

Clinical Data Management

*Slides from: Elizabeth Ness, RN, MS
Nurse Consultant (Education)
Center for Cancer, NCI*

PII slides from NIH HIPAA training



National Institutes
of Health

Agenda

- Overview of Clinical Data Management (CDM)
- Case Report Form (CRF) development
- Clinical Trials Management Systems (CTMS)
- Managing discrepancies
- Quality Control (QC) and Quality Assurance (QA) activities

Clinical Data Management

- A multi-disciplinary activity that includes:
 - Research nurses
 - Clinical data managers
 - Investigators
 - Support personnel
 - Biostatisticians
 - Database programmers
- Various activities involving the handling of information outlined in protocol

Clinical Data Management Activities

- Data acquisition/collection
- Data abstraction/extraction
- Data processing/coding
- Data analysis
- Data transmission
- Data storage/security
- Data privacy
- Data QA

Good Clinical Practice (GCP) Guidelines

- Trial management; data handling, record keeping (2.10, 5.5.3 a-d)
- Subject and data confidentiality (2.11; 5.5.3 g)
- Safety reporting (4.11)
- Quality control (4.9.1; 4.9.3; 5.1.3)
- Records and reporting (5.21; 5.22)
- Monitoring (5.5.4)

21 CFR Part 11

- Applies to all data (residing at the institutional site and the sponsor's site) created in an electronic record that will be submitted to the FDA
- Scope includes:
 - validation of databases
 - audit trail for corrections in database
 - accounting for legacy systems/databases
 - copies of records
 - record retention

Data Management Plans (DMP)

- Living paper or electronic records to document processes and procedures of how data will be handled
- Promotes consistent, efficient and effective data management practices at a study level

Components of a DMP

- Roles and responsibilities of all research team members who will handle data
- Description of the data that will be collected, data dictionaries or form annotations
- List of standards or terminologies that are used along with their version(s)
- How the data is acquired, processed and stored
- Where data will be stored
- Data handling rules
- Data sharing or access policy and processes

Case Report Form (CRF)

Patient X1 Page 13 (Pharmacoki for Course 1) Page 1 of 1.

Visit Date: 12-Dec-2006 Blank Comment: _____

Phkm1 Phkm2 COM

Type: PK1 Blank

PHARMACOKINETICS - 1

Course # _____ Day in Course _____

Study Agent _____ Specimen Sampled _____

Start Date _____ Start Time _____ Stop Date _____ Stop Time _____

Sample ID #	Planned Interval	Sample Collected?	Actual Start Date	Time Interval (min)	Parent Study Agent
	PRE-DOSE				
	0				
	0.5				

PharmaForm - E-Form

File

New Patient Screening

Protocol: MP_914_98_5

Subject ID: talbert-1

Visit ID: Screening

Date of Birth: Month 12 Day 31 Year 98

Sex: Female Male

Ethnicity: Caucasian Hispanic
 African American
 Other _____

Height (in cms): 100

Weight (in kgs): 100

Body surface area: 40.39

Ideal Body Weight: 6.551

Save Submit to Exit

CONCOMITANT MEASURES/MEDICATION NCI/DCTD/CTMS CASE REPORT FORM (Include all supportive measures instituted while on study)

Date Completed: (dy/mth/yr)	Protocol #:	Institution:	Sheet #:	Patient ID:
Start Date (dy/mth/yr)	Agent Or Procedure	Total Daily Dose	Schedule	
Stop Date (dy/mth/yr)		Units	Reason	
1.				
2.				
3.				

What is a CRF?...

- Most frequently used data collection tool in clinical trials
- Paper based or electronic
- Include data entry forms used by:
 - Patients
 - Health care providers
- GCP Guidelines:
 - “A printed, optical, or electronic document designed to record all of the protocol-required information to be reported to the sponsor on each trial subject.”

...What is a CRF?

