SCHOOL OF MEDICINE
VANDERBILT UNIVERSITY

Prepared Statement of Bradley Malin, Ph.D.
Associate Professor of Biomedical Informatics, Vanderbilt University Medical Center
Associate Professor of Computer Science, Vanderbilt University
National Committee on Vital & Health Statistics; Subcommittee on Privacy, Confidentiality, & Security
Hearing on De-Identification and the Health Insurance Portability and Accountability Act
Tuesday, May 24, 2016

Good morning and let me begin by thanking the committee for the opportunity to present testimony this morning. My name is Bradley Malin, and I am an Associate Professor of Biomedical Informatics in the School of Medicine, as well as an Associate Professor of Computer Science in the School of Engineering, at Vanderbilt University. I also serve as the Vice Chair for Research in the Department of Biomedical Informatics at the Vanderbilt University Medical Center.

I was asked to provide several remarks regarding policy interpretations of the de-identification guidance associated with the Health Insurance Portability and Accountability Act of 1996 (HIPAA). To provide context, it should be recognized that, between 2009 and 2012, I served as a paid expert consultant to the Office for Civil Rights at the U.S. Department of Health and Human Services. During that time, I assisted in the facilitation of a public workshop on best practices in de-identification (1), as well as drafting the guidance itself based on material presented at the workshop (2). However, I speak to you today, not as a representative (now should my testimony be viewed as any official statement) of the federal government, but instead as a scientist and practitioner of de-identification who has worked in this field for over 15 years. In particular, during this time, I assisted in the establishment and management of (i) a large de-identified electronic medical record system (over 2 million patient lives), the Synthetic Derivative (3), at the Vanderbilt University Medical Center, (ii) its extension into BioVU (3), a de-repository of over 200,000 individuals DNA, and (iii) various oncology clinical trials datasets shared by pharmaceutical companies to the public through the Project Data Sphere program (4).

To keep my remarks on point, I will focus my attention on the specific questions that were posed to the panelists prior to this hearing.

1. **What issues do you see as being the most pressing when you consider de-identification and HIPAA?**

To provide an appropriate response, it should be acknowledged that, while HIPAA has a single de-identification definition, it allows for two different implementations, each of which has its own set of disputable points. The first implementation is what is generally referred to as "Safe Harbor," whereby the covered entity suppresses (i) 18 enumerated types of identifying attributes, as well as (ii) any additional information for which the covered entity holds actual knowledge that the data shared could identify the individual to whom the data corresponds. It is fairly straightforward to realize Safe Harbor, which was the intention of its design; however, there are several interpretative challenges.

- **An unbounded set of features related to one's persona.** Though there no explicit identifiers will remain in such data, there may be many features that could lead to unique representation of an individual, such as (i) size of family, (ii) personal income, (iii) general title of occupation, (iv) marriage status, (v) language spoken, (vi) native country, (vii) level of education. Technically, this list could be infinite in length because Safe Harbor only lists what must be excluded. While there are no explicit identifiers that are permitted, the wide variety of information that can be disclosed is clearly quite substantial. Moreover, it is growing daily as an artifact of our ability to collect ever more detailed information about an individual's activities and experiences.
- **Unknown semantics in natural language text.** Safe Harbor does not provide clear guidance on how to handle natural language text, such as history and physical reports or the communications between a clinician's office and a patient. There have been numerous natural language processing (NLP) technologies developed over the past several decades (more on this in a moment) to find and redact terms in such text that fall into the 18 prohibited categories (5,6), but there is a potential for a large quantity of semantic information to be communicated, which may implicate an individual indirectly. This is particularly a concern for electronic medical record systems that document the history and experiences of a patient over decades, as well as emerging and evolving *ad hoc* communications (such as those generated through online patient portals) that are becoming increasingly more conversational and intimate. Again, there are no explicit identifiers, but given the volume of data and high variety of information communicated in natural language, there is certainly potential for undesirable disclosure. This begs the question of if natural language, with an open vocabulary, meets Safe Harbor requirements.

The second HIPAA de-identification implementation option is an "expert determination" or what is sometimes referred to as the "statistical standard". Under this approach, an expert uses generally accepted statistical or scientific methods to (potentially amend) and determine that the data shared has a small risk of uniquely identifying an individual to an anticipated recipient. This is a more prudent approach than Safe Harbor because it requires a formal assessment of the data that will be disclosed. However, there are issues here that need to be considered. For brevity, I will focus on the most pressing.

- **Fuzzy notion of an "anticipated recipient".** This is a notable concept because it acknowledges that not all recipients of health data have the same capabilities and motivations. It is clearly the case, for instance, that a trusted scientific investigator who has appropriate training in ethical behavior and good practices in information security is a different type of recipient than any individual who accesses the Internet to download data (possibly through some anonymous proxy server). However, what is challenging is that even the most trustworthy individuals may be subject to hacking or make mistakes in data management more generally. As such, it begs the question of when a risk assessment should consider only the explicitly-stated anticipated recipient, but also the unanticipated recipient (who, after a hacking event would be realized as an anticipated recipient). In many cases, it is prudent to consider both types of recipients, but with different probabilities of their realization (i.e., the anticipated recipient has a probability of 1, while the unanticipated has a probability proportional to the likelihood that there will be a security breach).

