Pre-print
Winter, J.S., & Davidson, E. (2022) "Harmonizing regulatory regimes for the governance of patient-generated health data." *Telecommunications Policy*, *46*(5). https://doi.org/10.1016/j.telpol.2021.102285

Jenifer Sunrise Winter, University of Hawaii at Manoa

Elizabeth Davidson, University of Hawaii at Manoa

Harmonizing Regulatory Regimes for the Governance of Patient-generated Health Data

Abstract

Patient-generated health data (PGHD), created and captured from patients via wearable devices and mobile apps, are proliferating outside of clinical settings. Examples include sleep trackers, fitness trackers, continuous glucose monitors, and RFID-enabled implants, with many additional biometric or health surveillance applications in development or envisioned. These data are included in growing stockpiles of personal health data (PHI) being mined for insight by health economists, policy analysts, researchers, and health system organizations. Dominant narratives position these highly personal data as valuable resources to transform healthcare, stimulate innovation in medical research, and engage individuals in their health and healthcare. Large tech companies are also increasingly implicated in these areas, through mobile health application sales and data acquisitions. Given the many possible uses and users for PGHD, ensuring privacy, security, and equity of benefits from PGHD will be challenging. This is due in part to disparate regulatory policies and practices across technology firms, health system organizations, and health researchers. Rapid developments with PGHD technologies and the lack of harmonization between regulatory regimes may render existing safeguards to preserve patient privacy and control over their PGHD ineffective, while also failing to guide PGHD-related innovation in socially desirable directions. Using a policy regime lens to explore these challenges, we examine three existing data protection regimes relevant to PGHD in the United States that are currently in tension with one another: federal and state health-sector laws, regulations on data use and reuse for research and innovation, and industry self-regulation of consumer privacy by large tech companies. We argue that harmonization of these regimes is necessary to meet the challenges of PGHD data governance. We next examine emerging governing instruments, identifying three types of structures (organizational, regulatory, technological/algorithmic), which synergistically could help enact needed regulatory oversight while limiting the friction and economic costs of regulation that may hinder innovation. This policy analysis provides a starting point for further discussions and negotiations among stakeholders and regulators to do so.

Keywords: Patient-generated health data, PGHD, governance, privacy, Big Tech, regulation

1. Introduction

Mobile health apps and wearable monitoring devices are growing in popularity, and patient-generated health data (PGHD) – health data created and captured by or from patients or other non-clinical actors through these devices – are proliferating outside of clinical settings. PGHD include, but are not limited to, biometric data (e.g., from remote monitoring), self-reported health measures, health-related activity data (e.g., exercise, diet, sleep), symptoms, and health and treatment history. PGHD are distinguished from clinical data in that the patients are the primary capturers of their own data, and patients currently have some control over how, and with whom, to share these data (Office of the National Coordinator for Health Information Technology, 2014). Existing examples of PGHD applications include sleep trackers, fitness

trackers (e.g., Fitbit and Apple Watch), fertility apps, Internet of Medical Things devices such as continuous glucose monitors, smart thermometers and EKG monitors (e.g., AliveCOR), and consumer DNA tests to examine health predictors (e.g., 23andMe), with many additional applications existing or in development. Individuals are opting to use these technologies for a variety of reasons, including self- monitoring for health maintenance or improvement, the intention to link PGHD to physicians for continuity of care, or out of altruistic concern for public welfare (e.g., sharing health data for medical research). The COVID-19 pandemic has brought discussion about PGHD to the fore, as use of related health-tracking and monitoring devices proliferate (e.g., blue- tooth enabled pulse oximeters, disease exposure mobile apps), and corporations and governments seek to use such personal health information to monitor and control the spread of the disease in the pandemic.

Dominant narratives about health data analytics position these highly personal data as valuable resources that are poised to transform healthcare, stimulate innovation in medical research, and provide feedback, autonomy, and control related to personal health to individuals, i.e., the Quantified Self movement, with personal health behavior being shaped by health analytics to "nudge" individuals towards certain actions (Swan, 2013). However, as the variety and volume of PGHD grows exponentially, the accuracy and utility of the PGHD generated by such devises and apps, even for intended uses (such as self-monitoring health), are questionable. In the United States, many consumer PGHD technologies do not undergo Food and Drug Administration (FDA) or other regulatory scrutiny for efficacy or accuracy, as "the availability and development of the technologies has, in many instances, outpaced the publication of trials designed to evaluate health outcomes, usability, interoperability, and benefits and harms of these technologies" (Agency for Healthcare Research and Quality, 2019, p.1). Software as a medical device (SaMD), which is proliferating via mobile health apps, is ill-suited to the U.S. regulatory framework designed for hardware-based medical devices (Shuren, Patel & Gottlieb, 2018).

In addition to concerns about safety, accuracy, and efficacy of PGHD, the governance of these personal, even intimate, health data collected, stored, and shared by mobile devices is a growing area of concern (e.g., Montgomery, Chester & Kopp, 2018). Where and how PGHD are collected, aggregated, and maintained vary. The technology firms that provide PGHD mobile applications or monitoring devices often initially aggregate these data on their own IT infrastructure using cloud services, patients may store data locally on a personal device, or both may occur. Patients may sometimes share PGHD with clinicians, directly or via the technology vendor's or clinician's health data portal. Clinicians may, or may not, integrate these data into the patient's electronic health record (Abdolkhani, Gray, Borda & DeSouza, 2019; Genes et al., 2018). The extent to which PGHD (anonymized or not) are shared with other third parties, for instance, with employers who sponsor workplace wellness programs, insurers who include use of health monitoring apps in their insurance policies (Aetna, 2020; Bari & O'Neill, 2019), or retailers or marketing firms, is unknown. Thus, in addition to integration into clinical health settings, PGHD are undoubtedly included in growing stockpiles of personal health data that may be mined for a multitude of purposes by third parties, including their use in health systems analytics and in development of artificial intelligence/deep learning. Concerns about how these data could lead to unjust discrimination through differential pricing, increased health premiums, or denial of health insurance – or whether innovations from their use will enrich only the lives of a privileged few – are emerging (Winter & Davidson, 2019b).

Governing the vast array of personal health data to facilitate patient care, health research and health system innovation while preserving individual privacy and autonomy is challenging, and will be even more so with PGHD, which are the least regulated domain of digitalized personal health data in the U.S. (U.S. Department of Health and Human Services, 2018). In this research, we consider the question of *how PGHD regulations might be harmonized to better protect personal privacy while also fostering needed innovations of health services for societal benefit*. In doing so, we assess whether current approaches for regulating access to protected health information are sufficient to address the emerging domain of personally generated health data. We consider whether these regimes are adequate and synergetic to address emerging regulatory needs for PGHD and whether enhanced federal health privacy laws or stricter omnibus data protection laws more akin to GDPR (Davis, 2019) would address this new environment. Using a regulatory regime lens, we examine three overlapping regimes of health data governance and assess how these regimes address PGHD to identify areas of gaps and overlaps. We then outline three types of governance structures (organizational, regulatory, technological/algorithmic) that could help in harmonizing existing regimes and addressing gaps. This policy analysis contributes to thinking about balancing tradeoffs between personal privacy and societal innovation in health that PGHD might enable. Finally, we consider how integrating analysis of regulatory regimes and data governance may provide insights into other areas with similar challenges.

2. Regulatory regimes and data governance

Regime theory (Krasner, 1982) conceptualizes regimes as "sets of implicit or explicit principles, norms, rules, and decision-making procedures around which actors' expectations converge" (Krasner, 1982, p. 186). While regime theory was originally intended to theorize international relations, it sets a conceptual foundation for broader discussion of regulatory governance (Maggetti & Ewert, 2018). For example, Levi-Faur (2011) argues that "the notion of a regulatory regime encompasses the norms, the mechanisms of decision making, and the network of actors that are involved in regulation" (2010, p. 20). The concept has evolved to include a broad scope, including obligations and prohibitions with compliance and enforcement mechanisms, incentives, and non-state actors employing regulation to self-regulate or regulate others (Maggetti & Ewert, 2018). Thus, we consider regulatory regimes as various systems of regulation and enforcement mechanisms created to address specific activities. Following this line of thought, May and Jochim (2013) conceptualize policy regimes as "the governing arrangements for addressing policy problems" (p. 428). Here, the policy regime construct is both descriptive and analytic. Descriptively, the policy regime lens provides a conceptual map for identifying and analyzing governing arrangements, particularly distributed issues that do not yet have comprehensive efforts to address them.

In the digital economy, data generated and consumed by information systems and mobile technologies are key resources for economic and social innovation. Data governance regimes relate to the management of data, with particular legal specifications focusing on protecting personal data. Data governance regimes typically focus on privacy and security of data, particularly personally identifiable data. These regimes vary substantially across and within countries. Yeh (2018), for example, compared the data privacy regimes of the European Union and United States and argued that data brokers (i.e., third-party aggregators and sellers of data)

are best regulated under a comprehensive legal framework such as the GDPR. At the same time, movements towards "data localization" have also presented challenges to innovations that are made possible by utilizing growing data resources, by setting requirements for subjects' data to be stored and processed within a nation, crossing borders only when data transfer and use requirements are met (Taylor, 2020).

Regulatory regimes directed towards governing the movement and uses of personal health data – *data protection regimes* – vary greatly across national contexts. Some nations' legal frameworks have sections specific to health data. For example, the GDPR (and the laws of many countries following this lead) has stricter provisions for data concerning health, including genetic and biometric data, than for other types of personally identifiable data. Other jurisdictions, such as the United States, have sector-specific health-related laws (e.g., the Health Information Portability and Accountability Act and the Genetic Information Nondiscrimination Act). The existing regulatory frameworks for personal data protection provide alternative broad and flexible rules for accessing PGHD for scientific research and innovation. Recognizing potential threats to personal health data privacy, Marelli, Testa, and Van Hoyweghen (2021) call for the "re-purposing" of health big data governance for research to ensure necessary safeguards for individuals and society are articulated and maintained. As we explore in this paper, the more recent entry of Big Tech into business activities involving access to health data, and PGHD in particular, brings more challenges to the efficacy of existing data protection regimes.

