Keep Moving Forward: Health Informatics and Information Management beyond the COVID-19 Pandemic

Barbara L. Massoudi1, Diana Sobolevskaia2
1 Rollins School of Public Health, Emory University
2 Georgia State University

Summary
Objectives: To identify gaps and challenges in health informatics and health information management during the COVID-19 pandemic. To describe solutions and offer recommendations that can address the identified gaps and challenges.

Methods: A literature review of relevant peer-reviewed and grey literature published from January 2020 to December 2020 was conducted to inform the paper.

Results: The literature revealed several themes regarding health information management and health informatics challenges and gaps: information systems and information technology infrastructure; data collection, quality, and standardization; and information governance and use. These challenges and gaps were often driven by public policy and funding constraints.

Conclusions: COVID-19 exposed complexities related to responding to a world-wide, fast moving, quickly spreading novel virus. Longstanding gaps and ongoing challenges in the local, national, and global health and public health information systems and data infrastructure must be addressed before we are faced with another global pandemic.

Keywords
Health information management; public health; information governance; health information system; informatics; SARS-CoV-2; coronavirus; COVID-19

Yearb Med Inform 2021:75-83
http://dx.doi.org/10.1055/s-0041-1726499

1 Introduction
Coronavirus disease 2019 (COVID-19) was responsible for a global pandemic beginning in late 2019. The outbreak of COVID-19 began in Wuhan, China on December 31, 2019. COVID-19 rapidly spread across the world and on March 11, 2020 the World Health Organization (WHO) declared a global pandemic [1]. As of December 2020, more than 82 million cases of COVID-19 worldwide have been recorded, along with more than 1.8 million deaths [2].

As the number of infected patients and deaths continued to rise globally, clinicians, scientists, epidemiologists, informaticians, and researchers worldwide struggled to understand all facets of COVID-19. The quickly evolving and growing pandemic revealed and compounded existing problems in the US healthcare system as well as delivery systems worldwide [3]. For example, resource constraints (beds, staff, supplies) on hospitals and health systems became apparent as the number of COVID-19 cases surged. Hospitals and health systems had to quickly scale up their surge capacity [4, 5].

The impacts of the unprecedented challenges of COVID-19 in low-income and developing countries were enormous [6, 7]. In the US, data reflect underlying health and healthcare inequities and disparities, as well as ongoing data and information gaps and limitations. Along with the clinical and public health impacts and concerns of dealing with the pandemic, there has been a concurrent and underlying data and information crisis. Despite the large investment in health information technology (IT) over the past two decades [8], in the US, there was inadequate capacity to prepare for, respond to and mitigate this pandemic.

The pandemic created the need for unique case definitions and ICD codes for COVID-19, but it also exposed existing needs for more timely, robust, and actionable data that are not unique to COVID-19 but were exacerbated by the urgency of the pandemic. These data are needed to: describe and characterize the pandemic and its spread; track and monitor the capacity of healthcare settings and facilities; assess the natural history of the disease and outcomes of patients; and confirm the amount and availability of critical supplies and personal protective equipment at provider sites. Further, critical gaps emerged about the way data are collected, gathered, reported, shared, and analyzed [9]. This paper focuses on the health information management challenges elevated during COVID-19 which require resolution to help ensure a strategic, progressive, and comprehensive approach to respond to ongoing and future public health emergencies.

Due to this paper’s focus on the data and information gaps and challenges experienced during the global COVID-19 pandemic, health information management and health informatics are being considered together. These two fields, while differing in origin, have undergone a convergent evolution in recent years, evolving to adapt to a similar environment, and now share many common foci [10]. For purposes of this paper, public health informatics is considered an important component of the broader field of health informatics [11, 12].
The main objectives of this survey are to:

- Identify gaps and challenges in health information management during the COVID-19 pandemic;
- Describe potential solutions that can help address the gaps and challenges;
- Provide recommendations for policy, investment, and research actions.