There is one additional issue that I wish to raise, which specifically pertains to *omics data derived from biological material. High-dimensional data derived from biological materials, such as genome sequencing records are not explicitly in the Safe Harbor list. Yet over the past two decades, there has been a number of detective-like investigations illustrating how high-dimensional omics data, devoid of features precluded by Safe Harbor, can be linked back to named individuals (7). Still, these attacks tend to assume either that (i) the recipients of such data already know the identity of the corresponding record (such that they aim to learn clinical information) or that (ii) they can probabilistically guess the identity of the corresponding person through a series of transformations and integrations of the data with other resources. The Office of Human Research Protections is currently determining whether they will revise their definition of human subjects in the Common Rule to explicitly state that such data is inherently identifiable. At the present time, as my colleague Dr. Ellen Wright Clayton have I elaborate upon, the evidence suggests that such data is not readily exploitable in a systematic fashion (8). This is not to say that such data is impregnable to attacks in the future, but that it should retain its status as de-identified at this time. It may be prudent to preclude such information from sharing under Safe Harbor, but ensure that it is appropriately assessed it in the context of an expert determination that considers the data semantics and the environment into which it will be made available.

2. Is the current HIPAA de-identification guidance sufficient? Does it pose challenges to or does it advance the use of data (including aggregation, analysis, dissemination, sharing) in healthcare?

One of the principle challenges of the current guidance is that it is not quite clear on the definition of risk. This does not limit how data can be used, but there are many different ways in which risk can be perceived and assessed. To gain some intuition into the challenge, let me set up a simple scenario. Every dataset can be viewed as a *sample* of information about individuals who are drawn from *population*. Now, there are, at a minimum, three types of risk measures that are invoked in practice. The first, which is called journalist risk, corresponds to the most risky record in the population because it is assumed that the recipient does not know if a targeted individual is, in fact, in the dataset. The second, which is called prosecutor risk, corresponds to the most risky record in the sample because it is assumed that the recipient does know that the targeted individual is in the dataset. The third, which is called marketer risk, corresponds to the *average* risk over all records in the dataset. While there is precedence in terms of the risk values for resulting from these perspectives, there is often a question of which should be applied when? One possible approach is that all three should be considered, but that each should have a probability of being realized.

Moreover, once a risk definition has been agreed upon, it is not always clear how risk itself should be measured. Consider, it may be the case that different patient's records may have different value. And, as our research has demonstrated (9) - if an individual is rationally motivated, then they will likely go for the record yields the best return for their investment in attack. For instance, the medical record of the chairmen of a Fortune five company is likely to be of greater interest than others individuals (e.g., because of the potential for sale of such records to popular media outlets). However, it is tricky to determine the precise value of health data, such that the default approach to risk assessment is to treat all records as equal in terms of their worth. There is no specific guidance around how to prioritize records in terms of protection, that is, assuming that they even should be.

3. What are the points of confusion or challenges related to HIPAA? What are options for resolving these?

There are several points of confusion that could be further refined to help in risk assessments. The first is the question of time-limited determinations. The guidance states that a determination may be limited in its length by the expert. However, there is no indication of what such a length could, or should, be? For instance, it could be left as short as one day, one year, five years, or indefinite. There should further be some resolution on how to transfer a determination from one expect to another in the event that the former is no longer available to perform the assessment.

A second question is about how new types of methods relate to the various implementations.

- **Perturbation (e.g., date shifting)**: Safe Harbor states that dates no smaller than one year in length can be disclosed. However, does this mean that the dates must be reported at the level of year only? Or can they be perturbed, such that they are shifted by a random value between 1 and 365. We have published a method to provide support for such perturbation, but given the potential for a semantic break between certain information in the record (e.g., release and revocation of a drug from market) and the randomly shifted dates, it is unclear if such a method can be directly applied to achieve Safe Harbor (10). Clearly, there are similar issues that can transpire with the perturbation of an individual's geographic locale.

- **Hiding Plain Sight:** Natural language processing techniques are never guaranteed to have 100% performance. The amount of effort in bringing such techniques the final mile towards perfection leads to an exponential growth in terms of investment and computation (11). As an alternative, it has been suggested that (and we have developed software to support), rather than attempt to redact every potential identifier in natural language, we replace redacted terms with "fake" information to hide the real leaks (12). However, this could mean that real patient names could be revealed. The rate at which this could occur might be small (e.g., less than 1 in 100 instances), but it could be interpreted that this is a scenario in which data is put in harm's way. We have conducted evaluations with human readers (as well as machines that mimic the software tools used in redaction) to show that the likelihood of discerning the difference between real and fake identifiers is no better than random (12,13), but Institutional Review Boards (IRBs) must still agree to permit the application of such an approach. In this setting, it would be helpful to have guidance provided directly to IRBs, or allow for data privacy determinations to be performed outside of the IRBs.

4. What is your perspective of oversight for unauthorized re-identification of de-identified data?

There is limited oversight with respect to the use of de-identified data. However, it should be recognized that this is how HIPAA was architected. De-identified data is no longer subject to the rule and is thus not covered. There is clearly an opportunity to provide better oversight in the use of de-identified data, such as how audits have been ramped up with respect to compliance with the HIPAA Security Rule over the past several years. Still, doing so will require either guidance on self-regulation in the use of such data (by covered entities and the recipients) or a change in the scope of HIPAA to allow for oversight with respect to de-identified data.

5. What are the current gaps in rules (laws, regulations, self-governance regimes, best practices) across sectors, or with respect to de-identification?

Since de-identified data, by definition, is no longer covered by HIPAA, the notion of an unauthorized re-identification is meaningless. Technically, there is nothing to prevent re-identification attempts on such data. This is why the guidance noted that experts often tie the sharing of de-identified data to data use agreements (DUAs) that preclude re-identification attempts. Violating such a contract would not be a criminal penalty, but it is likely to deter ill behavior due to fear of civil penalties. However, again, the extent to which such an agreement is a viable deterrent will be due to the perceived value of the data and the moral scruples of the individual.

6. **What recommendations would you make to help keep policy at pace with or ahead of technology (e.g., outreach, education, technical assistance, a policy change, or guidance)**

I will begin by saying that, while my comments have pointed out some of the confusion and challenges around de-identification, I believe that it is generally working well. While re-identification and/or misuse of de-identified records may take place behind closed doors, the de-identification standard has been in place for over a decade and there is little evidence that re-identification is taking place and individuals are being harmed (14).