Tackling these complex challenges to negotiate and maintain a balance between innovative opportunities and privacy protection for PGHD requires us to first identify existing regimes and assess where they overlap or where lacunae may exist. Regulatory regimes vary across countries and regions and, in some cases, within particular domains. In this paper, we will focus on the national regulatory regimes in the United States, beginning with an analysis of the current state of regulations related to PGHD. This analysis provides the groundwork for examining regulation surrounding PGHD more generally, which will will address in more detail in the discussion.

3. Examining three data protection regimes relevant to PGHD

Most discussions about regulating personal health information center on the important concerns of maintaining privacy and security, while also promoting data interoperability and sharing data, in order to facilitate patient care and support medical research (Meystre, Lovis, Bürkle, Tognola, Budrionis, & Lehmann, 2017; Rosenbaum, 2010). These discussions occur primarily among health-sector policy makers, organizational leaders, and researchers and focus on regulated clinical data. In the United States, substantive discourse on personally generated health data that considers how these data differ from clinical data, and thus how PGHD should be governed, the types of value potentially afforded by these data, and the values and interests that shape their governance, has not yet emerged. This is due in part to legal and regulatory assumptions that non-clinical health data collected by an organization are de facto the assets (IP) of that firm, to govern and to utilize as they deem appropriate within applicable regulatory regimes such as consumer protection laws. Questions about what data protection regimes are relevant to PGHD governance, and whether individually or collectively these regulatory regimes provide adequate societal oversight of PGHD remain largely unexplored. To begin this

discussion, we review three existing data protection regimes applicable to PGHD in the U.S. and highlight data governance issues within each.

3.1 U.S. Federal and state health-sector laws

When considering the U.S. approach to privacy and data protection, it is important to acknowledge that it is an ambiguous concept with no consensus definition, and it has evolved over time (Acquisti, Brandimarte, and Loewenstein, 2015; Igo, 2018; Solove, 2010). The U.S. has not had a major federal privacy reform in several decades, relying instead on a loose collection of sector-specific laws. Additionally, states have their own laws, and where these offer more protection to the individual, they often take precedence over federal law.

In the United States healthcare sector, the Health Insurance Portability and Accountability Act of 1996 (HIPAA) mandated the protection and confidential handling of certain health information (protected health data) and addressed digital data developments evident at that time. HIPAA required the U.S. Department of Health and Human Services (HHS) to develop regulations and accountability mechanisms. HHS published two related rules. The Standards for Privacy of Individually Identifiable Health Information ("Privacy Rule"; U.S. Department of Health and Human Services, 2013) set national standards for data protection of some health information. The Security Standards for the Protection of Electronic Protected Health Information ("Security Rule") set national security standards for certain data in digital form. Within HHS, the Office for Civil Rights (OCR) is tasked to enforce both rules. Additional provisions to the HIPAA privacy and security regulations were introduced as part of the Health Information Technology for Economic and Clinical Health Act (HITECH) in 2009, and in 2013 the HIPAA Omnibus Rule updated these privacy and security regulations. (U.S. Department of Health and Human Services, nd). Other narrow slices of health data are also protected under the Fair Credit Reporting Act (limiting data use and sharing of medical information in consumer financial reports), the Genetic Information Nondiscrimination Act of 2008 (prohibiting discrimination in health insurance and employment based on genetic information), Family Educational Rights and Privacy Act (limited sharing of medical information found in school records), and a few other laws related to personal health information held by the government. However, overall HIPAA is viewed as the dominant regulation.

HIPAA applies only to specified covered entities (e.g., physicians, hospitals, health insurers) and their designated business associates. As a result, the health data regulated via HIPAA originate within, or are related to, clinical healthcare settings. For instance, physicians' notes in an electronic health record (EHR), prescriptions or laboratory orders exchanged with pharmacies or labs, insurance claims for reimbursement, health information exchange between clinicians, data in EHR-tethered patient portals, and so on, fall under the regulatory regime for health data and specifically HIPAA. Under HIPAA, protected health information that is personally identifiable must be secured and patient privacy protected (U.S. Department of Health and Human Services, 2013). Sharing such data is permissible for clinical care or insurance reimbursement, with patients' consent[1], within these limits. Further sharing of HIPAA-protected patient data, for instance, for health or health policy research, generally requires deidentification of the data. However, to ease administrative burdens and increase the value of health data

[1] Consent is generally required in order to activate insurance.

through new linkages and analyses, there are several potential changes to the HIPAA Rules in progress. This includes making "sharing PHI with other providers mandatory rather than permissible" to address organizational reluctance to sharing clinical data assets (HIPAA Journal, 2021, para. 7).

In the decades since the law was created, consumers' widespread use of the Internet for health information, proliferation of Internet of Things devices including mobile devices and health monitors, and advanced user profiling and data analytics software that can correlate non-health data to potential health conditions, have led to an environment where a great deal of health-related data not covered by HIPAA are collected. Of note, HIPAA applies to PGHD collected by patients only in certain limited circumstances, for instance if a physician prescribes to a patient a specific health monitoring device provided by a business associate, which then manages data on behalf of the clinician (HIPAA Journal, 2021). In some instances, the FDA has declared that wearable health devices should be classified as medical devices (e.g., continuous glucose monitors or pacemakers) and are therefore subject to HIPAA.

Since most PGHD are collected by patients themselves using mobile apps that are not FDA approved and are provided by technology firms (rather than healthcare providers), and these data are typically collected outside clinical encounters, PGHD fall outside of HIPAA (U.S. Department of Health and Human Services, 2016, 2018). Instead, PGHD are maintained primarily on the information technology (IT) infrastructure of vendors who provide the mobile apps and devices, and data are governed under the IT firm's own privacy policies and within the firm's intellectual property rights in data. Patients can view and may be able to download their data to devices they control, and they may be able share data with clinicians, thus replicating data in multiple locations. If PGHD are integrated into clinical EHR systems managed by covered entities, only that instance of integrated data would then fall under HIPAA regulations. Conversely, if a patient chooses to download clinical data from the clinical provider's EHR portal, for instance, into a personal health record app on a mobile phone, those clinical data are no longer covered by HIPAA. (The app vendor may voluntarily choose to comply with HIPAA as part of the firm's privacy policy.)

The ramifications of most PGHD not falling under HIPAA or other health data regulations are just now emerging. Most consumers are likely to assume that PGHD are protected in a similar manner as their clinical health data, and they may not appreciate that by transferring clinical data to PGHD apps and platforms they are foregoing HIPAA oversight of the data. Similarly, the extent to which intimate PGHD can be repurposed/resold among commercial firms is typically buried in the PGHD app provider's privacy policies; full disclosure might concern many consumers and limit their willingness to use these devices. For instance, in January 2021, menstruation and ovulation tracking app creator Flo Health entered into a settlement with the Federal Trade Commission due to its improper disclosure of data with third parties such as Google and Facebook (Federal Trade Commission, 2021).

3.2 PGHD reuse for research and innovation

The promises of transformational improvements in health care treatments and personalized healthcare, as well as in the cost, quality, and access to healthcare services, depend on researchers' access to health data, particularly individual-level health data, and on technology

innovations associated with data analytics and artificial intelligence. Researchers in many domains (health, life sciences, technology development, etc.) are anxious to obtain access to protected health data (Rosenbaum, 2010), and increasingly, to PGHD that might provide a holistic view of patient health practices and health outcomes. In the U.S., any health data used for research may be governed by general research regulations for ethical treatment of research subjects (U.S. Department of Health and Human Services, 2016). The "Common Rule" 45 CFR 46 specifies that research conducted by federal agencies or institutions receiving federal funding be reviewed and approved by Institutional Review Boards (IRBs), which assess the scope, research value, and regulatory compliance of human subjects research. Due to the broad reach of federal funding, the vast majority of U.S. university-affiliated and health system-affiliated researchers fall within this scope.

Research that involves health data is often subject to additional regulations. In particular, HIPAA privacy rules apply to all protected health data (discussed above), along with specific privacy protections such as 42 CFR Part 2 for data on substance abuse. The U.S. federal Center for Medicare and Medicaid Services (CMS), state government agencies, and other health data holders may also impose additional regulations on the access to and uses of health data under their governance (e.g., Medicare claims data from HHS/ CMS). Government policies and practices for granting research access to protected health data are developed through administrative rulemaking and are published publicly. The specific regulations that researchers are subject to when accessing individual-level protected health data depend on whether the data are anonymized or are patient identified (or identifiable). HIPAA specifies 18 data fields (name, DOB, medical record number, phone number, dates medical services were provided, and so on) that must be removed or masked to be fully "deidentified". Once data in a dataset are deidentified, data might be provided as a public use file (PUF) that is no longer subject to HIPAA, per se. Requests for fully or partially identifiable data are subjected to some form of IRB or other approval authority, and typically are also subject to data use agreements with the data governing body, which specify how data can be used (in addition to HIPAA rules) and when data must be disposed of. For instance, data use agreements might prohibit attempts to reidentify data, to link data across datasets, or to resell data to organizations that might do so.

