5-2023

The Health Care Industry is Ready for a Revolution: Its Privacy Laws are Not

Erin Rutherford

The Health Care Industry is Ready for a Revolution: Its Privacy Laws are Not

Erin Rutherford

I. INTRODUCTION

As the world is becoming increasingly digitized, its citizens are making increasingly more amounts of personal information publicly available, both knowingly and unknowingly. This information has the potential to revolutionize health care in the United States by improving both health outcomes and responses

to public health emergencies, as well as by addressing problems with accessibility and affordability. These benefits are attainable only by properly regulating the data with respect to the privacy interests of the individuals who generate the data. Using digitized data to improve public health responses and to utilize artificial intelligence to drive valuable care innovations is dependent upon cooperation from individuals. Trust encourages cooperation, and transparency encourages trust. If people do not trust that the system respects and protects their privacy interests, the entire system fails.

The response to the COVID-19 pandemic provides an illustration of the benefits of digitized health data and the need for cooperation. Prior to the development of the COVID-19 vaccines, contact tracing was one of the most effective measures available to slow the spread of the virus.[1] An efficient way to readily provide the necessary parties access to accurate and uniform digitized information regarding exposures could have minimized the need to implement widespread lockdowns and quarantines.[2] However, due to a general distrust regarding storage and subsequent use of location and health information, the full benefits of contact tracing were out of reach.[3]

To illustrate the need for a willingness to contribute information for innovation, consider the following example. Imagine a diagnostic method that detects colon cancer even before symptoms appear. It does so with high accuracy and without the need for invasive, time consuming, and costly lab tests, blood tests, colonoscopies, and image analyses. Instead, this method would need only to analyze an individual's routinely collected health data. Efficient and reliable diagnostic processes like the one just described will not only provide early detection, which is crucial for survival rates, but they will also reduce the cost of and increase the accessibility to quality health care.

These two examples demonstrate the power of digitized data to improve three major values of a health system: accuracy, efficiency, and accessibility. Achieving all three of these with one

---

1. Emily Berman et al., COVID-19 Surveillance 1 (Aug. 3, 2020) (unpublished manuscript) https://ssrn.com/abstract=3666300.

2. Leah R. Fowler et al., *Improving Data Collection and Management, in* COVID-19 POLICY PLAYBOOK: LEGAL RECOMMENDATIONS FOR A SAFER, MORE EQUITABLE FUTURE 45, 46–47 (Scott Burris et al. eds., 2021).

3. Berman et al., *supra* note 1, at 11.

solution is remarkable in that the means to each of these respective goals are often in conflict with one another.

Capitalizing on all the benefits and promise that innovative technologies have to offer to health care requires utilizing readily transmissible health data. As the health care industry specifically becomes increasingly digitized, health care data is essential not only for improving diagnoses and treatments, but also for improving quality, efficiency, and accessibility of care. Achieving the latter three requires use of health data in ways that might be unexpected to individuals and even to researchers and product developers until after obtaining the data. This presents a number of risks including data misappropriation, new avenues of discrimination, and increased vulnerability to security threats. Failure to address those problems and others will cause individuals to lose trust in the health care system and become hesitant to disclose information even for their own treatment purposes, let alone for secondary purposes that may not benefit them directly but instead are intended to serve a larger interest of the health care system or society as a whole.

Systemwide distrust and hesitation could stall innovation and more significantly, impair the integrity and function of the health care system. In order to both reduce frustration related to privacy of health information and encourage technological innovation, the US should implement a flexible privacy framework that prioritizes individual autonomy and allows practices to adapt as different contexts and individual interests require.

This paper highlights the costs and benefits associated with the gathering, storing, analyzing, and digitizing of health information; examines current privacy laws and their inadequacies in the new and constantly changing digital health world; and then provides a proposal framework to balance encouraging innovation while protecting individual autonomy. The article specifically proceeds as follows. This paper first discusses of the evolution of the health industry, from paper records to the wide array of sources generating health information today. Next, it considers the benefits to the ever-increasing amount of health information, which, while considerable can often be in tension with privacy and autonomy interests. It then examines the current privacy models applicable to the various sources of health information, and highlight the types of information left unprotected, as well as the

ways in which individuals lack control over their health information. Finally, this paper introduces a framework that is flexible enough to adapt to different societal interests but maintains the integrity of the health system by ensuring that the individual's interests remain the priority. Such a framework will allow individuals to be in control of their health information in a way that protects their individual interests and allows them to contribute to societal growth.

II. BACKGROUND

Health data comes from a number of sources, ranging in quality, uniformity, and identifiability. The more standardized and complete a data set is, the more beneficial it is for health care solutions.[4] In other words, data utility increases as individual privacy decreases.

Different contexts and subjective privacy interests can influence the degree of tension between data utility and privacy interests. In addition to privacy concerns and risks of reidentification, increased collection and use of health data creates and highlights other problems such as potential discrimination, inadequate consent practices, and cybersecurity risks. These problems not only threaten individual autonomy regarding one's health information, but they have consequences that can spill over into other areas of life as well.[5]

For example, widespread dissemination of health data allows those in possession of it to draw inferences about the people from whom it came and subsequently create consumer profiles about individuals or even whole classes of people.[6] From these inferences, organizations could illegally discriminate against people without their knowing it was even possible for the organization to do so. Additionally, marketing companies can capitalize on these consumer profiles by analyzing a person's health information and sending relevant advertisements straight to the individual. While customized advertisements

---

4. Charlotte A. Tschider, *The Healthcare Privacy-Artificial Intelligence Impasse*, 36 SANTA CLARA HIGH TECH. L.J. 439, 440 (2020).

5. Scott R. Peppet, *Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security, and Consent*, 93 Tex. L. Rev. 85, 112 (2014).

6. Mark A. Rothstein, *Is Deidentification Sufficient to Protect Health Privacy in Research?*, 10 AM. J. BIOETHICS 6 (2010).

might be desirable to some people, they are unnerving to others.[7] While it is possible that one may consent to sharing his health information for one purpose, that information could eventually also contribute to a purpose that the person would not have given consent.[8] The generalized consent models responsible for protecting these individual preferences often fall short of their duties as they typically involve long and complex forms that are presented in a context that does not actually provide the individual with a meaningful choice.[9] Additional problems arise with cybersecurity concerns, but this article will not address those problems.