Still, the notion of expert determination would benefit from having greater standardization and transparency. Different experts are likely to perceive risks in different ways. This is natural, but it would be useful to the community of experts, as well as covered entities working with such experts, to have more clarity around best practices. This may be achieved through some non-profit, industrial coalition, or even a government standards body.

References

1. Workshop on the HIPAA Privacy Rule's De-Identification Standard. Information available at: http://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/2010-de-identification-workshop/index.html
2. Office for Civil Rights. Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. U.S. Department of Health and Human Services. November 2012. Available at: http://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html
3. The Vanderbilt Synthetic Derivative and BioVU. Information available at: https://victr.vanderbilt.edu/pub/biovu/
4. Project Data Sphere. Information available at: https://www.projectdatasphere.org/projectdatasphere/html/home
5. Meyestre S, et al. Automatic de-identification of textual documents in the electronic health record: a review of recent research. BMC Medical Research Methology. 2010; 10: 70.
6. Stubbs A, Kotfila C, Uzuner O. Automated systems for the de-identification of longitudinal clinical narratives: Overview of 2014 i2b2/UTHealth shared task Track 1. Journal of Biomedical Informatics. 2015; 58 Suppl, S11-S19.
7. Naveed M, Ayday E, Clayton EW, Fellay J, Gunter CA, Hubaux JP, Malin BA, Wang X. Privacy in the genomic era. ACM Computing Surveys. 2015; 48(1): 6.
8. Clayton EW, Malin B. Assessing risks to privacy in biospecimen research. In Specimen Science: Ethics and Policy Implications, MIT Press. In press.
9. Wan Z, Vorobeychik Y, Xia W, Clayton EW, Kantarcioglu M, Ganta R, Heatherly R, Malin BA. A game theoretic framework for analyzing re-identification risk. PLoS One. 2015; 10(3): e0120592.

10. Hripcsak G, Mirhaji P, Low AF, Malin BA. Preserving temporal relations in clinical data while maintaining privacy. Journal of the American Medical Informatics Association. 2016; in press. Preprint available at: http://jamia.oxfordjournals.org/content/early/2016/03/23/jamia.ocw001.long
11. Carrell D, Cronkite D, Malin B, Aberdeen J, Hirschman L. Is the juice worth the squeeze? Costs and benefits of multiple human annotators for clinical text de-identification. Methods of Information in Medicine. 2016; in press.
12. Carrell D, Malin B, Aberdeen J, Bayer S, Clark C, Wellner B, Hirschman L. Hiding in plain sight: use of realistic surrogates to reduce exposure of protected health information in clinical text. 2013; 20: 342-348.
13. Li M, Carrell D, Aberdeen J, Hirschman L, Kirby J, Li B, Vorobeychik Y, Malin B. Optimizing annotation resources for natural language de-identification via a game theoretic approach. Journal of Biomedical Informatics. 2016; in press. Preprint available at: http://linkinghub.elsevier.com/retrieve/pii/S1532-0464(16)30010-7
14. El Emam K, Arbuckle L, Jonker E, Malin B. A systematic review of re-identification attacks on health data. PLoS One. 2011; 6(12): e28071.

National Committee on Health & Vital Statistics Hearing on De-Identification & HIPAA:

Policy Interpretations of Guidance

Bradley Malin, Ph.D.

Associate Professor of Biomedical Informatics & Computer Science

Director, Health Data Science Center

Vanderbilt University

May 24, 2016

Conflicts of Interest

- 2009 – 2012:
  - I was a paid consultant to the Office for Civil Rights at the U.S. Department of Health and Human Services and was involved in the development of the HIPAA de-identification guidance
- 2006 – 2016
  - I consulted for numerous companies involved in the generation of de-identified data (none of which are represented today)

De-identification May Be Safe

- In 20111, we reviewed all **actual** re-identification attempts
- Attacks on health data
  - 14 published re-identification attacks on any types of data
  - 11 of 14 were conducted by researchers as demo. attacks
  - 10 of the 14 attacks verified their results
  - Only 2 of 14 attacks followed any standard
  - Only case with health data subject to "Safe Harbor" had a success rate of 0.00013