In contrast to these formal procedures and policies for research access to protected health information, regulations for how and when PGHD are being (or should be) used in research are much less clear. If researchers subject to the Common Rule (but are not HIPAA covered entities) receive identifiable PGHD directly from research subjects, for instance, then their data use will be subject to 45 CFR 46, but not to HIPAA, since these PGHD are not covered by HIPAA. If researchers obtain PGHD directly from the technology vendor versus from the research subjects, the application of 45 CFR 46 is less clear. Many such arrangements are underway, as PGHD-app firms may work with researchers to enhance their products' offerings or in collaborative research endeavors. Individually identifiable data that are derived from pre-existing files, without the researcher's interaction with the subjects of that data, can fall outside the scope of "human subjects research" and thus outside IRB review. Whether use of PGHD would fall under the Common Rule in these circumstances would depend on the individuals' expectations for privacy on the technology vendor's IT platform. Expectations of privacy then depend on the vendor's stated privacy policy. Importantly, PGHD technology vendors selling products and services directly to consumers (that is, not as a business associate for a covered entity) likely are not

subject to either 45 CFR 46 or HIPAA regulation. Thus, in many instances, how PGHD are used in research will depend on the self-regulation of technology vendors in this market.

3.3 Consumer privacy and technology vendor self-regulation

Technology companies are racing to enter the lucrative healthcare sector, as providers of secure cloud services and AI-enabled analytics to health care organizations, developers of a variety of consumer health apps, and in medical research. The central role of data to the economy is growing (Krämer, Whalley, & Batura, 2019), and health data is among the most lucrative types of data (Humer & Finkle, 2014). While early forays by companies like Google and Microsoft into personal health data applications failed (Davidson, Østerlund, & Flaherty, 2015), over the past several years these companies have refocused their health initiatives while being joined by Apple, IBM, Amazon, Facebook, and many other less well-known high-tech firms. For instance, Apple's healthcare efforts now center on three main areas: 1) consumer products based on wearable and smartphone apps; 2) health research in partnership with health organizations enabled via Apple's open-source framework, ResearchKit; and 3) secure health records via its Health app, which was launched in 2014 and consolidates data from an individual's phone, watch, and health-related third-party apps (Dyrda, 2019). By 2020, Google Health had grown to over 500 employees, with many allocated from other parts of Alphabet (Farr, 2020), including its health AI venture, Google DeepMind Health. Google also bought leading fitness tracker Fitbit in late 2019.

The regulatory regime that governs these Big Tech firms in regards to consumer privacy generally differs substantively from HIPAA-related regulations that oversee protected health data. Between 1997 and 2007, there was an array of efforts to create industry- or government-supported self-regulatory guidelines for handling of personal data, but "the majority of the industry self-regulatory programs that were initiated failed in one or more substantive ways, and many disappeared entirely" (Gellman & Dixon, 2011, p. 2). Although federal law does not require companies to have, share, and enforce a privacy policy if they gather and use personal data, under the Federal Trade Commission Act, a firm may do business in jurisdictions that require one (e.g., as dictated by individual state laws or the laws of other countries). Where privacy policies are included, a firm's lack of specific language and provisions related to privacy would defy the law and may lead to enforcement as a deceptive practice (Federal Trade Commission, 2018). Aside from this, there is little control over policy content. A firm's privacy policy may make blanket statements that personal data may be shared with third parties to cover myriad future opportunities to monetize consumer data. The FTC is authorized to regulate privacy practices under the unfair and deceptive practices standard, though audit and enforcement are lax, and when a firm is absorbed by another via sale, previous privacy policies may be nullified. For instance, when popular fitness tracking app Fitbit was purchased by Google in 2019, there was concern about how Google would use the PGHD accumulated via Fitbit, since "current laws and regulations do little to hold Google and other companies to their promise" (Frazee, 2019, para. 4) for governance of consumer data.

As these Big Tech firms compete to gather PGHD and other data that infers health information, their efforts often bypass existing health regulation in the U.S. Most PGHD devices are not classified as medical devices by the Food and Drug Administration, and the self-reported

or inferred health data[2] collected and aggregated by Big Tech also does not fall under HIPAA. Unless a device is classified by the FTC as a medical device, HIPAA does not automatically apply, and consumers (patients) must rely on the good will and good intentions of the vendors' self-regulation and policies. When consumers transfer protected health data from a HIPAA-regulated context (such as their clinician's patient portal) to the technology firm's PGHD infrastructure (such as Apple's Health Kit), these instances of their PHI data are no longer HIPAA-protected. Some Big Tech firms also have access to detailed information about an individual's search history, social networks, interests, and consumer consumption patterns that are not explicitly health-related but can reveal aspects of a person's health. For example, information about a person's location over time, their associates, or demographic and community data may be used to infer health status. Essentially, "all your data is health data" through a process called digital phenotyping, which refers to "taking information from our digital behaviors — on websites, via our phones — and using it to gain insight into potential health issues" (Perez-Pozuelo et al., 2021; Warzel, 2019, para. 4).

3.4. Gaps and overlaps in PGHD data protection regimes

Figure 1 depicts key contexts in which PDHD arise and summarizes our analysis of how data protection regimes in the U.S. are applicable to these various domains. The three regulatory regimes discussed here each address some aspects of personal health data privacy and protection, and each provide some flexibility for data sharing and reuse to enhance health care delivery, health system improvement and health-related innovation. Critical analysis of health data protection regulations within each regime, and importantly, across the three regimes, helps to identify lacunae in current privacy and autonomy approaches for governing the evolving PGHD landscape.

[2] Inferred health or "digital phenotyping" uses data collected from web searches, purchase transactions, or IoT-enables non-medical devices (e.g., using data collected about how one uses brakes while driving to infer cognitive or motor difficulties) to make inferences about a subject's health.

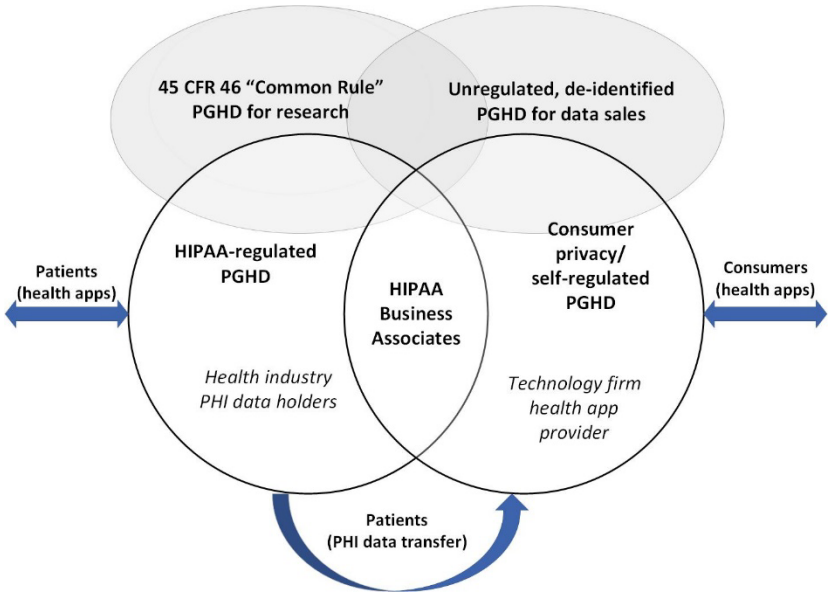

Figure 1. Domains of PGHD and Related Data Protection Regimes

Utilizing PGHD on individuals' daily health-related behaviors and experiences increases the potential for innovation if PGHD are integrated with clinical health data (e.g., Jim et al., 2020; Melstrom, Rodin, Rossi, Fu, Fong, & Sun, 2021; Ravuri, Kannan, Tso, & Amatriain, 2018). However, given uncertainties about regulatory oversight and potential liability, clinical providers who would find patients' PGHD useful to enhance patient care might hesitate to allow such data into their HIPAA-compliant EHRs or to release EHR-data for combination with PGHD and reuse elsewhere. For instance, the regulatory outcomes of combining PGHD with HIPAA-protected clinical data depend on which entities control the combined dataset. As noted above, integrating PGHD into a HIPAA-compliant data repository controlled by a HIPAA-covered entity or business associate (e.g., hospital, physician, contracted IT vendor) affords that instance of the PGHD additional privacy protections these data are not accorded elsewhere, whereas transferring clinical data to a PGHD repository controlled by a firm that is not a HIPAA-covered entity substitutes the technology vendor's consumer privacy policy for HIPAA protections. At best, this is confusing for health providers and for consumers, who may direct health data transfers without reading or comprehending detailed consumer privacy statements. At worst, consumers risk abdicating rights to control both their (previously) protected health data and the PGHD they generate through the vendor's mobile devices, applications, and services from uses that are not in the consumers' best interest, such as monetized health data used in economically discriminatory ways (Tanner, 2017).

The availability of and access to large-scale databases of health data for research purposes, when combined with rapid developments in information technologies such as artificial intelligence and machine learning, hold the promise of transformative innovation in healthcare for individuals and for healthcare systems (Flores et al., 2013; Miotto et al., 2017). In the regimes of national or regional privacy laws and data reuse for research and innovation, regulatory oversight is well outlined and discussed for protected (i.e., clinical) health information (Meystre et al., 2017). However, the growth of large health data sets derived from PGHD challenge taken-for-granted practices and approaches, since much PGHD fall under industry self-

regulations that "challeng[e] the regulatory requirements and research ethics norms that govern traditional approaches to human subject protections" (Rothstein et al., 2020, p. 197). As van den Broek & van Veenstra (2018) argue, the principles of data maximization and open-ended purpose at the heart of big data and machine learning innovation are at odds with data protection regulation – "it is impossible to set up a pure Market arrangement of governance for settings in which personal data are used" (p. 336). Moreover, consistent oversight of PGHD used in institution-based research will be difficult to enact and monitor, because researchers are spread across many institutions with varied resources for IRBs and IRBs may be staffed by volunteers with limited expertise in health data regulation. Regulatory guidelines from the 45 CFR 46 Common Rule are abstract and open to interpretation, for instance, about individuals' privacy expectations on health-related social media sites. Researchers in private, for-profit firms are not covered by the Common Rule, so that any IRB-like oversight will be at the discretion of the firm's management. Over-regulation that inhibits beneficial research as well as under-regulation of privacy-challenging research are more likely to occur in these circumstances.

