Digital Health COVID-19 Impact Assessment: Lessons Learned and Compelling Needs

Peter Lee, PhD, Microsoft; **Amy Abernethy, MD, PhD,** Verily; **David Shaywitz, MD, PhD,** Astounding HealthTech; **Adi V. Gundlapalli, MD, PhD, MS,** U.S. Centers for Disease Control and Prevention; **Jim Weinstein, DO,** Microsoft; **P. Murali Doraiswamy, MBBS, FRCP,** Duke University School of Medicine; **Kevin Schulman, MD,** Stanford University School of Medicine; and **Subha Madhavan, PhD,** AstraZeneca

January 18, 2022

> **About the NAM series on Emerging Stronger After COVID-19: Priorities for Health System Transformation**
> This discussion paper is part of the National Academy of Medicine's Emerging Stronger After COVID-19: Priorities for Health System Transformation initiative, which commissioned papers from experts on how 9 key sectors of the health, health care, and biomedical science fields responded to and can be transformed in the wake of the COVID-19 pandemic. The views presented in this discussion paper and others in the series are those of the authors and do not represent formal consensus positions of the NAM, the National Academies of Sciences, Engineering, and Medicine, or the authors' organizations.
> Learn more: **nam.edu/TransformingHealth**

Over the last decade, some of the digital technologies that have profoundly transformed industries from banking to media have, at last, arrived in health care. A medical records system that only 20 years ago consisted mainly of handwritten notes stored in patient charts is today almost entirely digital. Radiological images are acquired, stored, and viewed digitally. Prescriptions are transmitted and reimbursed electronically. Hospital bedside devices for monitoring patient status, and even the location of equipment such as hospital beds, are tracked electronically. In more advanced systems, distributed sensors monitor not only equipment but also the vital signs, weight, heart rhythm, and movement of patients. And, in what may prove to be the most transformational development of all, the promise of artificial intelligence (AI) is now revealing itself in en-

hancing the detection of diseases and reducing errors by intelligently assisting the interpretation of blood tests, electrocardiograms, images from radiology, pathology, ophthalmology, and beyond.

The medical impact of these technologies is also being felt outside the hospital, as affordable consumer technologies encourage a growing number of patients to exercise more (through activity-tracking smartwatches and Internet-connected home exercise machines), eat better (via nutrient-counting apps and self-improvement apps), and choose among a wide array of customized health care options (through the use of websites to obtain reviews of providers, fill prescriptions, and more). Telemedicine offers the opportunity to obtain care without the disruption of traveling to and from a doctor's office, and AI-powered chatbots

> **BOX 1 | Definition of Digital Health**
>
> Digital health technologies encompass a broad range of tools, including "mobile health (mHealth), health information technology (IT), wearable devices, telehealth and telemedicine, and personalized medicine. Digital health technologies use computing platforms, connectivity, software, and sensors for health care and related uses."
>
> SOURCE: U.S. FDA. n.d. What is Digital Health? Available at: https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health (accessed January 10, 2022).

are providing consumers with convenient 24/7 access to health care expertise.

Indeed, digital health technologies (see *Box 1*) are starting to approach the promise of health care delivery that is no longer sporadically provided, confined to the four walls of a hospital, and built around the convenience of the physician. Instead, they are allowing for a people-centered, collaborative approach to continuous health and wellness. The evolving digital foundation of a person-centered health care system is making it possible to envision a system that is more holistic, centers on the needs of the patient and their support structure, and embraces a longitudinal view of health, wellness, and social equity, in contrast to the mostly fragmented, reactive health care system that currently exists.

COVID-19 arrived in the context of such promise and demonstration of opportunity—the first global pandemic of the digital age. There have been many shining examples of how digital health solutions have helped in critical ways during the pandemic. Perhaps the most noticeable acceleration, both in the United States and other parts of the world, has been in the rapid adoption of telemedicine, but there have also been less visible digitally-dependent advances that are just as important across all sectors of health care, public health, and medical research. In many ways, the response to COVID-19 sparked years of advances in mere months.

However, the pandemic also revealed important limitations to digital health technologies and exposed significant challenges and equity concerns. One of the most significant lessons learned in the U.S. is that digital health's ability to help address the pandemic is dependent on a coherent and accessible data infrastructure. Despite the digitization of information made possible by the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act, various critical health care data sources are simply not yet ready to be put to use [122]. This can be crippling in key situations, because data fuel digital technologies that

ultimately support people—both those who require care and those entrusted with delivering it.

Synopsis of the Crisis, Through a Digital Lens

In the early days of the rise of COVID-19 in the U.S., the health care industry, as well as local and federal leaders, sought answers to many critical questions, including the following:

- Who are the most vulnerable people, and where is the infection spreading?
- How many COVID-19 patients does each health system serve, and what is the system's capacity for treating them?
- What capacity does each health system have for testing, and who should get tested and when?
- How can the health care community best triage patients who may be highly infectious?
- Does each health system have an adequate supply of personal protective equipment (PPE), intensive care units (ICUs), and ventilators—and most importantly, do they have the appropriately trained and adequately rested staff that are required to deliver care and monitor complex equipment?
- Should care be redirected to designated institutions, and should some of America's major referral systems be allowed to continue providing routine medical care, designating specific facilities for the pandemic, while others manage care for those who cannot afford to miss or go untreated for pre-existing chronic conditions?
- For each COVID-19 patient, what are the key data elements of treatment and outcome, and what does a population-scale analysis of these data elements tell us about best practices?
- For COVID-19 patients with comorbidities or already on a course of medication, what does population-scale analysis of these treatments and outcomes tell us about risks and treatment effectiveness with near real-time data?

- What are the best treatment protocols for people with COVID-19 and other diseases, especially when the pandemic has interrupted usual capabilities for in-person evaluation and care?
- How can clinical researchers conduct clinical trials and keep study participants safe when they cannot conduct in-person visits or evaluate treatment effects?

The U.S. health care community looked to the interconnected system of devices, digital platforms, and data to help address these questions, since surely the answers lurked within the petabytes of digital data being generated daily by the health care system. At first glance, the task seemed simple enough: these digital data only needed to be extracted, integrated, and disseminated in useful forms with the use of a wide range of digital tools such as telemedicine, biosensors, easy-to-use digital apps, machine learning, and AI. Although these tools are relatively commonplace in many industries, the health care industry has struggled to leverage them [123]. Despite notable strides in digital interoperability, health care interoperability still requires a significant architectural mobilization of largely ad hoc collaborations and new system deployments.

Telemedicine proved to be an example of successful digital impact. More generally, however, society's lived experience with emerging technologies was often far short of expectations [1]. The answers to the pressing questions listed above fell far outside the normal operational capabilities of health systems, and in crisis-response situations they often eluded stakeholders for critical periods of time, highlighting the tremendous gap between existing raw data and urgently needed aggregation and insights. Technology may provide the tools, but solutions require the capacity for successful implementation, turning the promise into real-world practice, especially for the most vulnerable patients and communities. Even the implementation of telemedicine, widely lauded as a success within the pandemic, was not distributed equally and resulted in variable access to care for seniors, as well as Black and Hispanic patients and rural communities separated by the widening digital divide [124].

As a result, during the initial stage of the pandemic in the U.S., decision-makers were essentially flying blind. Electronic health record (EHR) systems were mired in a sea of codes, few of which pertained to COVID-19, due to its novelty [2]. These systems were not connected to enterprise resource planning systems, and thus lacked the ability to correlate relevant patient encounters with human resources and physical capacity. The utilization of testing, PPE, beds, and ventilators varied within and across each and every health system (and often varied even across departments within a single hospital or clinic) [3]. Public health departments in charge of implementing rules, policies, public guidance, and contact tracing operations each operated within their own data silos—often taking the form of piles of spreadsheets—and were almost always unconnected and non-interoperable with any other health care information technology (IT) system [4,5]. In too many cases, the only effective communication of data between health care delivery systems and public health agencies was through a fax machine [6].

Scores of medical researchers diverted their attention to patient treatment and compassionate application of experimental treatments, often discovering critical, life-saving insights while providing this care. However, the sharing of these insights through effective digital channels was initially done in ad hoc ways (often through social media) and well outside the traditional channels used for medical research. These structured and unstructured data, the biomedical communities' ideas and experiences, and newly developed digital tools were trapped in the urgency of crisis response. Applying even rudimentary machine learning or AI tools in ways that would inform or persuade other clinicians or regulators was well beyond reach, in part because these tools required voluminous, ready-to-use data [7].

Legions of technologists rushed to address these crises in access, connectivity, and interoperability, and achieved some great successes through heroic and unprecedented collaborations, some involving thousands of health care and technology experts and their organizations. However, in the process, these efforts often resulted in the creation of yet more data silos and more digital platforms that not only struggled to interoperate with the rest of the health care ecosystem, but also contributed more staggering, ineluctable complexity. Ongoing challenges in vaccine distribution and monitoring are the most current and urgent example of the existing limitations of data visibility, fluidity, transparency, and access.

One is reminded of the poem by Samuel Taylor Coleridge, *The Rime of the Ancient Mariner.* It contains the famous verse, "Water, water, everywhere, nor any drop to drink [8]." Despite nearly complete digitization of health care data, and an abundance of tools available for data analysis, machine learning, AI, and visualization, the health care community expended

far more effort than should have been necessary to quench its thirst for high-quality, actionable data upon which these technologies, patients, and caregivers foundationally depended. Data were needed not only from health systems, but also from all other relevant sources—personal, social, infrastructural, biological, population-wide, and more.

How did the U.S. find itself in this situation, despite possessing unimaginably powerful digital capabilities? Imagine for a moment that we are setting out to build a house. We, of course, would need good tools, an adequate supply of lumber, and an understanding of the architecture of the house we are trying to build. However, if we lacked the components required to support the process of construction—skilled tradespeople, heavy equipment, building inspectors, and other infrastructure—it would be impossible to connect the tools and lumber to the architecture and realize a completed house. Furthermore, without a modularity that is both intentionally designed and defined, as we see in industry standards and building codes, the orchestration of architectural components such as electrical systems, plumbing, roofing, and heating would be wildly complex and unwieldy.

Even more important, innovators who make technological advances in those component systems would find it hard to survive in the marketplace because they would not have standard places to "plug in" their new ideas at industrial scale. Indeed, such innovators, out of desperation, would likely find themselves forced to stray into other domains, as well as make moves to protect themselves from new competitors in order to stay viable. In such a scenario, home construction would likely be a low-productivity, artisanal activity—much like early automobile production.

This is exactly the situation that is occurring in today's health care data ecosystem. In digital health, it is not enough to have the AI tools and the data (that is, the "digital lumber"). Addressing the nation's deficiencies will require an overall system architecture, with modular components that allow innovation to flourish, and an infrastructure to support that architecture all the way from design to coordinated implementation, safe deployment, managed evolution, and continuous feedback. To extend the metaphor, health data must advance from its current artisanal state and achieve industrialization.

These concepts of data *architecture, modularity,* and *infrastructure* are foundational needs in medicine and health care, just as they have been shown to be in areas such as global telecommunications, supply chains, and more. Achieving digital transformation in these areas requires not only technological advances, but also new organizational structures involving key public-private partnerships. The goal in this paper, then, is to present lessons learned from COVID-19 that may inform any plans made for meaningful progress along these lines. This discussion paper examines how the current digital health infrastructure and applications have both supported and hindered management of the COVID-19 pandemic, using the insights to extract lessons learned and develop a set of requirements and conditions for future progress. In parallel, and informed by this work, the National Academy of Medicine's Leadership Consortium: Collaboration for a Value and Science-Driven Health System is developing a comprehensive framework for advancing progress in digital health. While this paper focuses on the COVID-19 experience, its development has been coordinated with that broader proposal, particularly as it relates to the concept of a learning health system (LHS). As such, what is presented in this paper can be viewed as a practical living example—a "use case"—of the key desired current and anticipated features of the LHS model.

The first challenge President Joseph Biden directed the Office of Science and Technology Policy (OSTP) to address in 2021 was, "What can we learn from the pandemic about what is possible—or what ought to be possible—to address the widest range of needs related to our public health?" [9] The specific needs highlighted by President Biden—including the need to "dramatically improve our ability to rapidly address [biological] threats," the need to "dramatically speed our ability to develop and conduct clinical trials of therapies for other types of diseases like cancer," and the need to "enable the rapid sharing, with patient consent, of health information to build a smarter and more effective health care system"— are central issues that were also independently highlighted during the development of this paper [9]. Critically, this paper is foundationally rooted in the vital equity imperative captured by another challenge issued by President Biden : "How can we guarantee that the fruits of science and technology are fully shared across America and among all Americans? [9]"

Digital Health: Accomplishments and Opportunities Across the Health System Sectors

This paper is one of nine sectoral assessments that, together, provide a coordinated analysis of the health care system's response to the COVID-19 pandemic [10]. Across all of these sectoral assessments (the *Emerging Stronger After COVID-19* series), the achievements and

challenges of digital technologies are remarkably prominent. Put together, they reveal important accomplishments, as well as opportunities for improvements, in the ways that digital data technologies can and should be harnessed for crisis response and better resilience for managing day-to-day patient care functions in the future.

Table 1 summarizes the elements of digital health as they appear in the other sectoral assessments. The elements can be grouped under the following themes:

- Telehealth became real, practical, and essential during the COVID-19 response.
- Data proved critical for care coordination, forecasting, and quality improvement, but data collection was a time-consuming and sometimes a chaotic burden on clinicians and administrators.
- Data interoperability and scaling proved to be more aspirational than reality in health care delivery and public health assessments.
- Effective public-private partnerships proved essential in crisis response.
- The digital divide was occasionally bridged but continued to contribute to and often exacerbated health inequities.
- Digital tools, including AI, became key to advancing knowledge and coping with information overload.

Common across all the papers in the *Emerging Stronger After COVID-19* series is the recognition that to achieve the goal of a learning health care system built around and foundationally focused on sustaining the health of individuals and enabling the care of patients, there must be relevant, fluid data flowing within an agile yet robust infrastructure [10]. All tasks of digital health care, including "generat[ing] new health-related knowledge, monitor[ing] its application, predict[ing] response, and guid[ing] courses of action," depend on data—data quality, data analysis, and the patient-centered implementation of results [10]. When the data infrastructure is limited, the capabilities of technology and the effectiveness of the health care system as a whole are severely constrained. Data can and should be an effective mechanism to align learning and health care delivery with the requirements of individual patients, provided that their perspectives and needs are intentionally incorporated.

Lessons Learned During the COVID-19 Response

The limitations revealed by the digital health field's encounter with COVID-19 highlights a recurrent lesson in the history of technology development: The discovery of a new technology does not lead immediately to its gainful application [125]. Characteristically, the initial emergence of a powerful new technology, or series of technologies, requires a long period of subsequent, generally iterative, innovation, often by early adopters who begin to understand how to harness the potential of the technology and drive its reduction to practice [11]. Famously, when the electric generator initially replaced steam power, factory owners swapped out one power source for another, realizing only minimal gains in productivity. It was not until new innovators fundamentally restructured the workflow of factories, coupling technology innovation with business model innovation, that the large productivity gains enabled by electricity were achieved [12]. Harnessing the power of advanced information technologies such as AI will require an equally fundamental restructuring codified in numerous innovations.