2 Methods

To inform the paper, the authors conducted a literature review of relevant peer reviewed and grey literature articles focusing on health information and COVID-19 broadly, published from January 2020 to December 2020 in PubMed. Health informatics and health information search terms included “information exchange”, “information governance”, “health information exchange”, “health information systems”, “health information management”, “digital healthcare”, “electronic surveillance data”, “COVID-19”, “electronic medical records”, and “electronic health records”. Searches for grey literature (i.e., evidence not published in commercial publications), including, but not limited to, government publications, conference papers, and white papers were conducted using Google. Articles for review were drawn from peer-reviewed journals, conference papers, research performed by independent research and academic institutions, and letters to the editor. Other public documents reviewed included blogs and opinion pieces in peer-reviewed journals as well as articles in news media. A snowball approach was taken to identify additional sources of information to ensure as wide a review as possible to identify relevant materials. This included looking at the papers cited by identified papers as well as reviewing any papers citing source documents. Efforts were undertaken to identify relevant resources to depict the international aspects of this paper’s focus.

3 Results

The review of the literature identified many challenges, barriers, and areas for resolution in health information management. These include challenges in information systems and the technologies underlying them; barriers to the collection of complete and accurate information; and challenges related to data and information governance, information systems and technology. Given the magnitude and pace of the COVID-19 pandemic, there was a heightened awareness of the need for secure collection, transmission, exchange, maintenance, and analysis of public health data. Related to this is the need to urgently expand, enhance, and modernize current public health data collection and reporting and to advance interoperability of data flow between and among public health authorities (PHA) and clinical healthcare providers (HCPs) [13].

Broadly speaking, public health functions during a pandemic include early and rapid detection, modeling, and assessment of potential impact, and proactively planning a systematic and iterative response [14]. To fulfill these functions, PHAs must effectively carry out several activities including data collection and sharing; case identification and management; contact tracing; mapping; monitoring of cases and transmission; community risk sensitization and engagement; preventive education; and other timely preparedness and response activities [15]. However, just executing these actions is not sufficient to limit spread and mortality. The entire public health response needs to be coordinated and synergistic, built upon a robust health information system infrastructure where real-time data collection, reporting and analytics underlie every activity, and findings from contact tracing are analyzed along with transmission statistics and disease monitoring to assess the efficacy of mitigation strategies [16]. The COVID-19 pandemic has highlighted the urgency to ensure robust public health prevention, preparedness, and responsiveness within and across several jurisdictions, including at the community, local, state, national and international levels [17].

3.1 Information Systems and Information Technology Infrastructure

In the US, congressional hearings [18] identified information management challenges during COVID-19 including those related to case identification and reporting; syndromic surveillance; electronic laboratory reporting; and contact tracing. These challenges have further emphasized the need for a robust and interoperable public health data and information systems infrastructure [19,20] with an updated focus on information management and real-time, technology-driven, surveillance and reporting to respond effectively to public health emergencies [21,22].

Significant US Federal Government efforts, including the development and release of executive orders and a data-driven framework for pandemic response, are underway and directed toward addressing some of these long-standing issues [23, 24]. Additionally, the US Centers for Disease Control and Prevention (CDC) continues to refine, update, and implement its Data Modernization Initiative through funding from the CARES Act [25].

CDC’s Public Health Emergency Preparedness Capability 6, developed to provide a baseline for what PHAs need to do to be prepared for public health threats, is about information sharing [26, 27]. Industry analysts urged the development of a comprehensive pandemic preparedness strategy including data and analytics using health IT [27]. The lack of interoperability of systems and the inability to share data efficiently between organizations have been a major barrier during the pandemic.

In the US, prior public policy to promote use of health information technologies focused largely on adoption of technologies through incentives, without requiring widespread adoption and use of standards. PHAs and HCPs often do not use the same information systems, data formats, or even data standards, thus hindering interoperability. With COVID-19 specifically, there is a disconnect between the case definition used by PHAs to report cases and the format and standards used by electronic health record systems [28]. These issues existed long before COVID-19 [29], and they are not unique to the US.