Gaps in health data protection regimes are becoming more apparent as Big Tech firms move into the healthcare sector. Scholars, policy makers and privacy advocates have raised significant concerns about Big Tech's stockpiling and governance of personal data collected through day-to-day social and economic activities (e.g., Schneier, 2015). Given the sensitivity of health data, suspicions about Big Tech's intentions for PGHD and other health data come into conflict with concurrent expectations for health care innovations through advanced analytics and AI technologies these firms might develop and provide. Particularly concerning is the lack of transparency in data use agreements between HIPAA-regulated entities and Big Tech and the data use policies that Big Tech will craft and adhere to related to reuse of personal health data (including PGHD) across an array of AI development projects.

For instance, in November of 2019, a whistleblower working at Google (Anonymous, 2019; Pilkington, 2019) revealed that Google Health was partnering with a large health system nonprofit, Ascension Health, to analyze the PHI of approximately 50 million people across 21 U.S. states. Called Project Nightingale, this partnership raised immediate public, health industry, and governmental concern (Copeland, 2019; Copeland & Needleman, 2019; Loveland, 2020; Pifer, 2019; Price, 2020), and the Department of Health and Human Services Office of Human Rights opened an investigation to determine whether HIPAA violations had occurred. Some members of Congress also expressed grave concerns about the partnership and called for both parties to disclose their data use agreement (United States Senate, 2020). It appeared that both parties were technically compliant with HIPAA, as Google was designated as a business associate providing the IT infrastructure and records interface for Ascension. However, questions about Google's plans for using this protected health data to develop AI capabilities remain. Google has access to vast storehouses of PGHD from Fitbit, as well as consumer search behavior and mobile data. These personally identifiable data could be integrated with these protected health data in AI projects. The potential for dramatic innovation but also devasting losses of personal privacy and autonomy are thus evident in the Nightingale case. This case involving Google is not unique, nor is Google alone in its health industry aspirations, as other Big Tech firms are undertaking similar ventures.

The appeal of AI for health system innovation poses challenges across these data protection regimes because of the scale and scope of health data needed for deep learning algorithm training and the algorithms' opacity, which complicates compliance monitoring. As large amounts of data are necessary to train deep learning models (Xiao, Choi & Sun, 2018), these ventures will rely on access to a variety of PHI, PGHD, and other health-related data that are initially collected for other purposes and reused in new settings by new organizational actors operating under different regulatory regimes (Winter & Davidson, 2019b). Anonymizing health data can be helpful to reduce regulatory complexity while maintaining some personal privacy protection, but this is not a simple, comprehensive solution (Langarizadeh, Orooji, Sheikhtaheri & Hayn, 2018). Well-intentioned attempts to anonymize PGHD would likely fail to achieve that goal unconditionally, since PGHD contain many subtle identifying attributes even beyond what is specified by HIPAA criteria for anonymizing PHI. Moreover, the research and innovation value of health data are much higher when data are linked across data sets, and linkage requires some form of unique identifier across data collection contexts (e.g., location, home or school, work, various clinical setting), making re-identification technically feasible.

These developments highlight the pressing need to examine the degree to which existing data protection regimes can adequately address evolving PGHD (and health data generally) governance challenges and to highlight and address ineffective and overlapping approaches that frustrate potential uses of PGHD for socially desirable innovation (Mandl, Mandel, & Kohane, 2015). Our discussion above is a first step to stimulate this debate. Also critical is to develop new regulatory and technical approaches to enhance or replace inadequate regimes. We turn now to our analysis of recent developments that may contribute to these goals.

4. Towards harmonizing regulatory regimes for PGHD governance

Effective data protection regulation requires surfacing potential and actual conflicts between diverse stakeholders, assessing normative goals across sectors, and determining mechanisms to detect and prevent harm and enforce compliance/accountability. Regulatory strategies must then balance interests of multiple stakeholders with different expectations, accountabilities, and ethical codes arising under different regulatory regimes. We refer to these processes as harmonization of regulatory regimes. The lack of harmonization between regulatory regimes may render ineffective existing policies and interventions to preserve privacy and individuals' control over data about themselves, while also hindering beneficial innovations that depend on health data sharing and reuse. There are a number of potential benefits to regulatory harmonization, such as opening and/or stabilizing markets for products and services and facilitating cooperation across regional boundaries, Here we focus in detail on how effective national harmonization can identify overlaps and gaps in regulations and enforcement mechanisms in how PGHD are addressed, so as to (i) *increase consistency* of regulatory goals, rules, and practices across regulatory regimes, (ii) *minimize redundancy* in regulatory oversight in different contexts that adds unnecessary delays and costs, (iii) *reduce complexity* of regulatory mechanisms to encourage compliance, and (iv) *enable audits* to ensure compliance and identify lack of compliance. Such advances would better protect personal health data, reduce regulatory overlap and policy uncertainty, and stimulate innovation and investment (Faitelson, 2019).

As we detail above, developments with PGHD are challenging existing data protection regimes for health data governance. In the past, health data governance was largely handled

within health industry organizations with well-defined regulatory responsibilities. With the rise of PGHD, along with the many partnerships among health industry organizations and Big Tech, health data governance responsibilities are now shared among numerous types of institutions and firms, which operate in different regulatory regimes for data governance and with different norms and practices for protecting privacy and for sharing data (cf. Taylor, 2017). We suggest that three overlapping categories of data governance structures can contribute to regulatory harmonization goals and processes: organizational, regulatory, and technological/algorithmic (see Figure 2). We provide examples of emerging governance structures for each, consider how these structures may work synergistically, and consider how developing and adopting such structures might contribute to harmonizing regulatory regimes for PGHD governance while limiting the friction and economic costs of regulation that may hinder innovation.

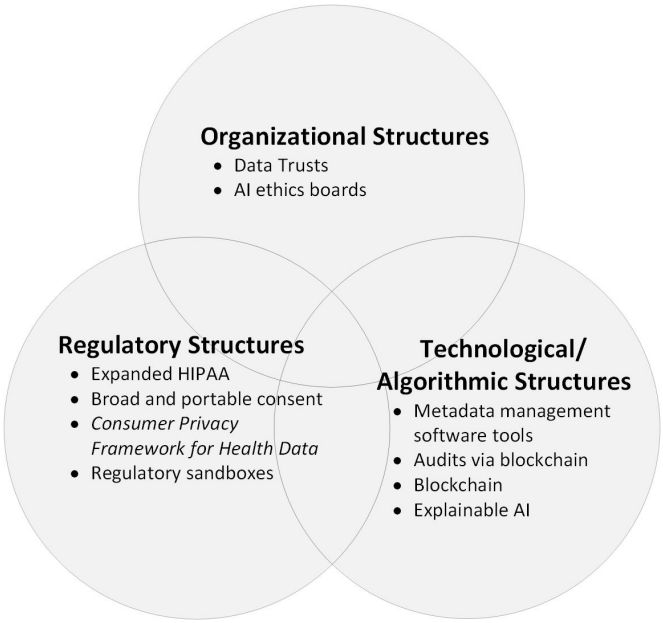

Figure 2. Categories of data governance structures

We argue that each approach, for which we outline examples below, can contribute to harmonization in some ways, but that multiple approaches are needed to achieve fuller coverage and to identify any conflicts, overlaps, or missing areas, including where existing provisions may be too ambiguous. Table 1 summarizes our arguments about beneficial harmonization for the types of governance structures we detail below.

Emerging data governance structures	*Increasing consistency*	*Minimizing redundancy*	*Reducing complexity*	*Enabling audits*
Organizational structures				
Data trusts	■		■	
AI ethics boards				■
Regulatory structures				
Broad and portable consent	■	■	■	
Consumer privacy framework for health data	■	■		
Regulatory sand boxes			■	■
Technological/Algorithmic Structures				
Metadata management software tools			■	■

Audits via blockchain				
Blockchain-enabled secure storage and sharing of health data				
Explainable AI				

Table 1. Harmonization of emerging data governance structures

4.1 Organizational structures for PGHD governance

A variety of organizational data governance structures are emerging to address the challenges of digitized health information. One example is data trusts (Delacroix & Lawrence, 2019; O'hara, 2019), agreements where data subjects agree to pool their rights to achieve common goals that require the use of the pooled data. Data trusts could increase consistency and reduce complexity of data governance and regulation of the data pool, for data subjects, providers, and users. For example, the social health site PatientsLikeMe aggregates and shares self-reported health data among community members seeking insight into their health conditions. Data are also made available to external researchers, with revenues generated from data sales used to support the community (Tempini, 2017). A fiduciary, or data trustee, makes agreements for use of the data on behalf of the data trust contributors. "This 'bottom-up' data trust model is resolutely complementary to top-down, regulatory constraints (including those of the GDPR)" (Delacroix & Lawrence, 2019, p. 243).

