# Data Do's and Don'ts
## Best Practices and Mistakes to Avoid

Fanny Ennever, PhD

Research Compliance Officer, BMC

CRRO Seminar October 9, 2019

# Learning Objectives

- Understand the importance of handling study data properly
- Understand requirements for data privacy and security
- Understand how to reduce risks of data confidentiality non-compliance
- Identify how to get answers about data best practices

# DO: Value Your Data

- Key to generating meaningful results
- Ethical obligation to study subjects
  - To protect
  - To use wisely

# DO: Plan Ahead

- What data you will collect, store, analyze, and transmit
- How you will protect the data

➢ Needed for INSPIR questions

# Special Subject Populations - NEW

| 12.0 | Subjects |
|---|---|
| **12.5 Special Populations** | |
| Please indicate if ANY (even one) of the following populations will be recruited (Note: Enrollment from any of these categories requires prior IRB approval):<br>☐ …<br>☐ Individuals whose HIV testing status is provided to the study team prior to consent being obtained (e.g., for recruitment)<br>☐ Individuals identified as a patient of a federally-assisted substance use disorder clinic (Project RESPECT, Office-Based Addiction Clinic, CATALYST Clinic, or others - see (?) Help Icon for full list) | |

# HIV Testing Status - NEW

12.7 Massachusetts law prevents disclosure of HIV testing status (both that the test has occurred and the results of the test) outside the covered entity without written informed consent from the patient. Therefore, HIV testing status cannot be shared with anyone on the study team who is not a member of the BMC workforce until consent for disclosure has been obtained. In the textbox, please add a statement that HIV testing status will not be disclosed to any non-BMC research staff member prior to written consent being obtained. If the study staff includes BU Medical Campus employees, please explain how you will operationalize the structure of your data access to prevent HIV testing status from being disclosed to the BU staff members prior to written consent being obtained.

# Patient of substance use disorder clinic - NEW

- If Yes, after you submit, the IRB will provide you with a form for you to get completed by the director of the substance use disorder clinic.

- You may pro-actively request the form prior to the submission by contacting the IRB (medirb@bu.edu; 617-358-5372

# Data Analysis

| **13.0 Design/Procedure** |
| --- |
| **13.3 Data Analysis** |
| Provide a description of your plan for data analysis. State the types of comparisons you plan (e.g. comparison of means, comparison of proportions, regressions, analysis of variance). Which is the PRIMARY comparison/analysis? How will the analyses proposed relate to the primary purposes of your study? If you are doing qualitative research please state how comparisons will be made. |

# Sample size

| **13.4 Sample Size/Specimens** |
| --- |
| **How many subjects (or records, or specimens, or charts) will be enrolled in this study?** Be sure to include all subjects who will be consented - even those who will be disqualified following consent because they did not meet the enrollment criteria. |
| **Subjects under BMC/BU Medical Center PI** |
| ☐ |
| **For multi-center studies only - Total worldwide subjects, including subjects under BMC/ BU Medical Campus PI** |
| ☐ |

# Sample size justification

| **13.4 Sample Size/Specimens** (continued) |
| --- |
| **Sample Size Justification**<br><br>Describe how you will have access to a population that will allow recruitment of the necessary number of subjects. Indicate why you chose the sample size proposed. Provide your sample size calculations. If this is a pilot study, this justification does not necessarily require a formal sample size calculation, but should provide a rationale for choosing the sample size proposed (e.g. to estimate a mean to a certain accuracy, to determine if the response rate is above a certain percentage, etc.) |

# Sample size justification - continued

**13.4 Sample Size/Specimens** (continued)

**Sample Size Justification** (continued)

Note: Once the IRB approves a certain study sample size then you may not enroll beyond that sample size without first obtaining approval from the IRB. Explain how many evaluable subjects you will need to end up with to answer your study question and how many subjects you will need to enroll and consent to achieve this number. The IRB counts study subjects starting when they are consented.

# Data & Safety Monitoring

**15.0 Data & Safety Monitoring**

**A data and safety monitoring plan (DSMP) is meant to assure that each clinical investigation has a system for oversight and monitoring of the conduct of the clinical investigation. This oversight is intended to ensure the safety of the participants and the validity and integrity of the data. A DSMP should be commensurate with the risks.**

# Data & Safety Monitoring - continued

**15.0 Data & Safety Monitoring** (continued)

**A DSMP can be as simple as the investigator reporting Unanticipated Problems, Adverse Events, and Protocol Deviations to the IRB. A DSMP can be as complex as having a Data and Safety Monitoring Board. A DSMP can include clinical trial monitoring.**

# Screening data

| 17.0 Screening Procedures |
| --- |
| **17.2** Describe what eligibility data will be stored and how it will be stored, who will have access, and when these data will be destroyed. For screening failures, please detail how and what data will be retained, if any, along with when these data will be destroyed. Please describe whether identifiers are being retained from those who screen out. Please note if contact information is being retained for future research. |
| |

