizens, healthcare providers, policy makers and governments, promoting coordination to ensure appropriate management of this crisis. Future research could focus on the assessment of the cost-effectiveness of digital health apps for the management of COVID-19 with the introduction of key performance indicators.

Digital solutions can help support integrated care in the era of COVID-19, on a larger scale. Healthcare systems need to focus on flexibility and interoperability to achieve sustainability

and to realize the vision for healthcare at the point of need. Turning vision into reality requires taking the responsibility to act now to ensure health systems and people continue to benefit.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

The work presented in this paper has been funded by the Center for eHealth Applications and Services, Institute of Computer Science, Foundation for Research and Technology–Hellas.

REFERENCES

1. Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Agents.* (2020) 55:105924. doi: 10.1016/j.ijantimicag.2020.105924
2. Haleem A, Javaid M, Vaishya R. Effects of COVID 19 pandemic in daily life. *Current Med Res Pract.* (2020) 10:78–9. doi: 10.1016/j.cmrp.2020.03.011
3. Mahmood S, Hasan K, Carras MC, Labrique A. Global preparedness against COVID-19: We must leverage the power of digital health. *JMIR Public Health and Surveillance.* (2020) 6:e18980. doi: 10.2196/18980
4. WHO (World Health Organization). *Digital Tools for COVID-19 Contact Tracing: Annex: Contact Tracing in the Context of COVID-19, 2 June 2020*. Geneva: World Health Organization (2020).
5. Council of Europe PA, Walter JP. Joint Statement on Digital Contact Tracing [Online]. (2020). Available online at: https://rm.coe.int/covid19-joint-statement-28-april/16809e3fd7 (accessed July 16, 2021).
6. Gopal G, Suter-Crazzolara C, Toldo L, Eberhardt W. Digital transformation in healthcare–architectures of present and future information technologies. *Clin Chem Lab Med.* (2019) 57:328–35. doi: 10.1515/cclm-2018-0658
7. Glynn LG, Valderas JM, Healy P, Burke E, Newell J, Gillespie P, et al. The prevalence of multimorbidity in primary care and its effect on health care utilization and cost. *Fam Pract.* (2011) 28:516–23. doi: 10.1093/fampra/cmr013
8. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet.* (2012) 380:37–43. doi: 10.1016/S0140-6736(12)60240-2
9. Rijken M, Hujala A, van Ginneken E, Melchiorre MG, Groenewegen P, Schellevis F. Managing multimorbidity: profiles of integrated care approaches targeting people with multiple chronic conditions in Europe. *Health Policy.* (2018) 122:44–52. doi: 10.1016/j.healthpol.2017.10.002
10. Czypionka T, Kraus M, Reiss M, Baltaxe E, Roca J, Ruths S, et al. The patient at the centre: evidence from 17 European integrated care programmes for persons with complex needs. *BMC Health Serv Res.* (2020) 20:1–14. doi: 10.1186/s12913-020-05917-9
11. WHO (World Health Organization) Regional Office for Europe (2013). *Roadmap Strengthening People-Centred Health Systems in the WHO European Region. A Framework for Action Towards Coordinated/Integrated Health Services Delivery (CIHSD)*. Copenhagen: Regional Office for Europe.
12. WHO (World Health Organization). *WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening*. Geneva: WHO (2019).
13. Kouroubali A, Kondylakis H, Kavlentakis G, Logothetides F, Stathiakis N, Petrakis Y, et al. An eHealth platform for the holistic management of COVID-19. *Stud Health Technol Inform.* (2020) 273:182–8. doi: 10.3233/SHTI200636
14. Katehakis DG, Kondylakis H, Koumakis L, Kouroubali A, Marias K. Integrated care solutions for the citizen: personal health record functional models to support interoperability. *EJBI.* (2017) 13:41–56. doi: 10.24105/ejbi.2017.13.1.8
15. Kouroubali A, Koumakis L, Kondylakis H, and Katehakis DG. An integrated approach towards developing quality mobile health apps for cancer. In: *Mobile Health Applications for Quality Healthcare Delivery*. Hershey, PA: IGI Global (2019). 46–71. doi: 10.4018/978-1-5225-8021-8.ch008.ch003
16. Katehakis DG, Kouroubali A, Karatzanis I, Manousos D, Kondylakis H, Kavlentakis G, et al. *Personal Health ICT Systems to Support Integrated Care Solutions, Technical Report*. Crete: Heraklion (2018).
17. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. *J Med Internet Res.* (2020) 22:e23170. doi: 10.2196/23170
18. European Commission. *Mobile Contact Tracing Apps in EU Member States.* (2021). Available onlne at: https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/travel-during-coronavirus-pandemic/mobile-contact-tracing-apps-eu-member-states_en (accessed May 17, 2021).
19. Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). *COVID-19 Dashboard.* (2020). Available online at: https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6 (accessed May 17, 2021).
20. Kouroubali A. *Implementation of Health Care Information Systems: Key Factors and the Dynamics of Change*. Cambridge, UK: University of Cambridge (2004).
21. Tsiknakis M, Kouroubali A, Vourvahakis D, and Orphanoudakis SC. Implementing a Regional Health Information Network: impact on health care performance and the management of change. In: Savage GT, Chilingerian JA, Powell M, Xiao Q, editors. *International Health Care Management*. Bingley: Emerald Group Publishing Limited (2005). p. 297–329. doi: 10.1016/S1474-8231(05)05011-1
22. Triantafyllidis A, Kondylakis H, Votis K, Tzovaras D, Maglaveras N, Rahimi K. Features, outcomes, and challenges in mobile health interventions for patients living with chronic diseases: a review of systematic reviews. *Int J Med Inform.* (2019) 132:103984. doi: 10.1016/j.ijmedinf.2019.103984
23. Webster P. Virtual health care in the era of COVID-19. *Lancet.* (2020) 395:1180–1. doi: 10.1016/S0140-6736(20)30818-7
24. Omboni S. Telemedicine during the COVID-19 in Italy: a missed opportunity? *Telemedicine and e-Health.* (2020) 26:973–5. doi: 10.1089/tmj.2020.0106
25. Kouroubali A, Starren J, Barrows RC, Clayton PD. Practical lessons in remote connectivity. *Proc AMIA Annu Fall Symp.* (1997) 335–339.
26. Triantafyllidis AK, Koutkias VG, Chouvarda I, Adami I, Kouroubali A, Maglaveras N. Framework of sensor-based monitoring for pervasive patient care. *Healthc Technol Lett.* (2016) 3:153–8. doi: 10.1049/htl.2016.0017
27. Kouroubali A, Katehakis DG. The new European interoperability framework as a facilitator of digital transformation for citizen empowerment. *J Biomed Inform.* (2019) 94:103166. doi: 10.1016/j.jbi.2019.103166
28. Kondylakis H, Spanakis EG, Sfakianakis S, Sakkalis V, Tsiknakis M, Marias K, et al. Digital patient: personalized and translational data management through the MyHealthAvatar EU project. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Milan: IEEE (2015). p. 1397–400.
29. Petrakis Y, Kouroubali A, Katehakis D. A mobile app architecture for accessing EMRs using XDS and FHIR. In: *2019 IEEE 19th International Conference on*

Bioinformatics and Bioengineering (BIBE). Athens: IEEE (2019). p. 278–283. doi: 10.1109/BIBE.2019.00057
30. HL7 Switzerland. *ValueSet Covid19 Patient Reported Outcome Observations.* (2020). Available online at: http://build.fhir.org/ig/hl7ch/covid-19-prom/branches/master/ValueSet-covid-19-prom.html (accessed May 17, 2021).
31. Katehakis DG, Kouroubali A. A Framework for eHealth interoperability management. *J Strateg Innov Sustainability.* (2019) 14:51–61.doi: 10.33423/jsis.v14i5.2521
32. WHO (World Health Organization) Global strategy on digital health 2020–2025 [Online]. (2020). Available online at: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf (accessed July 16, 2021).
33. Sust PP, Solans O, Fajardo JC, Peralta MM, Rodenas P, Gabaldà J, et al. Turning the crisis into an opportunity: digital health strategies deployed during the COVID-19 outbreak. *JMIR Public Health Surveill.* (2020) 6:e19106. doi: 10.2196/19106

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

POLICY AND PRACTICE REVIEWS
published: 05 August 2021
doi: 10.3389/fdgth.2021.677929

Toward a Common Performance and Effectiveness Terminology for Digital Proximity Tracing Applications

Wouter Lueks[1], Justus Benzler[2], Dan Bogdanov[3], Göran Kirchner[2], Raquel Lucas[4], Rui Oliveira[5], Bart Preneel[6], Marcel Salathé[7], Carmela Troncoso[1†] and Viktor von Wyl[8,9*†]*

[1] Security and Privacy Engineering Laboratory, School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [2] Robert Koch Institute, Berlin, Germany, [3] Cybernetica AS, Tartu, Estonia, [4] Medical School and Institute of Public Health (EPIUnit), Universidade Do Porto, Porto, Portugal, [5] Institute for Systems and Computer Engineering, Technology and Science & University of Minho, Porto, Portugal, [6] Department of Electrical Engineering, Katholieke Universiteit Leuven and IMEC, Leuven, Belgium, [7] Digital Epidemiology Laboratory, School of Life Sciences, School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Global Health Institute, Geneva, Switzerland, [8] Digital and Mobile Health Group, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland, [9] Institute for Implementation Science in Health Care, University of Zurich, Zurich, Switzerland

OPEN ACCESS

Edited by:
Constantinos S. Pattichis, University of Cyprus, Cyprus

Reviewed by:
Yun William Yu, University of Toronto Scarborough, Canada
Milena B. Cukic, Amsterdam Health and Technology Institute (AHTI), Netherlands

***Correspondence:**
Viktor von Wyl
viktor.vonwyl@uzh.ch
Carmela Troncoso
carmela.troncoso@epfl.ch

† These authors share last authorship

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *08 March 2021*
Accepted: *08 July 2021*
Published: *05 August 2021*

Citation:
Lueks W, Benzler J, Bogdanov D, Kirchner G, Lucas R, Oliveira R, Preneel B, Salathé M, Troncoso C and von Wyl V (2021) Toward a Common Performance and Effectiveness Terminology for Digital Proximity Tracing Applications. Front. Digit. Health 3:677929. doi: 10.3389/fdgth.2021.677929

Digital proximity tracing (DPT) for Sars-CoV-2 pandemic mitigation is a complex intervention with the primary goal to notify app users about possible risk exposures to infected persons. DPT not only relies on the technical functioning of the proximity tracing application and its backend server, but also on seamless integration of health system processes such as laboratory testing, communication of results (and their validation), generation of notification codes, manual contact tracing, and management of app-notified users. Policymakers and DPT operators need to know whether their system works as expected in terms of speed or yield (performance) and whether DPT is making an effective contribution to pandemic mitigation (also in comparison to and beyond established mitigation measures, particularly manual contact tracing). Thereby, performance and effectiveness are not to be confused. Not only are there conceptual differences but also diverse data requirements. For example, comparative effectiveness measures may require information generated outside the DPT system, e.g., from manual contact tracing. This article describes differences between performance and effectiveness measures and attempts to develop a terminology and classification system for DPT evaluation. We discuss key aspects for critical assessments of whether the integration of additional data measurements into DPT apps may facilitate understanding of performance and effectiveness of planned and deployed DPT apps. Therefore, the terminology and a classification system may offer some guidance to DPT system operators regarding which measurements to prioritize. DPT developers and operators may also make conscious decisions to integrate measures for epidemic monitoring but should be aware that this introduces a secondary purpose to DPT. Ultimately, the integration of further information (e.g., regarding exact exposure time) into DPT involves a trade-off between data granularity and linkage on the one hand, and privacy on the other. More data may lead to better epidemiological information but may also increase

the privacy risks associated with the system, and thus decrease public DPT acceptance. Decision-makers should be aware of the trade-off and take it into account when planning and developing DPT systems or intending to assess the added value of DPT relative to the existing contact tracing systems.

Keywords: contact tracing app, digital health, COVID-19, coronavirus, performance, effectiveness, digital proximity tracing

INTRODUCTION

Digital proximity tracing (DPT) is a novel health technology, designed to complement manual contact tracing (MCT) by using apps in national efforts to mitigate the Sars-CoV-2 pandemic (1). The primary purpose of DPT is to provide an instrument for fast, anonymous notification of other app users with potential exposure risks to an infected app user (2).

The current Sars-CoV-2 crisis is the first global public health crisis that sees massive, nationwide roll-outs of DPT apps (3). This is noteworthy because DPT has never undergone large-scale, real-world testing in its target population prior to release, as would normally be required for health technologies (4, 5). This is largely due to the urgency of the Sars-CoV-2 crisis, where many countries gave precedence to fast release over extensive testing. All the more, governments face pressure to justify the rapid DPT deployment and to demonstrate its impact on pandemic mitigation. Specifically, evaluations are needed to demonstrate whether single parts and the whole system of DPT perform well from a technical perspective, but also whether DPT helps to contain transmission chains (6). First data-driven studies of DPT effectiveness have started to emerge only recently (7–9).

Ideally, such evaluations of DPT should follow a standardized protocol to allow comparability across countries and settings, but also to facilitate learning from other countries' experiences. However, in direct exchanges and discussions with national health authorities and DPT developers, this group of authors noticed a substantial confusion among health authorities, politicians, and even DPT experts about the aims of DPT, the terminology, and goals for system evaluations. The metaphor of a "Babylonian confusion of tongues" is not too far to describe the current situation. This problem has been recognized, and international health authorities and different groups of academics have attempted to bring some structure into discussions about DPT development, deployment, and evaluation (10, 11). A starting point for discussion (including a glossary with relevant keywords) is presented in von Wyl et al. (6). Furthermore, Colizza et al. present some high-level recommendations regarding effectiveness evaluations of DPT apps (12). Specifically, they emphasize elements such as user update and adherence, speed of notification, but also transparency of DPT risk scoring algorithms and evaluations (12). In addition, experimental studies have been conducted to assess the performance of contact detection by DPT (13–16). The study by Rodriguez and colleagues is noteworthy because it also aimed at defining impact indicators for DPT apps, mostly for user behavior and exposure detection (16). But the study description does not reveal why and how the specific indicators were selected. Of further note, all these studies were conducted by researchers from different scientific backgrounds, and hence using their respective terminologies and performance indicator measures (e.g., from clinical research or computer science), with little consistency across studies.

The present viewpoint attempts to provide further clarifications on key aspects of DPT evaluations by bringing together DPT developers and public health experts from different countries to present a unified proposal for terminology and classification of measures to evaluate DPT. Thereby, we will focus on DPT apps that follow the privacy-preserving design principles outlined by the Decentralized Privacy-Preserving Proximity Tracing (DP-3T) (17) protocol for two reasons. First, the design principles serve—to our knowledge—as the basis for most currently deployed DPT apps [exceptions are, for example, the French TousAntiCovid app (18) or the TraceTogether app from Singapore (13)]. Second, decentralized, privacy-preserving DPT apps, as well as the voluntariness of their use, pose the greatest methodological challenges for monitoring and for designing effectiveness evaluation strategies due to the (intended) paucity of data.

The viewpoint is structured as follows. Section *Principles of Digital Proximity Tracing in Support of Manual Contact Tracing* describes the basic principles of DPT. Section *DPT Is a Complex Intervention* argues that DPT is a complex intervention, relying on the fast completion of clearly defined actions in the notification cascade by different health systems actors. Section *A Closer Look at DPT Steps and Their Influence on Intervention Outcomes* breaks the DPT notification cascade into its separate parts and describes how some basic questions and checks may easily be utilized in the DPT evaluation. Section *Basic Concepts for DPT Evaluations* introduces basic concepts and terminologies to describe and assess DPT systems from different viewpoints, namely system performance assessments and public health effectiveness evaluations. Section *Proposal for Classification of Different DPT Evaluation Measures* outlines a classification matrix to distinguish different types of indicator measures. Concrete indicator examples are provided and referenced in the **Supplementary Materials**. In section *Key Considerations for the Practical Implementation of Performance and Effectiveness Measures*, the viewpoint closes with some basic considerations for developing and implementing indicators for DPT evaluation.

PRINCIPLES OF DIGITAL PROXIMITY TRACING IN SUPPORT OF MANUAL CONTACT TRACING

The principles of DPT have been described extensively elsewhere (1). In brief, DPT enables participants to trace proximity contacts (exposures) that could pose relevant infection risks. If one of the proximity contact persons tests positive for Sars-CoV-2, the app will warn other users who were in close proximity during the infected person's time window of infectivity, thereby increasing the coverage and/or speed of the contact tracing process relative to MCT. The potential advantages of DPT compared with MCT are 3-fold. DPT can (#1) lead to faster exposure notifications than MCT, (#2) reach persons who are not personally known to an index case, and (#3) DPT is easily scalable and should still work when MCT reaches its capacity limits.

Most countries that have released DPT apps have opted for a privacy-preserving, decentralized architecture according to the DP-3T blueprint (17, 19). That is, proximity contacts are not sent to a central server, but stored and evaluated locally on smartphones. The only data that is sent to a central server are pseudonymous, random identifiers of persons with a confirmed SARS-CoV-2 infection. These "infectious" identifiers are downloaded by all other users and compared to the locally stored identifiers to find out which of the proximity encounters were with SARS-CoV-2 positive people. If the temporal aggregation of the matched encounters exceeds a minimal duration at a relevant proximity (depending on the estimated infectiousness of the positively tested person), users are notified and recommendations on further steps to take are provided.

The dominant choice of a privacy-preserving architecture across many countries highlights the emphasis of the primary DPT function ("warning people early in an anonymous manner") over purposes such as disease monitoring. DPT is, first and foremost, a notification tool aimed at breaking transmission chains, and its primary function does not necessitate the collection of personal data of index cases and their contacts. Nevertheless, debates in several countries suggest that DPT is sometimes also viewed as an opportunity to collect data for epidemiological monitoring, for example, to obtain additional information on time and setting of the events with high risk of exposure. Such *secondary functions* of DPT are beyond the scope of this article and are only discussed briefly where relevant for the broader context.

DPT IS A COMPLEX INTERVENTION

The preventive effect of DPT results from a timely warning of exposed persons so that they can enter quarantine and initiate further preventive measures. In this context, timely means a faster contact notification than is usual in MCT. In addition, DPT can also reach persons who would normally be missed by MCT (e.g., because they were chance encounters of the index case). As illustrated by **Figure 1**, the app notification process reflects an information flow in multiple steps to eventually produce specific actions leading to the prevention of further transmission (indicated by the #-signs in **Figure 1**). The effect of DPT depends on the interplay between health system actors (e.g., testing laboratories) and *app users*. It is not the app *per se* but the fast completion of the full notification cascade and subsequent actions that lead to the desired results (9). Of note, the distinction between app users and other actors is warranted because (voluntary) user actions are strongly influenced by behavioral aspects and incentives (20, 21), whereas actions required by other health system actors may depend more on automatization, technical interfaces, resources, or capacity (22).

We identify three high-level tasks (illustrated by colored boxes in **Figure 1**) that combined cover the entire notification cascade:

- **Proximity estimation**: This aspect pertains to the exchange of ephemeral (regularly changing) identifiers between users' devices, and the detection of significant proximity contacts (e.g., $<$1.5 m for more than 15 min).
- **Diagnosis and identifiers upload**: This aspect pertains to the upload of identifiers by index cases.
- **Notification of proximity contacts**: This aspect pertains to the notification of proximity contacts by their mobile device, and the subsequent actions taken by users as a response to this notification.

The dependency of the DPT intervention on its embedding in the overall pandemic mitigation response and the involvement of multiple actors fulfills the definition for complex interventions, as used in other fields of healthcare research (23, 24). In DPT, involved health system actors are setting-specific but may include testing laboratories or health authorities including MCT units or operators of infolines (**Figure 1**). **Figure 2** provides an even more detailed view of 10 individual steps in the DPT notification cascade. The red person illustrates the infected app user who gets tested, receives a positive test result, and triggers the app notification. The green person depicts a proximity contact who receives an app notification. Of note, in most DPT implementations, several of the steps in **Figure 2** involve free user choices whether or not to complete a specific task (e.g., step 6, authorization of key upload) without fears of retribution.

This insight that DPT is a complex intervention involving voluntary actions is relevant from a practical standpoint because it shifts the focus of discussion from single aspects (e.g., technical accuracy of Bluetooth measurements) to a broader systems perspective (25). However, in addition to measuring the final intended outcome of a complex intervention, the monitoring of individual components of a complex intervention is nonetheless important. Because complex interventions, DPT in particular, depend on a seamless, fast cascade of events (as shown in **Figure 1**), measurements characterizing speed and efficiency of specific system components and actors are useful to identify bottlenecks in the notification chain, as well as to act as leverage points for improving system behavior.

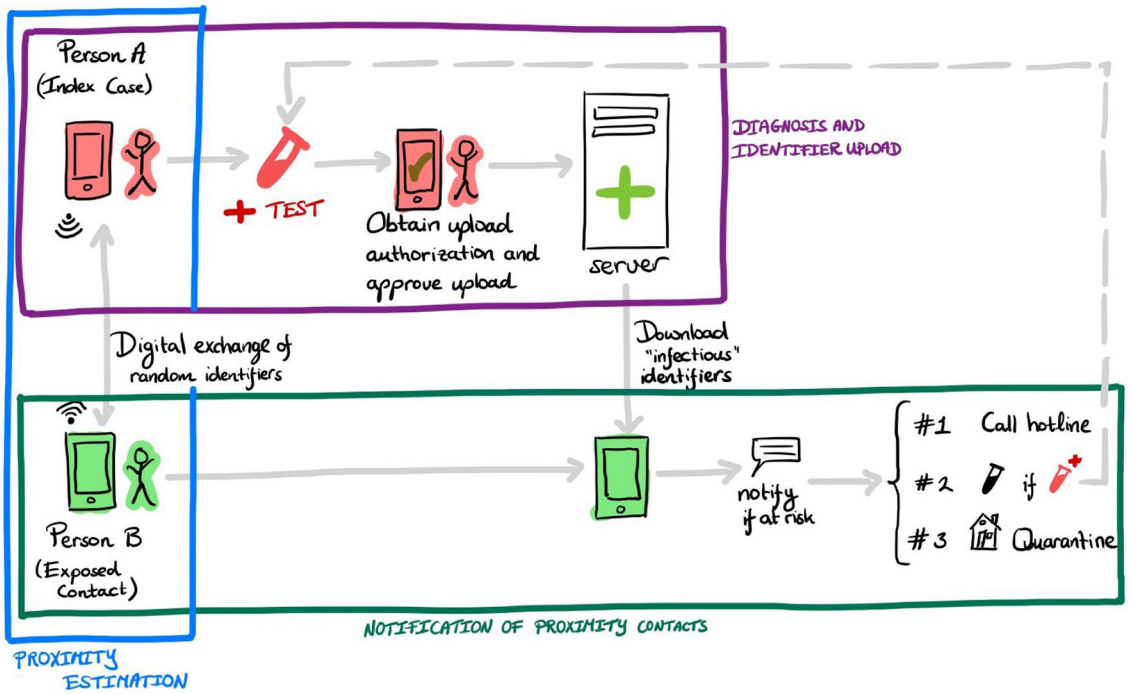


FIGURE 1 | Example of required steps in the notification cascade of digital proximity tracing in decentralized systems. An infected person A receives a positive test result for Sars-CoV-2 and (possibly automatically) an upload authorization. After consent by the user, the app uploads the random identifiers to a server. Person B's device regularly downloads "infectious" identifiers. If B was in close proximity to A (or other infected users) for a prolonged time, then B receives the app notification. Upon receiving this notification, B has several options, for example—in some countries—calling an information line (#1), getting tested for Sars-CoV-2 (#2), or entering quarantine voluntarily (#3). The colored boxes refer to the three main tasks involved in the DPT notification cascade, described in section *DPT Is a Complex Intervention*.

A CLOSER LOOK AT DPT STEPS AND THEIR INFLUENCE ON INTERVENTION OUTCOMES

To illustrate how system components can influence the outcome of the intervention, let us look at the three high-level tasks indicated in **Figure 1** above (*proximity estimation, diagnosis and identifiers upload, proximity contact notification*). The precise details of each of the tasks depend on national or regional choices regarding system design and configuration. However, all systems adhere to a similar, three-step structure. In the following, we will in greater detail describe the processes and country-specific variations, and identify questions that help to assess the system performance.

Proximity Estimation

All systems we consider in this paper rely on the Google/Apple Exposure Notification (GAEN) framework (26). Therefore, the accuracy of proximity estimation depends on the functioning of the GAEN framework and how well the chosen parameters reflect the desired measure of proximity (with slight regional variation).

Determining the accuracy of this component requires accurate ground-truth information about the exact distance and duration of a proximity contact. Collecting such information is nearly impossible in non-experimental settings without infringement of privacy (e.g., because it would require the use of video cameras to establish ground truth). Therefore, research groups (27–30) have used laboratory and simulated settings to replicate these scenarios with carefully constructed/measured ground-truth information. These measurements can then be used to answer questions such as:

- What parameter choices best reflect the desired distance/duration threshold?
- What distribution of false positives/false negatives does this choice induce for particular users' behaviors and under different environments?
- Does the choice of device model/manufacturer/platform influence the realized threshold?

To inform technical questions and setting-specific implementation decisions, the GAEN framework documentation is a suitable starting place (17).

Diagnosis and Identifier Upload

To enable proximity contacts to receive notifications, Sars-CoV-2-positive users must upload the random identifiers they broadcasted during the contagious window. Most countries use a system of upload authorization to allow uploads by confirmed index cases only, thereby preventing false warnings or

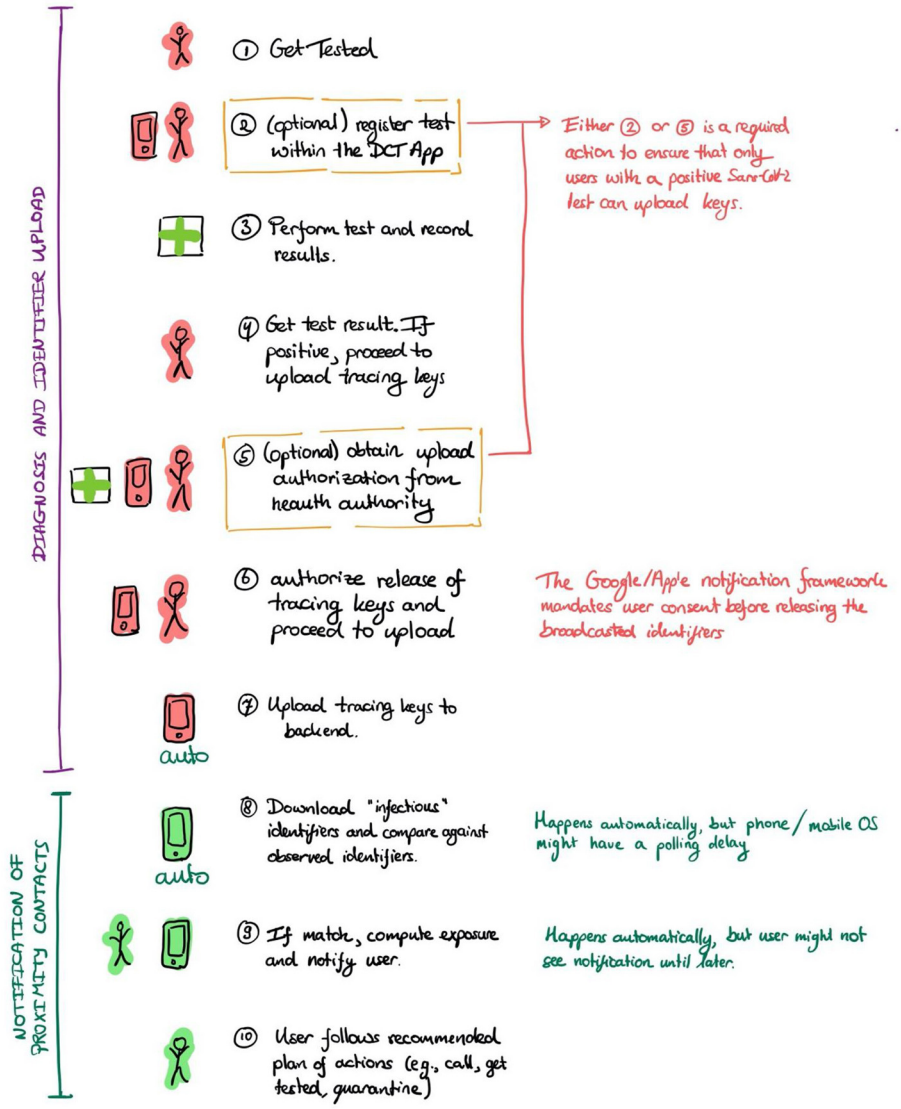


FIGURE 2 | Specific actions required in the notification cascade.

manipulation. As of today, automatic uploads of identifiers are not possible.

The following steps are considered critical (**Figure 2**):

- Get tested and receive a (positive) test result.
- Obtain upload authorization by a health authority (e.g., automatically when users registered the test in the app; or via an upload authorization code obtained through interaction with the health authorities after receiving a positive test).
- Upload identifiers by consenting to the release and upload of identifiers in the app.

The exact process of obtaining upload authorizations differs per country. Some countries, e.g., Belgium (31), and Germany (32), let users register Sars-CoV-2 tests in the app (who then also receive their test result through the app). Users with positive tests then automatically receive an upload authorization. In addition to or instead of such an automatic flow based on registered tests, other countries provide index cases with an upload authorization code that they enter into the app (Switzerland and Portugal use this as the only flow, German and Belgium use it as an alternative) or let users obtain authorizations via an eHealth system (e.g., in Estonia where the eHealth system also manages Sars-CoV-2 tests).

Moreover, the GAEN framework requires users to explicitly consent to release the random identifiers their phones have broadcast. Therefore, any app will request user action and explicit consent. However, the timing of consent provision varies across countries and can occur, for example, during laboratory test registration (e.g., in Germany) or upon receipt of a positive test result.

Based on the three critical processes described above, one could ask the following questions related to the operation of the system:

- How long between users getting tested and receiving a result?
- What fraction of users register tests in the app, if this function is available?
- What is the time from receiving the result to uploading identifiers?
- How many users with permissions to upload do not finish the process? At what point do they abandon the process?

While data collections to address these questions are highly setting-specific, publicly accessible examples of such monitoring systems already exist, for example, in Switzerland (9, 33), Germany (34), or in the Netherlands (35).

Notification of Proximity Contacts

Apps regularly retrieve uploaded positive identifiers from the system's backend server and then compare them (using the GAEN framework) against stored identifiers. If the framework detects a sufficiently long exposure, the app sends a notification to the end-user. Depending on the country, this notification instructs the user to contact the local health authority or a hotline, take a test, or self-quarantine. As with MCT, users may fail to follow these instructions.

To protect privacy (and to facilitate efficiency), uploaded identifiers are not immediately downloadable by other app end-users. Instead, the backend regularly releases a new batch of identifiers. For example, every 2 h.

The following factors influence the notification timeline.

- Time to publish. The time between receiving "positive" identifiers at the system's backend, and their publication, when these identifiers can be retrieved by other phones.
- Polling frequency. The frequency with which user's devices retrieve new "positive" identifiers and compute matches. This time is influenced by the mobile OS's scheduler as well as the internet connectivity of the phone.
- Time for the user to notice the notification and to subsequently act upon it.

The first and second factors are configurable parameters, constrained by the capabilities enabled by the GAEN framework (17). The effect of the system's scheduler can be tested in a laboratory setting.

The following questions can be asked concerning this notification task (some may not be possible to answer in privacy-preserving architectures):

- How many people receive a notification? How many of them later test positive?
- How many people follow through after receiving a notification?
- What is the time between positive upload and notification?
- How long between a notification is received and the user acts upon it?

The monitoring websites mentioned in section *Diagnosis and Identifier Upload* also provide good example metrics for the proximity contact notification step (33–35).

BASIC CONCEPTS FOR DPT EVALUATIONS

As illustrated above (**Figures 1**, **2**), DPT relies on a fast and seamless flow of information along the notification cascade. Blockades or inefficiencies in single steps can lead to bottlenecks and prevent the information flow, thus inhibiting the primary goal of DPT to warn other app users about potential risk exposures. The information flow of each DPT step can be described by at least three attributes: *speed, yield, and capacity*. *Speed* describes how fast an action is completed and can be measured in terms of time. *Yield* refers to a completion fraction, that is, a number of tasks executed as needed per 100. Yield sometimes also has a time connotation, that is the fraction of task completion for a given time frame (also referred to as *throughput*). *Capacity* relates to the task volume an actor can handle in a given time (which, in turn, may also influence speed and yield). For example, testing laboratories or manual contact tracers can only process a certain number of samples or cases, given the available resources such as machinery or personnel. Therefore, volumes can be described as percentages below or above the capacity limit.

The different DPT system components, as well as the basic attributes (*speed, yield, capacity*) and questions about the DPT functioning, lend themselves to the development of *Key Performance Indicators* (KPIs). For a better interpretation of these KPIs, a contextualization with the dynamic of the Sars-CoV-2 pandemic is often useful (e.g., to compare the number of key uploads with the incidence of new infections).

Overall, such KPIs provide valuable information on the procedural performance of the overall system, as well as possible bottlenecks in the notification cascade. If any of the components of the tasks in **Figure 1** (e.g., providing upload authorization) malfunctions, the delay will ripple to the whole notification cascade and undermine DPT's ability to perform its primary purpose of notifying exposed contacts. Therefore, KPIs can also be viewed as measures of technical and procedural preconditions for DPT to fulfill its primary purpose (9).

However, KPIs have the caveat that they often summarily reflect sequences of different actions (e.g., app usage, positive test, identifier uploads, download of identifiers, and proximity estimation). This composition complexity hinders the interpretation of these metrics as a consequence of a single factor.

Furthermore, KPIs are, in a strict sense, not revealing concerning how well DPT achieves its primary purpose of reducing viral transmissions, respectively its "*effectiveness,*" defined as the "ability to produce a desired result" (36). In epidemiological studies, the concept of effectiveness often stands for the real-world effect of an intervention against a comparator (comparative effectiveness) and is expressed as an exposure-outcome relationship.

It is important to note that, in the majority of countries, DPT is intended and designed to complement MCT. Therefore, in most settings, a "fair" DPT effectiveness evaluation would involve comparison between the use of DPT apps vs. non-use in addition to manual contact tracing. In those settings where DPT is implemented alongside MCT, effectiveness investigations should ideally center around the three postulated main advantages of DPT over MCT: DPT should lead to faster exposure notifications than MCT, DPT can reach persons who are not personally known to an index case, and DPT should still work when MCT reaches its resource limits (6). Specific effectiveness outcomes could focus, for example, on the time from the first exposure notification (either by MCT and/or DPT) to entering quarantine, or on the comparisons of the average number of persons who later test positive between groups who were notified by MCT or DPT. In both examples, the most obvious comparator is classic MCT. In settings where no MCT exists (or where it is no longer operable) and DPT is introduced in a staggered process, regional comparisons of Sars-CoV-2 incidence could be performed between geographic units that have introduced DPT at different times.

Sometimes DPT and DPT-related measurements are also discussed in the context of *epidemiological monitoring* (surveillance). Such discussions are tied to the hope of gathering relevant data and gaining insights about transmission dynamics. It is important to note that epidemiological monitoring is not part of the key functionality of DPT and necessitates an entirely different set of measurements and KPIs that go beyond the data requirements for privacy-preserving proximity tracing and it is not included in the following discussions.

PROPOSAL FOR CLASSIFICATION OF DIFFERENT DPT EVALUATION MEASURES

High-Level Distinction Between Key Performance Indicators and Public Health Effectiveness Metrics

Figure 3 proposes a *KPIs classification matrix* of different measure types and perspectives ("aspects") relevant for the DPT assessments. We acknowledge that the distinction between the different proposed types may not always be clean-cut in practice. In fact, one may need several of these metrics to assess how well the system performs each of the three tasks DPT systems must realize to fulfill their objective (see **Figure 1**). Nevertheless, the classification matrix may guide KPI development by illustrating different dimensions that comprehensive DPT monitoring systems should cover.

The horizontal dimensions of the classification matrix show different steps from a basic Input-Processing-Output (IPO) model perspective (37). Each step in the DPT notification cascade requires resources (e.g., technical infrastructure, money) and inputs (e.g., information), which are processed to create outputs (e.g., notifications). The different IPO steps can be examined from different viewpoints shown in the vertical classification matrix dimension, namely from a technical (app-) perspective, as well as from the viewpoint of different actors, including app end-users, but also laboratories or public health services (**Figure 1**).

Therefore, each matrix cell reflects a combination of IPO-step and viewpoint that can be useful to describe and evaluate the performance of specific steps or elements of DPT systems (KPIs). By taking a specific step in the notification cascade described in section *A Closer Look at DPT Steps and Their Influence on Intervention Outcomes*, the performance can be evaluated from different angles using the guiding questions such as: What are the resources needed to complete this step? How well does the information flow along the notification cascade work? Or how much desired output is generated by this step? Such KPIs can be formalized as raw numbers, proportions, or ratios to describe *speed, yield, and capacity* attributes (as described in section *Basic Concepts for DPT Evaluations*).

Separate and located below the classification matrix in **Figure 3** are the public health *effectiveness* measures. They are distinct from KPIs and aim to address a different question: does the DPT system achieve its intended primary aim of notifying exposed app users swiftly so they can take preventive measures? Measures of the DPT effectiveness can relate, for example, to the prevention of further transmission or comparative cost-effectiveness when compared to MCT (6). The health-centered effectiveness measures are different from KPIs, and yet not independent. The completion of the notification process is a precondition for achieving the DPT public health goal. In other words, many public health metrics are an integral of different processes in the app notification cascade, as they are a direct consequence of how effective the notification cascade tasks are executed.

A Worked Example of the KPIs Classification Matrix and Public Health Measures: The Proximity Notification Step

To further illustrate the use and usability of the KPIs classification matrix, we will—cell by cell—describe KPIs examples related to the proximity notification step (which could also be applied to classic MCT). The proximity notification step (section *Notification of Proximity Contacts*) is a crucial element in the notification cascade with a direct relation to the primary DPT goal: to warn proximity contacts as early as possible about potential transmission risk exposures.

The first matrix column represents resource-oriented metrics, which define resource needs for technical and non-technical implementation and include, for example, costs for PCR-tests (which are free for persons with a DPT notification in some countries), costs for quarantining of DPT-notified persons, or any other expenditures.

In the vertical dimension, the first matrix row reflects the app-oriented, technical perspective. During DPT development and operation, IT system design requires choices for parametrization of measurements and backend systems, which are resource-dependent and may impact speed, yield, or capacity of the DPT processes. Therefore, concrete KPIs examples for resource-oriented metrics from the technical perspective (*cell 1*) include costs for development and maintenance of the app itself, as

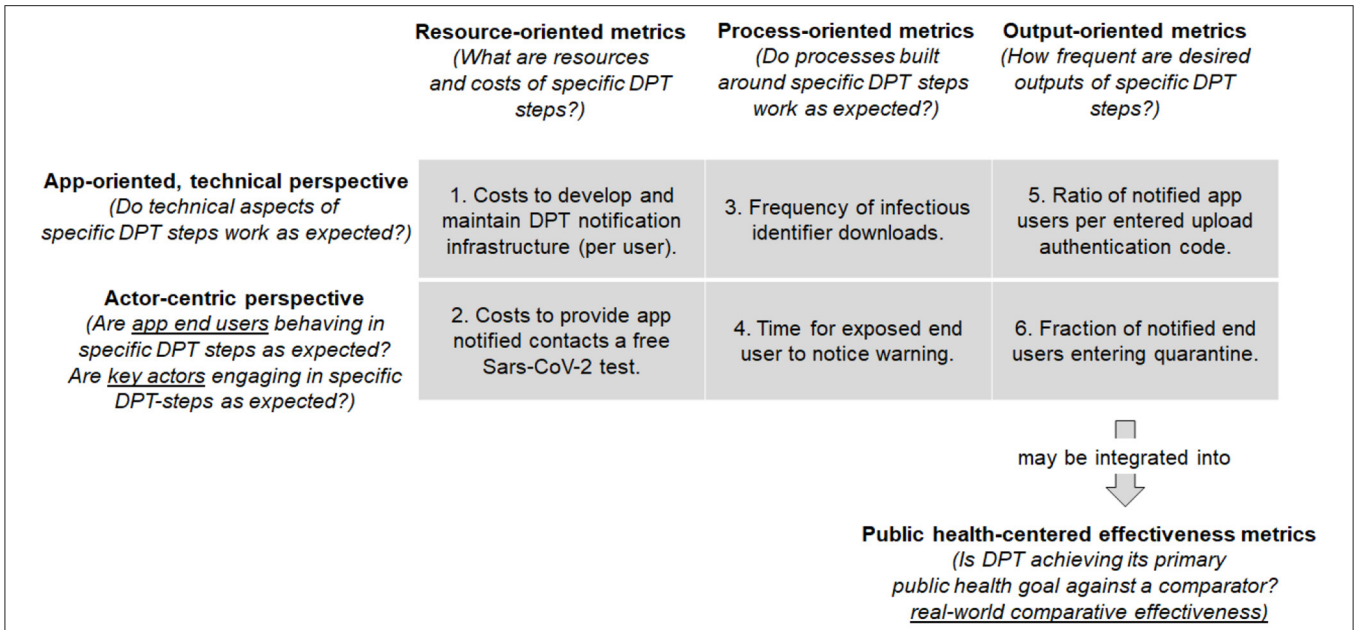


FIGURE 3 | Proposal for classification matrix of different indicator types to monitor DPT systems by steps of the Input-Process-Output model (resource-, process-, and output metrics) as well as by aspect (app-oriented, technical viewpoint vs. actor-centric perspectives). The examples in the matrix pertain to the notification step of proximity contacts (section *Notification of Proximity Contacts*). The output oriented metrics (5 & 6) may be integrated into public health-centered metrics.

well as for the technical infrastructure and the backend. Such expenditures are often scaled by the number of active users or the number of quarantine orders (or other indicators for prevented transmissions).

The second matrix row reflects the actor-centric perspective on the notification chain. Involved actors include laboratories that perform PCR-tests and communicate results to app end-users and health authorities, health authorities and other authorized parties that take calls from notified users, and—in settings with manual upload authorization—the parties who provide upload authorization codes to diagnosed users. Furthermore, among all involved actors, the app end-users play a central role. End-users need to decide whether to use the DPT app, but also to actively trigger (or at least consent to) the upload of identifiers in case of testing positive for Sars-CoV-2. Example KPIs that combine resource and actor perspectives (*cell 2*) include expenditures for user-linked actions, such as the costs for providing app-notified contacts a free Sars-CoV-2 test.

Process metrics are located in the second matrix column. Those KPIs describe interactions of the app and its users with other parts and actors of the health system. For the app to work as intended, several processes need to occur seamlessly so that all tasks can be carried out successfully and timely: from testing to prompt results communication, upload authorization code generation, and identifier upload, notification of exposed contacts, and these contacts taking action (e.g., calling the hotline or a doctor and receiving advice). Process metrics can be used to monitor how well the different conditions for app-functioning are met, respectively, whether the different system parts work as expected.

Examples of process-centric metrics that integrate the technical perspective (*cell 3*) include, for example, precision and recall of Bluetooth and exposure time measurements, which are usually assessed in experimental settings. Specific design choice evaluations may involve measurements of how well the GAEN/Bluetooth approximation reflects actual physical distance and time exposure, as well as backend configurations regarding the frequency of infectious key uploads or downloads of lists of infectious identifiers (which only happens a certain number of times per day). *Cell 4* represents process-oriented metrics from an actor perspective. Several steps in the notification cascade require human involvement, sometimes on a voluntary basis. Therefore, such actions may be strongly influenced by behavioral aspects, digital and health literacy, but also by incentives. KPIs examples include the fraction of positive app users who consent to or actively initiate the identifier uploads. KPIs used to describe such steps can often be based on yield (fraction of completed tasks) or speed (time to completion) attributes.

Finally, the third matrix column reflects output-oriented metrics. These KPI refer to desired outputs of DPT, which could be numbers or yields of DPT-notified users who undertake a recommended action (e.g., entering quarantine or getting tested for Sars-CoV-2). These metrics are related to public health goals but differ in that they focus on an intra-system perspective: they often do not encompass external comparators but focus on how a system has evolved. Technical aspects influence desired outcomes in various ways. In *cell 5*, notifications of exposed contacts are the desired outcome. An example KPI is the ratio of the number of exposure notifications over the number of upload authentication codes entered by positive tested users. The

ratio depends on various technical aspects, including calculation methods for proximity risk scores derived from Bluetooth attenuation measurements.

Furthermore, many outputs depend on end-user interactions. App-notified contacts are expected to follow certain procedures such as calling an infoline, getting tested, or entering quarantine (*cell 6*). Examples for such user-dependent outcome KPIs are the fraction app-notified users who voluntarily enter quarantine or who seek testing. It is often instructive to express these KPIs both as yield (fraction who completes an action) and speed attributes (time until an action is completed).

Finally, the public-health-oriented metrics reflect the real-world effectiveness of DPT apps as defined above and relate to the main pandemic mitigation goals. These DPT goals are the result of the interplay between the different technical and non-technical aspects of the notification cascade. That is, the KPIs provide information about the performance of the notification cascade required for DPT effectiveness. However, KPIs do not provide direct evidence for DPT effectiveness, for which a comparator group is required (which, however, can also be a comparator without any measures such as DPT or MCT as described in section *Basic Concepts for DPT Evaluations*). In the context of the proximity contact notification step, an effectiveness evaluation could, for example, compare the time from positive testing of the index case to quarantine of the proximity contact between DPT and standard MCT.

As outlined in the previous section, key performance indicators may contribute to effectiveness evaluations if (a) they reflect actions or outcomes that are of public health relevance and (b) if they are compared against a suitable "comparator group." Condition (b) is also the reason why the indicator in cell 6 is not (yet) an effectiveness measure. Monitoring this indicator in a time-series will be informative about longitudinal changes. However, an effectiveness analysis would require to compare the indicator 6 against a reference group, such as the fraction of persons entering voluntary quarantine with or without exposure notification. The latter can occur if exposed contacts receive an informal exposure notification, e.g., a message from a friend or a relative who tested positive for SARS-CoV-2.

KEY CONSIDERATIONS FOR THE PRACTICAL IMPLEMENTATION OF PERFORMANCE AND EFFECTIVENESS MEASURES

Definition of Expected Process Targets

To monitor the process performance of DPT, it is helpful to have an expectation of where indicators should stand at a given time. That is, to assess performance on the basis of absolute numbers (e.g., number of authorized uploads) or a yield (e.g., fraction of authorized over realized uploads), expectations or precise benchmark targets should be defined. Because DPT is still a very novel health technology, defining specific benchmarks can be challenging. Moreover, targets may not only be country—but also setting—or subgroup-specific (e.g., targeting a specific app coverage in the working population).

Target definitions can in part be informed by modeling results (e.g., concerning required DPT coverage to create an effect) (1, 38–40). But in many instances, only qualitative targets may be feasible because of a lack of suitable reference values. A possible approach to derive such qualitative targets is to describe "desired" effects in a hypothetical, perfectly functioning system. For example, KPIs to measure the full completion of actions, such as the fraction of positive users that upload their information to the server (e.g., by measuring the fraction of upload authorization codes that are redeemed). Other KPI measuring attributes such as speed may lack a clearly defined benchmark (for example, the time from app notification to quarantine), but could be formulated in terms of a comparison to other measures (for example, the same intervals in a manual process).

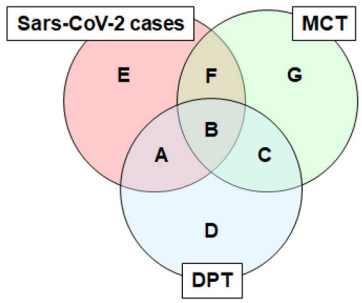
Practical Implications of the Distinction Between Performance and Effectiveness Measures

The distinction between KPIs and comparative effectiveness measures is more than just semantics. KPIs and effectiveness measures require different data and measurement approaches. Process metrics can be collected at different contact points in the notification cascade (**Figure 1**), for example at app download, during regular configuration updates, when upload authorization codes are generated, or when notified users call an infoline. But due to the privacy-preserving, decentralized nature of DPT apps (at least those that follow the DP-3T blueprint) these metrics provide only aggregated, non-identifiable data, and these data points cannot be easily connected into a unique data stream for a specific user.

By contrast, comparative effectiveness investigations need to establish a link to processes and data collections of its comparator, which will be MCT in most settings. But from a privacy perspective, it should be clear that investigations providing unquestionable evidence for DPT effectiveness can no longer be privacy-preserving and anonymous. For example, an ideal study of DPT/MCT effectiveness should be able to connect infected cases with exposed contacts and follow their notification and quarantining cascade in great detail. Often proposed key metric is the secondary attack rate (SAR), which measures how many exposed contacts of an infected person later test positive for Sars-CoV-2 (11). However, calculating precise SAR measures require an exact identification and linkage of cases and contacts, something that is not foreseen in privacy-preserving DPT apps. At the same time, DPT can also mitigate some shortcomings of MCT, for instance by improving notification speed and extending to exposed contacts who are unknown to the index case.

Choosing the Right Denominator

Selecting appropriate denominators for KPIs and public health metrics can pose challenges. While there is no universal best practice, we find that Venn diagrams can be a helpful tool to guide the search for suitable denominators (**Figure 4**). Venn diagrams are useful to illustrate the (non-)overlap between different populations of interest. In the context of DPT, these are the persons who are tested positive for Sars-Cov-2 (cases), those who



Example interpretation

- All persons with a positive Sars-CoV-2 test (cases): Segments A + B + F + E
- All persons who are contacted by manual contact tracing. Infected cases: Segments B + F Exposed contacts: Segments C + G
- All persons who received a DPT app notification: Infected cases: Segments A + B Exposed contacts: Segments D + C

The overlapping segments A, B, C, F indicate groups that belong to more than one group.

The interpretation of segments can be adjusted as needed, but the assumptions of the Venn diagram need to be clarified along the criteria listed in the main text.

FIGURE 4 | Example use of Venn diagrams to define denominators for KPIs and public health metrics (courtesy of https://pact.mit.edu/).

are notified by DPT, and those who are informed through manual contact tracing (MCT). DPT represents the whole population of notified app users. MCT includes all persons who were identified through manual contact tracing. The segments labeled from A-G represent population counts (e.g., corresponding to outcome-oriented metrics in **Figure 3**).

While Venn diagrams may facilitate conceptual discussions, the assumptions and context must be well-described. These assumptions include the following aspects.

- Goal: What type of high-level task or metric should be described by the Venn diagram?
- Population: What is the origin of the data for KPIs or public health metric analysis (case series, cohort study, population-based study, or administrative database)?
- Time horizon: What is the time perspective covered by the Venn diagram (cross-sectional, or cumulative over a longer period)?
- Evaluation time point: At what time point are classifications into the three groups (positive tested cases, DPT, MCT) established? Shortly after the time of exposure, when PCR test results are still pending? Cross-sectional at a given moment in time?
- Case definitions: Furthermore, what is the accuracy with which infection status can be determined (i.e. how to deal with infected, untested individuals)?
- Setting-specific assumptions: Finally, country-specific simplifications may be warranted based on the Test-Trace-Isolate-Quarantine strategy (e.g., whether all PCR-positive cases are automatically referred to MCT). Therefore, for some countries, one can assume that segments A + E are close to 0, whereas F + B are approximating the number of positive cases.