Data trusts are thus a potential organizational solution to some multi-stakeholder problems by holding user data and ensuring acceptable use on behalf of affected constituents. Harmonizing regulatory oversight through a PGHD trust would require the data trust participants to understand the external regulatory regimes that apply to PGHD and then to adopt and adapt data governance policies acceptable to the participants. For instance, in a consumer health community such as PatientsLikeMe, HIPAA and 45 CFR 46 regulations may not automatically apply to a firm seeking to purchase the community's data. Instead, the data trust agreement can specify whether PGHD can be sold, to whom, for what range of purposes (e.g., commercial innovation, research), and under which privacy regulation rules. Data trusts could foreseeably enable individuals to benefit financially from their PGHD (and other health data) contributions, rather than all value from data sales accruing to a technology vendor, as is currently the case with consumer PGHD devices.

Nonetheless, it will be difficult for a data trust to ensure that data use agreements are complied with or to take preventative actions against misuse, if data are replicated for use outside of the IT infrastructure controlled by the trust. Importantly, data trust arrangements do not negate the need to negotiate, document and effectively communicate data use policies both to those who contribute their data and to potential users of the data. For instance, data contributors would need to understand how their data would be used and if they could later revoke their consent, even after data have been provided to researchers or other users. Data trusts such as PatientsLikeMe have wrestled with this issue, as data-intensive infrastructures evolve and data are reused by different types of stakeholders (Tempini, 2017).

A second example of an emerging organizational governance structure is the AI ethics board. Some firms have established these boards as internal groups that provide governance, recommendations, and oversight regarding ethical AI development in the organization. Oversight of data uses (in this instance, PGHD) in AI/ML is an essential task of such boards. In some

cases, organizations have also tapped external review committees to review their work and provide assurance to the public and regulators that the company is behaving ethically. Doing so effectively would be a step towards greater transparency on health data generally as well as PGHD acquisition and use in AI initiatives. For instance, after widespread public outcry about its data sharing relationship with the UK National Health System in 2016, Alphabet's (Google's) DeepMind Health created two entities: 1) an external Independent Review Panel tasked to review DeepMind Health's activities and issue an annual report to the public; and 2) the DeepMind Ethics & Society Fellows to research issues such as privacy and fairness (Winter & Davidson, 2019a). However, rather than engendering greater trust, this latter group was perceived as a "gigantic conflict of interest" due to "a commercial AI giant researching the ethics of its own technology's societal impacts" (Lomas, 2017, para. 5). During the merger of DeepMind Health and Google in 2019, Google also created the Advanced Technology External Advisory Council (ATEAC) to guide ethical development of new technologies. ATEAC was shut down within a week due to controversial choices of members and principles that were hard to enforce, "not least because the enforcement mechanisms for violating the principles aren't well-defined, and, in the end, the entire enterprise remains a self-regulatory endeavor" (Johnson & Lichfield, 2019, para. 7).[3]

Despite such difficulties, AI/data governance ethics boards have important potential roles in auditing compliance to data protection regimes, including the firm's own publicly avowed policies, and in this way to hold firms accountable and bring public interests and visibility to health data and PGHD use practices. Deciding who should appoint such boards, who the boards are accountable to, and what latitude boards have will be key to their effectiveness. Whether such boards are established and how they operate thus might draw on a regulatory regime similar to the Common Rule and IRBs, rather than remaining in the self-regulatory regime that Big Tech currently enjoys. Feijóo et al. (2020) argue for a broader approach, a new technology diplomacy focused on a multi-stakeholder and multi-layered process to engage government, corporations, non-profits and research organizations towards the global alignment of AI governance and policy harnesses innovation while seeking to diminish negative impacts.

4.2 Regulatory structures for PGHD governance

Regulatory data governance structures are intended to protect health data privacy and security while not unduly posing stumbling blocks for innovation. In the U.S., proposed regulatory governance structures include an update to HIPAA focusing on expansion of covered entities, informed consent, and research use. Such revisions would bring the three regulatory regimes discussed earlier into closer alignment by increasing consistency across the regimes. A key limit of HIPAA today is that it only applies to "business associates" of the originating healthcare actor. For instance, in Google Health's partnership with Ascension, the former is

[3] Concerns about such efforts by Big Tech to self-manage these issues have recently reignited with the controversial exit of Google AI ethics researcher Dr. Timnit Gebru, who had co-authored a paper under review that emphasized several risks of large language models, which are central to Google's research and innovation (Hao, 2020). The subsequent firing of Margaret Mitchel, a founder of the Ethical AI team who had publicly defended Gebru (Metz, 2021), has led to widespread criticism of Google and questioning of its AI ventures.

classified as a business associate to the latter for HIPAA purposes, but when Google Fitbit sells a fitness tracking device and an individual uses it, data generated by the device are not covered by HIPAA. Expanding HIPAA to cover at least some additional domains of PGHD (Bari & O'Neill, 2019) would narrow the gap between HIPAA-regulated PHI governance and firm self-regulation of PGHD.

Expansive uses of PGHD and other health data for health research are critical to innovation, but such uses can challenge informed consent policies, because it is not possible to forecast all possible uses of health data or the risks posed at the time data is collected (Sharon, 2016). Moreover, keeping track of where and how health data are used from the point of origin through the many instances of use will require organizations that develop or utilize PGHD stores to greatly enhanced documentation and tracking capabilities to enable renewed consent and audits for compliance. New models of broad and portable consent for large-scale research projects using health data are being explored to address "the unavoidable question of who stands to benefit and in which way from research results" (Sharon, 2016, p. 568). Effective design and implementation of broad and portable consent mechanism could help align clinical and research data regulations to increase consistency, minimize redundancy, and reduce complexity of compliance across data protection regimes. An even broader expansion of HIPAA to incorporate PGHD that are accumulated by Big Tech firms, and later sold to researchers or commercial ventures, would further align commercial self-regulation of PGHD with clinical and research data protection regulations, and thus provide a more uniform, consistent regulatory approach to user consent with less redundancy across these three regulatory regimes.

The U.S. might also move from dependence on sectoral data regulation towards a new omnibus data protection law. The EU's General Data Protection Regulation (GDPR) has stimulated a global discussion about data privacy and protection and many jurisdictions are moving towards GDPR-compatible regimes. The GDPR, which was enforced in May 2018, and the passage of California's Consumer Privacy Act (CCPA) in June of 2018 (enforced in 2020), have renewed discussion about an omnibus U.S. federal privacy law. Big Tech firms, under increasing federal and international scrutiny, have also called for legal reform: "In November 2018, in response to a call for comments on a federal privacy law by the NTIA, numerous companies responded by stating that they were now in favor of a federal privacy law" (Solove, 2019, para. 3). However, at this time, key uncertainties will likely shape the outcome of U.S. federal privacy law for the near future: (i) The resolution of tensions between state and federal privacy laws given the current Congressional deadlock and new administration; (ii) the possibility that stricter state laws may be preempted by a less stringent federal general privacy law; (iii) the fallout of the Privacy Shield decisions related to U.S.-European personal data sharing and other GDPR enforcement, and (iv) pending antitrust actions towards large tech companies (e.g., Federal Trade Commission v. Facebook, Inc., 2020).

As federal privacy regulation reform is discussed and negotiated, the eHealth Initiative & Foundation (eHI) and the Center for Democracy and Technology (CDT) have sought to address regulatory gaps by working with industry leaders, civil rights groups, and government to create a proposed Consumer Privacy Framework for Health Data (eHealth Initiative and Foundation & the Center for Democracy and Technology, 2021). This self-regulatory model consists of a set of standards targeting non-HIPAA-covered personal health data and providing a set of data access,

use, and disclosure limitations for companies that collect or handle personal health information outside of HIPAA. The goal is to create a multi-stakeholder set of standards that will benefit consumers, corporations, and government and serve as a bridge towards future data protection regulation. Accountability mechanisms include annual assessments and audits of a random sample of members, suspension or dismissal from the program, and funneling non-compliant corporations to the attention of the FTC or state Attorneys General.

A third approach is the creation of regulatory sandboxes, that is, supervised environments managed by a data protection authority to pilot innovative products or business models in a live market with real consumers. Each participant has clear goals (e.g., improved health outcomes) and tests on a small scale with limited time and consumers (Centre for Information Policy Leadership, 2019). With regards to data protection, a regulatory sandbox "can simultaneously address two inevitable uncertainties—the uncertainties of innovation ('what is this going to deliver?') and the uncertainties of principles-based regulation ('will this processing be fair?')" (Centre for Information Policy Leadership, 2019, p. 5). This allows companies to test out new ideas without concern for regulatory challenge and enforcement, and individuals benefit from more scrutinized and customized privacy safeguards. Such approaches are promising to avoid overregulation as well. Regulations are not frictionless; they can stifle innovations in health care research, services, and products. Thus, evidence-based regulatory evaluations such as through a sandbox approach can help determine whether a new regulation is meeting its goals of reducing complexity while also enabling auditability with acceptable socio-economic costs and risks.

4.3 Technological/algorithmic structures for data governance

Technological/algorithmic governance structures are also emerging to address the challenges of digitized health information and growth in PGHD. Data use and movement often lacks transparency to data subjects or regulators, and thus noncompliance and associated harms are hard to detect. Metadata management software tools exist today that can identify data types (such as PHI) in data warehouses, reverse engineer data structures and trace data transformations across computerized systems to facilitate analysis of data lineage. If such metadata incorporates regulatory compliance rules, then systems developers can avoid building non-compliant systems, and auditors can more readily assess and monitor compliance. Acquiring, implementing, and maintaining metadata management software will require significant resources, and not all PGHD firms will voluntarily apply these practices without regulatory compliance incentives.

Concern about transparency in data use agreements has led to a movement towards fairness, accountability, and transparency (FAccT) in algorithms. One example of this is establishing audits via blockchain, proposed as a means to ensure repudiation of transactions does not happen and that transactions cannot be altered at a later date. For instance, as a result of regulatory censure and a loss of public trust over the DeepMind Health-NHS partnership, DeepMind announced it was using an automated audit of health data accessed using a new blockchain-like technology called the "Verifiable Data Audit" (DeepMind 2017; Hern 2017). This new governance structure purportedly would provide a real-time audit and verification of data access and use (Winter & Davidson, 2019a).