One way of evaluating the COVID-19 experience is to understand that the pandemic brought a unique group of latent use cases to the forefront in health care delivery and health policy. These use cases quickly overwhelmed the capacity of current technology. While the pandemic made these shortcomings visible in a very public fashion, the challenges of similar use cases have been described frequently over the last decade in the medical and gray literature. A national technology review should be conducted to understand the ways in which the current system failed and to consider ways to address the gaps that hold back our ability to use technology to fully achieve health care that is effective, efficient, equitable, enhances the patient experience, and saves lives.

Historically, health care solutions have evolved functionally, along disciplinary lines. The public health, clinician, payer, life science, and patient-advocacy communities (and the many sub-communities within these categories) have understandably focused on solving the problems they each view as most relevant. This has resulted in an assortment of often very different solutions in which patients and communities have not always been the focus. Many of these solutions, however, involved common data elements that were not aligned across organizations or disciplines. The grand challenge of responding to the pandemic and mak-

Sectors	Digital Health Challenges and Opportunities					
	Tele-health	Data for Coordination, Forecasting, and Quality	Data Interoperability	Effective Public-Private Partnerships	Health Inequities	AI Tools to Address Information Overload
Health Product Manufacturers	Remote monitoring of clinical trials was leveraged	Design of COVID-19 clinical trials was critically dependent on up-to-date pandemic information		R&D alliances enabled sharing, in digital form, pre-clinical and clinical protocols, plus predictive models, without jeopardizing IP or competitiveness (e.g. COVID-19 vaccine development)		
Clinicians & Professional Societies	A variety of virtual mental health services emerged	Care continuity between EHRs and telehealth systems was addressed; further, certification and accreditation, typically conducted in-person, shifted to a mostly virtual mode	Interoperability was critical for emergency preparedness; outdated infrastructure affected credibility through data backlogs, glitches and lost test results		Social and behavioral risk data emerged as a key to improve services; the lack of cultural training for providers impeded equitable access	Clinicians found it hard to manage the volume of pre-prints with relevant clinical information; there was a lack of remote learning for continuing medical education and moderated social media groups to share timely and relevant information
Payers		Risk of fraud/abuse potential in virtual care was reduced through digital analytics; opportunities remain to improve cost benchmarks for value-based payment arrangements and alternative payment models	The lack of standardized data elements affected uniform payments		Most payers had limited to no access to race and socioeconomic data	
Care delivery organizations	Virtual and bot-based triage clinics emerged; tele-ICU expanded	The support for demand forecasting and planning emerged as a digital need	Data interoperability was lacking between care delivery organizations		Supply chain issues for PPE and testing predominantly impacted home health care workers	

Sector						
Quality and Safety		The administrative burden of quality measurement and reporting became overwhelming	The lack of standardized data elements affected the usefulness of reporting (e.g. race and ethnicity)			
Patients, Families, and Communities	Expanded telehealth to support "all"		Digital means to address patient privacy concerns emerged as a critical need		The deepening digital divide became more pronounced: Access to essential technology and reliable internet created inequities; inequities also arose from a lack of language translation and other communication challenges due to sensory deficits of patients	
Public Health		Challenges encountered with electronic case reporting and disparities in the quality of data collected were addressed	Outdated infrastructure impeded facile exchange of information; no scalable way to keep track of millions of cases	The lack of deeper data on communities beyond county/ZIP code information impeded planning and forecasting		
Research	Virtual clinical trial monitoring emerged in practice		Variation in data systems and reporting capacity contributed to challenges for facile enrollment and follow-up for COVID-19 clinical trials			Innovative analytics, evidence accelerators, and disaster response research programs emerged; current research networks were repurposed to address the pandemic (e.g. PCORI)

TABLE 1 | Digital Health Challenges and Opportunities Revealed in the Eight Other Sector Papers
SOURCE: All sector papers outlined in the table can be found at nam.edu/TransformingHealth

ing digital health more robust is developing "yes, and" approaches and solutions that remain highly responsive to critical local needs, contribute to broader data needs beyond those related to the COVID-19 pandemic to serve patients and communities, and to enable improved analysis and upgrades in the quality and equity of care.

Achieving these goals is a tall order and one that is likely to be addressed most effectively by iterative experimentation, rather than a fixed prescription or defined recipe. It requires a constant focus on the needs of the individuals whose health must be maintained, and, when required, restored. Yet this pursuit must be guided by key learnings and transcendent principles learned from challenges with COVID-19 as well as the important, if limited, successes achieved.

Over the next several subsections, this paper contemplates digital health lessons learned during the response to COVID-19, integrating and expanding upon key themes surfaced in other papers in the *Emerging Stronger After COVID-19* series [10]. This is followed by a discussion on observations for the future and a final section on key priorities to inform a vision for a better future.

Data Without Architecture Leads to Data Siloes

As has been witnessed throughout all sectors of medicine and health care delivery, the COVID-19 pandemic created a critical and urgent demand for data to answer the questions posed in the opening section of this paper. This demand, in turn, elicited a determined response by technologists and technology companies to create and deploy systems to make that data available. Hundreds, and perhaps thousands, of new data systems were created and deployed with incredible speed to allow hospitals and health systems to forecast COVID-19 capacity and utilization, improve connectivity between health care delivery and public health operations, create evidence and share best practices for treating a novel disease, and more [13,14,15]. While perhaps less visible to the general public than the rapid rise of available telemedicine services, the deployment of these data systems was no less impressive and no less important to the response to the pandemic. These data system deployments successfully harnessed the technological progress of the past 50 years.

Although these elements of response were incredibly important and undoubtedly saved lives, the systems fell far short of what was and is actually needed as a foundation for coordinated national response and patient empowerment. In the middle of the pandemic,

with speed being of the essence, the lack of coordination around a common architecture meant that nearly all of these new data systems struggled to interoperate with each other. This happened either intentionally by design or, more often, because the task of interoperability was left to a nebulous day in the future in the name of haste. While enabling interoperability in any single component system might, in most cases, require relatively modest engineering effort, when multiplied by the thousands of data systems across counties, states, and nations, any aspiration towards a unified data asset became infeasible in practice. The tremendous scientific and technical advances in machine readability that power global supply chains, massive retail markets, internet search, social media, and more are a stark contrast to the creation of yet more inaccessible data silos in health care.

The lack of available data and digital interoperability was particularly acute when seeking to understand the number and trends of COVID-19 patient encounters, the capacity of the health system to treat those patients, and developments in the utilization of that capacity. In an effort to provide more visibility, the U.S. Centers for Disease Control and Prevention (CDC) published new data modules to standardize the reporting of encounter, capacity, and utilization data elements [16]. In March 2020, then Vice President Michael Pence issued a request to all health care delivery organizations in the U.S. to report such data on a daily basis [17]. While this step was tremendously helpful, it was disappointing to see the nation reduced to asking every over-stressed hospital and clinic to take on the additional burden of gathering data manually, filling out a spreadsheet, and emailing it in every day—a method that would hardly be tolerated in any modern manufacturing, supply chain, e-commerce, or logistics system.

The fundamental lesson is that without coordination around a comprehensive data architecture, as exists in other industries, all of the digital tools and data assets of the past two or more decades are decidedly less useful than expected. ***As the U.S. looks to the future, it will need to embrace the importance of data architecture for any coordinated national or international response to health crises, and find effective ways to define such an architecture and then create the infrastructure to put it into action.***

Right-Sizing Health Care Regulation Can Improve Patient Care in a Hurry

Improving and optimizing health care delivery requires the development and uptake of innovations, such as

digital health solutions, especially telemedicine, which clearly enabled health care delivery during the pandemic. One of the enormous challenges of innovation in health care is the lack of opportunity to test innovative solutions in the clinical environment. From the perspective of innovation theory, innovation occurs most often in unregulated spaces [18]. While there were many factors that contributed to the nation's rapid uptake of telemedicine in response to the pandemic, both in terms of technology innovations and policy prescriptions, a key moment was the public statement by the U.S. Department of Health and Human Services (HHS) Office of Civil Rights (OCR) that it would use its discretion with regard to enforcing Health Insurance Portability and Accountability Act (HIPAA) provisions during the public health crisis. Overnight, a debate that had raged for over two decades about virtual medical visits was resolved with a massive migration of clinical care to digital platforms and, consequently, no widespread reports of data or privacy breaches [19]. Indeed, it is possible that proactive policies that extend this type of posture in enforcement discretion may help speed innovation in other areas of health care as well [21].

The key lesson from the COVID-19 response is that such "right-sizing" of regulation can be extremely important and productive. For example, HIPAA was implemented in the 1990s to protect individuals from misuse of their data by specific third-party covered entities such as health care providers or insurers in a world where consumers had no access to their data and no insight into how their data were being used. While the privacy of health information remains a concern today, individuals also need to be able to share their health information with trusted people, organizations, and digital applications of their choosing that are not currently covered under HIPAA. Meanwhile, consumer privacy laws have been rapidly evolving, including the General Data Privacy Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA), which serve to protect information that is shared by consumers and patients with third parties [22,23]. The general consumer data privacy approach ensures that there is a consistent set of standards across the market, rather than a patchwork of different data privacy standards based on the use of the data or categorization of the service. Such modern standards could focus on questions like: Who owns the data? How can and will consumers use data? What will be the impact of data on each individual's, and the nation's, overall health and well-being? What are the key measures of success?

The expansion of telemedicine was also facilitated in part by changes in clinical licensing requirements and assurance of payment. Medical licensure occurs at the state level, but in the midst of a crisis, state-based licensure regimes were relaxed to allow movement of essential clinical staff across state lines. National or inter-state cross licensure could facilitate access to digital services (especially for those in rural areas), removing one barrier to interstate provision of care. At the same time, the migration to telemedicine was accelerated by the implementation of site-neutral payment policies ensuring equal reimbursement for virtual and in-person visits. Once the pandemic recedes, business models predicated on facility fees for visits may find it financially challenging to provide virtual services.

Current Health Care Data Systems are Inadequate for Longitudinal Patient Care and Data Needs
A variety of technology-related challenges in health care delivery were highlighted during the pandemic. The first major challenge was related to connectivity— the ability to gather and aggregate data about individual patients over time. For example, a typical use case might be to discharge a patient with a documented COVID-19 infection from a hospital to a rehabilitation facility with a requirement for supplemental oxygen. Now, consider the need to monitor the patient throughout. The hospital has an EHR that includes the clinical data related to the hospital stay. The rehabilitation facility has a separate record system related to the stay in that facility. The patient might have an oxygen sensor that provides data that is recorded on their cell phone (if at all). There is no system today that aggregates data from these three sources.

Although this is a simple use case, this type of scenario illustrates how a health care system's capacity became stressed and how the ways in which patients accessed health care resources became less conventional as the pandemic progressed. Failures often occurred because the technology architecture focused on the needs of the individual providers, not the needs of the patient. An alternative architecture, such as a personal health record (PHR), would in principle enable a more patient-centered ability to capture and share data, and thereby provide significant clinical value. A PHR architecture puts patients at the center of health care information, including control over the access to those data. As a means of organizing health care information, a PHR does not have to require the patient's direct interaction but can be viewed as a "canvas"—an "organizing principle" for health data—that enables pa-

tients to aggregate and share data for their own health benefit. As highlighted in a situation such as COVID-19, where patients may be obtaining health care from a diverse set of providers and via a diverse set of delivery means (such as telehealth), the "portability" of a PHR should, in principle, benefit all stakeholders in a patient's care and safety.

PHR architectures have been proposed to the HHS Office of the National Coordinator for Health Information Technology (ONC) several times, but have never been adopted due to privacy concerns or concerns about data overload [126,127]. Nevertheless, a growing number of health care systems around the world today are seeking ways to adopt this approach. However, the fundamental lesson is that in order to evaluate a system architecture, it is important to prioritize the specific use cases for the technology, as well as the consumer's need and willingness to engage with it. This statement may seem simple, but clarifying that *the primary use case of health information technology is patient care from the patient perspective* helps to advance the deployment of technology that may better optimize the delivery of preventive services, primary care services, and chronic care that is safe and effective [24].

There are, of course, many important secondary use cases for technology as well, and they are readily apparent in improving health care quality, research, and public health. Yet these secondary use cases too often receive priority in policy discussions. These use cases are essential, even crucial, for a high performing health care system, but they have to be developed in ways that do not overshadow the primary use case of providing patient care from the patient perspective. To take one common example, there are often use cases that depend on a wireless networking environment, with high-speed data and video services available for patients. Unfortunately, access to high-speed broadband services today is limited in many high-needs environments in the U.S, creating a digital divide and disadvantaging crucial populations of people from receiving care. With a patient-centered approach, this infrastructure challenge immediately illuminates a key factor in health care disparities and digital technology today.

A second challenge in today's health care data systems pertains to productivity. Many physicians became overwhelmed with clinical duties in the midst of the pandemic but did not experience even temporary relief from burdensome documentation requirements. Challenges with electronic health records, including the user interface, length of notes, and administrative tasks—which the majority of clinicians are required to use for patient documentation—are all well documented [25,26,27,28]. From a clinician perspective, EHR systems are not optimized for patient throughput and have limited flexibility to respond to volume by adjusting data entry requirements. Further, most EHR systems have few tools to help clinicians locate and prioritize the essential information for each clinical encounter. This is not the case in other major industries that rely on digital documentation. For example, airplanes generate tremendous amounts of data, but pilots are not tasked with sorting through them. Pilots are provided with heads-up displays that share only the information required for immediate decision-making.

While the burden of clinical documentation and the lack of adequate user interfaces are hardly a new lesson learned during the COVID-19 response, the pandemic did highlight that an overstressed health care system was unable to cope with the need to collect even more data and, more important, effectively focus on the key data elements most pertinent for crisis response. Clearly, a major effort must be made to reduce this burden and thereby create greater capacity for smart data collection in times of greatest need.

A third challenge is the distinction between *data* and *services* in the technology environment. High-fidelity clinical data provide a record of clinical conditions, treatment, and response over time. This was the primary description of data in the paper record environment. However, many conceptualizations of the EHR were aimed at reducing the burden of paper record storage by creating electronic files. This solved the storage and transmission issue but did not lead to a transformation of care. Today, the value is in how data are used. Rather than simply being stored as a record, data have become the resource driving our most advanced digital technologies. Services are the benefits we receive from the data. In the broader economy, data are used to power technology such as machine learning (ML) in the pursuit of optimization of attributes of digital services for consumers. Services range from digital tools providing insights to complete experiences such as online banking, digital shopping, and streaming video entertainment.

The transformation from data to services can be illustrated by some of the shifts that have occurred in libraries. Libraries used to house vast collections of documents, which served as the data for researchers. A revolution in data storage occurred when new technologies allowed for paper records to be transferred

to microfiche, dramatically reducing storage costs for the library, and in the process making more data accessible by making it easier to store more documents. Nonetheless, the retrieval process remained the same: Researchers went to the library to access the data. There were no new services created by this advance in data storage that improved research access or productivity. A second, more significant, revolution occurred when we moved from microfiche records to machine-readable digital records. People could now use the data to power new digital services such as online search and retrieval for documents. Research projects that once required weeks or months of sitting at microfiche readers to review documents and curate reference lists can now be completed online, from any location with internet access, in minutes or days. Even more important, the patterns of use across individuals can be collected and used to inform even more productive and intelligent use of the data. The leap in productivity came from the digital services which allowed better access and curation of data for the user (PubMed and online access) [29].

Similarly, medical knowledge was previously contained in voluminous textbooks, laboriously curated and updated every few years with "new editions," which contained the latest synthesis of the medical literature. This resource was only available in the medical library, far away from the provision of patient care. Now, the data from these textbooks have been transformed to a continuously updated service called Up-To-Date which is available everywhere—from home, in the ward, in the clinic, and on clinicians' phones [30]. This digital service has been so transformative that the medical certification process—which used to require mastery of the content of medical textbooks—is now focused on the application of knowledge by including the Up-To-Date service as a tool on medical board re-certification exams.

