JOURNAL OF CLINICAL RESEARCH BEST PRACTICES

Vol. 13, No. 5, May 2017 "Happy Trials to You"

Research and Patient Privacy in the 21st Century: HIPAA and Beyond

By Rachel Abramovitz

As we near the end of the second decade of the 21st century, clinical research continues to play a vital role in healthcare. In some cases, patients obtain cutting edge treatments by participating in clinical research. However, medical breakthroughs must fight for headlines with news of privacy and information security breaches. Fast-paced technological and medical developments offer exciting and innovative solutions to individuals, but they also pose challenges. As is often the case, the key is finding the appropriate balance between the needs of the individual and those of society.

The 21st Century Cures Act

Recent passage of the 21st Century Cures Act emphasized the importance of research. Section 2063 of the Act directs the Department of Health and Human Services (HHS) to issue guidance allowing covered entities to grant researchers remote access to protected health information (PHI), for review preparatory to research. The Act specifies that remote access to PHI must meet minimum safeguards consistent with the HIPAA Security Rule. It also directs HHS to convene a working group to study and report on the uses and disclosures of PHI for research purposes. The working group is expected to opine as to whether or not HIPAA should be modified to permit greater use of PHI for research purposes. The Act permits HHS to exempt certain biomedical information from Freedom of Information Act disclosures if the individual is identified or there is at least a "very small risk, as determined by current scientific practices of statistical methods" of identification. Finally, the Act clarifies that Certificates of Confidentiality do not allow release of any information for which there is even a very small risk of identification.

Updated Common Rule

Almost six years following its notice of intent to modernize the Common Rule, HHS issued its new Final Rule on January 19, 2017. The new rule will take effect on January 19, 2018.[1]

The new rule defines an "identifiable biospecimen" as a biospecimen for which the identity of the subject is or may be "readily ascertained" by the investigator. Whether the subject's identity may be "readily ascertained" is assessed through examination of new technological developments and new analytical techniques that have the potential to re-identify information or biospecimens that were previously considered to be anonymized. If such technologies are identified, they will be placed on a list published in the Federal Register, along with recommendations on how to manage privacy and security accordingly. Whole-genome sequencing is anticipated to be one of the first technologies to be reviewed by HHS.

The new rule also requires HHS to issue guidance for IRBs to help them determine when the standards for protection of privacy of human subjects have been met. Under the new rule, IRBs may approve a research proposal in which investigators, for eligibility screening and recruitment purposes, may access the PHI of prospective human subjects without informed consent. Access may be obtained through oral or written communication with the prospective subjects, or by accessing pre-existing records containing PHI. The new rule also allows subjects to provide "broad consent" for secondary research, subject to certain requirements and IRB review.

© 2017 First Clinical Research and the Author(s)

Technology and Big Data

Lately, we find ourselves in an environment where technology is changing extremely rapidly. Patients are using new gadgets, apps, wearable devices, implanted medical devices, etc., in their healthcare, and individuals are using them in their day-to-day wellness activities. Some data are collected by the app/device owners on behalf of the individual/patient, while, in other cases, data are collected on behalf of the healthcare provider or even the device manufacturer.

In the context of clinical research, in addition to the standard collection of data through the patient's medical records, patients are often asked to utilize eDiaries, mobile apps, and other wearable devices to collect information for the purpose of the study.

"Big data" is also playing a growing role in population health management and epidemiological research. The amount of data collected with these new technologies can exceed that collected with previous methods by 1,000-fold or more.

"Big data" is a term that refers to extremely large data sets that can be analyzed to reveal patterns, trends and associations. While big data sets are generally de-identified and, therefore, not necessarily relevant to clinical research, big data plays an important role in epidemiological studies and secondary research. In the EDPS Opinion 7/2015, the EU data protection supervisor provides guidance on appropriate methods for anonymization of big data sets.[2] U.S. privacy laws that address these questions include HIPAA, the FTC Act, and state medical privacy and information security laws. Appropriately anonymized individual-level clinical trial data will most certainly provide researchers with significant amounts information that can be utilized for secondary research.[3]

Research Data, HIPAA and Other Privacy Laws and Regulations

Healthcare providers know they must protect individually identifiable patient information under HIPAA. However, study sponsors generally are not subject to HIPAA. Is information collected on behalf of a research sponsor (as opposed to the healthcare provider) for the purpose of conducting a clinical research study subject to HIPAA? HIPAA defines PHI as individually identifiable health information that is transmitted or maintained in any form by a covered entity or its business associates. Therefore, once the information is transmitted to and maintained by a healthcare provider that is a "covered entity" under HIPAA, the information is subject to HIPAA.