Feasibility of Integrating Measurements Directly Into DPT

Given the different data requirements for KPI monitoring and effectiveness studies, the question arises how the necessary information should be collected: by integration into DPT apps and corresponding backend systems or through separate research studies?

The addition of measurement capabilities to DPT apps can be a sensitive matter. First, DPT apps following the DP-3T blueprint are not designed as data collection instruments, but as privacy-preserving notification tools that keep their users anonymous. Adding more measurement capabilities (e.g., in the backend or the app itself) leads to a data granularity-privacy trade-off. The gain in knowledge has to be weighed against a greater likelihood for de-anonymization. Adding measurement capabilities may require an increased trust by end-users in the system operators. For example, collecting exact dates of exposures, notifications, and contacts with different actors (e.g., the infoline) may, in combination, imply that study subjects may no longer be sure that their identities remain concealed. The combination of these measures may already identify persons uniquely, especially in smaller populations. If the collection of such data is to take place, it must happen transparently and app users should provide informed consent. In addition, other privacy-preserving technologies that minimize the amount of data collected and limit the capability of linkage across databases can be employed[1].

The decision of whether and how to integrate additional measurements into DPT apps (beyond what is needed for notification) is one that each country needs to make separately. Such a decision must take into account specific legal considerations, overall acceptance of the DPT technology, and public expectations toward DPT privacy, as well as the individual and societal risks associated with the new data collection.

As an alternative to DPT-integrated measurements, dedicated (observational) research studies with volunteers and specifically designed databases should be considered. Given informed consent by participants, a linkage of information between DPT and MCT should be possible. For example, studies could be

[1]Examples of privacy-preserving building blocks that could be used to support measurements while minimizing risks for users are multi-party computation, differential privacy, anonymous authentication, and homomorphic encryption. Other privacy technologies could also be of interest.

integrated into contact tracing and specifically survey index cases and exposed contacts regarding app usages and notifications. Ireland, for example, employs separate case management systems, which collect numerous complementary data that could be valuable for effectiveness research. Alternatively, app users could be presented with short questionnaires including questions regarding usage and past exposures. However, a linkage of apps with survey advertisements (even on a voluntary basis) could be regarded as intrusive and fuel privacy fears. Therefore, the advantages and disadvantages of each survey recruitment method should be deliberated carefully.

CONCLUSION

The development of monitoring systems for DPT performance and effectiveness requires complex decisions. While there is no universal advice that could suit all settings and countries, it may help to obtain clarity on the distinctions between performance monitoring and effectiveness. Furthermore, decision makers should become aware that *not all measurements can and should be integrated into DPT apps* and connected backend systems. Separate studies or data collection systems may be needed to generate the necessary evidence for performance and effectiveness of DPT. The proposed indicator classification aims to support this process.

AUTHOR CONTRIBUTIONS

VW conceived the study and wrote the first draft. All authors provided conceptual inputs, analyzed the data, contributed to paper writing and revision, and approved the manuscript.

ACKNOWLEDGMENTS

The authors thank the regular participants of the DP-3T international exchange group for helpful discussions. VW thanks the MIT PACT group members https://pact.mit.edu/ for sharing their insights and for providing the inspiration to use Venn diagrams. We also thank Mina Stanikić for language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.677929/full#supplementary-material

REFERENCES

1. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*. (2020). 368:6491. doi: 10.1126/science.abb6936
2. Swiss National Covid-19 Science Task Force. *Digital Proximity Tracing. Policy Briefs*. (2020). Available online at: https://ncs-tf.ch/de/policy-briefs/digital-proximity-tracing-15-may-20-en/download (accessed July 19, 2021).
3. Dar AB, Lone AH, Zahoor S, Khan AA, Naaz R. Applicability of mobile contact tracing in fighting pandemic (COVID-19): Issues, challenges and solutions. *Comput Sci Rev*. (2020) 38:100307. doi: 10.1016/j.cosrev.2020.100307
4. Servick K. COVID-19 contact tracing apps are coming to a phone near you. How will we know whether they work? *Science*. (2020). doi: 10.1126/science.abc9379
5. Zastrow M. Coronavirus contact-tracing apps: can they slow the spread of COVID-19? *Nature*. (2020). doi: 10.1038/d41586-020-01514-2
6. von Wyl V, Bonhoeffer S, Bugnion E, Puhan MA, Salathé M, Stadler T, et al. A research agenda for digital proximity tracing apps. *Swiss Med Wkly*. (2020) 150:w20324. doi: 10.4414/smw.2020.20324
7. Menges D, Aschmann HE, Moser A, Althaus CL, von Wyl V. A data-driven simulation of the exposure notification cascade for digital contact tracing of SARS-CoV-2 in Zurich, Switzerland. *JAMA Netw Open*. (2021) 4:e218184. doi: 10.1001/jamanetworkopen.2021.8184
8. Wymant C, Ferretti L, Tsallis D, Charalambides M, Abeler-Dörner L, Bonsall D, et al. The epidemiological impact of the NHS COVID-19 app. *Nature*. (2021) 594:408–12. doi: 10.1038/s41586-021-03606-z
9. Salathé M, Althaus CL, Anderegg N, Antonioli D, Ballouz T, Bugnion E, et al. Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. *Swiss Med Wkly*. (2020) 150:w20457. doi: 10.4414/smw.2020.20457
10. World Health Organization & European Centre for Disease Prevention and Control. *Indicator Framework for the Evaluation of the Public Health Effectiveness of Digital Proximity Tracing Solutions*. World Health Organization (2021). Available online at: https://apps.who.int/iris/handle/10665/341818 (accessed July 19, 2021).
11. European Centre for Disease Prevention and Control. *Mobile Applications in Support of Contact Tracing for COVID-19 - A Guidance for EU EEA Member States*. (2020). Available online at: https://www.ecdc.europa.eu/en/publications-data/covid-19-mobile-applications-support-contact-tracing#no-link (accessed July 19, 2021).
12. Colizza V, Grill E, Mikolajczyk R, Cattuto C, Kucharski A, Riley S, et al. Time to evaluate COVID-19 contact-tracing apps. *Nat Med*. (2021) 27:361–2. doi: 10.1038/s41591-021-01236-6
13. Huang Z, Guo H, Lee Y-M, Ho EC, Ang H, Chow A. Performance of digital contact tracing tools for COVID-19 response in Singapore: cross-sectional study. *JMIR Mhealth Uhealth*. (2020) 8:e23148. doi: 10.2196/23148
14. Sattler F, Ma J, Wagner P, Neumann D, Wenzel M, Schäfer R, et al. Risk estimation of SARS-CoV-2 transmission from bluetooth low energy measurements. *NPJ Digit Med*. (2020) 3:129. doi: 10.1038/s41746-020-00340-0
15. Leith DJ, Farrell S. Measurement-based evaluation of Google/Apple Exposure Notification API for proximity detection in a light-rail tram. *PLoS ONE*. (2020) 15:e0239943. doi: 10.1371/journal.pone.0239943
16. Rodríguez P, Graña S, Alvarez-León EE, Battaglini M, Darias FJ, Hernán MA, et al. A population-based controlled experiment assessing the epidemiological impact of digital contact tracing. *Nat Commun*. (2021) 12:587. doi: 10.1038/s41467-020-20817-6
17. Troncoso C, Payer M, Hubaux J-P, Salathé M, Larus J, Bugnion E, et al. *Decentralized Privacy-Preserving Proximity Tracing*. (2020). Available online at: https://arxiv.org/abs/2005.12273 (accessed July 19, 2021).
18. TousAntiCovid. Available online at: https://bonjour.tousanticovid.gouv.fr/ (accessed April 20, 2021).
19. Wikipedia contributors. *COVID-19 Apps Wikipedia, The Free Encyclopedia*. (2020). Available online at: https://en.wikipedia.org/w/index.php?title=COVID-19_apps&oldid=994533835 (accessed December 22, 2020).
20. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Drivers of acceptance of COVID-19 proximity tracing apps in switzerland: panel survey analysis. *JMIR Public Health Surveill*. (2021) 7:e25701. doi: 10.2196/25701
21. Bonardi J-P, Brülhart M, Danthine J-P, Saxena A, Thöni C, Thoenig M, et al. How to make digital proximity tracing work: the view from

economics. (2020). Available online at: https://e4s.center/document/how-to-make-digital-proximity-tracing-work-the-view-from-economics/ (accessed July 19, 2021).
22. Kretzschmar ME, Rozhnova G, Bootsma MCJ, van Boven M, van de Wijgert J, Bonten MJM. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *Lancet Public Health*. (2020) 5:e452–9. doi: 10.1016/S2468-2667(20)30157-2
23. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med*. (2010) 8:63. doi: 10.1186/1741-7015-8-63
24. Ross J, Stevenson F, Dack C, Pal K, May C, Michie S, et al. Developing an implementation strategy for a digital health intervention: an example in routine healthcare. *BMC Health Serv Res*. (2018) 18:794. doi: 10.1186/s12913-018-3615-7
25. von Wyl V. Challenges for nontechnical implementation of digital proximity tracing during the COVID-19 pandemic: media analysis of the SwissCovid App. *JMIR Mhealth Uhealth*. (2021) 9:e25345. doi: 10.2196/25345
26. Apple/Google. *Privacy-Preserving Contact Tracing*. (2020). Available online at: https://www.apple.com/covid19/contacttracing (accessed July 19, 2021).
27. bt-measurements. Github. (2020). Available online at: https://github.com/DP-3T/bt-measurements (accessed December 22, 2020).
28. PT-System-Documents. Github. (2020). Available online at: https://github.com/admin-ch/PT-System-Documents (accessed December 22, 2020).
29. cwa-documentation. Github. (2020). Available online at: https://github.com/corona-warn-app/cwa-documentation (accessed December 22, 2020).
30. gaen-risk-scoring. Github. (2020). Available online at: https://github.com/lfph/gaen-risk-scoring (accessed December 23, 2020).
31. Sciensano Belgian institute for health. *Coronalert be Frequently Asked Questions: How do I Report an Infection?* Brussels: Sciensano (2020).
32. Hoerdt J. *What Requirements Must Be Met in Order for a Personal Test Result to Be Successfully Transmitted via the Corona-Warn-App?*. Corona-Warn-App open-source project (2020). Available online at: https://www.coronawarn.app/en/blog/2020-09-11-positive-test-result/ (accessed December 18, 2020).
33. Swiss Federal Office of Statistics. *SwissCovid App Monitoring*. (2020). Available online at: https://www.experimental.bfs.admin.ch/expstat/de/home/innovative-methoden/swisscovid-app-monitoring.html (accessed July 19, 2021).
34. RKI - Coronavirus SARS-CoV-2 - *Übersicht zu aktuellen und früherer Zahlen und Fakten zur Corona-Warn-App*. (2020). Available online at: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/WarnApp/Archiv_Kennzahlen/WarnApp_KennzahlenTab.html (accessed December 23, 2020).
35. Onderzoek: hoe weten we of CoronaMelder helpt tegen Corona? (2020). Available online at: https://www.coronamelder.nl/nl/faq/40-onderzoek-hoe-weten-we-of-coronamelder-helpt-tegen-corona/ (accessed December 2 3, 2020).
36. Definition of EFFECTIVENESS. Available online at: https://www.merriam-webster.com/dictionary/effectiveness (accessed December 2 2, 2020).
37. Wikipedia contributors. IPO model. Wikipedia, The Free Encyclopedia. (2020). Available online at: https://en.wikipedia.org/w/index.php?title=IPO_model&oldid=974000798 (accessed December 22, 2020).
38. Abueg M, Hinch R, Wu N. Liu L, Probert WJM, Wu A, et al. Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state. *npj Digit Med*. (2021) 4:49. doi: 10.1038/s41746-021-00422-7
39. Cencetti G, Santin G, Longa A, Pigani E, Barrat A, Cattuto C, et al. Digital proximity tracing on empirical contact networks for pandemic control. *Nat Commun*. (2021) 12:1655. doi: 10.1038/s41467-021-21809-w
40. López JAM, Arregui-Garcla B, Bentkowski P, Bioglio L, Pinotti F, Boëlle PY, et al. Anatomy of digital contact tracing: Role of age, transmission setting, adoption, and case detection. *Sci Adv*. (2021) 7:eabd8750. doi: 10.1126/sciadv.abd8750

Conflict of Interest: DB is an employee of Cybernetica AS.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

REVIEW
published: 06 August 2021
doi: 10.3389/fdgth.2021.707902

Data and Digital Solutions to Support Surveillance Strategies in the Context of the COVID-19 Pandemic

Patty Kostkova[1], *Francesc Saigí-Rubió*[2,3], *Hans Eguia*[2,4], *Damian Borbolla*[5], *Marieke Verschuuren*[6], *Clayton Hamilton*[6], *Natasha Azzopardi-Muscat*[6] *and David Novillo-Ortiz*[6]*

[1] UCL Centre for Digital Public Health in Emergencies (dPHE), Institute for Risk and Disaster Reduction, University College London, London, United Kingdom, [2] Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain, [3] Interdisciplinary Research Group on ICTs, Barcelona, Spain, [4] SEMERGEN New Technologies Working Group, Madrid, Spain, [5] Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States, [6] Division of Country Health Policies and Systems, Regional Office for Europe, World Health Organization, Copenhagen, Denmark

Background: In order to prevent spread and improve control of infectious diseases, public health experts need to closely monitor human and animal populations. Infectious disease surveillance is an established, routine data collection process essential for early warning, rapid response, and disease control. The quantity of data potentially useful for early warning and surveillance has increased exponentially due to social media and other big data streams. Digital epidemiology is a novel discipline that includes harvesting, analysing, and interpreting data that were not initially collected for healthcare needs to enhance traditional surveillance. During the current COVID-19 pandemic, the importance of digital epidemiology complementing traditional public health approaches has been highlighted.

Objective: The aim of this paper is to provide a comprehensive overview for the application of data and digital solutions to support surveillance strategies and draw implications for surveillance in the context of the COVID-19 pandemic and beyond.

Methods: A search was conducted in PubMed databases. Articles published between January 2005 and May 2020 on the use of digital solutions to support surveillance strategies in pandemic settings and health emergencies were evaluated.

Results: In this paper, we provide a comprehensive overview of digital epidemiology, available data sources, and components of 21st-century digital surveillance, early warning and response, outbreak management and control, and digital interventions.

Conclusions: Our main purpose was to highlight the plausible use of new surveillance strategies, with implications for the COVID-19 pandemic strategies and then to identify opportunities and challenges for the successful development and implementation of digital solutions during non-emergency times of routine surveillance, with readiness for early-warning and response for future pandemics. The enhancement of traditional surveillance systems with novel digital surveillance methods opens a direction for the most effective framework for preparedness and response to future pandemics.

Keywords: digital surveillance, digital epidemiology, data sources, outbreak, COVID-19

OPEN ACCESS

Edited by:
Pradeep Nair, Central University of Himachal Pradesh, India

Reviewed by:
Manh-Toan Ho, Phenikaa University, Vietnam
Mohamed Elhakim, World Health Organization, Djibouti

***Correspondence:**
David Novillo-Ortiz
dnovillo@who.int

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *10 May 2021*
Accepted: *30 June 2021*
Published: *06 August 2021*

Citation:
Kostkova P, Saigí-Rubió F, Eguia H, Borbolla D, Verschuuren M, Hamilton C, Azzopardi-Muscat N and Novillo-Ortiz D (2021) Data and Digital Solutions to Support Surveillance Strategies in the Context of the COVID-19 Pandemic. Front. Digit. Health 3:707902. doi: 10.3389/fdgth.2021.707902

INTRODUCTION

In London's Soho district in 1854, the father of public health John Snow removed the handle of the local community water pump to stop the spread of the famous cholera outbreak (1). Snow proved cholera is a water-borne disease rather than an airborne "bad air" infection by manually mapping data of citizens infected and deaths onto a map (2). This began what we today call epidemiology and surveillance.

Public health surveillance is a regular routine process of collecting data on diseases, cases and public health interventions (such as vaccination) to inform public health authorities about the situation in order to respond with appropriate public health measures. Public health surveillance also includes early warning systems alerting about upcoming outbreaks and emergencies. In order to enable rapid response, inform public health policy and strategies (3). The recent increase of big data and digital and mobile technology has enabled the rapid growth of "digital epidemiology."

Digital epidemiology provides numerous opportunities and challenges and now is an indispensable part of infectious disease surveillance systems. Epidemic intelligence is understood as the systematic collection and analysis of traditional (indicator-based epidemiological) and new data sources (event-based surveillance), which are used to identify new infection threats to provide early warning and rapid assessment of risk (4). Once a threat or an outbreak is detected and an event verified and risk assessed, a rapid response must be implemented to control the outbreak, including diagnosis, testing, contact tracing, and risk communication with the public.

These novel data sources arise from new digital solutions such as tracking devices, mobile applications (apps), and social media interventions; they can also contribute to infectious disease outbreak management. Digital epidemiology uses these devices and the data gathered to find new solutions to minimise disease spread as well as determine the population's behaviour and insights, e.g., behavioural reactions to public health interventions, contact tracing, and others (5). However, considerable computational and technical challenges arise from the rapid increase in relevant data from digital data sources: "Extracting meaningful information from this data deluge is challenging, but holds the unparalleled potential for epidemiology" (6). The general term for analysing and disseminating real-time health information from news and social media is also referred to as "infodemiology" (7, 8).

With the amount of data that can be collected through digital epidemiology, making sense of the data and determining whether it will adequately support epidemiological surveillance can pose difficulty. Applications and digital tools are being developed to process large volumes of unstructured data (big data) to help uncover useful information for problem solving. The term "big data" refers to the use and analysis of verified information that has been collected. This includes complex data that is rapidly collected in massive amounts, as long as the data is real and verifiable (9, 10). Digital sources of big data in healthcare include electronic medical records, genomics, imaging data, data from social networks, and sensor data (11).

Big data can be extracted from diverse real-time or static information sources that are often underutilised or not accessible, which could potentially increase the acquisition of new knowledge contributing to a better understanding of disease epidemiology. Algorithmic analysis to "train" data for classification or predictions for decision-making is a rapidly growing computer science domain called machine learning.

Digital epidemiological surveillance involves sources that are not typically used in traditional epidemiology, generating larger amounts of information that should be incorporated into public health systems as part of the response to traditional diseases, new emerging pathogens such as the COVID-19 virus we are currently fighting.

The main aim of this paper is to provide a comprehensive overview for the literature digital solutions and big data to support surveillance strategies in the context of the COVID-19 pandemic and beyond.

METHODS

This research is a review of original research on digital surveillance from January 2015 until May 2020 with implications for opportunities and strategies for new outbreaks such as the COVID-19 pandemics and future emergencies.

A review of literature going back 15 years was conducted as 2005 was chosen as the year of the dawn of wide spread mobile technology and big data. We focused on syntesis of aproaches in order to draw implications for surveillance opportunities for COVID-19 and beyond (we are aware this study is not presenting a review of COVID-19 strategies as it is too soon to conduct such an excercise).

The search was conducted in the electronic database MEDLINE (accessed by PubMed) for articles published between January 2005 and May 2020 using combinations of the following free terms and Boolean operators (AND and OR):

"COVID-19"[Title/Abstract] OR "COVID-19 diagnostic testing"[Supplementary Concept] OR "surveillance"[Title/Abstract] OR "Pandemics"[MeSH Terms] OR "epidemic control"[Title/Abstract] OR "self-diagnosis"[Title/Abstract] OR "self-evaluation"[Title/Abstract] OR "contact tracing"[Title/Abstract] AND ("digital health"[Title/Abstract] OR "information system"[Title/Abstract] OR "apps"[Title/Abstract] OR "eHealth"[Title/Abstract] OR "e-Health"[Title/Abstract] OR "electronic health record*"[Title/Abstract] OR "big data"[Title/Abstract] OR "machine learning"[Title/Abstract] OR "data science"[Title/Abstract] OR "artificial intelligence"[Title/Abstract] OR "mHealth"[Title/Abstract] OR "m-Health"[Title/Abstract] OR "social media"[Title/Abstract] OR "IoT"[Title/Abstract] OR "smartphone"[Title/Abstract] OR "Internet of things"[Title/Abstract]).*

The search was limited to English-, Portuguese-, and Spanish-language publications and was complemented using the snowballing technique to identify relevant articles in the reference lists of articles returned by our search (12). Additional search for grey literature was conducted regarding digital surveillance. Expanded grey literature searching included

internet search engine, targeted websites and social media. The search is subject to a selection bias as publications were limited to the three major languages, however, as the majority of scientific literature is published in the three major languages this bios is minimal.

Diverse studies covering the use of digital data sources for surveillance during a health emergency were included. Initial screening was based on titles and abstracts, and articles were independently evaluated. Abstracts lacking sufficient information to identify their inclusion or exclusion, were retrieved for full-text evaluation. Subsequently, two investigators independently evaluated the full-text articles and determined eligibility for inclusion or exclusion. Authorship, journal, or years were not blinded.

Study Selection

The initial research included complete publications and abstracts that were reviewed to determine whether they met the inclusion or exclusion criteria. Abstracts lacking information were retrieved for full-text evaluation. The inclusion criteria were (1) original research articles, (2) studies conducted during outbreaks or pandemic situations that measured the use of digital tools for contact tracing, (3) studies on the application of data and digital solutions to support surveillance strategies, and (4) studies covering the use of digital data sources for surveillance during a health emergency. The exclusion criteria were studies that described the use of technology outside an epidemic, big data studies that were not focused on epidemiological problems, and other surveillance interventions that were not related to the use of digital health solutions.

After the first review of the titles and abstracts, 280 studies were selected. For the grey literature, 11 electronic notes were reviewed. After reading the full texts of these studies, 130 were deemed to have met the search selection criteria, one of them was grey literature. Two authors (DB and FSR) screened all articles individually. Discrepancies were resolved through discussion with a third author (DN) when necessary. All the data were analysed qualitatively and quantitatively. The search and selection processes are summarised in **Figure 1**.

Based on the surveillance and early-warning & response processes established at WHO, ECDC and member states, four core components of digital epidemiology were identified in previous research and expanded[1]:

- digital surveillance, supporting public health routine surveillance;
- early warning and epidemic intelligence, striving to improve early warning tools that alert public health professionals to upcoming threats. After a threat has been verified, big data analysis might support public health experts in:
- rapid response, outbreak control, and increasing the impact of public health measures through digital interventions; and
- risk communication and public communication being advanced through mobile apps and social media, while the improvement in epidemic modelling that leverages real-time heterogeneous data improves public health policy through better assessment of how control measures impact healthcare and society.

While our systematic review illustrates how digital technology has been contributing to these four subdisciplines for over a decade. In **Table 1**, we map them to the objectives of COVID-19 public health surveillance defined by World Health Organization (WHO) (13)—requiring a predominantly public health and healthcare sector response—complemented by opportunities provided by digital epidemiology leveraging novel big data streams.

While reviewing the articles using an iterative process (with regular meetings), the digital data and innovation systems were described in line with the four digital epidemiology categories.

RESULTS

Our results provide an overview of the literature identified throughout our search process and highlight the existing opportunities and applications for the found disciplines in digital epidemiology.

Digital Surveillance

Surveillance is a core component of public health and preventive medicine and can be categorised as "active" (where health authorities make direct contact with the population or care providers to measure the actual conditions) or "passive" (when the health authorities get pre-designed reports about specific conditions, typically with the care providers reporting) (14).

With the use of new technology—for example social networks (15)—the traditional surveillance process can be not only enhanced but also reach a wider population (16). Passive surveillance, drawing from a wide range of novel data sources, is less accurate and subject to elevated noise in the data, either by proactively searching related information and or by gathering it passively, making it less resource intensive (17).

Before an infectious disease is confirmed by a laboratory, infected persons may exhibit symptoms, signs, behavioural patterns, or laboratory findings that can be tracked through a variety of data sources (18). This relates to what is called "syndromic surveillance," which may be boosted by the use of digital technologies.

Integrated or enhanced surveillance combines both systems (19): traditional and digital. The strength of syndromic surveillance systems is in early warning systems for emerging disease threats, such as dengue (20), while the strength of indicator-based (active) surveillance is the production of regular, robust, reliable surveillance reports (21). Syndromic surveillance threats could also capture undiagnosed infections (22, 23). A good example is metagenomics, as infected (and infectious) people who are asymptomatic may unwittingly spread the infection to others (24).

The use of online, mobile, and social media data streams for routine surveillance enhances traditional methods (3, 25). Also called "participatory surveillance," (26)—disease and symptoms reported directly by citizens themselves rather than by health

[1]https://dl.acm.org/doi/10.1109/WI-IAT.2011.311

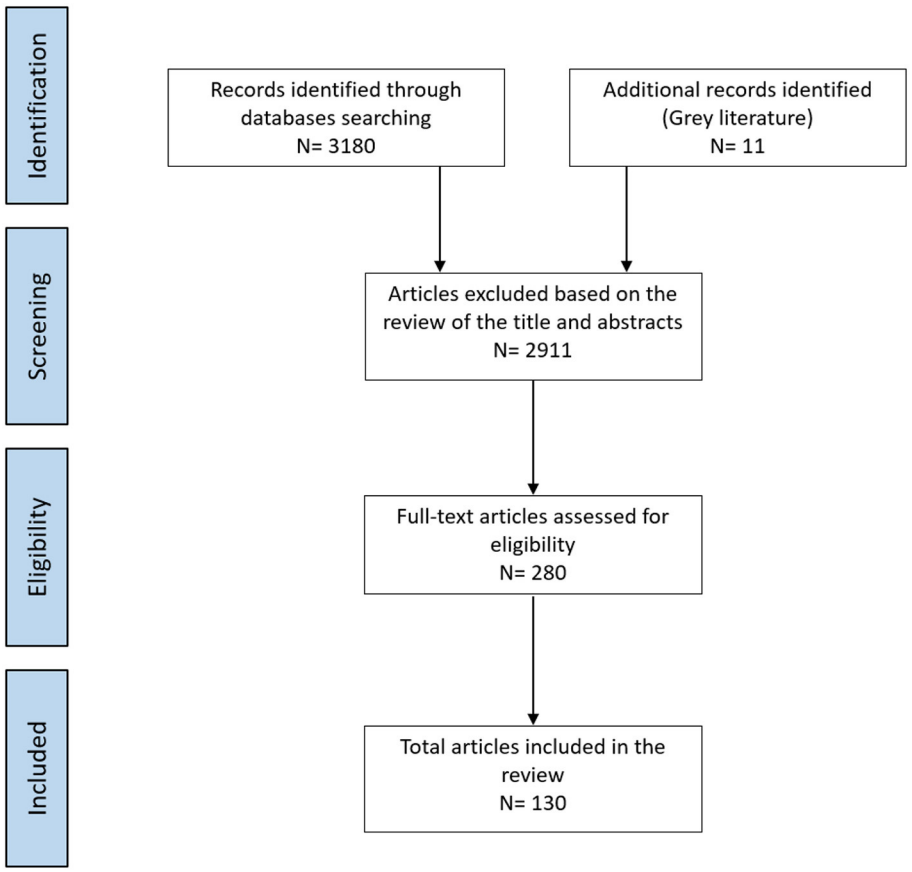


FIGURE 1 | Flow diagram of studies through the review.

authorities—the new generation of health trackers and sensors brings new opportunities (5). In this type of surveillance, the structured data collection across all the participating countries enable study and research of subgroups, e.g., the health status of the population outside the health care system (27, 28) or vaccine effectiveness in vaccinated groups and attitudes towards vaccination (such as influenza vaccination) (29, 30). The desirable benefits for participants, improving the long-term engagement in standard reporting, include real-time information on disease rates at local or neighbourhood levels (31).

Participatory systems have also been successfully implemented for mapping tropical diseases such as Zika and dengue and during mass gatherings like the 2016 Olympics in Rio de Janeiro, Brazil (32). Such a "hybrid" system (3, 33) (for example, using online surveys, communicating events, or using websites and applications) was successfully explored by combining Zika-related Google searches, Twitter microblogs, and the HealthMap digital surveillance system (34, 35) and cross-validating them with traditional disease surveillance data (36). Combining diverse data sources, search queries, social media data, digital data from internet-based sources (25), and website visits have proven effective for digital surveillance systems (37). With a new generation of trackers and sensors, this kind of individual surveillance will soon increase in scope, intensity, and significance (5), especially for emerging diseases and epidemic outbreaks, helping to monitor populations and potential security threats (5). Online surveillance-mapping tools have the potential to improve the early detection of infectious diseases compared to traditional epidemiological tools (38).

Online surveillance-mapping tools used in tandem with traditional epidemiological tools can improve the early detection of infectious diseases as was accomplished with Ebola (38) and Zika with critical results. Mobile phone data was used in Kenya to help identify the dynamics of human carriers that drive malaria parasite importation between regions (39); Google Trends was used for epidemiologic searching for Mayaro virus (40). All of these technologies are being employed for the COVID-19 epidemic (41). Mobile phone data and apps where users can report cases are also valuable tools for surveillance purposes (42). Due to the universal ownership of mobile devices with Bluetooth, these types of applications can be used all over the world, including low-connectivity and low-resource environments like rural Sierra Leone (43).

Many other instances of the digital surveillance of emerging infectious diseases have successfully used digital media for public health management (44, 45), including dengue (46), chikungunya

TABLE 1 | Mapping of four digital epidemiology disciplines to WHO objectives of COVID-19 surveillance.

The objectives of COVID-19 surveillance	Sub-disciplines of digital epidemiology
"Enable rapid detection, isolation, testing, and management of suspected cases"	Early warning and Epidemic intelligence digital surveillance
"Monitor trends in COVID-19 deaths"	Digital surveillance
"Identify, follow up and quarantine of contacts"	Digital surveillance
"Detect and contain clusters and outbreaks, especially among vulnerable populations"	Rapid response, outbreak control digital interventions
"Guide the implementation and adjustment of targeted control measures, while enabling safe resumption of economic and social activities"	Rapid response, outbreak control digital interventions
"Evaluate the impact of the pandemic on health-care systems and society"	Public communication and the impact on healthcare and society
"Monitor longer-term epidemiologic trends and evolution of SARS-CoV-2 virus"	Digital surveillance
"Contribute to the understanding of the co-circulation of SARS-CoV-2 virus, influenza and other respiratory viruses, and other pathogens"	Digital surveillance

(47), Ebola (48), monkeypox (49), and influenza (3, 50). The most common disease surveillance that utilises social media analytics for early detection and surveillance is influenza (51). For example, digital surveillance in influenza was used in Canada to demonstrate the correlation between influenza incidence and Google Ads click rates (52). It was also used to find the correlation between influenza incidence and Yahoo search trends (53). It is not only used at a national level but also in regional and local areas (54, 55).

Many examples show digital surveillance and epidemic intelligence noticeably using social networks, indicating the essential role social media plays, even according to the World Health Organization (WHO), for which more than 60% of initial disease epidemic reports derive from unofficial sources (56). Twitter, for instance, is used for surveillance because it can be used in outbreaks or emergencies, in monitoring diseases, in prediction, in gauging public reactions, in lifestyle analysis, in geolocation, and other general applications (57). Twitter was also used for Middle East respiratory syndrome surveillance in Korea (58), to analyse the H1N1 pandemic in 2009 (59), and to broadcast information about Ebola (60). In the literature, infectious disease surveillance, predicting disease spread, dissemination of public health information, and assessment of the public's views on public health outbreaks are some of the roles suggested for Twitter (34, 45, 51, 57–60).

Studies on digital information communication on social media sites are on the rise, as they have played an important role in real-time analysis and have been used for faster trend prediction (61); however, these studies are often small pilots that need more methodological rigour and scalability (62). More studies should be conducted using appropriate technologies such as web-based systems to help support data quality improvements and future reporting (63).

Early Warning and Epidemic Intelligence

Digital surveillance expands traditional epidemiology by adding information that previously did not exist. Important new sources of data include social networks, geographical location measured by GPS, wish lists, and consumption of mobile data, etc. These new types of information are not medically related (the classic epidemiological domain) but allow the expansion of early warning and prevention systems to prevent avoidable exposure to public health threats (64). Systematic review of digital data source their implication for health could be found in Li et al. (65)—some examples of novel data sources for the news of surveillance and epidemic intelligence are:

Online Media: Screening online news for mentions of specific diseases or conditions could be helpful to identify, for example, local food poisoning outbreaks. Special systems such as GPHIN (66), WHO EOIS (67), and HealthMap (68) screen all global media in multiple languages and were developed specifically to monitor epidemic events.

Online Searches: Search terms provide an invaluable geo-located monitoring tool for public information that could reveal public sentiment, shopping panics, or disease outbreaks, as demonstrated by Google Flu Trends in 2008 (40). While the opportunities are vast, commercial ownership of the search engines by Google and other tech giants prevents researchers and public health experts from exploring this resource for public health purposes. Online searches were also analysed on public medical websites such as the National Electronic Library of Infection and National Resource for Infection Control (69) to identify spikes in information needs resulting from the publication of major government guidelines (70).

Sensors, Digital Traces, and Internet of Things (IoT) Devices: Monitoring of population movements through citizens' digital traces via GPS-enabled phones, sensor networks, and credit/store cards seamlessly collects information about our moves, physical locations, purchases, online preferences, and payments and could provide early warnings of upcoming outbreaks. However, like Google, these datasets are mostly owned by commercial companies (supermarkets, pharmacies, etc.) and are not available for research or public health benefits.

Internet of things (IoT)-enabled devices and sensors allow real-time data streams from readings and measurements of environmental or smart home devices and could create opportunities for digital epidemiology such as mapping the spread of infection (71).

Mobile data and Mobility GPS data: Medical apps and games and health condition tracking devices support citizens in activities such as managing long-term conditions, increasing their physical activities, or losing weight (72) but create ethical challenges (64). GPS location and mobility data also

play a part in locations and directions (navigation and mapping apps), or contact tracing (e.g., COVID-19 contacts) using Bluetooth, GPS, cellular location tracking and QR codes (73). For example, geo-coded electronic health records (EHR) were successfully mapped to better visualise the spread of methicillin-resistant Staphylococcus aureus (CA-MRSA), identifying risks for CA-MRSA in children that would not have been uncovered using traditional EHR analyses (74). For accuracy and privacy, Bluetooth technology with a decentralised server architecture is recommended (75). Bluetooth, however, is not precise enough to avoid false positive contacts (10).

Social Media Streams: Unlike digital traces, collected seamlessly, the increase of Web 2.0, user-generated content actively shared via social networking tools, has seen an unprecedented explosion. The privacy settings of Facebook, Instagram, and other social networks allow users to restrict their profile content and activity. Consequently, the most important social media channel for research has undoubtedly become Twitter due to a relatively open data policy that allows researchers and IT developers access to tweets through an open, free API, returning a 1% random sample of raw tweets free of charge.

The use of social network data can also provide early warning because analysis can detect a peak in an outbreak up to 2 weeks before the official public health authorities (as occurred in 2009 for swine flu) (76). In 2019, earlier than the official reports, Twitter reported almost one-third of the total notifications related to avian influenza outbreak (77). This corroborates that social networks can serve as a method to obtain valuable information on a disease's behaviour or spread, even a week in advance of what the general practitioners in a particular locality could report (78). The evaluation of commonly used drugs for seasonal influenza on Twitter also provides surveillance ahead of the flu season (79).

Social media data allow disease tracking (80, 81) and can help make predictions that could prevent danger to the population (3). Social media analytics in correlation with traditional laboratory data can predict an outbreak; examples of this include the cases of influenza and cholera (82) or Ebola and Marburg filoviruses (83). Some researchers also conducted a content analysis to identify key trends during the 2009 H1N1outbreak that could also correlate with outbreak incidence data (84). Even concerning the coronavirus, Kogan et al. found that Twitter could be used as a kind of barometer, showing potential growth 2–3 weeks before the growth of coronavirus cases (by region) (85) and in other cases even 8–12 days before the outbreak (86). This is not new. An analysis of more than 500 million tweets worldwide found a significant association between the geographic locations of HIV-related tweets and HIV prevalence (87).

Twitter data in 2010 accurately tracked the spread of cholera, but researchers advised that this type of information is not always reliable and must not replace traditional epidemiological methods, as information and guidance is missed (88). Moreover, social network analyses need to be challenged and scrutinised by the "ground truth." In 2016, Mowery et al. conducted a study describing how epidemiological surveillance of influenza using Twitter incorrectly predicted the 2011–2012 flu season 3 months early (89).

Improvements in services and cost reductions in the health sector coupled with the need for early warnings for the onset of adverse health conditions are the main drivers of these developments and new sources of data (90). These novel data sources collect a massive amount of new information, and it can be difficult for big data tools to review the data obtained and present it as new useful knowledge for the prevention of a pandemic outbreak. However, the use of big data is proving crucial for the COVID-19 pandemic (91).

Rapid Response, Outbreak Control, and Digital Interventions

Once a threat or an outbreak is detected, an event verified, and risk assessed, a rapid response must be implemented to control the outbreak, which ranges from diagnosis, testing, and contact tracing to risk communication with citizens.

This iterative process includes discovering patterns and generating new information (92) that can be used to control the outbreak. Digital information and new technologies are providing a quick response and coordinated control and management of possible outbreaks. Through monitoring cases using mobile technology, contact tracing infected citizens, following up with patients, and providing medical advice, digital and mobile technology can successfully complement the efforts of medical and public health experts (93).

With social networks, big data enables the early epidemiological storey of an outbreak to be reconstructed (94). Data streams and non-classical datasets in the early stages of the outbreak can inform the design and implementation of effective public health measures (95).

Big data could effectively support the rapid response to and better control of an outbreak but could do so more quickly and generate more accurate predictions. The algorithms that accomplish this are known as artificial intelligence (AI). For example, Toronto's surveillance system was first to detect the COVID-19 epidemic outbreak in the first reported epicentre of Wuhan (96). Also, scientists from the John Hopkins University visualised the spread of the coronavirus in real-time (95). AI is changing the landscape in public health and clinical management with very promising results (97). For example, a project seeking information related to the COVID-19 called Evidence Navigator provides computer-generated evidence maps of scientific publications on the pandemic, which is updated daily in PubMed (98).

Machine learning (ML) is a dramatically growing computer science AI discipline investigating algorithms to find new results or predictions without looking for specific solutions. ML was used to analyse several projects using the internet to enhance epidemiological surveillance and disease prediction [e.g., malaria (99), dengue (100), and influenza (101)]. The research demonstrated a positive predictive value of the incidence of infectious diseases (101) or little predictive value (102). ML was successfully used to create an evidence-based guideline from the

information gathered from the Ebola virus epidemic and turn it into an application called the Ebola Care Guidelines app (103) to inform the general population and healthcare providers of updated, evidence-based guidelines in real time during a global pandemic (41).

In Africa, the rapid recognition of localised areas of higher transmission of Ebola and the resulting quantitative assessment could support the optimal deployment of public health resources (104). In Latin America and the Caribbean, the rapid integration of Zika virus prevention recommendations into sexual and reproductive health services made it possible to reduce the incidence of the Zika virus (105). In 2012, digital pens were used by the New Hampshire Department of Health and Human Services to rapidly acquire epidemiologic data during a gastrointestinal illness outbreak (106), providing rapid assessment, response, and control measures before the problem multiplied.

The novel digital information gathered about the coronavirus is being used in Taiwan in conjunction with their immigration database to classify travellers according to different risk types and to issue alerts in real time to prevent infections (107). This is often called a "digital fence." Russia, China, and Poland have used facial recognition software to monitor population compliance with government policies (42). Applications exist for self-testing, quarantine monitoring, and contact tracing that are being already used in many parts of the world such as India to support the rapid response to COVID-19 (108).

Online news surveillance and mapping tools have successfully provided early warnings, but their potential to improve response and risk assessment as the outbreak progresses over time has led to the adoption by WHO operations of the most robust systems: e.g., GPHIN, Medisys, SORMAS (Surveillance and Outbreak Response Management and Analysis System) (109) and HealthMap (68). News surveillance also means that real-time feedback and effective responses should function as an intervention. Locally appropriate technologies, such as web-based systems and mobile phones, can help support data quality improvements and reporting timeliness (63).

In response to the COVID-19 outbreak, the rapid development and deployment of point-of-care (POC) diagnostics for screening have shown to help slow the spread of the disease (110), demonstrating that telemedicine can be used as a means of surveillance. For example, Botswana is using telemedicine to control patients remotely (111), and obstetric departments in the US are monitoring coronavirus patients (112). Rapid deployment of an in-patient telemedicine response is feasible across many settings in response to the COVID-19 pandemic (113).

Public Communication and Evaluation of the Impact on Healthcare and Society

An essential piece at the centre of outbreak response success is public communication: a clear and concise message communicating the risk, measures, and policies taken by the government. Digital technologies complement traditional mass media and play an increasingly important role for sharing reliable, evidence-based information with the public and gaining citizens' buy-in (114).

While traditional media channels—newspapers and television—are still actively used for mass communication, the role of social media has grown dramatically. In particular, Twitter has demonstrated great potential to be used not only for tracking epidemics but also to inform citizens about the risks of pandemics in real time (57). Social media can be also used to study the public's risk perceptions (115). Mobile apps and wearable devices have further been used to monitor patient behaviour to provide personalised service and advice (116, 117). Twitter was analysed for risk communication potential during the epidemics of the Middle East respiratory syndrome (58), SARS (45), Ebola virus (60), Zika virus (34), H1N1 ("swine flu") (59), and H7N9 ("avian flu") (62). A study in Vietnam includes social media and science journalist for COVID-19 public policy[2]. The studies involving Twitter and other social networks summarise how they could be useful tools for disseminating information to the population about how to avoid the spread of the outbreak (118).

Specifically, role-play social media, a channel for gauging public attitudes, could effectively be used to disseminate risk communication in real time (78) but could be a double-edged sword in that also it is prone to misuse, misinformation, and the spread of fake news. How online channels inform healthcare professionals and the public and their genuine information needs to be part of any comprehensive government communication plan (70).

DISCUSSION

The results have shown many opportunities ranging from the use of social networks to the use of AI and big data for digital surveillance and reference early warning and epidemic intelligence, rapid response, outbreak control, risk communication, and public communication. The results present solutions that have been implemented in many countries with differing results because of their diverse socio-cultural realities. In August 2020, WHO updated the objectives for COVID-19 surveillance: (i) enable rapid detection, isolation, testing, and management of cases; (ii) monitor trends in COVID-19 deaths; (iii) identify, follow-up, and quarantine of contacts; (iv) detect and contain clusters and outbreaks, especially among vulnerable populations; (v) guide the implementation and adjustment of targeted control measures while enabling safe resumption of economic and social activities; (vi) evaluate the impact of the pandemic on healthcare systems and society; (vii) monitor longer term epidemiologic trends and evolution of SARS-CoV-2 virus; and (viii) contribute to the understanding of the co-circulation of SARS-CoV-2 virus, influenza and other respiratory viruses, and other pathogens (119).

With digital epidemiology's subdisciplines already assessed and ready for use, the population must be made aware of the measures to be followed to control the outbreak using (ix) public communication and evaluation of the impact on healthcare and society as culture might shape how people think about privacy and surveillance Risk communication and

[2]https://www.mdpi.com/2071-1050/12/7/2931/htm

public communication are now taken to a new level through mobile apps and social media, while the advances in epidemic modelling that leverage real-time heterogeneous data improve public health policy through better assessment of the impact of control measures on healthcare and society.

Mobile applications and mHealth approaches could be used for public health surveillance due to their multiple benefits as an efficient, almost universal presence (120), a contributor to high digital literacy, and a source that creates wide availability of data. With the number of mobile phones around 14 billion units worldwide and expecting an increase to almost 17 billion by 2023, the options available with this type of device must be considered. The use of mobile phones and Bluetooth to assess the follow-up of potential patients has always been a strong alternative (121). Apps with tracing functionalities using Bluetooth technology can support health authorities in the contact tracing process, identifying the possible contacts (known or unknown) of a confirmed/positive case, creating a network that will function as epidemic control if it is used by enough people (122). Some examples exist of mobile apps and mHealth being used already. For example, the Chinese government required citizens in 200 cities to use the Alipay app, assigning a risk code (green, yellow, red) to each person indicating to what extent they were allowed to move around the community. An algorithm incorporated information about the time they spent in risky locations and the frequency of contact with other people (42).

Although these technologies only use monitoring data temporarily, many people are reluctant to use them because they think that their data will be sold to private companies or they will continue to be monitored once the pandemic is over (123). Therefore, notification is necessary that the data obtained will be used only for monitoring purposes, that it will not be transferred to any public or private company, and that it will only be used in periods of a pandemic or used as anonymised data (124).

Mobile devices (including tablets, wearables, etc.) also can connect to social networks, which allows their users to be informed of the latest news about the pandemic. At the same time, they can be used to provide measures to avoid contagion. Social media's power of influence is vast, becoming an important tool public health during a pandemic, which is precisely a difference between digital and traditional epidemiology. Social networks can provide a different perspective than the traditional approach that relies on health reports by providing important correlations with abnormal disease trends that could indicate a potential outbreak.

The use of social media also offers significant opportunities to encourage citizen engagement in crisis management (125). However, not all the most-read information on social networks is true. Many celebrity "influencers" (people with many followers, such as actors, athletes, etc.) publish controversial posts such as false cures or prophylactics for the coronavirus or indicate that it is simply an invention (126). This is due to voluntary submission and a lack of gatekeeping (127). The advances in AI could help in this facet (128). This kind of digital technology, correctly applied, could benefit the healthcare landscape in public health and clinical management with promising results (97). It has already been used for some time to handle data related to the coronavirus (96).

Big data use for surveillance seems to be a beneficial tool for public health but it must follow a rigorous statistical analysis. Qin et al. used social media search indexes (SMSI) to predict the new number of COVID-19 cases that could be detected 6–9 days in advance (129), but the prediction is generated after an algorithm has evaluated the information. With all this generation of digital information, there is a risk of losing what is impactful to the epidemiological study. To accomplish an "understanding of all data" created by social networks, the use of big data has also been implemented.

While useful in digital epidemiology, big data can also serve as a standard to control possible adverse effects and can support traditional epidemiology by discovering additional facts from social media behaviour that may complement other data from the population. Furthermore, it can help overcome some challenges, such as geographic heterogeneity, insufficient representation in developing countries, and spatial/temporal uncertainty in the information obtained. On the other hand, a disadvantage of using internet search data or data from social networks for surveillance purposes is its—eventual—lack of representativeness, as well as possible fake results. In order to create big data, information must be collected and processed before it can be used; therefore, the WHO has called for requiring the detection and management of suspicious cases at entry points worldwide to generate this information to avoid the spread of COVID-19 (130). Information from symptoms checkers and laboratory data could help to estimate the predictive value of the respiratory symptoms on the community as well as present facts on the level of virus circulation (93).

The use of this technology would allow epidemiologists to evaluate millions of digital trails of people who constantly use their digital equipment such as mobile phones for social networking. However, an ideal prediction using machine learning and AI is not yet realised; the tracking and prediction of how COVID-19 will spread are not yet completely reliable. This could be due to two reasons: (1) the availability of COVID-19-related clinical data is a key barrier; and (2) AI requires data on COVID-19 to train itself (131). Therefore, the WHO recommends being cautious with the implementation of digital solutions until the utility of public health policies are better understood. Planned solutions should cover and allow (i) quality monitoring, transparency, and accountability, (ii) resource allocation optimisation, (iii) citizen participation and inclusion, and (iv) resilience and adaptation to exogenous events.

Several promising initiatives have been launched to gather and share both existing and new data using new AI models, including WHO's Global Research on Coronavirus Disease Database, along with the GISAID Initiative (Global Initiative on Sharing All Influenza Data); the COVID-19 Open Research Dataset Challenge of Kaggle data science platform; the around 20,000 related articles in ScienceDirect in its Novel Coronavirus Information Centre early-stage and peer-reviewed research on COVID-19; etc. Finally, the initiative formed by Microsoft, Facebook, Semantic Scholar, the Allen Institute for AI, and five other collaborators to make the COVID-19 Open Research Dataset (CORD-19) openly available, which contains about 44,000 scholarly articles for data mining, is worth mentioning.

Finally, there are a number of related subjects that are beyond the scope of this paper. Firstly, cost issues and culture aspects of surveillance were not part of our initial search strategies, however, it is important to note the cost to setting up and conducting digital surveillance, especially in low and middle income countries[4,5]. Culture attitudes (positive or negative) towards digital surveillance are also essential to consider for a successful deployment of a digital solution[6,7] as demonstrated in fight against COVID-19 in Vietnam[8]. Final limitation of this study includes our search strategy focusing on PubMed and therefore not covering studies published in computer science outlets such as conferences published in ACM and IEEE libraries, such as pioneering Twitter research for epidemic intelligence and eastly warning (see text footnote 1)—this is a subject for a future research.

CONCLUSIONS

We have highlighted the opportunities and challenges for digital epidemiology as a growing discipline that have contributed to surveillance of the COVID-19 pandemic and highlighted how digital epidemiology will become indispensable for fighting future public health and natural disasters and pandemics.

Prevention and control of the COVID-19 pandemic requires public health and epidemiology measures. The use of digital technology enhances traditional epidemiological means to contain outbreak and supports prevention, early warning, rapid response, and digital interventions such as remote care for patients or providing reliable information to the public. In addition, it is crucial that the technology is inclusive and user friendly (for example, social networks and specially designed apps). However, additional support strategies are required for vulnerable groups who are not active technology users.

The key opportunities and challenges for effective digital epidemiology systems for the 21st century lie in front of us. To improve capacity and preparedness for later epidemics, the repurposed and emerging systems for COVID-19 need to be fully developed and evaluated after the crisis with the goal of creating fully integrated, interoperable digital epidemiology solutions at national and international levels.

AUTHOR CONTRIBUTIONS

All authors contributed sufficiently and meaningfully to the conception, design, drafting, editing, and revising the manuscript.