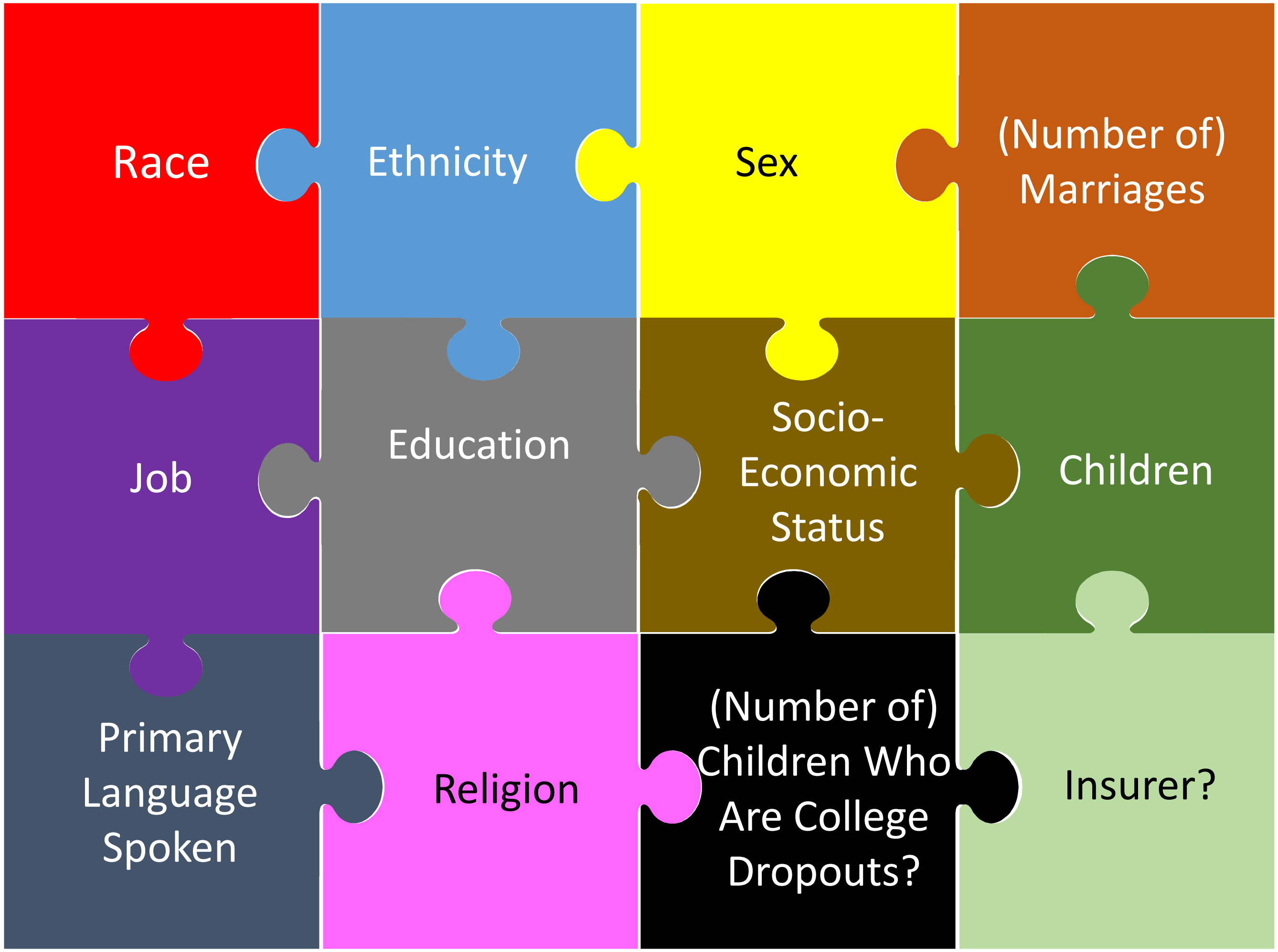
K. El Emam, et al. PLoS One. 2011.

A Case Study on Demographics (covered by Garfinkel)

- Details at http://www.ehcca.com/presentations/HIPAAWest4/lafky_2.pdf
- Challenge issued by U.S. Dept. Health & Human Services
- 15,000 Safe Harbor records provided to well-known academic team
- Team purchased public records from commercial broker
- Correctly identified 2 people

What issues do you see as being the **most pressing** when you consider de-identification and HIPAA?