In addition to HIPAA, a growing number of state privacy laws might apply to personal information collected by sponsors for clinical research purposes. The Federal Trade Commission's (FTC's) breach notification rule might apply when electronic health information is collected by an entity not covered by HIPAA, through computer programs and mobile apps.[4] If personal health data that is subject to the FTC rule is "unsecured" (i.e., not encrypted), then, in the event of a breach of that data, the provisions of the rule would apply.[5] Generally speaking, information is considered "secured" if the HIPAA Security Rule has been followed with respect to the handling of the information, i.e., access to the information is appropriately controlled and the information is encrypted when it is in transit and at rest.

With healthcare organization data being one of the primary targets of hackers, taking a proactive approach to privacy protection might help address study subjects' concerns about the privacy of their data.[6] According to a recent survey from the Office of the National Coordinator of Health Information Technology, about three-quarters of respondents stated that they are very or somewhat concerned about the privacy and security of their medical information.[7] When patients agree to participate in a clinical research study, their concerns

© 2017 First Clinical Research and the Author(s)

often become more acute, since, as a result of the research, they must authorize the sharing of their PHI with the sponsor, the FDA, and other entities.

Use of "De-identified Data" in Research

Many people believe data is "de-identified" when direct identifiers, such as first and last name, address and Social Security number, have been redacted. Under HIPAA, health information may be de-identified in two ways. The first way, also known as the "safe harbor method," requires the removal of all 18 identifiers specified by HIPAA.[8] The 18 identifiers include information that many might not consider to be a direct identifier, e.g., ZIP code, IP address, and all elements of dates (except for year) related to the individual, including dates of service (e.g., study visits), dates of treatment, date of birth, etc. Case report forms, for example, usually contain the subject's date of birth, as well as dates of service. Many case report forms also include subject initials. The NIH has provided guidance that initials are an identifier under HIPAA.[9] Therefore, most case report forms cannot be considered "de-identified" and should be treated as PHI.

The second method of de-identification requires an expert determination by a statistician that there is a very small chance that an individual's information could be re-identified. In November 2012, the Office of Civil Rights (OCR) issued guidance regarding the "Expert Determination" method, as set forth in the HIPAA Privacy Rule.[10] This method requires that a person with statistical expertise determine that the risk of re-identification of a particular data set is "very small." An "expert," according to OCR's guidance, "may be found in the statistical, mathematical or other scientific domains."[11] There is no specific "numerical level" of identification risk that is deemed to universally meet the "very small" requirement. Rather, the ability of someone to identify an individual based on a data set depends on a variety of factors. According to OCR, "the risk of identification that has been determined for one particular data set in the context of a specific environment may not be appropriate for the same data set in a different environment or a different data set in the same environment."[12]

Even if we follow one of the two methods of de-identification under HIPAA, statisticians recognize that big data sets are still susceptible to re-identification by using secondary data sets. In a high-profile and controversial study published in 2013, a Harvard researcher concluded that 42% of a sample of participants in a DNA study could be re-identified.[13] The data set used for the study was the Personal Genome Project (PGP). All direct identifiers, such as name, medical record number, and Social Security number, were excluded from that data set. The research team took just three key pieces of information from a subset of participants — ZIP code, date of birth, and gender — and, using voter registration and other public records, succeeded in identifying a significant number of individuals in that group. As noted above, ZIP code and date of birth are included in the list of 18 identifiers under HIPAA, so the data set in question was not de-identified using the HIPAA Safe Harbor method.