[4]https://www.nature.com/articles/s41562-017-0281-4
[5]https://springerplus.springeropen.com/articles/10.1186/s40064-015-1279-x
[6]https://www.nature.com/articles/s41599-018-0189-2
[7]https://www.frontiersin.org/articles/10.3389/fpsyt.2020.589618/full
[8]https://dl.acm.org/doi/10.1145/2597892

REFERENCES

1. Fine P, Victora CG, Rothman KJ, Moore PS, Chang Y, Curtis V, et al. John Snow's legacy: epidemiology without borders. *Lancet*. (2013) 381:1302–11. doi: 10.1016/S0140-6736(13)60771-0
2. Mapping A London Epidemic | National Geographic Society. (n.d.). Available online at: https://www.nationalgeographic.org/activity/mapping-london-epidemic/ (accessed July 9, 2021).
3. Aiello AE, Renson A, Zivich PN. Social media– and internet-based disease surveillance for public health. *Annu Rev Public Health*. (2020) 41:101–18. doi: 10.1146/annurev-publhealth-040119-094402
4. St Louis C, Zorlu G. Can Twitter predict disease outbreaks? *BMJ*. (2012) 344:e2353. doi: 10.1136/bmj.e2353
5. Samerski S. Individuals on alert: digital epidemiology and the individualization of surveillance. *Life Sci Soc Pol*. (2018) 14:13. doi: 10.1186/s40504-018-0076-z
6. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS. Digital epidemiology. *PLoS Comput Biol.* (2012) 8:e1002616. doi: 10.1371/journal.pcbi.1002616
7. Barros JM, Duggan J, Rebholz-Schuhmann D. The application of internet-based sources for public health surveillance (infoveillance): systematic review. *J Med Internet Res*. (2020) 22:e13680. doi: 10.2196/13680
8. Park HW, Park S, Chong M. Conversations and medical news frames on twitter: infodemiological study on COVID-19 in South Korea. *J Med Internet Res*. (2020) 22:e18897. doi: 10.2196/18897
9. Wyber R, Vaillancourt S, Perry W, Mannava P, Folaranmi T, Celi LA. Big data in global health: improving health in low- and middle-income countries. *Bull World Health Organ*. (2015) 93:203–8. doi: 10.2471/BLT.14.139022
10. Ward JS, Barker A. *Undefined By Data: A Survey of Big Data Definitions*. (2013). https://arxiv.org/abs/1309.5821v1 (accessed July 9, 2021).
11. Gaitanou P, Garoufallou E, Balatsoukas P. The effectiveness of big data in health care: a systematic review. *Commun Comput Inform Sci*. (2014) 478:141–53. doi: 10.1007/978-3-319-13674-5_14
12. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*. (2005) 331:1064–5. doi: 10.1136/bmj.38636.593461.68
13. World Health Organization (WHO). *Management of Ill Travellers at Points of Entry - International Airports, Seaports and Ground Crossings - in the Context of COVID-19 Outbreak* (2020). Available online at: https://www.who.int/publications/i/item/10665-331512 (accessed July 9, 2021).
14. Bergquist T, Pejaver V, Hammarlund N, Mooney SD, Mooney SJ. Evaluation of the secondary use of electronic health records to detect seasonal, holiday-related, and rare events related to traumatic injury and poisoning. *BMC Public Health*. (2020) 20:46. doi: 10.1186/s12889-020-8153-7
15. Hartley DM. Using social media and internet data for public health surveillance: the importance of talking. Milbank Q. (2014) 92:34–9. doi: 10.1111/1468-0009.12039
16. Salathé M. Digital epidemiology: what is it, and where is it going? *Life Sci Soc Policy*. (2018) 14:1. doi: 10.1186/s40504-017-0065-7
17. Yan SJ, Chughtai AA, Macintyre CR. Utility potential of rapid epidemic intelligence from internet-based sources. *Int J Infect Dis*. (2017) 63:77–87. doi: 10.1016/j.ijid.2017.07.020
18. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. *J Am Med Inform Assoc*. (2004) 11:141–50. doi: 10.1197/jamia.M1356
19. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, et al. *Public Health Surveillance: A Tool for Targeting and Monitoring Interventions*. (2006). Disease Control Priorities in Developing Countries. The International Bank for Reconstruction and Development/The World Bank.
20. Bringay S, Flamand C, Quenel P, Ardillon V, Carvalho L, Teisseire M. The Epidemiologic Surveillance of Dengue-Fever in French Guiana: When Achievements Trigger Higher Goals Influenza Surveillance View Project Air Pollution View Project Claude Flamand, Institut Pasteur International Network (2016).

21. Brownstein JS, Freifeld CC, Reis BY, Mandl KD. Surveillance sans frontières: internet-based emerging infectious disease intelligence and the healthmap project. *PLoS Med.* (2008) 5:e151. doi: 10.1371/journal.pmed.0050151
22. Henning KJ. Overview of syndromic surveillance what is syndromic surveillance? *September.* (2004) 24:2004. doi: 10.1037/e307182005-001
23. Muscatello DJ, Churches T, Kaldor J, Zheng W, Chiu C, Correll P, et al. An automated, broad-Based, near real-Time public health surveillance system using presentations to hospital emergency departments in new south wales, australia. *BMC Public Health.* (2005) 5:141. doi: 10.1186/1471-2458-5-141
24. Gilbert GL, Degeling C, Johnson J. Communicable disease surveillance ethics in the age of big data and new technology. *Asian Bioethics Rev.* (2019) 11:173–87. doi: 10.1007/s41649-019-00087-1
25. Milinovich GJ, Avril SMR, Clements ACA, Brownstein JS, Tong S, Hu W. Using internet search queries for infectious disease surveillance: screening diseases for suitability. *BMC Infect Dis.* (2014) 14:690. doi: 10.1186/s12879-014-0690-1
26. Fernandez-Luque L, Karlsen R, Melton GB. HealthTrust: a social network approach for retrieving online health videos. *J Med Internet Res.* (2012) 14:e22. doi: 10.2196/jmir.1985
27. Guerrisi C, Turbelin C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, et al. Participatory syndromic surveillance of influenza in Europe. *J Infect Dis.* (2016) 214: S386–92. doi: 10.1093/infdis/jiw280
28. Debin M, Turbelin C, Blanchon T, Bonmarin I, Falchi A, Hanslik T. Evaluating the feasibility participants' representativeness of an online nationwide surveillance system for influenza in france. *PLoS ONE.* (2013) 8:e73675. doi: 10.1371/journal.pone.0073675
29. Debin M, Colizza V, Blanchon T, Hanslik T, Turbelin C, Falchi A. Effectiveness of 2012-2013 influenza vaccine against influenza-like illness in general population: estimation in a french web-Based cohort. *Hum Vaccines Immunother.* (2014) 10:536–43. doi: 10.4161/hv.27439
30. Boiron K, Sarazin M, Debin M, Raude J, Rossignol L, Guerrisi C, et al. Opinion about seasonal influenza vaccination among the general population 3 years after the a(H1N1)Pdm2009 influenza pandemic. *Vaccine.* (2015) 33:6849–54. doi: 10.1016/j.vaccine.2015.08.067
31. Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, et al. Influenzanet: citizens among 10 countries collaborating to monitor influenza in europe. *JMIR Public Health Surveill.* (2017) 3:e66. doi: 10.2196/publichealth.7429
32. Leal Neto O, Cruz O, Albuquerque J, Nacarato de Sousa M, Smolinski M, Pessoa Cesse EÂ, et al. Participatory surveillance based on crowdsourcing during the rio 2016 olympic games using the guardians of health platform: descriptive study. *JMIR Public Health Surveill.* (2020) 6:e16119. doi: 10.2196/16119
33. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C. Big data for infectious disease surveillance and modeling. *J Infect Dis.* (2016) 214:S375–9. doi: 10.1093/infdis/jiw400
34. Masri S, Jia J, Li C, Zhou G, Lee MC, Yan G, et al. Use of Twitter data to improve zika virus surveillance in the United States during the 2016 epidemic. *BMC Public Health.* (2019) 19:761. doi: 10.1186/s12889-019-7103-8
35. McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting zika incidence in the 2016 Latin America outbreak combining traditional disease surveillance with search, social media, and news report data. *PLoS Neglect Trop Dis.* (2017) 11:e0005295. doi: 10.1371/journal.pntd.0005295
36. Li J, Xu Q, Cuomo R, Purushothaman V, Mackey T. Data mining and content analysis of the chinese social media platform weibo during the early cOVID-19 outbreak: retrospective observational infoveillance study. *JMIR Public Health Surveill.* (2020) 6:e18700. doi: 10.2196/18700
37. Seo DW, Shin SY. Methods using social media and search queries to predict infectious disease outbreaks. *Healthc Inform Res.* (2017) 23:343–8. doi: 10.4258/hir.2017.23.4.343
38. Bempong NE, De Castañeda RR, Schütte S, Bolon I, Keiser O, Escher G, et al. Precision global health - the case of ebola: a scoping review. *J Glob Health.* (2019) 9:010404. doi: 10.7189/jogh.09.010404
39. Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, Snow RW, et al. Quantifying the impact of human mobility on malaria. *Science.* (2012) 338:267–70. doi: 10.1126/science.1223467
40. Adawi M, Bragazzi NL, Watad A, Sharif K, Amital H, Mahroum N. Discrepancies between classic and digital epidemiology in searching for the mayaro virus: preliminary qualitative and quantitative analysis of google trends. *JMIR Public Health Surveill.* (2017) 3:e93. doi: 10.2196/publichealth.9136
41. Alwashmi MF. The use of digital health in the detection and management of COVID-19. *Int J Environ Res Public Health.* (2020) 17:2906. doi: 10.3390/ijerph17082906
42. Mello MM, Wang CJ. Ethics and governance for digital disease surveillance. *Science.* (2020) 368:951–4. doi: 10.1126/science.abb9045
43. Danquah LO, Hasham N, MacFarlane M, Conteh FE, Momoh F, Tedesco AA, et al. Use of a mobile application for Ebola contact tracing and monitoring in northern Sierra Leone: a proof-of-concept study. *BMC Infect Dis.* (2019) 19:810. doi: 10.1186/s12879-019-4354-z
44. Geneviève LD, Martani A, Wangmo T, Paolotti D, Koppeschaar C, Kjelsø C, et al. Participatory disease surveillance systems: ethical framework. *J Med Internet Res.* (2019) 21:e12273. doi: 10.2196/12273
45. Carrion M, Madoff LC. ProMED-mail: 22 years of digital surveillance of emerging infectious diseases. *Int Health.* (2017) 9:177–83. doi: 10.1093/inthealth/ihx014
46. Buczak AL, Baugher B, Moniz LJ, Bagley T, Babin SM, Guven E. Ensemble method for dengue prediction. *PLoS ONE.* (2018) 13:e0189988. doi: 10.1371/journal.pone.0189988
47. Del Valle SY, McMahon BH, Asher J, Hatchett R, Lega JC, Brown HE, et al. Summary results of the 2014-2015 DARPA Chikungunya challenge. *BMC Infect Dis.* (2018) 18:245. doi: 10.1186/s12879-018-3124-7
48. Viboud C, Sun K, Gaffey R, Ajelli M, Fumanelli L, Merler S, et al. The rAPIDD ebola forecasting challenge: synthesis and lessons learnt. *Epidemics.* (2018) 22:13–21. doi: 10.1016/j.epidem.2017.08.002
49. Silenou BC, Tom-Aba D, Adeoye O, Arinze CC, Oyiri F, Suleman AK, et al. Use of surveillance outbreak response management and analysis system for human monkeypox outbreak, nigeria, 2017-2019. *Emerg Infect Dis.* (2020) 26:345–9. doi: 10.3201/eid2602.191139
50. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. (2015). Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comput Biol.* 11:E1004513. doi: 10.1371/journal.pcbi.1004513
51. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a Chronology of innovation. *J Med Internet Res.* (2013) 15:e147. doi: 10.2196/jmir.2740
52. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res.* (2009) 11:e11. doi: 10.2196/jmir.1157
53. Porta M. Incidence (Syn: Incidence Number). In: Porta M, editor. *A Dictionary of Epidemiology*. New York, NY: Oxford University Press (2015). p. 144.
54. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clin Infect Dis.* (2009) 49:1557–64. doi: 10.1086/630200
55. Chunara R, Aman S, Smolinski M, Brownstein JS. Flu near you: an online self-reported influenza surveillance system in the USA. *Online J Public health Inform.* (2013) 5:4456. doi: 10.5210/ojphi.v5i1.4456
56. WHO | Epidemic Intelligence - Systematic Event Detection. WHO (2020).
57. Aramburu MJ, Berlanga R, Lanza I. Social Media Multidimensional Analysis for Intelligent Health Surveillance. *Int J Environ Res Public Health.* (2020) 17:2289. doi: 10.3390/ijerph17072289
58. Shin SY, Seo DW, An J, Kwak H, Kim SH, Gwack J, et al. High Correlation of middle east respiratory syndrome spread with google search and twitter trends in Korea. *Scie Rep.* (2016) 6:32920. doi: 10.1038/srep32920
59. Ahmed W, Bath PA, Sbaffi L, Demartini G. Novel insights into views towards h1N1 during the 2009 pandemic: a Thematic analysis of twitter data. *Health Info Libr J.* (2019) 36:60–72. doi: 10.1111/hir.12247
60. Liang H, Fung ICH, Tse ZTH, Yin J, Chan CH, Pechta LE, et al. How did Ebola information spread on Twitter: broadcasting or viral spreading? *BMC Public Health.* (2019) 19:438. doi: 10.1186/s12889-019-6747-8

61. Alessa A, Faezipour M. A review of influenza detection and prediction through social networking sites. *Theor Biol Med Model.* (2018) 15:2. doi: 10.1186/s12976-017-0074-5
62. Tang L, Bie B, Park SE, Zhi D. Social media and outbreaks of emerging infectious diseases: a systematic review of literature. *Am J Infect Control.* (2018) 46:962–72. doi: 10.1016/j.ajic.2018.02.010
63. Ohrt C, Roberts KW, Sturrock HJW, Wegbreit J, Lee BY, Gosling RD. Information systems to support surveillance for malaria elimination. *Am J Trop Med Hygiene.* (2015) 93:145–52. doi: 10.4269/ajtmh.14-0257
64. Kostkova P. Disease surveillance data sharing for public health: the next ethical frontiers. *Life Sci Soc Policy.* (2018) 14:16. doi: 10.1186/s40504-018-0078-x
65. Li L, Novillo-Ortiz D, Azzopardi-Muscat N, Kostkova P. Digital data sources and their impact on people's health: a systematic review of systematic reviews. *Front Public Health.* (2021) 9:645260. doi: 10.3389/fpubh.2021.645260
66. About - GPHIN - Canada.Ca. (n.d.). Available online at: https://gphin.canada.ca/cepr/aboutgphin-rmispenbref.jsp?language=en_CA (accessed June 18, 2021).
67. World Health Organization (WHO). (n.d.). *Early Detection, Verification, Assessment and Communication. Epidemic Intelligence from Open Sources (EIOS).* Available online at: https://www.who.int/initiatives/eios (accessed July 9, 2021).
68. Flu and Ebola Map | Virus and Contagious Disease Surveillance. (n.d.). Available online at: https://healthmap.org/pt/ (accessed July 9, 2021).
69. Welcome to National Resource for Infection Control (NRIC) | National Resource for Infection Control (NRIC). London: University College London (n.d.).
70. Kostkova P, Fowler D, Wiseman S, Weinberg JR. Major infection events over 5 years: how is media coverage influencing online information needs of health care professionals and the public? *J Med Int Res.* (2013) 15:e107. doi: 10.2196/jmir.2146
71. Peeri NC, Shrestha N, Rahman MS, Zaki R, Tan Z, Bibi S, et al. The sARS, mERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? *Int J Epidemiol.* (2020) 49:717–26. doi: 10.1093/ije/dyaa033
72. Kostkova P. Grand challenges in digital health. *Front Public Health.* (2015) 3:1. doi: 10.3389/fpubh.2015.00134
73. Soontornpipit P, Viwatwongkasem C, Taratep C, Teerawat W, Vanitchatchavan P. Development of the electronic surveillance monitoring system on web applications. *Proc Comput Sci.* (2016) 86:244–7. doi: 10.1016/j.procs.2016.05.110
74. Immergluck LC, Leong T, Matthews K, Malhotra K, Parker TC, Ali F, et al. Geographic surveillance of community associated MRSA infections in children using electronic health record data. *BMC Infect Dis.* (2019) 19170. doi: 10.1186/s12879-019-3972-9
75. How Apple Google Are Enabling Covid-19 Bluetooth Contact-Tracing | WIRED (2020). Available online at: https://www.wired.com/story/apple-google-bluetooth-contact-tracing-covid-19/ (accessed July 9, 2021).
76. De Quincey E, Kostkova P. Early warning and outbreak detection using social networking websites: the potential of twitter. In: *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 27 LNICST:21–24. Berlin, Heidelberg: Springer (2010). doi: 10.1007/978-3-642-11745-9_4
77. Yousefinaghani S, Dara R, Poljak Z, Bernardo TM, Sharif S. The assessment of twitter's potential for outbreak detection: avian influenza case study. *Sci Rep.* (2019) 9:18147. doi: 10.1038/s41598-019-54388-4
78. Szomszor M, Kostkova P, De Quincey. E. #Swineflu: twitter predicts swine flu outbreak in 2009. In: *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 69 LNICST:18–26. Berlin, Heidelberg: Springer (2011). doi: 10.1007/978-3-642-23635-8_3
79. Kagashe I, Yan Z, Suheryani I. Enhancing seasonal influenza surveillance: topic analysis of widely used medicinal drugs using twitter data. *J Med Internet Res.* (2017) 19:e315. doi: 10.2196/jmir.7393
80. Lampos V, De Bie T, Cristianini N. Flu detector-tracking epidemics on twitter. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6323 LNAI:599–602. Berlin, Heidelberg: Springer (2010). doi: 10.1007/978-3-642-15939-8_42
81. Lampos V, Cristianini N. Nowcasting events from the social web with statistical learning. *ACM Trans Intell Syst Technol.* (2012) 3:1–22. doi: 10.1145/2337542.2337557
82. Hossain L, Kam D, Kong F, Wigand RT, Bossomaier T. Social media in ebola outbreak. *Epidemiol Infect.* (2016) 144:2136–43. doi: 10.1017/S095026881600039X
83. Natesan M, Wu SW, Chen CI, Jensen SMR, Karlovac N, Dyas BK, et al. A smartphone-Based rapid telemonitoring system for ebola and marburg disease surveillance. *ACS Sensors.* (2019) 4:61–8. doi: 10.1021/acssensors.8b00842
84. Chew C, Eysenbach G. Pandemics in the age of twitter: content analysis of tweets during the 2009 h1N1 outbreak. *PLoS ONE.* (2010) 5:e14118. doi: 10.1371/journal.pone.0014118
85. Kogan NE, Clemente L, Liautaud P, Kaashoek J, Link NB, Nguyen AT, et al. An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time. *Sci Adv*. (2021) 7:eabd6989. doi: 10.1126/sciadv.abd6989
86. Li C, Chen LJ, Chen X, Zhang M, Pang CP, Chen H. Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. *Euro Surveill*. (2020) 25:2000199. doi: 10.2807/1560-7917.ES.2020.25.10.2000199
87. Young SD, Rivers C, Lewis B. Methods of using real-Time social media technologies for detection and remote monitoring of hIV outcomes. *Prev Med.* (2014) 63:112–5. doi: 10.1016/j.ypmed.2014.01.024
88. Bates M. Tracking disease: digital epidemiology offers new promise in predicting outbreaks. *IEEE Pulse.* (2017) 8:18–22. doi: 10.1109/MPUL.2016.2627238
89. Mowery J. Twitter influenza surveillance: quantifying seasonal misdiagnosis patterns. *Online J Public Health Inform.* (2016) 8:7011. doi: 10.5210/ojphi.v8i3.7011
90. Page T. A forecast of the adoption of wearable technology. *Int J Technol Diff.* (2015) 6:12–29. doi: 10.4018/IJTD.2015040102
91. Qiu HJ LX, Yuan XK, Huang YQ, Zhou QW, Wu R, et al. Using the big data of internet to understand the characteristics of coronavirus disease 2019: a Big data study. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi.* (2020) 55:569–75. doi: 10.1016/j.wjorl.2020.05.003
92. Garattini C, Raffle J, Aisyah DN, Sartain F, Kozlakidis Z. Big data analytics, infectious diseases and associated ethical impacts. *Philos Technol.* (2019) 32:69–85. doi: 10.1007/s13347-017-0278-y
93. European Commission (2020). *Mobile Applications to Support Contact Tracing in the EU's Fight against COVID-19 - Common EU Toolbox for Member States*. Brussels. Available online at: https://ec.europa.eu/health/sites/default/files/ehealth/docs/covid-19_apps_en.pdf (accessed July 9, 2021).
94. Sun K, Chen J, Viboud C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *Lancet Digital Health.* (2020) 2:e201–8. doi: 10.1016/S2589-7500(20)30026-1
95. Bragazzi NL, Dai H, Damiani G, Behzadifar M, Martini M, Wu J. (2020). How big data and artificial intelligence can help better manage the covid-19 pandemic. *Int J Environ Res Public Health.* 17:3176. doi: 10.3390/ijerph17093176
96. McCall B. COVID-19 artificial intelligence: protecting health-care workers and curbing the spread. *Lancet Digital Health.* (2020) 2:e166–7. doi: 10.1016/S2589-7500(20)30054-6
97. Shaban-Nejad A, Michalowski M, Buckeridge DL. Health intelligence: how artificial intelligence transforms population and personalized health. *Npj Digit Med.* (2018) 1:53. doi: 10.1038/s41746-018-0058-9
98. Antons D, Grünwald E, Cichy P, Salge TO. The application of text mining methods in innovation research: current state, evolution patterns, development priorities. *RandD Manage.* (2020) 50:329–51. doi: 10.1111/radm.12408
99. Ocampo AJ, Chunara R, Brownstein JS. Using search queries for malaria surveillance, thailand. *Malar J.* (2013) 12:390. doi: 10.1186/1475-2875-12-390

100. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a New model for neglected tropical disease surveillance. *PLoS Neglect Trop Dis.* (2011) 5:e1206. doi: 10.1371/journal.pntd.0001206
101. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature.* (2009) 457:1012–4. doi: 10.1038/nature07634
102. Daughton AR, Chunara R, Paul MJ. Comparison of social media, syndromic surveillance, and microbiologic acute respiratory infection data: observational study. *JMIR Public Health Surveill.* (2020) 6:e14986. doi: 10.2196/14986
103. Colubri A, Hartley MA, Siakor M, Wolfman V, Felix A, Sesay T, et al. Machine-learning prognostic models from the 2014–16 ebola outbreak: data-harmonization challenges, validation strategies, mHealth applications. *EClinicalMed.* (2019) 11:54–64. doi: 10.1016/j.eclinm.2019.06.003
104. Santermans E, Robesyn E, Ganyani T, Sudre B, Faes C, Quinten C, et al. Spatiotemporal evolution of ebola virus disease at sub-national level during the 2014 West Africa epidemic: model scrutiny and data meagreness. *PLoS ONE*. (2016) 11:147172. doi: 10.1371/journal.pone.0147172
105. Beare S, Simpson E, Gray K, Andjelic D. Rapid integration of zika virus prevention within sexual and reproductive health services and beyond: programmatic lessons from latin america and the caribbean. *Global Health Sci Pract.* (2019) 7:116–27. doi: 10.9745/GHSP-D-18-00356
106. Mathewson AA, Daly ER, Cavallo SJ, Alic A. Use of digital pens for rapid epidemiologic data collection during a foodborne outbreak investigation. *Disaster Med Public Health Prep.* (2015) 9:349–53. doi: 10.1017/dmp.2015.43
107. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. *JAMA*. (2020) 323:1341–2. doi: 10.1001/jama.2020.3151
108. Bassi A, Arfin S, John O, Jha V. *An Overview of Mobile Applications (Apps) to Support the Coronavirus Disease 2019 Response in India*. Mumbai: Indian Journal of Medical Research (2020). Wolters Kluwer Medknow Publications.
109. Official Website of SORMAS. (n.d.). Available online at: https://sormas.org/ (accessed July 9, 2021).
110. Pang J, Wang MX, Ang IYH, Tan SHX, Lewis RF, Chen I-P, et al. Potential rapid diagnostics, vaccine and therapeutics for 2019 novel coronavirus (2019-NCoV): a systematic review. *J Clin Med.* (2020) 9:623. doi: 10.3390/jcm9030623
111. Ncube B, Mars M, Scott RE. The need for a telemedicine strategy for botswana? A scoping review and situational assessment. *BMC Health Serv Res.* (2020) 20:794. doi: 10.1186/s12913-020-05653-0
112. Reforma LG, Duffy C, Collier AY, Wylie BJ, Shainker SA, Golen TH, et al. A multidisciplinary telemedicine model for management of COVID-19 in obstetric patients. *Am J Obstetr Gynecol.* (2020) 2:100180. doi: 10.1016/j.ajogmf.2020.100180
113. Vilendrer S, Patel B, Chadwick W, Hwa M, Asch S, Pageler N, et al. Rapid deployment of inpatient telemedicine in response to cOVID-19 across three health systems. *JAMIA*. (2020) 27:1102–9. doi: 10.1093/jamia/ocaa077
114. Mahmood S, Hasan K, Colder Carras M, Labrique A. Global preparedness against COVID-19: we must leverage the power of digital health. *JMIR Public Health Surveill.* (2020) 6:e18980. doi: 10.2196/preprints.18980
115. Lohiniva AL, Sane J, Sibenberg K, Puumalainen T, Salminen M. Understanding coronavirus disease (COVID-19) risk perceptions among the public to enhance risk communication efforts: a practical approach for outbreaks, Finland, February 2020. *Euro Surveill*. (2020) 25:2000317. doi: 10.2807/1560-7917.ES.2020.25.13.2000317
116. Bayham J, Kuminoff NV, Gunn Q, Fenichel EP. Measured voluntary avoidance behaviour during the 2009 A/H1N1 epidemic. *Proc Royal Soc B Biol Sci.* (2015) 282:1818. doi: 10.1098/rspb.2015.0814
117. Chapman GB, Coups EJ. Emotions and preventive health behavior: worry, regret, influenza vaccination. *Health Psychol.* (2006) 25:82–90. doi: 10.1037/0278-6133.25.1.82
118. Rufai SR, Bunce C. World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis. *J Public Health*. (2020) 42:510–6. doi: 10.1093/pubmed/fdaa049
119. World Health Organization (WHO). *Management of Ill Travellers at Points of Entry - International Airports, Seaports and Ground Crossings - in the Context of COVID-19 Outbreak* (2020). Available online at: https://www.who.int/publications/i/item/10665-331512 (accessed July 9, 2021).
120. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. *Nat Med.* (2020) 26:1183–92. doi: 10.1038/s41591-020-1011-4
121. Can a Coronavirus Tracking App Be Both Effective and Privacy-Centric? | VentureBeat. (2020). Available online at: https://venturebeat.com/2020/04/09/can-a-coronavirus-tracking-app-be-both-effective-and-privacy-centric/ (accessed July 9, 2021).
122. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science.* (2020) 368:6491. doi: 10.1126/science.abb6936
123. Vaughan A. The problems with contact-tracing apps. *New Sci.* (2020) 246:9. doi: 10.1016/S0262-4079(20)30787-9
124. *Coronavirus: An EU Approach for Efficient Contact Tracing Apps to Support Gradual Lifting of Confinement Measures | Shaping Europe's Digital Future* (2020). Available online at: https://ec.europa.eu/digital-single-market/en/news/coronavirus-eu-approach-efficient-contact-tracing-apps-support-gradual-lifting-confinement (accessed July 9, 2021).
125. Chen Q, Min C, Zhang W, Wang G, Ma X, Evans R. Unpacking the black box: how to promote citizen engagement through government social media during the COVID-19 crisis. *Comput Human Behav.* (2020) 110:106380. doi: 10.1016/j.chb.2020.106380
126. Bridgman A, Merkley E, Loewen PJ, Owen T, Ruths D, Teichmann L, et al. The causes and consequences of COVID-19 misperceptions: understanding the role of news and social media. *Harvard Kennedy Sch Misinform Rev*. (2020) 1. doi: 10.37016/mr-2020-028
127. Hasan K, Carras MC, Labrique A. Global preparedness against cOVID-19: we must leverage the power of digital health. *JMIR Public Health Surveill.* (2020) 6:18980. doi: 10.2196/18980
128. Arora VS, McKee M, Stuckler D. Google trends: opportunities and limitations in health and health policy research. *Health Policy.* (2019) 123:338–41. doi: 10.1016/j.healthpol.2019.01.001
129. Qin L, Sun Q, Wang Y, Wu KF, Chen M, Shia BC, et al. Prediction of number of cases of 2019 novel coronavirus (COVID-19) using social media search index. *Int J Environ Res Public Health*. (2020) 17:2365. doi: 10.3390/ijerph17072365
130. Management of Ill Travellers at Points of Entry – International Airports, Seaports and Ground Crossings – in the Context of COVID-19 Outbreak. (n.d.)
131. Akhtar M, Kraemer MUG, Gardner LM. A dynamic neural network model for predicting risk of zika in real time. *BMC Med.* (2019) 17:1–16. doi: 10.1186/s12916-019-1389-3

Author Disclaimer: DN-O, NA-M, and CH are staff members of the WHO. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the WHO.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 24 August 2021
doi: 10.3389/fdgth.2021.654234

Operating an eHealth System for Prehospital and Emergency Health Care Support in Light of Covid-19

Efthyvoulos Kyriacou[1], Zinonas Antoniou[2], George Hadjichristofi[3], Prokopios Fragkos[4], Chris Kronis[4], Theodosis Theodosiou[5] and Riana Constantinou[5]*

[1] Department of Electrical and Computer Engineering and Informatics, Cyprus University of Technology, Limassol, Cyprus, [2] eHealth Lab, Department of Computer Science, University of Cyprus, Nicosia, Cyprus, [3] Department of Computer Science and Engineering, European University Cyprus, Nicosia, Cyprus, [4] eHealth Lab, Department of Electrical and Computer Engineering and Informatics, Frederick University, Limassol, Cyprus, [5] Ambulance Department, State Health Services Organization, Ministry of Health, Nicosia, Cyprus

Introduction: The support of prehospital and emergency call handling and the impact of Covid-19 is discussed throughout this study. The initial purpose was to create an electronic system (eEmergency system) in order to support, improve, and help the procedure of handling emergency calls. This system was expanded to facilitate needed operation changes for Covid-19.

Materials and Methods: An effort to reform the procedures followed for emergency call handling and Ambulance dispatch started on the Island of Cyprus in 2016; along that direction, a central call centre was created. The electronic system presented in this work was designed for this call centre and the new organization of the ambulance services. The main features are the support for ambulance fleet handling, the support for emergency call evaluation and triage procedure, and the improvement of communication between the call centre and the ambulance vehicles. This system started regular operation at the end of 2018. One year later, when Covid-19 period started, we expanded it with the addition of several new features in order to support the handling of patients infected with the new virus.

Results: This system has handled 112,414 cases during the last 25 months out of which 4,254 were Covid-19 cases. These cases include the transfer of patients from their house to the reference hospital, or the transfer of critical patients from the reference hospital to another hospital with an intensive care unit or transfer of patients from one hospital to another one for other reasons, like the number of admissions.

Conclusion: The main purpose of this study was to create an electronic system (eEmergency system) in order to support, improve, and help the procedure of handling emergency calls. The main components and the architecture of this system are outlined in this paper. This system is being successfully used for 25 months and has been a useful tool from the beginning of the pandemic period of Covid-19.

Keywords: eEmergency, emergency medical systems, Covid-19 prehospital patient management, security, emergency dispatch

OPEN ACCESS

Edited by:
Björn Wolfgang Schuller, Imperial College London, United Kingdom

Reviewed by:
Luca Ragazzoni, Università degli Studi del Piemonte Orientale, Italy
Romain Jouffroy, Assistance Publique Hopitaux De Paris, France

****Correspondence:***
Efthyvoulos Kyriacou
efthyvoulos.kyriacou@cut.ac.cy

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *15 January 2021*
Accepted: *29 July 2021*
Published: *24 August 2021*

Citation:
Kyriacou E, Antoniou Z, Hadjichristofi G, Fragkos P, Kronis C, Theodosiou T and Constantinou R (2021) Operating an eHealth System for Prehospital and Emergency Health Care Support in Light of Covid-19. Front. Digit. Health 3:654234. doi: 10.3389/fdgth.2021.654234

INTRODUCTION

The effective and quick handling of an emergency incident can be crucial for a patient's survival and recovery. This handling is directly affected by the work of first responders (paramedics), the effective dispatch procedures of ambulance vehicles, as well as the type and content of data interchanged between an ambulance vehicle, the control center, and the hospital emergency department (1, 2).

Recent advances in electronic health systems, along with technology evolutions and research in computer science can significantly help the aforementioned issues (2, 3). Despite the advances, several logistical problems appeared in emergencies due to failures to coordinate their distribution (1). To address these issues, several emergency dispatch centers around the world are increasingly using several forms of emergency dispatch protocols when handling emergency calls (2, 3). The main objective behind these protocols is to ensure that an incident is appropriately evaluated and responded. In addition, there has been a consistent effort to implement several priority dispatch protocols through computer-based systems in an effort to automate processes and further minimize human error rates (1, 4–8).

The aim of this work is to present the design of Emergency Response protocols and the development and implementation of a secure emergency handling platform that was created to support the Ambulance Department of the Ministry of Health of Cyprus. The Ambulance Department was recently restructured. The main purpose of this work was to reform and optimize the procedures of emergency call handling and Ambulance dispatch. The work started in early 2016 with the creation of a central call center and additional ambulance base stations. The call center initially started working for one region and eventually covered the entire island. All the procedures, starting from emergency dispatch to incident handling, have been organized using specific protocols. Initial work for this system was presented in a conference paper by Kyriacou et al. (9). The appearance of the Coronavirus disease (COVID-19) and the spread of the virus in Cyprus was an extra challenge to the system, which acted as a catalyst for optimizing handling of incidents and the support from high tech solutions. We present the architecture of the system; outline the user functions and system operations from the perspective of usability, efficiency and security; and we discuss the impact of Covid-19 on the operation of the system.

METHODOLOGY

The presented system has been designed and developed according to the needs of the ambulance services so as to support the workflow followed while receiving an emergency call. This includes triage, ambulance dispatch and incident handling from paramedics (3). In extensive user requirements procedure was initially completed before the design and development phase of the system. Despite the initial definition of user requirements, the development team followed an incremental development model that was easily modified especially since there was a small group of easily accessible users. During the development phase of the system, the procedures were modified dynamically to match the reformation of the call center and ambulance services. Initially, the users and use cases were identified as described in section Use Cases.

Our proposed platform securely handles and stores Electronic Health Records. Operations were structured according to the European Union laws and directives (10).

A series of actors are involved in the use case scenarios of our system. Each actor has a specific role and executes different functions within the system. More specifically, system users have seven different types of roles: (1) Administrator, (2) Call center manager, (3) Chief of Ambulances, (4) Dispatcher, (5) Covid-19 Consultant (6) Paramedic, and (7) Doctor. The Administrator is responsible for the overall administration of the system and deals with aspects such as setting access controls and adding removing any of the other type of users. The Call Center Manager is responsible for the overall management of the activities related to the call center. He takes important decisions regarding call center and manages the Dispatchers. The Dispatchers are the personnel of the call center handling the calls and assigning Ambulances to incidents. The Chief of Ambulance handles the functions of the Ambulance unit and the Paramedics within. The paramedics are responsible for driving to the incident during an emergency event and handling the patient. The Doctors offer prehospital health care by going over submitted patient information and guiding the paramedics to treat the patient in transit.

Use Cases

In the approach we followed, we divided the system into two main parts. The first part refers to call handling and ambulance management in the call center, while the second part refers to any activities that take place between prehospital actors after ambulance dispatch.

As shown in **Figure 1A**, part one affects the procedures followed by the call center. The actors involved in this case are the dispatchers (trained paramedics that answer and handle the emergency call), the call center manager, the manager/administrator of the unit (ambulance services), and the Covid-19 consultant. The main use cases of this part are:

- Record a new emergency incident: This is the case when a dispatcher receives a new call. This case starts with initial evaluation of the incident based on the protocols of the center, which include triage and data recording. The availability of an ambulance is found, and the call is transferred (send an ambulance to the incident). Dispatchers are the actors involved in these cases.
- Organize transportation of a patient: Transportation needs include moving a patient from a hospital to another hospital, from their house to a hospital etc. In general, these transportations are not accident and emergency incidents related.
- Cancel emergency or transportation: An emergency or transportation is canceled for some reason and the allocated vehicle is then released.
- Handle Covid-19 cases: This scenario is the new addition to the system to handle Covid-19 incidents. These calls will

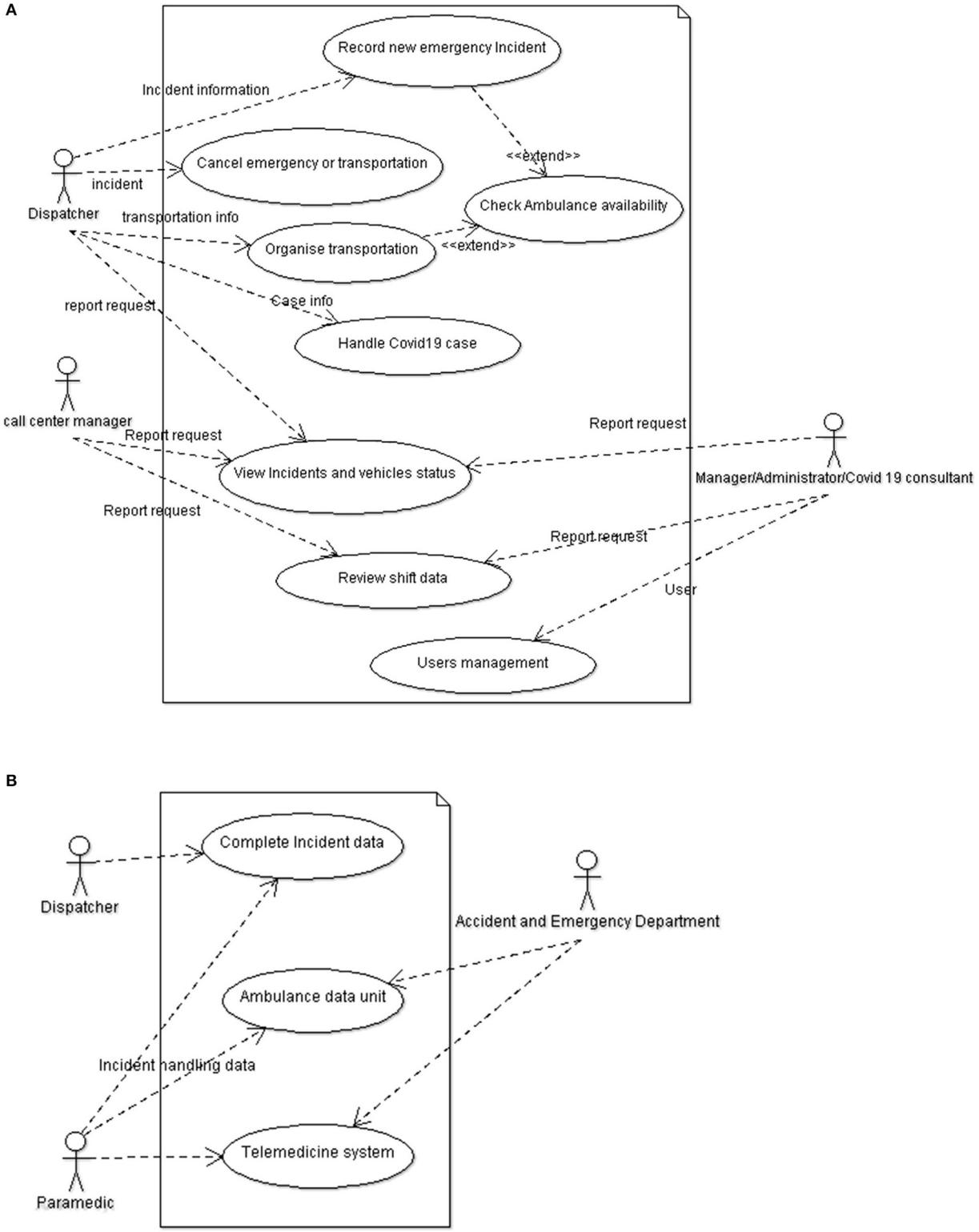


FIGURE 1 | Use case diagram displaying the use cases or functions of users that take place in the system. **(A)** Part one shows the functions of users within the developed system for emergency call handling and ambulance management prior to the dispatchment of the ambulance. **(B)** Part two shows the interactions of the Dispatcher, Paramedic, and Accident and Emergency Department unit that includes doctors after the ambulance is dispatched. Through these interactions data is exchanged between call center (*via* Dispatchers), the ambulance vehicles (*via* Paramedics), and Accident and Emergency Department unit (*via* doctors).

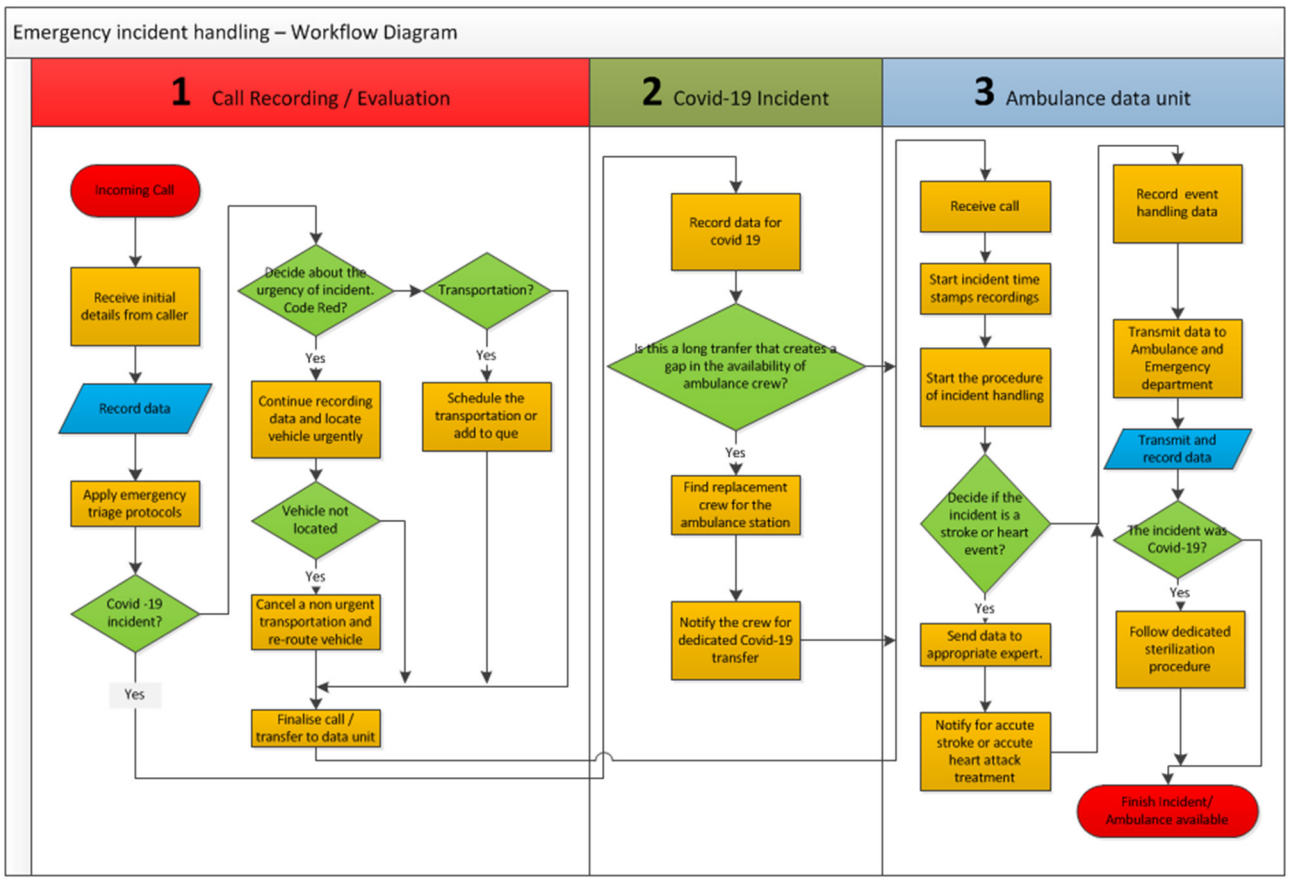


FIGURE 2 | Emergency procedures—workflow management system.

be either emergency calls for transferring a patient to a hospital, or transportation calls for transferring patients from a hospital to another hospital. These calls are handled using a different protocol for staff protective measures, and urgency of transportations compared to the typical cases.

- Reporting cases: The remaining three use cases described in **Figure 1A** (view incidents and vehicle status, review shift data, and user's management) are managerial cases where the call center manager and the manager of the ambulance services are involved.

Part two (see **Figure 1B**) affects the procedure of data exchange. The actors involved in this case are the emergency call center, the ambulance system, and the accidents and emergency department. The main use cases are:

- Real-time monitoring of the patients involved in an incident scene, through a portable telemedicine unit that transmits bio-signals. This allows on-call doctors at the First aid control center to make decisions a-priori for in hospital patient handling (e.g., support stroke or cardiac clinic decision for patients needing emergency angioplasty).
- Get initial call data from the call center and complete any data related to the transportation and handling of the patient by the ambulance paramedics.
- Get data related to the handling of Covid-19 patients and provide any needed guidance and support for patient handling.

Emergency Call Workflow Description

Emergency calls are handled based on a specific set of steps. These steps have been edited to incorporate the actions of a Covid-19 consultant. The workflow of actions that take place during an emergency call scenario is described as follows (see **Figure 2**):

- An emergency call starts at the call center.
- The dispatcher receives input from the caller and follows the triage protocol in order to assess the incident.
- The dispatcher records initial information about the incident, including Covid-19 related information, and if possible record basic information about the patients involved. If the patient is a possible Covid-19 or a verified case, the paramedics are informed in order to take the appropriate protective measures and follow the designated protocols for Covid-19.
- The dispatcher chooses an available ambulance and paramedics and dispatches them to the location of the incident.
- Information recorded is transmitted to the Ambulance Data Unit.

- Following the emergency dispatch procedure, paramedics arrive at the incident location, assess the patient's condition, and inform the Ambulance Dispatching Unit about the status of the patient/incident.
- If needed, they also transmit bio-signal information to the hospital so that the doctor on call can view, assess the severity of the situation, and guide the paramedics accordingly. Once the patient is transferred to the hospital, the data regarding the incident are also transferred to the data station of the accidents and emergency department.
- Soon after the delivery, the paramedics have to take all the necessary procedures for cleaning and sterilization of the vehicle. Then they inform the call center that the ambulance is available to service other requests. If the ambulance is carrying a Covid-19 patient, it is sterilized based on a Covid-19 specific protocol and then released to other service requests.

System Architecture

The system architecture is mainly divided into two major components. The first component is the system responsible for managing the incident call, ambulance vehicles, and initial information recorded, it covers part 1 of the user requirements. The second component covers the ambulance vehicle part which includes data exchange processes between ambulance vehicles, call center and hospitals. This refers to part 2 of the user requirements. This gives the ability to exchange data like incident initial description, incident handling procedure, and bio-signals of a patient, between the ambulance vehicles, the call center and the referring hospital/doctor.

Based on the aforementioned operations and user requirements we designed the system as shown in **Figure 3**. The procedures followed by the ambulance department are based on specific protocols that can differ for each case. These protocols were created by the Ambulance Department according to internationally accepted protocols and standards published from associations like NAEMT- National Association of Emergency Medical Technicians—https://www.naemt.org/. All these protocols are supported by the system. The detailed workflow of incident handling is depicted in **Figure 2** and described in section Emergency call Workflow description.

The recorded data can be made available to the accident and emergency department of a main hospital (the place where the ambulance transfers the patient), to the call center and to the ambulance vehicle. The ambulance vehicles have the technology to exchange data and vital bio signals from the scene and during transportation of a patient (11) using 4G wireless technology.

As can be seen in **Figure 3**, the ambulance vehicle is equipped with an Ambulance Data Unit and an Emergency Telemedicine Unit. These units are the evolution of the work presented in (11). The Emergency Telemedicine Unit is responsible for sending bio-signals (ECG 3-7 leads, spO_2, blood pressure, temperature, and respiration). A doctor on call can then use the Doctor's Mobile Unit (e.g., a tablet or pc), authenticate and get authorized into the server, and monitor a patient's condition.

The Ambulance Data Units are Android tablets equipped with 4G communication cards. The paramedics use the tablet (or pc) to get initial information about the incident (recorded during the emergency and the triage procedure from the call center staff) and concurrently report information regarding the incident. These include information about incident handling and procedures followed by the ambulance paramedics. Additionally, the information can include patient, or incident info. All this information together with information related to device and user authentication are hosted on the servers that act as a backbone of the system as shown in **Figure 3**.

The communication between the mobile units, the call center and the hospitals is feasible through TCP/IP protocols over mobile 4G wireless networks. Data from ambulances to the server are exchanged using web services. Access control to the system is enabled using a unique ID for each unit, Secure Socket Layer (SSL) for communication, Virtual Private Network (VPN) infrastructure through a firewall, and different access levels depending on the user group (see section Security Design).

System data storage is supported by a dedicated database server. The entities include system users, where each group of users has different access levels. The database system has the ambulance entity, where each vehicle and its status are recorded, the patient entity which stores general information about a patient, the incident entity, which stores any information regarding the incident (medication, handling procedure, initial diagnosis, final diagnosis, etc.), the bio-signals entity, which stores any real time bio-signals transmitted during an incident, and the logs entity which stores the activity done and the user that did it.

The main aspects of system architecture are described in the following section Security Design up to Ambulance Data Units.

Security Design

Based on the scenarios described, the system supports users having seven different types of roles: (1) Administrator, (2) Call center manager, (3) Chief of Ambulances, (4) Dispatcher, (5) Covid-19 Consultant, (6) Paramedic, and (7) Doctor. Each role requires different access controls.

Access control takes into account the user identity, as well as the identity of the connected devices. The Administrator has the highest authority in the system. He is responsible for registering and authorizing the Chief of Ambulances (shown as call center manager in **Figure 1A**), the Shift Managers, and the Dispatchers. They are each given a unique username and password and are authorized to handle various operations relevant to the Ambulance Department in which they belong. This authentication material is managed offline by the system administrator.