To operationalize and improve preparedness, prevention, detection, and response to public health emergencies, there is a need for new data and approaches for public health programs, including those made possible by health IT and digital tools. For example, case identification, investigation, and manage-
Worldwide, government policies have varied in terms of containing, suppressing, and mitigating COVID-19 [5, 47-49]. The emergency response process in Belgium was divided into four phases: information-evaluation-coordination-decision making at the regional level and the final decision making at the federal level [50]. South Korea’s rapidly controlled transmission, while implementing less stringent national social distancing measures than other countries, led to substantial interest in their “test, trace, isolate” strategy [51]. Additionally, there is a need to address the infrastructure and data deficiencies and limitations that have surfaced during COVID-19 as well as insufficient preparedness and surveillance capabilities complicated by health inequalities and the spread of misinformation [52].

3.2 Data Collection, Data Quality, and Standardization

Public health functions and processes in the US are decentralized and rely on information collected by thousands of federal, state, local, tribal and territorial PHAs. Data are submitted from state and local PHAs to federal programs in many ways and through numerous systems. Thus, in the US there are differences in the data definitions contained in federal, state, and local-level data reporting requirements, and across the globe lack of data comparability contributed to the challenges. The need for quality data (complete, accurate, timely, useful, understandable) is a foundational principle of health information management and is essential for healthcare and PHAs. The impact of the COVID-19 crisis differs across countries, but also across regions and municipalities within countries, both in terms of declared cases and related deaths. The information management challenges it presented were recognized across the globe.

Differences in data related to case counts, COVID-19-attributable hospitalizations and deaths, and testing have hindered the ability to detect threats and to learn from the global experience [53]. Efforts are needed to ensure that documentation and data about COVID-19 are captured consistently and accurately.

One of the key needs with the emergence of a new disease is to develop new standard diagnostic codes for the disease and cause of death, codes for new laboratory tests, new treatments, and other procedures. This need for new diagnostic and procedure codes was reiterated and addressed by multiple organizations and agencies including the WHO [54] and the CDC [55]. The Current Procedural Terminology (CPT®) procedure codes were revised by the American Medical Association [56] and Logical Observation Identifiers Names and Codes (LOINC®) codes for laboratory tests were revised as well [57]. Finally, the US Centers for Medicare and Medicaid Services issued new Official Guidelines for Coding and Reporting [55]. However, given the broad scale of the challenges around data collection and use for COVID-19, it is unlikely that the new COVID-19 codes alone will be able to fully address these problems.

The need to prioritize the creation of a national testing strategy along with logistical and laboratory certification barriers led to confusion and a reported lack of availability of certain laboratory testing [58]. This lack of laboratory-confirmed case data has also compromised the ability to make real-time policy decisions, allocate resources, and inform preparedness plans. The absence of timely information and data sharing has led to weak national and international coordination, affecting logistics in supply allocations, triage and testing, quarantine, and treatment options [15]. The need for timely and robust data is essential for public health functions including disease surveillance to inform operational requirements such as hospital capacity planning and resource management and for a broad range of research needs, including studies of virus mutations, patient risk factors, clinical outcomes, and drug trials [59].

Other activities include exploring how recent changes in the US regulatory environment (such as interoperability and increasing patients’ access to their data) may benefit public health and address access timely, accurate, and actionable information flow from EHRs to PHAs. Additional efforts are needed to improve data collection and data flows between HCPs and PHAs using health IT, including EHRs [30], with the
goal of seamless interoperability between providers and public health systems [31]. Although there has been steady growth in bidirectional information exchange between HCPs and PHAs, some PHAs are still unable to receive electronic messages from clinical entities [32, 33]. Areas of consideration for data collection looking forward include new clinical data sources such as pre-diagnosis signs and symptoms from EHRs; data from “non-traditional” sites of care (such as pop-up clinics and mobile testing sites); and earlier, more robust, consistent, near real-time syndromic surveillance capabilities [31, 34-35]. Additionally, as COVID-19 cases increased in the US, data were subject to time-consuming review before being made available. The HHS system that released data suffered from serious time lags [58, 60] and was limited in the information it provided, leading stakeholders to demand greater data transparency.