# Privacy

| 19.0 Privacy and Confidentiality |
| --- |
| **19.1 Privacy** (Privacy refers to an individual's control over who has access to him/herself) |
| ☑ The following measures will be used to protect the privacy of subjects and potential subjects:<br>• The information that will be obtained from and/or about subjects and potential subjects is the minimum necessary to conduct the study; and<br>• If any interventions and interactions occur with subjects and potential subjects, they will take place in private settings. |

# Identifiability – option 1

| **19.2 Confidentiality of the Data** *[option 1]* |
| --- |
| ☐ Study data/results, documents, CRFs, and other documents/files will be identified with a unique study ID #. The study ID # will be linked to a master-code list that contains all study ID #s and direct subject identifiers (i.e. name, address, DOB, MRN, etc). The master-code list will be maintained separately from study files and access limited to the researchers. |

# Identifiability – option 1 - NEW

**19.2 Confidentiality of the Data** (continued)

You have indicated above that Study data/results, documents, CRFs, and other documents/files will be identified with a unique study ID #. Please select one of the options below:

- ○ Study data/results, CRFs, and other documents/files for subjects who have been assigned a study ID # may also contain subject identifiers that by themselves or when combined with other identifiers, could result in identifying a subject (ex. maintaining paper medical records that contains a subject's name and MRN in a participant's research file.)
- ○ Study data, documents, CRFs, and other documents/files for subjects who have been assigned a study ID # will NOT contain any subject identifiers that by themselves or when combined with others identifiers, could result in identifying a subject.

# Identifiability – option 2

| **19.2 Confidentiality of the Data** *[option 2]* |
| --- |
| ☐ All study data, documents, CRFs, and other documents/files will be recorded as anonymous. There is NO master-code. There will be no reasonable way to link study data and documents to individual subjects, even temporarily AND subject identities cannot be reasonable ascertained via deductive disclosure. |

# Identifiability – option 2 - continued

| **19.2 Confidentiality of the Data** *[option 2, continued]* |
| --- |
| You have checked off that all data are recorded anonymously, meaning that there is no way to link any of the study data to the individual participants. Please provide a brief study-specific description of how the data will be recorded anonymously. |

# Identifiability – option 3

| **19.2 Confidentiality of the Data** *[option 3]* |
| --- |
| ☐ There is an alternate plan for how subject will be identified in study data, documents, CRFs, and other documents/files. Please specify in text box below. |

# Secure storage of data – options 1 & 3

| **19.2 Confidentiality of the Data** (continued) |
| --- |
| • Please describe in the text box below how you will secure the data (e.g. how the master-code will be stored relative to the study data).<br>• If the dataset contains protected health information (PHI) or Personal Information (as defined under Massachusetts law) and is being stored electronically, please provide explicit confirmation that it will be stored according to BMC and/or BU policy for secure storage of such data. Please see the (?) Help Icon to the right for the definitions of PHI and Personal Information and for the appropriate storage options for BMC and BU and specify which will be used. |

# Secure storage of data – Help

**BMC Options for Storage of PHI:**

- BMC network drive
- BMC Box.com account Contact PrivacyOfficer@bmc.org

**BU Options for Storage of PHI (Restricted Use Data):**

- BU REDCap and MyCap app (Iphone or Android research subject app)
- BU FreezerPro
- BU Qualtrics
- BUMC Network Y Drive
- BU Office365 Services, including SharePoint, Teams, OneDrive
- If a HIPAA Limited Data Set only, all services above, plus BU Shared Computing Cluster (SCC) [Confidential Use Data]

# Data release

| **19.3 Release of identifiable data** |
| --- |
| Is identifiable data being released outside of BMC/BU Medical Campus? (e.g. to sponsors, because of mandated reporting, etc).<br>◉ Yes<br>○ No |

# Data release (continued)

| **19.3 Release of identifiable data** (continued) |
| --- |
| If identifiable data is being transmitted outside of BMC/BU Medical Campus, please describe what information will be released and under what circumstances, to whom the information will be released, and how confidentiality will be maintained during data transmission. (e.g. encrypted email, encrypted flash drive, sponsor-provided data capture system). |

# Third-party programs - NEW

**19.3 Release of identifiable data** (continued)

Do you plan to share data with a third-party vendor or software application or program? Note: sponsors are not considered third parties.

- ⦿ Yes
- ○ No

# Third-party programs - NEW (continued)

**19.3 Release of identifiable data** (continued)

For BMC research, reach out to **privacyofficer@bmc.org**
For BU research, reach out to **buinfosec@bu.edu**

They will need to review their privacy and security practices, and obtain a BMC and/or BU HIPAA Business Associate Agreement before HIPAA data is shared with the third party/vendor or application.

Please be sure to make this contact promptly because this process can require significant information gathering and analysis.