III. EVOLUTION OF HEALTHCARE INDUSTRY

Gone are the days of physician house calls. Even the days of intimate, face-to-face visits in a physician's office seem to be dwindling. Technological advancements have already revolutionized the healthcare industry but the revolution (and the need for one) is far from over.[10] With this comes a change in how entities collect, store, and transmit health information.[11] Moreover, this entails a change in who is collecting, storing, and transmitting this information, how they are collecting it, and for what purposes.[12] What has not sufficiently changed amidst this revolution are privacy regulations.[13]

A. TRANSITION FROM PAPER RECORDS TO ELECTRONIC MEDICAL RECORDS

Prior to the digital era, a provider would record a patient's medical information by hand in free-form text that read as a patient's narrative.[14] Physicians obtained the narrative's data

---

7. Daniel J. Solove, *Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880, 1896 (2013).

8. Rothstein, *supra* note 6, at 7.

9. *See generally* Solove, *supra* note 7 (highlighting problems with a self-management regime established under current informed consent practices).

10. Tschider, *supra* note 4, at 441.

11. Tasha Glenn & Scott Monteith, *Privacy in the Digital World: Medical and Health Data Outside of HIPAA Protections*, CURRENT PSYCHIATRY REPS., Sept. 2014, at 1, 2.

12. *Id.*

13. Charlotte A. Tschider, *The Consent Myth: Improving Choice for Patients of the Future*, 96 WASH. L. REV. 1505, 1507 (2018).

14. Daniel A. Moros, *The Electronic Medical Record and the Loss of Narrative*, 26 CAMBRIDGE Q. HEALTHCARE ETHICS 328, 328 (2017).

through natural conversation with and observation of a patient, asking relevant questions accordingly.[15] With electronic health records (EHR) now well-established, structured templates found within the EHR guide the desired types and forms of data for any care team member, often not the physician, to record.[16] Additionally, some of this data may come directly from outside labs or other testing facilities without any human communication.[17] While providers may still electronically record notes in a narrative format, the bulk of a patient's EHR consists of structured data.[18]

Digitizing health records makes data collection more efficient and makes using the data collected effectively easier than before.[19] Additionally, electronic systems now allow for storing and easy sharing of vast amounts of health information.[20] This benefits not only health care and insurance providers, but health administrators and patients as well.[21] Providers and patients benefit because quick and easy data transfer allows for higher quality and safer continuity of care as patients move between various types of facilities and providers. Insurance providers benefit because easy access to treatment and payment information allows them to process insurance claims effectively and promptly. Public health administrators benefit because readily accessible data allows them to review health information pertinent to policymaking and health surveillance, and then respond quickly when necessary.[22] An additional benefit for patients is the level of transparency that easy access to their EHR provides. Patients now have access to their health information as soon as it appears in their EHR. The ability to see this information so readily gives patients an element of control over their care experience that was not possible with paper records.

---

15. *Id.*

16. April Moreno Arellano et al., *Privacy Policy and Technology in Biomedical Data Science*, 1 ANN. REV. BIOMEDICAL DATA SCI. 115, 120 (2018).

17. Moros, *supra* note 14.

18. *Id.*

19. Sylvestre Uwizeyemungu et al., *European Hospitals' Transition Toward Fully Electronic-Based Systems: Do Information Technology Security and Privacy Practices Follow?*, 7 JMIR MED. INFORMATICS 1, 2 (2019).

20. *Id.*

21. *Id.*

22. *Id.*

These benefits illustrate how digitized data vastly enhances the data's primary uses. Innovators, however, have looked beyond the digitized data's primary use possibilities and have recognized value in using data in ways that are less obvious; frequently referred to as "secondary uses." The structured format of the clinical data within an EHR makes it particularly useful for research, drug development, and device development.[23] Digitization of health records was not originally intended to serve these innovative, secondary functions. As a result, the law is now struggling to ensure both the market and opportunities for these secondary uses, while still continuing to ensure the priorities for the primary purposes of health records, namely patient privacy and improving quality of care.

Digitized health records allow for more long-term storage of health information, which is helpful for providing a complete picture of a patient's medical history. As data from EHRs of all individuals within a health system cumulate, however, storage becomes an issue.[24] The answer to this problem typically involves utilizing some kind of cloud storage system.[25] While cloud systems provide increased storage with quick and easy access, these systems also come with additional security and privacy concerns.[26] For example, if a patient wishes to obtain a digital copy of any information within her EHR, she would need to download the information from the provider's EHR system.[27] Once downloaded to a personal device, the owner of that device becomes responsible for maintaining its privacy as the data is no longer in the custody of a regulated entity and is now outside legal protection.[28] For those who are less technologically savvy, that responsibility bears more risk of inadvertent sharing of personal information.[29] Therefore, with greater patient accessibility to health information comes greater responsibility. Expecting individuals to knowingly self-manage privacy

---

23. Arellano et al., *supra* note 16, at 120.

24. N. Peek et al., *Technical Challenges for Big Data in Biomedicine and Health: Data Sources, Infrastructure, and Analytics*, 23 Y.B. MED. INFORMATICS 42, 44 (2014).

25. *Id.* at 45.

26. *Id.*

27. Glenn & Monteith, *supra* note 11, at 2.

28. *Id.*

29. *Id.*

interests regarding their health information is an unrealistic expectation in the digital era.[30]

B. SOURCES OF HEALTH CARE DATA

Digitizing the health care sector triggered an exponential increase in sources for health care information—providing interested parties with an unprecedented amount of data.[31] Much of this data originates from sources outside the protections of the dominant federal privacy law related to health information—the Health Insurance Portability Accountability Act (HIPAA).[32] At its initial passage, HIPAA's primary purpose was to facilitate transmission of protected health information related to insurance coverage.[33] As privacy was not an initial concern, HIPAA applies only to covered entities and business associates, the definitions of which are relatively narrow and do not sufficiently account for the technological advancements within the health care industry.[34] Further discussion of HIPAA and its limitations follows in a later section.