Many public health and governmental agencies continue to rely on manual processes to gather, structure, and submit data, resulting in incomplete and missing information and therefore loss of data integrity and an inability to predict and map infection surges [61]. Ongoing and renewed attention is needed to enhance the quality, timeliness, and completeness of government-produced health data, by strengthening existing and contemplating new information management practices [62]. With the reliance on university and non-profit data dashboards, COVID-19 also demonstrated the value of public-private sector collaborations to produce and disseminate near real time and comprehensive data.

Globally, reports indicated lack of standard procedures resulting in haphazard data collection, often lacking crucial information. Researchers in China identified similar challenges to those found in the US, including low efficiency of data collection, difficulty in guaranteeing data quality, low efficiency of data use, and lack of timely data sharing [63]. The pandemic demonstrated the limitations of the existing systems for epidemic surveillance and monitoring and exposed gaps in data collection during an emerging outbreak [53].

Even when data are available and able to be shared, data quality issues related to completeness, timeliness, and validity still limit the utility of the data [62]. The ability to manage the pandemic has been hindered by incompleteness and unavailability of critical data needed by health officials to count deaths, particularly of non-residents such as those in hospitals, prisons, or long-term care facilities [64].

There is a need to address vulnerable population and communities disproportionately impacted by the COVID-19 pandemic, including how COVID-19 relates to factors of age, race, ethnicity, disability, and sex. HCPs and PHAs also need to have a better understanding of how COVID-19 has disproportionately impacted various geographical areas, populations, and communities to increase the capability of PHAs and the healthcare system to develop appropriate responses. The absence of complete and timely data has hindered reaching this goal. In some instances, the lack of needed demographic information - race, ethnicity, telephone, occupation, and physical address - has hindered treatment and mitigation actions about potential exposure to COVID-19. Countries have documented that ethnicities, even when required to be reported, are often missing [65]. In the US, assessing whether there was a differential risk of contracting COVID-19 for people of color was nearly impossible because race and ethnicity information was not included in the CDC Persons Under Investigation COVID-19 case report forms 85% of the time [66]. Identified limitations included lack of standard race/ethnicity categories; variations in which racial/ethnic groups for which data are reported; and a high proportion of cases with unknown race/ethnicity [67].

Contact tracing, which depends on identifying and locating infected people for interviewing, was hindered because demographic information was often not available or not willing to be shared by healthcare facilities [68]. Contact tracing and testing efforts have been compromised because basic demographic data elements including telephone numbers and addresses to identify and communicate with patients are not complete or are missing or inaccurate. In the US, the Council of State and Territorial Epidemiologists (CSTE) reported that approximately 80% of COVID-19 test results were missing demographic information [69].

Challenges related to accurately identifying patients and matching patients to their information, data collection and reporting, data quality, information management, misinformation, and deliberately falsified disinformation, as well as limitations of the broad health information infrastructure have characterized the pandemic response [70, 71].

3.3 Governance and Use of Data and Information

Issues related to underfunding of and improvements desperately needed to the public health infrastructure in the US are not new [72, 73]. In response to this need, recommendations for sustained funding for public health preparedness and response programs have grown during COVID-19. Funding efforts have sought resources to support core functions and capabilities of a robust public health data and information technology infrastructure; a well-trained and appropriately resourced public health workforce; modernized data and surveillance capacities and developing evidence-based mitigation and treatments [21, 74]. Chronic underfunding has plagued PHAs across many parts of the world, not only in the US [75]. Despite previous efforts, investment and international assistance for pandemic preparedness remain insufficient [53].

The development, deployment and use of health IT and digital tools (for data monitoring, reporting, and contact tracing) have raised questions about legal and regulatory issues, including privacy and security, and data use and re-use.

