



## STANDARD OPERATING PROCEDURE ADDENDUM

### Security and Integrity of Human Research Data

#### A. Definitions

1. **Human research data:** information captured in the course of conducting human research as defined by federal regulations and IRB policies.
2. **HIPAA:** the Health Insurance Portability and Accountability Act of the United States.
3. **Personally Identifiable Information (PII):** information that can be used to uniquely identify a single person or group of individuals. Examples include an individual's name and Social Security Number, Driver's license, non-Driver's license identification number or financial account information.
4. **Protected Health Information (PHI):** as defined under the Health Insurance Portability and Accountability Act (HIPAA) and in Penn State Hershey Hospital Administrative Manual (HAM) Policy C-18, is PII concerning the health status, provision of care, or payment for care.

#### B. Policy

It is the policy of the Penn State University College of Medicine (PSU) Institutional Review Board (IRB) to review the data security and integrity plans for all research studies involving human subjects to ensure the protection of confidential information of research participants and to ensure the integrity of the data. This policy defines the following 3-level categorization schedule for human research data and accompanying sets of security and integrity measures to protect the human research data.

Level 1 – De-identified research information about people and other non-confidential research information.

Level 1 information is research information in which all of the eighteen HIPAA identifiers that could be used to identify an individual have been removed and is referred to as “de-identified research information”. Level 1 includes coded data which do not have any of the 18 HIPAA identifiers (The list linking code numbers to identifiers is considered to be Level 2 or Level 3 human research data and must be stored according to the requirements of that level.) See Appendix E for the eighteen categories of information that must be removed in order to de-identify data.

There are no specific data security requirements for the protection of de-identified research information or for other non-confidential research information, but researchers may want to protect such data for their own reasons, i.e., keeping data private until a paper about the data is published. The data security recommendations for Level 1 information are listed in Appendix A and the data integrity measures for Level 1 information are defined in Appendix D.

## Level 2 – Non-sensitive information about individually identifiable people

Level 2 information includes individually identifiable information, disclosure of which would not ordinarily be expected to result in material harm, and for which a subject has been promised confidentiality.

Data security requirements for Level 2 information are defined in Appendix B. The data integrity measures for Level 2 information are defined in Appendix D.

## Level 3 – Sensitive information about individually identifiable people

Level 3 information includes individually identifiable information that, if disclosed, could reasonably be expected to be damaging to a person's reputation, present risk of civil or criminal liability, psychological harm or other significant injury, loss of insurability or employability or social harm to individuals or groups.

Data security requirements for Level 3 information are defined in Appendix C. The data integrity measures for Level 3 information are defined in Appendix D.

A. The data security and integrity plan for a proposed research study involving human subjects is evaluated by the IRB in order to determine that the plan fulfills the requirements for the applicable security category(ies) and provides adequate measures to protect the integrity of the human research data.

B. The IRB has the authority to approve variances from the data security and integrity requirements that would apply to a study given its security category(ies), so long as the resulting plan complies with any legal requirements and does not increase the risk to the participants, or jeopardize the integrity of the research data. The IRBs may seek the advice and recommendations of appropriate Penn State Hershey technical experts in assessing the adequacy of provisions to maintain confidentiality of data and in approving a security category level.

C. If human research data are subject to security requirements specified in an information use agreement (such as data use or business agreements), grant, contract, or research protocol, those requirements must be met. Should the IRB have concerns that additional protections may be necessary, it may consult with Penn State Hershey technical experts and may impose requirements appropriate to the level of sensitivity of the information. If there are no security requirements specified in a data use agreement, grant, contract, or research protocol, the appropriate level of security and protection is determined by these data security and integrity policies.

D. Research study data security and integrity plans are reviewed by the expedited review process by an IRB Chair or designated, experienced IRB member.

E. In conducting the review the IRB may rely on the investigators' information provided in the data security and integrity plan, may request confirmation that there has been a satisfactory information security office review, or may take other actions as appropriate to the sensitivity of the information and the applicable security category.

F. Compliance with data security and integrity plans is monitored by the Research Quality Assurance Office as part of routine or directed post-approval reviews according to the standard operating procedures of that

office. Any variances from the approved data security and integrity plans are reported to the IRB according to the Research Quality Assurance Office's standard operating procedures.

G. Breaches in confidentiality of research data must be reported promptly to the IRB according to the IRB Standard Operating Procedure Addendum: Reporting and Review of Unanticipated Problems Involving Risks to Participants or Others.

