Data Security Considerations for Research

Institutional Review Board
Annual Education

May 8, 2012

PRIVACY vs. SECURITY

What's the Difference?:

**PRIVACY**

Refers to **WHAT** is protected — Health information about an individual and the determination of **WHO** is permitted to use, disclose, or access the information

**SECURITY**

Refers to **HOW** private information is safeguarded—Insuring privacy by controlling access to information and protecting it from inappropriate disclosure and accidental or intentional destruction or loss and assuring it is available when and where it is needed.

Protected Health Information (PHI)

🌟 Protected Health Information is any information that:

- is created or received by a health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse; and
- relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual

- All information whether maintained in electronic, paper or oral format
HITECH Act (ARRA)
Health Information Technology for Economic and Clinical Health

  - Applies to all electronic “unsecured PHI”
  - Requires immediate notification to the Federal Government if more than 500 individuals effected
  - Annual notification if less that 500 individuals effected
  - Requires notification to a major media outlet
  - Breach will be listed on a public website
  - Requires individual notification to patients

- Criminal penalties - apply to individual or employee of a covered entity

- Encryption is a safe harbor for breach of data

Breach Notification:
500+ Breaches by Location of Breach

September 2009 – March 2012
409 reports involving over 500 records/individuals

CMPs Increased
45 CFR 160.404 - Amount of a Civil Money Penalty

<table>
<thead>
<tr>
<th>For violations occurring prior to 2/18/2009</th>
<th>For violations occurring on or after 2/18/2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penalty Amount</td>
<td></td>
</tr>
<tr>
<td>Up to $100 per violation</td>
<td>$100 to $50,000 or more per violation</td>
</tr>
<tr>
<td>Calendar Year Cap</td>
<td></td>
</tr>
<tr>
<td>$25,000</td>
<td>$1,500,000</td>
</tr>
</tbody>
</table>

OCR may reduce a penalty if the failure to comply was due to a reasonable cause and not willful neglect, and the penalty would be excessive relative to the noncompliance.

OCR
Impact of HITECH Act on research activities:

- The primary implication for researchers is to ensure their research incorporates necessary safeguards when using health information.

- When possible, the investigator should implement measures that render the data into a form that no longer qualifies as "unsecured PHI".

- Data in this form obviates the need for notification if a breach occurs. Methods recognized in the HITECH Act for achieving this purpose include:
  - de-identification: once data has been de-identified in accordance with the HIPAA Privacy Rule, it no longer qualifies as PHI;
  - encryption of data with processes meeting National Institute of Standards and Technology (NIST) criteria is considered by the HITECH Act to render the data secured;
  - destruction of media on which PHI is stored following the standards in the HITECH Act.

Source: http://www.hhs.gov/ocr/privacy/hipaa/newsinitiatives.html
Key HIPAA Terms for Researchers

- Business Associate Agreement
- Use – Sharing within the Organization
- Disclosure – Sharing outside the organization
- Data Use Agreement / Limited Data Set
- Coded Data
- Waiver / Authorization
- De-identified – without 18 identifiers
- Password does not = Encryption

Why should research data be de-identified?

- De-identified health information is not considered to be PHI and therefore, subject to the HIPAA Security Rule reporting requirements if a privacy/security breach occurs.

18 identifiers as defined by the HIPAA Privacy Rule:

1. Name
2. Geographic Location (including city, state, zip)
3. Elements of Dates
4. Telephone Number
5. Fax Number
6. E-mail Address
7. Social Security Number
8. Medical Record or Prescription Numbers
9. Health Plan Beneficiary Number
10. Account Number
11. Certificate/license Number
12. VIN and Serial Numbers, License Plate Number
13. Device identifiers, serial numbers
14. Web URLs
15. IP Address Numbers
16. Biometric identifiers (finger prints)
17. Full face, comparable photo images
18. Unique Identifying numbers
De-identified vs. Coded Data

Coded data
- contains an assigned code so even though the information has been stripped of identifiers, the health information can be linked back to the individual by members of the research team.

De-identified data
- Stripped data with no code.
- It cannot be linked back to the individual.
- A re-identification code can be assigned to a de-identified dataset by a covered entity; however, members of the research team may not have the access to the means/method of re-identification.
- In the case where the research team has access to the re-identification means/method, the data is not considered to be de-identified.

Business Associate

- Business Associate - a person or entity that performs certain functions or activities that involve the use or disclosure of protected health information on behalf of, or provides services to, a covered entity.
- Business Associate - 45 CFR 160.103.
  - Claims processing or administration; data analysis, processing or administration; utilization review; quality assurance; billing; benefit management; practice management; and pricing.
  - Business associate services are: legal; actuarial; accounting; consulting; data aggregation; management; administrative; accreditation; and financial.
- Is an agreement needed?
- Who can sign agreement?
- Who had a record of BAA’s at my organization?
- Also in Security Rule & Definition revised by HITECH

Fundamentals of Information Security

- Confidentiality
  - Keeping data out of unauthorized hands

- Integrity
  - Preventing data from changing in unknown, unwanted ways

- Availability
  - Ensuring data are accessible when expected
What is Risk?

- A vague and general term. Speaking generically...

\[ \text{Likelihood of occurrence} \times \text{Impact if occurred} = \text{RISK} \]

- Business Risks help organizations focus in on the corresponding Technical Risks.

Business Risks Examples

- Theft of confidential information could result in financial or medical harm
  - Financial and insurance data have higher impact costs to institution and patient if lost, stolen or otherwise compromised.