1. Sources Within the Reach of HIPAA

a. Electronic Health Records

Passage of the Health Information Technology for Clinical Health (HITECH) Act of 2009 mandated health organizations to make health data electronic,[35] and the 21st Century CURES Act passed in 2016 mandated interoperability standards for all EHR systems in the United States.[36] Interoperability refers to the ability of electronic health information systems to share and transfer information.[37] Congress later focused on transparency

---

30. Solove, *supra* note 7, at 1880.

31. Alap Shah, *Top Ten Issues in Health Law 2021*, AM. HEALTH L. ASS'N (Oct. 1, 2021, 10:11 AM), https://www.americanhealthlaw.org/content-library/connections-magazine/article/bbcd0c57-0da4-4d16-adbe-c997605ecea2/Top-Ten-Issues-in-Health-Law.

32. Glenn & Monteith, *supra* note 11, at 2.

33. Tschider, *supra* note 13, at 1513 n.39.

34. *Id.* at 1515.

35. Charlotte Tschider, *AI's Legitimate Interest: Towards A Public Benefit Privacy Model*, 21 HOUS. J. HEALTH L. & POL'Y 125, 140 (2021).

36. *Interoperability and Patient Access Fact Sheet*, CTRS. FOR MEDICARE & MEDICAID SERVS., https://www.cms.gov/newsroom/fact-sheets/interoperability-and-patient-access-fact-sheet (last visited Oct. 1, 2021).

37. *Id.*

and patient control of their health; 2020 amendments to the CURES Act prohibit information blocking, meaning patients must have access to their health information as soon as the information enters the chart.[38] While recent laws encourage the use of technology by mandating that entities utilize technologies and software systems that make information readily transferrable and accessible to all authorized parties, they do not address updating and implementing corresponding privacy regulations.

Nonetheless, EHRs have become a rich source of clinical data.[39] Even though EHRs are primarily meant for patient care, the wide array of clinical records holds tremendous value for secondary uses, such as research and product development.[40] EHR data include demographics, laboratory values, dispensed medications, imaging and diagnostic data, clinical interventions, and some clinical notes in free-form text.[41] Considering the number of patients within an organization's health system, the number of diseases and conditions represented is extensive. Additionally, the records include repeated observations of patients giving the data a longitudinal quality that is difficult to obtain outside of a clinical setting.[42]

While EHR data's potential is exciting from an innovation standpoint, HIPAA protections limit its utility in two ways.[43] First, if an organization wishes to retain the data's maximum utility by including identifiable elements, HIPAA requires express, written consent of the individual.[44] Obtaining consent from each individual in a data set creates an administrative burden that is often nearly impossible to overcome. Second, an organization can circumvent the consent requirement by accepting the data in a deidentified format. The disadvantage here is that as data sets lose their identifiable elements, they lose some of their value. Some secondary uses benefit more from identifiability than others, so the magnitude of this cost varies according to the secondary use.[45] Although sometimes the exact

38. *Id.*
39. Peek et al., *supra* note 24, at 42–43.
40. *Id.*
41. *Id.*
42. *Id.*
43. Tschider, *supra* note 4, at 441.
44. Tschider, *supra* note 35, at 146.
45. *Id.*

secondary use is not known until well after the entity obtains the data set, making it difficult to weigh this cost at the time of acquisition.[46]

b. Administrative Data

A second source of HIPAA protected information comes from what is termed "administrative data." Administrative data refers to the information needed for "insurance or other claims of payment."[47] Typically, these data sets consist of less clinical data, limited to only information pertaining to diagnosis and treatment. Treatment data may include procedures, medications, and devices.[48] Additionally, these sources contain an individual's financial and insurance information.[49] Even despite the lack of clinical information, the data these records do contain are more standardized making it easier for parties to retrieve and analyze and are therefore still valuable.[50] As with EHR data, administrative data is subject to the same limitations presented by HIPAA protections for identifiability forcing entities to sacrifice utility by accepting deidentified information.[51]

2. Sources Outside of the Reach of HIPAA

a. Routine Interactions

Along with the obvious new sources of health information discussed above, the digital era created a number of less obvious sources. Despite their inconspicuous nature, these sources provide diverse and numerous bits of health information.[52] Sources of routine interactions provide information that is volunteered, sometimes linked to an individual's identity, and subject only to the Federal Trade Commission's privacy regulations or additional privacy policies of the respective organization.[53]

---

46. Tschider, *supra* note 4, at 441.
47. Peek et al., *supra* note 24, at 43.
48. *Id.*
49. *Id.*
50. *Id.*
51. *Id.*
52. Glenn & Monteith, *supra* note 11, at 2.
53. Tschider, *supra* note 35, at 150.

Consider a routine financial transaction. Individuals offer up health information every time they use a credit card to pay for a medical appointment, over the counter medications, health foods and supplements, home testing products, or items related to disabilities.[54] Often, when someone is not feeling well or a sudden pain strikes, one of the first things that person will do is conduct an internet search. Searching online for symptoms and possible treatments leaves a trail of health information that data brokers are mining the web to collect.[55] Here, a single search may not reveal identifiable information.[56] However, in conjunction with other internet activity the individual uses that device for, it may provide enough information for data brokers to link unidentifiable information to something that reveals the individual's identity.[57] Researchers also find a wealth of information through social media and other social networking resources.[58] Blog posts, Facebook likes, Tweets, memberships in online groups, and discussion boards often include voluntarily given identifiable health information.[59] For example, imagine a person who is in a support group for living with Parkinson's. That person may often check discussion boards and review various treatments, potentially revealing what treatments he is currently using or considering. Perhaps that individual finds it therapeutic to share his story, including his daily symptoms and struggles. All of this information has potential value for product development, marketing medications, disease progression, disease demographic, and other uses.