In addition to data access regulation, blockchain can advance patient-driven interoperability by making PGHD more readily available (Gordon & Catalini, 2018). Blockchain

technology has also been proposed as a means for individuals to manage their own medical records and data. Because blockchain maximizes security and accessibility, the technology can be used in many different areas of the healthcare system, such as for storing and sharing medical records and insurance information both in healthcare venues and in mobile applications and remote monitoring systems, and for clinical trials (Chen et al., 2019). In 2016, the government of Estonia established a national cloud-based, blockchain-secured electronic health records system to secure personal health data while making it available to the individual and his or her healthcare providers (Park, 2019). Such technological developments to enhance the efficacy of regulatory control over all personally identified health data, and PGHD in particular, are sorely needed. Google's continued forays into clinical health data acquisitions demonstrate that the allure of profit from development of AI intellectual property may outweigh concerns about regulatory consequences. Even in a highly regulated environment such as the UK, where Google DeepMind's partnership with the UK National Health Services led to public and regulatory censure (Powles & Hodson, 2017), Google (via its DeepMind Health operation, which has since been absorbed into Google Health) continued to acquire patient data (Winter & Davidson, 2019a). In the U.S. in 2019, the University of Chicago Medical Center and Google were also sued over their partnership to share personal health data (Wakabayashi, 2019).

As deep learning algorithms become more sophisticated, disparate sources of data may be linked to enhance predictive analytics (Bates et al., 2014; Siegel, 2016). Big tech firms rely on these vast amounts of data from disparate sources to develop large language models, a core part of their business, and PGHD are a lucrative target. PGHD "can be combined with personal information from other sources– including healthcare providers and drug companies–raising such potential harms as discriminatory profiling, manipulative marketing, and data breaches" (Montgomery, Chester & Kopp, 2018, p. 42). Assessment of whether firms are compliant with health laws is becoming increasingly difficult due to AI/deep learning's opacity (Winter & Davidson, 2019b). Thus, another example of a promising technological/algorithmic governance structure is "explainable AI". Explainable AI employs mathematics to simplify the so-called black box so that humans can understand the path an AI took to reach an outcome (Abdul et al., 2017; National Institutes of Standards and Technology, 2020), thereby illuminating how personal data including PGHD were used (or misused) and enabling audits for compliance to data protection regulations.

5. Discussion and Conclusions

The growing volume of personally generated health data reflects innovations in health care that can help individuals better manage their health, healthcare providers to advise and treat their patients, third-party payers to incentivize subscribers to adopt healthy practices, and health system leaders and policymakers to improve health outcomes and health service efficiency. To help bring about these outcomes, data protection regulations are critical public policy tools to help balance stakeholders' interests in PGHD and thus to guide developments and innovations in socially desirable ways for all affected parties. In this paper, we have argued that the diverse regulatory regimes that stakeholders operate within contribute to confusion about what regulations apply to PGHD in different contexts, what compliance entails, and how compliance can be audited by regulatory authorities. Our analysis of regulatory regimes in the U.S. examines three data protection regimes (sector-level health data regulation; regulations on research uses of

health data; and self-regulation by commercial firms under consumer privacy regulations) to illustrate how existing regulations governing use of clinical health data are less (or not) effective for governance of PGHD. Comparing across these regimes, we identify lacunae in existing regulations that may enable misuse of PGHD if not addressed.

Our analysis demonstrates how regulatory regimes may be examined and actions taken to harmonize across regimes with respect to governance of PGHD. However, just as algorithms for anonymizing PGHD do not provide a simple, comprehensive solution for data privacy, we argue that regulations alone will also be inadequate to address the challenges of governing PGHD effectively. A variety of data governance structures, tools, and techniques working in tandem will be required to more fully address these challenges. We identify three types of structures (organizational, regulatory, technological/algorithmic) and provide examples within each type, which synergistically could help enact needed regulatory oversight while limiting the friction and economic costs of regulation that may hinder innovation.

This analysis provides a starting point for further discussions and negotiations among stakeholders and regulators on harmonizing regulatory regimes related to PGHD. Although our focus in this paper is limited to a specific regulatory domain in the United States, we believe this work has broader implications for personally generated health data governance policy internationally. The challenges of PGHD regulation are not unique to the U.S., and similar regulatory concerns are arising in many countries (Tanenbaum, 2020). Our analytical approach, drawn from regime theory (May & Jochim, 2013; Maggetti & Ewert, 2018) and applied in the analysis of data protection regulation, will be useful to investigate other national and international contexts of PGHD governance and for other domains of personally identifiable data. Many countries are considering how to integrate software as a medical device (i.e., health apps) and other PGHD-generating technologies into their existing regulatory approaches. Strict general privacy laws, such as the GDPR offer some protection. However, with the EU's strong push to digitize health care, including developing AI and machine learning to improve care and enhance innovation, these protections may not be enough to balance potential societal and economic benefits of widespread, even open, uses of health data in AI developments with the privacy rights promised to individuals (Winter & Davidson, 2019b).

The COVID-19 pandemic vividly highlights the importance of harmonizing regulation of PGHD, to enable various stakeholders to agilely leverage such data resources to address emergent health crisis. PGHD, integrated with mobile location data from smart phones and clinical data (such as test results), could be invaluable to public health authorities for contract tracing and epidemiologists for understanding and predicting communicable disease spread. For instance, tracking apps that generate PGHD with a range of data use permissions have been deployed, and both Big Tech firms and regional and national governments are working on "vaccination passports" that only authorize travel to vaccinated individuals, but also track potential COVID-19 exposure and outbreaks for contact tracing purposes. Patient-generated data on side effects of vaccines or COVID-19 health outcomes (e.g., "long COVID") – collected and/or reported via mobile health app – expand and enrich health data generated in clinical settings such as hospitals. At the same time, widespread collection of PGHD and mobile health apps to address the COVID-19 pandemic heightens threats to individuals' privacy and autonomy to control their data, particularly if these data are widely reused by researchers and commercial

firms in the wake of this public health crisis and for purposes other than public health initiatives. Gasser et al. (2020) outline a number of concerns about the use of data collected to combat the pandemic, including potential repurposing of health and location data, digital inequality and discrimination developing from how these data are used in future, and a lack of accountability for the safety, security, and governance of data in hastily deployed PGHD applications. Acknowledging the diverse regulatory regimes in which PGHD are aggregated and used is a first step towards harmonizing how these COVID-19 data resources are regulated. For instance, while Australia, the EU, and China have centralized their COVID-related data, the U.S. has no centralized federal database of vaccination status. Instead, several large pharmacies and IT corporations (with IBM and CLEAR as frontrunners) are aggregating these data through mobile travel apps (Turner-Lee, Lai, & Skahill, 2021). The gaps in regulatory oversight of PGHD as well innovation-stifling overlaps discussed above, if not addressed effectively, will present substantive challenges to leveraging such data resources to address future crises. The urgent need for regulatory harmonization of PGHD has never been stronger.

Regulatory harmonization is also important for economic development and innovation in the health sector generally. It is particularly important on an international scale to ensure that concerns about privacy and the complexity of the regulatory environment do not overpower the potential for innovation. There is a global race for AI dominance across sectors, where countries with fewer restrictions on data use, such as China and Russia, have a notable advantage due to access to large data sets. The societal importance and economic prospects for firms, and nations, that dominate AI health innovations present strong incentives to prioritize innovation over privacy. For instance, the rapid deployment of AI to address the COVID-19 pandemic, including successes in vaccine development, drug discovery, and disease diagnosis and monitoring (Arshadi et al., 2020; Harmon et al., 2020; National Institutes of Health, 2020) has been lauded and may accelerate or legitimize the desire for fewer restrictions on health data use. Developing agile, easily deployable approaches to regulating PGHD use across regulatory regimes will be vital to maintaining some balance of these goals. The organizational, regulatory, and technological structures discussed above (see Figure 2) outline some key approaches that require elaboration, development, trial, and rapid deployment in order to facilitate socially acceptable access to PGHD resources.

The implications of our analysis of data protection regimes for PGHD extend beyond health data governance per se. Graef, Husovec, and van den Boom (2020) note that, in addition to data portability provisions in the GDPR, other data access regimes are emerging in several sectors of the EU–energy, automotive, payment and digital content/services, and so on. Each of these domains represent important sources of personally identifiable activity data, which can be mined in multiple ways by the firm that originally collects data and then packaged and sold (with some restrictions under GDPR) across organizational and sector boundaries (Spiekermann, 2019). For example, connected vehicles with hundreds of sensors networked with automotive firms and other service providers are being deployed in many countries. In addition to security concerns such as hacking (i.e., taking control of) a moving vehicle, questions about who has access to personal data generated by connected vehicles are being raised, and even whether it will be possible to distinguish personal and non-personal data under the law (Graef, Husovec, and van den Boom, 2020). Across numerous sectors, development of IoT devices and networks is producing many novel data types that will be collected at scale for the first time in human

history, so that their potential uses (and misuses) are as yet unknown. Further research to assess gaps and overlaps in regulatory regimes, for instance within and across sectors and contexts, such as financial, transportation, educational, or household data, is needed to assess the effectiveness of existing regulatory regimes for balancing privacy and innovation, and to formulate approaches to enhance or replace data protection regulations that are ineffective to do so. Our research outlines an approach to carry out this work, which entails examining the prevailing regimes, identifying gaps and overlaps in governance, and identifying potential actions to harmonize across regimes using three types of governance structures (organizational, regulatory, technological/algorithmic).