- Theft of research data could lead to theft of intellectual property
  - Plenty of examples of 'advanced persistent threats' targeting world class institutions for espionage. Major impact to researcher.

Business Risks Examples

- Internal user mistake of data set can lead to loss of data.
  - If integrity research data is unknowingly manipulated, months or years of effort could be lost.

- Crash of system hosting data set can lead to significant loss of time, or research
  - The availability of systems is becoming more important for furthering research.
Not all Risks “owned” by the same group.

Data Flows in many ways... and each of these ways can induce risk.

Risk Management Overview
Before We Start...

• Firewalls are not a silver bullet

• “inside” and “outside” the firewall are misleading. Firewalls must allow traffic through.
  – Web (as we will see) is primary example

---

Use Case 1: Use of mobile device to store/transmit PHI/PII

Technical Risks

• Loss of laptop/mobile device with PHI
  – Hard to be called “technical” but accounts for 52% of reportable breaches from 2009-2012.
  – #1 biggest risk on today for PHI
• Poor password management exposes PHI
  – Sticky note with laptop password; physical access
• Client device with PHI “hacked”
  – Can occur through multiple methods including:
    • Drive-By Downloads
    • “Clickjacking”
• Email contains PHI and stored on laptop

---

Use Case 1: Use of mobile device to store/transmit PHI/PII

Solutions

• Full Disk Encryption!
  – OCR specifically states it wants to push this agenda for all systems that contain/process/transmit PHI.
  – An encrypted lost device means no breach.
  – Windows: BitLocker, PGP, GuardianEdge
  – Macintosh: FileVault2 (OSX Lion only), PGP
• Anti-Virus
  – Symantec, Sophos, Microsoft Security Essentials
• Host-Based Intrusion Prevention
  – Symantec SEP
• Proper use of Passwords
  – Try Password Vaults (KeePass, or PasswordSafe, both free)
Use Case 2: Sharing PHI via the Web

**Technical Risks**

- Loss of unencrypted data to Internet due to lack of access control
  - Google search engines & caching
- Information stolen in transit
  - No encryption = easily “sniffed” traffic
- Vulnerable services on server exposed to Internet
  - Allow for exploitation; both web and web application
  - #2 biggest risk on the Internet today in regards to PHI

**Solutions**

- Encrypt data & transmission protocols:
  - Leverage SSL encryption for HTTPS
  - Leverage AES encryption for flat files when sharing, to ensure end-to-end encryption
- Enforce Access Control
  - Minimize number of users on the system
  - Re-certify access privileges on scheduled basis
- Enable access logging and audit trails!
  - BIG OCR point. Enable audit logs showing access to PHI.
  - Review, and demonstrate you are reviewing, audit logs

**Solutions (round 2)**

- Physical security!
  - “Locked in an office” is not security. Servers should be in server rooms with badge readers, video cameras, and access logs. Otherwise it must be full disk encrypted (in essence the same security level as desktop)
  - Remember, theft is #1 problem. Theft from offices is very common
- Security “aware” system administration
  - If running a server you should know security basics.
    - www.sans.org/reading_room
Use Case 3: Use of the Cloud for Storage

Technical Risks

- Unencrypted data on cloud
  - Publicly available data are viewable (public Google/Yahoo calendars)
  - Data are viewable by cloud service providers (Dropbox)
- Cloud “encryption” solutions does not encrypt from “cloud eyes”
  - If the cloud provider encrypts the data, they hold the key to decrypting.
  - Free providers use your data for data mining, part of the EULA (Google does this)
- Regulatory issues of non-approved disclosures (BAA)

Use Case 3: Use of the Cloud for Storage

Solutions

- Either 1) establish Business Associate Agreement, or
  - Legal document that extends HIPAA compliance to vendor
  - Required by HIPAA/HITECH, even if BA only stores and doesn’t access.
- 2) encrypt all data resident on cloud with key not available to cloud provider
  - The KEY is “key”.
- Use “Private” cloud services with appropriate contracts

General Solutions and Observations
Password Reminders

- All devices (cell phones, laptops, USB drives etc.) must have a password
- Do not share your password
- Do not log on to a computer and then allow someone else to do their work
- Log off application before leaving the area

btw, passwords of “123456” or “qwerty” or “trustno1” are in the top 10 passwords.

What Is My Role in Protecting Medical Information?

Good Security Standards follow the “90 / 10” Rule:

- 10% of security safeguards are technical
- 90% of security safeguards rely on the computer user (“YOU”) to adhere to good computing practices
  - Example: The lock on the door is the 10%.
  - You remembering to lock,
  - check to see if it is closed,
  - ensuring others do not prop the door open,
  - keeping controls of keys is the 90%.
  - 10% security is worthless without YOU

An example of how CUMC executes it’s Risk Management Program
Just to drive home the point...

OCR really is serious about encryption

Appropriate Safeguards Prevent Breaches

- Evaluate the risk to e- PHI when at rest on removable media, mobile devices and computer hard drives
- Take reasonable and appropriate measures to safeguard e- PHI
  - Store all e- PHI to a network
  - Encrypt data stored on portable/movable devices & media
  - Employ a remote device wipe to remove data when lost or stolen
  - Train workforce members on how to effectively safeguard data and timely reporting of incidents

Questions?
Erik Decker
Assistant Director, Information Security
Columbia University Medical Center
ed2513@columbia.edu

Karen Pagliaro-Meyer
Privacy Officer
Columbia University Medical Center
kpagliaro@columbia.edu

FOR ADDITIONAL INFORMATION:
http://privacyruleandresearch.nih.gov