When individuals openly share their personal health struggles, product developers and health officials are able to infer what types of health problems are prevalent in the public and respond accordingly.[60] Consider a particularly severe flu season. Perhaps more and more people begin searching flu symptoms and flu treatments. Maybe afflicted individuals are turning to social media to share their current state. Epidemiologists can use this information to explore possibilities for the unusually severe flu season. For example, maybe less

---

54. Glenn & Monteith, *supra* note 11, at 2.
55. *Id.*
56. *Id.*
57. Peek et al., *supra* note 24, at 43.
58. *Id*.
59. *Id.*
60. *Id*.

people received a flu shot that year, or perhaps the flu shot did not contain the necessary components to protect against the prevalent strain of the flu. Now, those responsible can adjust and make improvements for next year's flu shot. Additionally, public health officials can use the information to warn the public to take extra precautions against the flu. In regard to discussion boards or support groups for a particular disease, researchers are able to track which medications and methods of treatment are having what kind of results. They can use this information to market new medications or treatments to those individuals, as well as work to improve the existing medications and treatments to avoid some of the negative side-effects people are discussing. Furthermore, researchers can look at the make-up of individuals in the group and deduce classes of people more commonly affected by that particular disease and look into why the disease more commonly or more severely affects that class of people, spurring a deeper understanding of the disease's mechanisms.

b. Wearables and Third-Party Apps

Health related wearables continue to grow in popularity. Consumers often purchase wearables for self-initiated health improvements. Additionally, physicians are increasingly prescribing smart wearables.[61] Key to the operation of all wearables is accompanying mobile apps. Wearables monitor various human activities and report that information back to the app. Individuals can then turn to the app and analyze the information collected. In the event that the wearable is prescribed, an individual's doctor may also have access to the information.[62] Examples include WebMD, FitBit, smart hearing aids, and smart insulin pumps.[63]

Contrary to popular belief, these mobile apps are not subject to HIPAA, even in the event that they communicate with a prescribed device.[64] Third party companies, that do not qualify as an entity subject to HIPAA, are the typical developers of mobile health apps even when the apps accompany use of a

61. *See* Glenn & Monteith, *supra* note 11, at 3 (discussing mobile medical apps).
62. *Id.*
63. *Id.*
64. *Id.*

wearable or device.[65] As a result, the information they generate about the individuals using their products can be used by the app developer for any reason they want, and the information is not subject to HIPAA protections. For example, the app developer may want to use the information for product development or for marketing purposes, offering the individual specific advertisements according to their information and product usage. Furthermore, it is possible that the app developer has a financial agreement to share the information with a third-party data broker that will then use the information for reasons it does not have to disclose.[66] Another possibility receiving increasing attention in the wake of *Dobbs v. Jackson* is how data from period tracking apps may be disseminated and used.[67]

Because these apps, along with other features of a smart phone, hold such a vast amount of personal information, the Supreme Court of the United States has held that in the criminal context it is unconstitutional to search the contents stored on a person's phone without a search warrant.[68] Yet despite this recognition of the value of the information stored in these apps, in a non-criminal context, if the information is obtained legally, there is no constitutional limitations on the data's use.[69] Instead, the data is subject only to a patchwork of privacy laws depending on the source of the information.[70] Such is the case with information obtained by a third-party app that collects health information, leaving both consumers and collectors in the dark as to the data's potential uses.

c. Medical Device Manufacturers

An additional source of health information outside HIPAA protections comes from medical devices that providers use in health care facilities for treatment and diagnostics.[71] These

65. *Id.*
66. *Id.* at 4.
67. *Dobbs v. Jackson Women's Health Organization*, 142 S.Ct. 2228 (2022) (holding that there is no constitutional right to abortion, which gives states the ability to criminalize it).
68. Berman et al., *supra* note 1, at 24.
69. *Id.*
70. *Id.*
71. Tschider, supra note 35, at 140.

devices are dependent machine learning systems, which require data input prior to their use in order to ensure accuracy.[72]

Charlotte Tschider provides an example (concerning robotic surgery) in her article *AI's Legitimate Interest: Toward a Public Benefit Privacy Model*. The process begins with a patient receiving care from a primary care physician, who then refers the patient to a surgeon specially trained in robotic surgery.[73] In order to prepare for a safe and effective surgery, the specialist must input the patient's health information into the machine learning system of the surgical robot.[74] Additional safety measures require that the robot collect data while it is operating.[75] All this information serves to make the treatment as precise and safe as possible. Now, however, the company that manufactures that surgical robot also has access to the information that the robot collected. Yet, because this medical device manufacturer is outside the definition of the entities subject to HIPAA regulations, HIPAA does not protect the information.[76] It is likely that the treating physician obtained the patient's consent to share information with the robot prior to the treatment, but it is unlikely the patient either felt that a meaningful choice existed or understood the potential secondary uses of the collected information.[77] A later section will further discuss the flaws in current consent models.

IV. BALANCING PRIVACY WITH INNOVATION

A. BENEFITS

Artificial intelligence (AI) promises to deliver benefits that seem unimaginable. Implementing AI can achieve improved research methods and analyses, more effective and more targeted product development, increased efficiency in diagnoses and treatments, and generally higher quality care.[78] High quality AI depends on utilizing high quality data sets to power its self-learning processes. For an industry that faces constant

72. *Id.*
73. *Id.*
74. *Id.*
75. *Id.*
76. 45 C.F.R. § 160.103 (2013).
77. Katherine Stewart, Transference as a Means of Building Trust in World Wide Web Sites, ICIS 1999 PROCEEDINGS, PAPER 47, 460 (1999).
78. Tschider, *supra* note 68, at 141-143

criticism for its shortcomings, particularly when it comes to cost of and access to healthcare, a revolution of the kind AI promises is tantalizing. With the increased utilization of technological advancements, particularly ones that incorporate AI, the health sector has positioned itself to provide safer, more accurate, and cost-effective methods of diagnosis and treatment.