Data are of limited value when they are static, or when they are irretrievable in machine-readable and "clean" condition. Yet in health care, the value of data is limited by design, regulation, and practice. The core of provider-based EHR systems was developed using decades-old architectures (and a half-century old computer language), all developed long before the ability to apply ML to the data. These EHR systems were built as data repositories as their original use case. In essence, "data" is not an action word in health care.

Further, there is little conceptual or business linkage between data and services. Again, consider a simple use case of helping a patient to manage their diabetes at home. After a decade of intensive research, clinicians now have closed-loop pumps and sensor systems that can assist Type I diabetes with daily glucose monitoring efforts. However, the majority of diabetes patients have Type II and use smart glucose meters to collect data that are uploaded to free-standing smartphone apps but do not integrate with clinical records or clinical decision support tools and hence are invisible to the care team. Obviously, these data need to be made available to the care team readily or else the utility of collecting them is very limited in improving overall health. The clear lesson here is the need for high-quality clinical data to support high-quality clinical services for patients.

Technological, Geographical, Social, and Political Barriers Impede Critical Public Health Response
As described in detail in the "Public Health COVID-19 Impact Assessment: Lessons Learned and Compelling Needs" discussion paper (and summarized in *Table 1*), COVID-19 presented the U.S. health care and public health communities with pressing and critical questions about the magnitude and management of the COVID-19 pandemic that were only partially addressed.

The successful use of data included U.S. visibility into nationwide case counts of COVID-19, laboratory testing results, data on special subpopulations, and, more recently, on vaccinations [31]. Monitoring national public health progress became a routine pastime, with most major news outlets summarizing epidemiology data in near real time. Progress was also made on timely data about hospitalization rates, including ICU care, and non-traditional public health data such as human mobility patterns that would help understand and monitor the pandemic. The adoption and adaptation of existing health information technology and data infrastructures were critical to collecting and analyzing these public health data. Newer technologies such as application programming interfaces (APIs) and cloud-native applications facilitated progress.

While technology has enabled progress, there remain clear geographical, socioeconomic, legal, and political barriers to collecting, organizing, integrating, analyzing, and then disseminating local, regional, and national data owned by various groups and subject to state and local jurisdictional policies and regulations. The major lessons learned in this pandemic are magnified versions of prior lessons such as the lack of visibility on nationwide public health data and the lack of truly interoperable health information technology

systems in health care settings. Even when data are available and exchanged between health care providers and the public health sector, there are gaps in the completeness, timeliness, and granularity of the data available for monitoring the pandemic.

For example, while the U.S. was able to create data systems that summarized laboratory testing information to manage the pandemic, the datasets were not complete enough to incorporate SARS-CoV-2 test performance into calculations of disease prevalence. Since SARS-CoV-2 test sensitivity and specificity varies widely and because differing tests were used by different laboratories in different locales, two contiguous counties in the same state could appear to have the same disease prevalence while, in reality, disease prevalence may have been different due to testing patterns and testing performance. Similarly, information about affected patients was frequently missing (e.g., race/ethnicity, comorbidities, likelihood of exposure) as was information on rates of asymptomatic COVID-19 positive patients. These gaps were critical because pandemic management approaches (e.g., recommendations for shelter-in-place) were based on observed disease prevalence, change over time, and a subpopulation's risk of exposure to the disease. The key observation is that better data integration across a variety of data types and data sources is critical for public health decision-making and timely and equitable action.

Innovative technology solutions can also advance public health and population-based management of the pandemic. One byproduct of the pandemic is that data visualization has become a widely accepted public health practice. Graphical data analysis techniques allow anyone to easily understand and respond to the pandemic, from sophisticated epidemiologists to individual Americans trying to figure out how to plan their day [32]. Software has also advanced traditional public health activities like monitoring vaccine performance over time. As a part of COVID-19 vaccine monitoring, the CDC has advanced a program called "v-safe" [33]. Vaccinated individuals use a Quick Response (QR) code or other registration strategy to sign up for a national monitoring database intended to document post-vaccination symptoms and identify potential safety concerns. The technology platform is secure and private, and provides a mechanism to drill into new potential safety signals of concern and to follow up. An enhancement to the Vaccine Adverse Event Reporting System (VAERS) could be the ability to link incidence reports with public health immunization records to verify that the person reporting the side effect actually received a vaccine.

Digital platforms have also been used for case investigation and contact tracing. These systems allow for notification and monitoring of persons exposed or infected with SARS-CoV-2 [34]. Digital tools provide for symptom monitoring and clinical and public health referral of persons who may need additional support for testing, isolation, and quarantine. While data visualization and person-focused software to facilitate participation in monitoring appear to be basic tools, they are truly innovative when incorporated into public health tasks. The key lesson learned here is that incorporating appropriate technology innovations into public health is a critical task for the nation's safety, security, and overall health.

Modernized, integrated, real-time public health data systems at every level of government will revolutionize the nation's response to health threats. There is a clear need for a national public health data ecosystem that functions well in the inter-pandemic phase and can then seamlessly adapt and scale for a future pandemic or other public health emergency [35]. Preserving the privacy and confidentiality of individuals while collecting and disseminating public health data remains a foundational principle. Modernization would ideally reduce the burden of health care providers in reporting conditions to local public health officials and of public health reporting to the federal government [36]. To fully realize this, standards and approaches to reporting a minimal set of public health data need to be universally adopted and enforced.

Realizing this vision requires sustained investment and guidance to state, local, tribal, and territorial health departments, the creation of advanced tools and capabilities at all levels, and the realization of best-in-class innovation with research, the private sector, and public health partners. Investments to date have laid the groundwork and spurred real progress, but much work remains to be done. In addition, there is a critical need to build and support a public health workforce that is skilled in informatics and data science to establish and maintain the ecosystem. This can be accomplished by reskilling, upskilling, recruiting, and retaining a data science workforce with the skills required to meet 21st century health threats. Finally, developing equitable governance while preserving privacy will require consensus building and cross-sector partnerships.

Operations Infrastructure, such as Supply Chains, are Critical and Data-Dependent

Health care's digital infrastructure was critical to all aspects of managing the health of the population during the pandemic, including optimizing a wide range of day-to-day operations, such as ensuring food availability in grocery stores or determining whether university students could attend classes in person. Like other aspects of managing the pandemic, a critical feature of the pandemic response was the need for readily analyzable data to inform and refine the operations of hospitals, universities, businesses, and other organizations. The early days of the pandemic offered dramatic examples of challenges, including deficiencies in supply chains in a range of industries, including hospitals. Physician executives were compelled to step out of their normal roles to help optimize supply chains for PPE and ventilators. The initial scramble for supplies was characterized by confusion but also remarkable improvisation such as in the activation of presidential emergency use authorizations and whiskey manufacturing plants repurposing their operations to make hand sanitizers with 80% alcohol [37]. Because many health care systems did not stockpile inventory, they were left vulnerable due to shortages when the need surged. For example, early in the pandemic, two-thirds of health care workers in the U.S. did not have enough masks, and about 70% of workers had to wear the same mask for more than one day, putting them at even greater risk of COVID-19 infection [38].

From lessons learned in managing supply chains during the COVID-19 pandemic, the U.S. health care system has an opportunity to optimize approaches to sourcing, inventory management, analytics and technology to better understand vulnerabilities and to address them. A flexible, resilient and pandemic-ready supply chain would include the following features:

- Sharing of accurate, timely, and real-time data between providers and suppliers to improve transparency in inventory tracking across individual health systems and allow for the equitable, trustworthy distribution of hospital supplies.
- Expanding investments in safety stockpiles that would reduce reliance on just-in-time orders and provide a sense of probabilities on supply availability.
- Using predictive modeling AI that incorporates information on individual part manufacturing and sourcing from current and potential suppliers.

- Improving supply-chain analytics by integrating data with user workflows for efficient data mining by product, geography, and timeline.
- Adopting internet-of-things (IoT) connectivity and digitization that will allow hospitals to better track products throughout the supply chain and assess vulnerability (e.g. single-sourced supplies, financially fragile suppliers) from shipping all the way to the point of care.
- Connectedness to and visibility by government actors responsible for making allowances and shifts in response to critical demands (e.g., the U.S. Food and Drug Administration's ability to provide emergency authorization for new manufacturers).

Supply-chain optimization can affect the quality of care through multiple factors, including by saving time for key personnel, allowing physicians to spend more time with the right patients, reducing the time spent looking for supplies, and allowing for better recall management to reduce patient safety risks. This case example has broad applicability across a range of daily operations, from better management of regional hospitals and ICU beds to ensuring a safe national food supply.

COVID-19 Spurred Progress and Exposed Key Gaps in Access to Digital Therapeutics

The pandemic highlighted the immense opportunities to leverage digital capabilities in service of improved health. Simple-to-use but sophisticated software applications that can run on personal computer devices, such as Somryst, an FDA-cleared prescription application for the management of chronic insomnia in adults, offered digital solutions directly to many patients who needed them [39,40,41]. Shelter-in-place and quarantine rules, which limited a patient's interface with the health care system, amplified the demand for care and increased the likelihood that patients and clinicians would gain experience with digital tools. As mentioned previously, relaxed regulations with regard to telehealth played a critical role in expanding access. Remote digital sensors such as heart rate monitors and pulse oximetry were well-utilized during the pandemic, offering the opportunity for more equitable home-based care monitoring. AI-powered conversational chatbots were deployed by the CDC and thousands of hospitals and clinics to enable patients to self-assess their potential COVID-19 symptoms. The use of digital therapeutics—software applications intended to deliver therapeutic relief—also expanded, especially for mental health in-

terventions, where the FDA provided emergency use authorization of relevant digital therapeutics without review.

COVID-19 also highlighted the large unmet gaps in digital solutions for mental health. Pre-COVID-19, about 51 million U.S. adults (20% of the population) lived with a mental illness, and almost two-thirds of lost workdays in the U.S. were caused by mental illness [133]. While most mental illnesses are treatable, nearly half of all people with mental illnesses do not receive any services; suicide was the tenth leading cause of death in individuals before COVID-19 [134]. The reasons for this care gap are multi-fold, ranging from stigma, lack of access, shortage of therapists, inadequate funding, and lack of parity between care for physical and mental health conditions. While the full impact of pandemic mitigation strategies on mental health may not be known for some time, early studies indicate an uptick in mental health disorders due to the pandemic [42]. For example, surveys conducted by Kaiser Family Foundation in January 2021 indicate that four in ten U.S. adults during the pandemic experienced symptoms of anxiety or depressive disorder—an increase from pre-pandemic levels of one in ten adults [43]. Furthermore, accumulating evidence suggests the problems have been amplified for youth, marginalized communities, and people of color [44,45,46].

The collision between the profound mental health needs revealed and intensified by the pandemic and the profusion of digital health tools not only highlights the promise of digital health tools to help address mental health needs, but also reveals important limitations in these tools. Digital tools for mental health can be divided into two broad categories: (1) lower-risk triaging and health care delivery digital tools such as telemedicine (via apps) and crisis counseling (via text messaging), and (2) digital diagnostics and therapeutics intended to diagnose or treat mental illness.

Triaging and health care delivery tools were able to be utilized almost immediately after the onset of CO-VID-19, and offered a vital point of immediate medical and personal connection. For instance, the Crisis Text Line (CTL), which provides free, 24/7 counseling to people experiencing a mental health crisis via text messaging, reports that more than half of its users (65%) had not spoken to anyone else before contacting CTL [47]. At the same time, the scaling of this category of tools has been hampered by issues related to privacy, cross-state licensing, bandwidth, and limited reimbursement.

Digital therapeutics also offered the possibility of immediate assistance to many suffering from mental health challenges during the pandemic, and there were quite a few from which to choose. Pre-pandemic, there were more than 10,000 apps claiming to help with stress, depression, anxiety, and insomnia. However, nearly all these apps operated as "wellness apps," in that they were not subject to FDA oversight if they did not make overt medical claims. Remarkably, over the past decade, only three mental health digital therapeutics (for the treatment of insomnia, substance abuse, and attention deficit hyperactivity disorder) have sought and executed the studies to gain formal FDA clearance [48,49,50]. Therefore, it is difficult to evaluate scientifically the efficacy of the vast majority of apps that claim to help with mental illness. For example, a study of 73 apps addressing a range of mental well-being issues found that many indirectly claimed effectiveness through scientific phrasing, but only two provided direct evidence from a trial [51].

Wellness and FDA-cleared apps are generally not integrated with clinician EHR systems, which presents another challenge in organizing and centralizing patient data. Cognizant of the mental health impacts of the pandemic, the FDA waived the requirement that mental health-focused apps and digital therapeutics (such as symptom checking and triaging apps, and low risk therapy or counselling apps for anxiety, depression, or sleep) must submit a 510(k) premarket notification before distribution to the public [52]. A large number of mental health symptom checker and triaging apps took advantage of this opportunity, broadening the market of available apps. A mental health digital therapeutic (to treat ADHD symptoms), which was in the process of a de novo submission, was also able to come to market much sooner under this provision [50]. Several other companies with digital therapeutics in the pipeline are also actively looking to take advantage of this temporary relaxation [52]. Last but not least, pandemic-related shifts to contactless clinical trials have spurred the use of digital tools (e.g., wearables) to monitor psychiatric symptoms as well as the use of digital therapeutics to augment pharmacotherapy.

Given the widespread adoption of consumer-oriented digital tools during the rise of COVID-19, a key lesson learned is the importance of evaluating the performance of digital tools to better understand where they contributed the most and what factors correlated with success. Equally important is the critical evaluation of which patients were not well served by digital

approaches, and what might be done to remedy these deficiencies. The availability of large real-world data-sets would better harmonize effectiveness standards between the FDA and payers (e.g., the U.S. Centers for Medicare & Medicaid Services (CMS)). It will also be valuable to consider how the data gathered by these digital approaches might be best leveraged to advance public health, while protecting patient privacy and data rights. Such insights will hopefully also inform continued post-pandemic authorization of these devices as well as the establishment of new regulatory pathways.

Access to reliable digital tools and effective and well-integrated apps offer the potential to radically change how patients cope with mental health challenges—not only during pandemics but also in their daily lives in inter-pandemic times. The successes seen during the pandemic highlight these possibilities and raise the question of whether some pandemic-related exceptions should be made routine. A thorough "after-action" report is also critical to ensure that patients with mental health challenges and their providers can make informed, discerning care choices; to enable regulators to refine and optimize review and approval protocols; and to enable innovators to build on what is proven to work to develop even more effective approaches for the future. Scaling such solutions is important for the mental health space as well as the health care system writ large.

Advancements in Clinical Evidence Generation were Essential but Rudimentary
COVID-19 generated an urgent need for medical science to understand and respond to the novel SARS-CoV-2 virus. Emerging digital technologies were pressed into service across the range of evidence-generation activities, including not only preclinical discovery and traditional clinical trials, but also extending to real-world data obtained from the observation and instrumentation of clinical treatments. The success of these approaches varied, highlighting their exceptional potential. The rapid sequencing of the SARS-CoV-2 virus genome and near-instantaneous global sharing of these data comes to mind, as well as the remaining work of leveraging and disseminating data from EHRs around the manifestation of COVID-19 symptoms and progression of the disease.