Safe Harbor: Unlimited Features



Race

Ethnicity

Sex

(Number of) Marriages

Job

Education

Socio-Economic Status

Children

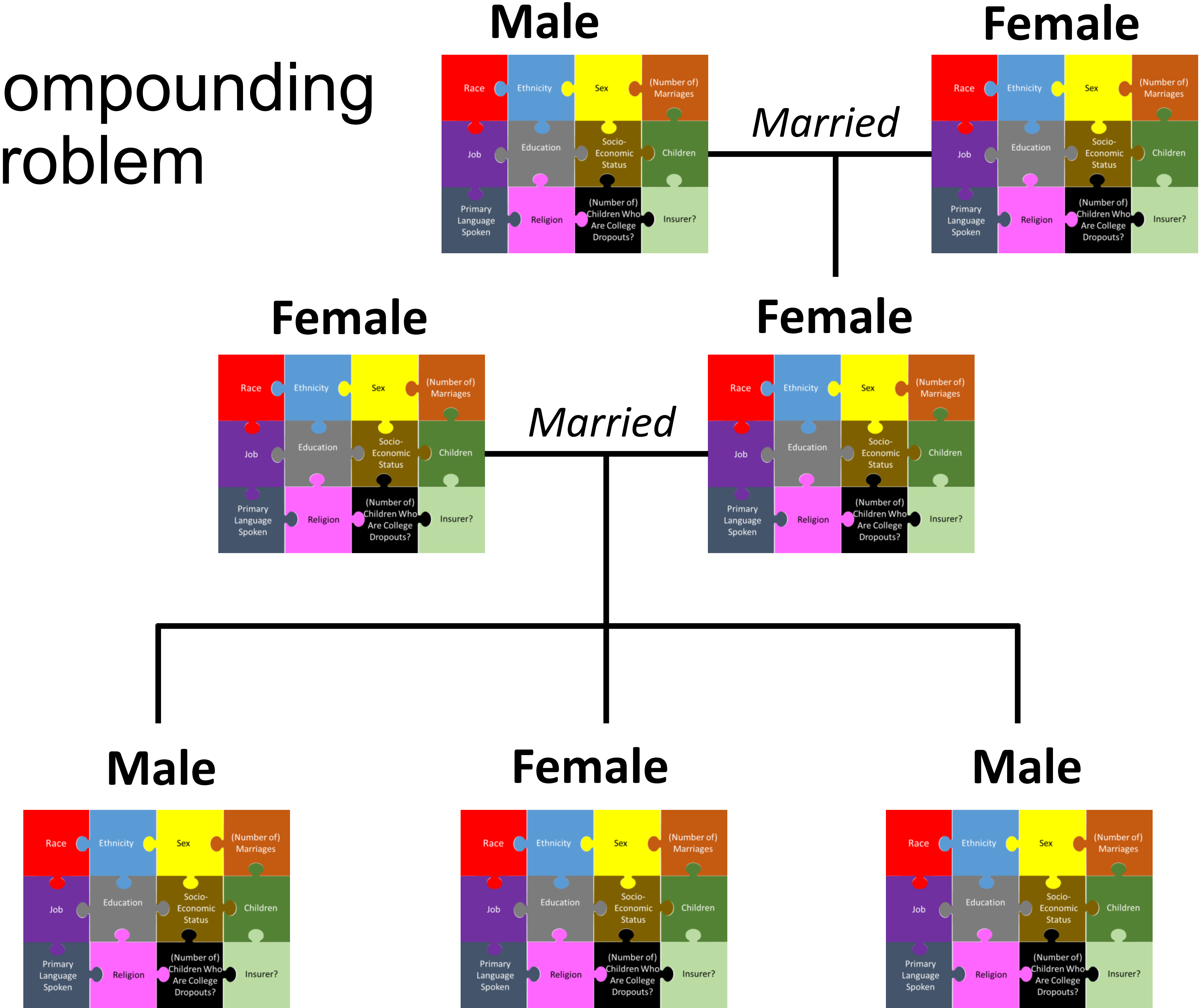
Primary Language Spoken

Religion

(Number of) Children Who Are College Dropouts?

Insurer?

A Compounding Problem



Male

Race
Ethnicity
Sex
(Number of) Marriages
Job
Education
Socio-Economic Status
Children
Primary Language Spoken
Religion
(Number of) Children Who Are College Dropouts?
Insurer?

Married

Female

Race
Ethnicity
Sex
(Number of) Marriages
Job
Education
Socio-Economic Status
Children
Primary Language Spoken
Religion
(Number of) Children Who Are College Dropouts?
Insurer?

Female

Race
Ethnicity
Sex
(Number of) Marriages
Job
Education
Socio-Economic Status
Children
Primary Language Spoken
Religion
(Number of) Children Who Are College Dropouts?
Insurer?

Married

Female

Race
Ethnicity
Sex
(Number of) Marriages
Job
Education
Socio-Economic Status
Children
Primary Language Spoken
Religion
(Number of) Children Who Are College Dropouts?
Insurer?

Male

Race
Ethnicity
Sex
(Number of) Marriages
Job
Education
Socio-Economic Status
Children
Primary Language Spoken
Religion
(Number of) Children Who Are College Dropouts?
Insurer?

Female

Race
Ethnicity
Sex
(Number of) Marriages
Job
Education
Socio-Economic Status
Children
Primary Language Spoken
Religion
(Number of) Children Who Are College Dropouts?
Insurer?

Male

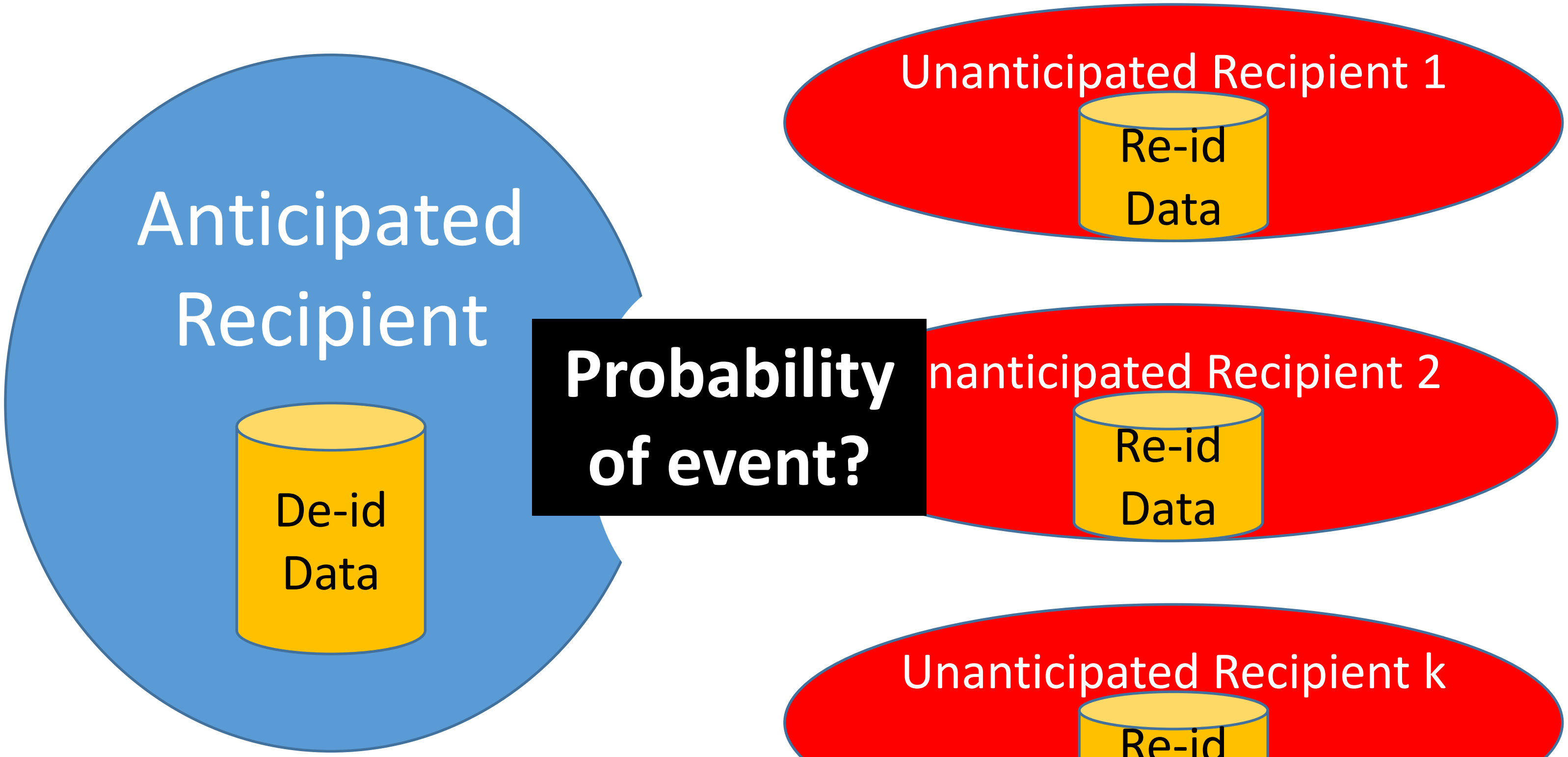
Race
Ethnicity
Sex
(Number of) Marriages
Job
Education
Socio-Economic Status
Children
Primary Language Spoken
Religion
(Number of) Children Who Are College Dropouts?
Insurer?