Additional guidance on de-identification of personal information can be found in the National Institute of Standards and Technology (NIST) IR 8053 guidance, which also provides several examples of re-identification of data.[14]

Research using existing data is exempt from federal human subject research regulations and IRB review if the data are publicly available or if the information is used by the "[research] investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects."[15] When would the information be considered "de-identified"? If the data set complies with HIPAA's de-identification Safe Harbor rule, i.e., when all 18 identifiers have been removed, the data is de-identified and the research can be conducted without IRB oversight. However, it is unlikely that a data set that has been de-

identified in this fashion will be of much use to researchers, since all dates of service related to the individual must be removed from the data set for it to be considered de-identified under the Safe Harbor method.[16] If the data set does contain certain identifiers, researchers might consider applying the "Expert Determination" method. A third option is to go the usual route and conduct the research in accordance with Code of Federal Regulations (CFR) Part 46 human subject regulations, including IRB oversight and the use of standard privacy protections.

HIPAA and Patient-Generated Health Data

Most wearable devices and smartphone health applications are designed for consumer or patient use, but they can also generate important patient information for healthcare providers. When patient information is collected on behalf of the patient (and not for a doctor), the general consensus is that the resulting data is not PHI. However, if that information is later transferred to a healthcare provider for inclusion in the patient's medical record or in a research data set, the information will likely then become PHI and, hence, subject to HIPAA.

Wearable fitness tracking devices, health monitoring devices, and smartphone applications are being developed to provide users with an amazing range of health-related data on heart rate, calories burned, blood sugar, cholesterol, movement and a host of other health-related parameters.[17] Although most popular, current products are designed to help individuals manage their fitness, these products will increasingly be used by both patients and physicians to help manage chronic diseases and detect changes in health conditions. In addition to countless digital health products under commercial development, the Apple ResearchKit, an iPhone software platform, enables researchers to create their own iPhone apps for specific clinical studies.[18]

Information collected through these devices can be used to improve the medical standard of care. For instance, a group of researchers at the Mayo Clinic College of Medicine asked patients who had undergone cardiac surgery to wear Fitbit activity trackers after the procedure. The researchers definitively concluded that patients who moved more the day after surgery were likely to be discharged sooner from the hospital.[19]

As noted above, when individually identifiable health information is created or received by a healthcare provider, that information falls under the definition of PHI.[20] So, when a mobile app collects individually identifiable health information and provides that information to a healthcare provider, the app must be HIPAA compliant. The app developer might say the information is being collected on behalf of the individual, who then chooses whether to share it with their doctor. In that case, although HIPAA might not apply, FTC regulations and state laws might apply. Mobile health app developers may utilize the FTC's mobile health apps interactive tool to assess which laws and regulations likely apply to their app.[21]

Putting aside the question of the legal status of the data prior to its provision to the researcher, once information is entered in a patient's record by a healthcare provider, that information becomes PHI, regardless of the status of the information at the time it was generated.

The Food & Drug Administration (FDA) has only recently begun to address the privacy and security of information created, stored and transmitted by conventional medical devices like blood pressure monitors and implantable defibrillators. Recent high-profile incidents, such as FDA's announcement of security vulnerabilities in networked, programmable infusion pumps, give cause for concern.[22]

© 2017 First Clinical Research and the Author(s)

State Privacy Laws

In the context of state law, the definition of "Personal Information" varies from state to state, but, for the most part, it includes an individual's first name or first initial and last name, plus one or more of the following data elements: (a) Social Security number, (b) driver's license number or state-issued ID card number, or (c) account number, credit card number or debit card number combined with any security code, access code, PIN or password needed to access an account. States and territories that currently include medical and/or health information in the list of data elements include Arkansas, California, Florida, Illinois, Missouri, Montana, Nevada, New Hampshire, North Dakota, Rhode Island, Virginia, Wyoming and Puerto Rico. The good news is that if a data set does not include first name/initial and last name, state privacy laws probably do not apply. However, if a data set does include names, privacy counsel can advise on whether state laws apply. HIPAA pre-empts state law when it is more stringent. In some cases, such as California, the state law is more stringent and, therefore, pre-empts HIPAA.

Information Security Best Practices

When creating, receiving, maintaining or transmitting personal health information, appropriate security controls must be utilized to minimize the risk of a breach. The HIPAA Security Rule, along with the most recent NIST guidance documents, serve as best practice guidance for entities to follow in securing their data, even if those entities are not subject to HIPAA. The HIPAA Security Rule provides a step-by-step method for creating and maintaining electronic health information securely. It includes administrative, physical and technical safeguards, and implementation specifications for each safeguard. An entity that complies with the HIPAA Security Rule, generally speaking, will likely be materially in compliance with applicable state privacy laws, as well as with the FTC breach notification rule.[23] If health information is not properly secured, unauthorized access could occur. Unauthorized access to PHI is a HIPAA breach. Unauthorized access to personal information, as that term is defined by applicable state law, would thus likely be a state privacy breach in states that include medical and/or health information in their definition of "personal information." Unauthorized access to PHI subject to the FTC Breach Notification Rule is an FTC health information breach.

