Towards Unified Data Security Requirements for Human Research

Susan Bouregy, Ph.D., CIP
Chief HIPAA Privacy Officer
Vice Chair, Human Subjects Committee
Yale University
susan.bouregy@yale.edu
March 21, 2013

Social Science, Behavioral and Educational Research at Yale

- Approximately 600 active protocols
- Broad range of studies
  - Cognitive development in children
  - The role of faith for individuals in conflict zones
  - Randomized trials of economic interventions in developing countries
  - Video ethnographies of marginalized communities

Yale as a HIPAA Covered Entity

- Hybrid Entity
  - Faculty practice, health care clinic, self-insured health plan
  - 20,000 faculty, staff, students, etc are required to comply
- Research conducted by faculty, staff and students in the covered entity are required to comply.
- Research conducted by faculty, staff and students outside the covered entity only deal with HIPAA when attempting to access health information from a covered entity.
- Most SBE projects are conducted outside the HIPAA covered entity.

Section V. Strengthening Data Protections to Minimize Informational Risks

- Harmonizing concept of individually identifiable
- Require data security protections indexed to identifiability
- Use HIPAA security and breach notification standards as model for protection scheme

Informational Risk

- What could happen if participant's spouse, parent, boss, friends, police department, found out what he/she said?
- Potential for deductive disclosure
- Context dependent
- Population dependent

Proposed Changes

- Adopt the HIPAA standards for purposes of the Common Rule regarding what constitutes individually identifiable information, a limited data set, de-identified information.

Definitions

- HIPAA: individually identifiable health information: identifies the individual or with respect to which there is a reasonable basis to believe the information can be used to identify the individual

- CR: individually identifiable private information: the identity of the subject is or may readily be ascertained by the investigator or associated with the information

De-identification

- HIPAA:
  - Strip 18 defined identifiers
  - Statistical determination of very small risk of re-identification
- CR:
  - Undefined

Impacted Data Sets

- Household surveys or ethnographic interviews that include zip codes of the respondents
- Cognitive development data including dates of birth
- Epidemiological data set including date of vaccination
- Linguistic studies of endangered languages with limited numbers of speakers identified by country

- Data security requirements would apply
- Could not be deemed exempt

Issues

- No single accepted term in the literature
- Currently confusion on the part of IRBs and investigators regarding de-identified vs anonymous
- More information would be considered identifiable under HIPAA definition
- Focus on “individually identifiable” ignores community risks

Proposed Changes

- Mandate data security and information protection standards that would apply to all research that collected, stored, analyzed or otherwise reused identifiable or potentially identifiable information.
- Data security and information protection standards would be scaled appropriately to the level of identifiability of the data.

Key HIPAA Security Elements

- Encryption of data at rest (laptops, desktops, thumbdrives, smart phones etc.)
  - Export control issues in some locations
- Secure transmission of data (email encryption, secure file transfer)
  - Not user friendly
- Strong physical security
  - Can be practical issues in remote field locations
- Access controls and logging
  - Cloud storage issues

Issues

- Suitability of IRB for determinations of appropriate data security plan
- Proposed rule applies standards to all data including that from “excused” research and would apply to all institutions that have some federal funding
- Not all identified data is risky
- Not all studies promise confidentiality
- Some participants request attribution
- Costly

Identifiability vs. Sensitivity

- Identified interviews with current or former combatants regarding actions in local communities
- Identified data on participation in local elections in the US
- Identified data on participation in elections in emerging democracies

Recommend Guidance for IRBs and PIs

- IRB best suited for determining risk of harm
- PI best suited for determining what is manageable in the field.
- Provide guidance on solutions for low, medium and high risk data

Proposed Incorporation of HIPAA Breach Notification Requirement

- Breach: the acquisition, access, use, or disclosure of PHI in a manner not permitted under subpart E of this part which compromises the security or privacy of the PHI

- Presumed to be a breach unless the covered entity demonstrates that there is a low probability that the protected health information has been compromised based on a risk assessment

IRB Adverse Event and Unanticipated Problem Reporting

- Data breaches qualify as AE/UAPIRSO
- Consideration of notice to participants as part of risk mitigation strategy
- Common considerations:
  - Extent of possible harm
  - Ability to further mitigate harm based on awareness
  - Autonomy considerations

Issues

- HIPAA breach standards are more stringent, require reporting more incidents
- Costs of investigation and notice
- Incident fatigue
- Utility of providing awareness of events for which there is no preventative action that can be taken
- Utility of providing information transnationally when the risk is local/contextual
- Providing notice in studies conducted under a waiver of consent
- Harm arising from notice itself based on association with the study.

Conclusions

- Applaud an effort to harmonize terminology around identifiability of data
- Applaud an effort to provide a mechanism for IRBs to minimize informational risks
- Informational risks are not sufficiently correlated to identifiability alone to allow indexing data security needs to presence of identifiers
- The costs of data security and breach notification requirements must be justified by the anticipated risks to the data and benefit of the notice
- The diversity of SBE research requires that the risk mitigation strategy be flexible