Doctors using the Doctor Mobile Unit are considered as external entities. Authentication is executed both on the user's ID and the device's ID connected to the system. This setup enables multiple users to use the same Doctor Mobile Units. Further, a Doctor Mobile Unit is tied to a specific location and thus once this device is authenticated it allows for authorization of specific operations within the system. Following an incident, the authentication of Doctors' ID on the Doctor Mobile Unit

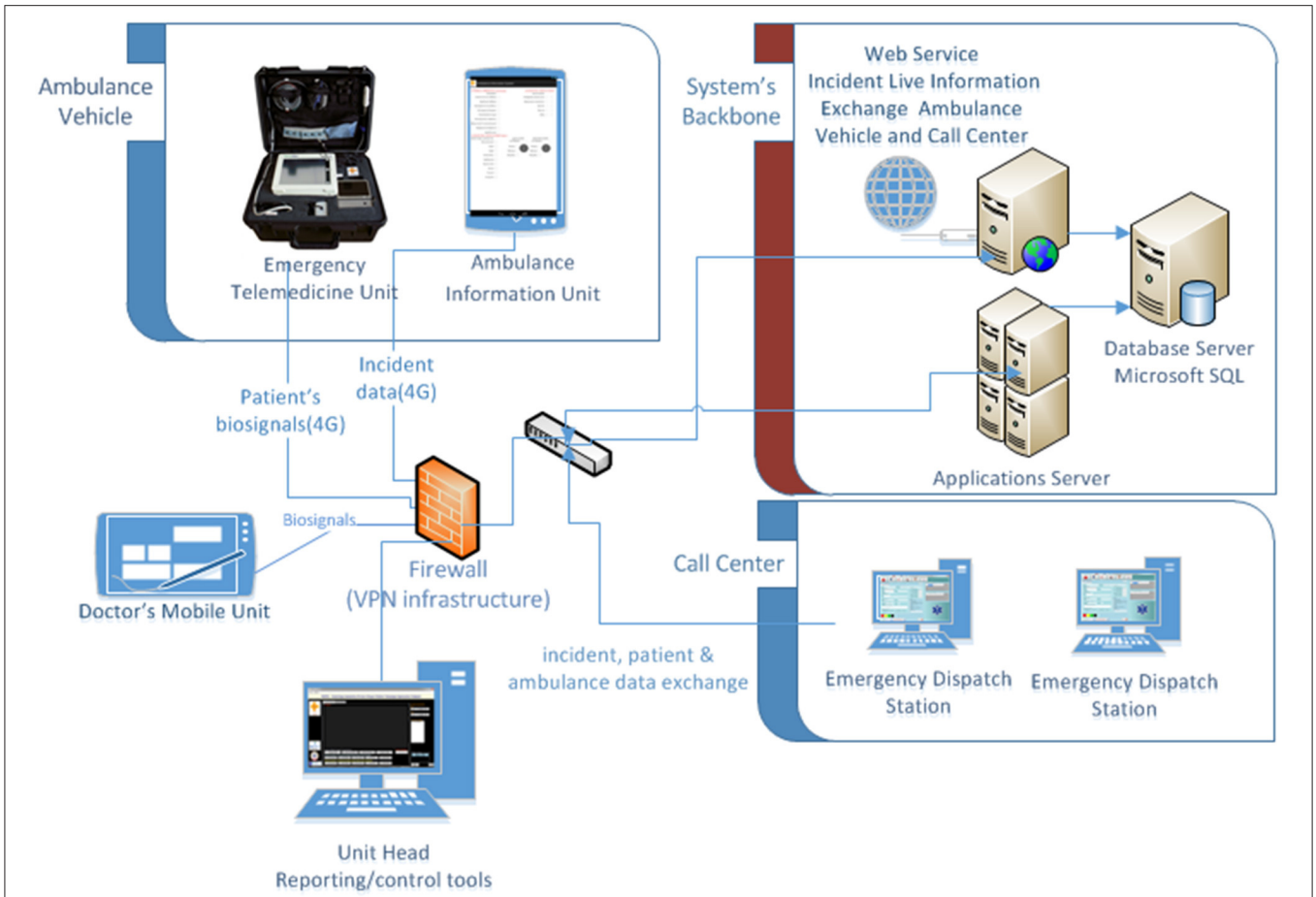


FIGURE 3 | eEmergency health care system architecture.

authorizes them to get access only to their assigned incidents. Doctor Mobile Units can be Android devices, which have a unique ID.

The Ambulance Data Units used by paramedics utilize double authentication. Thus, during an emergency both the user and the unit are identified and authorized for access. Introducing per user authentication in addition to per unit authentication, prevents someone stealing the device from altering information regarding an incident or submitting wrong information. Currently, there are no physical security controls to prevent anyone from stealing the Ambulance Information Unit. One such control that can be introduced in the future is the installation of a small safe deposit box with a digital key within the ambulances.

Each ambulance unit is assigned a specific Android device and is used by different groups of paramedics depending on the scheduled shifts. Within each group, an individual is granted access and assigned the responsibility for data entry. In this manner, the system holds accountable a specific individual that worked during a specific shift for filling up incident-related information. A strength of this setup is that only one individual needs to be trained to handle the device. A weakness, however, is that in case that individual needs to be replaced or is absent, it creates complexities in terms of finding a trustworthy substitution.

Duration of access to the system is controlled through session keys. Session keys are dynamically generated using hash functions after a user gets authenticated into the system. The validity period of a session key depends on the length of time that the user needs to access the system and the criticality of the operation. Frequent refreshing would imply a more critical operation.

A summary of the various access levels in the system per user type is offered in **Table 1**. The Administrator, the Chief of the Ambulance Services and the Shift Manager have the most important role in the system. Separation of duties is utilized, and different tasks are assigned to these two users. The Administrator is responsible for overseeing that the network operations are being carried out effectively and the system is always available. He/she registers any type of users including the Chief of Ambulance Services, the Shift Manager, the Covid-19 Consultant, Dispatchers, Doctors, and Paramedics. However, the Chief of Ambulance and the Shift Managers have the general responsibility of overseeing the correct operation of the Ambulance Department. They can also add or remove new

TABLE 1 | Role based access control to the different modules of the system.

	Role Roles of participants in the emergency unit						
Database Entities	Administrator	Shift manager	Chief of ambulances	Dispatcher	Covid-19 Consultant	Paramedic	Doctor
User	R/W(A,L,M)	R/W(A,L,M)***	R/W(A,L,M)**	-	-	-	-
Ambulance	-	R/W(A,L,M)	R/W(A,L,M)	W(M)****	R/W(Ap)	W(M)****	-
Patient	-	R	R	R/W(A,M)	R/W(Ap)	R/W(A,M)	R/W(Ap)
Incident	-	R	R	R/W(A,M)	R/W(Ap)	R/W(M)	R*
Biomedical signals	-	-	-	-	-	W(A)	R
Logs	R	R	R	-	R/Ap		-

R, read; W, write; -, no permission.
W(A, add; L, Lock; M, Modify; Ap, append).
**For Incident tied to patient.*
***For adding Dispatchers and Shift Managers only.*
****For adding Dispatchers only.*
***** Modify the status of the Ambulance (available, on Call, In Service).*
Modify implies Append, unless Append is stated instead.

ambulances joining the system. Similar to the Administrator, they may add dispatchers, for the scenario where the Administrator is unavailable. In addition, the Chief of Ambulance is the only one who can add or remove Shift Managers. Dispatchers can create a new incident, update an existing incident, but cannot delete any incident for any reason. They can view the ambulance entity and alter the status of an ambulance based on the emergency schedule. The status of the ambulance could be "On Call," "Available," or "In Service." Dispatchers can add or modify the records of new patients and link the patient to a newly recorded incident. Furthermore, the dispatchers should be able to view/edit all incidents recorded by their co-workers at the Ambulance Emergency Call Center. The dispatchers should also be able to see information that is being sent by Ambulance crew at real time regarding the incident status and alter the status of an ambulance in the current shift.

The role of the Covid-19 consultant is to provide guidance to the Ambulance Center Personnel regarding all aspects of its operations as impacted by Covid-19. Thus, the access control of the consultant is restricted to mainly append on records of ambulances, patients, incidents, and logs. This access enables consultants to introduce comments, such as the need for certain personnel to get tested for Covid-19, or the need for certain procedures to be altered or followed. Furthermore, it allows them to keep track of all the incidents handled and provide recommendations on how future incidents need to be handled.

Doctors can only read and append to patients' records. They also have read access for the incidents in which their patient was involved. Paramedics can add a patient to the system or modify a patient's records. They can also update the status of the patient's incident.

A logging mechanism is used throughout the entire system to track the events so as to provide accountability. Every instance of our database has a log file associated with it, which is accessible for reading only by the Chief of Ambulance, the Covid-19 consultant and the Administrator. Log files include incident_log, ambulance_log, patient_log, and logs capturing the activity of users, such as the doctor_log, the paramedic_log, and the dispatcher_log. Typically, the Chief of Ambulance will look at the logs to investigate issues regarding the operation of the Ambulance Unit, such as the effectiveness of an operation. On the other hand, the Administrator may look at the logs to investigate any malicious or erroneous behavior of the users in the system. In light of Covid-19 the Covid-19 consultant uses the log to keep track of Coronavirus incidents and assess the effectiveness of handling events. Items, which might be of interest are the number of Coronavirus patients per location, the severity of the various cases, and the methodology of treatment. In order to promote safety, the personnel and ambulances involved per Covid-19 incidents is also accounted for. Paramedics handling Covid-19 incidents are frequently checked for a possible infection and where possible are prevented from participating in non-corona virus related incidents. Thus, a separation of responsibilities is established based on the type and severity of incident handled. Lastly, the sanitization schedule of the ambulances is set up based on the number and frequency of incidents.

Note that nobody in the system has access to delete any records. An option would be to start deleting records after a certain number of years. However, that period would depend on the policies adopted by the Ambulance Department.

In addition to access control, another important security property is privacy. Privacy is tied to the confidentiality of data, which is achieved through the encryption of transmitted data. All external entities connecting to the system, such as the ambulance data units, utilize the FortiClient VPN by Fortinet (12) with SSL VPN activated to provide privacy. Confidentiality is even more critical with Covid-19, as it is important to protect the Identity of infected people and their location. Note that encryption is not used in the Communication between the bio-signal server and the Emergency Telemedicine Unit. The induced encryption overhead would be operation prohibitive.

Data Exchange

One of biggest challenges in this system was the accurate exchange of medical data and availability of data to all the parties involved (call center, ambulance vehicles, reference hospital/doctor). Toward this direction, there was a great need to enable the seamless tablet devices' communication with the emergency handling platform by implementing the System Communication module as shown in **Figure 3**. This facilitates real-time monitoring of the patients involved in an incident scene during his/her transportation from the incident scene to the hospital. Thus, it allows the Emergency Department healthcare providers' to be well-prepared for the treatment of the patient as soon as he/she arrives at the hospital premises.

The Emergency handling platform was installed in two-tier client-server architecture in which the functional process logic, data access, user interface, and computer data storage were developed and maintained as independent modules on two separate virtual servers (see **Figure 3**). Both virtual servers are located on a physical server located at the Ambulance Department premises. This is installed in a dedicated server room with all the appropriate security measures, data backup devices, and power failure support devices.

The first tier corresponds to the Application and Presentation Layer. This layer uses an application server that contains the functional business logic of the system, which drives an application's core capabilities and the user interface. The communication with the specific Android application that is installed on a tablet device is enabled through Application Programming Interface (API) calls using Representational State Transfer (REST) API. API is a set of clearly defined methods of communication among various components. It is a predefined set of rules that allow programs to talk to each other. We created the APIs on the server and verified that the client android application can talk to it. REST determines how an API looks like. REST is a set of rules that we followed when we were creating our APIs. One of these rules states that the android application is able to get a piece of data (called a resource) when it links to a specific URL. Each URL is called "a request" that triggers an operation on the server while the data sent back to the android application client is called "a response."

The second tier corresponds to the Data Layer. This specific layer handles a Microsoft SQL database management system that provides secured access to application data. Data is communicated through SQL stored procedures. A stored procedure is a prepared SQL query that is stored on the SQL server so that it can be reused repeatedly by just calling and executing the stored procedure. Stored procedures enhance security of the system and minimize injections by ensuring that operations being performed are allowed by the user. That means that *ad hoc* queries and data modifications can be disallowed. This prevents users from maliciously or inadvertently destroying data or executing queries that impair performance on the server or the network. Furthermore, stored procedures can reduce network traffic by combining multiple operations into one procedure call.

The system supports auditing mechanism by performing security audits on every API call using audit trails and logs that offer a back-end view of system usage. Audit trails and logs record key activities, showing system threads of access, modifications, and transactions.

Data is accessed on the Ambulance Data Units (Android app)via secure API calls using in JavaScript Object Notation (JSON) format that are implemented in ASP.NET user-interface (view), data (model), and application logic (controller) (MVC) environment. Once a paramedic user opens the android application, he/she must provide valid credentials in order to be authorized and authenticated by the system to access data. The System administrator is responsible to create and provide valid credentials to paramedics.

API calls support the following functions so as to facilitate the communication with the app: (a) login, (b) get paramedic information, (c) get active ambulance call, (d) insert information regarding the patient(s) of an active call, (e) get ambulance calls that were performed during the paramedic's shift, (f) update call information, and (g) send SOS alert. A related data exchange flow that depicts interaction between the user and the two-tier client-server architecture using REST APIs can be seen in the corresponding swim lane diagram in **Figure 4**.

The system allows a paramedic to update information about an ambulance call that was performed the last 6/12 h, based on the corresponding shift duration, to update any information or add missing information. Due to legal issues that might be raised, ambulance response datetime and ambulance station departure datetime fields, once recorded they cannot be updated. Moreover, the system allows a paramedic to send an SOS alert to the Ambulance call center directly from the app to call for help (this can be used in cases where the paramedics face a thread from a patient or other people located at the incident scene).

Ambulance Data Units

Ambulance data units are assigned specific Android devices (Lenovo Tab E10) which are designated to ambulance vehicles depending on the scheduled shifts. Each Android device functions as an intermediary between the paramedics/ambulance and the server responsible for data exchange. The application running on each device is designed to record and update information of incidents recorded for the first 6/12 h. Once 6/12 h have passed since the creation of the incident, the information gets hidden on the device, and thus no modification is allowed. Tablet access is restricted only on the application for data exchange, this secures any unwanted access to internet data or applications which can cause serious problems to our system.

As depicted in **Figure 2** in the Ambulance Data Unit (section Results and Discussion), the application is divided into various activities, which indicate when the fundamental transactions occur between the user and the system. Once a connection to the server has been established, the user is prompted to enter his/hers credentials. The system will then authenticate the user.

Once authenticated, the user is presented with different information according to their role (generic or not generic). Generic users in the system exist for administrative use only, whereas non-generic users are regular users, which are disparate

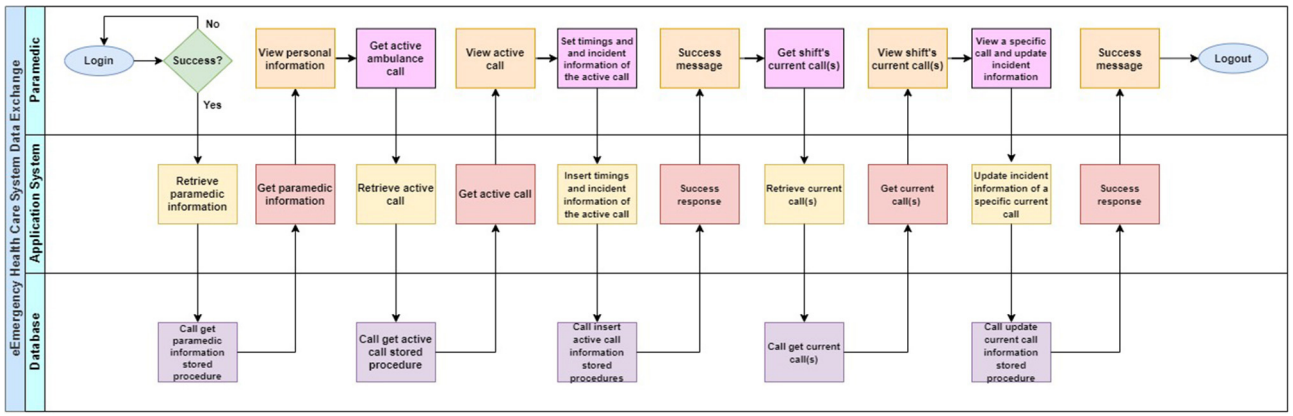


FIGURE 4 | eEmergency health care system data exchange swim lane diagram.

entities within the system. Before accessing incident data, non-generic users will be prompted to enter their assigned station and shift period. For generic users, their ID and ambulance number is also required.

Incident data is represented in the form of a table and is retrieved every time the user visits the incident selection screen. Based on the date of creation, ambulance station name and patients, users are able to recognize the incident of their choice. Once an incident is selected, the user can navigate the menu responsible for handling the records for each patient (see **Figure 5**). Each incident is composed of patient incident records, which can either be created or updated by the users. Through the ambulance data unit, paramedics are able to record all data regarding the incident. Screens used for recording of trauma, drug administration and eye response evaluation are shown in **Figure 5**. Deletion is possible only for records, which have been stored statically/temporarily in the application, and thus do not coincide with the contents of the database. If a patient record has already been stored within the incident, it can be deleted temporarily. The deleted data can be retrieved upon reentering the selected incident.

Each patient incident record contains sensitive data, which is sent in the form of a status report to a previously designated email address. Users may also choose to send the report to more email addresses, each corresponding to an ambulance unit. Completing such information requires the utmost attention from the users.

RESULTS AND DISCUSSION

The presented system has been completed and has been in daily practice routine for about 25 months. The system was adopted for daily practice after using it for 5 months as a test phase. During these 5 months, the system was thoroughly tested, and any errors were corrected. Additionally any minor additions requested by the users were added.

Each incident in the system is tagged with various priority levels, which are denoted with color codes. The color codes are Red (Immediate response extremely urgent case), Yellow (case with serious problems personnel can response up to 20 min), Purple (denotes transportation of patient. This can be a prescheduled transportation or a transportation from one hospital to another), Green (case which can wait up to 3 h) and black (meaning the person is dead). The number of incidents per priority level can be seen in **Table 2**.

The test phase 1 was between 1st December 2018 until 30th April 2019. During this phase, the system recorded 23,386 cases out of which 6,739 cases (29% of total cases) were code red cases as shown in **Table 2**. Soon after the test phase 1 the system entered the regular use in daily practice. This is shown as phase 2. It started from May 1st 2019 until the 29th of February 2020. This new timestamp was added when Covid-19 incidents started appearing in Cyprus. The number of incidents recorded during phase 2 from the system were 46,574 out of which 14,584 were code red (31% of total cases).

Soon after covid-19 cases started appearing in Cyprus (around the beginning of March 2020). The system was dynamically adopted in order to cover the new cases. This period started from March 1st 2020 until now January 14th 2021 (when this paper was submitted).The pandemic Covid-19 period is also divided into two periods. The first wave of the pandemic starting from March 1st 2020 until July 31st 2020 and the second wave starting from August 1st 2020 until the 14th of January 2020.

Comparing the different periods of observations it can be seen that the percentages of priority levels of each phase are similar except from the red and yellow codes. Red was reduced while yellow was increased during the covid-19 and the lockdown periods that took place. This is something observed through systems of other countries also as shown in (13). Red incidents still have a high percentage because of incidents like stroke and heart attacks also shown in (14). Trauma cases (red code) were reduced which explains the small reduction of this group. The increase in yellow codes was expected as covid-19 calls increased (these calls need special attention but are not at the urgency of level red).

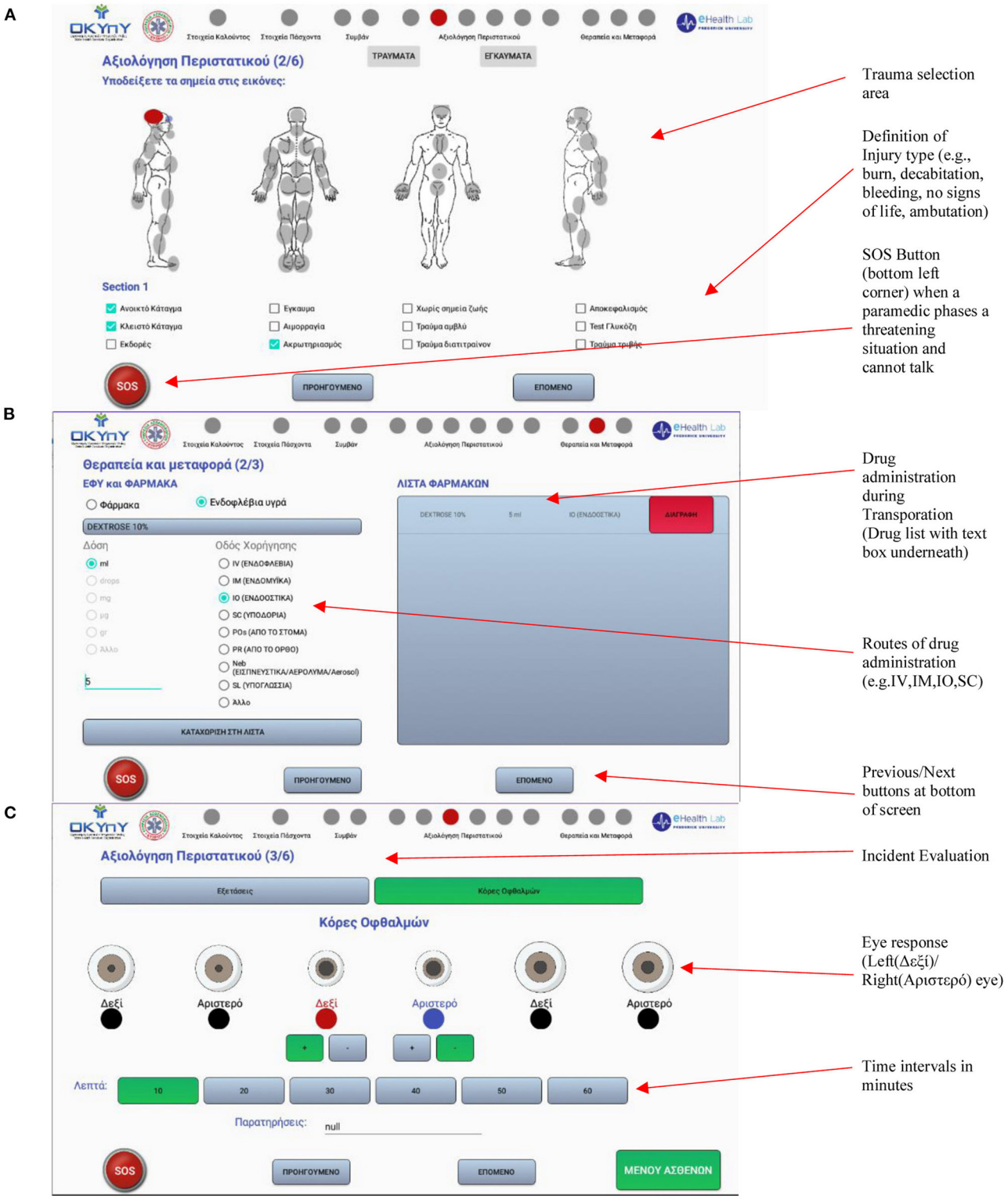


FIGURE 5 | Some of the key screenshots of the GUI used to record incident data **(A)**. Incident Evaluation for traumas **(B)**. Medicine administered during transportation **(C)**. Incident evaluation based on for eye response.

TABLE 2 | Total number of recorded incidents from 1st December 2018 until 14 January 2021.

Priority level	Phase 1 (test)	%Phase 1	Phase 2	%Phase 2	Phase 3 (Covid-19 first wave)	%Phase 3	Phase 4 (Covid-19 second wave)	%Phase 4	Total
Red	6,739	29%	14,584	31%	5,086	27%	5,423	23%	31,832
Yellow	9,739	42%	19,713	42%	8,561	45%	10,980	47%	48,993
Purple	6,237	27%	11,169	24%	4,923	26%	6,281	27%	28,610
Green	481	2%	771	2%	400	2%	449	2%	2,101
Black	190	1%	337	1%	179	1%	172	1%	878
Total	**23,386**		**46,574**		**19,149**		**23,305**		**112,414**

The division of the incidents is based on priority levels. Bold values emphasize the total values.

TABLE 3 | Recorded incidents during Covid-19 period, starting from March 1st 2020 until January 14th 2021.

	Covid-19 first wave						Covid-19 second wave					
	March 2020	April 2020	May 2020	June 2020	July 2020	August 2020	September 2020	October 2020	November 2020	December 2020	January 2020	Total
Covid-19 cases	230	613	100	53	88	403	256	2,474	6166	11,636	6,105	28,124
Emergency calls	3,739	3,174	4,061	3,823	4,352	4,202	4,285	3,990	3,882	4,767	2,179	42,454
Covid-19 calls	312	294	201	72	89	192	153	552	705	1,317	582	4,469
Percentage of Covid-19	8.3%	9.3%	4.9%	1.9%	2.0%	4.6%	3.6%	13.8%	18.2%	27.6%	26.7%	10.5%

Since March 2020, when the pandemic period started in Cyprus, 4,152 emergency incidents regarding Covid-19 cases have been recorded. The Covid-19 cases have been continuously increasing because of the increased incidents recorded during the second wave of the pandemic, which proved to be much tougher to handle than the first wave. Based on the statistics published by the government of Cyprus available on the page of the World Health Organization (WHO) (15), the total cases reported are 28,124 in total until January 14th 2021. In detail, the total cases compared to the total emergency incidents and covid-19 emergency calls are shown in **Table 3**.

Handling of covid-19 cases created the need for communication and synchronization of the Ambulance Department database and the Cyprus' National Contact Point eHealth (NCPeH) database. NCPeH enables seamless cross-border care and secure access to patient health information between European healthcare systems (16). This will correlate the Patient Summary record of a given patient to be updated automatically when a patient that is a verified Covid-19 incident will be transferred by an ambulance vehicle. Additionally, the Patient Summary will be updated accordingly to include all the corresponding data of that patient, including bio-signals, current medication, active problems and other medical information that will be recorded by the paramedic using the Ambulance data units.

CONCLUSION

The presented system has been successfully used for around 25 months without any major technical problems. One of the main challenges of the system was to follow the restructuring of operation procedures of the Ambulance Department. The restructure caused a lot of changes such as the creation of a central call center, the creation of extra ambulance stations, introduction of newly established protocols and others.

The design and implementation of the new digital system was a challenge for our group. We had to follow the restructure procedure and adapt according to the changes. The general aim was to support the Ambulance Department of the Ministry of Health in Cyprus. The key contribution of this work is the design of an integrated system that supports emergencies by allowing the dynamic assignment of paramedics and ambulances to incidents during an emergency and monitoring the procedure from the moment a call is received until a patient is delivered to a hospital. The main task is to minimize mistakes and create a more effective emergency handling system.

Appearance of Covid-19 and the start of the pandemic period was another challenge that we had to phase with our ehealth system. The stress on the emergency health care system was high and all these procedures needed accurate and immediate data exchange. This was also supported with success.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because they contain classified medical data. Requests to access the datasets should be directed to efthyvoulos.kyriacou@cut.ac.cy.

AUTHOR CONTRIBUTIONS

EK: system concept, system development, and project management. ZA: system restructure, web service development for ambulance, and mobile devices communication. GH: system security. PF: ambulance data unit and communication module development. CK: call centre core system development. RC: system concept, system user requirements, and system evaluation. TT: system user requirements and system evaluation. All authors contributed to the article and approved the submitted version.

FUNDING

This personnel cost was self-funded by internal funds of the participating partners. Government of Cyprus covered equipment. Additional Covid-19 equipment was provided as donation from several bodies and individuals.

REFERENCES

1. Kyriacou E, Panayides A, Constantinides P. m-Health e-emergency systems. In: Eren H, Webster G, editor. *The Handbook of Electronic Medicine, Electronic Heath, Telemedicine, Telehealth and Mobile Health.* Limassol: CRC press (2015).
2. Bohm K, Kurland L. The accuracy of medical dispatch - a systematic review. *Scand J Trauma Resusc Emerg Med.* (2018) 26:94. doi: 10.1186/s13049-018-0528-8
3. Ageron F, Debaty G, Gayet-Ageron A, Belle L, Gaillard A, Monnet M-F, et al. Impact of an emergency medical dispatch system on survival from out-of-hospital cardiac arrest: a population-based study. *Scand J Trauma Resusc Emerg Med.* (2016) 24:53. doi: 10.1186/s13049-016-0247-y
4. Deakin CD, Sherwood DM, Smith A, Cassidy M. Does telephone triage of emergency (999) calls using advanced medical priority dispatch (AMPDS) with Department of Health (DH) call prioritisation effectively identify patients with an acute coronary syndrome?", NHS Trust. *Emerg Med J.* (2006) 23:232-5. doi: 10.1136/emj.2004.022962
5. Clawson J, Olala CHO, Heward A, Scott G, Patterson B. Accuracy of emergency medical dispatchers' subjective ability to identify when higher dispatch levels are warranted over a Medical Priority Dispatch System automated protocol's recommended coding based on paramedic outcome data. *Emerg Med J.* (2007) 24:560-3. doi: 10.1136/emj.2007.047928
6. Bürger A, Wnent J, Bohn A, Jantzen T, Brenner S, Lefering R, et al. The effect of ambulance response time on survival following out-of-hospital cardiac arrest. *Dtsch Arztebl Int.* (2018) 115:541-8. doi: 10.3238/arztebl.2018.0541
7. Hsia RY, Huang D, Mann NC, Colwell C, Mercer MP, Dai M, et al. A US national study of the association between income and ambulance response time in cardiac arrest. *JAMA Netw Open.* (2018) 1:e185202. doi: 10.1001/jamanetworkopen.2018.5202
8. Constantinides P, Barret M. A narrative approach to understanding coordination practices in emergency response. *Inform Organ.* (2012) 22:273–94. doi: 10.1016/j.infoandorg.2012.07.001
9. Kyriacou E, Constantinou R, Kronis C, Hadjichristofi G, Pattichis C. eEmergency system to support emergency call evaluation and ambulance dispatch procedures. *In: IEEE 20th Mediterranean Electrotechnical Conference (MELECON).* (2020). p. 354-7.
10. Overview of the National Laws on Electronic Health Records in the EU Member States and Their Interaction With the Provision of Cross-Border eHealth Services. Milieu, Time.lex, EU report under EU Health Programme(2014).
11. Kyriacou E, Pavlopoulos S, Berler A, Neophytou M, Bourka A, Georgoulas A, et al. Multi-purpose healthcare telemedicine systems with mobile communication link support. *BioMed Eng Online.* (2003) 2:7. doi: 10.1186/1475-925X-2-7
12. FortiClient. *Forticlient VPN.* Available online at: https://www.forticlient.com/ (accessed January 15, 2021).
13. Weiner S, Cash R, Hendricks M, El Ibrahimi S, Baker O, Seethala R. Ambulance calls for substance-related issues before and after COVID-19. *Prehosp Emerg Care.* (2020). doi: 10.1080/10903127.2020.1845420. [Epub ahead of print].
14. Holmes JL, Brake S, Docherty M, Lilford R. Emergency ambulance services for heart attack and stroke during UK's COVID-19 lockdown. *Lancet.* (2020) 395:E93-4. doi: 10.1016/S0140-6736(20)31031-X
15. World Health Organization. *Covid 19 Dashboard.* (2021). Avaialble online at: https://covid19.who.int/. (accessed January 15, 2021).
16. Antoniou Z, Constantinnou I, Neofytou M, Neokleous K, Schiza E, Panayides A, et al. *Deployment of Generic Cross Border eHealth Services in Cyprus.* IEEE Communication Society eHealth Technical Committee Newsletter (2017).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Digital Health

PERSPECTIVE
published: 20 September 2021
doi: 10.3389/fdgth.2021.639827

Factors to Consider in the Use of Vital Signs Wearables to Minimize Contact With Stable COVID-19 Patients: Experience of Its Implementation During the Pandemic

Esther Monica Pei Jin Fan[1], Shin Yuh Ang[1], Ghee Chee Phua[2], Lee Chen Ee[3], Kok Cheong Wong[4], Franklin Chee Ping Tan[5], Lydia Wan Har Tan[6], Tracy Carol Ayre[1], Chee Yong Chua[7], Benedict Wee Bor Tan[8] and Khung Keong Yeo[9,10]*

[1] Nursing Division, Singapore General Hospital, Singapore, Singapore, [2] Department of Respiratory and Critical Care Medicine, Singapore General Hospital, Singapore, Singapore, [3] Organisational Transformation, SingHealth, Singapore, Singapore, [4] Nursing Division, Changi General Hospital, Singapore, Singapore, [5] Office for Service Transformation, SingHealth, Singapore, Singapore, [6] Office of Innovation, Changi General Hospital, Singapore, Singapore, [7] Emerging Services and Capabilities Group, Integrated Health Information Systems, Singapore, Singapore, [8] Division of Digital Strategy, SingHealth, Singapore, Singapore, [9] Department of Cardiology, National Heart Centre Singapore, Singapore, Singapore, [10] Duke-National University of Singapore Medical School, Singapore, Singapore

OPEN ACCESS

Edited by:
Pradeep Nair, Central University of Himachal Pradesh, India

Reviewed by:
Milena B. Cukic, Amsterdam Health and Technology Institute (AHTI), Netherlands
Akihiro Nomura, Kanazawa University, Japan

***Correspondence:**
Esther Monica Pei Jin Fan
esther.monica.fan.p.j@sgh.com.sg

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *10 December 2020*
Accepted: *16 August 2021*
Published: *20 September 2021*

Citation:
Fan EMPJ, Ang SY, Phua GC, Chen Ee L, Wong KC, Tan FCP, Tan LWH, Ayre TC, Chua CY, Tan BWB and Yeo KK (2021) Factors to Consider in the Use of Vital Signs Wearables to Minimize Contact With Stable COVID-19 Patients: Experience of Its Implementation During the Pandemic. Front. Digit. Health 3:639827. doi: 10.3389/fdgth.2021.639827

The COVID-19 pandemic has created a huge burden on the healthcare industry worldwide. Pressures to increase the isolation healthcare facility to cope with the growing number of patients led to an exploration of the use of wearables for vital signs monitoring among stable COVID-19 patients. Vital signs wearables were chosen for use in our facility with the purpose of reducing patient contact and preserving personal protective equipment. The process of deciding on the wearable solution as well as the implementation of the solution brought much insight to the team. This paper presents an overview of factors to consider in implementing a vital signs wearable solution. This includes considerations before deciding on whether or not to use a wearable device, followed by key criteria of the solution to assess. With the use of wearables rising in popularity, this serves as a guide for others who may want to implement it in their institutions.

Keywords: COVID-19, vital signs wearables, vital signs monitoring, digital health, digital solution

INTRODUCTION

A pneumonia of unknown cause was detected in Wuhan, and was first reported to World Health Organization (WHO) on 31st December (1). The disease spread quickly and was soon characterized by WHO as a pandemic on 11th March 2020 (2). The first case of COVID-19 infection in Singapore was detected on 23 January 2020. By 18 November, the number of cases has risen to 58,135, with 28 fatalities (3). Consequently, the healthcare industry met with various challenges. The need for healthcare facilities and healthcare workers (HCWs) rose rapidly (4). Demand for equipment, personal protective equipment (PPE), medications, and consumables rose so quickly that supply chains struggled to meet them.

With rising cases of COVID-19, facilities were converted/created to care for them (4). There was a need to monitor more patients with less HCWs while preserving

PPE. It was also necessary to ensure that healthcare remains cost effective. Additionally, easing the strain on HCWs to avoid burnout was a major consideration given the extended duration of this pandemic. Exploration of the use of vital signs wearables, which begun a few years ago, was accelerated during this period in attempt to meet these needs.

This paper presents an overview of factors to consider in implementing a vital signs wearable solution during an infectious disease outbreak. In the age where the use of wearables is expected to rise, these learnings may prove useful for those implementing them in the future.

PURPOSE AND SUITABILITY OF WEARABLES

Vital signs wearables are devices worn for continuous and non-invasive monitoring of vital signs (5). Once attached to the patient, remote and tetherless monitoring occurs (5), reducing the contact of nurses with infectious cases and reducing the workload of performing vital signs measurements manually.

Before deciding which wearable device to use, the purpose and suitability of wearables in the specific clinical environment should be considered. This depends on the severity and contagiousness of disease, as well as availability of manpower. Contagiousness of disease refers to how easily it spreads. It is influenced by multiple factors including but not limited to the infectious period, mode of transmission, and ability of the pathogen to survive outside of a host. Refer to **Figure 1** for a decision guide on the suitability of wearables.

Patients with higher severity of illness are unlikely to be highly mobile. Hence traditional bedside monitors rather than wearables may be more suited. Traditional bedside monitors are not affected by poor WIFI/Bluetooth signal strength [a common limitation for wearables (6)], this is a more reliable form of monitoring for patients requiring close monitoring.

For patients with mild illness, wearables may be considered as they are mobile (6). In a situation with high manpower and no contagious disease, spot monitors may suffice. If manpower is low with no contagious disease, wearables could be used for mass and remote monitoring, relieving nurses of the task of manually taking parameters.

If the disease is highly contagious, wearables could be deployed regardless of manpower availability to minimize patient-nurse contact, reducing the exposure of the nurse to the contagion while preserving PPE.

In our case, the use of wearables was in an isolation setting with low severity of illness. Patients were confirmed or suspected COVID-19 cases with low risk for complications. They presented with mild symptoms, had no other medical conditions and could independently perform activities of daily living (ADLs). Due to the increased need for nurses as well as the expected mild illness, the nurse: patient ratio in our setting was lower than that of a general ward. Additionally, COVID-19 is highly contagious. Hence, the decision was made to use wearables with the purpose of patient monitoring while minimizing nurse-patient contact and to preserve PPE, not for early detection of deterioration.

CRITERIA TO DETERMINE SUITABILITY OF THE WEARABLE VITAL SIGNS MONITORING SOLUTION

Once decided that a wearable solution is suitable, a myriad of factors influence the selection of the specific solution. Key criteria to consider are: device functions, fidelity of the product, operational requirements, cyber security, cost effectiveness and sustainability. Refer to **Figure 2** for an overview of the criteria involved.

Device Functions

Device functions refers to the specifications of the device in terms of its physiological measures [e.g., heart rate (HR)], as well as its form factor. Required functions largely depend on the nature of disease.

Measurements

There are solutions for capturing full sets of vital signs: HR, respiration rate (RR), oxygen saturation (SpO_2), temperature, and blood pressure (BP). An example would be the ViSi mobile by Sotera. However, it requires the patient to be strapped onto multiple devices which is not ideal (discussed further in the next section). Therefore, prioritization of vital signs is paramount.

In our case, as COVID-19 is a respiratory disease, monitoring of RR (7), and SpO_2 is important for quick recognition of deterioration (8). HR is also essential as changes in HR occurs as a compensatory mechanism in the early stage of clinical deterioration (9). Therefore, HR, RR, and SpO_2 were prioritized to require close monitoring and the solution we selected measured those parameters.

BP and temperature were measured for our patients at regular intervals (e.g., 4 hourly/6 hourly), rather than continuously. This is because BP is usually not the first vital sign to respond during a deterioration (9), and frequent temperature monitoring for adults with normal thermoregulation is usually not mandated (10). These measurements were timed to be performed when the nurse entered the room for other purposes. This could be done without compromising on patient safety as our patient population was at low risk for complications.

Form Factor

Wearable devices come in many forms including smart watches, chest patches, and pulse oximeters (11). An ideal wearable should have maximum functionality with minimum burden (12). In our population where patients were ADL independent, devices that do not restrict movement were preferred. Comfort of the device ensures compliance on the patients' part. An uncomfortable device may lead to frequent removal, adding burden on nurses to repeatedly troubleshoot the lack of vital signs readings.

Considering the prioritized measurements and form factor, the Masimo SafetyNet^{TM} solution was used in our setting (refer to **Figure 3**). The Radius PPG^{TM} senses the patient's vital signs. The readings are then reflected on the Masimo SafetyNet^{TM} application as well as on a clinician portal at the nurses' counter.

The Radius PPG^{TM} is designed to provide accurate pulse oximetry in the presence of motion and low perfusion (13). It

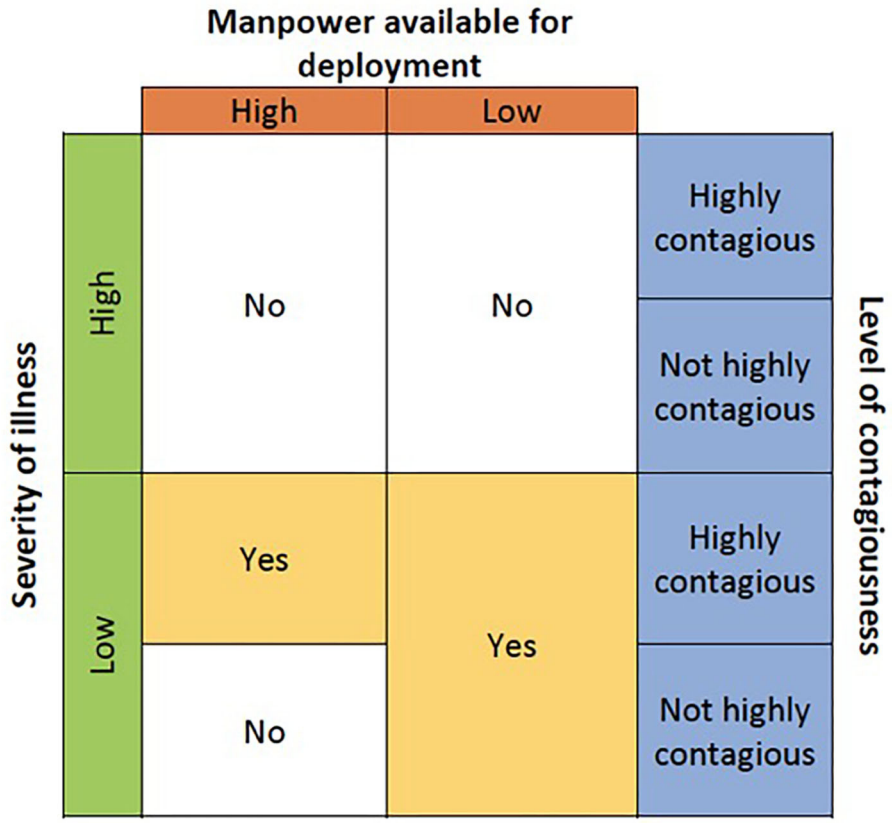


FIGURE 1 | Decision on whether a vital signs wearable is suitable.

also provides a RR derived from phlethysmography (13). An automated measurement of RR is beneficial as RR changes are seen early in deterioration, yet it is often deemed least important by nurses and it is tedious to manually count it (9). The Radius PPG^{TM} is light weight. In a survey of 37 patients, 83.8% agreed or strongly agreed that it was comfortable to wear, and 89.2% agreed or strongly agreed that it did not restrict their movement.

One limitation of the Radius PPG^{TM} was that it had to be removed before a shower. However, it was easy for patients to replace it afterwards following the instructions on a poster provided.

Fidelity of the Product

Wearables available in the market range from commercial grade to medical grade to research grade (11). A device is considered medical grade if it fulfills the regulatory requirements of the region where it is used. For instance, Food and Drug Administration (FDA) in the United States which evaluates effectiveness of the device and its risk for harm (14), European CE mark that affirms the device meets high safety, health and environmental protection requirements (14) and Health Sciences Authority in Singapore.

Duration to Implementation

For use in a healthcare setting, a medical grade device is required. As time is required for validation of new devices as well as for obtaining regulatory requirements, quick deployment during a pandemic demands for wearables that are medical grade. As a note of caution, devices marketed to be medical grade may only have some (not all) of their parameters clinically validated. For instance, Everion by Biofourmis is marketed to be medical grade (15). However, only HR and SpO_2 are clinically validated vital signs while heart rate variability and RR are not (15). Care should be taken to ensure all vital signs prioritized by the medical team have been clinically validated.

Availability of Supply

Surges in demand for medical devices coupled with supply chain disruptions caused some medical grade devices to be unavailable. For example, Canada faced a supply mismatch in pulse oximeters during this pandemic (16). In some cases, a commercial grade device may be deployed due to the lack of a better option. In such cases, a safety net should be in place. The institution should make available some medical grade devices (not necessarily wearables) for rechecking purposes if the patient's vital signs were recorded to be out of range on the commercial grade device, or if the patient reports to be unwell.

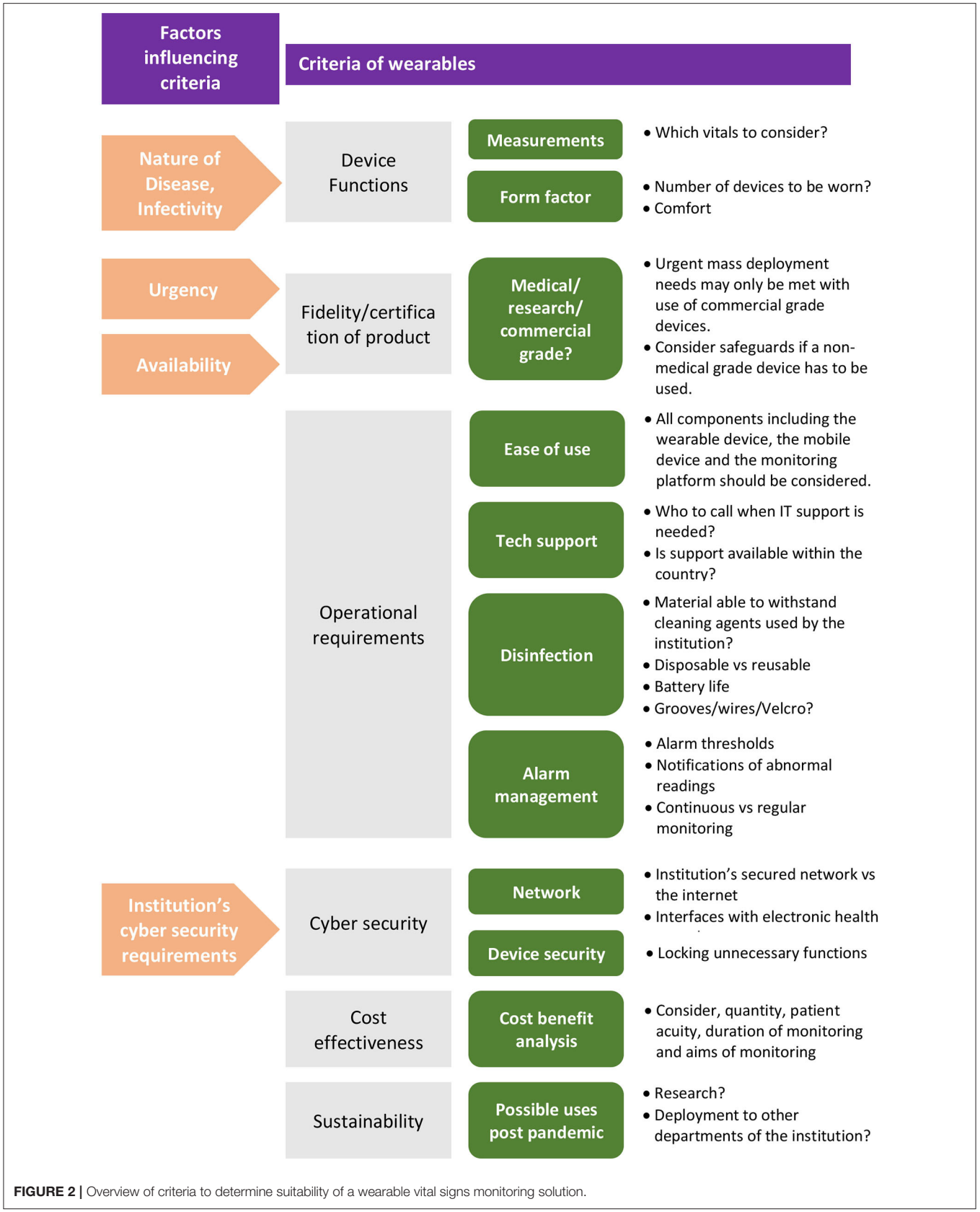


FIGURE 2 | Overview of criteria to determine suitability of a wearable vital signs monitoring solution.

FIGURE 3 | Masimo SafetyNetTM (1) (*Reproduced with permission from Masimo*).

Operational Requirements

Ease of Use

To facilitate training and prevent errors, the solution should be simple. It should enhance the workflow instead of creating an additional burden.

Some practical questions for considerations are as follows:

- Is the wearable device easy to apply?
- Can the patient easily reapply it if it has to be removed for a shower?
- Is the monitoring dashboard clear? Can it be customized?
- How are notifications of abnormal vital signs displayed?
- Is there a sound to alert nurses of an abnormal vital sign?
- Is the monitoring dashboard viewed from an existing intranet environment or will it require separate devices connected to the internet for viewing? (Each of these decisions will require its own cybersecurity assessment).

IT Support

A support structure should be emplaced. Nurses should have ready access to help when technical difficulties are faced. The IT support should consist of staff within the hospital (who can respond quickly), as well as personnel from the company (who will be able to troubleshoot more technical issues). Use of products with a local support office is preferred.

It is also advisable to involve the IT and informatics team from the start of the project.

In our implementation, a common reason for troubleshooting was that the vital signs were not reflected on the clinician portal. Initially, the most common reason was that the battery of the Radius PPG^{TM} ran out. Subsequently, we learnt that if the mobile device was not in use for a prolonged duration, the Bluetooth of the mobile device goes to sleep cutting off the connection between the sensor and the mobile device. This was the main limitation experienced during this implementation as the nurse would need to enter the room to turn on the application in order to continue monitoring the patient. Our team was informed that all current mobile devices turns off Bluetooth after prolonged inactivity. Hence, this is a limiting factor to consider for the use of any wearables relying on Bluetooth connection to a mobile device till future developments resolves this.

Disinfection

The device has to withstand disinfection procedures as per institution's guidelines. In our institution, disinfection with Ultraviolet (UV) treatment or Hydrogen Peroxide Vaporization (HPV) is required for areas or items used by patients who are COVID-19 positive to prevent cross contamination.

Wearables may be disposable or reusable (with rechargeable batteries/disposable batteries). Reusable wearables need to be removed from the room for charging at regular intervals. However, if the patient is not discharged by then, the device which have not undergone UV or HPV treatment cannot be removed from the room for charging. Hence, disposable wearables are preferred. Disposable wearables vary in their battery life. A longer battery life reduces frequency at which they need to be replaced. However, the battery life should not be much longer than the expected length of stay to minimize waste.

If reusable devices are used, disposable batteries would be preferred over rechargeable batteries for the same reason mentioned above. Ease of cleaning should also be considered. Wireless devices without grooves and without materials difficult to disinfect (e.g., Velcro) would be preferred.

For the Masimo SafetyNet^{TM} solution, the Radius PPG^{TM} is disposable with a reusable chip. In our setting, the reusable chip was wiped with 70% isopropyl alcohol (as per manufacturer's instructions) and undergone UV treatment as per

our institution's requirements. It was a small chip without many grooves and it was easy to clean.

Alerts and Alarms

A platform displaying each patient's vital signs at a remote location (e.g., the nurses' station) will be beneficial. The platform should alert nurses to any abnormalities.

Safeguards must be in place to ensure that no deteriorating patient is undetected. Customisable alarm thresholds are necessary to prevent unacceptably high number of alarms (17), preventing alarm fatigue. Customisable dashboards to support operational processes will also be beneficial.

Alarm management is challenging when continuously monitoring patients who are ADLs independent. Traditional vital signs thresholds were set for vital signs taken at rest. However, patients who are ADLs independent may be moving or talking causing artifacts which are one of the biggest problems in data evaluation (5). Although some studies suggest that continuous monitoring with automated alerts improves patient outcomes (6), alarm fatigue could be counter-productive. To prevent alarm fatigue, patients who are relatively well with low risk for complications should have regular rather than continuous monitoring.

Even though most wearable solutions offer continuous monitoring, the purpose for wearables in our situation was not meant for that purpose. As mentioned, our aim was to minimize contact between nurses and patients. Therefore, staff should not be additionally burdened to continuously monitor the patients just because the wearables are able to do so. Rather, adjustments to work processes should be made to maximize the benefits of technology without increasing the burden on staff. For instance, protocol may require nurses to check the wearables recordings at fixed intervals rather than continuously.

Cyber Security

Cyber security is the practice of defending computers, servers, mobile devices, electronic systems, networks, and data from malicious attacks. As healthcare information are highly sensitive, confidentiality is paramount. All patient identifiers and health information should be protected (18). Therefore, the implementation of the wearables necessitates the following:

- Device security of any mobile devices that are used to collect, store, or transmit information;
- Secure data transmission and storage- Data transmission and data at rest have to follow relevant security guidelines (e.g., Health Security Instruction Manual). Data stored in the cloud has to be anonymised to reduce exposure risks of 3rd party product (18).
- Proper account provisioning and management; patient re-identification governance process; data backup and device fidelity are also important hygiene considerations.

Other important risks include malicious hacking to corrupt or alter data collected, introduction of malware that impairs the performance of the device, or the devices being used as portals or mediums for cyber criminals to gain access to enterprise digital assets such as the Electronic Medical Records (EMR) system.

