Table of Contents

THE DICHOTOMY BETWEEN SAFEGUARDING DATA PRIVACY AND PROMOTING INDIVIDUALIZED HEALTHCARE USING ARTIFICIAL INTELLIGENCE: HOW MODERN REGULATIONS MISS THE MARK AND NEW STANDARDS CAN RECONCILE THE TWO

Mohammad Reza Kameli

I. Introduction

As a “testing ground” for determining the potential of Artificial Intelligence (AI) to improve healthcare, Israel, through a governmental initiative, plans to bring together millions of individuals’ private data into one integrated system.[i] According to government officials, the goal is “to make health care less expensive, more effective and better tailored to individuals everywhere.” This decision was primarily a response to the need for an amalgam of data—the bigger they are in volume, the better—that Computer Data Support Systems (CDSS) rely on in conducting predictive analytics.[ii] Predictive analytics refers to the utilization of electronic algorithms that can anticipate the course of medical events in real-time[iii]. It is this unique tool that has made the incorporation of AI into the healthcare industry very appealing and, at the same time, controversial. For Israeli researchers, software developers, and healthcare professionals, utilizing AI data can be beneficial, potentially allowing for more tailored medicine, and the ability to read diagnostic materials more accurately; however, these advances also come with both ethical and legal challenges regarding the handling of individuals’ private data, particularly those pertaining to a patient’s health, medical history, and medical conditions.[iv]

This paper intends to discuss the most crucial factor in the productive development of CDSS and, by extension, individualized healthcare that countries around the world, including the U.S., need to consider: the demand for a constant flow of broad and diverse collections of patient and non-patient data into these systems. The majority of this paper strives to address the legal concerns within existing legal frameworks concerning data privacy and the protection of patients’ medical records. Lastly, this paper provides an outlook on possible solutions that can balance seemingly conflicting pillars of the nascent structure of AI-based healthcare.

II. AI's Role in Healthcare

In healthcare, AI concentrates on improving outcomes through analyzing both consumer and patient data for providing diagnoses, stimulating medical R&D, and translating medical device images into reliable health data[v], a significant portion of which is conducted via CDSS. Currently, the refined and synthesized product of these sophisticated processes is translated in the field of medical practice into suggestions to physicians, surgeons, and healthcare providers that help them verify the variant set of risks for thousands of patients separately and accurately. In addition, CDSS can update their evaluations every second, taking into account real-time data regarding the internal and external factors affecting individuals, and they come up with proposals as to what procedures and forms of care are best suited for each patient.[vi]

Now, for the operations, as mentioned above, to take place reliably and perpetually, the AI-based algorithms that CDSS rely upon must receive updated factual data from patients and non-patients alike. These data points can come from various sources such as medical literature, insurance-claims data, electronic health records, clinical trials, and information recorded via individuals' smart devices regarding their health conditions, physical activities, and daily routines.[vii] That said, one of the challenges to the integrity and robustness of CDSS is the possible bias of the amassed data that subsequently flows into these systems, which could be a consequence of limited sources from which data are collected as well as their low volumes.[viii] This will, in turn, jeopardize every piece of medical advice and diagnoses that these machines come up with.

With such an understanding of the importance of collecting diverse, broad, and comprehensive data for the successful operation of CDSS, it becomes crucial to look into the relevant privacy issues. When it comes to predictive analytics, data privacy concerns arise in at least two areas: (1) gathering vast amounts of data to develop the necessary algorithms, and (2) sharing the data to conduct oversight.[ix] These concerns are heightened when the information comprises an individual's "health data," which generally refers to information about individuals' physical or mental health conditions as well as payment for and the delivery of health services to them.[x]

One of the problems that may arise in the context of dealing with health data is the potential inability of CDSS suppliers to ensure their safe handling. This is, in part, a result of the nature of the work that AI companies perform, which stands in contrast to that of healthcare providers. That is, while healthcare providers are bound by high ethical and legal standards of care regarding the safeguarding of individuals' data, AI companies dedicate their resources to intra-company collaboration and R&D to enhance the algorithms upon which their software rely.[xi] As such, the vulnerabilities and risks associated with the open nature of CDSS suppliers' work include but are not limited to the potential breach of patients' privacy, misuse of the systems, and even errors of analysis.[xii] For example, the value of an individual's medical records on the black market is ten times the value of their credit card information.[xiii] These factors stress the importance of understanding the current legal frameworks AI companies must operate within.