Preclinical Discovery
The ability of the international science community to share information about SARS-CoV-2 so quickly represents a prominent example of digitally enabled bio-

medical progress. The combination of next-generation sequencing (NGS) capabilities and powerful open-source data-sharing platforms such as NextStrain enabled scientists to characterize the molecular structure of the virus and rapidly share it with colleagues around the world. This shared understanding provided critical insights into how the virus was spreading and how it was evolving over time, while also enabling researchers to identify potential viral vulnerabilities. In addition to genetic sequencing data, researchers also used open platforms to share information related to the characterization of the immune system response to the virus, the chemical structures of potential antiviral compounds, 3-D structural data for models of SARS-CoV-2 proteins, transcriptional data, and histopathological images from infected tissue.

The conspicuous success of data generation and sharing in the preclinical area, like other examples of digital success, reflects in large measure the work done and progress achieved prior to the pandemic. Stemming from learnings tracing back at least to the human genome project two decades ago, these scientific communities now have deep experience sharing data, both technically and culturally, and have established standards and tacit conventions that facilitate this process [53,54,55]. Critically, these efforts are enabled by features of the datasets themselves—on balance, these data tend to be highly structured, consistent, reliable and complete; generated by instruments; and ready for analysis. Analysis of clinical data, in contrast, must contend with the idiosyncrasies of health care delivery and the management of patient privacy, presenting additional thorny challenges. Ultimately, a key lesson for the future is the importance of aligning and integrating these varied data sources in ways that build on the progress already achieved by the research community.

Bioinformatics
There is no question that genomics is a "big data" science, involving millions, possibly billions, of genomics, proteomics, metabolomics, and other -omics and related phenotypic data datasets, and continuing on an exponential growth trajectory. Bioinformatics brings crucial context to these data through tools such as machine learning algorithms and predictive analytics that can help understand the research and clinical data from a gene-centric approach to a multi-scale systems-level approach. For example, sequencing the SARS-CoV-2 genome and its bioinformatic analysis was the essential first step towards developing a vaccine against COVID-19 and provides a roadmap for track-

ing the emergence and spread of variants of the virus. However, to unleash the opportunities in bioinformatics requires coordinated community efforts.

Clinical Trials

The pandemic created profoundly disruptive threats to clinical trial efforts unrelated to the virus, as well as studies seeking to better understand and manage COVID-19. The successful execution of so many clinical studies under such difficult circumstances owes much to both the availability of emerging digital technologies and the inspirational resilience and adaptability of researchers, regulators and, especially, patients. As with preclinical research, most of the required technologies were already in place, at least provisionally; the needs created by the public health emergency merely served to accelerate implementation and adoption of these approaches in clinical trials.

Modern digital dashboards for clinical trials proved especially helpful in quickly understanding and effectively responding to the intensified needs created by the global health emergency [56,57]. A particularly important application of digital technology was in remote patient evaluation, an effort promoted by guidance issued by the FDA in March 2020 [58]. This effort encouraged trial sponsors to consider virtual assessments such as telemedicine visits as a means of ensuring subject safety. Remote monitoring was also enabled through the use of at-home sensors (such as pulse oximetry devices). Prior to the pandemic, the potential to use such remote devices had been highlighted in previous FDA guidance, but adoption was limited as many sponsors worried that the advantages did not outweigh the potential regulatory risks and potential inequities [135]. Remote approaches also helped sponsors monitor individual study sites when travel was prohibitively difficult.

The ability of the clinical research enterprise to continue during the pandemic reflects the pre-pandemic transition to a digital infrastructure within industry, regulatory bodies, clinical research organizations, academia, and health care organizations. At the same time, the environment of crisis response means that a full grasp of the key lessons learned is likely not yet apparent. Thus, a thorough post-pandemic review will be essential to evaluate the novel digital approaches used in clinical studies and to examine in a systematic and disciplined manner the impact on patient safety and the integrity of study data.

Real World Data

Data collected from the routine care of patients—termed "real world data" (RWD)—were of critical importance for advancing understanding of COVID-19, including not only diagnosis and treatment, but also the design and conduct of clinical trials. The digitization of health care over the last decade offered the promise of RWD that could be made available in near-real time from both the EHR and from insurance claims data (or "administrative data.") Other potential sources of RWD included biosensor information (e.g., accelerometer data in a watch), biological information (e.g., SARS-CoV-2 genomic information), socioeconomic data (e.g., personal or neighborhood resources), social media summaries (e.g., discussions on Reddit or Twitter), and immune profiles in response to infection. One critical realization during the pandemic was the need for reliable RWD sources to describe an evolving clinical scenario in near-real-time—writing the novel as the story unfolded.

Over the past decade, the medical research and business communities have had high expectations for RWD. Some of these are realized today and others are more future-looking. For COVID-19, RWD were used to refine clinical research studies, including choice of population, endpoints, and sample size, as well as assumptions involving the anticipated mortality by age, comorbidity, and disease severity—all factors that could otherwise be especially challenging to estimate in the earliest phases of a new disease [59]. RWD were also used to complete longitudinal study datasets, as demonstrated in the RECOVERY trial and as contemplated for studies in the iSpy platform trial network [60,61,62,63].

A critical question for any dataset used in clinical research is whether it is sufficiently reliable to meet the evidentiary task, and RWD is no exception. In recent years, intensive effort has been devoted to developing standardized approaches for documenting dataset characteristics and how these can be matched to a clinical research task [64,65]. The pandemic showed just how important this foundational work is in assessing the completeness, variable reliability, and provenance of data for the crisis response [66].

The pandemic also highlighted important challenges in the use of RWD. An initial hurdle in the early days of the pandemic was learning how to work with RWD datasets, understanding their reliability, and developing common definitions for key parameters describing severity of disease, such as whether a patient was re-

ceiving supplemental oxygen. It was also important to contextualize data within the changing contours of the pandemic over time. For instance, the mortality rate of hospitalized patients in New York City decreased dramatically between spring and late summer 2020. The mismatch between the urgent need for RWD and its limited availability also led to some acknowledged setbacks. Two papers published in leading medical journals were subsequently withdrawn after the improbably robust RWD dataset upon which these articles were based was called into question [67]. Greater familiarity with RWD might have identified this conspicuous red flag prior to publication.

The COVID-19 Evidence Accelerator was established by the Reagan-Udall Foundation in collaboration with Friends of Cancer Research to catalyze the effective sharing of RWD methods and insights using a public-private partnership model [68]. The Evidence Accelerator was initially set up to help address a pre-specified set of questions around the natural history of the pandemic. The goal was to bring together data holders, analytic teams, technology innovators, government bodies, and others to solve problems, and the forum provided a legal space for cross-organization information sharing and problem solving. Perhaps not surprisingly, consortia that already had developed and implemented standardized models were able to adapt to the pandemic with particular speed; examples include Observational Health Data Sciences and Informatics (OHDSI), the U.S. Department of Veterans Affairs (VA), and University of California Health [69,70,71]. Many datasets required curation but were otherwise quickly adapted to help answer questions related to the pandemic. Additional work was required by data aggregators, who had collected large volumes of data that required further aggregation and analysis in order to make sense of the information.

A key determinative factor in driving value from RWD appeared to be data empathy—a deep familiarity with the clinical context of the data and experience in using it in a health care context [72]. One particularly important lesson has been the renewed awareness of the gap between raw clinical data and derivable insight, along with the recognition that achieving such actionable understanding typically requires more than a technology fix—it needs insight into the nuances of both the clinical data available and the research question to be addressed. Extracting value from RWD, the health care community has learned, requires meaningful collaboration between clinical experts (who understand the often very local clinical context), statisticians (who recognize the evidentiary requirements for medical research), and data scientists fluent in large data sets and the techniques, including data curation and AI, for organizing and understanding them [73].

Digital Future of Evidence Generation

Digital technologies played a critical role in accelerating global understanding of the virus and the urgent development and critical evaluation of a range of potential countermeasures. The successes, from the global sharing of viral sequence data to the ability to conduct robust clinical trials during unprecedented circumstances, highlight the transformative potential of digital technologies, as well as the value in establishing both tools and culture in service of these technologies in advance of a Public Health Emergency. Critical lessons learned include the importance of:

- Incentivizing data sharing across the landscape from basic biological discoveries (e.g., SARS-CoV-2 viral sequence) to longitudinal clinical data (e.g., use of EHR data to complete long-term follow up of a person on a research study).
- Incentivizing data interoperability (e.g., the ability to merge variables across datasets from various sources) as well as documentation and improvements in data quality.
- Considering all evidence generation tasks as important, from basic description of the pandemic to determining treatment effectiveness.
- Leveraging digital technologies to more efficiently conduct evidence generation tasks more efficiently (e.g., telemedicine for remote monitoring or RWD for longitudinal follow up of patients enrolled in clinical trials).
- Matching the evidence generation task to the dataset and analytic approach.
- Conducting a post-pandemic "after action report" to carefully evaluate the benefits and risks of clinical research innovations deployed during the pandemic (e.g., remote patient evaluation using telehealth solutions) and to develop approaches to allow meaningful innovations to persist.
- Ensuring that all required types of expertise participate in evidence generation tasks, including clinicians.
- Leveraging public-private partnerships to advance solutions to quickly develop the science and explore innovations related to digital solutions for evidence generation tasks.

Harnessing AI and Other New Capabilities Depend on a Coherent Data Infrastructure

The advances of the past decade in AI—particularly ML and data science—have captured the attention of the field. Today, AI is infused into nearly all aspects of life, including the smart diagnostics that predict when automobiles and home appliances are about to break down, the analytics that facilitate connections to relevant social circles and retailers, intelligent decision supports that help power global supply chains and ensure that foods and medicines get to where they are needed, and much more [74,75,76]. Over the next five years, scientists expect AI systems to provide practical capabilities that may transform our understanding and abilities in human language, biology, climate modeling, social systems, and more, unleashing new waves of scientific and technological advancement [77,78,79].

While it can be hard at times to separate hype from reality, our nation is indeed living through a fundamental transformation, fueled by the ever-increasing ability of AI to absorb massive quantities of data—generated through a combination of human thought and activity, increasingly ubiquitous sensors, and simulations of progressively higher fidelity—and then to distill that data into knowledge that has practical significance for societies, communities, organizations, and individuals. This is a transformation that the authors of this paper want and believe society will demand in medicine and health care delivery.

From the perspective of data and AI, the COVID-19 pandemic has presented many opportunities to put this transformative vision into action. Examples of positive impact have together shown a bright future for AI. These examples include the widespread impact of intelligent chatbots, progress toward the accelerated discovery and development of therapeutics, and life-saving decision support systems based on intelligent forecasting [80,81,82].

At the same time, the COVID-19 experience brought into clearer view how challenging it can be to access the benefits what are still relatively new capabilities in data and AI. A consistently underappreciated aspect of digital systems is that the act of acquiring, aggregating, and normalizing data, while potentially an arduous task in and of itself, is only the first step in creating and deploying an intelligent system that can operate at enterprise-grade quality and scale. Other steps include establishing:

1. ML training infrastructure to process data,
2. Application infrastructure for deployment and user engagement,
3. Feedback infrastructure to monitor for faults and enable continuous improvement,
4. Ethics and compliance infrastructure to ensure fairness and accountability, and
5. DevOps infrastructure to manage and evolve the overall system.

Further, there is an overarching challenge of scale. AI requires massive amounts of data, but any single institution might have relatively limited data for any specific problem. Hence, the management of data sharing agreements, usage rights, and chains of custody are themselves significant infrastructure needs. In other words, there is a significant infrastructure foundation necessary for any operational data and AI system. Data repositories such as the NCATS National COVID Cohort Collaborative Data Enclave established to address the COVID-19 pandemic could offer a rich trove of data and effective model for cooperatively exchanging data [128].

As AI has emerged from laboratories into operational deployments, various forms of AI lifecycle concepts codify the continuously evolving nature of these systems (see *Figure 1*).

AI and data are the tools and raw materials needed to power analysis, but the full operational lifecycle of any intelligent system depends on a comprehensive and coherent infrastructure. A similar lesson of lifecycle management applies to other emerging capabilities in information technology. For example, cryptographically secure distributed ledgers (e.g., blockchain technology) offers the capability to protect transactions while relieving participating parties from the need to work out a myriad of multilateral contractual arrangements. This has direct implications for decentralized patient identity systems (to enable, for example, secure vaccine credentials), supply chain management, and reimbursements for patient-owned data. As with AI, such capabilities are dependent on systems that are built on architectures that support the end-to-end requirements of all key stakeholders.

As the health care community thinks about the future, infrastructure preparedness requires an early understanding of not only the overall data and AI architecture, but also how bias is introduced and perpetuated by an AI system. Fortunately, a great deal of tooling is now available, with more emerging all the time, to help standardize aspects of this lifecycle's components. These include modern data standards, open API frameworks, DevOps platforms for AI development teams, and cloud-based computing infrastructure that
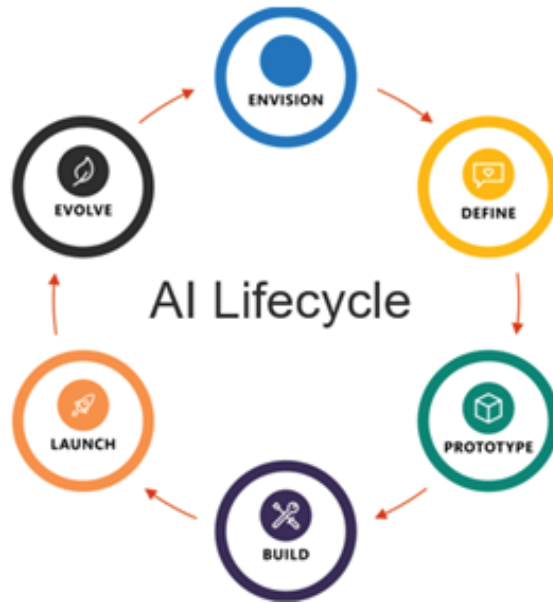
FIGURE 1 | Lifecycle for Continuous Management and Refinement of AI Models
SOURCE: Microsoft, Inc.

is robust enough and specifically tuned to power all the above [83,84,85]. At the same time, there is a need for guidelines, standards, mechanisms, and governing structures to ensure that equity is ingrained throughout the AI lifecycle.

Priorities for the Future of Digital Health

The learned lessons from the COVID-19 experience help envision a better future and help prepare for a set of practical next steps. The COVID-19 experience has showed, more vividly than ever, the overwhelming need for readily analyzable and aggregated health care data, supported by systems to make sense of it, implement findings, and improve the work over time. Both the U.S. health care system's successes and failures in pandemic response have provided greater clarity on what our nation will need to focus on for future public health crises.

Insights gleaned from hard experience are pertinent not only to pandemic response. They are also relevant to the future of digital health broadly, and in particular they reinforce key elements of the concept of an LHS. The lessons learned during COVID-19 provide insights that update the LHS concept—insights relevant to person-centered care, business incentives, and cybersecurity.

Unleashing the Potential of a Learning Health System

Since the 2000s, the Institute of Medicine (now the National Academy of Medicine, or NAM) has advanced the vision of an LHS, "in which science, informatics, incentives, and culture are aligned for continuous improvement, innovation, and equity with best practices and discovery seamlessly embedded in the delivery process and new knowledge captured as an integral by-product of the delivery experience" [86]. While the pandemic exposed health care's fault lines, its enormity and the resulting pace of scientific innovation brought the potential and need for an LHS into sharper focus.