There’s only one church in the region with a deacon… but who knew?

This patient has a really interesting history… and is married to the deacon of the local church.

Experts: Anticipated Recipient?

- Wonderful concept → tailor risk to capabilities



*omics Data is High-Dimensional!

Only 1 person with these 500 SNPs

```
AG AT CA AA CA TG GC AA TA AT CA GT AG GG AT AT CC AC AG TT AA TA CT
CC AG CA GC AG AT AA CA GG TA GG GC GA AA GA GA GC CA TA CT CC CC CC
CA TT AG TT GT AA GG GG AT AC [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] TC AA CC AC AT
CG TG AC AT GG [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] TC TC AC AC
AT AT GA CA TG CT TA AC CT AG TC CC AC TT CA AA CG TG TA AA TT GA TA
TC CT AC AT CA CT CA CC TC AG TC TA CA CC CG TG TA TT AG GT CC TT TA
TG TA CC GA GG CC GG GG AG TT GG TT CA CG TT CC CC AA CG CC TC AT TG
AT AT AT CC CC GC CA CA TC AA GT CT GG CT CC AC CT TA GT TA AC GC CA
GT GA AA CT TG GC GC CA TT CG CA AG GG AG CC AC AT AG GC GT GG TA GG
GA AG TT TT [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] AT [illegible] [illegible] [illegible] [illegible] [illegible] CT [illegible] TG AT TT
TT AG CT GA [illegible] [illegible] [illegible] GC [illegible] [illegible] [illegible] [illegible] GG [illegible] TC [illegible] [illegible] [illegible] [illegible] TC CG GT GT
CG AC AA AT GG GT CC GT TT CG CA GA AT TT GC TA CA TA AC AC AC AG TC
TG AC TG TA GT AC GT TG TG AA GG TG CG TG CT AG TG CG GG CC AA CC GG
AG GA GG GT CG TC TG CC GT AC TT CC AG CT AT AT TG TT AG GG TA AA AG
AA GT CA AC GC CC TT TC GA GC GT GG AT GT CG TC AC AT GA AC AT CT GA
[illegible] CA GA CA GA CA AT AT TG AC GG CC CA AT [illegible] CT TT [illegible] GT CT AG CT CG
CA [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] TA [illegible] [illegible] AG [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
CG GG TC AG TA GC TC CT GA AG GG AC TA AT TA CT TC TG TT GT GT GC GT
CC CT TC GC TA GG AC GA AT CT TG GA TG GG AA AA TA AT CT CA AA AC CC
TC AT CG AG GG TG CC GC GT AG AC CT TC CA AG GC GA TA TG AC TA AA CG
AA GA CA GA TA TG [illegible] AT [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] CC TG TA TG TT
GT CC CT AC TT [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
```

But who is this?

Premature to Designate as Identifier

Lin, Owen, & Altman. Science 2004.

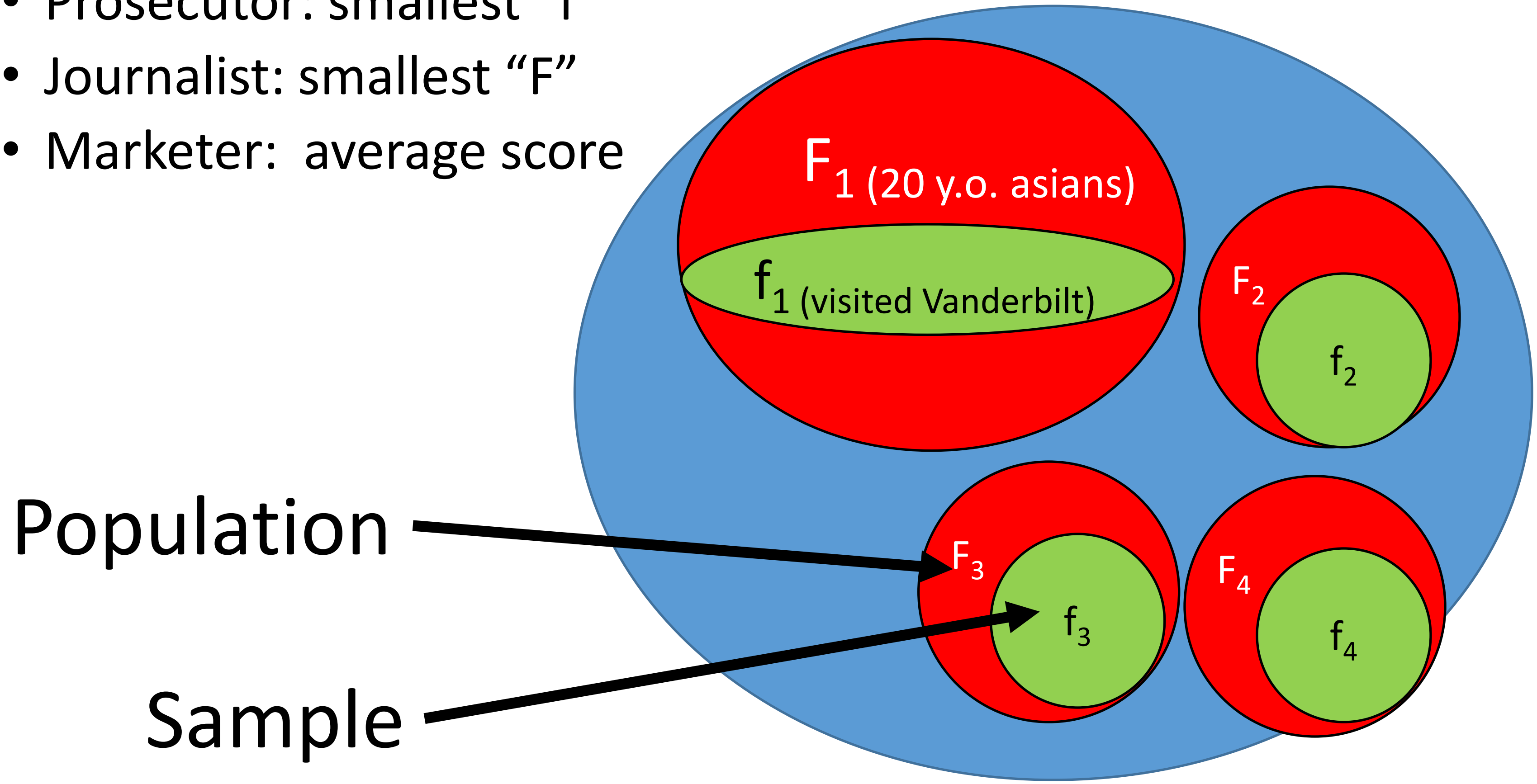
Is the current HIPAA de-identification guidance sufficient?