In our institution, EMR and other enterprise IT systems are connected to a private, secured network, not the Internet, as governed by the public healthcare IT policies. Ideally, the wearables solution should sit within this secured network for enhanced cybersecurity and work processes. If the solution was within the secured network and integrated with the EMR, readings from the wearables would be directly charted into the EMR without transcription errors or additional effort from nurses. In addition, full patient identification (e.g., name and registration number) may be viewed for easy patient identification.

However, most wearable solutions are designed to store data in a public cloud. Hence, they require internet access. Furthermore, time is required to architect a secured solution to interface data from the wearables solution to the EMR system. These reasons ruled out our preference of sitting the wearable system in the secured network.

Working with our Chief Information Security Officer (CISO) and IT teams, we arrived at a quick implementation of an internet enabled solution. To maintain cyber security, the wearables solution was a stand-alone system with no patient identifiers within it. Pseudo IDs were used to mitigate risks associated with cybersecurity. All functions in the mobile device except those required for the solution to work were locked down to prevent usage habits from sabotaging security of the device or software system.

Another possible scenario without syncing the wearables solution with the EMR systems would be for vital signs to be measured and self-charted by patients onto a platform that can be accessed by the nurses. This is not recommended as there are some major limitations. A similar approach was carried out in some community isolation facilities (CIFs) in Singapore. CIFs isolated patients with very mild symptoms not requiring hospital stay. These patients were provided with vital signs monitoring devices (not wearables) and were required to self-chart their vital signs. Challenges faced were that some patients confused the PR with SpO_2 and entered "PR = 99 bpm, SpO_2 = 60%," instead of the other way round. Patients may also measure their vital signs after physical activity, leading to a high number of false alerts being sent to clinicians.

Cost Effectiveness

It is unclear when the pandemic will end. Hence, the solution needs to be cost effective. Severity of illness, quantity required, aims of monitoring using the wearables should be taken into consideration in determining its cost effectiveness.

Sustainability

To prevent wastage, potential uses of the wearables after the pandemic should be contemplated during the selection of the solution. Suggestions for future use of wearables would be for

research purposes or for feasibility trials in the management of other groups of patients (such as outpatients or patients enrolled in a hospital at home program). If the use of that wearable device proves successful, plans could be made to integrate the wearables system with the EMR within the secured network, and to implement its use across the institution.

CONCLUSION

The use of vital signs wearables can be expected to rise with the ongoing advancement in technology. Although this list of considerations is not exhaustive, this may be a starting point for those looking to implement a wearables solution in their area.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

EF, GP, BT, and KY were involved in the implementation of the solution with support from all other authors. EF, LT, LC, CC, and BT were involved in the drafting of the manuscript. SA and KY provided mentorship in the drafting of the final manuscript and supervised the project. All authors were involved in the conception of the original idea.

REFERENCES

1. World Health Organization. *Rolling Updates on Coronavirus Disease (COVID-19)*. (2020). Available online at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen
2. World Health Organization. *WHO Timeline- Covid-19*. (2020). Available online at: https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19
3. Ministry of Health Singapore. *Updates on COVID-19 (Coronavirus Disease 2019) Local Situation*. (2020). Available online at: https://www.moh.gov.sg/covid-19
4. Fan PEM, Fazila A, Ang SY, Elena BMA, Norhayati BA, Chiang JL, et al. Preparation and response to COVID-19 outbreak in Singapore: a case report. *Infection Dis Health*. (2020) 25:216–8. doi: 10.1016/j.idh.2020.04.002
5. Jacobsen M, Dembek TA, Kobbe G, Gaidzik PW, Heinemann L. Noninvasive continuous monitoring of vital signs with wearables: fit for medical use? *J Diabet Sci Technol*. (2021) 15:34–43. doi: 10.1177/1932296820904947
6. Leenen JP, Leerentveld C, van Dijk JD, van Westreenen HL, Schoonhoven L, Patijn GA. Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: systematic review. *J Med Internet Res*. (2020) 22:e18636. doi: 10.2196/18636
7. Quaresima V, Ferrari M. More on pulse oximetry for monitoring patients with COVID-19 at home. *Ann Am Thoracic Soc*. (2020) 17:1496. doi: 10.1513/AnnalsATS.202006-701LE
8. Luks AM, Swenson ER. Pulse oximetry for monitoring patients with COVID-19 at home. potential pitfalls and practical guidance. *Ann Am Thoracic Soc*. (2020) 17:1040. doi: 10.1513/AnnalsATS.202005-418FR
9. Mok W, Wang W, Cooper S, Ang ENK, Liaw SY. Attitudes towards vital signs monitoring in the detection of clinical deterioration: scale development and survey of ward nurses. *Int J Qual Health Care*. (2015) 27:207–13. doi: 10.1093/intqhc/mzv019
10. Cahill K. *Royal Prince Alfred Hospital Patient Observation (Vital Signs) Policy-Adult*. Sydney South West Area Health Service (2014). Available online at: https://support.biofourmis.com/hc/en-us/articles/212369329-Everion-data-quality-
11. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. *Personalized Med*. (2018) 15:429–48. doi: 10.2217/pme-2018-0044
12. Bodine K, Gemperle F. Effects of functionality on perceived comfort of wearables. In: *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings*. Toronto, ON: Citeseer (2003). p. 57.
13. Masimo. *Masimo SafetyNet*. Alexandria, VA (2021).
14. Conley D. *Two Paths for Medical Device Approval: FDA vs. CE. Health Management*. New York, NY: AnnalsATS (2015).
15. Biofourmis. *Everion Data Quality*. New York, NY (2021).
16. Government of Canada. *Medical Device Shortage: Saturometer (Pulse Oximeter)*. England (2020).
17. Downey C, Randell R, Brown J, Jayne DG. Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: pilot cluster randomized controlled trial. *J Med Internet Res*. (2018) 20:e10802. doi: 10.2196/10802
18. Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. *Clin Pharmacol Therapeut*. (2018) 104:42–52. doi: 10.1002/cpt.966

Conflict of Interest: KY declares the following conflicts of interest: 1. Research funding, unrelated to this project: Biofourmis, Holmusk, Bayer, Medtronic, Astra Zeneca, and Shockwave Medical. 2. Consultancy: Abbott Vascular, Boston Scientific, Medtronic, Amgen, Bayer, Novartis, and Medopad. 3. Speaker or Honararia: Shockwave Medical, Abbott Vascular, Boston Scientific, Medtronic, Philips, Alvimedica, Biotronik, Amgen, Astra Zeneca, Orbus Neich, and Bayer.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 24 November 2021
doi: 10.3389/fdgth.2021.678009

Retrospective Analysis and Forecasted Economic Impact of a Virtual Cardiac Rehabilitation Program in a Third-Party Payer Environment

Arash Harzand[1], Aaron C. Weidman[2], Kenneth R. Rayl[2], Adelanwa Adesanya[3], Ericka Holmstrand[2], Nicole Fitzpatrick[2], Harshvardhan Vathsangam[3] and Srinivas Murali[4]*

[1] Emory University School of Medicine, Atlanta, GA, United States, [2] VITAL Innovation, Highmark Health, Pittsburgh, PA, United States, [3] Moving Analytics, Inc., Los Angeles, CA, United States, [4] Cardiovascular Institute, Allegheny Health Network, Pittsburgh, PA, United States

OPEN ACCESS

Edited by:
Constantinos S. Pattichis, University of Cyprus, Cyprus

Reviewed by:
Martijn Scherrenberg, Jessa Hospital, Belgium
Nathaniel Moulson, University of British Columbia, Canada

***Correspondence:**
Arash Harzand
aharzan@emory.edu

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *08 March 2021*
Accepted: *15 October 2021*
Published: *24 November 2021*

Citation:
Harzand A, Weidman AC, Rayl KR, Adesanya A, Holmstrand E, Fitzpatrick N, Vathsangam H and Murali S (2021) Retrospective Analysis and Forecasted Economic Impact of a Virtual Cardiac Rehabilitation Program in a Third-Party Payer Environment. Front. Digit. Health 3:678009. doi: 10.3389/fdgth.2021.678009

Background: Participation in cardiac rehabilitation (CR) is recommended for all patients with coronary artery disease (CAD) following hospitalization for acute coronary syndrome or stenting. Yet, few patients participate due to the inconvenience and high cost of attending a facility-based program, factors which have been magnified during the ongoing COVID pandemic. Based on a retrospective analysis of CR utilization and cost in a third-party payer environment, we forecasted the potential clinical and economic benefits of delivering a home-based, virtual CR program, with the goal of guiding future implementation efforts to expand CR access.

Methods: We performed a retrospective cohort study using insurance claims data from a large, third-party payer in the state of Pennsylvania. Primary diagnostic and procedural codes were used to identify patients admitted for CAD between October 1, 2016, and September 30, 2018. Rates of enrollment in facility-based CR, as well as all-cause and cardiovascular hospital readmission and associated costs, were calculated during the 12-months following discharge.

Results: Only 37% of the 7,264 identified eligible insured patients enrolled in a facility-based CR program within 12 months, incurring a mean delivery cost of $2,922 per participating patient. The 12-month all-cause readmission rate among these patients was 24%, compared to 31% among patients who did not participate in CR. Furthermore, among those readmitted, CR patients were readmitted less frequently than non-CR patients within this time period. The average per-patient cost from hospital readmissions was $30,814 per annum. Based on these trends, we forecasted that adoption of virtual CR among patients who previously declined CR would result in an annual cost savings between $1 and $9 million in the third-party healthcare system from a combination of increased overall CR enrollment and fewer hospital readmissions among new HBCR participants.

Conclusions: Among insured patients eligible for CR in a third-party payer environment, implementation of a home-based virtual CR program is forecasted to yield significant cost savings through a combination of increased CR participation and a consequent reduction in downstream healthcare utilization.

Keywords: mobile health, virtual care, cardiac rehabilitation, economic impact, coronary artery disease

INTRODUCTION

Cardiac rehabilitation (CR) is recommended for patients with coronary artery disease (CAD) to reduce the risk of hospital readmission and cardiovascular (CV) death after an acute myocardial infarction (MI) or coronary procedure (1, 2). Despite the benefits, fewer than 20% of eligible patients enroll in center-based CR (CBCR) (3). Limited program availability, distance, and high cost make attending CBCR—typically delivered over 12 weeks with thrice weekly sessions—burdensome to patients and these factors all contribute to low participation (4). There is a clear need to develop effective and patient-centric alternatives to expand CR access for eligible patients (4). This need has only been magnified by the emergence of new safety considerations for higher-risk patients from travel and social exposure as well as the need for many CBCR programs to operate on more restricted schedules and reduced patient appointments during the ongoing COVID pandemic.

One proposed alternative is virtual, home-based cardiac rehabilitation (HBCR) which combines self-led exercise training with health coaching and remote patient monitoring, often through a mobile health platform (5). HBCR has been shown to be non-inferior to facility-based CR in large meta-analyses and is supported by clinical practice guidelines with a Class IIa (i.e., reasonable alternative) recommendation for patients who are unable or unwilling to participate in a facility-based program (6–8). Aside from fully integrated, risk-bearing healthcare systems such as Kaiser Permanente (KP) and the Department of Veterans Affairs (VA), widespread adoption of HBCR has remained poor among systems operating within third-party payer environments due to limited reimbursement (5, 6, 9–11).

MULTIFIT is an evidenced-based HBCR model with demonstrated success in reducing readmissions and adverse events in CAD patients and this program has been widely adopted among integrated care networks such as KP (**Figure 1**) (9, 10). MULTIFIT is a 12-week, nurse-mediated case-management program that delivers guideline-based recommendations for comprehensive CAD risk factor modification through remote encounters (12). During each virtual visit, a nurse manager provides counseling on and sets goals for exercise, smoking cessation and dietary modification using patient-derived algorithms that have been previously described (9). Among CAD patients within KP of Northern California, enrollment in MULTIFIT led to significant reductions in hospital readmission (49%), recurrent MI (58%), and all-cause mortality (53%) (13). More recently, MULTIFIT has been successfully implemented within the VA through a mobile health platform with high retention (90% at 30 days, 62% at 90 days) and high patient satisfaction (80%) (5). Still, the feasibility of delivering MULTIFIT as a virtual CR program within commercial payer environment remains unknown. Here, we evaluated the potential clinical and economic feasibility of implementing a virtual HBCR program based on MULTIFIT in a third-party payer system.

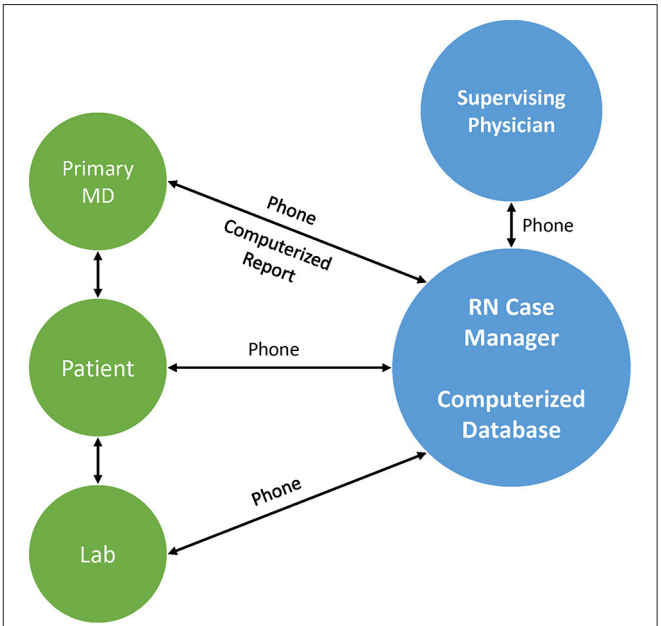


FIGURE 1 | The MULTIFIT model of home-based cardiac rehabilitation.

METHODS

Among eligible patients admitted for a CAD related diagnosis or procedure, our objectives in this retrospective cohort study were to (1) describe the current rates of CBCR enrollment and hospital readmission within 12 months after discharge, (2) calculate healthcare delivery costs including CBCR delivery, all-cause and CV related readmission, and (3) estimate the economic impact of implementing a MULTIFIT-based virtual HBCR program within a third-party payer system. For the economic feasibility calculation, we hypothesized that virtual HBCR adoption would generate economic savings through a combination of increased CR adoption from new patients that would have not otherwise participated in CR and subsequent downstream savings from reduced hospitalizations. The additional *cost* for new HBCR participants would therefore be outweighed by the *savings* that would result from improved downstream clinical outcomes.

Data Sources

We retrospectively analyzed data from Highmark Health, a non-profit health care organization providing insurance coverage primarily throughout Pennsylvania, West Virginia, and Delaware. Claims data were derived from commercial, Medicare Advantage, and Medicare Supplemental insurance plans. Costs included in this report refers to allowed amounts, or the maximum possible payment that the insurance plan could be obligated to make to medical providers for services rendered.

Population

Our population of interest included patients aged 18+ years with an index hospitalization for a CAD-related primary diagnostic or procedural code (e.g., CABG or PCI) between October 1, 2016, and September 30, 2017 (see **Supplementary Table 1** for full list of codes). Patient insurance claims were examined post-discharge for an additional 12-months (through at latest September 30, 2018) to measure CR enrollment and readmission rates. We included all members who met eligibility criteria regardless of their primary state of residence; most Highmark members live in Pennsylvania, Delaware, and West Virginia, but members span all 50 US states. We only included patients with continuous enrollment in an eligible insurance plan for the 12 months following discharge to obtain complete follow up information. Federal Employee Program (FEP) members were excluded due to limitations in use of their claims data.

CR Participation Rates

CR participation rate was defined as the percent of patients who had at least one insurance claim for an outpatient CR session during the 12-months follow discharging. For each patient, we also calculated the days between discharge and the initial CR session (CR enrollment lag) and the total number of CR sessions attended. We used *z*-tests to compare CR participation rates across subgroups (e.g., men vs. women).

Healthcare Utilization and Hospital Readmission Rates

All-cause and CV-related readmission rates were calculated at monthly intervals up to 12-months post-discharge. We defined each hospital readmission by the presence of at least one inpatient insurance claim using billing codes indicating claim type (e.g., inpatient, outpatient) and place of service (e.g., hospital inpatient, urgent, or emergency care); emergency department and observation admissions were excluded. We identified CV-related encounters using "Major Diagnostic Category" codes in the insurer database. We used *z*-tests to compare the cumulative proportion of CR participants and non-participants who experienced hospital readmission at each monthly interval. We used *t*-tests to compare the mean number of hospital readmission episodes among patients who were readmitted at least once between the CR participant and non-participant groups.

CR and Healthcare Costs

Per-patient healthcare costs for CBCR delivery were calculated from claims data. Per-patient HBCR cost was set at $1,550, based on the commercial price of the Movn platform (Moving Analytics, Los Angeles, CA; note that this value is akin to an allowed amount from an insurance perspective—it is the maximum amount the insurer would typically have to pay for the medical service—and is therefore comparable to our calculated cost of CBCR). To provide an initial check on generalizability for CR costs, we obtained national CBCR costs from claims incurred between January 1, 2017, and December 31, 2018, using the Blue Health Intelligence (BHI®) National Data Warehouse Analytical DataMart (ADaM), a database of insurance claims populated with contributions from private payers within the Blue Cross Blue Shield Insurance Network across all 50 US states. Episodic costs for all-cause and CV-related readmissions during the 12-month follow up period were also calculated using claims data.

Clinical and Economic Forecasts

We conducted a clinical and economic analysis of HBCR implementation by forecasting scenarios in which: (a) some of the patients who did not originally participate in CBCR chose to participate via HBCR, and (b) this group of patients showed a lower rate of hospital readmission (14). We used previously published MULTIFIT data as a guide to define these scenarios and hypothesized a net savings as a result of (a) and (b) (12, 15, 16). We further conducted sensitivity analyses to examine the robustness of these forecasts to potential fluctuations in per-patient hospital readmission costs based on the data observed in the current study.

Study Design and Oversight

Moving Analytics provided funding for the study through a contract with the VITAL Innovation Program at Highmark Health. The study was designed by Highmark Health in consultation with the sponsor. Highmark Health had unrestricted access to the study data and was responsible for data collection and analysis. Data were analyzed by two authors who are employed by Highmark Health (AW and KR). The sponsor and its affiliated authors (AH, AA, & HV) did not have access to the study data and participated in data interpretation only. The first and second authors (AH and AW) drafted the manuscript. The trial was approved by the IRB at Highmark Health and informed consent was waived for the retrospective claims analysis.

RESULTS

CR Participation and Delivery Costs

Figure 2 displays a flowchart for patient identification and CR participation. We identified 7,264 unique inpatients who were eligible for CR (mean age = 67 ± 12.9 years; 67% men). CR participation was only 37% within 12 months of discharge (n = 2,663) with a mean enrollment lag of 49 ± 49 days. CR participants attended a mean of 23 ± 13 sessions with a mean per-patient allowed cost of $2,922 ± $2,413. Comparatively, data from BHI® ADaM demonstrated a mean annual cost of $2,870 per patient-member, speaking to the generalizability of our observed mean CBCR delivery cost.

Among subgroups, men participated at a higher rate than women (40 vs. 29%; $z = 9.37$; $p < 0.001$). CR participation

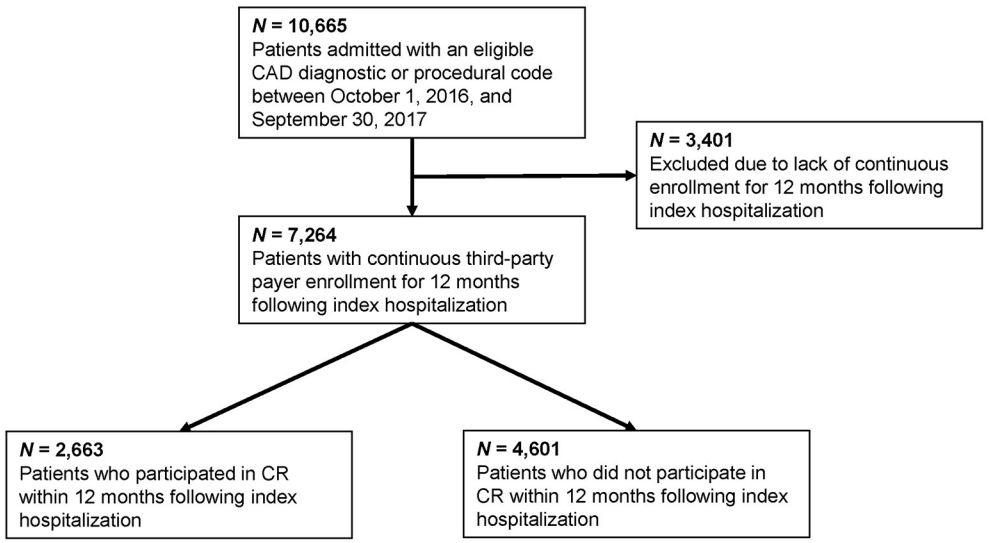


FIGURE 2 | Patient flowchart.

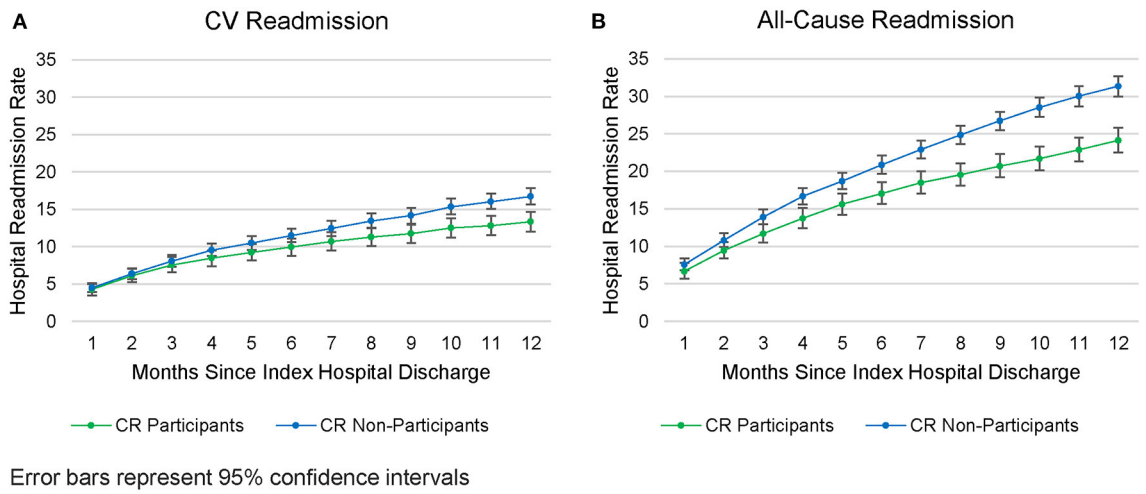


FIGURE 3 | Readmissions rates from the CAD population. **(A)** CV Readmission. **(B)** All-cause Readmission.

was higher among patients <60 years of age (43 vs. 34% in those >60 years; $z = 7.04$; $p < 0.001$); consistent with this finding, participation was also higher in patients with commercial compared to senior-oriented (e.g., Medicare Advantage) insurance plans (42 vs. 31%; $z = 9.42$; $p < 0.001$; see **Supplementary Tables 2, 3**).

Readmission Rates and Cost

Cardiovascular-related readmission rates were lower among CR participants than non-participants beginning at 6-months post-discharge (11.5 vs. 10%; $z = 2.01$, $p = 0.045$; see **Figure 3A**), a trend which continued until the 12-month time point (13 vs. 17%, $z = 3.84$, $p < 0.001$). An even earlier divergence was seen for all-cause readmission rates, which were lower among CR participants (vs. non-participants) beginning at 3-months post-discharge (11.7 vs. 13.9%; $z = 2.70$, $p = 0.007$; see **Figure 3B**) a difference that persisted until 12-months (24 vs. 31%, $z = 6.61$, $p < 0.001$). Furthermore, among patients who were readmitted, CR participants were readmitted less frequently than non-participants [1.48 ± 0.95 vs. 1.74 ± 0.35 readmissions; $t(2,084) = 4.42$; $p < 0.001$; see **Supplementary Table 4** for the most frequent readmission etiologies]. The mean per-patient, per readmission allowed amount did not differ significantly between the CR and non-CR groups [$32,164 vs. $30,213; $t(2,084) = 0.61$; $p = 0.55$].

Economic Analysis

HBCR would offer a viable option for the 63% of patients in our sample who declined facility-based CR. If some patients were to take this option, overall CR participation rate would increase, and hospital readmissions and associated costs would likely decline. We projected the economic impact of offering HBCR in a third-party payer population by modeling nine hypothetical scenarios which varied on: (a) the rate at which the 4,601 patients who previously declined CR enrolled in HBCR, and (b) the hospital readmission rate among this group of new HBCR participants.

To best anchor our economic forecasts for the Movn virtual program, we primarily relied on published data from prior MULTIFIT trials which represented the evidence base upon which Movn was developed. For HBCR enrollment, the "Pessimistic" scenario was set to reflect extremely low uptake (i.e., 10%), as this would imply a nearly complete rejection of the hypothesis that Movn would significantly increase enrollments. Subsequently, we set the "Conservative" scenario to closely match the lowest reported enrollment rate from prior MULTIFIT studies [41% from Levin et al. (16), and the "Optimistic" scenario was set to closely match the highest reported enrollment rate in prior MULTIFIT studies (80% in Landis et al. and 83% in DeBusk et al.)] (12, 15).

We applied similar rationale to determine our projected hospital readmission rates, again basing these values on the current state of readmissions at Highmark as well as previously published data from MULTIFIT. For each of the 3 scenarios (Pessimistic, Conservative, Optimistic) we performed 3 separate projections for possible readmission rates set to "Low," "Moderate," and "High." We set the most pessimistic possible outcome (i.e., "High" rate of readmissions) to closely match the current state of hospital readmissions following center-based CR enrollment at Highmark (i.e., 24% based on our current data). The "Low" readmission rate was set to closely match the best/lowest reported readmission rate – we were limited with the existing published MULTIFIT studies as none previously reported on readmissions rates, hence we relied on published rates from the 2015 Cochrane review (which is now updated to the 2017 version by Anderson et al.). In both systematic reviews, the lowest reported readmission rate following an HBCR intervention was 8% (from Jolly 2007) which we rounded up to 10% for the "Low" readmission rate projection. The "Moderate" readmission rate projection was set to a value in between the "Low" and "High" estimates (i.e., 17%) (14, 17).

For each of the nine scenarios (3 CR participation rates*3 readmission rates), we compared the *projected* annual cost to the *current* annual medical cost of $72.1 million in our sample, which included (a) $28.5 million for delivery of facility-based CR to 2,663 patients and (b) $43.6 million for hospital readmission-related medical treatment for 2,087 patients. A projected cost that fell below $72.1 million would indicate a projected net savings under HBCR implementation.

We projected a median savings of $4.5 million across scenarios (see **Table 2**). Projected net savings varied considerably, increasing as the rate of HBCR increased and as the 12-month hospital readmission rate decreased. For example, at 10% HBCR participation and 24% readmission rate (identical to the current observed rate of 24%), projected savings were just $98,000, whereas projected savings were $17.3 million at 80% participation and 10% readmission rate.

TABLE 1 | Variables used in economic impact analysis.

Category	Scenario 1 (Pessimistic)	Scenario 2 (Conservative)	Scenario 3 (Optimistic)
HBCR CR participation rate	10%	40%	80%
Patients			
Facility-based CR	2,663	2,663	2,663
HBCR	460	1,840	3,681
Non-CR	4,141	2,761	920
CR cost			
Facility-based	← $2,922 [$2,830,$3,013] →		
HBCR	← $1,550 [Fixed] →		
Readmission cost			
CR Patients	← $32,164 [$28,787, $35,540] →		
Non-CR Patients	← $30,213 [$26,276, $34,151] →		
Readmission rate			
Low	10%	10%	10%
Moderate	17%	17%	17%
High	24%	24%	24%

Cost values are mean [95% Confidence Interval].
← → indicates that a value is constant across all three scenarios.

Sensitivity Analyses

Episodic hospital readmission costs showed considerable variability across patients, which could affect the result of our economic analysis (see **Table 1**). As a sensitivity analysis, we therefore ran two additional economic forecasts to account for potential fluctuation in hospital readmission cost in future cardiac care samples and/or in other payer environments. Under a "best case" economic outcome, we set the per patient readmission cost at $28,787 for CR patients (i.e., the low end of the 95% confidence interval for this value) and at $34,151 for non-CR patients (i.e., the high-end of the 95% confidence interval). Not surprisingly, this scenario yielded healthy projected annual savings, with a median of 7.9 million, never falling below $1.0 million, and ranging as high as $23.1 million (see **Table 2**).

Conversely, under a "worst case" economic outcome, we set the per-patient readmission costs at $35,540 for CR patients and $26,276 for non-CR patients (i.e., the high- and low-ends of the 95% confidence intervals, respectively). This scenario yielded a more mixed financial picture. At 10 and 17% readmission rates, we still projected savings that were often substantial (range: $300,000–11.6 million). In contrast, at 24% readmission rate, we projected a net *cost* to the healthcare system (range: $ –$800,000 to –$6.7 million; see **Table 2**). Note that the projected net cost at 24% readmission rate became *larger* with higher participation in HBCR, because HBCR patients in this scenario would incur costs both to pay for their cardiac rehab and to cover costs for their relatively high hospital readmission rate.

TABLE 2 | Results of economic impact analysis.

Category	Scenario 1 (Pessimistic)	Scenario 2 (Conservative)	Scenario 3 (Optimistic)
HBCR participation rate (anticipated)	10%	40%	80%
Projected annual cost			
Facility-based CR patients		← $28.4 →	
Non-CR patients	$39.3	$26.2	$8.7
HBCR patients			
Readmission rate			
10%	$1.5	$8.8	$17.5
17%	$3.2	$12.9	$25.8
24%	$4.3	$17.1	$34.1
Projected total annual cost			
Readmission rate			
10%	$69.2	$63.4	$54.7
17%	$70.9	$67.5	$63.0
24%	$72.0	$71.7	$71.3
Projected net savings			
Readmission rate			
10%	$2.8 [$2.1, $3.6]	$8.7 [$5.8, $11.6]	$17.3 [$11.6, $23.1]
17%	$1.1 [$0.3, $2.0]	$4.5 [$1.2, $7.9]	$9.0 [$2.4, $15.7]
24%	$0.01 [$ −0.08, $1.0]	$0.4 [$-3.4, $4.2]	$0.8 [$-6.7, $8.3]

Projected costs/savings are in millions of dollars. Projected net savings compares the projected total annual cost under each modeled scenario to the current total annual cost of $72.1 million. 95% confidence intervals for projected net savings are taken from the above sensitivity analysis.

DISCUSSION

Our study shows that multidisciplinary CR remains significantly underutilized in population of 7,264 commercially and Medicare-insured beneficiaries with a CAD-related index hospitalization within a large third-party payer environment. Notably, the observed participation rate of 37% falls far below national initiatives to achieve ≥70% participation by 2022 (18), suggesting that inventive efforts, such as virtual HBCR, are needed to increase participation. Even if facility-based programs become increasingly available, ongoing concerns regarding CAD patients' safety during the COVID pandemic may continue to restrict availability (19).

We also forecasted the impact of offering HBCR to the 4,601 patients in our sample who originally *did not* participate in facility-based CR, by modeling (a) the rate at which the 4,601 patients who previously declined CR choose to participate in HBCR, and (b) the hospital readmission rate among this group of new HBCR participants. We projected a median annual savings of $4.5 million through a combination of increased CR participation and reduced hospital readmissions among new HBCR participants. These modal projections reflect the most plausible assumptions of a 40% HBCR participation rate and a 12-month hospital readmission rate for HBCR participants of 17% (see **Table 2**). Sensitivity analyses suggested that savings ranging from $98,000 to $17.3 million may theoretically be seen depending on local variations in HBCR enrollment and readmission as seen in prior studies (12, 14–16). Conversely, sensitivity analyses also helped to identify a boundary condition to the projected net savings, namely if the readmission rate among HBCR participants remains at or above our observed rate of 24% among CBCR participants, and if per-episode costs hospital readmission are higher than expected among all CR participants.

One benefit of HBCR therefore is cost-effectiveness, given that home-based programs provide a relatively inexpensive alternative to facility-based care ($1,550 vs. an average of $2,922 per-patient in our analysis). Several previous trials have shown a reduction in cardiac morbidity and hospital readmission as a result of facility-based CR participation (20, 21). Yet, these positive clinical outcomes can be partially offset by high investment and capital expense required for personnel, equipment, and space, thereby preventing the health system from realizing maximal benefits of traditional CR across large populations (22).

Another potential benefit of HBCR is a reduction in enrollment wait times as delays in enrollment are associated with mitigated benefits of CR following acute MI and CABG surgery (23, 24). Prior studies have shown that for every day that passes after hospital discharge, there is a ∼1% decrease in CR participation (5). Among the factors associated with longer wait times include being employed and longer drive times to CR, both of which reduce the convenience of participation (24). Early enrollment within 21 days of a qualifying event is therefore an important quality metric that results in increased CR participation and maximizes its potential benefits for eligible patients (6). Indeed programs that feature early enrollment in home-based CR have been shown to improve functional status and quality of life, both of which are linked to improvements in long-term outcomes following CR (6, 25). In the present study we observed an average latency of 49 days between hospital discharge and enrollment in facility-based CR, with <25% of CBCR participants enrolling within 21 days following discharge, thus highlighting the need for novel interventions to improve these metrics.

Limitations and Considerations

Our conclusions should be weighed against several limitations that warrant discussion, including the retrospective nature of our analysis of current CR utilization trends and our focus on one healthcare system (albeit one that encompasses members from many US states). More broadly, our conclusions regarding the potential clinical and economic benefits of HBCR are based on forecasts, which themselves make use of parameters (e.g., HBCR participation rate) were based on prior research involving MULTIFIT. All of these may limit the generalizability of our findings to other health systems and patient populations.

Prior studies have also shown challenges with HBCR adoption, given the large commitment and buy-in required

from health systems, providers, and patients for a successful implementation (6). Additionally, given that HBCR is delivered via smartphone-based application, successful completion of HBCR requires a level of technological fluency that cannot be assumed among the older adults most typically referred for CR. In line with this concern, CR participation in our sample was higher among patients younger than age 60 compared to those aged 60+ years, who in turn had higher hospital readmission rates and consequent medical costs. This suggests that converting older patients—who have a higher risk of cardiac disease coupled with a potential hesitance or inability to attend in-person CR—into HBCR participants would likely contribute heavily to an economic savings realized through HBCR adoption; as an important counterpoint, however, in a population where CR participation was already high among older (vs. younger) adults, an economic analysis such as the one we conducted may not show as much potential benefit of HBCR implementation. Nonetheless, in contexts where participation is already low among older adults, overcoming technological barriers among this population will be an essential hurdle for HBCR programs. Our study can therefore serve as a template for other systems to independently assess the potential clinical and economic impact of implementing an HBCR program under realistic conditions.

Furthermore, our economic analysis utilized allowed amounts—the maximum possible cost to an insurer for medical services (e.g., hospital readmissions). Allowed amounts can diverge from actual paid amounts due to several idiosyncratic factors (e.g., variable hospital facility costs; patient insurance types; deductible progress throughout the year). Using allowed amounts to make economic calculations therefore achieves more consistency and is standard in the health insurance industry. To ensure that our results were interpretable, we used allowed amounts across the board in our economic analysis, including for facility-based CR, hospital readmissions, and Movn HBCR.

Finally, the per-patient price of $1,550 for delivery of HBCR is subject to change across populations and healthcare systems. We set this price *ad-hoc* for our economic analysis based on the current commercial price offered by Moving Analytics (rather than choosing a *post-hoc* price based on what would make our economic analysis appear favorable). Importantly, salaries for staff directly involved in HBCR administration (e.g., nurses; physicians) are bundled into this overall per-patient cost and facility costs are moot given the virtual nature of HBCR, all of which should limit variability in per-patient cost. Yet variability in future healthcare settings is still possible, due to fluctuations in costs related to identifying eligible patients and coordinating HBCR referrals, which fall outside of the per-patient cost used in our analysis. Future economic analyses could therefore yield somewhat different results from the ones presented above due to this variability.

CONCLUSIONS

Our work suggests that there is significant opportunity to improve the CR implementation and delivery while generating substantial economic savings for third-party payers through adoption of a virtual, home-based CR program. We observed low participation rate for facility-based CR among eligible insured patients—a rate that could further dip during the ongoing COVID pandemic—and we projected significant cost savings from transitioning non-participants to HBCR and observing a subsequent reduction in hospital readmission rates. Introducing HBCR in a commercial payer environment may therefore prove to be beneficial to cardiovascular care more broadly by leading to improved outcomes among patients with CAD.

DATA AVAILABILITY STATEMENT

The raw data for this report was generated from insurance claims belonging to a private company and is therefore proprietary. Due to resulting privacy concerns, the data cannot be made publicly available. Further inquiries can be made to the corresponding author/s.

ETHICS STATEMENT

The studies described in this report were reviewed and approved by the IRB at Highmark Health and informed consent was waived for the retrospective claims analysis. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

AH and AW drafted the manuscript. Data were analyzed by two authors who are employed by Highmark Health (AW and KR). The sponsor and its affiliated authors (AH, AA, and HV) did not have access to the study data and participated in data interpretation only. All authors contributed to the article and approved the submitted version.

FUNDING

Moving Analytics provided funding for the study through a contract with the VITAL Innovation Program at Highmark Health who facilitated the design and executed the analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.678009/full#supplementary-material

REFERENCES

1. Thomas RJ, King M, Lui K, Oldridge N, Piña IL, Spertus J, et al. AACVPR/ACC/AHA 2007 performance measures on cardiac rehabilitation for referral to and delivery of cardiac rehabilitation/secondary prevention services. *Circulation.* (2007) 116:1611–42. doi: 10.1161/CIRCULATIONAHA.107.185734
2. Anderson L, Thompson DR, Oldridge N, Zwisler AD, Rees K, Martin N, et al. Exercise-based cardiac rehabilitation for coronary heart disease. *J Am Coll Cardiol.* (2016) 67:1–12. doi: 10.1016/j.jacc.2015.10.044
3. Dunlay SM, Pack QR, Thomas RJ, Killian JM, Roger VL. Participation in cardiac rehabilitation, readmissions, and death after acute myocardial infarction. *Am J Med.* (2014) 127:538–46. doi: 10.1016/j.amjmed.2014.02.008
4. Balady GJ, Ades PA, Bittner VA, Franklin BA, Gordon NF, Thomas RJ, et al. Referral, enrollment, and delivery of cardiac rehabilitation/secondary prevention programs at clinical centers and beyond. *Circulation.* (2011) 124:2951–60. doi: 10.1161/CIR.0b013e31823b21e2
5. Harzand A, Witbrodt B, Davis-Watts ML, Alrohaibani A, Goese D, Wenger NK, et al. Feasibility of a smartphone-enabled cardiac rehabilitation program in male veterans with previous clinical evidence of coronary heart disease. *Am J Cardiol.* (2018) 122:1471–76. doi: 10.1016/j.amjcard.2018.07.028
6. Thomas RJ, Beatty AL, Beckie TM, Brewer LC, Brown TM, Forman DE, et al. Home-based cardiac rehabilitation: a scientific statement from the American association of cardiovascular and pulmonary rehabilitation, the American heart association, and the American college of cardiology. *Circulation.* (2019) 140:E69–89. doi: 10.1161/CIR.0000000000000663
7. Members TF, Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, et al. 2016 European guidelines on cardiovascular disease prevention in clinical practice. The sixth joint task force of the European society of cardiology and other societies on cardiovascular disease prevention in clinical practice (Constituted by Representatives of 10 Societies and by Invited Experts. Developed with the special contribution of the European association for cardiovascular prevention & rehabilitation. *Giornale Italiano Cardiol.* (2017) 18:547–612. doi: 10.1007/s12529-016-9583-6
8. Hageman D, Fokkenrood HJ, Gommans LN, van den Houten MM, Teijink JA. Supervised exercise therapy versus home-based exercise therapy versus walking advice for intermittent Claudication (Review) summary of findings for the main comparison. *Cochrane Database Syst Rev.* (2018) 2018:88. doi: 10.1002/14651858.CD005263.pub4
9. Miller NH, Warren D, Myers D. Home-based cardiac rehabilitation and lifestyle modification: the multifit model. *J Cardiovasc Nurs.* (1996) 11:76–87. doi: 10.1097/00005082-199610000-00009
10. Pheatt N, Brindis R, Levin E. Putting heart disease guidelines into practice: kaiser permanente leads the way. *Perman J.* (2003) 7:18–23. Available online at: https://www.thepermanentejournal.org/files/Winter2003/guides.pdf
11. Wakefield BJ, Drwal K, Paez M, Grover S, Franciscus C, Reisinger HS, et al. Creating and disseminating a home-based cardiac rehabilitation program: experience from the veterans health administration. *BMC Cardiovasc Disord.* (2019) 19:242. doi: 10.1186/s12872-019-1224-y
12. DeBusk RF, Miller NH, Superko HR, Dennis CA, Thomas RJ, Lew HT, et al. A case-management system for coronary risk factor modification after acute myocardial infarction. *J Cardiopulm Rehabil.* (1994) 14:407–8. doi: 10.1097/00008483-199411000-00011
13. Fonarow GC, Gawlinski A, Moughrabi S, Tillisch JH. Improved treatment of coronary heart disease by implementation of a cardiac hospitalization atherosclerosis management program (CHAMP). *Am J Cardiol.* (2001) 87:819–22. doi: 10.1016/S0002-9149(00)01519-8
14. Anderson L, Sharp GA, Norton RJ, Dalal H, Dean SG, Jolly K, et al. Home-based versus centre-based cardiac rehabilitation. *Cochrane Database Syst Rev.* (2015) 6:CD007130. doi: 10.1002/14651858.CD007130.pub4
15. Landis D, Georgiou A, Apple J, Durand J. Cardiovascular risk reduction at unitedhealthcare of north carolina: the first 12 months. *J Clin Outcom Manag.* (2000) 7:40–43.
16. Levin EG, Coll SV, Dlott R. Success of a home-based cardiac rehabilitation program in a population of nearly 4 million. *J Am Coll Cardiol.* (2017) 69:1731. doi: 10.1016/S0735-1097(17)35120-3
17. West JA, Miller NH, Parker KM, Senneca D, Ghandour G, Clark M, et al. A comprehensive management system for heart failure improves clinical outcomes and reduces medical resource utilization. *Am J Cardiol.* (1997) 79:58–63. doi: 10.1016/S0002-9149(96)00676-5
18. Ritchey MD, Maresh S, McNeely J, Shaffer T, Jackson SL, Keteyian SJ, et al. Tracking cardiac rehabilitation participation and completion among medicare beneficiaries to inform the efforts of a national initiative. *Circ Cardiovasc Qual Outcomes.* (2020) 13:1–11. doi: 10.1161/CIRCOUTCOMES.119.005902
19. *Cardiac Rehabilitation and Implications During the COVID-19 Era*. Available online at: https://www.acc.org/latest-in-cardiology/articles/2021/01/04/14/03/cardiac-rehabilitation-and-implications-during-the-covid-19-era (accessed May 25, 2020).
20. Dendale P, Berger J, Hansen D, Vaes J, Benit E, Weymans M. Cardiac rehabilitation reduces the rate of major adverse cardiac events after percutaneous coronary intervention. *Eur J Cardiovasc Nurs.* (2005) 4:113–16. doi: 10.1016/j.ejcnurse.2004.11.003
21. Lisspers J, Sundin Ö, Öhman A, Hofman-Bang C, Rydén L, Nygren Å. Long-term effects of lifestyle behavior change in coronary artery disease: effects on recurrent coronary events after percutaneous coronary intervention. *Health Psychol.* (2005) 24:41–8. doi: 10.1037/0278-6133.24.1.41
22. Shields GE, Wells A, Doherty P, Heagerty A, Buck D, Davies LM. Cost-Effectiveness of cardiac rehabilitation: a systematic review. *Heart.* (2018) 104:1403–10. doi: 10.1136/heartjnl-2017-312809
23. Johnson DA, Sacrinty MT, Gomadam PS, Mehta HJ, Brady MM, Douglas CJ, et al. Effect of early enrollment on outcomes in cardiac rehabilitation. *Am J Cardiol.* (2014) 114:1908–11. doi: 10.1016/j.amjcard.2014.09.036
24. Marzolini S, Blanchard C, Alter DA, Grace SL, Oh PI. Delays in referral and enrolment are associated with mitigated benefits of cardiac rehabilitation after coronary artery bypass surgery. *Circulation.* (2015) 8:608–20. doi: 10.1161/CIRCOUTCOMES.115.001751
25. Babu AS, Maiya AG, George MM, Padmakumar R, Guddattu V. Effects of combined early in-patient cardiac rehabilitation and structured home-based program on function among patients with congestive heart failure: a randomized ontrolled trial. *Heart Views.* (2011) 12:99. doi: 10.4103/1995-705X.95064

Conflict of Interest: AH, AA, and HV report ownership interest in Moving Analytics. AA and HV are senior officers and receive salary support from Moving Analytics.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Digital Health

MINI REVIEW
published: 23 December 2021
doi: 10.3389/fdgth.2021.637944

COVID-19 Prognostic Models: A Pro-con Debate for Machine Learning vs. Traditional Statistics

Ahmed Al-Hindawi[1†], *Ahmed Abdulaal*[2†], *Timothy M. Rawson*[3,4], *Saleh A. Alqahtani*[5,6], *Nabeela Mughal*[1,2,7] *and Luke S. P. Moore*[1,2,7*]

[1] Chelsea and Westminster NHS Foundation Trust, London, United Kingdom, [2] Faculty of Medicine, Imperial College London, London, United Kingdom, [3] Health Protection Research Unit for Healthcare Associated Infections and Antimicrobial Resistance, Imperial College London, London, United Kingdom, [4] Centre for Antimicrobial Optimisation, Imperial College London, London, United Kingdom, [5] King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia, [6] Johns Hopkins University, Baltimore, MD, United States, [7] North West London Pathology, Imperial College Healthcare NHS Trust, London, United Kingdom

The SARS-CoV-2 virus, which causes the COVID-19 pandemic, has had an unprecedented impact on healthcare requiring multidisciplinary innovation and novel thinking to minimize impact and improve outcomes. Wide-ranging disciplines have collaborated including diverse clinicians (radiology, microbiology, and critical care), who are working increasingly closely with data-science. This has been leveraged through the democratization of data-science with the increasing availability of easy to access open datasets, tutorials, programming languages, and hardware which makes it significantly easier to create mathematical models. To address the COVID-19 pandemic, such data-science has enabled modeling of the impact of the virus on the population and individuals for diagnostic, prognostic, and epidemiological ends. This has led to two large systematic reviews on this topic that have highlighted the two different ways in which this feat has been attempted: one using classical statistics and the other using more novel machine learning techniques. In this review, we debate the relative strengths and weaknesses of each method toward the specific task of predicting COVID-19 outcomes.

Keywords: COVID-19, Coronavirus, machine learning, artificial intelligence, linear regression

OPEN ACCESS

Edited by:
Wendy Chapman,
The University of Melbourne, Australia

Reviewed by:
Jian Guo,
RIKEN Center for Computational Science, Japan
Alec Chapman,
The University of Utah, United States

****Correspondence:***
Luke S. P. Moore
l.moore@imperial.ac.uk

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Health Informatics, a section of the journal Frontiers in Digital Health

Received: *04 December 2020*
Accepted: *15 November 2021*
Published: *23 December 2021*

Citation:
Al-Hindawi A, Abdulaal A, Rawson TM, Alqahtani SA, Mughal N and Moore LSP (2021) COVID-19 Prognostic Models: A Pro-con Debate for Machine Learning vs. Traditional Statistics. Front. Digit. Health 3:637944. doi: 10.3389/fdgth.2021.637944

INTRODUCTION

The novel coronavirus SARS-CoV-2 (COVID-19) has placed a significant strain on global healthcare systems. A particular challenge for COVID-19 is the difficulty in predicting individuals who will progress from a viral upper respiratory tract infection to more severe complications (including a dysregulated host response, coinfections, or thrombotic complications). Patients who progress often require critical care and are at significant risk of mortality. With the emergence of potential treatments for both the viral and inflammatory phases of COVID-19, the ability to predict those at high risk and deliver appropriate, prompt therapy could have a significant impact on patient outcomes.

Yet, to help address these critical questions, there is an ever-increasing multimodal pool of "big-data", with clinical, physiological, radiological, and laboratory parameters to develop, test, and optimize our decision-making pathways. As we consider which input variables may have the greatest influence on patient outcomes, we have a range of techniques, both from classical statistics through to novel artificial intelligence techniques, which we can apply to our clinical questions.

Several prognostic models have already been developed and reported for COVID-19 (1). Many have been developed using traditional statistics, yet machine learning has also been applied to prognostication against a variety of different clinical outcomes (2–4). These machine-learning models bring together statistics and computational programming, with the aim of data analysis without the intrinsic biases inherent in human approaches (**Box 1**). Before the COVID-19 pandemic, the application of machine learning to infectious diseases had been gaining traction, but to date, very few machine learning programmes are used clinically for prediction and prognostication (5–7). In contrast, prognostic scoring systems developed using traditional statistical methods have been widely implemented in front-line healthcare, including for infectious diseases (8–10). Perhaps foremost among these classical statistics is linear regression, itself a precursor of supervised machine learning, where a model is trained on a set of data with known outcomes with the aim of using what this model learns to predict on data it has not seen before, thus providing clinical insights. What classical statistical methods perhaps lack, however, is flexibility in exploring "unknown" clinical associations, particularly useful in the context of emerging infections, a need that algorithms like neural networks may address.

To explore the potential strengths and weaknesses of both traditional statistical methods and machine learning, we present a pro-con debate looking at the current state of the art in these fields, in the context of the wider clinical need for COVID-19 prognostication.

IN DEFENSE OF CLASSICAL STATISTICS

Classical Statistics Are the Foundation of Evidence-Based Medicine

The artificial intelligence (AI) "revolution" in healthcare continues to be promulgated in both scientific and consumer media; yet few, if any, of these innovations have been adopted in day-to-day clinical medicine. Meanwhile, linear regression models like the Acute Physiology and Chronic Health Evaluation (APACHE) score in Intensive Care, CURB-65 score for pneumonia, and the Model of End-Stage Liver Disease (MELD) are in daily clinical use and influence decision making across the globe (8–10). Linear regression models underpin these prognostic scores, acting as the foundation of evidence-based medicine randomized controlled trials. Therefore, before AI techniques are adopted at a large scale into clinical prognostication, we must consider in some detail how they compete with or are perhaps synergistic with, classical statistical techniques.

Neural Networks Are Opaque and Obfuscate

One of the more recent AI methods to challenge classical statistics has been the resurgence of an approach termed neural networks (**Box 1**). Neural nets have been investigated for use in clinical medicine since 1976 but suffered a lull due to computational restrictions (11). The recent renewed clinical interest in neural networks has been heralded by the development of convolutional neural networks (CNN) with the concurrent optimisation of matrix multiplication on graphics processing units (GPU), leading to fast training times and faster inference on easy and cheap to acquire hardware. The development of programming frameworks has reduced the barrier of entry for the experimentation in neural networks, leading to the democratization of this technology from what was once a difficult subject (12, 13).