Moreover, as AI has found its way into nearly every part of the healthcare industry, operational processes also stand to benefit from AI advancements. Recent changes in the Stark and Anti-Kickback statutes, which govern financial arrangements with physicians, emphasize a focus on value-based healthcare as opposed to fee-for-service payment models.[79] A key characteristic of value-based care involves reducing the cost of care without reducing the quality of care.[80] AI is perfectly suited to help achieve this goal. Experts estimate that AI could "save as much as $71 billion annually through virtual nursing assistants, administrative workflow assistance, fraud detection, and dosage error reduction."[81] Moreover, accompanying AI-powered monitoring equipment "could save an additional $14 billion per year."[82] Organizations intending to implement AI do not necessarily intend to replace human workers. Rather, much of AI's utilization involves working with humans mainly by improving the accuracies and efficiencies of human job functions.[83]

For AI to achieve all of these desirable outcomes, it must continuously receive high quality data sets. As Charlotte Tschider has explained, AI's success is directly correlated to the quality of data it operates on.[84] AI utilizes algorithms that range from relatively simple to quite complex. Most commonly, AI operates on machine learning, which involves the use of large datasets to directly or indirectly create algorithms to deliver decisional conclusions.[85] Machine learning facilitates a self-

---

79. Mathew Larson et al., *Health Care Fraud*, 58 Am. Crim. L. Rev. 1073, 1009–10, 1014–15 (2021); *See* 42 C.F.R. §§ 411, 1001, 1003 (2020).

80. *See* Tschider, *supra* note 35, at 141 (describing value-based solutions with the equation "Value = Quality + Outcomes)/Cost").

81. *Id.* at 140.

82. *Id.*

83. *Cf. Id* at 141 (noting AI can decrease medical coding inaccuracies, increase healthcare provider efficiency, and reduce costly mistakes).

84. *Id.* at 136.

85. *Id.* at 132.

learning process that essentially is the basis for AI's remarkable capabilities. The more data that goes into the algorithms, the more accurate the decisional results. Additionally, self-learning processes in AI require diverse data sets to ensure that the algorithms are self-learning in a way that will allow them to treat any patient accurately and safely, not just someone from the particular demographic whose data the AI had access to.[86] As the decisional stakes increase, so does the need for data volume and data quality. Data quality generally increases as identifiability increases, although it does depend on context.[87] AI's need for data is not solely for improving each individual care experience, but also for improving societal health outcomes. Individuals may be willing to share health information with a higher degree of identifiability if they felt the benefits of doing so warranted relinquishing some of their own privacy rights. Yet a decision to lower one's privacy standards for the sake of technological gain needs to be one that is made with full understanding and confidence that despite relinquishing some privacy interests, the individual still has control over one's health information. Current privacy laws do not provide this level of autonomy.

B. DRAWBACKS

1. Risks

AI's need for large data sets creates a number of tensions.[88] One such tension is the associated risks involved with sharing health information. As accessible health information becomes more voluminous, it becomes increasingly more likely that the information could end up in unintended places.[89] Consequences of this misappropriation include companies marketing to individuals based on their medical histories or, more significantly, companies creating consumer profiles that allow employers, insurance companies, landlords, etc. to discriminate

86. *Id.* at 137–38.

87. *Id.* at 126-27.

88. Charlotte Tschider, *The Healthcare-Artificial Intelligence Impasse*, 36 SANTA CLARA COMPUTER & HIGH TECH. L. J. 439, 440 (2020).

89. April Moreno Arellano et al., *Privacy Policy and Technology in Biomedical Data Science*, 1 ANNUAL REV. OF BIOMEDICAL DATA SCIENCE 115, 119 (2018).

against individuals.[90] Consumer profiles allow for inferences that could lead to hidden discrimination in the event that employers, insurers, lenders, etc. use them to dictate economically important decisions.[91] Beyond individual discrimination, entities can start inferring things about a class of people as a whole, perhaps assigning different stereotypes to any individual within that class. For example, an entity might infer that a particular race is more susceptible to a certain disease. Insurance companies may surreptitiously rely on such a stereotype for individuals within that race, regardless of that individual's personal health record. This means of discrimination can even happen unintentionally when the insurance company is basing decisions off a facially neutral data point, but that facially neutral data point is only predictive of a certain outcome because of its correlation with a facially discriminatory data point.[92]

An additional risk includes the possibility of reidentification. Health information is often linked to financial and security information. If an entity can trace the information back to an individual's identity, not only are that person's health interests vulnerable to exploitation, but his financial and proprietary interests are as well.[93] Even "using various measures to deidentify health records," reidentification is still possible.[94] While one estimate indicates that if an entity fully complies with the HIPAA deidentification requirements, the rate of reidentification is 0.04%,[95] that rate increases substantially if the health data is cross-referenced against data in "voter registration records, hospital discharge records, commercially available databases," or other computer networking databases.[96]

Concerns also arise with potential downstream uses of health data. It is possible that a patient may consent to sharing

---

90. Scott R. Peppet, *Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security & Consent*, 93 TEX. L. REV. 85, 112 (2014).

91. *Id.*

92. Anya E.R. Prince & Daniel Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data*, 105 IOWA L. REV. 1257, 1270 (March 2020).

93. Rothstein, *supra* note 6, at 6.

94. *Id.* at 5.

95. *Id.* at 6.

96. *Id.* at 5–6.

her health information for improving robotic surgery techniques, while not realizing that her information could eventually be used for a purpose that does not align with her personal morals.[97] Imagine a patient consents for an entity to use her information for studying a particular genetic disorder. As a result, researchers eventually develop tests to detect the disorder while in-utero—which therefore prompts an increase in abortions, to which that patient religiously objects. The concern is that broad consent practices that do not have the capability to consider subjective interests can deprive an individual of true autonomy.

2. Loss of Privacy

A second tension that stems from AI's need for data is the tug of war between sharing data and valuing privacy. Privacy is a deeply rooted value, particularly in medicine. The eighth principle of the Hippocratic Oath states: "And about whatever I may see or hear in treatment, or even without treatment, in the life of human beings—things that should not ever be blurted out outside—I will remain silent, holding such things to be unutterable . . . "[98] Modern codes of ethics for physicians capture a similar sentiment.[99]

Beyond health care, privacy is important for individuals, relationships, and society.[100] Privacy is a form of autonomy in that it allows an individual to control what information about herself she wishes to share and with whom.[101] Furthermore, privacy allows individuals the freedom and ability to control the direction of their lives in that it provides spheres for people to form their own opinions, beliefs, interests, and decisions without worrying an outside party might be watching and casting judgment.[102]

Privacy in relationships also fosters the necessary component of trust,[103] and trust in relationships is *critically* important for a successful health care system. Without trust, it

---

97. *Id*. at 8. ("Currently, individuals have no control over the use of their deidentified information and specimens.")