## **C. Procedure**

This procedure provides guidance for submission, review and approval of data security and integrity plans.

### **I. Investigator Responsibilities**

A. Investigators are responsible for: (1) disclosing the nature of the confidential data they collect so the IRB can assess the data security risk; and (2) preparing study data security and integrity plans and procedures in accordance with the appropriate security category(ies) requirements. For all research involving human research data, a data security and integrity plan must be submitted to the IRB as part of the initial IRB application, or as part of the Continuing Progress Report (if requested).

B. Upon confirmation by the IRB of the appropriate security level(s), investigators are responsible for implementing and monitoring the data security and integrity plans over the course of their projects. If human research data are stored electronically, investigators are responsible for ensuring that computers and other devices that are used to store human research information are set up correctly and operated in a manner that meets the requirements of that level. Researchers may consult with Information Technology (IT) to help them understand and meet the requirements.

C. Investigators are responsible for ensuring that all research team members with access to Level 2 or Level 3 human research data have signed a confidentiality agreement which is available in policy C-01 HAM. Note: Penn State College of Medicine and Penn State Hershey Medical Center staff, faculty and students have a signed confidentiality agreement on file in their applicable Human Resource department.

D. Investigators are responsible for ensuring that all research team members with access to human research data have completed training in the protection of human research subjects according to the IRB educational policy.

E. Investigators are responsible for reporting breaches in confidentiality of research data (such as loss of or inappropriate access to Level 2 or 3 human research data) promptly to the IRB according to the IRB Standard Operating Procedure Addendum: Reporting and Review of Unanticipated Problems Involving Risks to Participants or Others. In addition, investigators must report incidents involving identifiable health information to the Penn State Hershey Privacy Officer at (717) 531-2081.

### **II. IRB Responsibilities**

A. The IRB Chair or his/her designee review the data security and integrity plan for research studies during initial review and when requested during continuing review.

B. The possible determinations the IRB can make regarding the data security and integrity plans include:

1. Data security plan and integrity plan approved as submitted; or
2. Modifications required.

C. The IRB Chair or designee, documents his/her initial determinations regarding the data security and integrity plan on the IRB reviewer checklist.

D. The approval memo for the study confers the final approval of the data security and integrity plan by the IRB. For exempt research involving human research data, the exemption determination memo confers acceptance of the data security and integrity plan by the IRB.

### **III. HSPO Responsibilities**

A. An experienced IRB Coordinator reviews the submission for completeness and assigns it to an IRB Chair or designee for review.

## Appendix A

### Level 1 – De-identified research information about people and other non-confidential research information

#### Examples of Level 1 information

- De-identified data collected for a research study with no regulatory or contractual requirements
- Data consisting of publicly available information
- Coded data which do not include any of the 18 HIPAA identifiers (Appendix E). The list linking code numbers to identifiers is considered to be Level 2 or Level 3 human research data and must be stored according to the requirements of that level.

#### Data security recommendations for hardcopy (paper) data storage

- Research data forms should be stored securely in a controlled environments, e.g., at a Penn State College of Medicine or Penn State Hershey Medical Center facility.
- Conveyance of hardcopy research data forms should be double-wrapped so that damage to the outer container alone will not expose data and the delivery should occur using a secure chain of possession, such as commercial carrier or hand-delivery by a member or agent of the research team.

#### Data security recommendations for electronic data storage

- Good computer use practice that meets the following requirements should be used when storing Level 1 research information and access should be limited to those individuals who have a specific research need to access the information. These requirements include making use of complex passwords, not sharing accounts, and limiting system accounts to those with a specific need.
- All portable media are physically secured when not in use either in a locked office or using lock-down cables
- Servers must have access controls
- Data may be transferred by unprotected e-mail
- Electronic devices may be disposed of following deletion of files or disposal of documents in regular trash

## Appendix B

### Level 2 – Non-sensitive information about individually identifiable people

#### Examples of Level 2 information

- Data that include identifiable non-health, non-sensitive information collected as part of non-health-related survey research, interview or focus group research or education research
- Information that “re-identifies” an otherwise de-identified (Level 1) dataset and, in tandem results in a Level 2 dataset; (for example, a study may employ coded data that are meaningless outside the context of the study; the list that maps those coded data to meaningful identifiers is Level 2 data)