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Number 1827592. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Abdolkhani, R., Gray, K., Borda, A., & DeSouza, R. (2019). Patient-generated health data management and quality challenges in remote patient monitoring. *JAMIA Open, 2*(4), 471-478.

Abdul, A., Vermeulen, J., Wang, D., Lim, B.Y. & Kankanhalli, M. (2017). Trends and trajectories for explainable, accountable, and intelligible systems: An HCI research agenda. CHI 2018, 582, 1-18.

Acquisti, A., Brandimarte, L., & G. Loewenstein. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509-514.

Aetna. (2020). Employee health and wellness programs. Retrieved from: https://www.aetna.com/insurance-producer/health-wellness/wellness-programs.html

Agency for Healthcare Research and Quality. (2019, Dec. 3). Automated-entry patient generated health data for chronic conditions: The evidence on health outcomes. Evidence-based Practice Center Technical Brief Protocol. Revised March 17, 2020.

Anonymous. (2019, November 14). I'm the Google whistleblower. The medical data of millions of Americans is at risk. *The Guardian.* Retrieved from: https://www.theguardian.com/commentisfree/2019/nov/14/im-the-google- whistleblower-the-medical-data-of-millions-of-americans-is-at-risk

Arshadi A K, Webb J, Salem M, et al. (2020). Artificial intelligence for COVID-19 drug discovery and vaccine development. *Frontiers in Artificial Intelligence, 3*, 65. doi:10.3389/frai.2020.00065

Bari, L., & O'Neill, D. P. (2019). Rethinking patient data privacy in the era of digital health. Health Affairs Blog. Retrieved from: https://www.healthaffairs.org/do/10.1377/hblog20191210.216658/full/

Bates, D.W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs, 33*(7), 1123-1131.

Centre for Information Policy Leadership. (2019, Mar. 8). Regulatory sandboxes in data protection: Constructive engagement and innovative regulation in practice. Retrieved from: https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_regulatory_sandboxes_in_data_protection_-_constructive_engagement_and_innovative_regulation_in_practice 8_march_2 019_.pdf

Chen, H. S., Jarrell, J. T., Carpenter, K. A., Cohen, D. S., & Huang, X. (2019). Blockchain in healthcare: A patient-centered model. *Biomedical Journal of Scientific & Technical Research, 20*(3), 15017.

Copeland, R. (2019, Nov. 11). Google's 'Project Nightingale' gathers personal health data on millions of Americans. *The Wall Street Journal*. Retrieved from: https://www.wsj.com/articles/google-s-secret- project-nightingale-gathers-personal-health-data-on-millions-of-americans-11573496790

Copeland, R., & Needleman, S.E. (2019, Nov. 12). Google's 'Project Nightingale' triggers federal inquiry. *The Wall Street Journal*. Retrieved from: https://www.wsj.com/articles/behind-googles-project-nightingale-a-health-data-gold-mine-of-50-million-patients-11573571867

Davidson, E. J., Østerlund, C. S., & Flaherty, M. G. (2015). Drift and shift in the organizing vision career for personal health records: An investigation of innovation discourse dynamics. *Information and Organization, 25*(4), 191-221.

Davis, J. (2019). Google Ascension partnership fuels overdue HIPAA privacy debate. Retrieved from: https://healthitsecurity.com/news/google-ascension-partnership-fuels- overdue-hipaa-privacy-debate

DeepMind. (2017). Trust, confidence and verifiable data audit. Retrieved from: https://deepmind.com/blog/trust-confidence-verifi-able-data-audit/

Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data trusts: Disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law, 9*(4), 236-252.

Dyrda, L. (2019, Nov. 8). Tim Cook: 4 key thoughts on Apple's healthcare offerings today and in the future. *Becker's Health IT*. Retrieved from: https://www.beckershospitalreview.com/healthcare-information-technology/tim-cook-4-key-thoughts-on-apple-s-healthcare-offerings-today-and-in-the-future.html

eHealth Initiative & Foundation, & Center for Democracy and Technology (2021, Feb. 9). Proposed consumer privacy framework for health data. Retrieved from:

https://cdt.org/wp-content/uploads/2021/02/2021-02-09-CDT-and-eHI-Proposed-Consumer-Privacy-Framework-for-Health-Data-d-FINAL.pdf

Faitelson, Y. (2019). Why U.S. GDPR-style privacy laws a good for business. *Forbes*. Retrieved from: https://www.forbes.com/sites/forbestechcouncil/2019/12/19/why-u-s-gdpr-style-privacy-laws-are-good-for-business/?sh=6acdb7258756

Farr, C. (2020, Feb. 11). Google Health, the company's newest product area, has ballooned to more than 500 employees. CNBC. Retrieved from: https://www.cnbc.com/2020/02/11/google-health-has-more-than-500- employees.html

Federal Trade Commission. (2018). Privacy & data security. Update 2018. Retrieved from: https://www.ftc.gov/system/files/documents/reports/privacy-data-security-update-2018/2018-privacy-data-security-report-508.pdf

Federal Trade Commission. (2021, Jan 13). Developer of popular women's fertility-tracking app settles FTC allegations that it misled consumers about the disclosure of their health data. Retrieved from: https://www.ftc.gov/news-events/press-releases/2021/01/developer-popular-womens-fertility-tracking-app-settles-ftc

Federal Trade Commission v. Facebook, Inc. (2020, December 9). Retrieved from: https://www.ftc.gov/system/files/documents/cases/1910134fbcomplaint.pdf

Feijóo, C., Kwon, Y., Bauer, J. M., Bohlin, E., Howell, B., Jain, R., et al. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. *Telecommunications Policy*, *44*(6), 101988. doi: 10.1016/j.telpol.2020.101988

Frazee, G. (2019, Nov. 1). Google bought Fitbit. What does that mean for your data privacy? PBS News Hour. Retrieved from: https://www.pbs.org/newshour/economy/making-sense/google-bought-fitbit-what-does-that-mean-for-your-data-privacy

Gasser, R., Ienca, M., Scheibner, J., Sleigh, J. & Vayena, E. (2020). Digital tools against COVID-19: Taxonomy, ethical challenges, and navigation aid. *The Lancet Digital Health 2020*(2), e425–34. doi: 10.1016/S2589-7500(20)30137-0

Gellman, R., & Dixon, P. (2011). Many failures: A brief history of privacy self-regulation in the United States. World Privacy Forum. Retrieved from: http://www.worldprivacyforum.org/www/wprivacyforum/pdf/WPFselfregulationhistory.pdf

Genes, N., Violante, S., Cetrangol, C., Rogers, L., Schadt, E. E., & Chan, Y. F. Y. (2018). From smartphone to EHR: A case report on integrating patient-generated health data. *NPJ Digital Medicine, 1*(1), 1-6.

Gordon, W. J., & Catalini, C. (2018). Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. *Computational and Structural Biotechnology Journal*, 16, 224-230.

Graef, I., Husovec, M., & van den Boom, J. (2020). Spill-overs in data governance: Uncovering the uneasy relationship between the GDPR's right to data portability and EU sector-specific data access regimes. *Journal of European Consumer and Market Law, 9*(1), 3-16.

Hall, M. A., & Studdert, D. M. (2021). "Vaccine passport" certification — Policy and ethical considerations. *New England Journal of Medicine, 385*(e32). doi: 10.1056/NEJMp2104289

Hao, K. (2020, Dec. 4). We read the paper that forced Timnit Gebru out of Google. Here's what it says. *MIT Technology Review*. Retrieved from: https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/

Harmon, S. A., Sanford T. H., Xu, S., et al. (2020). Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nature Communications, 11*(1), 1-7.

Hern, A. (2017). Google's DeepMind plans bitcoin-style health record tracking for hospitals. *The Guardian*. Retrieved from: https://www.theguardian.com/technology/2017/mar/09/google-deepmind-health-records-tracking-blockchain-nhs-hospitals

HIPAA Journal. (2021, Jan. 18). Possible HIPAA updates and HIPAA changes in 2021. Retrieved from: https://www.hipaajournal.com/hipaa-updates-hipaa-changes/

Humer, C., & Finkle, J. (2014). Your medical record is worth more to hackers than your credit card. Reuters. Retrieved from: https://www.reuters.com/article/us-cybersecurity-hospitals/your-medical-record-is-worth-more-to-hackers-than-your-credit-card-idUSKCN0HJ21I20140924

Igo, S. E. (2018). *The known citizen: A history of privacy in modern America*. Cambridge, MA: Harvard University Press.

Jim, H. S., Hoogland, A. I., Brownstein, N. C., Barata, A., Dicker, A. P., Knoop, H., Gonzalez, B. D., Perkins, R., Rollison, D., Gilbert, S. M., Nanda, R., Berglund, A., Mitchell, R., & Johnstone, P. A. (2020). Innovations in research and clinical care using patient- generated health data. *CA: A Cancer Journal for Clinicians, 70*(3), 182-199.

Johnson, B., & Lichfield, G. (2019, April 6). Hey Google, sorry you lost your ethics council, so we made one for you. *MIT Technology Review*. Retrieved from: https://www.technologyreview.com/2019/04/06/65905/google-cancels-ateac-ai- ethics-council-what-next/

Krämer, J., Whalley, J., & Batura, O. (2019). The data economy and data-driven ecosystems: Regulation, frameworks and case studies. *Telecommunications Policy, 43*(2), 113-182.

Krasner, S. (1982). Structural causes and regime consequences: Regimes as intervening variables. *International Organization, 36*, 185–205.

Langarizadeh, M., Orooji, A., Sheikhtaheri, A., & Hayn, D. (2018). Effectiveness of anonymization methods in preserving patients' privacy: A systematic literature review. *eHealth*, *248*, 80-87.