In short order, the health care system had to define the illness, identify those who needed treatment, design and advance medical interventions, manage a complex health care system, and refine its approach based on evolving data. Different types of data were put to work for specific goals (e.g., EHR data to define the likelihood of needing mechanical ventilation and to manage ICU beds; viral genomic sequences to inform vaccine development). These parallel streams of data were also combined, informing more refined actions including better optimization of health care delivery and personalization of care. Other digital tools, including data management capabilities, analytic and visualization solutions, telemedicine, and clinical decision support systems, enabled implementation of the pandemic response strategy. Even newer capabilities such as social media tools allowed efficient transmission of

clinical observations among clinicians (e.g., observations about COVID-19 venous and arterial thrombotic complications via Twitter).

A key observation during the pandemic was the need for and value of accessible, interoperable, and readily analyzable data to support the response to COVID-19. This need is central to the functioning of the health system beyond the pandemic. Early features of an LHS became evident with the nascent data streams available in 2020, at once highlighting the power of the LHS vision and urgently motivating its concrete realization for the benefit of health care delivery, therapeutic innovation, and public health. Improved health care data sources would allow optimization of health care delivery through continuous data analysis, review, adaptation, and reevaluation. Individual health care systems or hospitals could quickly learn from each other. As a practical example, early in the pandemic, clinicians questioned the appropriate timing of oxygen delivery and mechanical ventilation, and whether ventilated patients recovered more quickly when in a prone position. Given the thousands of patients being diagnosed and treated across the U.S. and worldwide, more systematic learning could have been achieved if there were relevant, readily combined datasets.

The development and precise delivery of novel therapeutic approaches, both pharmacological and non-pharmacological, would also most likely be accelerated through the effective combination of biological and clinical data. Already, scientists are seeking to identify the best ways to manage patient illness based on the knowledge of the SARS-CoV-2 genome and its various clinical manifestations, building on the success of integrative initiatives such as the All of Us program in the U.S. and the United Kingdom (UK) Biobank [87,88].

There is an especially striking need to extend LHS principles to public health, built on a foundation of accessible data with continuous learning enabled through real-time analysis, implementation, and refinement, which would directly inform the management of the pandemic. The UK's identification in late 2020 of a new SARS-CoV-2 variant with markedly increased transmissibility, B.1.1.7, highlights the value of such comprehensive data integration: the UK's critical insight was explicitly enabled by the routine genetic sequencing of virus samples, combined with public health data on transmission rates [89].

The U.S. has implemented a number of critical policy steps to build on LHS capabilities in line with the NAM vision outlined in 2006, such as the 2009 HITECH Act and the 2016 21st Century Cures Act [90,91,92]. Deliv-ering on the vision will also require thoughtful leadership, increased focus on the nation's fundamental data infrastructure, and a more inclusive perspective that regards technologists not as technicians, tool makers, or service providers, but rather as essential partners who must be at the table to successfully design and implement an LHS.

Business Solutions are Needed

In examining the failures of health information technology in the COVID-19 crisis, one observation is that the gaps in the system are significant, despite a decade of public investment in technology to support health care delivery. Many evaluations, including this paper, highlight the lack of an overarching data architecture to support the use cases for technology that developed during the pandemic. But in most settings, data architectures are designed to support business use cases that are core to the financial or business success of the organization. Thus, the failure of the data architecture is a symptom of a more general lack of an integrating business rationale for data fluidity in health care, a gap that profoundly influences stakeholder dynamics [129].

Despite a decade or more of discussion about interoperability, there has been limited discussion of what business solutions would be required to drive the market forward to true data availability. These business solution discussions have two diverging challenges. The first is that many of the core aims for patients are not supported by, or indeed conflict with, the fee-for-service business model of today's health care delivery. Practically, for many health care systems, the lack of technology interoperability is a key business proposition that helps to maintain patients in the system. The second challenge is that at the business level, most health care systems are struggling with limited EHR design flexibility, as well as maintenance and operational issues, that reduce interest and capacity to consider the further complexity required to accommodate a common data architecture.

The data availability and quality challenges encountered during the pandemic were predominantly encountered in the health care delivery and public health spaces. These challenges stand in contrast to clinical evidence generation for new COVID-19 treatments or vaccines or biological discovery sciences, where there was much greater data availability. When the business imperative to generate high quality data was aligned with the commercial success of a medical product, investments were made in standardized data collection

(although in non-interoperable data silos). Meanwhile, the expectation that high-quality and readily analyzable data would be a "free" by-product of health care delivery is misaligned with reality—both because the business incentives in health care delivery do not reinforce data interoperability and, even when data are interoperable, the costs of data curation and cleaning are not factored into the system.

In other areas of the economy faced with similar data silo challenges, like banking, the rationale for change became a public-private partnership with the Federal Reserve to reduce transaction costs by standardizing data elements [102]. The opportunity to reduce costs became the business rationale to drive standardization across firms and environments. In this case, the Federal Reserve served as a catalyst for this change, helping to address the activation energy needed to drive to the new transaction model. Such an overarching business catalyst and business solution are lacking in the health care system [136].

Business solutions could address real economic challenges faced by health care systems. For example, there are hundreds of different health care systems deploying hundreds of technologists, each performing required maintenance on their EHR systems. Could a different data architecture reduce these hidden costs, which currently amount to billions of dollars annually? [103] Similarly, since 1996, there have been efforts to reduce transaction costs at a federal level, yet overall, billing and other transaction costs remain high (relative to other health care systems globally), contributing to the high cost of health care in the U.S. [104,105,130].

Solving the data architecture challenges identified in this report requires the development of business models focused on novel use cases to spur the investment and effort that will be required; these business cases should help focus national leadership on an actionable path forward for the public and private sector. The need for active coordination becomes apparent as a means to ensure that use cases are supported by core business models to the data architecture and that they are added in a coordinated fashion that does not create further non-interoperable data silos.

Cybersecurity Must Be High on the Agenda
As 2020 came to a close, the U.S. was hit by a large-scale cyberattack, likely conducted by a nation-state [106]. First discovered as a "supply chain attack" perpetrated through SolarWinds software patches, it has become clear that this event involved much more than just SolarWinds software and affected U.S. private in-

dustry and government systems, including HHS—a critical aspect of our nation's response to the pandemic [107].

At a time when digital health solutions are becoming an integral part of health and its management, this cyberattack demonstrates our vulnerabilities. While the SolarWinds attack was most likely perpetrated for espionage and nation-state activities, more mundane criminal and predatory behaviors have also been witnessed during the pandemic. For example, ransomware has been found in the cold storage units needed to maintain COVID-19 vaccines at appropriate temperatures, COVID-19 vaccines are being sold on the dark web, and the European Medicines Agency was hacked and commercial vaccine-related regulatory documents were stolen [108,109,110]. In fact, sustained cyber assault has been witnessed across the entire vaccine clinical development and supply chain, both by criminal and potentially nation-state actors [111,112].

The cybersecurity risks in digital health were well-known prior to the pandemic. Medical devices such as pacemakers are vulnerable to cyberattack [113]. Ransomware attacks on hospitals are all too commonplace; in October 2020, the FBI warned of increasing attempts [114]. Medical device manufacturers have a responsibility to design their products to limit cybersecurity risk and monitor them accordingly. Health care delivery organizations should also be attentive to network security and the responsibility of individuals using the systems to ward against phishing and other schemes. Regulations such as HHS's cybersecurity guidances and safety communications are intended to thwart cybersecurity breaches [115,116]. Yet, as the U.S. moves toward an IoT model, the threat amplifies.

One concern that warrants attention is that of adversarial data manipulation; for example, data arising from funduscopic examinations, chest X-rays, and of dermatological exams [117]. The national response to COVID-19 depends on the ability to describe the evolving public health emergency, address it, and monitor actions—all of which depend on data. Medical misinformation may be malicious and is easily amplified via social media. Bad actors can erode or scramble a public health response by injecting incorrect data into systems ("data poisoning"). For example, misleading information about changes in the pandemic can lead to public health recommendations that worsen the pandemic (e.g., when the data erroneously suggest that the pandemic is under control) or cause public mistrust (e.g., when the data erroneously suggest that public health interventions are not working as intend-

ed). Malicious data manipulation could impair clinical trial results about vaccine efficacy (e.g., leading to delayed uptake of effective vaccines) or erroneously suggest safety challenges with vaccines, thereby slowing uptake and/or sending researchers on unnecessary hunts to understand safety signals. Adversarial ML algorithms are a practical byproduct of adversarial data manipulation. These sorts of adversarial threats are likely to grow as the use of data and the dependence upon algorithms intensifies.

Cybersecurity must be much more than an afterthought or an add-on; it is integral to the public health response to a crisis like COVID-19 and to the development of digital health capabilities. Cybersecurity leadership (e.g., the role of a Chief Information Security Officer) should be incorporated into digital health planning and implementation, and cybersecurity training for all health care employees should be intensified and frequently refreshed. Cyber vulnerabilities must be identified and proactively addressed.

The FDA's partnership with the "white hat" hacking community, initiated in 2018, and MITRE's *Medical Device Cybersecurity Regional Incident Preparedness and Response Playbook*, represent promising approaches [118,119]. Finally, the discovery of the SolarWinds supply chain attack may necessitate the rebuilding of government and other IT systems, which could represent an important opportunity to advance the nation's information architecture [120]. The success of cybersecurity countermeasures will require leadership, vision and collaboration, including public-private partnerships.

Digital Health Training is More Important Than Ever
Ongoing training is critical to ensure that the individuals accessing digital systems and associated data are qualified to make optimal use of them. Future generations must be ready to adapt to the evolving challenges and opportunities in health care and life sciences. They must also adopt appropriate digital health tools to support communities with new discoveries and knowledge and new business processes and business models, with the ultimate goal of improving clinical and economic outcomes. To reach a broad audience, such training in digital technologies can be delivered through public-private partnerships, including—but not limited to—degree-granting programs, micro-certifications, and massive open online courses. These programs can be structured around emerging roles—from efforts geared to understanding higher-level applications of digital technology in the clinical enterprise or health care system to more focused development

of technical skills and methods. Learning objectives for these programs should be specific to their overall goals and audience, and can include consideration of broad knowledge of new and emerging digital technologies applied to health care and life sciences. Objectives can also consider methods and approaches on the use of informatics in the discovery and management of new knowledge relating to health, drug development, and understanding diseases. Ideally, all programs should be motivated by efforts to improve human health and enable health care professionals to acquire cross sectional training and experience in ethics, business, and policy to apply biomedical data, information, and knowledge effectively for scientific inquiry, problem solving, and decision-making.

Ensuring All Individuals Get the Care That They Need
Perhaps the most important and enduring observation reinforced by COVID-19 is that fulfilling the promise of emerging technology does not depend on the intrinsic capabilities of technology itself, but rather on people—the people who develop technology, of course, but also those who implement it and, most importantly, those who are served by it.

For many individuals, technology was a form of self-empowerment in supporting personal health management and prevention, such as using COVID-19 trackers to inform shelter-in-place decisions or meditation apps to manage anxiety. Digital technologies influenced all aspects of health management for the U.S. population as a whole. Nonetheless, some fundamental aspects of person-centered care bear remembering:

- As innovations are adopted, the most vulnerable people are often excluded from early access or application. For example, access to telemedicine requires both access to a computer or mobile device and to broadband, which means that access to telemedicine disproportionately benefited more affluent people, especially among individuals younger than 65 years and without comorbidities [93]. Policymakers need to bear in mind the consequences of structural racism and digital redlining as they craft solutions for the future.
- Bias in underlying data can lead to erroneous conclusions and widen health disparities. This was evident during the pandemic when early digital descriptions of the pandemic missed patterns showing that Black Americans were at highest risk of contracting and dying of CO-

VID-19, thereby squandering the chance for early targeted interventions for this population [94,95].

- Bias and conflict of interest in AI models have the potential to amplify discrimination. The risk that bias in the underlying data will bias the results of the AI model is well documented. In COVID-19, unrecognized low blood oxygen levels were three times more frequent in Black patients than white patients; however, the risk models predicting risk of severe disease may underestimate the impact for Black patients. Attempts to adjust or equalize these models can in and of themselves lead to furthering discriminatory decision-making [94,95,96].

- Vaccine prioritization has proven to be more challenging than expected during the COVID-19 pandemic, with the initial rollout of the vaccine in early 2021 based on a first-come, first-serve model within the planned phases/tiers [97]. To make vaccine distribution work for all Americans, a vaccine prioritization scheme should be based on the CDC's identified medical risks, but should also take into account demographic risks (race, ethnicity, location). The framework put forth by the National Academies of Sciences, Engineering, and Medicine offers a model [98]. A robust priority scoring system would incorporate data provided by patients when they register to receive the vaccine, matched data from their EHRs, as well as guidelines from the CDC to ensure an equitable distribution that builds and strengthens herd immunity against COVID-19.

- User-centered design of software applications and interfaces is critical. If the people the software is serving cannot figure out how to use an application, then all good intentions are wasted. For example, many of the first recipients of the vaccine were 65 years and older and experienced difficulty attempting to register for a vaccine. If technical challenges persisted, or if many in this demographic group had not found another way to sign up for a vaccine, then more members of this vulnerable group would have gone unvaccinated [131,132].

- The level of trust that people and communities place in new technologies varies widely. During COVID-19, in particular, health care decision-makers were often guilty of overlooking the role of grassroots community leaders in the important elements of awareness, activation, and engagement of people to work together in pandemic response. One step toward building trust is by having technologists work with community leaders.

- Finally, the health care workforce deserves special mention. Burnout among clinicians, and especially nurses, has been pervasive in the pandemic [99]. Digital health tools, especially EHRs, can exacerbate burnout. Telemedicine removes the human contact with patients that may in itself be therapeutic to exhausted clinicians [100]. Even so, digital tools have the potential to help reduce burnout—such as clinical triage algorithms, remote digital monitoring devices, and voice-to-text capabilities to support clinical documentation—although many of these approaches are in their very early stages [101].

The future development of digital health solutions should be planned in ways to keep this core "customer" at the forefront. Further, technology solutions should not be developed or assessed in isolation—they must co-evolve with the people that the solution is intended to impact.

The Path Forward: Stewarding a Seamless Digital Health Infrastructure

American consumers today live in a world of digital services that offer choice, value, and convenience, and yet health care stubbornly remains one of the few remaining sectors of the economy not designed for the consumer. This has devastating consequences for both individuals and populations, as illustrated by the documented health care disparities of the pandemic. The COVID-19 pandemic highlighted the latent ability for digital innovation throughout the health care system, with successes such as the rapid adoption of telemedicine. However, as this paper has shown, these isolated successes occurred in the context of a health data architecture that was absent or, at best, largely dysfunctional. The result is that relevant information that was needed to understand and respond to the pandemic failed to be delivered effectively and consistently.

While advanced digital technology exists in pockets throughout the health care system, the component systems are largely (and unnecessarily) disconnected and lack the incentives for connectivity and relevant interfaces. Referring back to the home-building analogy discussed in the introduction, it is as though there are no framing standards, thus requiring expensive

customization to get doors and windows to fit; there is no common set of modular architectural components, thus requiring complex plans (and even legal agreements) to be negotiated between plumbers, electricians, and roofers, even though their work hardly relates to each other; and lastly, there is no infrastructure to support independent inspection, rendering the safety and performance of the house unknowable. This lack of a coherent data architecture, modularity, and infrastructure has proved a persistently unsurmounted barrier to the United States' transition to a functional digital health care system. This has constrained progress toward the fulfillment of what may be called health care's Quadruple Aim: better health, higher quality, lower cost, and more engaged people. Innovation is needed to achieve these four goals.