While contact tracing is a core public health function and has been used to control the spread of infectious diseases for decades, the more widespread use of digital tools, including smart phones, to potentially expand tracing capabilities raised privacy questions [76-78]. Some researchers have noted the usefulness of this data to public health and of data sharing across geographic borders [79]. Others have raised the need to assess the potential public health benefits of digital tools and health IT against the imperative to protect and maintain the privacy of patient health data [80-85].
In the US, there was an urgent and ongoing need to reduce confusion about the applicability of Health Insurance Portability and Accountability Act of 1996 (HIPAA) requirements during the COVID-19 pandemic. To address these issues, in the US, federal authorities provided explanations and issued formal guidance on how patient health information may be used and disclosed in response to the COVID-19 nationwide public health emergency, based on the HIPAA public health exception. And the US CDC confirmed that all aspects of COVID-19 case investigation and contact tracing must be voluntary, confidential, and culturally appropriate [86]. Public health responses need to ensure that key elements of privacy are not abused or disregarded during emergencies [77, 81].

Additionally, clinical data contained in EHRs that could be useful to better understand individual level risk and protective factors was often not available to public health, because healthcare facilities sought to meet the “minimum necessary” requirements of the HIPAA legislation [87]. Despite decades of education in the provider community, ongoing providers’ lack of compliance with legal requirements for public health reporting beyond the pandemic is a much bigger issue [88].

Concerns and recommendations about the ethical use of patient data during COVID-19 arose globally. While research shows that mobile position data contact tracing is important for epidemic control if it conforms to relevant data privacy regulations [89-91], frequently digital contact tracing appeared to conflict with patient data privacy regulations. Efforts to assess the impact of digital technologies on COVID-19 detection, prevention and treatment have also revealed challenges regarding data privacy and data access. In response to these concerns, international organizations released guidance around the use of these approaches. The WHO issued “Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing” [92]. The United Nations (UN) Personal Data Protection and Privacy Principles, the UN Secretary-General’s policy brief on human rights and COVID-19, and relevant health and humanitarian standards, data collection, use and processing by the UN System Organizations [93] all informed this new area of inquiry into privacy protection. Researchers in China identified similar challenges to those found in the US, including data privacy protection issues [63]. Some countries’ solutions to these challenges used sophisticated technological approaches to preserve privacy while benefitting from mobile phone-based surveillance [94, 95].

Similar to digital contact tracing concerns, scientific, legal, ethical and privacy considerations have been raised in other areas during the pandemic including the development of vaccine passports and digital vaccination certificates [96]. Questions also arose regarding privacy, data access and sharing during tele-health and telemedicine visits [97].

### 3.4 Summary of Health Information Management Challenges and Gaps

The main gaps and challenges to an effective pandemic response related to health information management and health informatics are: (1) lack of standards for information exchange between providers and PHAs; (2) problems in data collection and data quality, especially in terms of completeness and timeliness; and (3) governance, public policies and regulations. The latter included both the lack of policies to facilitate efficient data sharing, contact tracing and data governance, and also included providers’ concerns about privacy regulations that led to incomplete data sharing. Governance and public policy barriers related to chronic underfunding of the public health infrastructure, and lack of sufficient investments in resources (including trained personnel) and infrastructure. Lack of international coordination was also identified as a significant challenge. There are numerous overlapping and intertwined legal, ethical, scientific, technical, technological, health equity, and privacy dimensions affecting how health information was managed or mishandled during the COVID-19 global pandemic. To operationalize many of the data and information systems’ recommendations, other longstanding systemic issues in health information management will also require attention.

### 3.5 Proposed Solutions and Recommendations

Several solutions and recommendations that can be implemented to address needs emerging from the pandemic response related to the management of health information are listed here. While not the focus of this paper, it is important to acknowledge the public health workforce gaps and challenges, especially the need for training and skills in epidemiology, data science, and informatics. Public policies and investment strategies should consider the need for more robust public health workforce capability and capacity. Below are several recommendations for consideration by national and global public-sector agencies and private sector organizations.