The combination of readily available neural network programming frameworks, large curated clinical datasets, easy-to-learn programming languages, and CNNs have opened a wide window into the regression and classification of highly uncorrelated data, such as clinical radiographic images of computed tomography scans or X-rays (4). While such advances in AI seem potentially attractive, particularly for clinical prognostication, AI systems have been found to have learnt spuriously correlated data, such as a skin cancer classification neural network learning that the presence of a ruler in the image of the lesion accurately classified the presence of melanoma (14, 15).

Neural networks learn exquisite correlations between input variables and the output of interest. It can be argued that the above deficiencies of neural networks are secondary to faults in the dataset, but these faults are very hard to find. While the danger posed by this can be mitigated to some extent by supervision of systems, one might argue that this in some ways defeats the object of AI. Beyond this philosophical argument, in practical terms, supervised systems are difficult to clinically correlate as learnt latent (hidden) variables are difficult to interrogate, difficult to visualize, and impossible to prove coverage of data. In addition, it is currently not mathematically proven that new data entering a system is appropriately represented within the model's internal mechanisms, and reliance on cross-validation is a poor marker of this. It is believed that with sufficient "big data" the neural network may learn an implicit representation of its learning dataset to be sufficiently applied to out-of-sample data, but this is currently impossible to demonstrate, unlike regression models that have closed-form solutions to approximate out of sample performance. This leaves us back at the starting criticism of neural networks, where their hidden mechanics may provide outcomes we as clinicians think useful, but are based upon inputs with no plausible biological relevance.

Regression Analysis Can Provide Causality and Is Easily Interpretable

In contrast with classical statistical methods, such as multivariable regression, there are decades of research and validation, and when appropriately used can provide robust, simple yet genuine insights into clinical prognosis (8–10). Coefficients in classical multivariate regression have a literal translation, the bigger the coefficients the more important that variable is related to the outcome of interest. Negative coefficients are negatively correlated with the outcome of interest. This allows clinicians to tailor clinical decision-making based on the patient's

BOX 1 | Data-science tools potentially applicable to COVID-19 prognostication.

Artificial Intelligence (AI)	An overarching umbrella term used to denote software that demonstrates intelligence such as learning or problem solving.
Machine Learning (ML)	A specific subset of artificial intelligence that deals with the creation and validation of models that learn through experience, whether that is supervised (through informing the model of the correct answer during learning) or unsupervised.
Neural Network (NN)	A common framework that is inspired by the way neurons work in brains. A "neuron" receive inputs from other neurons, sums their outputs adds a bias element (and optionally normalizes the output to given range) and sends its output to another neuron. Useful in supervised tasks where the neuron's inner workings can be tuned to output a specific result given a set of inputs.
Deep learning/Deep neural networks (DL/DNN)	The finding that layering multiple neurons on-top of each other results in more accurate and precise neural networks.
Convolutional neural networks (CNN)	A subset of deep neural networks that use the convolutional operator as their basis for learning data features. These have revolutionized working with image and video datasets including the diagnosis of COVID-19 on radiography.
Graphics Processing Unit (GPU)	Historically used to render graphics in 3D intensive applications like computer games and computer aided design (CAD) where GPUs contain specific matrix multiplication machinery. This matrix multiplication machinery has been repurposed for General Processing on the GPU (GPGPU) leading to quick optimisation of neural networks and very efficient inference.
Matrix	An array of numbers arranged in a rectangle that can be together a single unit. In the training of ML techniques, these matrices typically represent the weights and biases of each neuron.
Tensor	A multidimensional array, similar to a multidimensional matrix, where each dimension would represent a different quality of the data. An example is a set of images batched together with a tensor of NxCxWxH where N is the number of images, C is the number of channels in the image, W is the width and H is the height of the image.

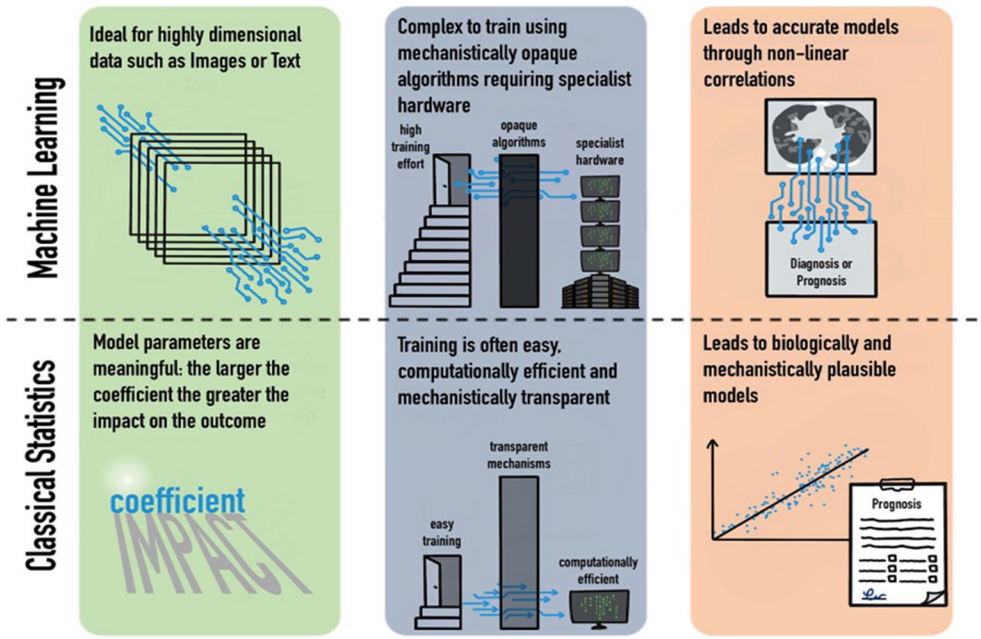


FIGURE 1 | Strengths and weaknesses of machine learning and classical statistics in their domains, training requirements, and outputs.

personal factors making precision, individualized, medicine a reality.

Finding out which variables are related to the outcome of interest from linear regression is inherent in their method, while neural network methods require multiple ablation studies to hint at which variables are correlated to an outcome. Training of linear regression is simple, and ordinary least squares is efficient, fast to train, and is mechanistically transparent. Multilevel, hierarchical, regression models have been successfully trained on tens of thousands of parameters and prior domain knowledge can be inserted into the models using Bayesian techniques (16, 17). Causality (rather than just correlation) can also be demonstrated using classical statistical methods through directed acyclic graphs, a big win if genuine knowledge of the world is required rather than just improved accuracy performance (18).

Classical Statistics Have Direct Applicability in COVID-19 Prognostication

For COVID-19, many publications have used neural networks to claim unprecedented accuracy for the prediction and

Table 1 | Uses and strengths of classical statistics vs. machine learning in COVID-19 prognostic modeling.

Classical statistics	Machine learning
In active clinical use	Useful for non-correlated data, like text or images
Foundation of modern medicine	Increasing in use across all areas of medicine
Mechanistically transparent	Requires little data pre-engineering
Easy to interrogate and can provide causality	Explainability is an active area of research with SHAP values explaining per patient predictions
Best non-biased estimator of mortality for COVID-19	Provided state of the art prognostic models across many domains

classification of COVID-19 outcomes including mortality, ICU admission, and length of stay (2–4, 19). Such an explosion of models has led to the publication of two "living" systematic reviews of those models, the findings of which have been pretty clear: many of those models exhibit "high bias" and are of little clinical use (1, 20). Gupta et al. applied many of the published models to their COVID-19 patient data, highlighting that the best performing, non-biased, model is a simple, well-specified, linear regression composed of age and oxygen saturations alone (20). This makes sense with clinical intuition, where older patients with COVID-19 have higher mortality, and patients who present in worse respiratory failure, as evidenced by lower hemoglobin oxygen saturation (SpO_2), also have higher mortality. This is not ground-breaking, but it is a transparent finding, which proves that clinical intuition is biologically plausible and is mechanistically probable. While such a simple prognostic model may not add to our understanding, it does perhaps allow us to finesse our pathways and risk stratification more efficiently care for patients when our healthcare services are at near-maximal capacity.

The Machine Learning Revolution Is Inevitable

Why Is Machine Learning So Powerful?

Consider a computerized tomography (CT) scan of the chest for a patient with COVID-19. The principal finding will be atypical or organizing pneumonia in up to 97% of patients with a severe infection (21–24). However, the images produced by the CT scanner are large, highly dimensional images, and therefore the data within them must be highly structured in some way so as to represent organizing pneumonia, and not random noise, or indeed a picture of something else.

The manifold hypothesis aims to explain this phenomenon. It posits that natural data lies on a low-dimensional manifold within the high-dimensional space where it is encoded (25). In other words, data pertaining to a particular class (for example, CT images of the chest) are a highly structured subset of all possible inputs for that class (i.e., all possible images/pixel values which can exist in the same size of image). This means that machine learning algorithms only need to learn a few key features from the data to be effective. This is analogous to physicians carefully picking a few important variables in multivariable regression analysis to answer a particular research question. The key difference is that the best possible features from any given highly dimensional dataset may turn out to be complicated functions of the original variables. The function of machine learning algorithms is to find these complex key features within a forest of data, which is a task that is not possible with classical statistical techniques.

Minimizing Bias While Maximizing Data Utilization

Bias, defined as a feature of a statistical technique or of its results whereby the expected value of the results differs from the true underlying quantitative parameter being estimated, is of paramount importance during all phases of model development, including training and validation. Christodolou et al. conducted a metaregression analysis that failed to demonstrate the improved discriminative performance of machine learning algorithms over logistic regression for clinical prediction models (6). While the area under the receiver operating curve (AUC) was on average no different between the two techniques when comparisons had a low risk of bias, machine-learning algorithms had improved performance among studies where there was a higher risk of bias, a potential advantage of machine learning algorithms over human-led statistical analysis. However, the systematic review was unable to report on measures of calibration due to poor reporting of this metric in the studies considered. There is a clear need therefore that future machine learning prognostic studies report calibration metrics and include a full report of all modeling steps, with particular adherence to the TRIPOD guidelines (26).

Predictive models in healthcare that utilize large datasets and a large number of parameters have demonstrated improved performance with machine-learning algorithms. A predictive model designed to forecast the development of acute kidney injury (AKI) analyzed data from 703,782 adults across 172 inpatient and 1,062 outpatient sites and considered 3,599 clinically relevant features that were provided to the baseline at each step (27). In all stages of AKI, classical logistic regression yielded lower precision-recall and receiver operator areas under the curve (PR AUC and ROC AUC, respectively) than Random Forest and Gradient Boosted Trees, which themselves yielded lower PR AUCs and ROC AUCs than deep learning approaches, such as intersection recurrent neural networks and long-short-term-memory networks (27).

In a systematic review and critical appraisal of current predictive models for COVID-19, Wynants et al. noted that all the 145 predictive models considered were at some risk of bias for a variety of reasons, ranging from lack of accounting for censoring (leading to selection bias), to using small sample sizes and subjective variables, and not reporting on calibration measures. They echo the importance of using the TRIPOD guidelines in future predictive work (1). When using the TRIPOD guidelines to develop statistical and machine learning predictive models for COVID-19 prognosis, including the use of Cox regression analysis to account for censoring, reporting the validation, discrimination, and calibration of both techniques;

and comparing both model ROC AUCs on the same dataset, there is evidence that machine-learning techniques outperform classical methods, even in moderately sized datasets (19).

Machine Learning; Not Quite as Opaque as Initially Thought

A key advantage frequently attributed to classical regression analysis is that each variable in the regression is assigned a coefficient by the model. The direction and magnitude of this coefficient directly relate to the direction and magnitude of the association between the variable considered and the outcome investigated. In contrast, the renewed interest in neural networks has been met with a steady criticism that such networks are non-transparent and that their predictions are not traceable by humans due to their multilayer, non-linear structure (28). However, explainable deep learning has recently become an active area of intense research which has produced three principled branches of explanatory methods, each with two subdivisions. Namely, visualization methods through perturbation or back-propagation, distillation methods through model translation or local approximation, and intrinsic techniques such as the use of attention mechanisms or joint training (29, 30).

Lundberg et al. utilized Shapley additive explanation, which is a variant of explanation through back-propagation work proposed by Shrikumar et al. which predicts near-term risk of hypoxaemia during anesthesia care, whilst explaining the patient- and surgery-specific factors leading to that risk in real-time (31, 32). Indeed, this technique can be applied to arbitrarily complex network architectures and has been used with success in deep learning prognostic models for COVID-19 to highlight salient patient characteristics leading to individual mortality predictions (2).

Optimizing Workflow Is Essential With Clinical "Big-Data"

Machine-learning algorithms can be easily implemented into end-to-end programmes capable of taking any desired data type as their input and producing relevant results (e.g., by scanning a dermatologic image through a phone app to produce a prediction of whether a skin lesion is malignant). While the important hazards of using inaccurate or potentially biased data cannot be overstated, such systems have nonetheless been able to outperform panels of expert specialists (33).

Furthermore, machine-learning algorithms can be used to predict multiple endpoints from a single feature set, which is difficult with classical statistical analysis. For example, Hofer et al. developed and validated a neural network from 59,981 surgical procedures capable of predicting postoperative mortality, AKI, and reintubation from a single feature set (34). Their model achieved a greater ROC AUC for their outcomes than the well-established ASA physical status score alone. This feature is particularly applicable to COVID-19, where predictive models need to be able to respond to changing management paradigms, changing outcomes, and evolving diseases complications.

CONCLUSION

There is little doubt that our ability to collect increasingly multimodal, highly dimensional clinical data will increase dramatically in the next few years, as typified by the formation and mandate of government bodies such as the United Kingdom's NHSX unit. Machine-learning techniques can produce models which are capable of utilizing a large array of multimodal data to produce multiple predictions simultaneously. This has been demonstrated by its promising use in the COVID-19 pandemic to produce ever more accurate predictions. However, the application of these complex models does not obviate the need for classical statistical analysis; causality and biological mechanistic plausibility remain in the realm of classical statistics (**Figure 1**; **Table 1**). Each technique has its merits, and blind application of either method has significant scientific ramifications.

AUTHOR CONTRIBUTIONS

AA-H and LM: conceptualization. AA-H, AA, TR, and LM: writing—original draft. AA-H, AA, TR, SA, NM, and LM: writing—review and editing. All the authors have read and approved this manuscript and agree as to its contents.

FUNDING

This work was funded by the Chelsea Infectious Diseases Research (CINDER) Fund at CW+ Charity.

ACKNOWLEDGMENTS

LM and TR acknowledge support from the National Institute of Health Research (NIHR) Imperial Biomedical Research Center (BRC) and the National Institute for Health Research HPRU in Healthcare Associated Infection and Antimicrobial Resistance [HPRU-2012-10047] at Imperial College London in partnership with Public Health England.

REFERENCES

1. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ*. (2020) 369:m1328. doi: 10.1136/bmj.m1328
2. Abdulaal A, Patel E, Charani S, Denny N, Mughal, Moore L. Prognostic modeling of COVID-19 using artificial intelligence in the united kingdom: model development and validation. *J Med Internet Res.* (2020) 22:e20259. doi: 10.2196/20259
3. Togaçar M, Ergen B, Cömert Z. COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Comput Biol Med*. (2020) 121:103805. doi: 10.1016/j.compbiomed.2020.103805
4. Ardakani A, Kanafi AR, Acharya UR, Khadem N, Mohammadi A. Application of deep learning technique to manage COVID-19 in routine clinical practice

using CT images: results of 10 convolutional neural networks. *Comput Biol Med*. (2020) 121:103795. doi: 10.1016/j.compbiomed.2020.103795
5. Peiffer-Smadja N, Dellière S, Rodriguez C, Birgand G, Lescure FX, Fourati S, et al. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? *Clin Microbiol Infect*. (2020) 26:1300–9. doi: 10.1016/j.cmi.2020.02.006
6. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. (2019) 110:12–22. doi: 10.1016/j.jclinepi.2019.02.004
7. Rawson TM, Hernandez B, Moore LSP, Blandy O, Herrero P, Gilchrist M, et al. Supervised machine learning for the prediction of infection on admission to hospital: a prospective observational cohort study. *J Antimicrob Chemother*. (2019) 74:1108–1115. doi: 10.1093/jac/dky514
8. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: A severity of disease classification system. *Crit Care Med*. (1985) 13:818–29.
9. Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology*. (2001) 33:464–70. doi: 10.1053/jhep.2001.22172
10. Lim WS. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax*. (2003) 58:377–82. doi: 10.1136/thorax.58.5.377
11. Mitchell T. Introduction to machine learning. *Mach Learn*. (1997) 7:2–5.
12. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. (2015) 19 arXiv:1603.04467v2. Available online at: https://arxiv.org/abs/1603.04467 (accessed December 3, 2020).
13. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: an imperative style, high-performance deep learning library. In: Wallach H, Larochelle H, Beygelzimer A, d Alché-Buc F, Fox E, Garnett R, editors. *Advances in Neural Information Processing Systems*. Curran Associates, Inc. (2019).vol. 32, pp. 8026–8037.
14. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. (2017 542:115–8. doi: 10.1038/nature21056
15. Narla B, Kuprel K, Sarin R, Novoa, Ko J. Automated classification of skin lesions: from pixels to practice. *J Invest Dermatol*. (2018) 138:2108–10. doi: 10.1016/j.jid.2018.06.175
16. Harris S, Singer M, Rowan K, Sanderson C. Delay to admission to critical care and mortality among deteriorating ward patients in UK hospitals: a multicentre, prospective, observational cohort study. *Lancet*. (2015_385:S40. doi: 10.1016/S0140-6736(15)60355-5
17. Zampieri FG, Damiani LP, Bakker J, Ospina-Tascón GA, Castro R, Cavalcanti AB, et al. Effects of a resuscitation strategy targeting peripheral perfusion status versus serum lactate levels among patients with septic shock. A Bayesian reanalysis of the ANDROMEDA-SHOCK Trial. *Am J Respir Crit Care Med*. (2020) 201:423–429. doi: 10.1164/rccm.201905-0968OC
18. Pearl J. *Causality: Models, Reasoning, and Inference*. Cambridge University Press (2009).
19. Abdulaal A, Patel A, Charani E, Denny S, Alqahtani SA, Davies GW, et al. Comparison of deep learning with regression analysis in creating predictive models for SARS-CoV-2 outcomes. *BMC Med Inform Decis Mak*. (2020) 20:299. doi: 10.1186/s12911-020-01316-6
20. Gupta RK, Marks M, Samuels THA, Luintel A, Rampling T, Chowdhury H, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. *Eur Respir J*. (2020) 56:2003498. doi: 10.1183/13993003.03498-2020
21. Kooraki S, Hosseiny M, Myers L, Gholamrezanezhad A. Coronavirus (COVID-19) outbreak: what the department of radiology should know. *J Am Coll Radiol*. (2020) 17:447–51. doi: 10.1016/j.jacr.2020.02.008
22. Cleverley J, Piper J, Jones MM. The role of chest radiography in confirming covid-19 pneumonia. *BMJ*. (2020) 370:m2426. doi: 10.1136/bmj.m2426
23. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*. (2020) 296:E32–40. doi: 10.1148/radiol.2020200642
24. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. (2020) 382:1708–20. doi: 10.1056/NEJMoa2002032
25. Fefferman C, Mitter S, Narayanan H. Testing the manifold hypothesis. *J Am Math Soc*. (2016) 29:983–1049. doi: 10.1090/jams/852
26. Collins Gary S, Reitsma Johannes B, Altman Douglas G, Moons Karel GM. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD)', *Circulation*. (2015) 131:211–219. doi: 10.1161/CIRCULATIONAHA.114.014508
27. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*. (2019) 57267:116–9. doi: 10.1038/s41586-019-1390-1
28. Benitez JM, Castro JL, Requena I. Are artificial neural networks black boxes? *IEEE Trans. Neural Netw*. (1997) 8:1156–64. doi: 10.1109/72.623216
29. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks *via* gradient-based localization. *Int. J. Comput. Vis*. (2020) 128:336–59. doi: 10.1007/s11263-019-01228-7
30. *Transparency by Design: Closing the Gap Between Performance and Interpretability in Visual Reasoning*. Available at: https://www.computer.org/csdl/proceedings-article/cvpr/2018/642000e942/17D45XeKgpp (accessed December 03, 2020).
31. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat. Biomed. Eng*. (2018) 2:749–60. doi: 10.1038/s41551-018-0304-0
32. Shrikumar A, Greenside P, Kundaje A. Learning important features through propagating activation differences. In: *Proceedings of the 34th International Conference on Machine Learning—Volume 70*. Sydney, NSW, Australia (2017). p. 3145–53 (Accessed November 14, 2020).
33. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol*. (2018) 29:1836–42. doi: 10.1093/annonc/mdy166
34. Hofer S, Lee C, Gabel E, Baldi P, Cannesson M. Development and validation of a deep neural network model to predict postoperative mortality, acute kidney injury, and reintubation using a single feature set. *NPJ Digit Med*. (2020) 3:58. doi: 10.1038/s41746-020-0248-0

Author Disclaimer: The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, or the UK Department of Health.

Conflict of Interest: All authors have completed ICMJE forms for Disclosure of Potential Conflicts of Interest and declare the following: NM has received speaker fees from Beyer (2016) and Pfizer (2019–2021) and received educational support from Eumedica (2016) and Baxter (2017). LM has consulted for/received speaker fees from DNAelectronics, (2015–18), Dairy Crest (2017–2018), Profile Pharma (2018–2019), bioMerieux (2013–2021), Eumedica (2016–2021), Umovis Lab (2020), Pfizer (2018–2021), Shionogi (2021), Pulmocide (2021), Sumitovant (2021), and Kent Pharma (2021) and received research grants from the National Institute for Health Research (2013–2019), Leo Pharma (2016), CW+ Charity (2018–2021), and LifeArc (2020–2021).

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 13 January 2022
doi: 10.3389/fdgth.2021.681608

Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU

Elham Jamshidi[1†], Amirhossein Asgary[2†], Nader Tavakoli[3†], Alireza Zali[1], Soroush Setareh[2], Hadi Esmaily[4], Seyed Hamid Jamaldini[5], Amir Daaee[6], Amirhesam Babajani[7], Mohammad Ali Sendani Kashi[8], Masoud Jamshidi[9], Sahand Jamal Rahi[10] and Nahal Mansouri[11,12*]*

[1] Functional Neurosurgery Research Center, Shohada Tajrish Comprehensive Neurosurgical Center of Excellence, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [2] Department of Biotechnology, College of Sciences, University of Tehran, Tehran, Iran, [3] Trauma and Injury Research Center, Iran University of Medical Sciences, Tehran, Iran, [4] Department of Clinical Pharmacy, School of Pharmacy, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [5] Department of Genetic, Faculty of Advanced Science and Technology, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran, [6] School of Mechanical Engineering, Sharif University of Technology, Tehran, Iran, [7] Department of Pharmacology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [8] Master of Business Administration (MBA)-University of Tehran, Tehran, Iran, [9] Department of Exercise Physiology, Tehran University, Tehran, Iran, [10] Swiss Institute for Experimental Cancer Research (ISREC), School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, [11] Division of Pulmonary Medicine, Department of Medicine, Lausanne University Hospital (CHUV), University of Lausanne (UNIL), Lausanne, Switzerland, [12] Laboratory of the Physics of Biological Systems, Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

OPEN ACCESS

Edited by:
Phuong N. Pham, Harvard Medical School, United States

Reviewed by:
Hao Wang, Shenzhen University General Hospital, China
Constantinos S. Pattichis, University of Cyprus, Cyprus

***Correspondence:**
Nahal Mansouri
nahal.mansouri@chuv.ch
Sahand Jamal Rahi
sahand.rahi@epfl.ch

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Health Informatics, a section of the journal Frontiers in Digital Health

Received: *16 March 2021*
Accepted: *22 December 2021*
Published: *13 January 2022*

Citation:
Jamshidi E, Asgary A, Tavakoli N, Zali A, Setareh S, Esmaily H, Jamaldini SH, Daaee A, Babajani A, Sendani Kashi MA, Jamshidi M, Jamal Rahi S and Mansouri N (2022) Using Machine Learning to Predict Mortality for COVID-19 Patients on Day 0 in the ICU. Front. Digit. Health 3:681608. doi: 10.3389/fdgth.2021.681608

Rationale: Given the expanding number of COVID-19 cases and the potential for new waves of infection, there is an urgent need for early prediction of the severity of the disease in intensive care unit (ICU) patients to optimize treatment strategies.

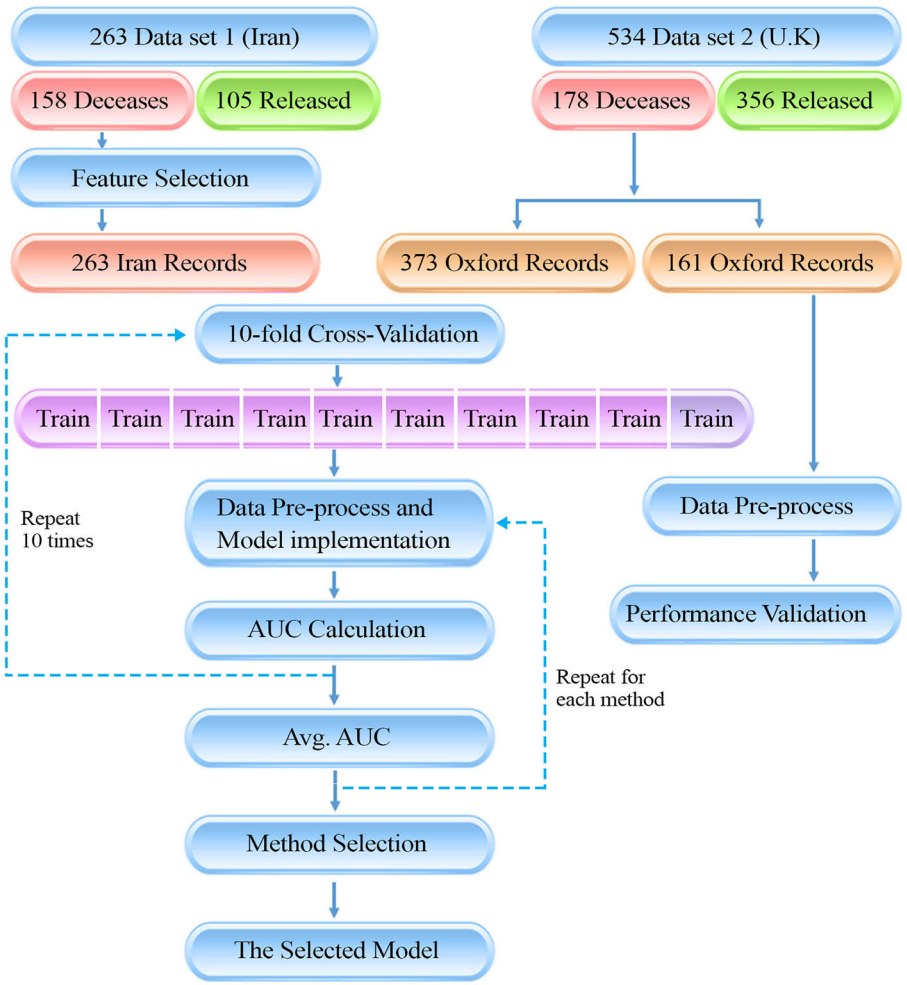
Objectives: Early prediction of mortality using machine learning based on typical laboratory results and clinical data registered on the day of ICU admission.

Methods: We retrospectively studied 797 patients diagnosed with COVID-19 in Iran and the United Kingdom (U.K.). To find parameters with the highest predictive values, Kolmogorov-Smirnov and Pearson chi-squared tests were used. Several machine learning algorithms, including Random Forest (RF), logistic regression, gradient boosting classifier, support vector machine classifier, and artificial neural network algorithms were utilized to build classification models. The impact of each marker on the RF model predictions was studied by implementing the local interpretable model-agnostic explanation technique (LIME-SP).

Results: Among 66 documented parameters, 15 factors with the highest predictive values were identified as follows: gender, age, blood urea nitrogen (BUN), creatinine, international normalized ratio (INR), albumin, mean corpuscular volume (MCV), white blood cell count, segmented neutrophil count, lymphocyte count, red cell distribution width (RDW), and mean cell hemoglobin (MCH) along with a history of neurological, cardiovascular, and respiratory disorders. Our RF model can predict patient outcomes with a sensitivity of 70% and a specificity of 75%. The performance of the models was confirmed by blindly testing the models in an external dataset.

Conclusions: Using two independent patient datasets, we designed a machine-learning-based model that could predict the risk of mortality from severe COVID-19 with high accuracy. The most decisive variables in our model were increased levels of BUN, lowered albumin levels, increased creatinine, INR, and RDW, along with gender and age. Considering the importance of early triage decisions, this model can be a useful tool in COVID-19 ICU decision-making.

Keywords: SARS-CoV-2, COVID-19, artificial intelligence, ICU—intensive care unit, machine learning (ML)



GRAPHICAL ABSTRACT | The presenting diagram, is showing the flow of our data gathering and method for the study. There are two data sources which ultimately have been used in a 10-fold cross-validation method to train the machine learning models. Finally, the model with the highest AUC was selected as final model.

INTRODUCTION

As of September 6, 2021, COVID-19 has caused more than 219 million infections worldwide and resulted in more than 4.55 million deaths. Complications are more common among elderly patients and people with preexisting conditions, and the rate of intensive care unit (ICU) admission is substantially higher in these groups (1, 2).

ICU admissions rely on the critical care capacity of the health care system. Iran, which is the primary testbed for this study, was one of the first countries hit by COVID-19. The ICU admission rate involves about 32% of all hospitalizations, and the ICU mortality rate is about 39% (3). With the potential of new waves of COVID-19 infections driven by more transmissible variants, ICU hospitalization numbers are expected to rise, leading to shortages of ICU beds and critical management equipment.

There is also the risk of a global shortage of effective medical supplies, making the judicious use of these medications a top priority for healthcare systems.

An individual-based prediction model is essential for tailoring treatment strategies and would aid in expanding our insights into the pathogenesis of COVID-19. A number of risk assessment scores are available to predict the severity of different diseases in ICU patients (4). Predictors of the need for intensive respiratory or vasopressor support in patients with COVID-19 and of mortality in COVID-19 patients with pneumonia have been identified (5, 6). To date, no general mortality prediction scores have been available for ICU admitted COVID-19 patients, irrespective of the patients' clinical presentation. Additionally, existing risk scales rely on parameters measured by health care providers such as blood pressure, respiratory rate, and oxygen saturation, which are subject to human error and operator bias especially under challenging and stressful conditions when numbers of COVID-19 patients surge (7). Thus, it remains vital to develop more unbiased risk-assessment tools that can predict the most likely outcomes for individual patients with COVID-19.

Recent advances in artificial intelligence (AI) technology for disease screening show promise as computer-aided diagnosis and prediction tools (8–11). In the era of COVID-19, AI has played an important role in early diagnosis of infection, contact tracing, and drug and vaccine development (12). Thus, AI represents a useful technology for the management of COVID-19 patients with the potential to help control the mortality rate of this disease. Nevertheless, an AI tool for making standardized and accurate predictions of outcomes in COVID-19 patients with severe disease is currently missing.

Beyond the general benefits of data-driven decision-making, the pandemic has also exposed the need for computational assistance to health care providers, who under the pressure of severely ill patients may make mistakes in judgment (7, 13, 14). Stressful conditions and burnout in health care providers can reduce their clinical performance, and a lack of accurate judgment can lead to increased mortality rates (15, 16). Artificial intelligence can help healthcare professionals determine who needs a critical level of care more precisely. Indeed, the effective use of AI could mitigate the severity of this outbreak.

Here, we propose a personalized machine-learning (ML) method for predicting mortality in COVID-19 patients based on routinely available laboratory and clinical data on the day of ICU admission.

Abbreviations: ACE2, Angiotensin-Converting Enzyme 2; AI, Artificial Intelligence; BUN, Blood Urea Nitrogen; COVID-19, coronavirus disease of 2019; CIC, clinical impact curve; Cr, creatinine; CRP, C reactive protein; DC, decision curve; ICU, Intensive care unit; INR, International Normalized Ratio; IFN, interferon; IL-6, Interleukin 6; IQR, interquartile range; KS, Kolmogorov- Smirnov; LR, Logistics regression; LIME, local interpretable model-agnostic explanation; LIME-SP, local interpretable model-agnostic explanation submodular-pick; ML, Machine learning; MCH, mean corpuscular hemoglobin; MCV, mean corpuscular volume; RF, Random forest; RDW, Red blood cell distribution width; ROC, receiver operating characteristic curve; RT-PCR, reverse transcription-polymerase chain reaction; WBC, white blood cells count.

METHODS

Data Resources

This is an international study involving patients from Iran (dataset 1) and the United Kingdom (U.K., dataset 2). We retrospectively studied 797 adult patients with severe COVID-19 infection confirmed through reverse transcription-polymerase chain reaction (RT-PCR). Two hundred sixty-three patients were admitted to ICUs at different hospitals in Tehran, Iran between February 19 and May 1, 2020, and 534 patients were admitted to ICUs and Emergency Assessment Units based on the Oxfordshire Research Database. The study was performed after approval by the Iran University of Medical Sciences Ethics Committee (approval ID: IR.IUMS.REC.1399.595).

Development of Mortality Prediction Model Using

The Mortality prediction model was aimed to predict whether patients were deceased or got released at the end of the admission period. Due to the generalizability and accessibility of the predictors recorded for patients in dataset 1 (Iran), and to reduce the model's feature space dimensionality, we merely used this dataset (consisting of 263 patients) for feature selection and model development. Only parameters with the highest predictive values were used in the modeling, leading to more robustness and generalizability of the model (17). Aside from that, further ML comparisons and validation was done with both the dataset 1 and 2.

Statistical Analysis and Feature Selection

On the day of the ICU admission, 66 parameters were assessed for each patient including 11 demographic characteristics (e.g., age and gender), past medical history and comorbidities (including nine different preexisting conditions), and 55 laboratory biomarkers. These parameters are listed in **Table 1**. Sixty-nine percent of measurements were reported on the day of admission, 27% were reported 1 day after, and 4% were reported within 2 days of ICU admission because of sampling limitations and laboratory practice. We excluded patients whose laboratory data were obtained more than 2 days after the date of admission to the ICU.

The aim was to predict a patient's survival. For the selection of parameters with the highest predictive value, under the null hypothesis of distributions being the same between the two groups, the two-sample Kolmogorov-Smirnov test (KS), shown in **Supplementary Figure 1**, was used for numerical parameters (age and laboratory biomarkers), and the Pearson chi-squared test (χ^2), shown in **Supplementary Figure 2**, was used for categorical parameters (e.g., gender and comorbidities).

All selected predictors were available in the second dataset (Oxfordshire, U.K.). Henceforth, the datasets have been merged, solely possessing previously selected predictors in common.

To investigate multicollinearity, Variance Inflation Factor (VIF) was calculated for each predictor and reported in **Supplementary Table 1**. A cut-off of 10 has been used to omit predictors that are showing collinearity, which includes none of the included predictors.

TABLE 1 | Machine learning methods with their parameters.

Method	Parameter	Value
Random forest	Number of trees	50
	Min. number of samples at a leaf node	0.1% of all samples
	Criterion	Gini
Logistic regression	C	1.0
Gradient boosting	Number of boosting stages to perform	10
	Fraction of samples used for fitting individual base learners	0.8
	Min. number of samples at a leaf node	10% of all samples
	Number of iterations with no change required for early stopping	3
	Max. number of features considered when looking for a split	3
Support vector machine	C	1.0
	Kernel type	RBF
	Kernel coefficient	1/number of features
Artificial neural network	Number of hidden layers	3
	Output space dimensionality for each hidden layer	32, 16, 8
	Activation function for each layer	Tanh, tanh, tanh, sigmoid

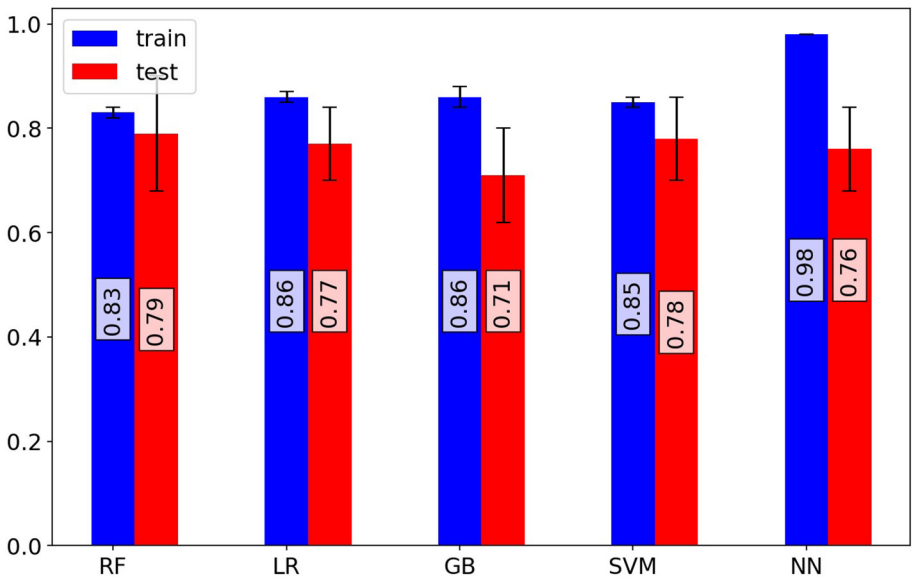


FIGURE 1 | Investigation of model performance. Mean area under the receiver operating characteristic curve (ROC-AUC) of random forest, logistic regression, gradient boosting classifier, support vector machine classifier, and artificial neural network models for training and test sets of cross-validation iterations. The random forest model shows superior performance on validation sets. The random forest model predicts patient outcomes with a 70% sensitivity and 75% specificity.

Data Preprocessing

Due to the difference in the measurement units and the necessity of units to be uniform, measurements of numerical parameters were unified between the two data sets by applying appropriate conversion factors, resulting in admissible input parameters for the model.

Data processing was carried out in four steps: First, because of incomplete laboratory data and in order to reduce difficulties associated with missing values, 771 patients out of the 797 total patients were selected as they had the data of at least 70 percent of all the biomarkers. Patients that did not have enough data present for biomarkers were removed. Second, samples were randomly separated into 10 independent sets with stratification over outcomes for 10-fold cross-validation to ensure the generalizability of the models (18). Of the 10 subsets, a single subset was retained as a validation set for model testing and the remaining nine subsets were used as training data. The cross-validation process was then iterated 10 times with each of the 10 subsets being used as the validation data exactly once. Third, numerical parameters were standardized by scaling the features to mean zero and unit variance. Last, missing biomarker values were imputed using the k-nearest neighbor (k-NN) algorithm,

FIGURE 2 | Feature importance in random forest model. The importance of the random forest features using local interpretable model-agnostic explanation submodular-pick with six submodules. Each submodule is related to a patient subpopulation (six subpopulation in this case) and represents decision criteria for them in the model. Negative values (blue) indicate favorable parameters suggesting a better prognosis, and positive values (red) indicate unfavorable parameters suggesting a worse prognosis.

and a binary indicator of missingness for each biomarker was added to the dataset (17, 19). Standardization and imputation were performed separately on each cross-validation iteration by using training set samples.

Machine Learning Model

Random forest (RF), logistic regression (LR), gradient boosting (GB), support vector machine (SVM), and artificial neural network (NN) methods were used to build classification models using the Python scikit-learn package. Methods along with their parameters are listed in **Table 1**. The performance of each method on training and validation sets in each cross-validation iteration was compared using a receiver operating characteristic curve (ROC), which is shown in **Supplementary Figure 3**. Area Under the Curve of ROC for each method is represented in **Figure 1**. Additional evaluation metrics for each model are also reported

TABLE 2 | Characteristics of intensive care unit patients with COVID-19 in our data.

	Survived ($N = 105$)		Died ($N = 158$)		Sig. key: <0.1 (*), <0.01 (**), <0.001 (***)			
	Number (%)	**Available data (%)**	**Number (%)**	**Available data (%)**	**Total (number)**	**X^2 statistics**	**X^2 *p*-value**	**Sig.**
Gender		105 (100)		158 (100)		1.70	0.19	
Male	63 (36.4)	..	110 (64.6)	..	173	..	..	
Female	42 (46.7)	..	48 (53.3)	..	90	..	..	
Comorbidity	..	105 (100)	..	158 (100)	..	..	..	
Autoimmune disorder	2 (33.3)	..	4 (66.7)	..	6	0.10	0.74	
Cancer	6 (42.9)	..	8 (57.1)	..	14	0.05	0.82	
Cardiovascular disorder	25 (29.1)	..	61 (70.9)	..	86	4.22	0.04	*
Diabetes mellitus	35 (38.0)	..	57 (62.0)	..	92	0.13	0.71	
Thrombosis	2 (40.0)	..	3 (60.0)	..	5	0.0003	0.99	
Hypertension	32 (34.0)	..	62 (66.0)	..	94	1.35	0.24	
Hepatic failure	2 (40.0)	..	3 (60.0)	..	5	0.0007	0.99	
Neurological disorder	8 (16.0)	..	42 (84.0)	..	50	11.93	<0.001	***
Respiratory disorder	7 (24.1)	..	22 (75.9)	..	29	3.01	0.08	*
	Median (IQR)	**Available data (%)**	**Median (IQR)**	**Available data (%)**	**Normal range**	**KS Statistics**	**KS *p*-value**	
Age (years)	58.0 (47.0–73.0)	105 (100)	72.5 (64.0–80.75)	158 (100)	..	0.35	<0.001	***
pH	7.42 (7.375–7.457)	87 (82)	7.4 (7.33–7.441)	129 (81)	7.31–7.41	0.18	0.05	*
pCO_2 (mm Hg)	38.4 (34.8–45.1)	87 (82)	40.2 (33.9–47.1)	125 (79)	35–40	0.09	0.66	
pO_2 (mm Hg)	37.05 (25.1–57.425)	86 (81)	39.9 (26.975–56.65)	124 (78)	42–51	0.08	0.81	
HCO_3 (meq·L)	25.5 (22.825–28.575)	86 (81)	24.2 (21.2–27.55)	123 (77)	22–26	0.15	0.14	
O_2 saturation (%)	72.7 (48.3–89.2)	85 (80)	73.5 (50.2–88.95)	123 (77)	−2.0 to 2.0	0.08	0.87	
Base excess (mEq/L)	2.2 (−0.55 to 4.65)	87 (82)	0.6 (−3.1 to 3.275)	126 (79)	..	0.18	0.06	*
Total buffer base (mEq/L)	49.1 (46.65–51.75)	87 (82)	47.5 (43.75–50.375)	126 (79)	..	0.20	0.01	*
Base excess in the extracellular fluid (mEq/L)	2.2 (−0.4 to 4.9)	87 (82)	0.35 (−3.175 to 3.75)	126 (79)	..	0.21	0.01	*
White blood cells count (x1000·mm^3)	7.4 (5.0–11.225)	104 (99)	9.7 (7.1–13.45)	155 (98)	4.0–10.0	0.23	0.002	**
Band (%)	3.0 (2.0–5.5)	23 (21)	3.0 (2.0–6.0)	38 (24)	..	0.05	1	
Segment (%)	78.0 (70.65–83.0)	87 (82)	82.8 (77.05–86.95)	119 (75)	..	0·25	0.002	**
Lymphocyte (%)	14.0 (10.0–20.225)	86 (81)	10.7 (6.85–15.4)	119 (75)	..	0.25	0.002	**
Monocyte (%)	6.0 (4.0–8.5)	45 (42)	5.0 (3.35–7.0)	59 (37)	..	0.17	0.34	
Basophil (%)	0.3 (0.2–0.8)	13 (12)	0.1 (0.0–0.1)	13 (8)	..	0.61	0.01	
Red blood cells count (mill·mm^3)	4.335 (3.83–4.908)	102 (97)	4.185 (3.64–4.748)	154 (97)	4.2–5.4	0.12	0.28	
Hemoglobin (g·dl)	12.6 (10.95–13.8)	103 (98)	12.2 (10.2–13.75)	155 (98)	12.0–16.0	0.07	0.81	
Hematocrite (%)	37.0 (32.85–41.2)	103 (98)	36.6 (31.45–40.75)	155 (98)	36–46	0.06	0.93	

(Continued)

TABLE 2 | Continued

	Survived (*N* = 105)		Died (*N* = 158)		Sig. key: <0.1 (*), <0.01 (**), <0.001 (***)			
	Number (%)	**Available data (%)**	**Number (%)**	**Available data (%)**	**Total (number)**	**X^2 statistics**	**X^2 *p*-value**	**Sig.**
Mean corpuscular volume (fL)	85.0 (81.4–88.65)	103 (98)	88.0 (84.65–91.9)	155 (98)	77–97	0.24	<0.001	***
Mean corpuscular hemoglobin (Pgm)	28.7 (26.6–29.85)	103 (98)	29.6 (27.8–30.55)	155 (98)	26–32	0.21	0.006	**
Mean corpuscular hemoglobin concentration (%)	33.1 (32.45–34.4)	103 (98)	33.3 (31.95–34.15)	155 (98)	32–36	0.09	0.59	
Platelet count (x1000·mm^3)	196.0 (151.5–260.0)	103 (98)	179.0 (125.0–255.0)	155 (98)	140–440	0.17	0.04	*
Red cell distribution width (%)	13.95 (13.2–14.825)	88 (83)	14.6 (13.75–16.0)	131 (82)	11.0–16.0	0.23	0.006	**
Platelet distribution width (FL)	12.8 (11.5–14.0)	85 (80)	13.2 (11.4–14.7)	120 (75)	10.0–17.0	0.13	0.32	
Mean platelet volume (FL)	9.7 (9.175–10.5)	84 (80)	10.0 (9.3–10.7)	120 (75)	8.5–12.5	0.13	0.30	
Platelet larger cell ratio (%)	24.4 (19.85–29.3)	83 (79)	26.7 (21.05–30.825)	120 (75)	17–45	0.17	0.07	*
C-reactive protein (mg·l)	48.0 (24.0–48.0)	56 (53)	48.0 (48.0–48.0)	67 (42)	<6	0.23	0.05	*
Erythrocyte sedimentation rate (mm · hr)	42.0 (27.5–68.5)	55 (52)	59.0 (33.75–75.25)	56 (35)	<20	0.24	0.06	*
Albumin level (g·dl)	3.3 (3.0–3.7)	48 (45)	2.9 (2.6–3.2)	71 (44)	3.5–5.5	0.37	<0.001	***
Serum calcium level (mg·dl)	8.8 (8.3–9.2)	72 (68)	8.6 (7.9–9.2)	97 (61)	8.6–10.6	0.13	0.37	
Inorganic P level (mg·dl)	3.3 (2.45–4.4)	59 (56)	4.0 (2.95–5.4)	87 (55)	2.5–5.0	0.20	0.08	
Serum Na level (mg·dl)	137.5 (135.0–140.0)	102 (97)	139.0 (135.0–142.0)	155 (98)	136–145	0.17	0.03	*
Serum K level (mg·dl)	4.3 (3.925–4.6)	102 (97)	4.4 (4.0–4.85)	155 (98)	3.7–5.5	0.11	0.37	
Serum Mg level (mg·dl)	2.25 (2.0–2.5)	66 (62)	2.4 (2.0–2.7)	96 (60)	1.8–2.6	0.14	0.32	
Uric acid level (mg·dl)	6.7 (4.05–9.0)	15 (14)	8.2 (5.95–9.95)	31 (19)	3.4–7.0	0.37	0.10	
Fasting plasma glucose (mg·dl)	124.0 (105.0–177.0)	65 (61)	154.0 (120.5–246.5)	99 (62)	..	0.21	0.04	*
Blood urea nitrogen (mg·dl)	16.0 (11.25–22.5)	102 (97)	30.0 (21.0–52.5)	156 (98)	5.0–23.0	0.47	<0.001	***
Creatinine (mg·dl)	1.1 (0.9–1.4)	102 (97)	1.5 (1.2–2.2)	156 (98)	0.5–1.5	0.31	<0.001	***
Aspartate aminotransferase (IU·L)	40.0 (29.0–55.0)	83 (79)	45.0 (31.5–82.5)	112 (70)	5.0–40.0	0.17	0.10	
Alanine aminotransferase (IU·L)	26.0 (16.0–38.5)	83 (79)	25.0 (18.0–45.0)	113 (71)	5.0–40.0	0.11	0.54	
Lactate dehydrogenase (U·L)	710.0 (561.0–1019.0)	57 (54)	859.0 (623.5–1256.0)	95 (60)	225–500	0.17	0.20	
Creatine phosphokinase (IU·L)	233.0 (89.0–546.5)	59 (56)	204.0 (83.0–434.0)	91 (57)	24–195	0.08	0.90	
Creatine phosphokinase-MB (U·L)	30.0 (22.5–41.0)	35 (33)	30.0 (24.0–49.0)	41 (25)	5–25	0.10	0.96	

(Continued)

TABLE 2 | Continued

	Survived ($N = 105$)		Died ($N = 158$)			Sig. key: <0.1 (*), <0.01 (**), <0.001 (***)		
	Number (%)	Available data (%)	Number (%)	Available data (%)	Total (number)	X^2 statistics	X^2 *p*-value	Sig.
Alkaline phosphatase (IU·L)	180.0 (132.5–248.5)	75 (71)	193.0 (155.75–264.25)	96 (60)	64–306	0.12	0.46	
Total bilirubin (mg·dl)	0.7 (0.5–0.9)	46 (43)	0.8 (0.6–1.45)	71 (44)	0.2–1.2	0.25	0.04	*
Direct bilirubin (mg·dl)	0.25 (0.2–0.3)	46 (43)	0.3 (0.2–0.6)	71 (44)	0–0.4	0.27	0.02	*
Prothrombin time (Sec)	14.9 (14.0–16.5)	81 (77)	16.0 (14.3–18.0)	121 (76)	12.0–13.0	0.23	0.009	**
Prothrombin time activity (%)	81.0 (72.0–89.0)	40 (38)	73.0 (54.5–85.5)	51 (32)	85–100	0.29	0.03	*
International normalized ratio (index)	1.2 (1.08–1.3)	81 (77)	1.3 (1.108–1.6)	120 (75)	1.0–1.1	0.29	<0.001	***
Partial thromboplastin time (Sec)	34.0 (30.0–40.0)	81 (77)	36.0 (31.0–45.75)	122 (77)	25–45	0.16	0.12	
D-Dimer (ng·ml)	1513.0 (1071.0–2207.0)	9 (8)	1875.0 (1558.0–5269.0)	11 (6)	..	0.28	0.68	
Interleukin-6 (pg·ml)	115.0 (76.0–147.25)	10 (9)	107.0 (47.0–301.0)	17 (10)	..	0.37	0.27	
Fibrin degeneration product (mg·L)	12.0 (12.0–12.0)	1 (0)	18.0 (18.0–18.0)	1 (0)	..	1	1	
Troponin (ng·L)	227.5 (78.375–1400.5)	12 (11)	49.0 (28.2–1027.0)	21 (13)	..	0.27	0.53	
Fibrinogen (mg·dl)	546.0 (477.5–727.0)	7 (6)	482.5 (226.75–615.75)	18 (11)	308–613	0.38	0.33	
Hemoglobin A1c (%)	7.2 (6.9–7.5)	5 (4)	6.4 (5.5–7.0)	9 (5)	..	0.55	0.22	

SIG. Key: * = <0.1, ** = <0.01, *** = <0.001.

in **Supplementary Table 2**. To prevent overfitting in the training process, the LR model was trained with an L2 regularization factor equal to one, and the RF was forced to hold more than 10% of samples in each of its terminal leaves (20, 21). The statistically significant difference between models' AUC curves has been affirmed by DeLong's test and corresponding DeLong's *p*-values assure the RF model's superiority and are shown in **Supplementary Figure 4**. To find the most influential parameters in the LR model prediction, we used regression coefficients, which are shown in the **Supplementary Figure 5**. Using the local interpretable model-agnostic explanation submodular-pick (LIME-SP) method, we identified different patterns among the whole feature space in the RF model (22). The LIME-SP method can interpret the model's predictions in different parts of the feature space by modeling a subset of model predictions in the feature space around the sample with the help of linear models that are more interpretable. In our study, LIME-SP was performed on 100 random samples to find six submodules with the most disparity in their selected markers, as shown in **Figure 2**. To identify meaningful clinical differences between patients, seven parameters with the highest predictive values were derived from each submodule.