98. HELEN F. NISSENBAUM, PRIVACY IN CONTEXT 172–73 ((Stanford ed., Stanford University Press 2010)

99. *Id*. at 173.

100. *Id*. at 81-87.

101. *Id*. at 81.

102. *Id*.

103. *Id*. at 84-85.

is unlikely there will be candid conversation in a patient/physician relationship, which in turn reduces the likelihood that the patient will receive the care he needs. Moreover, a lack of trust may prevent someone from seeking care in the first place.[104] While information shared within relationships vary according to the type of relationship,[105] patient/physician relationships are ones in which the patient shares information that he is not likely to share in other relationships. That only further emphasizes the need for trust to serve as the foundation of the patient/physician relationships.

Beyond importance of privacy in relationships, for AI's incorporation into the health sector to be successful, privacy models need to account for privacy as a societal interest. People need privacy for individual development, but this holds true for cultural development as well.[106] Cultural development depends on creativity and innovation, and freedom to think without fear of judgment is critical for an atmosphere that fosters creativity. To stunt individual creativity is to stunt societal progress.[107] Some of the most complex health care problems require a tremendous amount of creativity—a level of creativity that AI can facilitate. As mentioned previously, it is possible that some people would be willing to sacrifice their individual privacy interests to further society's interest in solving these complex problems. It is logical, then, that the successful incorporation of AI into the health sector requires guidelines that account for both individual interests and societal interests. Furthermore, these guidelines must provide enough flexibility to balance the two according to the relevant context. Unfortunately, current privacy laws lack this ability.

V. CURRENT PRIVACY MODELS FOR HEALTH CARE DATA

A. HIPAA

HIPAA is the chief federal privacy law that pertains to the privacy of health information.[108] Initially, the Health Insurance

104. Rothstein, *supra* note 6, at 7–8.
105. Nissenbaum, *supra* note 98, at 85.
106. Solove, *supra* note 7, at 1892.
107. *Id.*
108. Tschider, *supra* note 35, at 148.

Portability and Accountability Act mainly addressed improving the design and sale of health insurance, and also included provisions regarding the electronic processing of insurance claims and other types of medical information. As the need for improved privacy protections increased, HIPAA's privacy regulations have evolved. Parts 160 and 164 of the Code of Federal Regulations contain the key privacy provisions.[109]

These privacy regulations—known collectively as "the Privacy Rule"—generally require certain "covered entities" to maintain the confidentiality of defined types of "protected health information" (PHI). Part 160 defines PHI as "individually identifiable health information" maintained or transmitted in electronic media or any other form or medium.[110] Under HIPAA, a covered entity includes: "a health plan, a health care clearinghouse, and a health care provider who transmits any health information in electric form . . . "[111] In turn, a "business associate" is a person, business, or organization who, while not an employee of a covered entity, "creates, receives, maintains, or transmits protected health information for a function or activity . . . including claims processing or administration . . . or . . . provides" services such as legal, accounting, or management services on behalf of the covered entity.[112] The Privacy Rule prohibits covered entities from using or disclosing PHI unless one of the exceptions in part 164 apply. These exceptions include:

109. MARK A. HALL ET AL., HEALTH CARE LAW & ETHICS 124-27 (Rachel E. Barkow et al. eds., 9th ed. 2018).

110. *Id.*

111. 45 C.F.R. § 160.103 (2013).

112. *Id.*

1. PHI can be disclosed to the individual or the individual's personal representative

2. PHI can be used for treatment, payment, or health care operations.

3. PHI can be disclosed where the entity receives a more specific valid authorization

4. A specified subset of PHI can be disclosed without written authorization in certain situations after giving the patient an opportunity to object.

5. An individual's PHI can be disclosed without his or her authorization in special circumstances (e.g. where the disclosure is required by law).

6. A limited data set that excludes most identifying information can be disclosed for use in public health, research, and operations.

7. Covered entities are also permitted to disclose PHI incident to uses or disclosures otherwise permitted or required so long as the covered entity has followed the standards governing the minimum necessary disclosure of information and has put in place proper administrative, physical, and technical safeguards.[113]

In the event that a covered entity does transmit PHI, it must adhere to the "minimum necessary" standard.[114] Under this standard, dissemination is limited to the amount necessary to accomplish the intended purpose of the use, disclosure, or request.[115]

Additionally, HIPAA establishes a standard for when the health information is considered deidentified, at which point it loses its legal protections and can be disseminated for secondary purposes.[116] Health information is considered deidentified when there is "no reasonable basis" for believing that the patient can

113. HALL ET AL., *supra* note 109.
114. 45 CFR § 164.502 (2013).
115. *Id.*
116. *Id.*

be identified from the data.[117] Organizations can meet this standard through two different methods: expert determination or safe harbor.[118] Expert determination entails trained experts using statistical methods to determine that there would be a "very small risk" of reidentification.[119] The safe harbor method is the more commonly used method and entails the removal of 18 types of patient identifiers. Limited data sets only require removing sixteen of those eighteen types to achieve deidentified status.[120] Notably, there are no technological methods with the capability to deidentify free form text.[121]

The third exception stated above (the valid authorization exception) is relevant for secondary uses of identifiable information. In order to obtain consent from the patient for use of her identifiable health information, an organization must provide details regarding specifying third party interests and involvement, intended use of the data, and timeline of its use.[122] While consent may seem to offer sufficient privacy protections, current privacy models are expecting consent, and the individual, to do too much.[123]

Consent is often thought of as a means for providing the individual with a choice, however, for a number of reasons it falls short of legitimately granting individual autonomy.[124] The first reason is that individuals rarely truly understand what they are consenting to and what implications of that consent will follow.[125] Several factors can lead to a failure to understand. People often do not even read the form in front of them, particularly when an entity presents it in an electronic format (as often is the case in this digital era, with ever-changing privacy policies and terms of agreement).[126] Furthermore, it is unlikely an individual who does read the long, complex form would truly understand the terms.[127] Even if an entity makes an

117. Moreno Arellano et al., *supra* note 16, at 117.
118. *Id.*
119. *Id.*
120. Rothstein, *supra* note 6, at 4.
121. *Id.*
122. Tschider, *supra* note 35, at 148.
123. Solove, *supra* note 35, at 1884.
124. Tschider, *supra* note 13, at 1517.
125. Solove, *supra* note 7, at 1888-89.
126. Tschider, *supra* note 125, at 155.
127. Tschider, *supra* note 13, at 1522.

effort to write a consent form with plain language in the hopes of making it easy to understand, that often means that important information is left out or distilled to the point it is not completely accurate.[128]

A second reason why consent practices fall short of advancing individual autonomy is that they fail to provide meaningful choice.[129] When it comes to health care, individuals are usually giving consent to a party with much more bargaining power under a take it or leave it circumstance.[130] Providers or entities often request consent shortly before a course of treatment, diagnostic test, or use of a product or app. At this moment a patient is unlikely to decline and start all over with a search for an alternative provider. Additionally, consent for *treatment* accompanies consent for a *secondary use*, which can lead a patient to believe there is no option to refuse the secondary use without refusing the care he needs.[131]

It is certainly possible, though, that if the patient were to get an accurate and thorough description of the secondary use, he would support that use and willingly consent. But this illustrates the need for consent practices that vary according to the manner in which a party is seeking to obtain the consent, and the interest behind acquiring health data for a secondary purpose. For example, in the event that an entity is requesting a broad consent to store data for uses not yet known, it is critical that the individual understands any consequences of long-term data storage, such as (1) the data's eventual transfer to other parties, (2) whether or not the entity will notify the individual of any new party who has access to his information, and (3) if so, what this new party intends to use the information for. Without an understanding of short-term and long-term risks and benefits, an individual cannot truly make an informed decision and therefore lacks autonomy. And relying on consent in situations where the patient might understand the terms of agreement, but feels it is difficult to refuse, vitiates if not destroys autonomy.

---

128. *Id*.
129. *Id*. at 1519.
130. *Id*.
131. *Id*.

B. OTHER RULES

1. The Common Rule

The Department of Health and Human Services (HHS) established the Common Rule, which applies to all federally funded research. The aim of the Common Rule is to facilitate research while implementing steps to protect the human subjects.[132] Similar to HIPAA, the Common Rule offers legal protection to identifiable information but eliminates those legal protections once the information is considered deidentified. In contrast to HIPAA, the Common Rule sets a lower standard for deidentification stating that information or specimens "are not individually identifiable when they cannot be linked to specific individuals by the investigators either directly or through coding systems."[133] For the latter, the investigator must not be able to "readily ascertain" the identity of an individual through coding systems due to a number of possibilities including: if the key has been destroyed before the research begins, if the keyholder has agreed not to release the key to investigators under any circumstances, or if there are Internal Review Board approved policies or other legal requirements prohibiting the release of the key until the individuals are deceased.[134]

In addition to its own deidentification standard, the Common Rule also has its own consent requirements, though the aforementioned shortcomings of consent still persist. Under the Common Rule, consent forms must be publicly available for 60 days before the start of the clinical trial and must contain a clear and concise description of the research. The rule does allow for obtaining broad consent that covers current and future uses without specifying the future uses.[135] HHS also included exemptions from the Common Rule regulations including low risk studies, use of identifiable but publicly available data, and using previously obtained information if the subject gave broad, long-term consent.[136] Again, the privacy model here hinges on consent from the individual which places an expectation of

132. Rothstein, *supra* note 6, at 4.
133. *Id.* at 3.
134. *Id.*
135. *Id.*
136. *Id.*

information self-management on the individual not conducive for the digital health era.

2. Federal Trade and Commission Privacy Guidelines

As mentioned above, HIPAA definitions of covered entities and business associates limits the entities subject to its regulations. For all sources of health information that are outside of HIPAA's reach, the Federal Trade Commission (FTC) comes in.[137] Unfortunately, the FTC's privacy models operate under the same notice and consent regime discussed above, placing the onus on the individual to advocate for their privacy interests according to what they may or may not read or understand in an informed consent form.[138] While the FTC does have the ability to enforce more strict guidelines for fairness and deception as it pertains to privacy, however, it does so unpredictably.[139] Most of the FTC's guidelines are not specifically geared to health data, and even the ones that are, such as the Fair Information Practices (FIPs) and health app guidance, are non-binding.[140] Even if the guidance was binding, it pertains only to what health apps with certain types of health data must do in the event of a data breach. The guidance does not address trade practices of the health information.[141] Ultimately, FTC regulation is afflicted with the same inadequacies of a privacy self-management system as discussed above, leaving individuals without a meaningful choice and without a true understanding of the ways in which their information could be used.

VI. PRIVACY MODEL PROPOSAL

Realizing any of the potential benefits that digitized health care data presents depends upon a foundation of individual trust in the system. If individuals do not believe that the people who they entrust their intimate health information to will adequately protect it, they will be hesitant to share that information even

137. Tschider, supra note 35, at 150.

138. Solove, *supra* note 7, at 1882.

139. Tschider, *supra* note 35, at 161.

140. Tschider, *supra* note 35, at 162.

141. Press Release, Fed. Trade Comm'n, FTC Warns Health Apps and Connected Device Companies to Comply with Health Breach Notification Rule (Sept. 15, 2021) https://www.ftc.gov/news-events/press-releases/2021/09/ftc-warns-health-apps-connected-device-companies-comply-health.

for their own care and treatment needs, let alone for secondary purposes that may not directly benefit them. In order to establish this trust, individual autonomy must be the priority in any privacy model.

Currently, autonomy in health information privacy decisions hinges on privacy models that are too generalized and outdated for the digital health era. This places too much responsibility on the individuals and leaves them without an actual choice in what and how much information they share. The following is a proposal for a federal hierarchal framework that will apply to all entities who deal with health data.

1. Health care data shall only be collected and used with the individual's permission and fully informed consent.

2. The collections, disclosures, and dissemination of health care data may be allowed only where such action furthers a legitimate societal interest and does not interfere with #1.

3. Health care data may be used to benefit the collector, discloser, or disseminator of such data so long as it does not interfere with #1 and #2.

a. Health care data shall only be collected and used with the individual's permission and full informed consent

Individual autonomy in the health care industry means the individual controls what happens with their health information according to their subjective interests and privacy values. Ensuring the privacy interests of the individual fosters the sense of trust that is critical for the success of the health care system's primary and secondary interests.

Some people may enjoy receiving customized advertisements; others may feel such advertisements violate their privacy. Either way, the decision should be up to the individual. The individual needs to be able to make this decision with trust that entities will use their information according to their wishes only. They also need to be able to make this choice from a place of control, i.e. feeling like they have an actual choice in the matter. Current privacy models cannot account for these subjective preferences and foster a sense of trust or autonomy. If individuals trust the entity with which they are sharing their information, they will share information honestly and completely. Ultimately, this promotes not only their own quality

of care but builds the foundation for any downstream innovative uses as well.

b. The collections, disclosures, and dissemination of health care data may be allowed only where such action furthers a legitimate societal interest and does not interfere with #1

Innovative use of digitized health data has the ability to reform the health care industry in numerous ways, offering benefits both at the individual and the societal level. Societal benefits, however, are dependent upon individual trust in the system. It is entirely possible, for example, that people would be more than willing to share identifiable information if it led to a real possibility that researchers could finally cure a devastating disease such as Alzheimer's. In order for this to work in a way that does not infringe upon an individual's interests, the individual needs a thorough understanding as to the benefits of the secondary use as well as the risks involved. In accordance with HIPAA's data minimization principle, establishing this understanding would require an explanation of all the information to which a third party would have access and the degree of identifiability necessary for the intended use. Requiring a comprehensive explanation of both the benefits and the risks ensures that the secondary use serves a purpose that the individual finds worthy enough to warrant lessening their personal privacy protections. Without a thorough understanding of pros and cons, the individual cannot weigh them in a way that drives an autonomous decision.

Ensuring the individual's thorough understanding could create a burden on the part of the party wishing to obtain the data, but this added burden puts more responsibility on the party seeking the information to do their own weighing of pros and cons regarding their interests.[142] It is a good thing, however, for the entity to carefully consider how it wants to ensure that it meets the individual's needs in a way that also allows it to meet its own. Current privacy models place all this responsibility on the individual, yet the data collectors are better suited to consider the costs and benefits in a meaningful way and adjust accordingly. Adequate patient and consumer education is critical in guaranteeing that even in the event an individual's health information aids in achieving a more attenuated benefit, that

142. Tschider, *supra* note 35 at 178.

benefit is still within the individual's interest and therefore does not violate #1 of this proposed privacy model. Utilizing this framework would require prescriptive guidelines regarding the appropriate degree of minimization, identifiability, and notice/informed consent. These guidelines will vary according to the context. More downstream effects and uses involved, or ambiguous uses, would entitle the individual to receive a more granular level of information about the potential uses and the types of data they require.

c. Health care data may be used to benefit the collector, discloser, or disseminator of such data so long as it does not interfere with #1 and #2.

Individual preferences regarding collecting, sharing, and using their health information could be (1) no notice or consent needed, (2) consent but no notice needed, (3) notice and consent needed before any action commences, or (4) somewhere else along that spectrum.[143] Some individuals, for example, may be willing to relinquish health information to provide the collector or disseminator with some kind of purely internal benefit to the collector or disseminator. Others might not.

In order for entities to account for these differences, it is up to the government to put forth prescriptive rules that specify which contexts and uses require what kind of protections for collection, use, and transmission of health information.[144] Data uses intended only for an internal benefit, such as to improve marketing strategies, would be permissible as long as the person or entity obtaining consent to use the data made those intentions clear to the individual. This kind of secondary use would require full disclosure of the type of data and its degree of identifiability the entity is requesting. Further, the entity would need to completely separate the processes used to obtain consent for an internally beneficial secondary use and consent for the treatment or use of the drug or device. Other secondary uses, such as product development, may blur the lines of an internal benefit and a legitimate societal interest, illustrating the need for governmental clarifications as to what may qualify as a legitimate societal interest. Preventing coercive or underhanded acquisition of individual's health information through increased

143. Rothstein, *supra* note 6, at 9—10.
144. Arellano, *supra* note 16, at 124.

disclosure and consent requirements ensures trust in the health care system. This is particularly necessary for uses that will benefit neither the individual, nor society at large.

For any given use of health data, an entity must consider and determine what types of health data it needs to meet the anticipated purpose. From there, the necessary degree of identifiability can be determined. As the degree of identifiability increases, the entity must utilize more stringent privacy protections. Similarly, as the degree of medical necessity increases, the entity must ensure that it does not condition treatment on surrendering health information. Having these prescriptive rules within a flexible framework is crucial for innovators to achieve their goals without infringing on individual interests. Knowing what kind of protective measures are required will alleviate some of the burden and confusion on the part of innovators. Furthermore, it ensures that organizations are using data efficiently and in a way that promotes cultural development and prioritizes individual autonomy. Innovation without proper privacy protections is of no benefit at all if it comes at the cost of trust in the system. Individual privacy interests vary, as do opinions on the types of secondary uses that warrant relaxing those privacy interests. Privacy models regarding health care data need to have the ability to adapt to these differences to promote societal development in a way that prioritizes and adapts to individual interests, protecting the primary purposes of the health care industry and facilitating its improvement.

VII. CONCLUSION

A flexible privacy framework that prioritizes individual autonomy and allows practices to adapt to different contexts and interests will both reduce frustration related to privacy of health information and encourage technological development. Current laws are not capable of striking this balance. Many sources of health information are largely unregulated and even the regulations in place depend on consent practices that fall short of truly protecting privacy interests and advancing individual autonomy. Digitized health data holds tremendous innovative potential to both improve individual care experiences as well as achieve larger health care goals, such as decreased costs and increased accessibility. However, if law and policy makers do not harness this potential in a way that preserves individual trust

none of these potential benefits will come to fruition and the health system will crumble. Privacy models need to allow innovators to create revolutionary solutions without also allowing entities to violate privacy interests and create new avenues for discrimination.[145]

145. Portions of this article appeared on the *Indiana Health Law Review Blog* on Jan. 13, 2023.