Levi-Faur, D. (2010). Regulatory networks and regulatory agencification: Towards a single European regulatory space. *Jerusalem Papers in Regulation and Governance.* Working paper no. 1. Jerusalem: Jerusalem Forum on Regulation and Governance.

Loveland, L. (2020, Mar. 4). US senators question Ascension on its Google collaboration Project Nightingale. Retrieved from: https://www.mobihealthnews.com/news/us- senators-question-ascension-its-google-collaboration-project-nightingale

Maggetti, M., & Ewert, C. (2018). Comparative regulatory regimes and public policy. In *The Palgrave Handbook of Public Administration and Management in Europe* (pp. 635-651). Palgrave Macmillan, London.

Mandl, K. D., Mandel, J. C., & Kohane, I. S. (2015). Driving innovation in health systems through an apps-based information economy. *Cell Systems, 1*(1), 8-13.

Marelli, L., Testa, G., & van Hoyweghen, I. (2021). Big Tech platforms in health research: Re-purposing big data governance in light of the General Data Protection Regulation's research exemption. *Big Data & Society, 8*(1). doi:10.1177%2F20539517211018783

May, P. J., & Jochim, A. E. (2013). Policy regime perspectives: Policies, politics, and governing. *Policy Studies Journal*, *41*(3), 426-452.

Melstrom, L. G., Rodin, A. S., Rossi, L. A., Fu Jr, P., Fong, Y., & Sun, V. (2021). Patient generated health data and electronic health record integration in oncologic surgery: A call for Artificial Intelligence and machine learning. *Journal of Surgical Oncology*, *123*(1), 52-60.

Metz, C. (2021, March 5). Second Google A.I. researcher says she was fired. *The New York Times*. Retrieved from: https://www.nytimes.com/live/2021/02/19/business/stock-market-today#google-ethical-artificial-intelligence-team

Meystre, S. M., Lovis, C., Bürkle, T., Tognola, G., Budrionis, A., & Lehmann, C. U. (2017). Clinical data reuse or secondary use: Current status and potential future progress. *Yearbook of Medical Informatics, 26*(1), 38.

Montgomery, K., Chester, J., & Kopp, K. (2018). Health wearables: Ensuring fairness, preventing discrimination, and promoting equity in an emerging Internet-of-Things environment. *Journal of Information Policy, 8*, 34–77.

National Institutes of Health. (2020, Aug. 5). United States. NIH harnesses AI for COVID-19 diagnosis, treatment, and monitoring. Retrieved from: https://www.nih.gov/news-events/news-releases/nih-harnesses-ai-covid-19-diagnosis-treatment-monitoring

National Institutes of Standards and Technology. (2020, Aug.) Four principles of Explainable Artificial Intelligence. Draft NISTIR 8312. Retrieved from:

https://www.huntonprivacyblog.com/wp-content/uploads/sites/28/2020/08/NIST-Explainable-AI-Draft-NISTIR8312-1.pdf

Office of the National Coordinator for Health Information Technology. (2014). Patient-generated health data. Retrieved from: https://www.healthit.gov/sites/default/files/patient_generated_data_factsheet.pdf

O'hara, K. (2019). Data trusts: Ethics, architecture and governance for trustworthy data stewardship (Web Science Institute White Papers, Southampton). University of Southampton 1-27.

Park, A. (2019, June 25). What the US can learn from Estonia's cloud-based, blockchain-secured EHR system. *Becker's Health IT*. Retrieved from: https://www.beckershospitalreview.com/ehrs/what-the-us-can-learn-from-estonia-s-cloud-based-blockchain-secured-ehr-system.html

Perez-Pozuelo, I., Spathis, D., Gifford-Moore, J., Morley, J., & Cowls, J. (2021). Digital phenotyping and sensitive health data: Implications for data governance. *Journal of the American Medical Informatics Association,28*(9), 2002-2008.

Pifer, R. (2019, November 18). IT execs call for HIPAA overhaul in 'Project Nightingale' wake. https://www.healthcaredive.com/news/it-execs-call-for-hipaa-overhaul-in-project-nightingale-wake/567520/

Pilkington E. (2019, Nov. 12). Google's secret cache of medical data includes names and full details of millions – whistleblower. *The Guardian*. Retrieved from: https://www.theguardian.com/technology/2019/nov/12/google-medical-data-project-nightingale-secret-transfer-us-health-information

Powles, J., & H. Hodson. 2017. Google DeepMind and healthcare in an age of algorithms. *Health Technology 7*(4):351-367. doi.org/10.1007/s12553-017-0179-1

Price, L. (2020, Jan. 16). Project Nightingale: Google's four pillars for their secret patient data partnership. Retrieved from: https://www.healthcare.digital/single-post/2019/11/13/Project-Nightingale-Googles-four-pillars-for-their-secret-patient- data-partnership

Ravuri, M., Kannan, A., Tso, G. J., & Amatriain, X. (2018). Learning from the experts: From expert systems to machine-learned diagnosis models. *Machine Learning for Healthcare Conference* (pp. 227-243). PMLR.

Rosenbaum, S. (2010), Data governance and stewardship: designing data stewardship entities and advancing data access. *Health Services Research, 45*(5p2), 1442-1455.

Rothstein, M. A., Wilbanks, J. T., Beskow, L. M., et al. (2020). Unregulated health research using mobile devices: ethical considerations and policy recommendations. *The Journal of Law, Medicine & Ethics, 48*(1_suppl), 196-226.

Schneier, B. (2015). *Data and Goliath: The hidden battles to collect your data and control your world*. New York: WW Norton & Company.

Sharon, T. (2016). The Googlization of health research: from disruptive innovation to disruptive ethics. *Personalized Medicine*, *13*(6), 563-574.

Shuren, J., Patel, B., & Gottlieb, S. (2018). FDA regulation of mobile medical apps. *JAMA, 320*(4), 337-338.

Siegel, E. (2016). *Predictive analytics: The power to predict who will click, buy, lie, or die.* Hoboken, NJ: John Wiley & Sons.

Solove, D. J. (2010). *Understanding privacy*. Cambridge, MA: Harvard University Press.

Solove, D. J. (2019, April 22). Will the United States finally enact a federal comprehensive privacy law? TechPrivacy. Retrieved from: https://teachprivacy.com/will-us-finally-enact-federal-comprehensive-privacy- law/

Spiekermann, M. (2019). Data marketplaces: Trends and monetisation of data goods. *Intereconomics*, *54*(4), 208-216.

Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery. Big Data, 1(2), 85-99.

Tanenbaum, W. A. (2020). *Digital health, new technologies and emerging legal issues. Digital health 2020*. London: International Comparative Legal Guides.

Tanner, A. (2017). *Our bodies, our data: how companies make billions selling our medical records*. Boston, MA: Beacon Press.

Taylor, R. D. (2017). The next stage of US communications policy: The emerging embedded infosphere. *Telecommunications Policy, 41*(10), 1039-1055.

Taylor, R. D. (2020). 'Data localization': The Internet in the balance. *Telecommunications Policy, 44*(8), 1-15.

Tempini, N. (2017). Till data do us part: Understanding data-based value creation in data-intensive infrastructures. *Information and Organization, 27*(4), 191-210.

Turner-Lee, N., Lai, S., & Skahill, E. (2021). Vaccine passports underscore the necessity of U.S. privacy regulation. The Brookings Institution. Retrieved from: https://www.brookings.edu/blog/techtank/2021/06/28/vaccine-passports-underscore-the-necessity-of-u-s-privacy-legislation/

U.S. Department of Health and Human Services. (nd). HIPAA for professionals. Retrieved from: https://www.hhs.gov/hipaa/for-professionals/index.html

U.S. Department of Health and Human Services. (2013). Summary of the HIPAA Privacy Rule. Retrieved from: https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html

U.S. Department of Health and Human Services. (2016). Federal Policy for the Protection of Human Subjects ('Common Rule'), Retrieved from: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html

U.S. Department of Health and Human Services. (2018). Conceptualizing a data infrastructure for the capture, use, and sharing of patient-generated health data in care delivery and research through 2024. Retrieved from: https://www.healthit.gov/sites/default/files/onc_pghd_final_white_paper.pdf

United States Senate. (2020, March 2). Letter to Ascension [communication]. Retrieved from: https://www.warren.senate.gov/imo/media/doc/2020.03.02%20Letter%20to%20Ascension%20re%20Project%20Nightingale%20Partnership.pdf

Van den Broek, T. & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. *Technological Forecasting and Social Change, 129*, 330-338.

Wakabayashi, D. (2019, June 26). Google and the University of Chicago are sued over data sharing. *The New York Times*. Retrieved from: https://www.nytimes.com/2019/06/26/technology/google- university-chicago-data-sharing-lawsuit.html

Warzel, C. (2019, Aug. 13). All your data is health data: And Big Tech has it all. *The New York Times*. https://www.nytimes.com/2019/08/13/opinion/health-data.html

Winter, J. S., & Davidson, E. (2019a). Big data governance of personal health information and challenges to contextual integrity. *The Information Society, 35*(1), 36-51. Doi:10.1080/01972243.2018.1542648

Winter, J. S., & Davidson, E. (2019b). Governance of artificial intelligence and personal health information. *Digital Policy, Regulation and Governance (DPRG), 21*(3), 280-290. Doi:10.1108/DPRG-08-2018-0048

Xiao, C., Choi, E., & Sun, J. (2018). Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. *Journal of the American Medical Informatics Association, 25*(10), 1419-1428.

Yeh, C.-L. (2018). Pursuing consumer empowerment in the age of Big Data: A comprehensive regulatory framework for data brokers. *Telecommunications Policy, 42*(4), 282–292.