**Learning objectives**

*By the conclusion of this workshop you will be able to:*

1. Understand the importance of data management in human subjects research
2. Understand the concept of the data lifecycle, and name and describe its stages
3. Identify the key ways in which issues of data management will affect research at the planning stage, during data collection, and after data collection is complete
4. Define and understand the concept of confidentiality in human subjects research
5. Incorporate information on confidentiality into an informed consent document

Evaluation questions for LO1 (Canvas selects 2)

1. A good data management plan
   1. Should be directly linked to the purpose and methods of the research project (correct answer)
   2. Should ensure that data is never shared outside of the research team
   3. Should incorporate aspects of both qualitative and quantitative data
   4. All of the above
2. Which of the following is NOT a reason for developing a data management plan?
   1. They are often required by the organizations and agencies that fund research
   2. Human subjects data can contain personal information that could be compromised by hackers
   3. A management plan will ensure that all data is destroyed by the end of the research project, as required by law (correct answer)
   4. They can help protect the intellectual property of the researchers
3. T/F: Federal law requires that all U.S. colleges and universities have identical policies regarding the management of human subjects data (False)
4. T/F: Data management plans are generally necessary only when requested by the research funder (False)
5. T/F: Responsible conduct of research requires adherence to both professional norms and ethnical principles in the conduct of research (True)

Evaluation questions for LO2 (Canvas selects 2)

1. Destruction of data occurs at what stage of the data lifecycle?
   1. Plan and design
   2. Analyze and collaborate
   3. Evaluate and archive (correct answer)
   4. Access and reuse
2. Creating metadata should occur at what stage of the data lifecycle
   1. Plan and design (correct answer)
   2. Collect and create
   3. Evaluate and archive
   4. Share and disseminate
3. Which of the following actions does not occur during the plan & design stage of the data lifecycle?
   1. Creation of metadata
   2. Data analysis (correct answer)
   3. Development of data use agreements
   4. Collaboration among the research team
4. T/F: Determining how you will analyze and disseminate data before that data is collected is considered an unethical research practice (False)
5. T/F: Being able to access and reuse data collected by other researchers is an important component of the research process (True)

Evaluation questions for LO3 (Canvas selects 3)

1. Metadata describes
   1. The multiple different storage and encryption techniques that can be used to manage research data
   2. The entire body of research data created over the lifecycle of a research project
   3. Descriptive data that allow a more complete understanding of a dataset (correct answer)
   4. Human subjects data collected at a macro level (ie, county- or state-level data) rather than at an individual level.
2. Which of the following is NOT a typical component of metadata documentation:
   1. A list of pertinent institutional and federal policies (correct answer)
   2. Description of the data structure
   3. Who has access to the data and why
   4. Description of how the data were collected
3. A regulatory binder should contain:
   1. All data collected by the project, in a secure format
   2. Training and other documents needed by the entire study team (correct answer)
   3. Quantitative research data, but never qualitative data
   4. A complete collection of all federal regulations related to human subjects data
4. A data dictionary is:
   1. A complete list of everyone who is allowed to access some or all of the data
   2. The original data that should not be modified or analyzed
   3. A list of all relevant regulations that pertain to your study data
   4. An explanation of the variables, measurement units, and/or data types in your study data (correct answer)
5. Data is considered low disclosure risk if:
   1. It does not include the names of the research subjects
   2. If disclosure of the data would not likely cause the research subjects any risk or harm (correct answer)
   3. It is stored on a password protected computer
   4. If the research subjects have signed an informed consent waiver
6. Which of the following means would be best for storing high risk data?
   1. A Box account owned by your college or university (correct answer)
   2. A password protected computer
   3. A folder on Google docs shared only within the study team
   4. None of the above are acceptable ways to store high risk data
7. T/F: All members of the study team should have equal access to study data during all stages of the research (False)
8. T/F: It is often appropriate for some data users to have access to only some of the available research data (True)
9. T/F: Making a copies of your data is rarely allowed because it compromises data security (False)

Evaluation questions for LO4 (Canvas selects 3)

1. In which of the following situations might a researcher be required to disclose confidential data?
   1. The researcher wants to publish research findings in an academic journal that requires that data be made publicly available.
   2. A research subject is a student at the same university where the researcher works, and the university requests access to that student’s data.
   3. The research uncovers that one or more research subjects has engaged in an illegal activity (correct answer).
   4. None of the above.
2. Which of the following is NOT considered a breach of confidentiality when dealing with human subjects data:
   1. A USB drive that contains encrypted participant information is lost or stolen (correct answer)
   2. A laptop that contains participant information is lost or stolen
   3. Placing a paper document including confidential data in a recycle bin
   4. Two members of the research team share research data with each other over email.
3. Which of the following is NOT a necessary confidentiality consideration in a data management plan?
   1. How often study data will be monitored.
   2. Whether the data collected will include information that is legally classified as Protected Health Information (PHI).
   3. Who will have access to the data after it has been completely deidentified.
   4. Whether or not the participants will be U.S. citizens, who are entitled to different confidentiality rights than non-citizens (correct answer).
4. If your study includes collecting background data (such as demographic data or medical history) from research participants, the best practice is to:
   1. Collect as much background data as possible, even if you are not sure if you will need it, so you won’t risk violating the privacy of the participants by re-contacting them later on.
   2. Never collect any specific geographic data such as neighborhood of residence because such data high risk if disclosed.
   3. Only collect background data in face to face surveys, never in online surveys.
   4. Never collect background data unless you have a specific plan to use it an answering your research question (correct answer)
5. T/F: Confidentiality requires that human subjects data never be shared beyond the original research team (False)
6. T/F: De-identifying study data is a means of protecting confidentiality (True)
7. T/F: Using a code number to replace a research subject’s name when storing your data is not an effective means of protecting confidentiality (False)
8. T/F: Researchers may not be able to promise confidentiality if it is likely that the research will uncover that research subjects have been victims of abuse (True)

Evaluation questions for LO5 (Canvas selects 2)

1. A consent form for research participants should include information about how the participant’s information will be:
   1. Stored, protected, shared, and used (correct answer)
   2. Stored, protected, and shared
   3. Protected, used, and de-identified
   4. Stored, de-identified, and destroyed
2. Dr. Smith wants to publish research findings in a journal that requires that study data be uploaded to a public data repository. When Dr. Smith collected her study data, she didn’t mention anything about using a data repository in her informed consent forms. What can Dr. Smith do?
   1. Nothing. The original consent form is binding, so the study data cannot be uploaded to a repository, and Dr. Smith cannot publish in that journal.
   2. As long as the data are properly de-identified, then Dr. Smith is free to put the data in a repository as requested by the journal.
   3. Dr. Smith should consider re-contacting the research participants to inform them of her plans for the data and get an updated consent form (correct answer).
   4. Dr. Smith should contact the agency that funded the original data collection, as they will have the final say over what can be done with the data.
3. T/F: Because an informed consent form is a legally binding document, it cannot be changed or updated if the research plans change. (False)
4. T/F: Informed consent is a process that is ongoing throughout the study, so a single signed consent form may not be sufficient to protect confidentiality (True)