# Destruction of identifiers

| **19.4 Destruction of Identifiers** |
| --- |
| If the data are identifiable and/or if a master-code exists, when and how will the data be de-identified or the master-code be destroyed? |

# HIPAA – requested data

| 20.0 HIPAA Compliance |
| --- |
| **20.3 Selection criteria** (e.g. all Type 2 diabetics prescribed metformin, all men aged 50-75 with diagnosis of BPH) |
| |
| **20.4 Date range** (e.g. 11/14/98-12/1/13) |
| |
| **20.5 Data variables** (list or attach file) NOTE: If you are using the CDW to provide some or all of the data, the variables you list here will be utilized as your official data request by the CDW |
| |

# HIPAA – data protection

| 20.0 HIPAA Compliance |
| --- |
| **20.10 What is your plan to protect any identifiable information from use and disclosure by unauthorized parties?** |
| |
| **20.11 When and how will you destroy any identifiers linked to the data?** |
| |

# HIPAA – affirmations

| 20.0 HIPAA Compliance |
| --- |
| **20.10 Please affirm the items below:** |
| ☐ I agree that the protected health information will not be re-used or disclosed to any other person or entity, except as required by law, for the authorized oversight of the research study, or for other research for which the use or disclosure of protected health information would be permitted by the Privacy Regulation (45 CFR 164.512)<br>☐ I declare that the requested information constitutes the minimum necessary data to accomplish the goals of the research. |

# Repositories

| 27.0 Retention of Samples or Data |
| --- |
| **27.5 Provide the specific data points that will be retained (or attach it as a separate document). If samples will be retained, specify which data elements will be attached to the samples.** |
| |

# Repositories - continued

| **27.9 Confidentiality** |
| --- |
| Indicate the protections that will be put into place to protect subjects' privacy and confidentiality of their data<br>○ All samples/data will be de-identified at the time they are retained<br>○ All subject identifiers will remain at BMC/BU Medical Campus. Only coded samples/data will be given to other investigators/entities - BMC/BU Medical Campus investigators will maintain the master-code/key and will never release it to outsiders<br>○ Subject identifiers will be released to outside entities (e.g. NCI repositories) with subjects specific consent to do so<br>○ Access to subject identifiers will be limited to certain people (specify below)<br>○ Other (specify below) |

# DO: Follow the plan

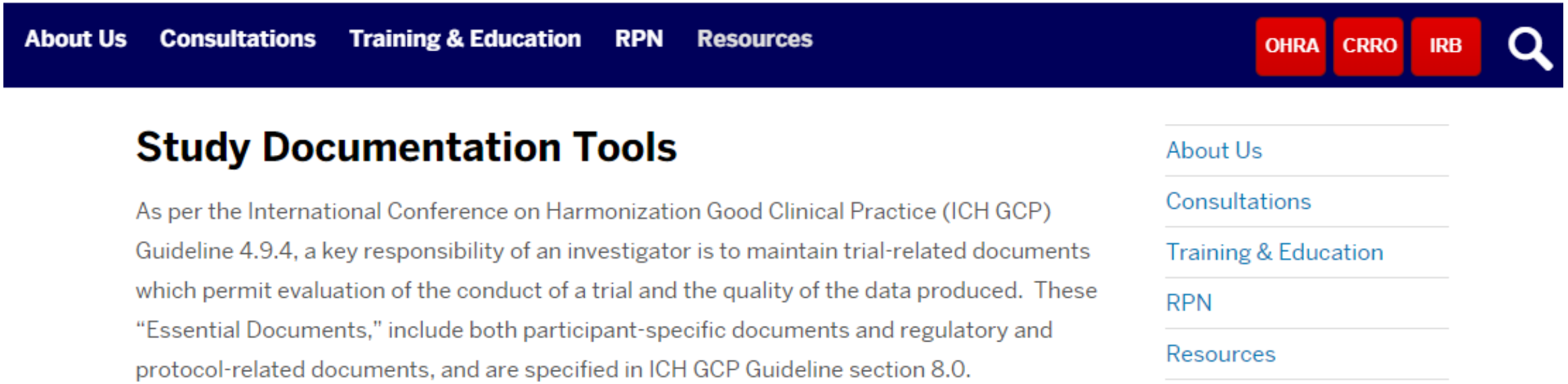
- Be aware of how data handling was described
- Encryption is your friend
  - Make sure all devices with data are encrypted
  - If transmitting password-protected files (digitally or physically), send password separately
- Control who has access
- Safeguard paper copies
- Beware of phishing

# DO: Ensure data quality

- Remember ALCOA: documentation should be
  - **A**ttributable
  - **L**egible
  - **C**ontemporaneous
  - **O**riginal
  - **A**ccurate

# DO: Use CRRO resources

http://www.bumc.bu.edu/crro/tools/



# DO: Know where to get help

Mary-Tara Roth, CRRO
mtroth@bu.edu
617-358-7679

David Corbett, Security, BU
corbettd@bu.edu
617-358-7879

Linda Rosen, Clinical Data Warehouse
Linda.Rosen@bmc.org
617-358-5337

Fanny Ennever, Research Compliance, BMC
Fanny.Ennever@bmc.org
617-638-8874

Diane Lindquist, Privacy, BU
dlindq@bu.edu
617-358-3124

Seraphina Rasmussen, Privacy, BMC
PrivacyOfficer@bmc.org
617-638-7958

# Questions???