HIPAA and Protecting Health Information in the 21st Century

In March 2018, the Trump administration announced a new initiative, MyHealthEData, to give patients greater access to their electronic health record and insurance claims information. The Centers for Medicare & Medicaid Services will connect Medicare beneficiaries with their claims data and increase pressure on health plans and health care organizations to use systems that allow patients to access and send their health information where they like.

MyHealthEData is part of a broader movement to make greater use of patient data to improve care and health. The movement seeks to make information available wherever patients receive care and allow patients to share information with apps and other online services that may help them manage their health. At the population level, this approach may help identify optimal treatments and ways of delivering them and also connect patients with health services and products that may benefit them. Analysis of deidentified patient information has long been the foundation of evidence-based care improvement, but the 21st century has brought new opportunities. With developments in information technology and computational science that support the analysis of massive data sets, the “big data” era has come to health services research.

For all its promise, the big data era carries with it substantial concerns and potential threats. Part of what enables individuals to live full lives is the knowledge that certain personal information is not on view unless that person decides to share it, but that supposition is becoming illusory. The increasing availability and exchange of health-related information will support advances in health care and public health but will also facilitate invasive marketing and discriminatory practices that evade current antidiscrimination laws. As the recent scandal involving Facebook and Cambridge Analytica shows, a further risk is that private information may be used in ways that have not been authorized and may be considered objectionable. Reinforcing such concerns is the stunning report that Facebook has been approaching health care organizations to try to obtain deidentified patient data to link those data to individual Facebook users using “hashing” techniques.

Given these concerns, it is timely to reexamine the adequacy of the Health Insurance Portability and Accountability Act (HIPAA), the nation’s most important legal safeguard against unauthorized disclosure and use of health information. Is HIPAA up to the task of protecting health information in the 21st century?

HIPAA Framework for Information Disclosure

HIPAA was considered ungainly when it first became law, a complex amalgamation of privacy and security rules with a cumbersome framework governing disclosures of protected health information. HIPAA has been derided for being too narrow—it applies only to a limited set of “covered entities,” including clinicians, health care facilities, pharmacies, health plans, and health care clearinghouses—and too onerous in its requirements for patient authorization for release of protected health information. Over time, however, HIPAA has proved surprisingly functional. Particularly after being amended in the 2009 HITECH (ie, the Health Information Technology for Economic and Clinical Health) Act to address challenges arising from electronic health records, HIPAA has accomplished its primary objective: making patients feel safe giving their physicians and other treating clinicians sensitive information while permitting reasonable information flows for treatment, operations, research, and public health purposes.

HIPAA’s Privacy Rule generally requires written patient authorization for disclosure of identifiable health information by covered entities unless a specific exception applies, such as treatment or operations. Researchers may obtain protected health information (PHI) without patient authorization if a privacy board or institutional review board (IRB) certifies that obtaining authorization is impracticable and the research poses minimal risk. The investigators can obtain a limited data set that excludes direct identifiers (eg, names, medical record numbers) without patient authorization if they agree to certain security and confidentiality measures. Importantly, data sets from which a broader set of 18 types of potentially identifying information (eg, county of residence, dates of care) has been removed may be shared freely for research or commercial purposes.

This has been a serviceable framework for regulating the flow of PHI for research, but the big data era raises new challenges. HIPAA contemplated that most research would be conducted by universities and health systems, but today much of the demand for information emanates from private companies at which IRBs and privacy boards may be weaker or nonexistent. Additionally, removing identifiers to produce a limited or deidentified data set reduces the value of the data for many analyses. Moreover, the increasing availability of information generated outside health care settings, coupled with advances in computing, undermines the historical assumption that data can be forever deidentified. Starting demonstrations of the power of data triangulation to reidentify individuals have offered a glimpse of a very different future, one in which preserving privacy and the big data enterprise are on a collision course.

It will be difficult to reconcile the potential of big data with the need to protect individual privacy. One reform approach would be data minimization (eg, limiting the upstream collection of PHI or imposing time limits on data retention), but this approach would sacrifice too much that benefits clinical practice. Another solution involves revisiting the list of identifiers to remove from a data set. There is no doubt that regulations should reflect up-to-date best practices in deidentification. However, it is questionable...
Limited Reach of HIPAA

HIPAA "attaches (and limits) data protection to traditional health care relationships and environments." The reality of 21st-century United States is that HIPAA-covered data form a small and diminishing share of the health information stored and traded in cyberspace. Such information can come from well-known sources, such as apps, social media, and life insurers, but some information derives from less obvious places, such as credit card companies, supermarkets, and search engines. For example, non-health information that supports inferences about health is available from purchases that users make on Amazon; user-generated content that conveys information about health appears in Facebook posts; and health information is generated by entities not covered by HIPAA when over-the-counter products are purchased in drugstores. Because HIPAA’s protection applies only to certain entities, rather than types of information, a world of sensitive information lies beyond its grasp.2

HIPAA does not cover health or health care data generated by noncovered entities or patient-generated information about health (eg, social media posts). It does not touch the huge volume of data that is not directly about health but permits inferences about health. For example, information about a person’s physical activity, income, race/ethnicity, and neighborhood can help predict risk of cardiovascular disease. The amount of such data collected and traded online is increasing exponentially and eventually may support more accurate predictions about health than a person’s medical records.2

Statutes other than HIPAA protect some of these non-health data, including the Fair Credit Reporting Act, the Family Educational Rights and Privacy Act of 1974, and the Americans with Disabilities Act of 1990.7 However, these statutes do not target health data specifically; while their rules might be sensible for some purposes, they are not designed with health in mind. For instance, the Family Educational Rights and Privacy Act of 1974 has no public health exception to the obligation of nondisclosure.7

Conclusions

Improved public understanding of these practices may lead to the conclusion that such deals are in the interest of consumers and only abusive practices need be regulated. Or it may create pressure for better corporate privacy practices. Some consumers may take steps to protect the information they care most about, such as purchasing a pregnancy test with cash. Shaping health information privacy protections in the 21st century requires savvy lawmaking as well as informed digital citizens.

To ensure adequate protection of the full ecosystem of health-related information, 1 solution would be to expand HIPAA’s scope. However, the Privacy Rules’ design (ie, the reliance on IRBs and privacy boards, the borders through which data may not travel) is not a natural fit with the variety of nonclinical settings in which health data are collected and exchanged.8

The better course is adopting a separate regime for data that are relevant to health but not covered by HIPAA. One option that has been proposed is to enact a general rule protecting health data that specifies further, custodian-specific rules; another is to follow the European Union’s new General Data Protection Regulation in setting out a single regime applicable to custodians of all personal data and some specific rules for health data. The latter has the appeal of reaching into non–health data that support inferences about health. Any new regulatory steps should be guided by 3 goals: avoid undue burdens on health research and public health activities, give individuals agency over how their personal information is used to the greatest extent commensurable with the first goal, and hold data users accountable for departures from authorized uses of data.

Rethinking regulation should also be part of a broader public process in which individuals in the United States grapple with the fact that today, nearly everything done online involves trading personal information for things of value. When such trades are made explicit, as when drugstores offered customers $50 to grant expanded rights to use their health data, they tend to draw scorn.9 However, those are just amplifications of everyday practices in which consumers receive products and services for free or at low cost because the sharing of personal information allows companies to sell targeted advertising, “deidentified” data, or both.

REFERENCES