III. Current Regulatory Frameworks Governing AI in Healthcare

In the past several years, state and national legislatures worldwide have adopted different sets of laws to promote data privacy, accountability for the use of private data by companies, and procedures to ensure the consent of individual citizens whose information is made subject to data analysis. For example, data protection in the European Union (EU) is currently governed by the General Data Protection Regulation (GDPR), which went into effect on May 25, 2018.[xiv] Requirements for consent are more stringent under the GDPR.[xv] For instance, Article 6 requires the "explicit consent" of data subjects before their health information can be processed.[xvi] Additionally, Article 6 requires that the processing of patient data be "necessary for the purposes of preventive or occupational medicine."[xvii] Also, as part of gaining consent, the regulation mandates healthcare providers and CDSS-manufacturing companies to describe to patients the chain of processes their data will undergo, and all the entities that will have access to the patients' information.[xviii]

Furthermore, as part of its efforts to foster accountability amongst organizations handling health data, Article 23 of the GDPR requires these entities to keep detailed records of the purpose of their predictive analytics operations.[xix] According to the GDPR, this standard is primarily enforced when these organizations deal with private patient data that could put their freedom and privacy in jeopardy.[xx]

Alternatively, the United States is primarily reliant on the Health Insurance Portability and Accountability Act (HIPAA) of 1996 as the central source of compliance.[xxi] Any technology company, including manufacturers of predictive analytic tools that render services to healthcare providers, are regulated by the rules of the Act.[xxii] The legal umbrella extends over these companies when they gain access to patient health data that is individually identifiable.[xxiii] In other words, HIPAA allows the use of identity redacted patient data without consent.

It is vital to note that within the United States, there is no single set of uniform law aimed at regulating privacy data.[xxiv] Many of the specific and more stringent requirements encompassed in the GDPR, namely detailed consent-seeking mechanisms, are lacking within HIPAA. In response, the Food and Drug Administration (FDA) has, in the past few years, started to think about ways to develop more sophisticated regulations that are in keeping with the growing market share of AI-based data processing tools within the healthcare industry.[xxv] Notwithstanding its intents, the FDA faces legal obstacles in achieving the aforementioned goals with regards to medical support tools, for its jurisdiction is constrained to medical devices. This is particularly an issue considering that the 21st Century Cures Act, enacted in 2016, legislatively exempts certain software from the definition of a medical device.[xxvi] These forms of software fall under general categories such as administrative support, general maintenance, and electronic patient record systems.[xxvii] However, there is much ambiguity as to which specific types of CDSS fall under these classifications.

That being said, the law must strike a balance between protecting patients' privacy rights and supporting the development of the positive life-altering by-products that AI systems can bring about. For example, the stringent consent requirements within the GDPR are inconsistent with the nature of CDSS programs. AI software is designed to take in data, perform algorithmic operations on that data, and then produce various types of analyses. AI software simultaneously learns from the variations in its input, allowing AI to understand the inconsistencies between its input and output, and process the errors that previously existed in its output in such a way that the AI can evolve its internal algorithms almost completely autonomously. This makes such systems operate like a completely contained black-box, and it becomes infinitely more challenging for a human observer to track the specific changes that the initial algorithms of the AI have undergone.[xxviii] It is the ever-evolving black-box feature of CDSS and other AI software that makes them so powerful, yet it is this same characteristic that renders it almost antithetical for a software or healthcare provider to ask a data subject to consent to the specific processes that their health data will be a part of, as is currently required by the GDPR. In other words, as these predictive analytics systems grow and become more individualized, they inevitably become more complex. It then becomes nearly impossible for AI developers to explain every operation that will be performed on a patient's data, as required by the GDPR. This limitation will disincentivize developers from creating more powerful systems that could potentially perform high-quality disease prevention, diagnosis, prognosis, and treatment.

Furthermore, article 23 of the GDPR requires that data processing centers keep thorough records of the purposes of the operations performed on private data.[xxix] For the same reasons mentioned above which go back to the black-box nature of AI software, it will not always be known why a CDSS focuses on a particular set of data out of all its input or for what reasons it processes them through the algorithms that it uses. Thus, again there is an incongruence in this law and other similar regulations adopted by different regulatory bodies around the world. Enforcing these requirements will, therefore, primarily serve as an impediment to AI technological progress in healthcare, rather than a means to promoting citizens' security.