In scale, the challenge described here is equivalent to some of the largest public challenges ever addressed, yet it is one that, in technical terms, ought to be within grasp through a new approach to the U.S. national digital infrastructure for the health care system—one not dissimilar to the need for an interstate highway system, a common financial system, or our modern flight control system. All of these efforts were well-served by public-private partnerships, with benefits that fostered ingenuity and innovation and greatly expanded economic activity for the country at large. The federal government plays a foundational role in enabling such solutions.

Solutions of this magnitude depend on basic paradigm changes—from a model based on organic and spontaneous evolution that occurs naturally with changes in technology to one that recognizes that unencumbered organizational inertia and incentives are basic deterrents to nurturing the common good. For this change to occur, advantage must be taken of certain aspects of the forces in play. For example, much of the raw material for continuous learning is already at hand. The rapid digitization of health care (98% of clinical health records are digital, compared to less than 15% in 2005) and the rapid advances in enabling digital technologies outside of health care have created many of the critical building blocks for a functional health data ecosystem. This represents not only a massive change in the technological foundation of health care, but is also conspicuous proof that the health care system itself can undergo comprehensive evolution. There is a need to align around basic governing rules and an approach to modularity so that consistent components can "plug in," work efficiently, and bring

unique elements to the overall design. A modular approach fosters competition around components, enabling improved quality, reduced costs, and the ability to connect and optimize relevant modules to address distinct human as well as biological challenges in different domains, including public health.

That alignment process requires coordinative and regulatory initiative from a governing locus with the reach to advance a modular architecture that is adaptable for scale, location, and function in achieving, wherever applied, optimal health system effectiveness, efficiency, equity, and continuous learning. This assumes consistent orientation to the following aspects:

- Focusing on creating the conditions for innovation and establishing the relevant ground rules—not dictating or excessively specifying what the "right" solutions should be;
- Ensuring a commitment to public trust, equality, and health;
- Facilitating vital private-public partnerships; and
- Embracing and strategically facilitating incremental innovation, recognizing that solutions will emerge gradually.

Given this pivotal opportunity, the federal government should catalyze innovation centered around the principles of modular design. Imagine the implementation of a newly empowered initiative aimed at digital health innovation, bold in ambition and aligned with President Biden's first challenge to his science advisor of what can we learn from the pandemic about what is possible – or what ought to be possible – to address the widest range of needs related to our public health, but flexible, humble, and pragmatic in approach [137].

Leadership for the National Health Data Architecture

The primary mandate of this sector assessment is to identify shortfalls in health system function and operations and characterize them in a fashion that prompts clarity on the solutions required to eliminate those shortfalls. Organizational approaches through the creation of a new entity or entities—for example, the creation of a new governmental office for digital health integration as an independent or White House level function—are essentially political questions that lie outside the scope of the review. The more fundamental questions relate not to an entity itself but to the capabilities, functions, and authorities that might be ascribed to a new or existing entity.

While the U.S. has many of the individual critical elements required for building and advancing an LHS, these components are rarely orchestrated in concert and cannot easily contribute to or build upon each other. Learning is much more haphazard than it should be. Success demands a system that leverages what is currently known about the current digital infrastructure and up-levels it for the future. Facilitative capacity can be substantial when invested with dynamic leadership and core expertise in areas such as health data architecture, AI/enterprise technology architecture, health care delivery, and cybersecurity.

The broad vision for the Office of the National Coordinator for Health Information Technology (ONC) at its outset implicitly embraced the notion of an expansive program of activities that engages agencies and offices across government and throughout the private sector—e.g., CMS, FDA, CDC, NIH, OCR, and the Office for Human Research Protections (OHRP). In principle, it can launch driver projects and new programs, as well as coordinate with and influence other governmental units such as the VA, the Federal Communications Commission (FCC), and the Federal Trade Commission (FTC), as appropriate; conduct public and private IT policies and initiatives, including prototypes and demonstrations for issues transcending individual agencies; and ensure transparent monitoring and reporting on progress.

Operational arrangements would take advantage of both the policy leverage of the White House and the domain expertise of HHS. Positioned to implement the strategic intent of the LHS with tactical and pragmatic approaches, the ONC was envisioned to regularly convene public-private partnerships and mobilize resources to support focused initiatives and address critical roadblocks or opportunities. The question is whether and how the necessary support might be mobilized.

Enhancing Government Agency Capacity and Decisions

Especially important is the ability to advance cross-agency and cross-sector programs and projects that create new capabilities and demonstrate what is possible. These efforts will be informed by the foundational expectation that solutions will be comprised of modular components—many of which will be connected by APIs, data pipelines, and the like. The role of the ONC or a related cross-government authority would neither be to mandate nor designate a single approach, but rather to nurture a diverse portfolio of modules and interstitial connectors. The goal should be to advance an overall national digital infrastructure and data architecture that promotes continuous innovation, interoperability, and improved data quality and safety.

Continuous dashboarding (e.g., Johns Hopkins COVID Tracking used to track COVID-19) represents a discrete use case, applicable to many health care areas such as other diseases, medical supply chain management, hospital bed availability, and vaccination programs [121]. Results from such dashboarding can be stratified and tailored for region and task. A review of challenges in accessing data and data quality concerns will inform new areas for intervention. A focus on health equity and underserved populations will ensure that critical data elements like race and ethnicity are incorporated in datasets and help to inform critical policy decisions. Making results publicly available will enhance transparency, reproducibility, and trust. Public access should also stimulate new businesses and follow-on innovation. Initiation could be expedited by leveraging the Biomedical Advanced Research and Development Authority (BARDA).

One important tool of an expanded capacity could be convening high-level public-private task forces, working groups, and actions to advance health and the core elements of an LHS. Examples of task force topics could include:

1. options and strategies for computing architecture, data architecture, and data interoperability;
2. elements of a national public health crisis pre-warning system;
3. the issues and options around a national "data trust" for public health;
4. how to operationalize and advance telemedicine and virtual care;
5. approaches to incentivize digital medicine solutions;
6. strategies for using digital capacity to create new evidence and appropriately update clinical care while ensuring data quality and protection; and
7. system-wide cybersecurity.

Achieving the Vision - A Challenging but Important Journey

The COVID-19 pandemic had disastrous consequences for the U.S. and the world, and digital health was a prominent but insufficient part of the national response. The visible rise of telemedicine was a great success story, but addressing the pandemic requires much more support and remediation of the digital health infrastructure. Digital tools helped the nation

make sense of available data and put analytic results to work. Now, however, the inadequacies of digital infrastructure, cultural barriers, out-of-date policies, and misaligned incentives must be addressed. The technological progress of the past two decades and the many successes apparent in the COVID-19 response demonstrate that interoperable health data are essential and that developing the appropriate architectural framework to support fluid data, while being a major undertaking, is achievable.

It will require a multi-pronged approach to build a federal response to the digital health challenge that has sufficient leverage to catalyze the adoption of relevant data standards and architectures. There are living examples where this has been achieved before—for example, the Federal Reserve was created and has provided a practical architecture into which innovators and the banking industry have been able to "plug in." Secure data architecture, parsimonious common data standards, business incentives, and regulatory enforcement together contribute to an infrastructure for a secure financial system. While health care has complexities that do not exist in banking—including the fact that health care data are far more complex than currency data—the existence of a well-functioning Federal Reserve system is a vivid reminder that well-formulated public-private approaches can drive progress.

The task, then, is to lay the groundwork and create the conditions for change of the magnitude required. Success will require advancing a digital infrastructure and a data architecture that promotes modularity, interoperability and innovation; acknowledging that a "one-size-fits-all" approach is unlikely except in clearly identified circumstances; and orchestrating and acting through federal and public-private entities.

The solution requires a focus on fundamental elements woven into the fabric of American health care, including technical infrastructure, data architecture, and modularity. Many of the building blocks are already in place. Now, the country needs to build on this foundation, focusing on incremental progress while prioritizing public need, trust, equality, and innovation. This is the digital infrastructure capacity required for the journey to health and health care that is effective, efficient, equitable, and adds value for every American.

References

1. Kelly, H. 2019. The most disappointing technologies of the decade: Face computers, space tourism and Juicero. *The Washington Post,* December 26. Available at: https://www.washingtonpost.com/technology/2019/12/26/most-disappointing-technologies-decade-face-computers-space-tourism-juicero/ (accessed February 19, 2021).

2. Glaser, J. 2020. It's Time for a New Kind of Electronic Health Record. *Harvard Business Review.* Available at: https://hbr.org/2020/06/its-time-for-a-new-kind-of-electronic-health-record (accessed February 19, 2021).

3. Weiner, M. 2019. Forced Inefficiencies of the Electronic Health Record. *Journal of General Internal Medicine* 34(11): 2299–2301. https://doi.org/10.1007/s11606-019-05281-3.

4. Llupia, A., A. Garcia-Basteiro, and J. Puig. 2020. Still using MS Excel? Implementation of the WHO Go.Data software for the COVID-19 contact tracing. *Health Science Reports* 3(2): e164 https://doi.org/10.1002/hsr2.164.

5. Hern, A. 2020. Covid: how Excel may have caused loss of 16,000 test results in England. *The Guardian,* October 6. Available at: https://www.theguardian.com/politics/2020/oct/05/how-excel-may-have-caused-loss-of-16000-covid-tests-in-england (accessed February 19, 2021).

6. Kliff, S., and M. Sanger Katz. 2020. Bottleneck for U.S. Coronavirus Response: The Fax Machine. *The New York Times,* July 13. Available at: https://www.nytimes.com/2020/07/13/upshot/coronavirus-response-fax-machines.html/ (accessed January 31, 2021).

7. Peiffer-Smadja, N., R. Maatoug, F.X. Lescure, E. D'Ortenzio, J. Pineau, and J. R. King. 2020. Machine Learning for COVID-19 needs global collaboration and data-sharing. *Nature Machine Intelligence* 2(6):293–294. https://doi.org/10.1038/s42256-020-0181-6.

8. Coleridge, S. T. 1834. *The Rime of the Ancient Mariner.* Poetry Foundation. Available at: https://www.poetryfoundation.org/poems/43997/the-rime-of-the-ancient-mariner-text-of-1834 (accessed February 19, 2021).

9. Biden, J. 2021. *A letter to geneticist Eric Lander from President-elect Biden. George Mason University.* Available at: https://science.gmu.edu/news/letter-geneticist-eric-lander-president-elect-biden (accessed February 19, 2021).

10. National Academy of Medicine (NAM). 2021. *Emerging Stronger After COVID-19: Priorities for Health System Transformation.* Available at: https://nam.edu/programs/value-science-driven-health-

care/emerging-stronger-after-covid-19-priorities-for-health-system-transformation/ (accessed November 21, 2021).

11. Interaction Design Foundation. 2020. *Understanding Early Adopters and Customer Adoption Patterns.* Available at: https://www.interaction-design.org/literature/article/understanding-early-adopters-and-customer-adoption-patterns (accessed February 19, 2021).

12. Harford, T. 2017. Why didn't electricity immediately change manufacturing? *BBC News,* August 21. Available at: https://www.bbc.com/news/business-40673694 (accessed February 19, 2021).

13. Sun, J., X. Chen, Z. Zhang, S. Lai, B. Zhao, H. Liu, S. Wang, W. Huan, R. Zhao, M. Alexander, and Y. Zheng. 2020. Forecasting the long-term trend of COVID-19 epidemic using a dynamic model. *Scientific Reports* 10(21122). https://doi.org/10.1038/s41598-020-78084-w.

14. Blumenthal, D., E. Fowler, M. Abrams, and S. Collins. 2020. Covid-19 — Implications for the Health Care System. *New England Journal of Medicine* 383(15):1483-1488. https://www.nejm.org/doi/full/10.1056/NEJMsb2021088.

15. Joyner, M., R. Carter, J. Senefeld, S. Klassen, J. Mills, P. Johnson, E. Theel, C. Wiggins, K. Bruno, A. Klompas, E. Lesser, K. Kunze, M. A. Sexton, J. C. Diaz Soto, S. E. Baker, J. R. A. Shepherd, N. van Helmond, N. C. Verdun, P. Marks, C. M. van Buskirk, J. L. Winters, J. R. Stubbs, R. F. Rea, D. O. Hodge, V. Herasevich, E. R. Whelan, A. J. Clayburn, K. F. Larson, J. G. Ripoli, K. J. Anderssen, M. R. Buras, M. N. P. Vogt, J. J. Dennis, R. J. Regimbal, P. R. Bauer, J. E. Blair, N. S. Paneth, D. Fairweather, R. S. Wright, and A. Casadevall. 2021. Convalescent Plasma Antibody Levels and the Risk of Death from Covid-19. *New England Journal of Medicine.* https://www.nejm.org/doi/full/10.1056/NEJMoa2031893.

16. Centers for Disease Control and Prevention (CDC). 2021. *National Healthcare Safety Network (NHSN).* Available at: https://www.cdc.gov/nhsn/index.html (accessed February 19, 2021).

17. Centers for Medicare and Medicaid Services (CMS). 2020. *Trump Administration Engages America's Hospitals in Unprecedented Data Sharing.* Available at: https://www.cms.gov/newsroom/press-releases/trump-administration-engages-americas-hospitals-unprecedented-data-sharing (accessed February 19, 2021).

18. Christensen, C. M. 1997. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Boston, MA: Harvard Business School Press.

19. Perrin, P. B., B. S. Pierce, and T. R. Elliott. 2020. COVID-19 and telemedicine: A revolution in healthcare delivery is at hand. *Health Science Reports* 3(2):e166. https://doi.org/10.1002/hsr2.166.

20. U.S. Department of Health and Human Services (HHS). 2021. *Enforcement Highlights.* Available at: https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/data/enforcement-highlights/index.html (accessed February 19, 2021).

21. Keesara, S., A. Jonas, and K. Schulman. 2020. Covid-19 and Health Care's Digital Revolution. *New England Journal of Medicine* 382(23):e82. https://www.nejm.org/doi/10.1056/NEJMp2005835.

22. General Data Protection Regulation (GDPR). n.d. *Home.* Available at: https://gdpr-info.eu/ (accessed February 19, 2021).

23. California Consumer Privacy Act (CCPA). 2018. *Home.* Available at: https://oag.ca.gov/privacy/ccpa (accessed February 19, 2021).

24. Schulman, K., and B. D. Richman. 2019. Toward an Effective Innovation Agenda. *New England Journal of Medicine* 380(10):900-901. https://www.nejm.org/doi/full/10.1056/NEJMp1812460.

25. Ratwani, R. M., E. Savage, A. Will, R. Arnold, S. Khairat, K. Miller, R. J. Fairbanks, M. Hodgkins, and A. Z. Hettinger. 2018. A usability and safety analysis of electronic health records: a multi-center study. *Journal of the American Medical Informatics Association* 25(9):1197–1201. https://doi.org/10.1093/jamia/ocy088.

26. Downing, N. L., D. W. Bates, and C. A. Longhurst. 2018. Physician Burnout in the Electronic Health Record Era: Are We Ignoring the Real Cause? *Annals of Internal Medicine* 169(1):50-51. https://doi.org/10.7326/M18-0139.

27. Holmgren, A. J., N. L. Downing, D. W. Bates, T. D. Shanafelt, A. Milstein, C. D. Sharp, D. M. Cutler, R. S. Huckman, and K. A. Schulman. 2020. Assessment of Electronic Health Record Use Between US and Non-US Health Systems. *JAMA Internal Medicine* 181(2):251-259. https://doi.org/10.1001/jamainternmed.2020.7071.