- Establish and convene global collaborations to enhance ways to operationalize the need for multilateral responses to public health crises such as COVID-19, including global data sharing using reliable and comparable data [98]. Such engagements could consider various approaches to reduce fragmented efforts both globally and nationally [99, 100] and increase coordinated public health responses;
- Address the need for sustained funding and support for public health data and information systems as well as innovative digital tools;
- Identify lessons learned and leverage best practices gleaned from the COVID-19 pandemic to:
  - Accelerate and enhance efforts to address ongoing global shortcomings of public health data and information systems;
  - Initiate activities that support increasing adoption of consensus-based open-source health IT and data standards, such as FHIR, for public health emergency preparedness and response functions, including prediction and prevention; syndromic surveillance; electronic data collection and reporting; case identification and investigation; immunization administration and tracking; laboratory reporting; and collecting and reporting data for situational awareness;
○ Improve public health IT and data systems to help ensure the efficient and secure collection, transmission, exchange, maintenance, and analysis of the wide array and diverse data that are needed (such as case reporting, hospitalizations, and deaths, and data about testing, treatments, supplies, and vaccinations);
○ Focus efforts to develop, implement and evaluate ways to support real time bidirectional exchange of data between HCPs and PHAs, especially during public health emergencies;
○ Enhance health data collection and reporting practices and processes regarding demographic information for patient identification and matching;
• Advance interoperability of real-time and consistent data flow between and among PHAs and HCPs [101]
○ Address current data infrastructure gaps to improve and increase consistency in data collection, storage, analysis, and dissemination;
○ Explore ways to extend and enhance different data systems standards and interfaces and to enhance data interoperability, quality, and governance [100];
○ Identify common data elements, including minimal or essential data elements, along with standards and implementation specifications to help improve the collection, analysis, interpretation, and reporting of data and information during a public health emergency;
• Identify and leverage global lessons learned from COVID-19 to implement strategies, public policies, practices, and processes to anticipate and mitigate future emergencies and threats:
○ Review existing practices around privacy, confidentiality and security of healthcare and public health data to enhance data sharing and exchange during emergencies;
○ Educate HCPs about the HIPAA public health exception, its application, and their duty to report notifiable conditions to their PHAs;
○ Develop public policies to address the need for real time access, exchange, and use of public health data for case identification and tracking, contact tracing, and treatment, and for research, evaluation and analytics while ensuring data governance, quality, and oversight;
○ Develop policies to address coordination and guidance at the global, national, and local levels;
○ Develop strategies to improve the electronic exchange of health information between federal, state, local, territorial, and tribal PHAs and HCPs and facilities;
○ Consider approaches for de-identification and/or anonymization solutions for the public health use case, especially during emergencies and global pandemics.

4 Conclusion
COVID-19 exposed the global complexities and challenges related to a world-wide, fast moving, quickly spreading novel virus. While effective testing and vaccination will likely bring COVID-19 under control before the end of 2021, there are many other infectious diseases that could pose a serious threat to human health, and for which many have no vaccine or effective medications [102]. Longstanding gaps and ongoing challenges in the local, national, and global health information systems and data infrastructure must be addressed before we are faced with another global health crisis. Enhanced interoperable healthcare and public health information systems, information management, and standardized data collection and reporting are essential components for ongoing healthcare and public health emergency preparedness and resilience.

References


78. Sookman B. AI And Contact Tracking: How to Protect Privacy While Fighting The COVID-19 Pandemic. [Internet]. 2020. Macdonaldlaurier.ca. Available at: https://macdonaldlaurier.ca/publications/files/pdf/20200416_COVID-Privacy_Sookman_COM


102. 10 infectious diseases that could be the next pandemic [Internet]. Gavi.org; 2020 [cited 2021 March 7]. Available from: https://www.gavi.org/vaccineswork/10-infectious-diseases-could-be-next-pandemic

Correspondence to: Barbara L. Massoudi, MPH, PhD
Department of Biostatistics & Bioinformatics
Rollins School of Public Health at Emory University
E-mail: bmassoudi@emory.edu