IV. Looking Ahead: Possible Legislative Approaches

Privacy concerns surrounding AI must be balanced in an even-handed way, which will help lead to coherent policy and regulatory decisions in addressing the incorporation of AI into healthcare. As a prominent legal practitioner in the field of technology, Yavar Bathaee, notes, the law "is built on legal doctrines that are focused on human conduct, which when applied to AI, may not function."[xxx] In order to make a shift from the current state of affairs towards laws and regulations that can serve as enablers for AI software developers and healthcare providers, there are a number of contemporaneous steps that can be taken. These steps will provide a combination of improved health data safeguards, uniform systems of data sharing, and shared liability doctrines as deterrents that will play an important role in addressing the data privacy concerns that citizens and governments share.

First, the security of patient health data can be enhanced, through laws and regulations that require the adoption of and adherence to certified data procurement, handling, and storage procedures that can reduce the risks that data processing companies are prone to. For example, the FDA emphasizes that data processing companies should detect and monitor cybersecurity vulnerabilities in their infrastructure as part of their customer support for clients in the healthcare industry.[xxxi] Under this new guise, computer scientists at John Hopkins developed a new security management system and were able to detect breaches within 5 minutes, compared to 75 minutes with their old system[xxxii].

This new system identified more "false positives," leaving time for detecting and addressing actual vulnerabilities in the systems. The transition to this new enhanced security management system was not only conducive of higher confidence in the privacy of the stored data, but also higher efficiency in terms of resource allocation.[xxxiii] This indicates that a significant increase in patient data safety can be achieved through enhanced governmental guidance.

Furthermore, one of the primary areas of concern about health data processing is the sharing of confidential data. AI companies often share data with one another or with healthcare providers to improve their algorithms, complement their databases, and render the services that their software is designed for. This is seen as potentially adding a layer of vulnerability to the ability of each of these entities to ensure the safekeeping of these data. One of the ways to address this concern is to set clear ground rules for interoperability amongst the systems that they use and require that they be adhered to through enforceable regulations.[xxxiv] That is, if these parties are required to work together to ensure that their systems can operate in tandem, the risk of breaches in data are significantly reduced. Reducing the risk is achieved by eliminating the role of middlemen and external data storage entities that are particularly prone and often vulnerable to cyberattacks. An additional step that these companies can take is to effectively redact the identity associated with data prior to sharing it. This can be done through "encryption and anonymization protocols that could be updated to combat the threat of machine learning re-identification."[xxxv] Aggregating the data before sharing them, reinforces these efforts.[xxxvi]

Finally, legislators and regulators can make concerted efforts to effectively incentivize data-processing companies and healthcare providers that utilize CDSS to enhance their standards and willingly engage in adhering to the aforementioned initiatives. One of the ways that such a goal can be achieved is by imposing shared liability on the parties that are part of the chain of collecting, processing, handling, and using patient health data. In other words, such a shared liability imposition will encourage these companies to be sensitive to each other's systems and procedures. This will encourage them to work together to improve the interoperability and uniformity of their security protocols, thereby ensuring confidentiality of the shared data. It will also make them think twice before working with companies or healthcare providers that have weak, obsolete, or non-conforming data-security systems.

V. <u>Conclusion</u>

Israel serves as an example of how, through uniformity of operations, safe and reliable data-centric collaboration amongst AI software manufacturers and healthcare providers becomes possible. The potential for achieving a sophisticated individualized healthcare system using AI is too important of a prospect to simply forego or impede for, albeit rightfully placed, legal and ethical concerns regarding privacy. A balance between privacy rights and protection and improving healthcare must be found within the legal system. Challenges that prevent the adoption of AI into healthcare need to be addressed to prevent harm to citizens and maintaining popular trust in these ventures. As was mentioned, only through a broad and multi-faceted data-collection mechanism can CDSS truly evolve into reliable and, in fact, effective tools. This can be accomplished only through the joint effort of all the entities within the chain of processing, storage, and use of patient-health data. Therefore, laws and regulations requiring these entities to adopt certified uniform security protocols is a step in that direction.

Additionally, by imposing shared liability in the handling of patient health data, governments can persuade the relevant actors to serve not only as checks on one another but also as models. That is, they will promote methods that comply with the base-line regulatory requirements and dissuade partners from adopting or continuing to use obsolete, non-conforming security and data-sharing platforms and protocols.