28. Tseng, P., R. S. Kaplan, B. D. Richman, M. A. Shah, and K. A. Schulman. 2018. Administrative Costs Associated With Physician Billing and Insurance-Related Activities at an Academic Health Care System. *JAMA* 319(7):691-697. https://doi.org/10.1001/jama.2017.19148.

29. Collins, M. 2019. *The New PubMed is Here.* National Library of Medicine Technical Bulletin. Available at: https://www.nlm.nih.gov/pubs/techbull/nd19/nd19_pubmed_new.html (accessed February 19, 2021).

30. UpToDate. n.d. *Home.* Available at: https://www.uptodate.com/home (accessed February 19, 2021).

31. Centers for Disease Control and Prevention (CDC). 2021. *United States COVID-19 Cases and Deaths by State.* Available at: https://covid.cdc.gov/covid-data-tracker/#cases_casesper100k-last7days (accessed December 31, 2020).

32. John Hopkins University (JHU). 2021. *COVID-19 United States Cases by County.* Available at: https://coronavirus.jhu.edu/us-map (accessed December 31, 2020).

33. Centers for Disease Control and Prevention (CDC). 2021. *V-safe.* Available at: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/vsafe.html (accessed February 19, 2021).

34. Centers for Disease Control and Prevention (CDC). 2019. *Digital Contact Tracing Tools.* Available at: https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/digital-contact-tracing-tools.html (accessed February 3, 2021).

35. Sittig, D. F. and H. Singh. 2020. COVID-19 and the Need for a National Health Information Technology Infrastructure. *JAMA* 323(23):2373-2374. https://doi.org/10.1001/jama.2020.7239.

36. Centers for Disease Control and Prevention (CDC). 2021. *Data Modernization Initiative.* Available at: https://www.cdc.gov/surveillance/surveillance-data-strategies/data-IT-transformation.html (accessed December 31, 2020).

37. Distillery Trail. 2020. *Tennessee Distillers Shift Spirit Production to Hand Sanitizer to Support States Essential Personnel.* Available at: https://www.distillerytrail.com/blog/tennessee-distillers-shift-spirit-production-to-hand-sanitizer-to-support-states-essential-personnel/ (accessed December 31, 2020).

38. Lagu, T., R. Werner, and A. Artenstein. 2020. Why don't hospitals have enough masks? Because coronavirus broke the market. *The Washington Post,* May 21. Available at: https://www.washingtonpost.com/outlook/2020/05/21/why-dont-hospitals-have-enough-masks-because-coronavirus-broke-market/ (accessed December 31, 2020).

39. Ritterband, L. M., F. P. Thorndike, and K. S. Inger-soll. 2017. Effect of a Web-Based Cognitive Behavior Therapy for Insomnia Intervention With 1-Year Follow-up: A Randomized Clinical Trial. *JAMA Psychiatry* 74(1):68-75. https://www.doi.org/10.1001/jamapsychiatry.2016.3249.

40. Kuhn, E., N. Kanuri, J. E. Hoffman, D. W. Gavert, J. I. Ruzeck, and C. B. Taylor. 2017. A randomized controlled trial of a smartphone app for post-traumatic stress disorder symptoms. *Journal of Consulting and Clinical Psychology* 85(3):267-273. https://www.doi.org/10.1037/ccp0000163.

41. Jarorski, B. K., K. Taylor, K. M. Ramsey, A. Heinz, S. Steinmetz, I. Pagano, G. Moraja, and J. E. Owen. 2021. COVID Coach: Exploring Usage of a Public Mental Health App Designed for the COVID-19 Pandemic. *Journal of Medical Internet Research* 23(3):e26559. https://www.doi.org/10.2196/26559.

42. Taquet, M., S. Luciano, J. R. Geddes, and P. J. Harrison. 2020. Bidirectional associations between COVID-19 and psychiatric disorder: retrospective cohort studies of 62 354 COVID-19 cases in the USA. *The Lancet Psychiatry* 8(2): 130-140. https://doi.org/10.1016/S2215-0366(20)30462-4.

43. Panchal N., R. Kamal, C. Cox, and R. Garfield. 2021. The Implications of COVID-19 for Mental Health and Substance Use. *Kaiser Family Foundation.* Available at: https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/ (accessed February 25, 2021).

44. Fegert, J. M., B. Vitiello, P. L. Plener and V. Clemens. 2020. Challenges and burden of the Coronavirus 2019 (COVID-19) pandemic for child and adolescent mental health: a narrative review to highlight clinical and research needs in the acute phase and the long return to normality. *Child and Adolescent Psychiatry and Mental Health* 14(20). https://doi.org/10.1186/s13034-020-00329-3.

45. Singh, S., D. Roy, K. Sinha, S. Parveen, G. Sharma, and G. Joshi. 2020. Impact of COVID-19 and lockdown on mental health of children and adolescents: A narrative review with recommendations. *Elsevier Psychiatry Research* 293(113429). https://doi.org/10.1016/j.psychres.2020.113429.

46. Moreno, C., T. Wykes, S. Galderisi, M. Mordentoft, N. Crossley, N. Jones, M. Cannon, C. U. Corell, L. Byrne, S. Carr, E. Y. H. Chen, P. Gorwood, S. Johnson, H. Karkkainen, J. H. Krystal, J. Lee, J. Lieberman, C. Lopez-Jaramillo, M. Mannikko, M. R. Phillips, H. Uchida, E. Vieta, A. Vita, and C. Arango.

2020. How mental health care should change as a consequence of the COVID-19 pandemic. *The Lancet Psychiatry* 7(9):813-824. https://doi.org/10.1016/S2215-0366(20)30307-2.

47. Crisis Text Line. 2021. *Home.* Available at: https://www.crisistextline.org/ (accessed February 19, 2021).

48. Pear Therapeutics. 2020. *Pear Therapeutics Obtains FDA Authorization for SOMRYST™, a Prescription Digital Therapeutic for the Treatment of Adults with Chronic Insomnia.* Available at: https://peartherapeutics.com/pear-therapeutics-obtains-fda-authorization-for-somryst-a-prescription-digital-therapeutic-for-the-treatment-of-adults-with-chronic-insomnia/ (accessed February 19, 2021).

49. Pear Therapeutics. 2017. *PEAR Obtains FDA Clearance of the First Prescription Digital Therapeutic to Treat Disease.* Available at: https://peartherapeutics.com/fda-obtains-fda-clearance-first-prescription-digital-therapeutic-treat-disease/ (accessed February 19, 2021).

50. Food and Drug Administration (FDA). 2020. *FDA Permits Marketing of First Game-Based Digital Therapeutic to Improve Attention Function in Children with ADHD.* Available at: https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-first-game-based-digital-therapeutic-improve-attention-function-children-adhd (accessed February 19, 2021).

51. Larsen, M. E., K. Huckvale, J. Nicholas, J. Torous, L. Birrell, E. Li, and B. Reda. 2019. Using science to sell apps: Evaluation of mental health app store quality claims. *NPJ Digital Medicine* 2(18). https://doi.org/10.1038/s41746-019-0093-1.

52. Food and Drug Administration (FDA). 2020. *Enforcement Policy for Digital Health Devices For Treating Psychiatric Disorders During the Coronavirus Disease 2019 (COVID-19) Public Health Emergency.* Available at: https://www.fda.gov/media/136939/download (accessed February 19, 2021).

53. National Institutes of Health, Office of the Director (NIH). 2020. *Final NIH Policy for Data Management and Sharing.* Available at: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html (accessed February 19, 2021).

54. Centers for Disease Control and Prevention (CDC). 2014. *Standards to Facilitate Data Sharing.* Available at: https://www.cdc.gov/nchhstp/programintegration/sc-standards.htm (accessed February 19, 2021).

55. American Hospital Association (AHA). 2019. *Sharing Data, Saving Lives: The Hospital Agenda for Interoperability.* Available at: https://www.aha.org/system/files/2019-01/Report01_18_19-Sharing-Data-Saving-Lives_FINAL.pdf (accessed February 19, 2021).

56. Mijuk, G. 2018. *Drug development gets big data analytics boost.* Novartis. Available at: https://www.novartis.com/stories/discovery/drug-development-gets-big-data-analytics-boost (accessed January 31, 2021).

57. Duran, C., M. Bonam, and A. Campbell. 2020. Digital health: Improving patient outcomes *through digitally transformed R&D.* AstraZeneca. Available at: https://www.astrazeneca.com/what-science-can-do/labtalk-blog/uncategorized/digital-health-improving-patient-outcomes-through-digitally-transformed-rd.html (accessed January 31, 2021).

58. U.S. Food and Drug Administration (FDA). 2021. *FDA Guidance on Conduct of Clinical Trials of Medical Products during the COVID-19 Public Health Emergency.* Rockville, MD: FDA Dockets Management.

59. Wierzba, O., B. Ziegler, J. F. Bobier, T. Delano, and J. B. Vermersch. 2020. *COVID-19 Opens a New Era for Real-World Evidence in Pharma.* Boston Consulting Group. Available at: https://www.bcg.com/publications/2020/covid-19-opens-a-new-era-for-real-world-evidence-in-pharma (accessed February 19, 2020).

60. RECOVERY (Randomized Evaluation of COVID-19 Therapy). 2021. *Home.* Available at: https://www.recoverytrial.net/ (accessed February 19, 2020).

61. Health Data Research UK (HDR UK). 2020. *Filling in the gaps: smart use of health data lies behind the RECOVERY trial's success.* Available at: https://www.hdruk.org/news/filling-in-the-gaps-smart-use-of-health-data-lies-behind-the-recovery-trials-success/ (accessed February 19, 2020).

62. Quantum Leap Healthcare Collaborative. 2019. *I-SPY Trials.* Available at: https://www.ispytrials.org/ (accessed February 19, 2020).

63. Plump, A. and D. Reese. 2020. Using a global network of adaptive clinical trials to fight Covid-19. *STAT News,* July 23. Available at: https://www.statnews.com/2020/07/23/using-a-global-network-of-adaptive-clinical-trials-to-fight-covid-19/ (accessed February 19, 2020).

64. Franklin, J. M., A. Pawar, D. Martin, R. J. Glynn, M. Levenson, R. Temple, and S. Schneeweiss. 2020. Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. *Clinical Pharmacology & Therapeutics* 107(4):817-826. https://www.doi.org/10.1002/cpt.1633.

65. Brown, J. S., J. C. Maro, M. Nguyen, and R. Ball. 2020. Using and improving distributed data networks to generate actionable evidence: the case of real-world outcomes in the Food and Drug Administration's Sentinel system. *Journal of the American Medical Informatics Association* 27(5):793-797. https://www.doi.org/10.1093/jamia/ocaa028.

66. Miksad, R. A., and A. P. Abernethy. 2018. Harnessing the Power of Real-World Evidence (RWE): A Checklist to Ensure Regulatory-Grade Data Quality. *Clinical Pharmacology & Therapeutics* 103(2):202-205. https://www.doi.org/10.1002/cpt.946.

67. Piller, C. and K. Servick. 2020. Two elite medical journals retract coronavirus papers over data integrity questions. *American Association for the Advancement of Science.* Available at: https://www.sciencemag.org/news/2020/06/two-elite-medical-journals-retract-coronavirus-papers-over-data-integrity-questions (accessed March 11, 2021).

68. COVID-19 Evidence Accelerator. 2020. *Home.* Available at: https://evidenceaccelerator.org/ (accessed February 21, 2021).

69. Observational Health Data Sciences and Informatics (OHDSI). 2021. *Home.* Available at: https://www.ohdsi.org/ (accessed January 31, 2021).

70. U.S. Department of Veteran Affairs (VA). 2021. *Novel Coronavirus Disease (COVID-19).* Available at: https://www.publichealth.va.gov/n-coronavirus/ (accessed March 11, 2021).

71. Deardorff, A. 2021. *COVID-19 Research and Information Resources. University of California San Francisco.* Available at: https://guides.ucsf.edu/COVID19/data (accessed March 11, 2021).

72. Shaywitz, D. 2021. *What's Your DEQ? Why Data Empathy is Essential for Health Data Impact.* Timmerman Report. Available at: https://timmermanreport.com/2021/01/whats-your-deq-why-data-empathy-is-essential-for-health-data-impact/ (accessed January 31, 2021).

73. Shaywitz, D. 2020. *Learning From COVID-19: The Lessons For Real World Data.* Timmerman Report. Available at: https://timmermanreport.com/2020/10/learning-from-covid-19-the-lessons-for-real-world-data/ (accessed March 11, 2021).

74. Tancev, G. 2020. *Predictive Maintenance within the Internet of Things.* Towards Data Science. Available at: https://towardsdatascience.com/strategies-for-predictive-maintenance-within-the-internet-of-things-c66adcba037d (accessed February 2, 2021).

75. Google Cloud. 2020. *Using Machine Learning on Compute Engine to Make Product Recommendations.* Available at: https://cloud.google.com/solutions/recommendations-using-machine-learning-on-compute-engine (accessed February 2, 2021).

76. Rice, J. 2019. *Artificial Intelligence and Machine Learning in Supply Chain Planning.* MIT Center for Transportation & Logistics. Available at: https://ctl.mit.edu/pub/report/artificial-intelligence-machine-learning-roundtable (accessed February 2, 2021).

77. GPT3 Examples. n.d. *Home.* Available at: https://gpt3examples.com/ (accessed February 2, 2021).

78. DeepMind. 2020. *AlphaFold: a solution to a 50-year-old grand challenge in biology.* Available at: https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology (accessed February 2, 2021).

79. Wheeling, K. 2020. Machine Learning Improves Weather and Climate Models. *Eos: Science News by the American Geophysical Union.* Available at: https://eos.org/research-spotlights/machine-learning-improves-weather-and-climate-models (accessed February 2, 2021).

80. Lyons, K. 2020. CDC uses Microsoft healthcare chatbot service to create coronavirus symptom checker. *The Verge.* Available at: https://www.theverge.com/2020/3/21/21189227/cdc-microsoft-chatbot-coronavirus-symptom-checker (accessed February 2, 2021).

81. Brothers, W. 2020. *A Timeline of COVID-19 Vaccine Development.* BioSpace. Available at: https://www.biospace.com/article/a-timeline-of-covid-19-vaccine-development/ (accessed February 2, 2021).

82. Magrabi, F., E. Ammenwerth, J. B. McNair, N. F. D. Keizer, H. Hyppönen, P. Nykänen, M. Rigby, P. Scott, T. Vehko, Z. S. Wong, and A. Georgiou. 2019. Artificial Intelligence in Clinical Decision Support: Challenges for Evaluating AI and Practical Implications. *Thieme Medical Publishers* 28(01): 128-134. https://doi.org/10.1055/s-0039-1677903.

83. Fast Health Interoperability Resources (FHIR). n.d. *Home.* Available at: http://www.fhir.org/ (accessed February 2, 2021).

84. SMART Health IT. 2020. *Standards and Specifications.* Available at: https://docs.smarthealthit.org/ (accessed February 2, 2021).

85. Baroni, K. 2018. *Getting AI/ML and DevOps working better together.* Microsoft Azure. Available at: https://azure.microsoft.com/en-us/blog/getting-ai-ml-and-devops-working-better-together/ (accessed February 2, 2021).

86. National Academy of Medicine (NAM). n.d. *The Learning Health System Series.* Available at: https://nam.edu/programs/value-science-driven-health-care/learning-health-system-series/ (accessed December 31, 2020).