Evaluation Criteria of the Model

To specify the evaluation dataset required for the validation of the model's performance, 30% of the records available in dataset 2 (the U.K, 161 patients; equal to 20% of the records) were randomly selected and assigned to the validation set to be used to blindly test the methods, and externally confirm the exactitude of the model. We have additionally included a data processing pipeline to summarize our methodology.

RESULTS

In dataset 1 (Iran), all the available patient records were used to train the models. The median age of patients was 69 years with an interquartile range (IQR) of 54–78. The minimum and maximum ages were 20 and 98 years, respectively. One hundred fifty-three patients (65.1%) were men, and 82 (34.9%) were women. One hundred five (39.9%) were discharged from the ICU after recovery and 158 (60.1%) patients died. The most frequent comorbidities among the patients were hypertension, diabetes, and cardiovascular disorders in 94, 92, and 86 patients, respectively. Among the 158 deceased patients, neurological disorders were the most prevalent comorbidity (42 patients, 84%). The statistical analysis and the availability of each parameter in our dataset are summarized in **Table 2**.

In the RF model, the optimum point between overfitting and efficiency was found by selecting 10 laboratory biomarkers out of 55 with the lowest KS *p*-values and three out of nine comorbidities with the lowest χ^2 *p*-values, besides demographic characteristics.

The selected numerical parameters for modeling were as follows: age, blood urea nitrogen (BUN), serum creatinine level (Cr), international normalized ratio (INR), serum albumin, mean corpuscular volume (MCV), red cell distribution width (RDW), mean corpuscular hemoglobin (MCH), white blood cell count

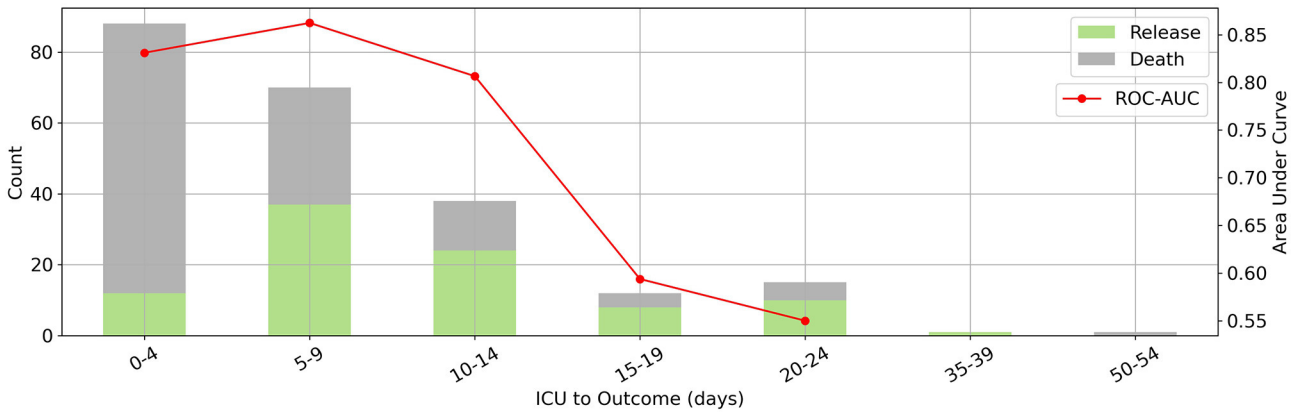


FIGURE 3 | The relation between prediction horizon and performance. Where x-axis denotes days from ICU admission to outcome. Distribution of days between intensive care unit admission and outcome (bars on the left vertical axis) and corresponding random forest model's area under the receiver operating characteristic curve scores for each bin (red line on the right vertical axis). Our model has the best performance to predict outcomes in a 15-day period.

(WBC), segmented neutrophil count, and lymphocyte count. In addition, selected categorical parameters were gender and a history of neurological, respiratory, and cardiovascular diseases. The distributions of selected numerical (age and biomarkers) and categorical (gender and preexisting conditions) variables are shown in **Supplementary Figures 6**, **7**, respectively.

Based on the ROC curves of the models (**Supplementary Figure 3**), the RF model outperformed other models and had superior efficiency. The higher efficiency of the RF model is also statistically significant in comparison to the other methods (**Supplementary Figure 4**). The better performance of RF could be explained by the complexity of the effects of COVID-19 and the varied etiologies underlying the deterioration of COVID-19 patients, for which the non-linear characteristics of the RF model was a more suitable option for predictions than the linear LR model. The RF model could predict a patient's outcome with a sensitivity of 70% and a specificity of 75%, whereas the sensitivity for the LR model was 65% and the specificity was 70%. Evaluation metrics for the models were also confirmed by the metrics reported as the results of the validating models.

By using the LIME technique, variables that provide the most information on the probability of each patient's death were identified. Among the six submodules identified with the highest disparity among 100 patients, albumin, BUN, and RDW were present in five of them. Age, MCH, and creatinine were present in four of the abovementioned submodules. This points out the importance of these measurements in the recorded parameters. Additionally, BUN (in three of these submodules), RDW (in two submodules), and age (in one submodule) were the most decisive ones.

This model could predict a patient's outcome reliably (AUC between 80 and 85) over a 15-day period, as shown in **Figure 3**. The mortality rate was highest between zero and 4 days. Given that the model was designed for first-day ICU admissions, moving away from this day reduced the accuracy of the predictions and the efficacy of the LIME method for clinical interventions, as expected.

To evaluate the clinical capability of the model, the decision curve (DC) and the clinical impact curve (CIC) were investigated (23). The DC framework measures the clinical "net benefit" for the prediction model relative to the current treatment strategy for all or no patients. The net benefit is measured over a spectrum of threshold probabilities, defined as the minimum disease risk at which further intervention is required. Based on the DC, CIC, and on the assumption of the same interventions for high-risk patients, our model indicated a superior or equal net benefit within a wide range of risk thresholds and patient outcomes, as shown in **Figure 4**.

Validation of the Model

In order to validate the performance of the model, similar records for 161 patients admitted to ICUs and Emergency Assessment Units were studied to externally confirm the prediction model (from dataset 2, U.K. cohort; see graphical abstract). The same Data preprocessing routine was applied to the additional validation data and ML methods with the same parameters as mentioned in **Table 1** were implemented. Models were blindly tested with the external validation data. Evaluation metrics for models are reported in **Supplementary Table 3**. Reported evaluation metrics indicate a 70% sensitivity for the RF model which accredits the certitude of the model. Validation results ensure the generalizability of the model and guarantee it's applicability for external data containing similar, globally accessible features.

DISCUSSION

The aim of this study was to develop an interpretable ML model to predict the mortality rate of COVID-19 patients at the time of admission to the ICU. To the best of our knowledge, this is the first study to develop a predictive model of mortality in patients

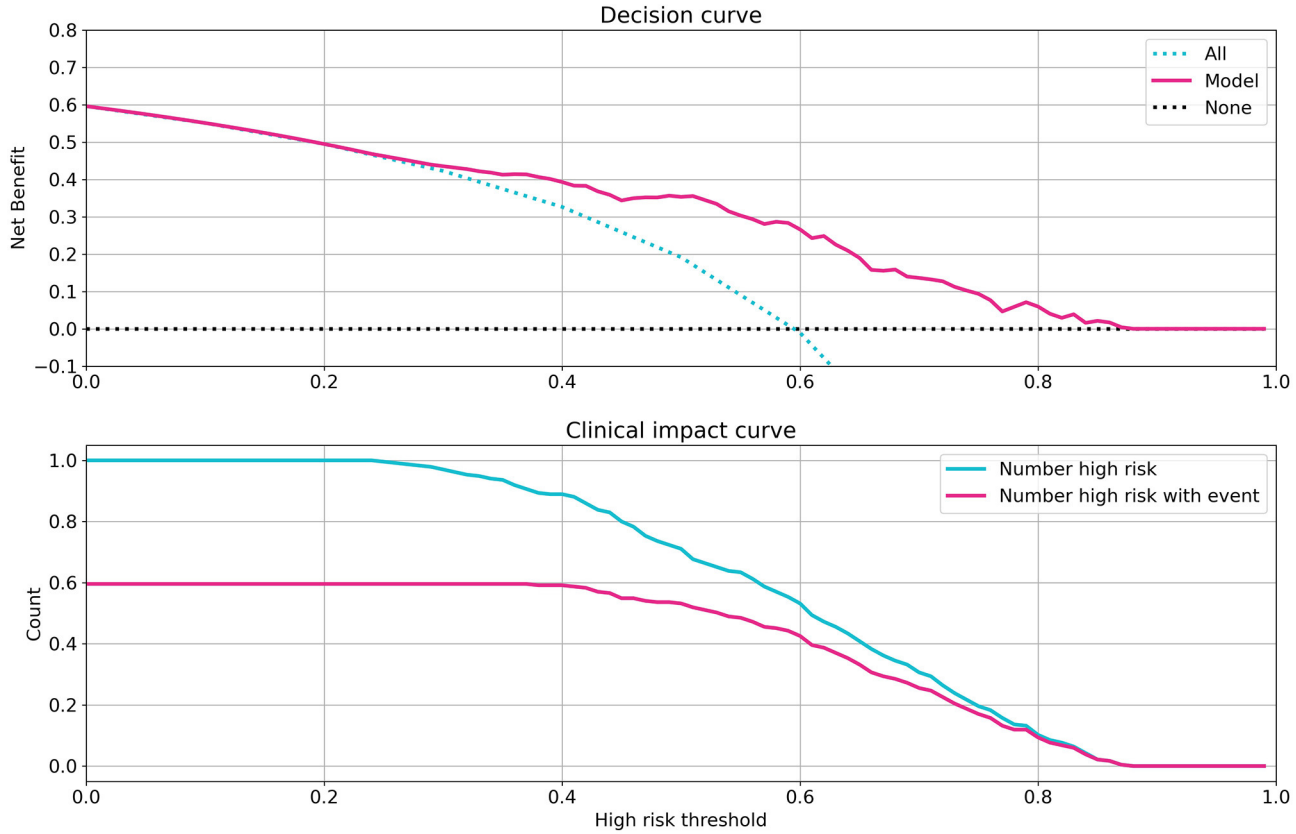


FIGURE 4 | Investigation of clinical impacts and benefits of the model. Decision curve **(Top)** and clinical impact curve **(Bottom)** of the random forest model. The decision curve compares the net benefits of an intervention in three scenarios: intervention for all patients (blue dotted line), intervention for no patients (gray dotted line), and intervention for high-risk patients based on the model prediction (red line). The clinical impact curve compares the number of patients classified as high risk by model and the number of patients with a really poor bad outcome who were classified as high risk, for all possible high-risk thresholds in model prediction from 0 to 1.

with severe COVID-19 infection at such an early stage using routine laboratory results and demographic characteristics.

Statistical analysis and feature selection tasks were performed merely by considering patients in dataset 1 (Iran dataset), which includes routine laboratory results, past medical histories and demographic characteristics, leading to selection among accessible and measurable predictors.

The most decisive parameters based on the two-sample KS test were, in decreasing order of importance, increased BUN, Cr, INR, MCV, WBC, segmented neutrophils count, RDW, MCH, and decreased albumin and lymphocyte levels. Moreover, based on a χ^2-test, age, gender, and a history of neurological, cardiovascular, and respiratory disorders were identified as parameters with high predictive values. Multicollinearity might affect the performance of the models and result in redundancy. Hence, variance inflation factor was calculated to find and remove highly correlated predictors. Selected predictors along with their references in the literature are listed in **Table 3**.

A number of studies have investigated the risk factors affecting COVID-19 infections (34, 35). Elevated inflammatory cytokines such as interleukin-6 (IL-6), granulocyte colony-stimulating factor (G-CSF), interferon gamma-induced protein 10 (IP-10), and interferon (IFN)-γ have been proposed as poor prognostic factors for COVID-19 patients (36–39). These markers, however, are not usually used as predictors of the severity of disease in clinical practice. Although using these cytokines in modeling may enable a more accurate prediction of the severity of COVID-19 infection, doing so impedes the model's clinical application, as most of the cytokines are not routinely checked at presentation to the ICU. In contrast, all 10 laboratory biomarkers identified in our model are commonly measured and are available to most clinical laboratories. Thus, the DC and CIC analyses indicated the notable clinical benefit of our model especially in a situation characterized by resource scarcity.

Only patients who had at least seven of the 10 selected biomarkers have been included in the training phase of the modeling and missing parameters were imputed using k-NN based on the data. As can be seen in the models' ROC curve, the RF algorithm outperformed other methods in predicting the outcome. Significance of this difference has been investigated using DeLong's test. Superior proficiency of the RF model is mainly due to the non-linear correlation between variables, manifesting the complexity of the problem.

TABLE 3 | Predictors with the highest predictive value, selected in this study, along with studies referring to them.

Predictor	Description	Literature references
Gender	Sex-dependent differences in clinical manifestation	(24, 25)
Age	Higher age affects COVID-19 poor outcomes	(25, 26)
Blood Urea Nitrogen	Assumed highest weights for prognosis	(27, 28)
Creatinine	A lower creatinine clearance levels increases the mortality	(29)
INR	INR >1.3 significantly increases mortality	(30)
Albumin	Assumed highest weights for prognosis	(27)
WBC	Abnormal white blood cell count increases mortality	(31)
Neutrophil count	affects COVID-19 poor outcomes	(32)
Lymphocyte count	Lymphocytes <10% increases mortality	(30)
RDW	RDW >14.5% increases mortality	(30)
MCH	Abnormal MCH increases mortality	(33)
Neurological disorders	Affects the COVID-19 outcome	(34)
Cardiovascular disorders	Affects the COVID-19 outcome	(34)
Respiratory disorders	Affects the COVID-19 outcome	(34)

Since a part of the data itself has been used for feature selection and a 10-fold cross-validation algorithm has been implemented to the data, an additional external validation was conducted to confirm the model's performance. Models were blindly tested with validation data, including records of measurements of the selected predictors for 161 patients, taken out of dataset 2 (U.K.).

Results of the validation assure that the model we developed could be applied globally and predict mortality of the patients with severe COVID-19 infection solely with universally accessible parameters (**Table 2**). As a result, physicians and healthcare systems are able to utilize this model, confident about high sensitivity and specificity in the outcome.

The application of the LIME-SM method allowed us to determine a patient-specific marker set that each patient's prognosis is based on. This technique explains the predictions by perturbing the input of data samples and evaluating the effects. The output of LIME is a list of features, reflecting each feature's contribution to a given prediction. Understanding the "reasoning" of the ML model is crucial for increasing physicians' confidence in selecting treatments based on the prognosis scores. Using the LIME method, the significance of variables with high predictive value was determined for each prediction made for an individual. The evaluation of the variables in the individual's personalized prediction can lead to supportive measures and help determine treatment strategies according to the interpretation of the individual prognosis.

As severe COVID-19 may result from various underlying etiologies, our model can help categorize patients into groups with distinct clinical prognosis, thus allowing personalized treatments. In addition to targeted therapies, the differentiation between patients may reveal disease mechanisms that coincide or that occur under specific preexisting conditions. Future cohort studies could explore these assumptions with increased sample sizes.

In this study, hypoalbuminemia and renal function were identified as the main factors with high predictive values for the model. These findings are in agreement with recent results showing that hypoalbuminemia is an indicator of poor prognosis for COVID-19 patients (40). It is well-documented that endogenous albumin is the primary extracellular molecule responsible for regulating the plasma redox state among plasma antioxidants (40). Moreover, it has been shown that albumin downregulates the expression of the angiotensin-converting enzyme 2 (ACE2) which may explain the association of hypoalbuminemia with severe COVID-19 (41). Intravenous albumin therapy has been shown to improve multiple organ functions (42). Therefore, early treatment with human albumin in severe cases of COVID-19 patients before the drop in albumin levels might have positive outcomes and needs to be further investigated.

Furthermore, increased levels of BUN and Cr are observed in our study, which is an indication of kidney damage. An abrupt loss of kidney function in COVID-19 is strongly associated with increased mortality and morbidity (43). There are multiple mechanisms supporting this association (44, 45).

One of the findings of this study is the identification of RDW (a measure of the variability of the sizes of RBCs) as an influential parameter. This result is in line with recently published reports (46). Elevated RDW, known as anisocytosis, reflects a higher heterogeneity in erythrocyte sizes caused by erythrocyte maturation and degradation abnormalities. Several studies have found that elevated RDW is associated with inflammatory markers in the blood such as IL-6, tumor necrosis factor-α, and CRP, which is common in severely ill Covid-19 patients (44). These inflammatory markers could disrupt the erythropoiesis by directly suppressing erythroid precursors, promoting apoptosis of precursor cells, and reducing the bioavailability of iron for hemoglobin synthesis.

Yan et al. recently identified LDH, lymphocyte, and high-sensitivity C-reactive protein (hs-CRP) as predictors of mortality in COVID-19 patients during their hospitalization. The blood results of hospitalized patients on different days after the initial ICU admission were used for their model (45). Since our goal was the prediction of mortality risk as early as possible for ICU

patients, this limited us to using only the laboratory results on day 0, in contrast. For patients with severe COVID-19 infection, early decision-making is critical for successful clinical management. Additionally, laboratory results from other days may not always become available. We also identified lymphocyte count as a predictor of mortality, as in the previous study; however, CRP levels and LDH did not reach statistical significance.

Although IL-6 has been found to be a good predictor of disease severity by other studies, it did not reach statistical significance in our model (47). IL-6 had a considerable KS statistical value, but because of the high number of missing values, its *p*-value was not significant compared to other markers. The fact that IL-6 is not always measured upon ICU admission is precisely why it is not suitable for our purposes.

In similar studies the impact of laboratory values was assessed. Booth et al. recruited two ML techniques, LR and SVM to design a prediction model for COVID-19 severity among 26 parameters. They indicated CRP, BUN, serum calcium, serum albumin, and lactic acid as the top five highest-weighted laboratory values. Their analysis showed that the SVM model displayed 91% sensitivity and specificity (AUC 0.93) for predicting mortality (27). In another study, Guan et al. used an ML algorithm to predict COVID-19 mortality retrospectively. They showed that CRP, LDH, ferritin, and IL-10 were the most important death predictors with a sensitivity of 85% (48). Zoabi et al. developed a ML-based predicting model that evaluate eight binary features: sex, age, known contact with an infected person, and five initial clinical symptoms including headache, sore throat, cough, fever, and shortness of breath. They showed that their model can predict the COVID-19 infection with 87.30% sensitivity and 71.98% specificity (49).

The missingness indicator of some markers in both LR and RF models has an impact on the predictions based on the regression coefficient and LIME, which can be the result of the model compensating for the imputation error. However, the missingness indicator may also indicate the existence of bias in biomarker reporting (50). Such biases (e.g., sampling bias) are an inevitable part of retrospective studies. They can be addressed using domain-adaptation techniques such as correlation alignment (CORAL) in future studies using additional data (51, 52). Another limitation of this study may be the lack of an objective criterion for ICU admission. Moreover, different treatment strategies can change the survival outcome for patients who may have had similar profiles when admitted to the ICU. In future studies, the accuracy of this model may be further improved by adding chest imaging data and by using a larger dataset. Possible targets for our ML framework include the prediction of other crucial information such as the patients' need for mechanical ventilation, the occurrence of cytokine release syndrome, the severity of acute respiratory disease syndrome, the cause of death, and the right treatment strategy.

In conclusion, we evaluated 66 parameters in COVID-19 patients at the time of ICU admission. Of those parameters, 15 metrics with the highest prediction values were identified: gender, age, BUN, Cr, INR, albumin, MCV, RDW, MCH, WBC, segmented neutrophil count, lymphocyte count, and past medical history of neurological, respiratory, and cardiovascular disorders. In addition, by using the LIME-SP method, we identified different submodules clarifying distinct clinical manifestations of severe COVID-19. The ML model trained in this study could help clinicians determine rapidly which patients are likely to have worse outcomes, and given the limited resources and reliance on supportive care allow physicians to make more informed decisions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

EJ, AA, SaJ, NM, and NT: conceptualization, methodology, project administration, writing—original draft, writing—review, and editing. SS: methodology, project administration, writing—review, and editing. AZ, HE, SeJ, AD, AB, MS, and MJ: data curation, investigation, writing—original draft, writing—review, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

IORD was supported by the Oxford NIHR Biomedical Research Centre.

ACKNOWLEDGMENTS

This work uses data provided by patients and collected by the UK's National Health Service as part of their care and support. We thank all the people of Oxfordshire who contribute to the Infections in Oxfordshire Research Database. Research Database Team: L. Butcher, H. Boseley, C. Crichton, D. W. Crook, D. Eyre, O. Freeman, J. Gearing (community), R. Harrington, K. Jeffery, M. Landray, A. Pal, T. E. A. Peto, T. P. Quan, J. Robinson (community), J. Sellors, B. Shine, A. S. Walker, and D. Waller. Patient and Public Panel: G. Blower, C. Mancey, P. McLoughlin, and B. Nichols.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.681608/full#supplementary-material

REFERENCES

1. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA*. (2020) 323:2052–9. doi: 10.1001/jama.2020.6775
2. Liu K, Chen Y, Lin R, and Han K. (2020). Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *J Infec.* 80, e14–e18. doi: 10.1016/j.jinf.2020.03.005
3. Abate SM, Ahmed Ali S, Mantfardo B, Basu B. Rate of Intensive Care Unit admission and outcomes among patients with coronavirus: a systematic review and meta-analysis. *PLoS ONE.* (2020) 15:e0235653. doi: 10.1371/journal.pone.0235653
4. Rapsang AG, Shyam DC. Scoring systems in the intensive care unit: a compendium. *Indian J Crit Care Med.* (2014) 18:220. doi: 10.4103/0972-5229.130573
5. Fan G, Tu C, Zhou F, Liu Z, Wang Y, Song B, et al. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *Eur Respir J.* (2020) 56:2002113. doi: 10.1183/13993003.02113-2020
6. Su Y, Tu G-W, Ju M-J, Yu S-J, Zheng J-L, Ma G-G, et al. Comparison of CRB-65 and quick sepsis-related organ failure assessment for predicting the need for intensive respiratory or vasopressor support in patients with COVID-19. *J Infect.* (2020) 81:647–79. doi: 10.1016/j.jinf.2020.05.007
7. Hartzband P, Groopman J. Physician burnout, interrupted. *N Engl J Med.* (2020) 382:2485–7. doi: 10.1056/NEJMp2003149
8. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA*. (2018) 320:2199–200. doi: 10.1001/jama.2018.17163
9. Stead WW. Clinical implications and challenges of artificial intelligence and deep learning. *JAMA*. (2018) 320:1107–8. doi: 10.1001/jama.2018.11029
10. Ghafouri-Fard S, Taheri M, Omrani MD, Daaee A, Mohammad-Rahimi H, Kazazi H. Application of single-nucleotide polymorphisms in the diagnosis of autism spectrum disorders: a preliminary study with artificial neural networks. *J Mol Neurosci.* (2019) 68:515–21. doi: 10.1007/s12031-019-01311-1
11. Ghafouri-Fard S, Taheri M, Omrani MD, Daaee A, Mohammad-Rahimi H. Application of artificial neural network for prediction of risk of multiple sclerosis based on single nucleotide polymorphism genotypes. *J Mol Neurosci.* (2020) 70:1081–7. doi: 10.1007/s12031-020-01514-x
12. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab Syndr Clin Res Rev.* (2020) 14:337–9. doi: 10.1016/j.dsx.2020.04.012
13. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. *JAMA Netw Open.* (2020) 3:e203976. doi: 10.1001/jamanetworkopen.2020.3976
14. Kannampallil TG, Goss CW, Evanoff BA, Strickland JR, McAlister RP, Duncan J. Exposure to COVID-19 patients increases physician trainee stress and burnout. *PLoS ONE.* (2020) 15:e0237301. doi: 10.1371/journal.pone.0237301
15. Motluk A. Do doctors experiencing burnout make more errors? *Can Med Assoc*. (2018) 190:E1216–7. doi: 10.1503/cmaj.109-5663
16. Dyrbye LN, Awad KM, Fiscus LC, Sinsky CA, Shanafelt TD. Estimating the attributable cost of physician burnout in the United States. *Ann Int Med.* (2019) 171:600–1. doi: 10.7326/L19-0522
17. Mladenić D. Feature selection for dimensionality reduction. In: Saunders C, Grobelnik M, Gunn S, Shawe-Taylor J, editors. *International Statistical and Optimization Perspectives Workshop "Subspace, Latent Structure and Feature Selection,"* Vol. 3940. Berlin; Heidelberg: Springer (2005). p. 84–102. doi: 10.1007/11752790_5
18. Wong TT. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognit.* (2015) 48:2839–46. doi: 10.1016/j.patcog.2015.03.009
19. Pujianto U, Wibawa AP, Akbar MI. K-Nearest Neighbor (K-NN) based missing data imputation. In: *2019 5th International Conference on Science in Information Technology (ICSITech)*. Yogyakarta: IEEE (2019). p. 83–8.
20. Pedregosa F, Cauvet E, Varoquaux G, Pallier C, Thirion B, Gramfort A. Learning to rank from medical imaging data. In: *International Workshop on Machine Learning in Medical Imaging*. Nice: Springer (2012). p. 234–41. doi: 10.1007/978-3-642-35428-1_29
21. Ng AY. Feature selection, L 1 vs. L 2 regularization, rotational invariance. In: *Proceedings of the Twenty-First International Conference on Machine Learning*. (2004). p. 78. doi: 10.1145/1015330.1015435
22. Ribeiro MT, Singh S, Guestrin C. "Why should i trust you?" Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA (2016). p. 1135–44. doi: 10.1145/2939672.2939778
23. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* (2006) 26:565–74. doi: 10.1177/0272989X06295361
24. Ancochea J, Izquierdo JL, Soriano JB. Evidence of gender differences in the diagnosis and management of coronavirus disease 2019 patients: an analysis of electronic health records using natural language processing and machine learning. *J Womens Health.* (2021) 30:393–404. doi: 10.1089/jwh.2020.8721
25. Aktar S, Talukder A, Ahamad MM, Kamal AHM, Khan JR, Protikuzzaman M, et al. Machine learning approaches to identify patient comorbidities and symptoms that increased risk of mortality in COVID-19. *Diagnostics.* (2021) 11:1383. doi: 10.3390/diagnostics11081383
26. Li M, Zhang Z, Cao W, Liu Y, Du B, Chen C, et al. Identifying novel factors associated with COVID-19 transmission and fatality using the machine learning approach. *Sci Total Environ.* (2021) 764:142810. doi: 10.1016/j.scitotenv.2020.142810
27. Booth AL, Abels E, McCaffrey P. Development of a prognostic model for mortality in COVID-19 infection using machine learning. *Modern Pathol.* (2021) 34:522–31. doi: 10.1038/s41379-020-00700-x
28. Kang J, Chen T, Luo H, Luo Y, Du G, Jiming-Yang M. Machine learning predictive model for severe COVID-19. *Infect Genet Evol.* (2021) 90:104737. doi: 10.1016/j.meegid.2021.104737
29. Paris S, Inciardi RM, Specchia C, Vezzoli M, Oriecuia C, Lombardi CM, et al. 554 machine learning for prediction of in-hospital mortality in COVID-19 patients: results from an Italian multicentre study. *Eur Heart J Suppl.* (2021) 23. doi: 10.1093/eurheartj/suab135.035
30. Kar S, Chawla R, Haranath SP, Ramasubban S, Ramakrishnan N, Vaishya R, et al. Multivariable mortality risk prediction using machine learning for COVID-19 patients at admission (AICOVID). *Sci Rep.* (2021) 11:12801. doi: 10.1038/s41598-021-92146-7
31. Blagojević A, Šušteršič T, Lorencin I, Šegota SB, Milovanović D, Baskić D, Baskić D, Car Z, et al. Combined machine learning and finite element simulation approach towards personalized model for prognosis of COVID-19 disease development in patients. *EAI Endorsed Trans Bioeng Bioinform.* (2021) 1:e6. doi: 10.4108/eai.12-3-2021.169028
32. Ye J, Hua M, Zhu F. Machine learning algorithms are superior to conventional regression models in predicting risk stratification of COVID-19 patients. *Risk Manag Healthc Policy.* (2021) 14:3159–66. doi: 10.2147/RMHP.S318265
33. Alves MA, Castro GZ, Oliveira BAS, Ferreira LA, Ramírez JA, Silva R, et al. Explaining machine learning based diagnosis of COVID-19 from routine blood tests with decision trees and criteria graphs. *Comput Biol Med.* (2021) 132:104335. doi: 10.1016/j.compbiomed.2021.104335
34. Jamshidi E, Asgary A, Tavakoli N, Zali A, Dastan F, Daaee A, et al. Symptom prediction and mortality risk calculation for COVID-19 using machine learning. *medRxiv*. (2021). doi: 10.1101/2021.02.04.21251143
35. Jamshidi E, Asgary A, Tavakoli N, Zali A, Dastan F, Daaee A, et al. Symptom prediction and mortality risk calculation for COVID-19 using machine learning. *Front Artif Intell.* (2021) 4:673527. doi: 10.3389/frai.2021.673527
36. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ. COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet.* (2020) 395:1033–4. doi: 10.1016/S0140-6736(20)30628-0
37. Conti P, Ronconi G, Caraffa A, Gallenga C, Ross R, Frydas I, et al. Induction of pro-inflammatory cytokines (IL-1 and IL-6) and lung inflammation by coronavirus-19 (COVI-19 or SARS-CoV-2): anti-inflammatory strategies. *J Biol Regul Homeost Agents.* (2020) 34:327–31. doi: 10.23812/CONTI-E
38. Jamshidi E, Babajani A, Soltani P, Niknejad H. Proposed mechanisms of targeting COVID-19 by delivering mesenchymal stem cells and their exosomes to damaged organs. *Stem Cell Rev Rep.* (2021) 17:176–92. doi: 10.1007/s12015-020-10109-3
39. Babajani A, Hosseini-Monfared P, Abbaspour S, Jamshidi E, Niknejad H. Targeted mitochondrial therapy with over-expressed MAVS protein from

mesenchymal stem cells: a new therapeutic approach for COVID-19. *Front Cell Dev Biol.* (2021) 9:695362. doi: 10.3389/fcell.2021.695362
40. de la Rica R, Borges M, Aranda M, Del Castillo A, Socias A, Payeras A, et al. Low albumin levels are associated with poorer outcomes in a case series of COVID-19 patients in Spain: a retrospective cohort study. *Microorganisms.* (2020) 8:1106. doi: 10.3390/microorganisms8081106
41. Liu B-C, Gao J, Li Q, Xu L-M. Albumin caused the increasing production of angiotensin II due to the dysregulation of ACE/ACE2 expression in HK2 cells. *Clin Chim Acta.* (2009) 403:23–30. doi: 10.1016/j.cca.2008.12.015
42. Caironi P, Gattinoni L. The clinical use of albumin: the point of view of a specialist in intensive care. *Blood Transfus.* (2009) 7:259–67. doi: 10.2450/2009.0002-09
43. Cheng Y, Luo R, Wang K, Zhang M, Wang Z, Dong L, et al. Kidney impairment is associated with in-hospital death of COVID-19 patients. *medRxiv*. (2020). doi: 10.1101/2020.02.18.20023242
44. Agarwal S. Red cell distribution width, inflammatory markers and cardiorespiratory fitness: results from the National Health and Nutrition Examination Survey. *Indian Heart J.* (2012) 64:380–7. doi: 10.1016/j.ihj.2012.06.006
45. Yan L, Zhang H-T, Goncalves J, Xiao Y, Wang M, Guo Y, et al. An interpretable mortality prediction model for COVID-19 patients. *Nat Mach Intell.* (2020) 2:283–8. doi: 10.1038/s42256-020-0180-7
46. Foy BH, Carlson JC, Reinertsen E, Valls RP, Lopez RP, Palanques-Tost E, et al. Elevated RDW is associated with increased mortality risk in COVID-19. *medRxiv*. (2020). doi: 10.1101/2020.05.05.20091702
47. Herold T, Jurinovic V, Arnreich C, Lipworth BJ, Hellmuth JC, von Bergwelt-Baildon M, et al. Elevated levels of IL-6 and CRP predict the need for mechanical ventilation in COVID-19. *J Allergy Clin Immunol.* (2020) 146:128–36. e4. doi: 10.1016/j.jaci.2020.05.008
48. Guan X, Zhang B, Fu M, Li M, Yuan X, Zhu Y, et al. Clinical and inflammatory features based machine learning model for fatal risk prediction of hospitalized COVID-19 patients: results from a retrospective cohort study. *Ann Med.* (2021) 53:257–66. doi: 10.1080/07853890.2020.1868564
49. Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *npj Digital Med.* (2021) 4:3. doi: 10.1038/s41746-020-00372-6
50. Horton NJ, Lipsitz SR, Parzen M. A potential for bias when rounding in multiple imputation. *Am Statist.* (2003) 57:229–32. doi: 10.1198/0003130032314
51. Sun B, Feng J, Saenko K. Correlation alignment for unsupervised domain adaptation. In: *Domain Adaptation in Computer Vision Applications.* Springer (2017). p. 153–71. doi: 10.1007/978-3-319-58347-1_8
52. Daume H. III, Marcu D. Domain adaptation for statistical classifiers. *J Artif Intell Res.* (2006) 26:101–26. doi: 10.1613/jair.1872

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 17 January 2022
doi: 10.3389/fdgth.2021.662343

Explainable Machine Learning for COVID-19 Pneumonia Classification With Texture-Based Features Extraction in Chest Radiography

Luís Vinícius de Moura[1], *Christian Mattjie*[1,2], *Caroline Machado Dartora*[1,2], *Rodrigo C. Barros*[3] *and Ana Maria Marques da Silva*[1,2]*

[1] *Medical Image Computing Laboratory, School of Technology, Pontifical Catholic University of Rio Grande do Sul, PUCRS, Porto Alegre, Brazil,* [2] *Graduate Program in Biomedical Gerontology, School of Medicine, Pontifical Catholic University of Rio Grande do Sul, PUCRS, Porto Alegre, Brazil,* [3] *Machine Learning Theory and Applications Lab, School of Technology, Pontifical Catholic University of Rio Grande do Sul, PUCRS, Porto Alegre, Brazil*

Both reverse transcription-PCR (RT-PCR) and chest X-rays are used for the diagnosis of the coronavirus disease-2019 (COVID-19). However, COVID-19 pneumonia does not have a defined set of radiological findings. Our work aims to investigate radiomic features and classification models to differentiate chest X-ray images of COVID-19-based pneumonia and other types of lung patterns. The goal is to provide grounds for understanding the distinctive COVID-19 radiographic texture features using supervised ensemble machine learning methods based on trees through the interpretable Shapley Additive Explanations (SHAP) approach. We use 2,611 COVID-19 chest X-ray images and 2,611 non-COVID-19 chest X-rays. After segmenting the lung in three zones and laterally, a histogram normalization is applied, and radiomic features are extracted. SHAP recursive feature elimination with cross-validation is used to select features. Hyperparameter optimization of XGBoost and Random Forest ensemble tree models is applied using random search. The best classification model was XGBoost, with an accuracy of 0.82 and a sensitivity of 0.82. The explainable model showed the importance of the middle left and superior right lung zones in classifying COVID-19 pneumonia from other lung patterns.

Keywords: coronavirus, radiomics, radiological findings, X-rays, machine learning, explainable models, SHAP

OPEN ACCESS

Edited by:
Phuong N. Pham, Harvard Medical School, United States

Reviewed by:
Elena Casiraghi, Università degli Studi di Milano, Italy
Lal Hussain, University of Azad Jammu and Kashmir, Pakistan

***Correspondence:**
Ana Maria Marques da Silva
ana.marques@pucrs.br

Specialty section:
This article was submitted to Health Technology Innovation, a section of the journal Frontiers in Digital Health

Received: *01 February 2021*
Accepted: *29 November 2021*
Published: *17 January 2022*

Citation:
Moura LVd, Mattjie C, Dartora CM, Barros RC and Marques da Silva AM (2022) Explainable Machine Learning for COVID-19 Pneumonia Classification With Texture-Based Features Extraction in Chest Radiography. Front. Digit. Health 3:662343. doi: 10.3389/fdgth.2021.662343

INTRODUCTION

The coronavirus disease-2019 (COVID-19) is a viral respiratory disease with high rates of human-to-human contagious and transmission and was first reported in 27 patients with pneumonia of unknown etiology on December 31st, 2019, in Wuhan, China. The causative agent, a beta coronavirus 2b lineage (1) named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was identified on January 7, 2020, by throat swab samples (2, 3). In December 2020, 1 year after the outbreak, the WHO reported 66,243,918 confirmed cases and 1,528,984 deaths by COVID-19 (4). On January 30, 2021, 1 year after WHO declared COVID-19 as an international public health emergency, confirmed cases achieved 101,406,059 and 2,191,898 deaths (4). August 2021, the mark of 198,778,175 confirmed cases was achieved, with 4,235,559 deaths related to

COVID-19 worldwide (4). In August 2020, at least five SARS-CoV-2 virus clade mutations were reported, which have increased the infectivity and viral loads in the population (5). One year later (August 2021), with more than 3,886,112,928 COVID-19 vaccines applied, the third wave of COVID-19 is being noticed in Europe due to the quick virus adaptations increasing the transmissibility and viral load, with four variants of concern and more than ten others identified (6).

The clinical aspects of COVID-19 are highly variable between individuals, varying in different levels of involvement from asymptomatic to lethal conditions. The incubation period goes from 2 to 14 days. A mildly symptomatic condition usually presents fever, dry cough, fatigue, muscle pain, and taste and smell changes, with few patients showing neurological and digestive symptoms (7). Severe symptomatic patients can develop dyspnea, acute respiratory distress syndrome, septic shock, and metabolic acidosis (1). In the United States of America (USA), the younger population, between 18 and 29 years old, are part of the group age with more cases of COVID-19, followed by people between 50 and 64 years (8). However, the number of deaths is higher in the elderly population. Over 30% of deaths are concentrated in 85+ years population, ∼60% between 50 and 84 years, and only 0.5% in youngers (18–29 years) (8). COVID-19 incidence is also higher in women (52% of cases), while men show to be more at risk of death (∼54%) (8). A study published in August 2020 (1), with 121 Chinese patients with COVID-19, showed that the risk of adverse outcomes in individuals with more than 65 years is 2.28 times higher, and initial clinical manifestation does not differ between non-severe and severe cases, as in survivors and non-survivors. The risk factors for predicting COVID-19 severity were cardiovascular and cerebrovascular diseases (1). Higher lactate dehydrogenase (LDH) and coagulation dysfunction also contribute to the severity of the disease and progression to death (1).

The diagnosis of COVID-19 uses two approaches: reverse transcription-PCR (RT-PCR) (7) and chest X-rays (CXRs) (9). However, with the possibility of RT-PCR false-negative results, clinical and laboratory tests have usually been added to the patient diagnosis investigation and imaging findings. The imaging techniques are typically CXR or CT, and the findings have been compared with those of typical pneumonia. COVID-19 pneumonia does not have a defined set of imaging findings resulting in heterogeneous positivity definitions. The diagnosis sensitivity ranges for chest CT and CXR are from 57.4 to 100% and 59.9 to 89.0%, specificity from 0 to 96.0 and 11.1 to 88.9%, respectively (10). Automatic methods of prediction of COVID-19 in CT have been evaluated in several articles, as described by Chatzitofis et al. (11) and Ning et al. (12). Even CT provides higher image resolution, CXR images are less costly, available in clinics and hospitals, implies a lower radiation dose, and have a smaller risk of contamination of the imaging equipment and interruption of radiologic services to decontamination (13, 14). Also, abnormality findings on CT are mirrored in CXR images (14, 15).

The interpretation of radiological lung patterns can reveal differences in lung diseases. For COVID-19-related pneumonia, the characteristic pattern in CXR includes a pleuropulmonary abnormality with the presence of bilateral irregular, confluent, or bandlike ground-glass opacity or consolidation in a peripheral and mid-to-lower lung zones distribution with less likely pleural effusion (14, 16). For other lung diseases, like typical pneumonia, radiological patterns are related to the disease origin: bacterial, viral, or another etiology. In general, CXR findings show segmental or lobar consolidation and interstitial lung disease (17). Specifically, for viral pneumonia caused by adenovirus, the radiological pattern is characterized by multifocal consolidation or ground-glass opacity. In addition, there are bilateral reticulonodular areas of opacity, irregular or nodular regions of consolidation for pneumonia by influenza virus (18). For bacterial pneumonia, there are three main classifications due to the affected region: lobar pneumonia, with confluent areas of focal airspace disease usually in just one lobe, bronchopneumonia, with a multifocal distribution with nodules and consolidation in both lobes, and acute interstitial pneumonia that involves the bronchial and bronchiolar wall, and the pulmonary interstitium (19).

Apart from typical visual interpretation, the lung disease patterns can be studied through texture-feature analysis and radiomic techniques. However, due to radiologists' unfamiliarity with COVID-19 patterns, computer-aided diagnosis (CAD) systems help to differentiate COVID-19-related and other lung patterns.

Radiomics is a natural extension of CAD that converts the medical images into mineable high-dimensional data, allowing hypothesis generation, testing, or both (20). Computer-based texture analysis can be present in radiomics and reflects the tissue changes quantitatively from a healthy state to a pathological one. The extracted features can feed a classification model. The process has been widely explored to help radiologists achieve a better and faster diagnosis and is being applied to analyze and classify medical images to detect several diseases such as skin cancer (21), neurological disorders (22), and pulmonary diseases, like cancer (23) or pneumonia (24). For example, COVID-19-related pneumonia studies have been using various methods to differentiate the disease from typical pneumonia healthy individuals or even several lung diseases. Many approaches use deep learning (DL) framework, basing the feature extraction and classification in convolutional neural networks (CNN) models (25–28). However, DL models are not inherently interpretable and cannot explain their predictions intuitively and understandably. Hand-crafted feature extraction yet has mathematical definitions and can be associated with known radiological patterns.

The interpretability of results using medical images has been highly required. Therefore, explainable AI (Artificial Intelligence) predictions have been developed, such as the Shapley Additive

Abbreviations: 1st order, First-order features; AP, anteroposterior; AUC, area under the curve; COVID-19, Coronavirus SARS-CoV-2 diseases; CNN, Convolutional Neural Networks; CXR, Chest X-ray; GLCM, Gray-level co-occurrence matrix; GLRLM, Gray-level Run Length Matrix; GLSZM, Gray-level Size Zone Matrix; GLDM, Gray-level Dependence Matrix; PA, posteroanterior; PCR, Protein C-reactive; ROC, receiver operating characteristic; SHAP, Shapley Additive Explanations; SHAP-RFECV, Shapley Additive Explanations with Recursive Feature Elimination with Cross-validation.

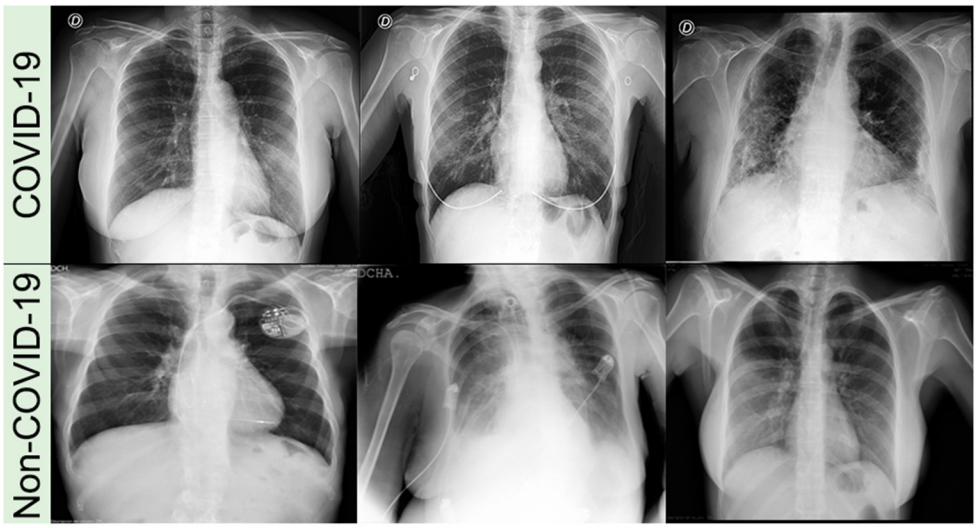


FIGURE 1 | Examples of CXR images from COVID-19 and non-COVID-19 datasets.

Explanations (SHAP) framework (29). Our goal is to investigate the radiomic features and classification models to differentiate chest X-ray images of COVID-19-based pneumonia and diverse types of lung pathologies. We aim to provide grounds for understanding the distinctive radiographic features of COVID-19 using supervised ensemble machine learning methods based on trees in an interpretable way using the SHAP approach to explain the meaning of the most important features in prediction. Our study analyzes the lung in three zones in both lobes, showing the middle left and superior right zones' importance in identifying COVID-19.

MATERIALS AND METHODS

We retrospectively used a public dataset of CXR images of COVID-19-related pneumonia and lung images of patients with no COVID-19 to investigate the radiomic features that can discriminate COVID-19 from other lung radiographic findings. We extracted first- and second-order radiomic features and divided our analysis into two main steps. First, we performed k-folds cross-validation to select the algorithm with the best performance classifying COVID-19 pneumonia and non-COVID-19. Second, we performed an explanatory approach to select the best set of radiomic features that characterize COVID-19 pneumonia.

Image Dataset

We used two public multi-institutional databases related to COVID-19 and non-COVID-19 to train and evaluate our model. BIMCV (Valencian Region Medical ImageBank) is a large dataset, with annotated anonymized X-ray and CT images along with their radiographic findings, PCR, immunoglobulin G (IgG), and immunoglobulin M (IgM), with radiographic reports from Medical Imaging Databank in Valencian Region Medical Image Bank. BIMCV-COVID+ (30) comprises 7,377 computed radiography (CR), 9,463 digital radiography (DX), and 6,687 CT studies, acquired from consecutive studies with at least one positive PCR or positive immunological test for SARS-Cov-2. BIMCV-COVID- (31) has 2,947 CR, 2,880 DX, and 3,769 CT studies of patients with negative PCR and negative immunological tests for SARS-Cov-2. Both databases have all X-ray images stored in 16-bit PNG images in their original high-resolution scale, with sizes varying between 1,745 × 1,465 pixels and 4,248 × 3,480 pixels for patients with COVID-19 and between 1,387 × 1,140 pixels and 4,891 × 4,020 pixels for non-COVID-19. **Figure 1** shows an example of CXR images of both datasets (BIMCV-COVID+ and BIMCV-COVID–).

We used only CXR images of anteroposterior (AP) and posteroanterior (PA) projections of adult patients (≥18 years old). Some images from the dataset do not have information regarding projection (AP, PA, or lateral) in the DICOM tag, and some projection tags are mislabeled. All images with missing the projection DICOM tag were discarded and, from the selected AP or PA projections, they were visually inspected and discarded if mislabeled.

For COVID-19 positive patients, we selected only the first two images between the first and last positive PCRs. For non-COVID-19 participants with more than two X-ray acquisitions, only the first two images were selected. Therefore, we have 2,611 images from patients with COVID-19, and we randomly chose 2,611 images from non-COVID-19 to ensure a balanced dataset. Demographic information regarding selected patients is shown in **Table 1**. In addition, since some CXR images were stored with an inverted lookup table, images with photometric interpretation equal to "monochrome1" were multiplied by minus one and summed with their maximum value to harmonize the dataset.

Previous studies already used the BIMCV-COVID dataset to evaluate lung segmentation (32), data imbalance corrections (33),

TABLE 1 | Demographic data of our study.

	COVID-19	Non-COVID-19
Age	62 ± 16	64 ± 19
Sex (male)	1,358	1,268
Sex (female)	1,253	1,343

DL classification models (34–36), and other imaging challenges (37, 38).

Lung Segmentation

We applied a histogram equalization in each CXR image (39, 40) to normalize the intensity values and reduce the dataset's features variability. Pixel values were normalized to 8-bits per pixel and resampled to 256 × 256 pixels. We segment lungs using an open-source pretrained U-Net-inspired architecture segmentation model to generate lung masks[1]. The model was trained in two different open CXR databases: JSRT (Japanese Society of Radiological Technology) (41) and Montgomery County (42). The databases used for training came from patients with tuberculosis, so they are not specific for COVID-19-based pneumonia lung segmentation.

We applied the opening morphological operator in each mask to remove background clusters and fill holes of the resulting lung mask, using a square structuring element, 8-connected neighborhood. The opening morphological operation smooths an object's contour, breaks narrow isthmuses, and eliminates thin protrusions. The mathematical details of opening morphological operation can be found in Gonzalez and Woods (43).

We removed all clusters with <5 pixels and all connected regions with <75 pixels. The lung mask is stretched back to the original image size and applied to the original image before processing. We normalized all segmented lung images considering only pixels inside the lung mask, between 0 and 255.

We split the image between the left and right side using the centroid of two areas; if the centroid is located within the first half of the matrix size (from left to right), it is considered as part of the right lung (the radiological image in CXR is mirrored). Next, each lung's height is divided into upper, middle, and bottom zones, determined by the extremities' distance, divided into thirds. The segmentation workflow is shown in **Figure 2**.

Radiomic Features

We used the PyRadiomics library to extract the first and second-order statistical texture-based features for each lung mask. The mathematical formulation of the features can be found in Zwanenburg, Leger, and Vallières (44). The radiomic features are divided into five classes (45):

First-order features (18 features): These are based on the first-order histogram and related to the pixel intensity distribution.

Gray-level co-occurrence matrix or GLCM (24 features): This gives information about the gray-level spatial distribution, considering the relationship between pixel pairs and the frequency of each intensity within an 8-connected neighborhood.

Gray-level run length matrix or GLRLM (16 features): This is like GLCM; it is defined as the number of contiguous pixels with the same gray level considering a 4-connected neighborhood, indicating the pixel value homogeneity.

Gray-level size zone matrix or GLSZM (16 features): This is used for texture characterization; it provides statistical representation by estimating a bivariate conditional probability density function of the image distribution values and is rotation-invariant.

Gray-level dependence matrix or GLDM (14 features): This quantifies the dependence of gray image level by calculating the connectivity at a certain distance when its difference in pixel intensity is <1.

