Report on Research Compliance Volume 16, Number 8. August 01, 2019

Despite Safeguards, U. Chicago-Google Research Sparks Lawsuit

By Theresa Defino

More than four years ago, a patient was twice hospitalized at the University of Chicago Medical Center, his care tracked and documented, as usual, in UC's electronic medical records (EHR) system. But, years later, and without his knowledge or consent, Matt Dinerstein claims those stays made him an unwitting participant in a research project involving UC and Google that has resulted in at least one published paper.

And now, one lawsuit as well.

Dinerstein's suit, filed June 26 in U.S. District Court for the Northern District of Illinois, Eastern District, seeks to be certified as class action litigation. Dinerstein was admitted to UC medical center for a total of six days in June 2015. UC and Google began a research collaboration in May 2017.

Quality of Deidentification At Issue

The suit terms UC's sharing of protected health information (PHI) with Google a "massive medical data grab" and claims—without providing any details—that UC "engaged in a cover up to keep the breach out of the public eye so as to avoid the public backlash."

Google and UC called the suit baseless and said they would mount a defense.

UC, the suit claims, was "happy to turn over [to Google] the confidential, highly sensitive and HIPAA-protected records of every patient who walked through its doors between 2009 and 2016."

Both Google and UC "violated HIPAA by sharing and receiving medical records that included sufficient information for Google to re-identify the patients," the suit alleges. "Both were aware at the time of the transfer that the medical records contained information outside of HIPAA's safe harbor provisions, that a competent expert determination was not made, and that the thousands of patients had not given proper consent to allow Google to take possession of the records for the purpose of creating a commercial product."

Little information is available about how data were scrubbed of PHI.

Just one paper appears to have resulted from the UC-Google project so far. Titled "Scalable and accurate deep learning with electronic health records," the paper was published last year in the Nature Partner Journals: *Digital Medicine.*

Using "de-identified EHR data from two U.S. academic medical centers with 216,221 adult patients hospitalized for at least 24 hours," authors from UC, Google, Stanford Medicine and the University of California San Francisco (UCSF) wrote that they "demonstrate[d] that deep learning methods using this representation are capable of accurately predicting multiple medical events from multiple centers without site-specific data harmonization."

According to the paper, "All electronic health records were de-identified, except that dates of service were maintained in the UCM [University of Chicago Medicine] dataset. Both datasets contained patient demographics,

provider orders, diagnoses, procedures, medications, laboratory values, vital signs, and flowsheet data, which represents all other structured data elements (e.g. nursing flowsheets), from all inpatient and outpatient encounters. The UCM dataset additionally contained de-identified, free-text medical notes. Each dataset was kept in an encrypted, access-controlled, and audited sandbox. Ethics review and institutional review boards approved the study with waiver of informed consent or exemption at each institution."

A few more details are contained in a May 2018 article about the collaboration with Google posted on UC's website. Its Center for Research Informatics (CRI) "has a team of data warehouse staff dedicated to providing de-identified data for research," the post states. "The team has built a reputation for providing high-quality data for research while going to great lengths to protect patient privacy and security."

The post adds that "all patient identifiers, such as names, dates of birth, Social Security numbers and any other unique characteristic or code, were stripped from the data before Google was given access." The research, UC said, "was conducted according to our rigorous standards" and was "supervised" by the institutional review board (IRB) of UC's Biological Sciences Division.

This document is only available to subscribers. Please log in or purchase access.

Purchase Login