87. National Institutes of Health (NIH). 2021. *All of Us Research Program.* Available at: https://allofus.nih.gov/ (accessed February 21, 2021).

88. UK Biobank. 2021. *Home.* Available at: https://www.ukbiobank.ac.uk/ (accessed February 21, 2021).

89. European Centre for Disease Prevention and Control (ECDC). 2021. *SARS-CoV-2 increased circulation of variants of concern and vaccine rollout in the EU/EEA, 14th update.* Available at: https://www.ecdc.europa.eu/sites/default/files/documents/RRA-covid-19-14th-update-15-feb-2021.pdf (accessed February 21, 2021).

90. Institute of Medicine. 2007. *The Learning Healthcare System: Workshop Summary.* Washington, DC: The National Academies Press.https://doi.org/10.17226/11903.

91. U.S. Department of Health and Human Services (HHS). 2017. *HITECH Act Enforcement Interim Final Rule.* Available at: https://www.hhs.gov/hipaa/for-professionals/special-topics/hitech-act-enforcement-interim-final-rule/index.html (accessed February 21, 2021).

92. U.S. Food and Drug Administration (FDA). 2020. *21st Century Cures Act.* Available at: https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act (accessed February 21, 2021).

93. Singh, J. 2020. The Digital Divide can Kill —Can we Bridge it? *Medium.* Available at: https://jagsing-hmd.medium.com/the-digital-divide-can-kill-can-we-bridge-it-93b136bdaf35 (accessed January 1, 2021).

94. Maybank, A. 2020. Why racial and ethnic data on COVID-19's impact is badly needed. *American Medical Association.* Available at: https://www.ama-assn.org/about/leadership/why-racial-and-ethnic-data-covid-19-s-impact-badly-needed (accessed March 11, 2021).

95. The Economist. 2020. *A lack of data on race hampers efforts to tackle inequalities.* Available at: https://www.economist.com/leaders/2020/11/21/a-lack-of-data-on-race-hampers-efforts-to-tackle-inequalities (accessed March 11, 2021).

96. Racial Bias in Pulse Oximetry Measurement. 2020. *New England Journal of Medicine* 383(25):2477-2478. https://www.nejm.org/doi/10.1056/NEJMc2029240.

97. Kates, J., J. Tolbert, and J. Michaud. 2021. The COVID-19 "Vaccination Line": An Update on State Prioritization Plans. *Kaiser Family Foundation.* Available at: https://www.kff.org/coronavirus-covid-19/issue-brief/the-covid-19-vaccination-line-an-update-on-state-prioritization-plans/ (accessed February 10, 2021).

98. National Academies of Sciences, Engineering, and Medicine. 2020. *Framework for Equitable Allocation of COVID-19 Vaccine.* Washington, DC: The National Academies Press.https://doi.org/10.17226/25917.

99. Mental Health America (MHA). 2020. *Mental Health of Healthcare Workers In COVID-19.* Available at: https://mhanational.org/mental-health-healthcare-workers-covid-19 (accessed February 21, 2021).

100. Physician's Weekly. 2018. *3 Ways Telemedicine Reduces Provider Burnout.* Available at: https://www.physiciansweekly.com/3-ways-telemedicine-reduces-provider-burnout/ (accessed March 29, 2021).

101. Miliard, M. 2019. Clinician burnout: Physicians name the technologies they think could best solve it. *Healthcare IT News.* Available at: https://www.healthcareitnews.com/news/clinician-burnout-physicians-name-technologies-they-think-could-best-solve-it (accessed March 11, 2021).

102. The World Bank. 2019. *Public-Private Partnerships.* Available at: https://www.worldbank.org/en/topic/publicprivatepartnerships/overview#2 (accessed February 21, 2021).

103. Centers for Medicare & Medicaid Services (CMS). 2020. *National Health Expenditure Data – Historical.* Available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical (accessed February 21, 2021).

104. Office of the Assistant Secretary for Planning and Evaluation (ASPE). 1998. *Health Insurance Portability and Accountability Act of 1996 – Implementation of Administrative Simplification Requirements by HHS.* Available at: https://www.commonwealthfund.org/publications/issue-briefs/2020/jan/us-health-care-global-perspective-2019 (accessed February 21, 2021).

105. Tikkanen, R. and M. K. Abrams. 2020. U.S. Health Care from a Global Perspective, 2019: Higher Spending, Worse Outcomes? *The Commonwealth Fund.* Available at: https://www.commonwealthfund.org/publications/issue-briefs/2020/jan/us-health-care-global-perspective-2019 (accessed February 21, 2021).

106. Sanger, D. E. 2021. Russian Hackers Broke Into Federal Agencies, U.S. Officials Suspect. *The New York Times,* December 13. Available at: https://www.nytimes.com/2020/12/13/us/politics/russian-hackers-us-government-treasury-commerce.html (accessed January 3, 2021).

107. Perlroth, N. 2021. Microsoft Says Russian Hackers Viewed Some of Its Source Code. *The New York Times,* December 31. Available at: https://www.nytimes.com/2020/12/31/technology/microsoft-russia-hack.html (accessed January 3, 2021).

108. Seals, T. 2020. *Food-Supply Giant Americold Admits Cyberattack.* Threat Post. Available at: https://threatpost.com/food-supply-americold-cyberattack/161402/ (accessed January 3, 2021).

109. O'Murchu, C. 2020. Vaccines for sale on dark web as criminals target pandemic profits. *Financial Times.* Available at: https://www.ft.com/content/8bfc674e-efe6-4ee0-b860-7fcb5716bed6 (accessed January 3, 2021).

110. BBC News. 2020. *Pfizer/BioNTech vaccine docs hacked from European Medicines Agency.* Available at: https://www.bbc.com/news/technology-55249353 (accessed January 3, 2021).

111. Barrett, B. 2020. Hackers Are Targeting the Covid-19 Vaccine 'Cold Chain'. *WIRED.* Available at: https://www.wired.com/story/hackers-targeting-covid-19-vaccine-cold-chain/ (accessed January 3, 2021).

112. Sanger, D. E., and S. LaFraniere. 2020. Cyberattacks Discovered on Vaccine Distribution Operations. *The New York Times,* December 3. Available at: https://www.nytimes.com/2020/12/03/us/politics/vaccine-cyberattacks.html (accessed January 3, 2021).

113. Peterson, A. 2013. Yes, terrorists could have hacked Dick Cheney's heart. *The Washington Post,* October 21. Available at: https://www.washingtonpost.com/news/the-switch/wp/2013/10/21/yes-terrorists-could-have-hacked-dick-cheneys-heart/ (accessed January 3, 2021).

114. Bajak, F. 2020. FBI warns ransomware assault threatens US healthcare system. *Associated Press News.* Available at: https://apnews.com/article/fbi-warns-ransomware-healthcare-system-548634f03e71a830811d291401651610 (accessed January 3, 2021).

115. U.S. Department of Health and Human Services (HHS). 2020. *Cyber Security Guidance Material.* Available at: https://www.hhs.gov/hipaa/for-professionals/security/guidance/cybersecurity/index.html (accessed February 25, 2021).

116. U.S. Food and Drug Administration (FDA). 2020. *Cybersecurity.* Available at: https://www.fda.gov/medical-devices/digital-health-center-excellence/cybersecurity (accessed January 3, 2021).

117. Finlayson, S., H. W. Chung, I. S. Kohane, and A. L. Beam. 2018. Adversarial Attacks Against Medical Deep Learning Systems. *ARXIV preprint* arXiv:1804.05296, https://arxiv.org/abs/1804.05296 (accessed January 4, 2021).

118. MITRE. 2018. *Medical Device Cybersecurity Regional Incident Preparedness and Response Playbook.* Available at: https://www.mitre.org/sites/default/files/publications/pr-18-1550-Medical-Device-Cybersecurity-Playbook.pdf (accessed February 25, 2021).

119. Kelly, S. 2019. FDA 'white hat' hacking tools may combat medical misinformation. *MedTech Dive.* Available at: https://www.medtechdive.com/news/fda-white-hat-hacking-tools-may-combat-medical-misinformation/556987/ (accessed January 3, 2021).

120. Williams, J. 2020. What You Need to Know About the SolarWinds Supply-Chain Attack. *SANS Institute.* Available at: https://www.sans.org/blog/what-you-need-to-know-about-the-solarwinds-supply-chain-attack/ (accessed February 21, 2021).

121. John Hopkins University (JHU). 2021. *Tracking.* Available at: https://coronavirus.jhu.edu/data (accessed February 21, 2021).

122. U.S. Department of Health and Human Services (HHS). 2009. *HITECH Act of 2009, Public Law 111-5 (February 2009).* Available at: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf (accessed

November 22, 2021).

123. Glaser, J. 2019. What Banking Can Teach Health Care About Handling Customer Data. *Harvard Business Review,* October 14. Available at: https://hbr.org/2019/10/what-banking-can-teach-health-care-about-handling-customer-data (accessed November 22, 2021).

124. Gilson, S. F., C. A. Umscheid, N. Laiteerapong, G. Ossey, K. J. Nunes, and S. D. Shah. 2020. Growth of Ambulatory Virtual Visits and Differential Use by Patient Sociodemographics at One Urban Academic Medical Center During the COVID-19 Pandemic: Retrospective Analysis. *JMIR Medical Informatics* 8(12):e24544. https://doi.org/10.2196/24544.

125. Shaywitz, D. 2019. A Drug is not an Outcome: Extending Translation Through Implementation Using Real-World Data. *Clinical Pharmacology & Therapeutics.* https://doi.org/10.1002/cpt.1382.

126. Witry, M. J., W. R. Doucette, J. M. Daly, B. T. Levy, and E. A. Chrischilles. 2010. Family Physician Perceptions of Personal Health Records. *Perspectives in Health Information Management.* Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2805556/ (accessed November 22, 2021).

127. Liu, L. S., P. C. Shih, and G. R. Hayes. 2011. Barriers to the adoption and use of personal health record systems. *iConference '11: Proceedings of the 2011 iConference.* https://doi.org/10.1145/1940761.1940811.

128. National Center for Advancing Translational Sciences. n.d. *N3C Data Access Forms and Resources.* Available at: https://ncats.nih.gov/n3c/resources/data-access (accessed November 22, 2021).

129. Shaywitz, D. 2021. *Aspiring Healthtech Companies Require A Nuanced Understanding of Healthcare's Human Dynamics.* Timmerman Report, November 7. Available at: https://timmermanreport.com/2021/11/aspiring-healthtech-companies-require-a-nuanced-understanding-of-healthcares-human-dynamics/ (accessed November 22, 2021).

130. Tollen, L., E. Keating, and A. Weil. 2020. How Administrative Spending Contributes to Excess U.S. Health Spending. *Health Affairs Blog.* https://doi.org/10.1377/hblog20200218.375060.

131. Andrews, T. M. 2021. For younger generation, securing vaccine appointments for parents can be 'full-time job.' *The Washington Post,* February 2. Available at: https://www.washingtonpost.com/technology/2021/02/02/vaccine-appointment-parents-kids/ (accessed November 22, 2021).

132. Weil, J. Z., G. S. Schneider, and E. Cox. 2021. Confusion and chaos: Inside the vaccine rollout in D.C., Maryland, and Virginia. *The Washington Post,* February 9. Available at: https://www.washingtonpost.com/local/vaccine-rollout-dc-maryland-virginia/2021/02/07/11d8c4a0-656d-11eb-8c64-9595888caa15_story.html (accessed November 22, 2021).

133. National Institute of Mental Health. n.d. *Mental Illness.* Available at: https://www.nimh.nih.gov/health/statistics/mental-illness (accessed January 10, 2022).

134. Centers for Disease Control and Prevention. n.d. *WISQARS™ — Web-based Injury Statistics Query and Reporting System.* Available at: https://www.cdc.gov/injury/wisqars/index.html (accessed January 10, 2022).

135. Vyas, D. A., L. G. Eisenstein, and D. S. Jones. 2020. Hidden in Plain Sight – Reconsidering the Use of Race Correction in Clinical Algorithms. *New England Journal of Medicine* 383: 874-882. https://doi.org/10.1056/NEJMms2004740.

136. Weinstein, J. N., W. B. Weeks, and J. S. Skinner. 2021. The Federal Health Authority, a Federal Reserve System for Health Care: COVID-19 Has Exposed a Need for Change. *JAMA Health Forum* 2(1):e201503. https://doi.org/10.1001/jamahealthforum.2020.1503.

137. View From Arch Street. 2021. *Biden Challenges Incoming Science Advisor: Five Questions on Science in the National Interest.* January 15. Available at: https://viewfromarchstreet.com/2021/01/15/biden-challenges-incoming-science-advisor-five-questions-on-science-in-the-national-interest/ (accessed January 10, 2022).

DOI

Suggested Citation

Author Information

Peter Lee, PhD, is Corporate Vice President of Research and Incubations at Microsoft. **Amy Abernethy, MD, PhD,** is President of Clinical Research Business at Verily. **David Shaywitz, MD, PhD,** is the Founder of Astounding HealthTech. **Adi V. Gundlapalli, MD, PhD, MS,** is Chief Public Health Informatics Officer at the U.S. Centers for Disease Control and Prevention. **Jim Weinstein, DO,** is Senior Vice President at Microsoft Healthcare. **P. Murali Doraiswamy, MBBS, FRCP,** is Professor of Psychiatry and Medicine at Duke University School of Medicine. **Kevin Schulman, MD,** is Professor of Medicine at Stanford University School of Medicine. **Subha Madhavan, PhD,** is Head of Data Science, Oncology R&D, at AstraZeneca.

Acknowledgments

Conflict-of-Interest Disclosures

Dr. Abernethy discloses receiving personal fees from Flatiron Health (Roche Group), AthenaHealth, SignalPath, and CareDx no later than January 2019. Dr. Shaywitz discloses being the founder of Astounding HealthTech, a consultancy providing advisory services to senior biopharmaceutical R&D leaders related to the implementation of emerging digital and data technologies. Dr. Doraiswamy discloses receiving personal fees from Brain Forum, VitaKey, Neuroglee, Nutricia, Transposon, Lumos, Otsuka, Lilly and Clearview; receiving grants from Avanir, Lilly, the National Institutes of Health, the U.S. Department of Defense, DARPA, ONR, Cure Alzheimer's Fund, Steve Aoki Foundation, Bausch Health, and Wrenn Trust; shareholdings in Advera Health, Transposon, Marvel Biome, UMethod and Evidation; and holding pending patents in wearables for infection detection and biomarkers of cognitive dysfunction. Dr. Madhavan is currently employed by AstraZeneca Pharmaceuticals and is the Head of Data Science in Oncology R&D. She maintains an adjunct faculty position at Georgetown University and teaches health data science in the graduate school. Dr Schulman reported receiving grants from NBA Inc; being a board member and shareholder of Grid Therapeutics; being a managing director of and shareholder in Faculty Connection, LLC; being on 10 the board of advisors for and shareholder in Prealize, being a board member and shareholder in Reserve Therapeutics (Pharmaceutical Services Company); passive investor in Altitude Ventures Inc and Excelerate Health Ventures; receiving personal fees from Novartis for consulting on study design and from Cytokinetics, HealthQuest, Business Roundtable, Motley Rice, Frazier Healthcare Partners, ISMIE, and Business School Alliance for Health Management President; being a senior associate editor for Health Services Research; being on an advisory board for CivicaRx (uncompensated); and being a board member for Catalysis (uncompensated) outside the submitted work.

Correspondence

Disclaimer