Model Selection

We performed 10-folds cross-validation after radiomic feature extraction to guarantee unbiased metrics results and error generalization. We realized that a normalization between 0 and 1 and used SHAP- Recursive Feature Elimination with Cross-Validation (RFECV) feature selection in 9-folds of the dataset. Each machine learning model was trained using the selected features with the hyperparameter optimization method, and randomized search with cross-validation from the sci-kit learn library (46, 47). The method uses a range of values for each parameter in the model. It tests a given number of times with different combinations and splits of training data, measuring the model performance in the validation set. We chose to run 1,000 iterations for each model, using an intern 5-folds stratified cross-validation. The parameter values explored in each model are shown in **Table 2**. We chose the best parameters based on the best performance of recall in cross-validation. After hyperparameter optimization, model evaluation is performed in the last fold.

Feature Selection

Feature selection is the process of selecting the most relevant features for a given task. The process reduces the computational cost regarding the training and evaluation of the machine learning model and improves the generalization (48). Moreover, some features may be irrelevant or redundant, negatively impacting the modeling, adding biases (49). To avoid these issues, we decided to make a feature selection before our modeling.

We used the SHAP-RFECV from the Probatus python library to perform the feature selection. Further information about the SHAP-RFECV algorithm and its applications can be found on the Probatus webpage[2]. The method uses a backward feature elimination based on the SHAP value of feature importance. The designed model is trained with all features initially and uses cross-validation (10-fold) to estimate each feature's SHAP importance value. At the end of each round, the features with the lowest importance are excluded. Then, the training is done again until the number of features chosen by the user is reached. We decided to remove 20% of the lowest importance features in each iteration for faster reduction of features in early iterations

[1]https://github.com/imlab-uiip/lung-segmentation-2D.

[2]https://ing-bank.github.io/probatus/.

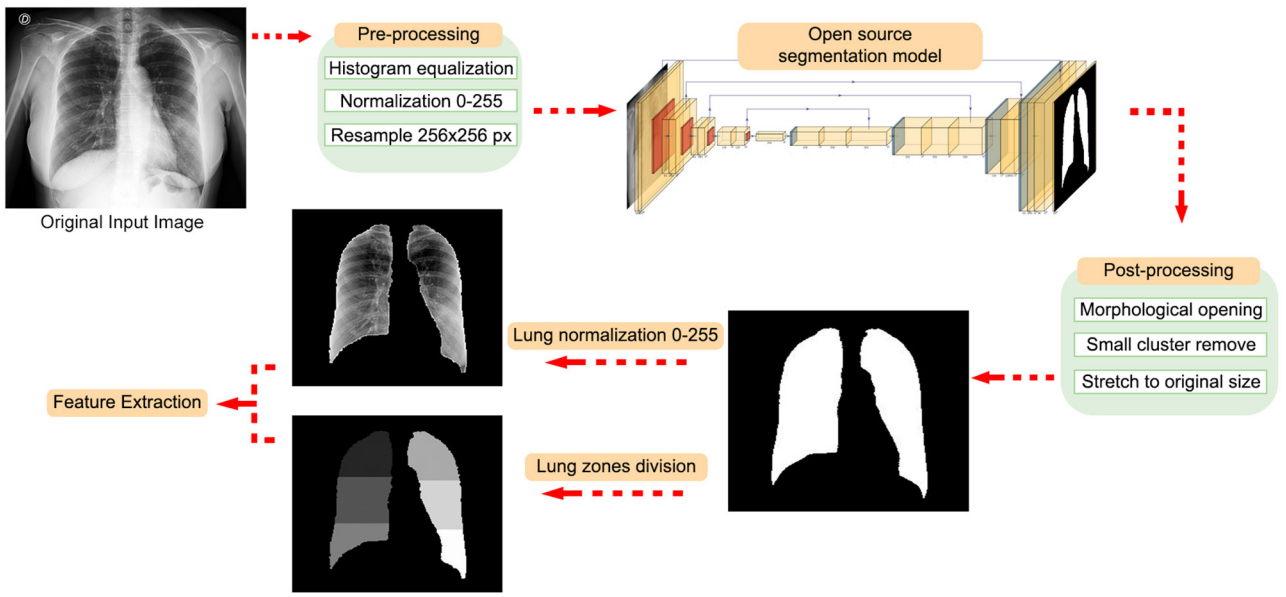


FIGURE 2 | Workflow of the segmentation process.

TABLE 2 | Parameter values explored by random search in XGBoost and Random Forest Classifier models.

XGBoost		Random Forest Classifier	
min_child_weight	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	n_estimators	10, 50, 100, 200
max_depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	max_depth	None, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
lambda*	0–1	criterion	gini, entropy
gamma*	0–1	min_sample_split	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
eta*	0–1	min_sample_leaf	1, 2, 3, 4, 5
objective	binary:logistic		
tree_method	gpu_hist		

**10 random values in the range.*

and more precise results in the later ones until it reaches the minimum value of 20 features.

Machine Learning Models

We trained two ensemble classification models based on tree-based models using the scikit-learn (47) library and XGBoost (XGB) (50) on Python version 3.6.5. The classification methods used in our article are the XGB and the Random Forest (RF).

XGBoost is a scalable ensemble model based on an extreme gradient for tree boosting. It is based on regression trees, which in contrast to decision trees, contains a continuous score on each of the leaves. The input data is sorted into blocks of columns that are categorized by the corresponding feature value. The split search algorithm runs in the block seeking the split candidates' statistics in all leaf branches. It uses decision rules into trees to determine for each leaf the example will be placed. The final prediction is calculated by summing up the score in the corresponding leaves (51). The proper algorithm and mathematical formulation are addressed in Chen (50).

Random Forest is a tree-based ensemble learning algorithm that induces a pre-specified number of decision trees to solve a classification problem. Each tree is built using a subsample of the training data, and each node searches for the best feature in a subset of the original features. The assumption is that by combining the results of several weak classifiers (each tree) *via* majority voting, one can achieve a robust classifier with enhanced generalization ability. The mathematical formulation of RF is described in Breiman (52).

Performance Evaluation

Accuracy, sensitivity, precision, F1-Score, and the area under the curve (AUC) of the receiver operating characteristic (ROC) were used for model evaluation. The final model was selected based on the best sensitivity achieved in the cross-validation. Each metric

is calculated as follows (53):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$Sensitivity = \frac{TP}{TP + FN}$$
$$Precision = \frac{TP}{TP + FP}$$
$$F1\ Score = 2 \times \frac{precision \times sensitivity}{precision + sensitivity}$$

where: TP = true positive, TN = true negative, FP = false positive, and FN = false negative.

Explanatory Approach

Despite their complexity, approaches to making AI models "interpretable" have gained attention to enhance the understanding of machine learning algorithms. Explaining tree models is particularly significant because the pattern that the model uncovers can be more important than the model's prediction itself. The SHAP approach is an additive feature attribution method that assigns an "importance value" to each feature for a particular prediction (29, 54). The SHAP approach satisfies three important properties for model explanation: (i) local accuracy because the explanation model and the original one has to match, at least, the output for a specific input; (ii) consistency, because if the model changes, it is because a feature's contribution increases or stay the same regardless of other inputs, so the input's attribution should not decrease; and (iii) missingness, meaning the missing values of features in the original input have no impact.

In our work, we chose the SHAP approach with tree models because it is calculated in each tree leaf and gives interpretability for local explanations, which reveals the most informative features for each subset of samples. Local explanations allow the identification of global patterns on data and verify how the model depends on its input features. It also increases the signal-to-noise ratio to detect problematic data distribution shifts, making it possible to analyze the behavior of the entire dataset, which is composed of medical images from different databases (54).

The SHAP approach uses an extension of Shapley values from the game theory to calculate the feature importance. The SHAP values are calculated based on the prediction difference when using all features and when using just a few ones. It addresses how the addition of one feature improves or not the prediction (55). In tree-based models, the SHAP values are also weighted by the node sizes, meaning the number of training samples in the node. Finally, the feature importance assumes that the features with large absolute SHAP values have more importance than others with smaller absolute values. To access the global importance, we average the feature importance across all data (56).

We use the SHAP-RFECV approach on the final model to evaluate the most important features and how they affect the model's prediction.

RESULTS

We extracted 88 features for each lung zone. Next, we applied RF and XGB models to analyze the performance with 10-fold cross-validation. **Table 2** shows the parameters used for both models, and **Table 3** has the performance metrics.

The XGBoost model was selected for hyperparameters optimization and feature selection using SHAP due to its higher classification performance. **Table 4** shows the hyperparameters, and **Table 5** shows the selected features.

The SHAP feature importance values are shown in **Figure 3** for the 20 selected features. **Figure 4** shows the effect of each feature in the model prediction.

Figures 5, **6** show the decision's plot of one COVID-19 pneumonia case and one non-COVID-19 pneumonia case, respectively. The plot shows the SHAP values related to each feature's importance and how they predict the classification for two random individuals. For example, the positive SHAP value in **Figure 5**, $f_{(x)} = 2.207$, is related to the model prediction identifying the CXR as a COVID-19. Similarly, the negative SHAP value in **Figure 6**, $f_{(x)} = -1.052$, is related to the model prediction identifying the CRX as a patient with non-COVID-19.

DISCUSSIONS

In the latest year, numerous studies have been developed applying different computer-aided methods to aid in diagnosing and in the prognosis COVID-19 (11, 12, 14, 15, 25–28, 57–75). However, most studies do not use explainable methods (57). Our approach uses the hand-crafted radiomics features approach and ensemble tree-based machine learning classification models to differentiate COVID-19-induced pneumonia from other lung pathologies and healthy lungs in CXR images. We use ensemble tree-based models since they are more accurate than artificial neural networks in many applications (54). Looking for explainability, we use the SHAP approach to unveil why specific features with low or high SHAP values are associated with the disease.

TABLE 3 | Mean cross-validation metrics of XGBoost and Random Forest models.

Model	Accuracy	F1-Score	Sensitivity	Precision
XGB	0.82	0.82	0.82	0.82
RF	0.77	0.78	0.81	0.75

TABLE 4 | Hyperparameters for XGBoost.

XGBoost	
min_child_weight	6
max_depth	3
lambda	1
gamma	0.50230907476997
eta	0.50515969241175
objective	binary:logistic
tree_method	hist

TABLE 5 | Chosen features by XGBoost model.

XGBoost	Abbreviation
Bottom Left - First Order - Maximum	BL-1st-M
Bottom Right - First Order - Energy	BR-1st-E
Bottom Right - First Order - Kurtosis	BR-1st-K
Bottom Right - GLCM - Cluster Prominence	BR-GLCM-CP
Bottom Right - GLCM - Difference Variance	BR-GLCM-DV
Middle Left - First Order - Kurtosis	ML-1st-K
Upper Left - First Order - Range	UL-1st-R
Upper Left - GLCM - Idmn	UL-GLCM-LDMN
Upper Left - GLRLM - Run Entropy	UL-GLRLM-RE
Upper Right - First Order - Robust Mean Absolute Deviation	UR-1st-RMAD
Upper Right - GLCM - Cluster Prominence	UR-GLCM-CP
Upper Right - GLCM - Cluster Shade	UR-GLCM-CS
Upper Right - GLCM - MCC	UR-GLCM-MCC
Upper Right - GLRLM - Gray Level Non-Uniformity	UR-GLRLM-GLNU
Upper Right - GLRLM - High Gray Level Run Emphasis	UR-GLRLM-HGLRE
Upper Right - GLSZM - Gray Level Non-Uniformity Normalized	UR-GLSZM-GLNUN
Upper Right - GLSZM - Gray Level Variance	UR-GLSZM-GLV
Upper Right - GLSZM - Large Area High Gray Level Emphasis	UR-GLSZM-LAHGLE
Upper Right - GLSZM - Size Zone Non-Uniformity	UR-GLSZM-SZNU
Upper Right - GLSZM - Small Area High Gray Level Emphasis	UR-GLSZM-SAHGLE

Moreover, we choose to analyze different lung zones, previously segmented, looking for regions of interest in the disease.

The computer-aided methods using non-segmented images may lead to biases (58, 59) since the models may associate elements from outside the lungs, such as bones and muscles, not related to the disease, with the presence of COVID-19. Restricting the region of interest ensures that the features extracted are associated with the radiological information present in the lung zone. Our study uses automatic lung segmentation and divides the lung into six zones, which are independently analyzed. The approach allows small structures in the analysis, which could be suppressed by analyzing the entire lung at once.

Nowadays, COVID-l9 radiological studies are focused on CT findings, which have better sensitivity than CXR. However, CT is more expensive and scarcer than conventional X-rays, requiring complicated decontamination after scanning patients with COVID-19. Therefore, the American College of Radiology (60) recommends CT to be used sparingly and reserved for hospitalized patients with COVID-l9 symptomatic with specific clinical indications. A portable chest X-ray equipment is suggested as a viable option to minimize the risk of cross-infection and avoid overload and disruption of radiological departments. Moreover, studies have shown that CXR COVID-19 findings mirror the CT findings (14, 15), with less radiation dose and higher availability in clinics and hospitals.

The use of feature importance techniques can improve clinical practice in different medical fields. Hussain et al. (76) showed the benefits of multimodal features extracted from congestive heart failure and normal sinus rhythm signals. The application of feature importance ranking techniques was beneficial to distinguish healthy subjects from those with heart failure. The same group also found that the use of the synthetic minority oversampling technique can improve the model performance when dealing with imbalanced datasets (77).

Meaningful Texture-Based Features in COVID-19 Pneumonia

The main contribution of this study is the findings of a group of meaningful radiomic features in differentiating COVID-19 from other lung diseases using CXR using an explainable machine learning approach. The most relevant features are presented in **Figure 3**. In addition, SHAP values summary plots can be used to try and explain how each feature is increasing or decreasing the model output, meaning the probability of classifying a CXR image as COVID-19 pneumonia or not.

As seen in **Figure 3**, the first-order kurtosis on the middle-left lung was, by far, the feature with the highest importance for classification. Kurtosis is, in general, a measure of the "peakedness" of the distribution of the values (78). Since it is a first-order feature, it is directly related to the pixel values on the CXR image. COVID-19 induces consolidation and ground-glass opacification, which increases these pixel values in the lung region and may induce a distribution with lighter tails and a flatter peak, resulting in lower kurtosis values (78). These lower values were associated with the disease, as can be seen in **Figure 4**.

In CXR images, the heart partially overlaps with the lungs in the left middle region, influencing feature importance. Despite being the most important feature, kurtosis was the only feature selected from this region. The second and third most important features are both related to gray-level homogeneity. Higher values of Run Entropy in GLRLM, which were associated with COVID-19 in the upper left region, indicate more heterogeneity in the texture patterns. Higher values of gray-level non-uniformity in GLRLM in the upper right region, also associated with the disease, indicate lesser similarity between intensity values, corroborating the previous feature findings. COVID-19-induced consolidations tend to be diffuse or patchy, which may explain these features associations (70).

We won't go over other features individually due to their decreasing and similar importance. However, it is important to note that from the 20 selected features, half are extracted from the upper right zone. Our previous study (61) showed similar results, where the two most important features were also from the upper right lung region. Moreover, no middle right zone features were selected.

Comparison With Related Work

Saha et al. (62) created EMCNet, an automated method to diagnose COVID-19 and healthy cases from CXR images. Their

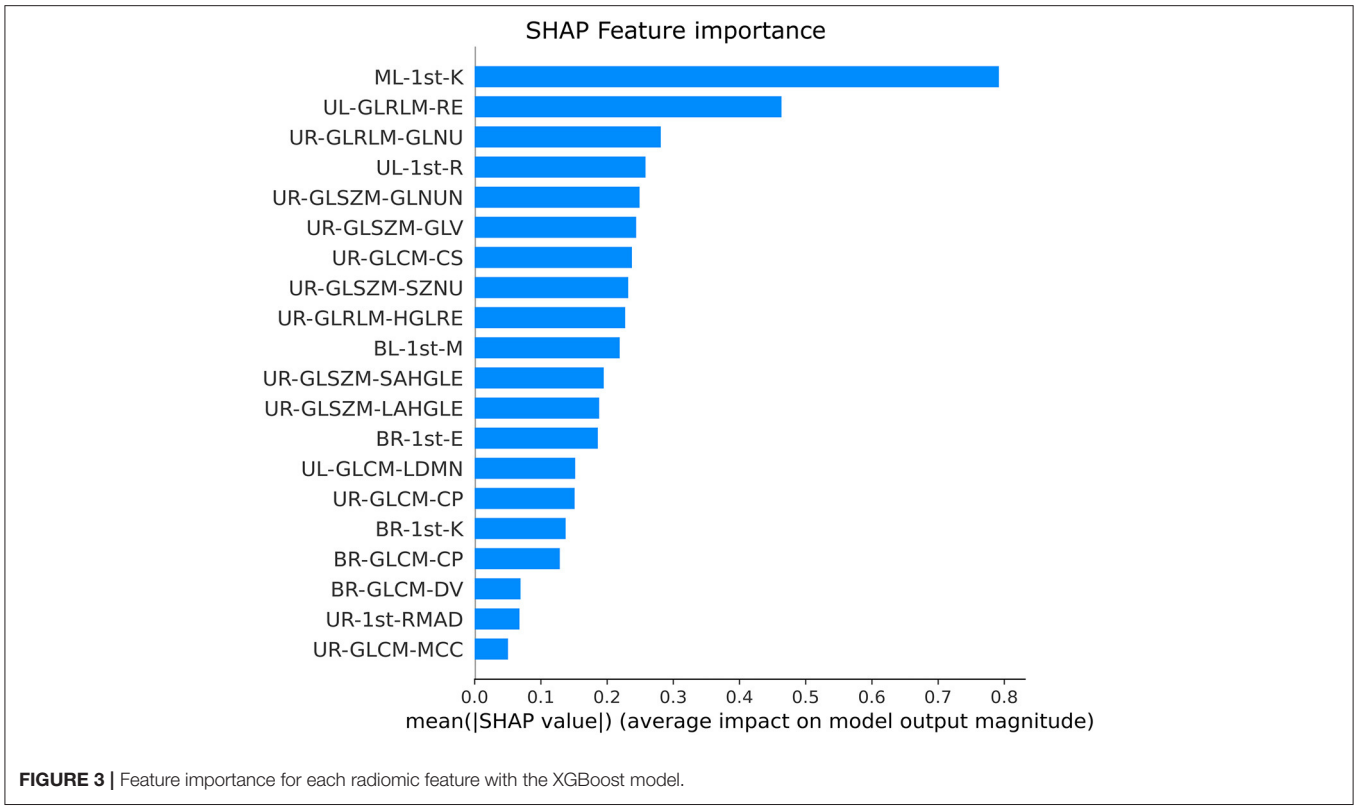


FIGURE 3 | Feature importance for each radiomic feature with the XGBoost model.

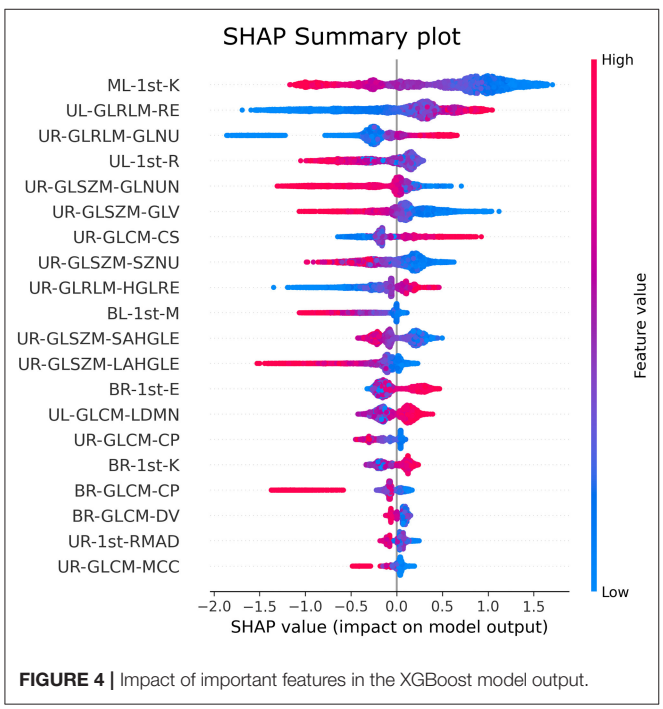


FIGURE 4 | Impact of important features in the XGBoost model output.

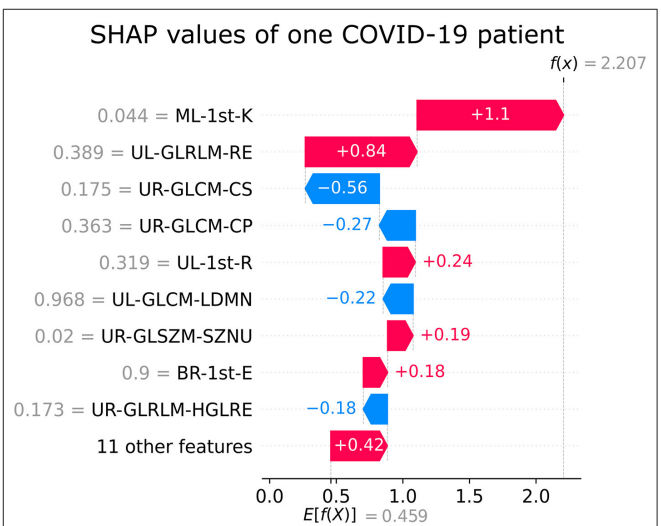


FIGURE 5 | Example of SHAP values affecting XGBoost model output for a single COVID-19 CXR image.

method uses a simple CNN to extract 64 features from each image and then classify binary with an ensemble of classifiers composed of Decision Tree, Random Forest, Support Vector Machines, and AdaBoost models. They used 4,600 images (2,300 COVID-19 and 2,300 healthy) from different public datasets (38, 79–81) and applied resizing and data normalization. As a result, they achieved 0.989 for accuracy, 1.00 for precision, 0.9782 for sensibility, and 0.989 in F1-score. In comparison with our study, we both use almost the same number of CXR

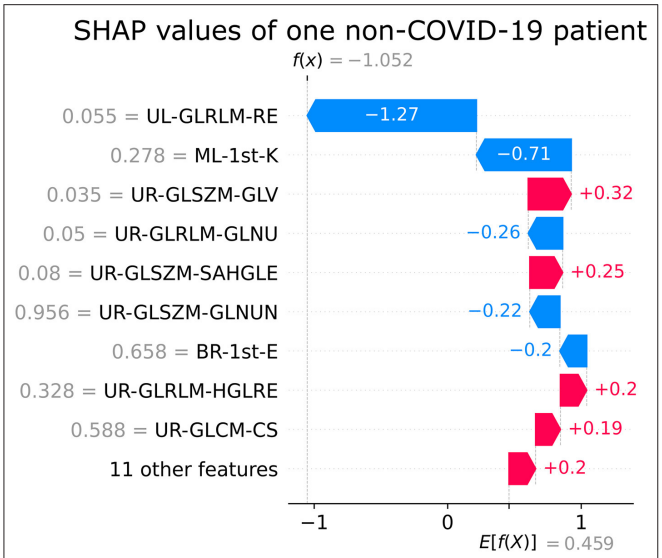


FIGURE 6 | Example of SHAP values affecting XGBoost model output for a single non-COVID-19 CXR image.

images, but we only use two databases, mostly from the same facilities. However, the authors' data included pediatrics patients, leading to biases due to age-related characteristics. Duran-Lopez et al. (34) proposed COVID-XNet, a DL-based system that uses pre-processing algorithms to feed a custom CNN to extract relevant features and classify between COVID-19 and normal cases. Its system achieved 92.53% sensitivity, 96.33% specificity, 93.76% precision, 93.14% F1-score, 94.43% balanced accuracy, and an AUC value of 0.988. Class Activation Maps were used to highlight the main findings in COVID-19 X-ray images and were compared and verified with their corresponding ground truth by radiologists.

Cavallo et al. (63) made a texture analysis to evaluate COVID-19 in CXR images. They used a public database selecting 110 COVID-19-related and 110 non-COVID-19-related interstitial pneumonia, avoiding the presence of wires, electrodes, catheters, and other devices. Two radiologists manually segmented all the images. After normalization, 308 textures were extracted. An ensemble made by Partial Least Square Discriminant Analysis, Naive Bayes, Generalized Linear Model, DL, Gradient Boosted Trees, and Artificial Neural Networks models achieved the best results. The ensemble model performance was 0.93 of sensitivity, 0.90 of specificity, and 0.92 of accuracy. We have not attained these high metrics. However, we used much more images (5,222 vs. 220), and like in the previous study, they used mixed-age data, where COVID-19 images were retrieved from adult patients and non-COVID-19 images from pediatric ones.

Rasheed et al. (64) proposed a machine learning-based framework to diagnose COVID-19 using CXR. They used two publicly available databases with a total of 198 COVID-19 and 210 healthy individuals, using Generative Adversarial Network data augmentation to get 250 samples of each group. Features were extracted from 2D CXR with principal component analysis (PCA). Training and optimization were done with CNN and logistic regression. The PCA with CNN gave an overall accuracy of 1.0. Using just 500 CXR images from COVID-19 and non-COVID-19 individuals from two different datasets and training the model with data augmentation techniques may suggest the possibility of overfitting or that classification can be differentiating the two datasets and not the disease itself.

Brunese et al. (65) developed a three-phase DL approach to aid in COVID-19 detection in CXR: detect the presence of pneumonia, discern between COVID-19 induced and typical pneumonia, and localize CXR areas related to COVID-19 presence. Different datasets for different pathologies were used, two datasets of patients with COVID-19 and one of the other pathologies. In total, 6253 CXR images were used, but only 250 were from patients with COVID-19. Accuracy for pneumonia detection and COVID-19 discrimination was 0.96 and 0.98, respectively. Activation maps were used to verify which parts of the image were used by the model for classification. They showed a high probability of prediction in the middle left and upper right lungs, agreeing with our findings.

Kikkisetti et al. (66) used portable CXR from public databases with CNN and transfer learning to classify the images between healthy, COVID-19, non-COVID-19 viral pneumonia, and bacterial pneumonia. They used two approaches, using all CXR and only segmented lungs. CNN heatmaps showed that with the whole CXR, the model used outside the lungs information to classify. It is a tangible example of the importance of segmenting the CXR images, especially when using data from different locations, to avoid biases created by the annotations in X-rays. They achieved an overall sensitivity, specificity, accuracy, and AUC of 0.91, 0.93, 0.88, and 0.89, respectively, with segmented lungs. However, they used pediatric data mixed with adults. It is interesting to note that in their CNN heatmaps, the lower and middle portion of the left lung showed a high importance in their classification, in agreement with our results. We have three features from these locations, which are essential in the COVID-19 classification, and two were selected from these regions. However, one feature is, by far, the most important for classification. Moreover, our database is almost five times larger.

Yousefi et al. (67) proposed a computer-aided detection of COVID-19 with CXR imaging using deep and conventional radiomic features. A 2D U-net model was used to segment the lung lobes. They evaluated three different unsupervised feature selection approaches. The models were trained using 704 CXR images and independently validated using a study cohort of 1,597 cases. The resulting accuracy was 72.6% for multiclass and 89.6% for binary-class classification. Since unsupervised models were used, it is impossible to check if the most important features are like ours. Unfortunately, the lobes were not investigated separately for further comparisons.

Casiraghi et al. (82) developed an explainable prediction model to process the data of 300 patients with COVID-19 to predict their risk of severe outcomes. They collected clinical data and laboratory values. The radiological scores were retrospectively evaluated from CXR by either pooling radiologists' scores or applying a deep neural network. Boruta

and RF were combined in a 10-fold cross-validation scheme to produce a variable importance estimate. The most important variables were selected to train an associative tree classifier, with AUC 0.81–0.76, sensitivity 0.72–0.66, F1 score 0.62–0.55, and accuracy from 0.74 to 0.68. The PCR achieved the highest relative relevance, together with the patient's age and laboratory variables. They noted that the radiological features extracted from radiologists' scores and deep network were positively correlated, as expected. Moreover, their radiological features were also correlated with PCR and have an inverse correlation with the saturation values. However, they did not evaluate radiomic features extracted from CXR for comparison with our work.

Even though the image characteristics are different in CT and CXR, they share the same physical interaction with tissues using X-rays. Caruso et al. study (68) used CT texture analysis to differentiate patients with positive COVID-19 and negative ones. Sensitivity and specificity were 0.6 and 0.8, respectively. The feature with the highest correlation with patients with positive COVID-19 compared with negative ones was kurtosis, with lower values associated with the disease. In our work, the most important classification feature was first-order kurtosis with the same lower values behavior associated with the disease.

Similarly, Lin et al. (69) developed a CT-based radiomic score to diagnose COVID-19 and achieve a sensibility of 0.89. They also found that the GLCM MCC feature was important during the classification. In our study, we also found the same feature in the upper right lung. Finally, Liu et al. (70) analyzed the classification performance in CT images using two approaches (only clinical features and clinical with texture features), increasing their sensitivity to 0.93. Their results show the importance of clinical information, if available. They also found cluster prominence features as important, but their analysis was made using a wavelet filter.

Shiri et al. (71) made an analysis using CT images and clinical data to develop a prediction model of patients with COVID-19. The model with the highest performance achieves a sensibility and accuracy of 0.88. Interestingly, one of the radiomic features used in their work, GLSZM – SAHGLE, was also selected in our model in the upper right lung zone.

Limitations

One of the main limitations of AI studies is the currently available COVID-19 and non-COVID-19 CXR databases (72). Even though databases have many data, most have several missing and mislabeled data. Moreover, most COVID-19 public databases do not include non-COVID images from the same medical center, requiring other databases from different facilities. Medical centers use various scanners and protocols, leading to different image patterns if no previous harmonization is executed. In studies using multiple databases with other pathologies, computer-aided methods may learn to differentiate the database pattern rather than the lung pathologies (73). Finally, a limitation of the databases used in this work is they do not include clinical data from patients.

The most important concern about some studies is that some databases of typical pneumonia CXR have images from pediatric and adult patients. They are primarily used due to the differentiation between viral and bacterial pneumonia. However, this may increase biases due to age-related characteristics (71). Usually, imaging acquisition protocols for pediatric patients are made with less radiation due to radiological protection.

Our model reached an accuracy of 0.82, a sensitivity of 0.82, a precision of 0.82, and F1-score metrics of 0.82 using CXR images and SHAP RFECV. Other COVID-19 studies using CXR images and machine learning models reached accuracy between 0.92 and 0.99, and sensibility between 0.91 and 0.99 (62, 63). However, the limitations of the studies with higher scores discussed in the previous section should be considered; some studies used non-segmented images (62–64, 66), and other studies used data from pediatric patients (62, 64, 65) (1–5 years old) mixed with adult data.

The main limitations of our study are the absence of clinical data to improve our models and the lack of statistical analysis to have more confidence about the importance of the features. Moreover, we limited the features extraction and other features that could be included, like the neighboring gray-tone difference matrix. Finally, we did not evaluate the effect of features' extraction applying different pre-processing filters.

Future Directions

The sudden onset of COVID-19 generated a global task force to differentiate it from other lung diseases. With the main symptoms related to atypical pneumonia, the number of CXR and chest CT datasets has rapidly increased. More than 1 year and a half from the COVID-19 onset, the available datasets of chest images are more extensive, so it is possible to have confidence in classifying the disease from other pulmonary findings. However, before the 2020 pandemic, most lung radiomic signature analysis studies focused on identifying and classifying nodules and adenocarcinomas. Therefore, pneumonia radiomic signature is not well-established, even for typical pneumonia.

We still do not know how COVID-19 affects the immune system and the lungs, but some cases do not have any CXR alterations, they have only parenchymal abnormalities (75). For further studies, PCR positive COVID-19 individuals without visible CXR modifications should be analyzed using radiomic features to evaluate small regions looking for pulmonary tissue texture variations. In addition, when available, further studies should include clinical information to allow the evaluation of the benefits in diagnosis when using both radiomics and clinical data.

A big challenge in using large CXR image databases is maintaining label information such as projection (i.e., lateral, AP, PA). The further effort in global data curation could confirm projection without the need for visual confirmation.

Conclusions

This article presents the SHAP approach to explain machine learning classification models based on hand-crafted radiomics texture features to provide grounds for understanding the characteristic radiographic findings on CXR images of patients with COVID-19. The XGB ML model is the best discriminant method between COVID-19 pneumonia and healthy and

other lung pathologies using radiomic features extracted from lung CXR images divided into six zones. The explainable model shows the importance of the middle left and superior right lung zone in classifying COVID-19 pneumonia from other lung patterns. The method can potentially be clinically applied as a first-line triage tool for suspected individuals with COVID-19.

The rapid increase of COVID-19 pneumonia cases shows the necessity of urgent solutions to differentiate individuals with and without COVID-19 due to its high spreadability and necessity of prompt management of ill individuals. Furthermore, the lack of knowledge about the disease made it necessary to find explainable radiological features to correlate with the biological mechanisms of COVID-19.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://bimcv.cipf.es/bimcv-projects/bimcv-covid19/.

AUTHOR CONTRIBUTIONS

LM: study design, data mining and pre-processing, model development, writing, and analysis. CM and CD: study design, literature review, writing, and analysis. RB and AM: study design. All authors contributed to the article and approved the submitted version.

FUNDING

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brasil (CAPES) – Finance Code 001 and PUCRS by financial support with research scholarship.

ACKNOWLEDGMENTS

We thank the professionals working on the first line in the fight against the COVID-19 and all researchers involved in understanding the disease and creating vaccines. We thank Google Collaboratory for the free availability of computational resources to develop this study.

REFERENCES

1. Zou L, Dai L, Zhang Y, Fu W, Gao Y, Zhang Z, et al. Clinical characteristics and risk factors for disease severity and death in patients with coronavirus disease 2019 in Wuhan, China. *Front Med*. (2020) 7:532. doi: 10.3389/fmed.2020.00532
2. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, et al. World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). *Int J Surgery*. (2020) 76:71–6. doi: 10.1016/j.ijsu.2020.02.034
3. Aljondi R, Alghamdi S. Diagnostic value of imaging modalities for COVID-19: scoping review. *J Med Internet Res*. (2020) 22:e19673. doi: 10.2196/19673
4. WHO. *Coronavirus Disease (COVID-19) Dashboard*. (2020). Available online at: https://covid19.who.int/ (accessed January 31, 2021).
5. Zuckerman N, Pando R, Bucris E, Drori Y, Lustig Y, Erster O, et al. Comprehensive analyses of SARS-CoV-2 transmission in a public health virology laboratory. *Viruses*. (2020) 12:854. doi: 10.3390/v12080854
6. WHO. *Tracking SARS-CoV-2 Variants*. (2021). Available online at: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/ (accessed August 2021).
7. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19). *JAMA*. (2020) 324:782. doi: 10.1001/jama.2020.12839
8. CDC. *COVID Data Tracker*. (2020). Available online at: https://covid.cdc.gov/covid-data-tracker/#demographics (accessed December 7, 2020).
9. Gupta D, Agarwal R, Aggarwal A, Singh N, Mishra N, Khilnani G, et al. Guidelines for diagnosis and management of community- and hospital-acquired pneumonia in adults: Joint ICS/NCCP(I) recommendations. *Lung India*. (2012) 29:27. doi: 10.4103/0970-2113.99248
10. Salameh J.-P., Leeflang MM, Hooft L, Islam N, McGrath TA, van der Pol CB, et al. Thoracic imaging tests for the diagnosis of COVID-19. *Cochrane Database Syst Rev*. (2020) 9:CD013639. doi: 10.1002/14651858.CD013639.pub3
11. Chatzitofis A, Cancian P, Gkitsas V, Carlucci A, Stalidis P, Albanis G, et al. Volume-of-interest aware deep neural networks for rapid chest CT-based COVID-19 patient risk assessment. *Int J Environ Res Public Health*. (2021) 18:2842. doi: 10.3390/ijerph18062842
12. Ning W, Lei S, Yang J, Cao Y, Jiang P, Yang Q, et al. Open resource of clinical data from patients with pneumonia for the prediction of COVID-19 outcomes via deep learning. *Nat Biomed Eng*. (2020) 4:1197–207. doi: 10.1038/s41551-020-00633-5
13. American College of Radiology. *ACR Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection*. (2020). Available online at: https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection (accessed January 31, 2021).
14. Wong HYF, Lam HYS, Fong AH.-T., Leung ST, Chin TW-Y, Lo CSY, et al. Frequency and distribution of chest radiographic findings in patients positive for COVID-19. *Radiology*. (2020) 296:E72–8. doi: 10.1148/radiol.2020201160
15. Shamout FE, Shen Y, Wu N, Kaku A, Park J, Makino T, et al. An artificial intelligence system for predicting the deterioration of COVID-19 patients in the emergency department. *NPJ Dig Med*. (2021) 4:80. doi: 10.1038/s41746-021-00453-0
16. Smith DL, Grenier J.-P., Batte C, Spieler B. A characteristic chest radiographic pattern in the setting of COVID-19 pandemic. *Radiol Cardiothor Imaging*. (2020) 2:e200280. doi: 10.1148/ryct.2020200280
17. Franquet T. Imaging of pneumonia: trends and algorithms. *Eur Respir J*. (2001) 18:196–208. doi: 10.1183/09031936.01.00213501
18. Koo HJ, Lim S, Choe J, Choi S.-H., Sung H, Do K.-H. Radiographic and CT features of viral pneumonia. *RadioGraphics*. (2018) 38:719–39. doi: 10.1148/rg.2018170048
19. Vilar J, Domingo ML, Soto C, Cogollos J. Radiology of bacterial pneumonia. *Eur J Radiol*. (2004) 51:102–13. doi: 10.1016/j.ejrad.2004.03.010
20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. (2016) 278:563–77. doi: 10.1148/radiol.2015151169
21. Moura LV, Machado Dartora C, Marques da Silva AM. Skin lesions classification using multichannel dermoscopic Images. In: *XII Simpósio De Engenharia Biomédica - IX Simpósio De Instrumentação e Imagens Médicas. Zenodo*. (2019).
22. Wu Y, Jiang J-H, Chen L, Lu J-Y, Ge J-J, Liu F-T, et al. Use of radiomic features and support vector machine to distinguish Parkinson's

disease cases from normal controls. *Ann Trans Med*. (2019) 7:773–3. doi: 10.21037/atm.2019.11.26
23. De Moura LV, Dartora CM, Marques da Silva AM. Lung nodules classification in CT images using texture descriptors. *Revista Brasileira de Física Médica*. (2019) 13:38. doi: 10.29384/rbfm.2019.v13.n3.p38-42
24. Sharma H, Jain JS, Bansal P, Gupta S. Feature extraction and classification of chest X-ray images using CNN to detect pneumonia. In: *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. IEEE (2020). p. 227–31.
25. Attallah O, Ragab DA, Sharkas M. MULTI-DEEP: A novel CAD system for coronavirus (COVID-19) diagnosis from CT images using multiple convolution neural networks. *PeerJ*. (2020) 8:e10086. doi: 10.7717/peerj.10086
26. Ragab DA, Attallah O. FUSI-CAD: coronavirus (COVID-19) diagnosis based on the fusion of CNNs and handcrafted features. *PeerJ Comput Sci*. (2020) 6:e306. doi: 10.7717/peerj-cs.306
27. Al-antari MA, Hua C.-H., Bang J, Lee S. Fast deep learning computer-aided diagnosis of COVID-19 based on digital chest x-ray images. *Appl Intelligence*. (2020) 51:2890–907. doi: 10.21203/rs.3.rs-36353/v2
28. Hemdan EED, Shouman MA, Karar ME. COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. *arXiv [Preprint]*.arXiv:2003.11055 (2020). Available Online at: https://arxiv.org/abs/2003.11055 (accessed Dec 20, 2021).
29. Lundberg S, Lee S-I. A Unified Approach to Interpreting Model Predictions. *arXiv [Preprint]*.arXiv:1705.07874 (2017). Available Online at: https://arxiv.org/abs/1705.07874 (accessed Dec 20, 2021).
30. de la Iglesia Vayá M, Saborit JM, Montell JA, Pertusa A, Bustos A, Cazorla M, et al. BIMCV COVID-19+: A Large Annotated Dataset of RX and CT Images From COVID-19 Patients. *arXiv [Preprint]*.arXiv: 2006.01174 (2020). Available Online at: https://arxiv.org/abs/2006.01174 (accessed Dec 20, 2021).
31. Saborit-Torres JM, Serrano JAM, Serrano JMS, Vaya M. BIMCV COVID-19-: A Large Annotated Dataset Of RX And CT Images From Covid-19 Patients. IEEE DataPort. *arXiv [Preprint]*.arXiv: 2006.01174 (2020). Available Online at: https://arxiv.org/abs/2006.01174 (accessed Dec 20, 2021).
32. Yao Q, Xiao L, Liu P, Zhou SK. Label-Free Segmentation of COVID-19 lesions in Lung CT. *IEEE Trans Med Imaging*. (2021) 40:2808–19. doi: 10.1109/TMI.2021.3066161
33. Calderon-Ramirez S, Yang S, Moemeni A, Elizondo D, Colreavy-Donnelly S, Chavarría-Estrada LF, et al. Correcting data imbalance for semi-supervised COVID-19 detection using X-ray chest images. *Applied Soft Computing*. (2021) 111:107692. doi: 10.1016/j.asoc.2021.107692
34. Duran-Lopez L, Dominguez-Morales JP, Corral-Jaime J, Vicente-Diaz S, Linares-Barranco A. COVID-XNet: a custom deep learning system to diagnose and locate COVID-19 in chest X-ray images. *Appl Sci*. (2020) 10:5683. doi: 10.3390/app10165683
35. Ahishali M, Degerli A, Yamac M, Kiranyaz S, Chowdhury MEH, Hameed K, et al. Advance warning methodologies for COVID-19 using chest X-ray images. *IEEE Access*. (2021) 9:41052–65. doi: 10.1109/ACCESS.2021.3064927
36. Aviles-Rivero AI, Sellars P, Schönlieb C-B, Papadakis N. GraphXCOVID: explainable deep graph diffusion pseudo-labelling for identifying COVID-19 on chest X-rays. *Pattern Recog*. (2022) 122:108274. doi: 10.1016/j.patcog.2021.108274
37. DeGrave AJ, Janizek JD, Lee S.-I. AI for radiographic COVID-19 detection selects shortcuts over signal. *Nat Mach Intelligence*. (2021) 3:610–9. doi: 10.1038/s42256-021-00338-7
38. Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. COVID-19 Image Data Collection: Prospective Predictions Are the Future. *arXiv [Preprint]*.arXiv: 2006.11988 (2020). Available Online at: https://arxiv.org/abs/2006.11988 (accessed Dec 20, 2021).
39. Simpson G, Ford JC, Llorente R, Portelance L, Yang F, Mellon EA, et al. Impact of quantization algorithm and number of gray level intensities on variability and repeatability of low field strength magnetic resonance image-based radiomics texture features. *Physica Medica*. (2020) 80:209–220. doi: 10.1016/j.ejmp.2020.10.029
40. Mali SA, Ibrahim A, Woodruff HC, Andrearczyk V, Müller H, Primakov S, et al. Making radiomics more reproducible across scanner and imaging protocol variations: a review of harmonization methods. *J Personal Med*. (2021) 11:842. doi: 10.3390/jpm11090842
41. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, et al. Development of a digital image database for chest radiographs with and without a lung nodule. *Am J Roentgenol*. (2000) 174:71–4. doi: 10.2214/ajr.174.1.1740071
42. Jaeger S, Candemir S, Antani S, Wáng Y-XJ, Lu P-X, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant Imaging Med Surg*. (2014) 4:475–7. [10.3978/j.issn.2223-4292.2014.11.20]10.3978/j.issn.2223-4292.2014.11.20
43. Gonzales RC, Woods RE. *Digital Image Processing*. 3rd ed. Pearson, Upper Saddle River (2007).
44. Zwanenburg A, Leger S, Martin Vallières SL. Image biomarker standardisation initiative. Reference Manual. *arXiv [Preprint]*.arXiv: 1612.07003 (2016). Available Online at: https://arxiv.org/abs/1612.07003 (accessed Dec 20, 2021).
45. Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339
46. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *J Mach Learn Res*. (2012) 13:281–305. Available online at: http://jmlr.org/papers/v13/bergstra12a.html
47. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*. (2011) 12:2825–30. Available online at: http://jmlr.org/papers/v12/pedregosa11a.html
48. Pal SK, Mitra P. *Pattern Recognition Algorithms for Data Mining*. New York: CRC Press (2004).
49. Zhao ZA, Liu H. *Spectral Feature Selection for Data Mining*. Chapman and Hall/CRC, New York (2011).
50. Chen T, Guestrin C. XGBoost. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM (2016). p. 785–94.
51. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. *arXiv [Preprint]*.arXiv: 1603.02754 (2016). Available Online at: https://arxiv.org/abs/1603.02754 (accessed December 20, 2021).
52. Breiman L. Random forests. *Mach Learn*. (2001) 45:5–32. doi: 10.1023/A:1010933404324
53. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*. (2020) 21:6. doi: 10.1186/s12864-019-6413-7
54. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intelligence*. (2020) 2:56–67. doi: 10.1038/s42256-019-0138-9
55. Sundararajan M, Najmi A. The many shapley values for model explanation. In: III HD, Singh A, editors. *Proceedings of the 37th International Conference on Machine Learning*. PMLR. Proceedings of Machine Learning Research. vol. 119 (2020). p. 9269–78.
56. Lundberg SM, Erion GG, Lee S.I. Consistent Individualized Feature Attribution for Tree Ensembles. *arXiv [Preprint]*.arXiv: 1802.03888 (2019). Available Online at: https://arxiv.org/abs/1802.03888 (accessed Dec 20, 2021).
57. El Asnaoui K, Chawki Y. Using X-ray images and deep learning for automated detection of coronavirus disease. *J Biomol Struct Dynam*. (2020) 39:3615–26. doi: 10.1080/07391102.2020.1767212
58. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp*. (2018) 2:36. doi: 10.1186/s41747-018-0068-z
59. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, et al. Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE Rev Biomed Eng*. (2021) 14:4–15. doi: 10.1109/RBME.2020.2987975
60. American College of Radiology. *ACR Recommendations for the Use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection*. (2020). Available online at: https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection (accessed June 22, 2020).
61. de Moura LV, Dartora CM, de Oliveira CM, Barros RC, da Silva AMM. A novel approach to differentiate COVID-19 pneumonia in chest X-ray. In: *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE (2020). p. 446–51.

62. Saha P, Sadi MS, Islam Md M. EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers. *Inform Med Unlocked*. (2021) 22:100505. doi: 10.1016/j.imu.2020.100505
63. Cavallo AU, Troisi J, Forcina M, Mari P, Forte V, Sperandio M, et al. Texture analysis in the evaluation of Covid-19 pneumonia in chest X-ray images: a proof of concept *Study*. (2020) 17:1094–102. doi: 10.21203/rs.3.rs-37657/v1
64. Rasheed J, Hameed AA, Djeddi C, Jamil A, Al-Turjman F. A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images. *Interdiscip Sci Comput Life Sci*. (2021) 13:103–17. doi: 10.1007/s12539-020-00403-6
65. Brunese L, Mercaldo F, Reginelli A, Santone A. Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. *Comput Methods Prog Biomed*. (2020) 196:105608. doi: 10.1016/j.cmpb.2020.105608
66. Kikkisetti S, Zhu J, Shen B, Li H, Duong TQ. Deep-learning convolutional neural networks with transfer learning accurately classify COVID-19 lung infection on portable chest radiographs. *PeerJ*. (2020) 8:e10309. doi: 10.7717/peerj.10309
67. Yousefi B, Kawakita S, Amini A, Akbari H, Advani SM, Akhloufi M, et al. Impartially validated multiple deep-chain models to detect COVID-19 in chest X-ray using latent space radiomics. *J Clin Med*. (2021) 10:3100. doi: 10.3390/jcm10143100
68. Caruso D, Pucciarelli F, Zerunian M, Ganeshan B, De Santis D, Polici M, et al. Chest CT texture-based radiomics analysis in differentiating COVID-19 from other interstitial pneumonia. *La Radiol Med*. (2021) 126, 1415–1424. doi: 10.1007/s11547-021-01402-3
69. Lin L, Liu J, Deng Q, Li N, Pan J, Sun H, et al. Radiomics is effective for distinguishing coronavirus disease 2019 pneumonia from influenza virus pneumonia. *Front Public Health*. (2021) 9: 663965. doi: 10.3389/fpubh.2021.663965
70. Liu H, Ren H, Wu Z, Xu H, Zhang S, Li J, et al. CT radiomics facilitates more accurate diagnosis of COVID-19 pneumonia: compared with CO-RADS. *J Trans Med*. (2021) 19:29. doi: 10.1186/s12967-020-02692-3
71. Shiri I, Sorouri M, Geramifar P, Nazari M, Abdollahi M, Salimi Y, et al. Machine learning-based prognostic modeling using clinical data and quantitative radiomic features from chest CT images in COVID-19 patients. *Comput Biol Med*. (2021) 132:104304. doi: 10.1016/j.compbiomed.2021.104304
72. Yi PH, Kim TK, Lin CT. Generalizability of deep learning tuberculosis classifier to COVID-19 chest radiographs. *J Thor Imaging*. (2020) 35:W102–4. doi: 10.1097/RTI.0000000000000532
73. Maguolo G, Nanni L. A critic evaluation of methods for COVID-19 automatic detection from X-ray images. *arXiv [Preprint]*.arXiv: 2004.12823 (2020). Available Online at: https://arxiv.org/abs/2004.12823 (accessed Dec 20, 2021).
74. Oh Y, Park S, Ye JC. Deep learning COVID-19 features on CXR using limited training data sets. *IEEE Trans Med Imaging*. (2020) 39:2688–700. doi: 10.1109/TMI.2020.2993291
75. Hwang EJ, Kim H, Yoon SH, Goo JM, Park CM. Implementation of a deep learning-based computer-aided detection system for the interpretation of chest radiographs in patients suspected for COVID-19. *Korean J Radiol*. (2020) 21:1150. doi: 10.3348/kjr.2020.0536
76. Hussain L, Aziz W, Khan IR, Alkinani MH, Alowibdi JS. Machine learning based congestive heart failure detection using feature importance ranking of multimodal features. *Math Biosci Eng*. (2021) 18:69–91. doi: 10.3934/mbe.2021004
77. Hussain L, Lone KJ, Awan IA, Abbasi AA, Pirzada J.-R. Detecting congestive heart failure by extracting multimodal features with synthetic minority oversampling technique (SMOTE) for imbalanced data using robust machine learning techniques. *Waves Random Complex Media*. (2020) 2020:4281243. doi: 10.1080/17455030.2020.1810364
78. Decarlo LT. On the meaning and use of kurtosis. *Psychol Methods*. (1997) 2:292–307. doi: 10.1037/1082-989X.2.3.292
79. Chung A. *Actualmed COVID-19 Chest X-Ray Dataset Initiative* (2020). Available online at: https://github.com/agchung/Actualmed-COVID-chestxray-dataset
80. Kermany D, Zhang K, Goldbaum M. *Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification*. Mendeley Data, (2018). p. 2.
81. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-Ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE (2017). p. 3462–71.
82. Casiraghi E, Malchiodi D, Trucco G, Frasca M, Cappelletti L, Fontana T, et al. Explainable machine learning for early assessment of COVID-19 risk prediction in emergency departments. *IEEE Access*. (2020) 8:196299–325. doi: 10.1109/ACCESS.2020.3034032

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership