Privacy and Security of Electronic Health Records

Brian P. Hogan

IST618 – Information Policy

Syracuse University

2020

Author Note

Paper prepared for Information Policy taught by Professor Ian MacInnes, Ph.D.

Background

Ensuring the safety, privacy, and security of patient Electronic Health Records (EHR) is fractured when updated sensitive data is iterated across networks. Sophisticated machine learning algorithms challenge data security algorithms by their ability to assemble scrambled data pressuring governments and the private sector to legislate confidentiality loopholes. The system is challenged to provide an increasingly cost effective, data mining resilient platform to build confidence in standards of care and patient confidentiality. There are clear pros and cons managing health information but the system is wrought with ethical dilemmas caused by data discrepancy mishaps. The following reviews policy adjustments to broaden EHR's transparency amongst patients, companies, agencies, and governments responsible for implementing EHR accountability best practices.

Advantages and Disadvantages of Electronic Health Record Information

The American Reinvestment and Recovery Act of 2009 (ARRA) stimulus package responded to the American Great Recession by providing a set of federal statutes to save jobs, allocate $19 billion in resources, and substantially augment patient heath record information management systems. According to USFHealth, a key ARRA advantage was the creation of a "meaningful use" healthcare digital records addressing security, patient privacy, and data quality (USFHealth, 2020). Quality medical data "improves epidemic surveillance, decreases the length of patient stay, achieves work efficiency by reducing nonvalue added activities…and helps to make for timely decisions at reduced cost" (Adane, Gizachew, and Kendie, 2019, pg. 67). It also facilitates large scale medical research from sanitized EHR structured data (Cowie, et al., 2017). An ultimate advantage of ARRA was viewed as achieving interoperability amongst healthcare institutions to facilitate "medical staff's understanding of the disease, diagnosis, and decision-making process" (Adane, et al., 2019, pg. 70). Large scale medical research also casts the advantage of applying Natural language processing (NLP) to generate new phenotypes from free text note fields by pairing new disease vectors with medical codes.

Disadvantages of EHRs center on protected health data. While great efforts have been invested in securing medical information there is no single solution or automatic classifier

available to administer this complex process across EHR subtypes. When large, or very large, EHR corpuses are accumulated adversarial machine learning has been found effective de-anonymizing data by re-identifying owners when combined with public records spawning unforeseen privacy violations. According to Friedrich, Kohn, Wiedemann, and Beimann's research, anonymizing methods need word classifier enhancements with stringent "aggressive vectorization," such as coding a *25-year* old person as "25" and "yo" (2019). Advanced ML parsing capabilities, such as Python's deep learning *Keras* framework, can readily deanonymize records by matching "25yo," to public driver license data. Government needs to consider embracing a clearer set of rules to help regulate such findings given the research community is self-regulating by performing tests and augmenting algorithms to ensure patient privacy.

According to Swire, market imperfections regarding privacy protection vary and the "size and type of compliance cost is the degree of precision in the regulation" (1997, pg. 6). As medical privacy research improves discrete actions, health-care organizations can embrace regulation and structure rules building increased compliance for data owners. It is not necessarily a government failure when private companies are deploying effective anonymizing methods but government policy needs to allow for technology improvements. Swire suggests industry member benefits of self-regulation as they want to promote the "reputation of the industry as a whole" (Swire, 1997, pg. 8). However, if industry leaders have the best intentions, technology advancements can quickly erode desired outcomes and limit the effectiveness of self-regulation without a governing body's stringent oversight.

In summary, EHR interoperability is paramount to help healthcare achieve cost reductions across delivery platforms. But, the governments reliance on the private sectors ability to self-regulate patient privacy concerns is not sustainable nor equitable to its constituents. Privacy concerns require EHR systems improve care diagnosis, delivery without compromising trust, and ensure scalable medical information management systems.

Concerns and Ethics Surrounding Electronic Health Records

Privacy concerns on EHRs include accessing records on public computers, users failing to properly log off, lost or dropped printed patient records, the inappropriate faxing or mailing of records, and other cases such as divorcees who access spousal records, such as blood test data,

for litigation. Anther significant EHR privacy concern centers on discrimination. According to Garcia-Murillo and MacInnes, “a primary reason people are concerned with their privacy is we are afraid of being discriminated against” (2014, pg. 17). Use of medical record information in legal affairs or insurance discrimination concerns all citizens. With insurance coverage and rates under constant review sharing certain medical information, such as substance abuse, can result in derogatory information and persistently high insurance costs. The misappropriation of healthcare information generates significant personal consequences and requires robust adverse event avoidance mechanisms.