#### Data security requirements for hardcopy (paper) data storage

- Research data forms and/or linking code lists must be stored securely in a controlled environments, e.g., at a Penn State College of Medicine or Penn State Hershey Medical Center facility unless a variance is granted by the IRB.
- Conveyance of hardcopy research data forms must be double-wrapped so that damage to the outer container alone will not expose data and the delivery must occur using a secure chain of possession, such as commercial carrier or hand-delivery by a member or agent of the research team.
- Paper and other non-electronic copies must be shredded when no longer needed

#### Data security requirements for electronic data storage

- Good computer use practice that meets the following requirements should be used when storing non-sensitive research information and access should be limited to those individuals who have a specific research need to access the information. These requirements include making use of complex passwords, not sharing accounts, and limiting system accounts to those with a specific need.
- Level 2 information must be stored on the following electronic devices unless a variance is granted by the IRB:
  - Secure file server operated, supported and maintained by the Information Technology (IT) department or the Department of Public Health Sciences (PHS).
  - A secure data base server operated, supported and maintained by IT or PHS.
  - Any personal computer managed by IT or PHS
  - Any removable media that is tracked, inventoried and systematically managed.
- Desktops and devices must be physically secured, including locked offices and/or locked facilities with access restricted to study personnel and their guests.
- Level 2 information stored on personal computer managed by IT or PHS or on removable media must be encrypted as outlined in HMC policy C-37 HAM.
- Level 2 information must be encrypted in order to transmit it electronically. Conveyance of portable media must occur using a secure chain of possession, such as commercial carrier or hand-delivery by a member or agent of the research team.
- Level 2 information may be re-classified as Level 1 if all HIPAA-specified identifiers are removed and there are no agreements or laws that regulate the use of the de-identified information.
- Data should be routinely backed up and the back-up copy physically secured.
- Devices must undergo secure deletion of the disc at the end of life of the device or prior to recycling.

## Appendix C

### Level 3 – Sensitive information about individually identifiable people

#### Examples of Level 3 information

- Data that include identifiable health information collected for a clinical trial
- Data that include identifiable sensitive non-health information, such as test scores or student record information, collected as part of an educational research project
- Data that include identifiable non-sensitive research information linked to social security numbers
- Information that “re-identifies” an otherwise de-identified dataset and, in tandem results in a Level 3 dataset; (for example, a study may employ coded data that are meaningless outside the context of the study; the list that maps those coded data to meaningful identifiers is Level 3 data)
- De-identified data collected for a research study with regulatory or contractual requirements for data security

#### Data security requirements for hardcopy (paper) data storage

- Hardcopy research data forms and/or linking code lists must be stored securely in a controlled environments, e.g., at a Penn State College of Medicine or Penn State Hershey Medical Center facility unless a variance is granted by the IRB.
- Hardcopy research data forms and/or linking code lists must be stored in a locked file cabinet or limited access storage area (e.g., a locked room) when not in use.
- Records must be maintained identifying who has or had keys that allow access to the hardcopy Level 3 information.
- Level 3 information must be de-identified before sharing the information with members of the research staff whenever the identifying information is not necessary. All 18 HIPAA-specified identifiers must be removed for de-identification.
- Conveyance of hardcopy research data forms must be double-wrapped so that damage to the outer container alone will not expose data and the delivery must occur using a secure chain of possession, such as commercial carrier or hand-delivery by a member or agent of the research team.
- Paper and other non-electronic copies must be shredded when no longer needed

#### Data security requirements for electronic data storage

- Level 3 information must be stored on the following electronic devices unless a variance is granted by the IRB:
  - Secure file server operated, supported and maintained by the Information Technology (IT) department or the Department of Public Health Sciences (PHS).
  - A secure data base server operated, supported and maintained by IT or PHS.
  - Any personal computer managed by IT or PHS. All data must be encrypted.
- Any removable media that is tracked, inventoried and systematically managed may only be used for either long-term archival storage of Level 3 information or conveyance to another party.
- Level 3 information stored on personal computer managed by IT or PHS or on removable media must be encrypted as outlined in HMC policy C-37 HAM.
- A device not explicitly listed above is not deemed acceptable for storage of Level 3 information unless a special exception is granted by the IRB.