Conclusion

Healthcare providers' use of big data and patient-generated data poses interesting questions and challenges related to privacy and regulatory compliance, particularly HIPAA compliance. Using these forms of data for research purposes might create an additional set of considerations. As healthcare providers begin to develop policies, procedures and institutional guidelines to facilitate uses and disclosures of these relatively new forms of data, they should consider not only clinical care purposes, but also research purposes.

References

1. See: https://www.federalregister.gov/documents/2017/01/19/2017-01058/federal-policy-for-the-protection-of-human-subjects.
2. See: https://edps.europa.eu/sites/edp/files/publication/15-11-19_big_data_en.pdf.
3. See EU Opinion 05/2014 on Anonymization Techniques (WP216) and NISTIR 8053 De-Identification of Personal Information.
4. https://www.ftc.gov/news-events/press-releases/2009/08/ftc-issues-final-breach-notification-rule-electronic-health.

© 2017 First Clinical Research and the Author(s)

5. 16 CFR § 318.3.
6. http://www.hipaajournal.com/experian-healthcare-organizations-main-targets-hackers-2017-8690/.
7. Healthcare Information and Management Systems Society (HIMSS), Healthcare IT News, E. McCann, *Privacy Remains Top Priority for US Consumers*, Jan. 2015.
8. The 18 identifiers under HIPAA are: (A) Names; (B) All geographic subdivisions smaller than a State, including street address, city, county, precinct, ZIP code, and their equivalent geocodes, except for the initial three digits of a ZIP code …; (C) All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, …; (D) Telephone numbers; (E) Fax numbers; (F) Electronic mail addresses; (G) Social Security numbers; (H) Medical record numbers; (I) Health plan beneficiary numbers; (J) Account numbers; (K) Certificate/license numbers; (L) Vehicle identifiers and serial numbers, including license plate numbers; (M) Device identifiers and serial numbers; (N) Web Universal Resource Locators (URLs); (O) Internet Protocol (IP) address numbers; (P) Biometric identifiers, including finger and voice prints; (Q) Full face photographic images and any comparable images; and (R) Any other unique identifying number, characteristic, or code [that could reasonably be used to identify an individual].
9. See: https://privacyruleandresearch.nih.gov/research_repositories.asp.
10. 45 C.F.R. 164.514(b).
11. OCR Guidance, *supra* note 14.
12. *Id.*
13. A Tanner, *Harvard Professor Re-Identifies Anonymous Volunteers in DNA Study*, Apr. 25, 2013, Forbes, http://www.forbes.com/sites/adamtanner/2013/04/25/harvard-professor-re-identifies-anonymous-volunteers-in-dna-study/.
14. http://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf.
15. 45 C.F.R. 46.101(b)(4).
16. See: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/#dates.
17. https://www.apple.com/ios/whats-new/health/ (last accessed Feb. 22, 2015).
18. A Duhaime-Ross, Apple's New ResearchKit: 'Ethics Quagmire or Medical Research Aid?, The Verge, Mar. 10, 2015, http://www.theverge.com/2015/3/10/8177683/apple-research-kit-app-ethics-medical-research.
19. E Dwoskin & J Walker, *Can Data From Your Fitbit Transform Medicine*, Wall St. Journal, http://www.wsj.com/articles/health-data-at-hand-with-trackers-1403561237.
20. 45 CFR § 160.103 – definitions.
21. https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool.
22. Food & Drug Administration, Safety Communication, May 13, 2015, http://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/ucm446809.htm.
23. See 74 FR 42964.

Acknowledgement

The author would like to acknowledge the contributions of Leah A. Voigt to an early version of this article.

Author

Rachel Abramovitz, LL.B. , LL.M.C. is a principal at Blustein Abramovitz, Law Offices. Contact her at 1.646.665.1578 or rachel@ablaw.nyc.

© 2017 First Clinical Research and the Author(s)