[i] Dov Lieber*, Israel Prepares to Unleash AI on Health Care,* Wall St. J. (Sept. 15, 2019, 10:01 PM), https://www.wsj.com/articles/israel-prepares-to-unleash-ai-on-health-care-11568599261.

[ii] *See* I. Glenn Cohen et al., *The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care*, 33 Health Affairs 1139, 1141 (Jul. 2014).

[iii] *See id*. at 1139-1147.

[iv] *Id.*

[v] Andrea Kulkarni, *AI in Healthcare: Data Privacy and Ethics Concerns* – Lexalytics (Nov. 12, 2019), https://www.lexalytics.com/lexablog/ai-healthcare-data-privacy-issues.

[vi] *See* I. Glenn Cohen et al., *The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care*, 33 Health Affairs, 1139 (2014).

[vii] W. Nicholson Price II, *Artificial Intelligence in Health Care: Applications and Legal Issues*, SciTech Law., Fall 2017, at 10.

[viii] Mariana Alpalhao Goncalves, *Liability arising from the use of Artificial Intelligence for the purposes of medical diagnosis and choice of treatment: who should be held liable in the event of damage to health?* (Tilburg University 2018).

[ix] See Nicholson, *supra* note 5.

[x] Ciera Logan, *Data Privacy and Bias Concerns in AI Health Tech*, HIPAA & Health Information Technology HIPAA & Health Information Technology (Oct. 3, 2019), https://hipaahealthlaw.foxrothschild.com/2019/10/articles/hit-health-information-technol/data-privacy-and-bias-concerns-in-ai-health-tech/.

[xi] Colin Mitchell & Corrette Ploem, *Legal challenges for the implementation of advanced clinical digital decision support systems in Europe*, 3 (Suppl 3) J. of Clinical and Transnat'l Res. 424, 424 - 430 (2018).

[xii] *Id.*at 425.

[xiii] *See How Artificial Intelligence Can Overcome Healthcare Data Security Challenges and Improve Patient Trust*, Health Catalyst (Sept. 18, 2019), https://www.healthcatalyst.com/insights/improving-healthcare-data-security-with-AI.

[xiv] Lincoln Tsang et al., *The Impact of Artificial Intelligence on Medical Innovation in the European Union and United States*, 29 Intell. Prop. & Tech. L. J. (2017).

[xv] Regulation of the European Parliament and of the Council (EU) 679/2016 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/.

[xvi] GDPR art. 6.

[xvii] *Id.*

[xviii] GDPR art. 22.

[xix] GDPR art. 23.
[xx] *Id.*
[xxi] *See* Lincoln Tsang, *supra* note 12.
[xxii] *Id.*
[xxiii] *Id.*
[xxiv] *Id.*
[xxv] *New FDA Guidance Clarifies Exemptions for Digital Health Software*, Akin Gump Strauss Hauer & Feld LLP (Jan. 2, 2018), https://www.akingump.com/en/news-insights/new-fda-guidance-clarifies-exemptions-for-digital-health.html (last visited Oct. 7, 2019).
[xxvi] *Id.*
[xxvii] *See* Lincoln Tsang, *supra* note 12.
[xxviii] Jocelyn Paulley, *What are the data protection challenges of using AI in healthcare?* Open Access Government (2019), https://www.openaccessgovernment.org/data-protection-ai-in-healthcare/67922/.
[xxix] GDPR art. 23.[xxx] Hannah Sullivan & Scott Schewikart, *Are Current Tort Liability Doctrines Adequate for Addressing Injury Caused by AI?*, 21 AMA J. Ethics., 160, 160-166 (2019).
[xxxi] *See* Lincoln Tsang, *supra* note 12.
[xxxii] *See How Artificial Intelligence Can Overcome Healthcare Data Security Challenges and Improve Patient Trust*, Health Catalyst, *supra* note 11.
[xxxiii] *Id.*
[xxxiv] Jessica Davis, *Healthcare Needs More than HIPAA, Legislation to Improve Security*, HealthITSecurity (2019), https://healthitsecurity.com/news/healthcare-needs-more-than-hipaa-legislation-to-improve-security.
[xxxv] Jordan Harrod*, Health Data Privacy: Updating HIPAA to match today's technology challenges*, Science in the News (2019), http://sitn.hms.harvard.edu/flash/2019/health-data-privacy/.
[xxxvi] *See* Sullivan, *supra* note 26.