Does it pose challenges to or does it advance the use of data (including aggregation, analysis, dissemination, sharing) **in healthcare?**

Risk…

… means something different to everyone

- Can be modeled in various ways
  - Prosecutor: smallest “f”
  - Journalist: smallest “F”
  - Marketer: average score



F_1 (20 y.o. asians)

f_1 (visited Vanderbilt)

F_2

f_2

F_3

f_3

F_4

f_4

Population

Sample

Does all Data Have the Same Value?

- Risk is proportional to the “worth” of the data to the recipient

Record A	$$$
Record B	$$
Record C	$

- But sets pricing?

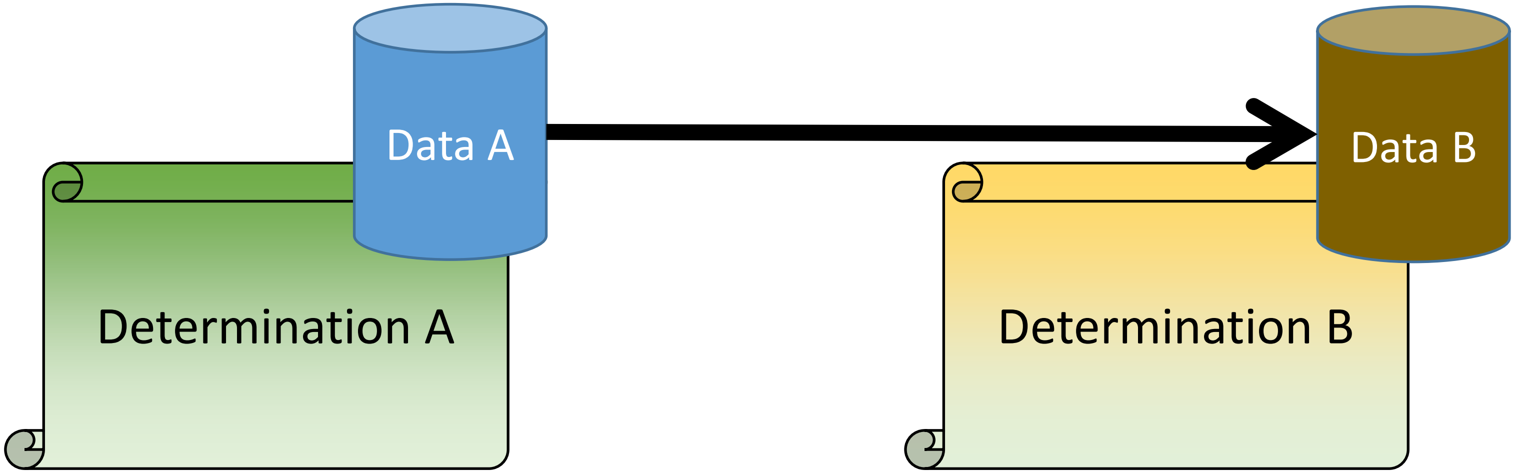
What are the points of confusion or challenges related to HIPAA?

What are options for resolving these?

Time-Limited Determinations

- Experts are permitted to limit the warranty of a certification
- What is an appropriate about of time?
- What happens when time-limitation expires?
    - Can we continue to use data already shared?
    - What happens if original certifier is no longer available?

Who is Responsible for Integrating Multiple Certifications?