The growth of EHR fosters a “collect it all” mentality. Schneier, a digital privacy expert, cautions that without technologists as part of the policy planning process both the government and private sector may not have the correct tools to facilitate and ensure privacy (Schneier, 2017). Given research around EHR data re-identification and gaps associated with assimilated knowledge perhaps machine learning requires a new regulatory agency to help manage advanced programmatic findings. Machine and deep learning technology are complex technologies, require significant training, and scientists with aptitudes across statistics and data warehouse disciplines. Production environments are also subject to automation resulting in an incremental loss of human oversight. According to Schneier, perhaps a balance of disconnected systems is required unless health data can be encrypted without limiting interoperability (2019, pg. 8). By agreeing with Schneier, concern should focus on the government’s management of health privacy information because even though population research is a notable endeavor it requires vast storage of citizen biometric data which may fundamentally conflict with the American value of an *unalienable Right to Liberty.*

Enabling values, such as health and education, are critical to support health equity initiatives (Schultz, 2006, pg. 8). Philosopher John Rawls would argue EHR platforms advance societal transformation and promote justice by deepening an individual’s social contract (ibid, pg. 9). While EHR research should expand equality by growing population health with *best off from worst off* the system remained subjugated by privacy dilemmas (ibid, pg. 10).

Electronic Health Records Policy Recommendations

According to Dinev, Hart and Mullen, American society is surveillance oriented and such perceptions have strengthened since Edward Snowden's 2008 revelations (2008). Policy surrounding EHR would benefit with specific data encryption methods and algorithmic techniques found most effective maintaining patient confidentiality. According to Hannah, the company Google envisions being able to save your life with a person's known or unknown information, such as machine learning projections of cancerous cell growth patterns derived from thousands of patients (2020). However, the implications are clear that any public and private patient EHR information will be harvested from whatever source versus record storage abandonment.

Privacy policies also need to tailor how healthcare information can be transacted with customer feedback loops notifying when, where, and how their biometric information is consumed. Currently, privacy "fails to convey a sense of urgency needed to address the serious health privacy consequences of comprehensive and longitudinal records" (Rothstein and Tovino, 2019, pg. 775). Privacy needs to re-establish this urgency to help mitigate drawbacks centering on a costly administrative oversight. Ensuring compliance is expensive but can be offset by providing a transparent privacy ownership platform governed by a central managing structure.

Conclusion

A patient's ability to access their electronic health records provides an opportunity to review their care from each visit fostering health ownership and greater engagement with care providers. This speaks to the importance of EHR portals helping patients become an active participant in their healthcare and expanding health equity across diverse patient populations. Moving forward, healthcare technology need to continuously improve technology platforms and enable EHR's interoperability. Achieving interoperability, protecting health information, and providing secure data structures will help providers and researchers discover novel insights by properly using machine learning to advance societal health equity.

References

Adane, K., Gizachew, M., Kendie, S., (2019). The role of medical data in efficient patient care delivery: a review. *Risk Management and Healthcare Policy*. URL: https://dovepress.com/by 173.48.128.71

Cowie, M. Blomster, J., Curtis, L., Duclanx, S. Ford, I., Fritz, F., Goldman, S., Jammohamed, S., Kreuzer, J., Leenay, M., Michel, A., Ong, S., Pell, J., Southworth, M., Stough, W., Thoeones, M., Zannad, F., Zalewski, A., (2017). Electronic health records to facilitate clinical research. *Clinical Research in Cardiology*, 106(1): 1-9. URL: https://doi-org.libezproxy2.syr.edu/10.1007/s00392-016-1025-6

Dinev, T., Hart, P., Mullen, M. (2008). Internet privacy concerns and beliefs about government surveillance – an empirical investigation. *Journal of Strategic Information Systems*. DOI: 10.1016/j.jsis.2007.09.002

Friedrich, M., Kohn, A., Wiedemann, G., Biemann, C., (2019). Adversarial learning of privacy-preserving text representations for de-identification of medical records. *National Center for Biomedical Computing*. URL: https://arxiv.org/abs/1906.05000

Garcia-Murillo, M, MacInnes, I., (2014). Cosi Fan Tutte: Why a right to be forgotten should not be pursued.

Garcia-Murillo, M., (2020). IST 618-Information Policy: Ethics lecture.

Hannah, K., (2020). Can we ever trust Google with our health data? FT.com. Retrieved on 3/15/20 from https://search.proquest.com/docview/2341964459?accountid=14214

Rothstein, M., Tovino, S., (2019). Privacy risks of interoperable electronic health records: segmentation of sensitive information will help. *The Journal of Law, Medicine, and Ethics*. DOI: 10.1177/1073110519897791.

Schneier, B., (2017). Click here to kill everyone. New York Magazine.

Schultz, R. A. (2005). Contemporary issues in ethics and information technology: IGI Global (Chapter 2)

Swire, P., (1997). Markets, self-regulation, and government enforcement in the protection of personal information, in privacy and self-regulation in the information age by the U.S. department of commerce. *SSRN Electronic Journal*. DOI 10-2139/ssrm11472

USFHealth (website), 2020. Federal Mandates for Healthcare: Digital Record-Keeping Requirements for Public and Private Healthcare Providers. Retrieved on 3/22/20 from URL: https://www.usfhealthonline.com/resources/healthcare/electronic-medical-records-mandate/