IMAGINE THE UNIVERSITY OF UTAH

Non-HIPAA Covered Entities: Data in Registries

Leslie Francis

Distinguished Alfred C. Emery Professor of Law

Distinguished Professor of Philosophy

Director, Center for Law & Biomedical Sciences

University of Utah

Outline

- Non-HIPAA covered entities: primary examples
- Registries: types, data sources
- Some preliminary data about registries and the inadequacy of protections

Non-HIPAA Covered Entities: Primary Examples

- Providers who do not have any records in electronic form (some counselors); near-providers (massage therapists)
- Social media (e.g. Facebook; Patients Like Me)
- Web search history (e.g. WebMD)
- Wearables (e.g. FitBit)
- Personal record storage (e.g. exercise logs; calorie intake logs, PHRs)
- Recreational genetics (e.g. 23 and Me)
- Registries (e.g. CF Foundation Patient Registry)

Non-HIPAA Covered Entities: Why the Problem?

- Health information may be as detailed and as sensitive as information possessed by HIPAA-covered entities
- May receive PHI from HIPAA-covered entities, without patients realizing that the PHI has been transferred or is no longer HIPAA-protected
- Protections (primarily FTC, state law) uneven at best for some
- Privacy policies often difficult to find, hard to read
- Important provisions may be dispersed among Terms of Conditions or at other places on the website
- Users may have little information about or control over how data are used or transferred by these entities, especially if there has been a representation that data have been de-identified

The Social Media Argument: People like to share because they judge they are getting benefits

- Fair enough, but . . .
- People may not realize what information is being collected, or how much, even on social media.
- Many data transfers from HIPAA-covered entities to non-HIPAA covered entities occur without either effective notice or choice for patients
- Registries are an example

Registries

- Repositories of patient data collected for specific purposes
- May be limited to patients with specific conditions (e.g. rare genetic diseases), specific known exposures (e.g. to a toxin), specific treatments (e.g cardiac device)
- May be funded from public $$, charitable contributions, pharmaceutical companies, professional organizations

Registry Landscape: Vast

- Public health: tumors, birth defects
- Disease specific
  - Charitable (e.g., CF Foundation)
  - Pharmaceutical company sponsored (e.g. MastCell Connect for mastocytosis; sponsor blueprint Medicines)
- Patient generated (e.g. Genetic Alliance)
- Researcher-created (e.g. SEER)
- Medical association sponsored (e.g. ACC)

Data sources for registries

- Data originally collected for clinical care, in electronic form, within the HIPAA-covered entity
- Data originally collected for clinical research within the HIPAA-covered entity
- Data entered by patients themselves
- Data entered by family members of patients

Registry data collection and use

- May be by one-time patient consent to entry on an ongoing basis
- May be by one-time surrogate consent to entry (e.g. parents); although many require adult consent to continuing data collection not all do
- May collect data directly from clinical records or from patients themselves
- Typically require patient consent for participation in particular studies using identifiable data, but not for uses or transfers of data in de-identified form
- May sell de-identified data to support registry operations

Identified or de-identified?

- Once de-identified, no longer HIPAA PHI
- Use of de-identified data not human subjects research
- Risk of re-identification
  - Has been primary subject of discussion regarding de-identified data
  - But re-identification is not the only, or even the primary, concern

Concerns beyond re-identification

- Inferences from conjoined data sets
  - Novel or surprising
  - Stigmatizing
  - May apply to others not included in original data sets
- Uses of data that are disapproved
  - Sense of contribution to something that is wrong
  - Loss of identity
- Uses of data that could cause economic harm
  - Job costs for groups: changes in workplace policy
  - Benefits loss: redlining

Data Downstream: protections?

- If de-identified, typically an agreement not to re-identify
- If identified
  - Data use agreements
  - Patient authorization (HIPAA)
  - Patient consent (research data)
  - IRB review
- Enforcement? Contract law, laws applicable to certain positions (e.g. public health employees, university faculty)
- How monitored? We don’t really know in many cases

Pilot study of registry governance

- NIH website list of registries
- Selected those with contact information, data collection on an ongoing basis: 59
- Successful contact with 30; 20 agreed to discuss governance
- Varied in size from 200 to 800,000 participants
- IRB approved questionnaire

Preliminary findings: governance

- All registries had identified staff, decision-making bodies
- Only half had an advisory board or second body of advisors to guide technical, scientific, or ethical decision making
- Fewer than a third were transparent about their decision-making process

Preliminary findings: privacy and security

- Half of the registries publicly specified uses of the data they were collecting
- Fewer than half permitted participants to access their data
- One-third gave specific information about data storage; this included one that stored data on a google format and another that stored data on servers outside of the US
- Only ONE registry had a protocol for addressing data breaches

Acknowledgements

- Research reported in this publication was supported by Utah Center for Excellence in ELSI Research (UCEER). UCEER is supported by the National Human Genome Research Institute of the National Institutes of Health under Award Number P20HG007249 (or RM1HG009037). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health
- I am grateful to Michael Squires, my UCEER RA, for interviews with registries