- Level 3 information may not be stored, temporarily cached or otherwise accessed in a way that creates a local copy of the data on so-called personal devices such as Personal Digital Assistants, USB-based portable devices (e.g., thumb drives, flash drives, or jump drives) or non-Penn State owned and managed devices of any kind (e.g., home computers, personal laptop computers, public computers).
- Remote displaying is permitted for remote access using applications, such as Citrix or Remote Desktop, where there are no persistent data copies when the programs are remotely displayed. Applications, such as most Email clients, which open an attachment by making a local copy of that document are not acceptable because a local cached copy of the document can persist on the user's computer indefinitely.
- Desktops and devices must be physically secured, including locked offices and/or locked facilities with access restricted to study personnel and their guests.
- Electronic devices must be set to automatically log-off and lock after defined periods of inactivity.
- Access Controls/Authorizations
  - The principal investigator must maintain a list of the individuals or the categories of people who are permitted to have access to Level 3 information.
  - Users' access to Level 3 electronic data must be removed if they no longer have a reason under the access policy to access the information, e.g., they change jobs or leave the institution.
  - Access to Level 3 electronic information must be logged. The logs must include the identity of the user, the time and the function (login or logout).
- Electronic Transmission: Level 3 information must be encrypted with at least the same level of encryption necessary to transmit other HIPAA-regulated data as outlined in HMC policy C-37 HAM. Level 3 information may only be transferred electronically in an encrypted state such as encrypted e-mail attachments or encrypted CD's or through web-based secure file transfer. Conveyance of portable media must occur using a secure chain of possession, such as commercial carrier or hand-delivery by a member or agent of the research team.
- Level 3 information must be de-identified before sharing the information with members of the research staff whenever the identifying information is not necessary. All 18 HIPAA-specified identifiers must be removed for de-identification.
  - The removal of identifying information from Level 3 information may change the classification of the data to Level 1.
- Level 3 information may be re-classified as Level 1 information if all of the following conditions are met:
  - All HIPAA-specified identifiers are removed;
  - It is reasonable to expect that individuals cannot be identified through deductive means (the advice of a biostatistician should be sought to ensure this requirement); and
  - There are no agreements, contracts, rules or laws that regulate the use, storage, transmission, handling or disclosure of the de-identified information.
- Data should be routinely backed up and the back-up copy physically secured.
- Devices must undergo secure deletion of the disc at the end of life of the device or prior to recycling.

## Appendix D

### Data Integrity for Human Research Data

The following are examples of measures that may be used in data security and integrity plans to ensure the integrity of the human research data.

- Data entry performed twice by two different individuals when transcription errors are possible.
- Edit checks (time-of-entry contextual and programmatic evaluation of entered data)
- Random, internal quality and assurance auditing by a person other than the individual who performed the original entry

If an institutionally-supported computer is being used to store human research data the principal investigator must ensure that backup copies of human research data are periodically created and stored in a safe and recoverable location. If the human research data is stored on an IT or PHS supported server, backups can be assumed.

- Backup copies if necessary should be maintained in a location that would not be affected if the primary location were destroyed by a catastrophic event.
- The frequency and storage location of backups should be commensurate with the value of the human research data.
- Backups for Level 3 information must be protected according to the requirements described for original Level 3 information.

## Appendix E

### 18 Identifiers as specified by the HIPAA Privacy Rule

The following is a list of elements considered to be identifiers according to HIPAA regulations (45 CFR 164 Security and Privacy regulations, 164.514 b(2)). These elements may be identifiers of the research participant or of the relatives, employers or household members of the participant.

1. Names
2. All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current publicly available data from the Bureau of the Census:
  - (1) The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and
  - (2) The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000.
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older
4. Telephone numbers
5. Fax numbers
6. Electronic mail addresses
7. Social security numbers
8. Medical record numbers
9. Health plan beneficiary numbers
10. Account numbers
11. Certificate/license numbers
12. Vehicle identifiers and serial numbers, including license plate numbers
13. Device identifiers and serial numbers
14. Web Universal Resource Locators (URLs)
15. Internet Protocol (IP) address numbers
16. Biometric identifiers, including finger and voice prints
17. Full face photographic images and any comparable images, and
18. Any other unique identifying number, characteristic, or code

Version Date: December 12, 2011

**Most recent changes:**

- December 12, 2011 - Revised definitions of Level 1, 2 and 3 data with regard to coded datasets and code lists.

**Revision History:**

- December 12, 2011
- Original 11/08/2011
-