It’s 3 a.m. . . . . . . . .

Do you know where your health data are?
No Records Responsive to Your Request

Barriers to State Oversight of Health Data Sharing
Google to Store and Analyze Millions of Health Records

The tech company's deal with Ascension is part of a push to use artificial intelligence to aid health services.

Inside Google’s Quest for Millions of Medical Records

The company has struck deals that grant it access to troves of patient data. “We want to be helpful.”
What does Google want health data for, anyway?

“[Google] was planning to build a search tool for medical professionals that would employ machine-learning algorithms to process data and make suggestions about prescriptions, diagnoses, and even which doctors to assign to, or remove from, a patient’s team.” (As per the Atlantic)

- Insight: a lot of problems in healthcare are actually well-suited for the application of data-driven / ML approaches
  - Diagnosis, treatment recommendations, admissions / discharges, bed assignments, ...
Growing Demand for Health Data?

- “Health Data Analytics” sector
  - ~$12 billion in 2018, projected to grow to $80 bil by 2026
- Access to data is a huge competitive advantage when trying to train predictive models for care delivery
  - Top-tier algorithms are usually published in papers / available to the research community --- the “secret sauce” is usually in the parameters learned through extensive training.
Thank God for HIPAA!

- Protected under the Health Insurance Portability and Accountability Act
- Key principle: Health providers can only share protected health data for reasons related to providing care
  - When a business wants to receive health data in order to provide a service (lab tests, EHR, diagnostic tools) they must enter into a Business Associate Agreement
  - BAAs specify the purposes for and the ways in which the business can use the data
With One Little Problem...

- Anonymized health data is not protected under HIPAA
  - HIPAA has fairly robust standards for anonymization
    - (...Query whether anything can ever be anonymized)
- Certain BAAs allow the business to create anonymized health data from protected health data
  - Once anonymized, it can be used for anything (no protections!), inc. sold for profit
What can Google do with that data? & Are they the only one?
The original research questions

- Investigate the BAAs entered into by different health care providers
  - Which companies are the agreements with?
  - What are the processing purposes for each?
  - What other data uses are allowed by the contracts?
Methods

- Freedom of Information Act Requests!
  - FOIA laws for federal and state gov’ts
- Advantages of FOIA
  - Entity has to respond (or indicate that they cannot) within a certain timeframe
  - Entity has to turn over actual records (rather than just relating their impressions / PR line)
  - Should get a broader swath of records for each request
- Disadvantages of FOIA
  - Limited to gov’t orgs and public hospitals
Methods

- FOIA requests to:
  - State health agencies
  - County / municipal health agencies
  - Public hospitals & public hospital groups

- Submitted through different mediums
  - Online portal, email, phone call, request to use mail (?) or fax (???)
Results

- We submitted 28 public records requests in total
  - Six cities, seven hospitals, three counties, and nine state-level agencies
  - In 11 states (and the District of Columbia)
- Of the 22 requests that yielded a response, 20 were rejections, and only two yielded responsive documents:
  - Lemuel Shattuck Hospital (MA) And Cook County Health (IL)

The Department is unable to identify any records responsive to your request.

Please be advised that there are no records responsive to your above-referenced request. With this response, your above-referenced request is now closed.

You do not appear to be a citizen of Arkansas.
Findings (from rejections)

- Almost every state-level agency indicated that they did not hold the records \textit{and} that they were not sure who would, other than the hospitals themselves
  - This indicates that there is little to no oversight of BAAs & their provisions
  - Hospitals are the entrants into these contracts \textit{and} their oversight bodies
  - Makes it extremely challenging to document any kind of “state of the world” (would you need to FOIA every single hospital?)
Findings (from successes)

- Both hospitals returned only a few BAAs, and indicated that returning any more would be onerous
  - Indicates that hospitals likely do not have a centralized or searchable database of their own agreements and the provisions / limitations of each
  - May actually violate some provisions of HIPAA (which requires that patients be able to request that their information be corrected or deleted)
Findings (from successes)

- For both hospitals, the language of BAAs was extremely ambiguous --- even with a close reading, we struggled to determine whether the BAA would permit the BA to anonymize data for their own business functions
  - Use of jargonistic phrases like “health care operations”
  - Reliance on state law whose application may be ambiguous
  - Indicates that mere collection of records may not be enough for oversight, unless BAAs are more explicit about the specific uses of data that are allowed or prohibited
Case Study: LabCorp

- One of the BAs from Lemuel Shattuck Hospital
- “Health care operations”
  - One of the seemingly-innocuous phrases that actually has a definition within HIPAA which encompasses business functions like anonymization
  - We’re still not sure whether LabCorp can anonymize data for its own use
  - However…
- Quest, LabCorp’s main competitor, makes millions of dollars per year from the sale of de-identified data
- LabCorp had also recently announced a partnership with Cioux Health to create a de-identified database of COVID-19 test results for use in epidemiological research and predictive modelling
Some of the challenges

- Absolute lack of oversight seems... dangerous
  - However, there are reasons for the federated system of health oversight
- Do we want to impose limits on health data sharing?
  - Fine balance between protecting personal privacy and stifling innovation in an area where innovation is actually... really good.
What would you recommend?
Recommendations

Problem:

- Lack of any oversight mechanism, and general confusion about who is responsible for oversight

Solutions:

- Require state Depts of Health to collect and maintain a database of these agreements as a part of their annual auditing work
- Require Depts of Health to have at least one employee whose responsibilities include collection, oversight, and auditing of BAAs
Recommendations

Problem:

- Lack of record-keeping at hospitals, making it difficult for them to answer questions about the use of patient data by third parties

Solutions:

- Fed and state Depts of Health should clarify the reporting requirements for hospitals, including a requirement to standardize the process of querying the third parties to whom a specific patient’s data has gone
Recommendations

Problem:

- Lack of transparency around what kinds of data use are permitted by BAAs, caused by patchwork of Fed and state laws as well as HIPAA jargon

Solutions:

- Fed and state Depts of Health should introduce a taxonomy of health data uses and provide template language for allowing / restricting each
- Hospitals should store each BAA with information about the uses permitted and disallowed under this taxonomy,
  - Would make it possible to answer questions like “How many of your agreements allow for the anonymization and reuse of data?”
Acknowledgements

HUGE thanks to my partner Maggie Baughman, who wrote this with me through fits of sleep-deprivation-induced delirium and got all the hard parts right.

HUGE thanks to Prof Mayer, who helped us focus in on the most parts of the questions we were asking and the most-likely-to-succeed methods we had proposed.
No Records Responsive to Your Request:
Barriers to State Oversight of Hospital Data-Sharing Practices

Eyes on the PHIs
Barriers to State Oversight of Public Hospitals’ Data-Sharing Practices

Curses, FOIA’d Again:
Record Availability Barriers to State Oversight of Hospital Data-Sharing Practices