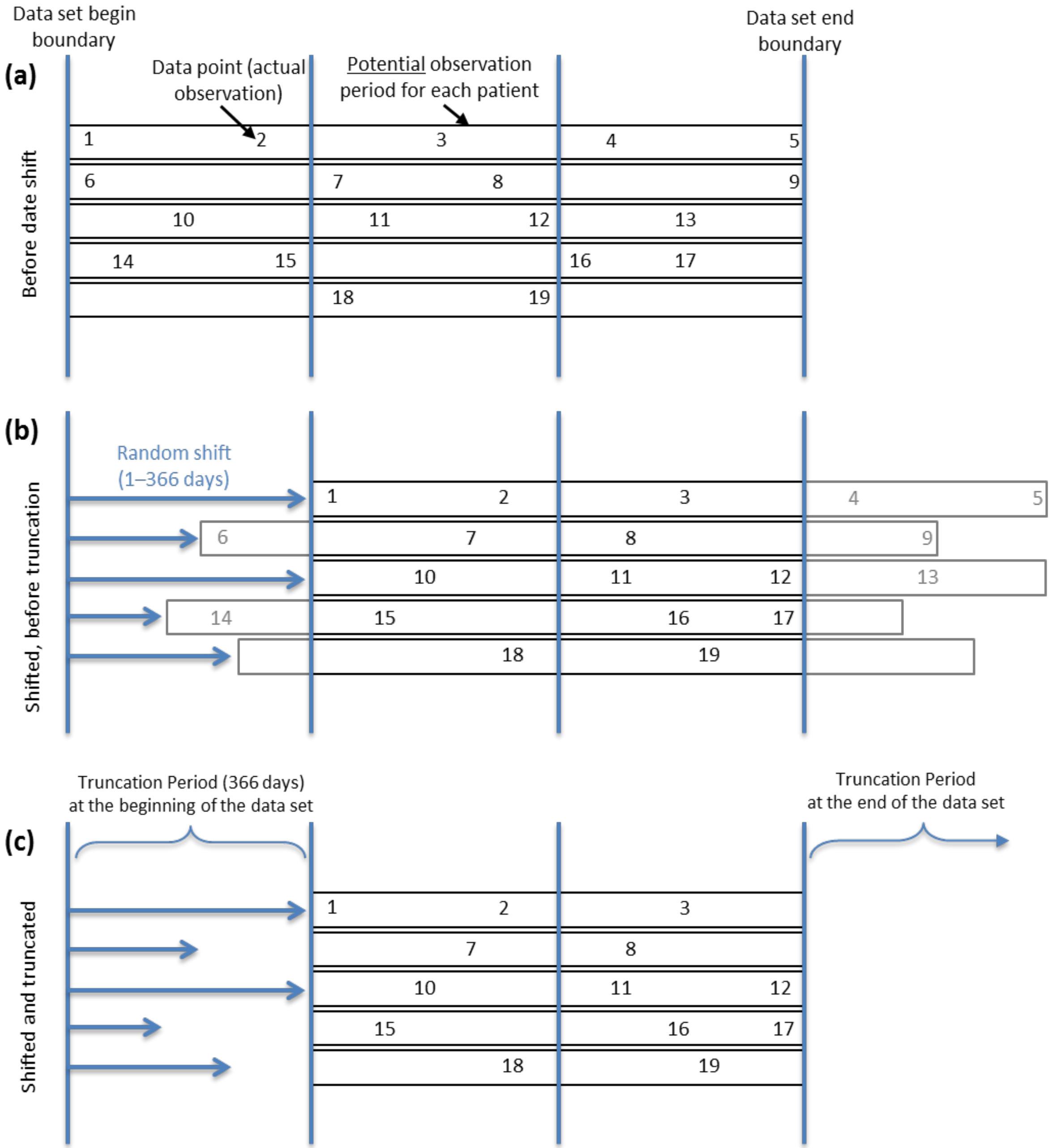
- B wants to integrate with A
- Is it Expert A’s responsibility?
- Is it Expert B’s responsibility?
- Is it both Expert A and Expert B’s responsibility?

Perturbation?

- Safe Harbor states that dates must be no more specific than one year
- Can dates be “shifted” and meet this requirement?

Hripcsak et al., JAMIA 2016

Shift & Truncate


Data set begin boundary
Data set end boundary
(a)
Data point (actual observation)
Potential observation period for each patient
Before date shift
1 2 3 4 5
6 7 8 9
10 11 12 13
14 15 16 17
18 19
(b)
Random shift (1–366 days)
Shifted, before truncation
1 2 3 4 5
6 7 8 9
10 11 12 13
14 15 16 17
18 19
Truncation Period (366 days) at the beginning of the data set
Truncation Period at the end of the data set
(c)
Shifted and truncated
1 2 3
7 8
10 11 12
15 16 17
18 19

Hripcsak et al., JAMIA 2016

Perturbation?

- Safe Harbor states that dates must be no more specific than one year
- But can dates be "shifted" and meet this requirement?
- Semantic Breaks: What happens if there are semantic breaks (e.g., drug released before date reported

Hripcsak et al., JAMIA 2016

Hiding in Plain Sight

Original PHI

**Redacted PHI & Leaked PHI

Surrogate PHI & Hidden PHI

Smith, **61** yo ...
daughter, **Lynn**, to ...
oncologist **Dr. White** ...
5/13/10 to consider ...
SWOG protocol **1811**, ...
was randomized **5/10** ...
to call **Mr. Smith** on ...
PLAN:**Dr White** and I ...

Evidence suggests small chance of exposure …
but will IRBs accept the risk?

Carrell et al., JAMIA 2013

What is your perspective of oversight for unauthorized re-identification of de-identified data?

De-identified data has no oversight

Several Options

- Oversight can be made contractual (under expert determination)
- Oversight can be made a government initiative (similar to HIPAA Security Rule) … but at a cost
- How can you show misuse of de-identified data?

What recommendations would you make to help keep policy at pace with or ahead of technology (e.g., outreach, education, technical assistance, a policy change, or guidance)?

Opportunities

- Providing for some oversight of de-identified data (should it really be a free for all?)
- Clearinghouse for best practices in de-identification
- Agreement (or authority) for setting “small” risk

Clarifications?

Questions?

b.malin@Vanderbilt.edu