- Allows for collection of study data in a standardized format:
 - According to the protocol
 - Complying with regulatory requirements
 - Allowing for efficient analysis
- Facilitates exchange of data across projects and organizations
- Accompanied by a completion/instruction manual

Protocol and CRF Relationship

- Protocol determines what data *should* be collected
- All data *must* be collected if specified in the protocol
- Data that will not be analyzed or used for regulatory compliance *should not* be captured on the CRF

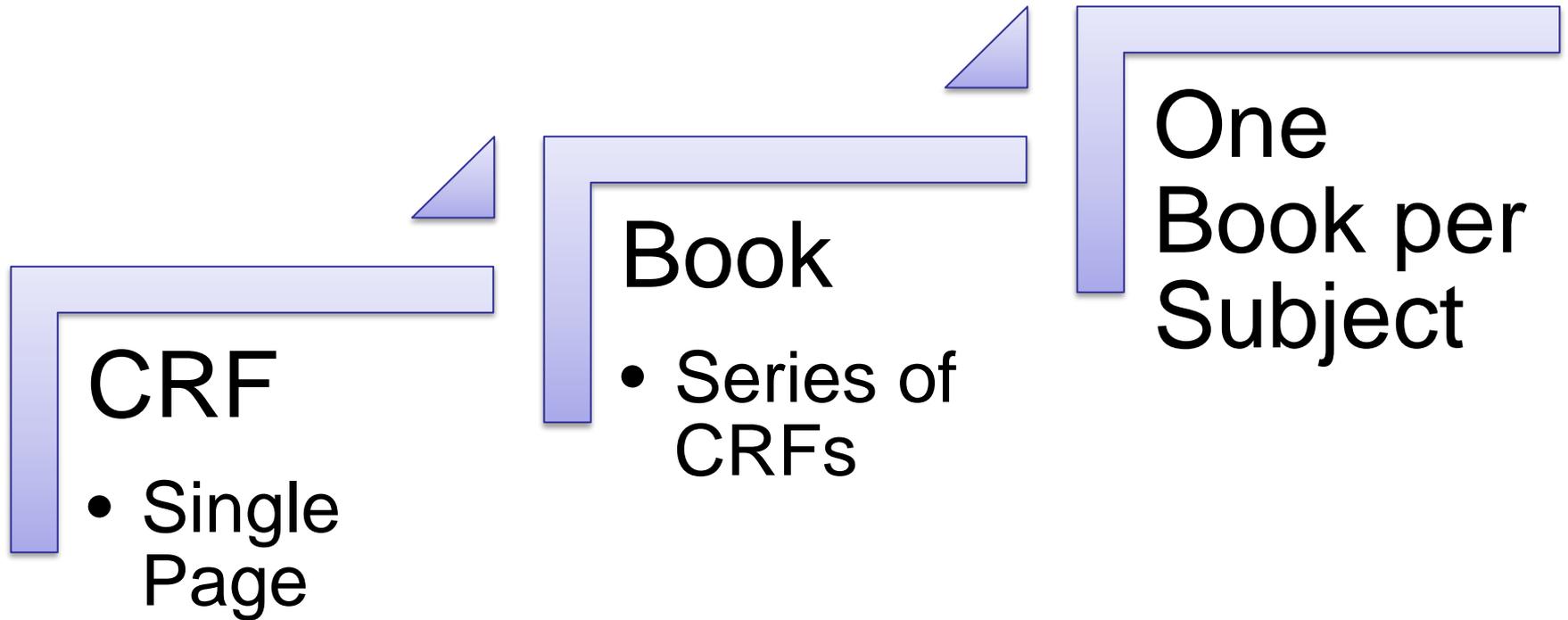
CRF Development Process

- Begins as soon in the protocol development process as possible
 - Work with protocol visit schedule
- All appropriate individuals need to be part of the process
- Interdisciplinary review is necessary

General Considerations for CRF Development

- Collect data outlined in the protocol
- Collect data required by the regulatory agencies
- Collect data with all users in mind
- Be clear and concise with data questions
- Avoid duplication
- Request minimal free text responses
- Collect data that allows for efficient computerization
- Develop version control procedures

CRF Terminology



Standard CRFs

- Allows rapid data exchange
- Removes the need for mapping during data exchange
- Allows for consistent reporting across protocols, across projects
- Promotes monitoring and investigator staff efficiency
- Allows merging of data between studies
- Provides increased efficiency in processing and analysis of clinical data

Common Data Element (CDE)

- Set of descriptors for a variable
- Standardize the way in which data elements or questions can be asked, collected, stored, exchanged and reported
- Need to allow for version control
- Metadata can be easier to revise and reuse
- Cancer Data Standards Registry and Repository (caDSR) maintained by the NCI
 - Provides a data element inventory at both the individual element and the CRF level

CDE Browser



Search Results [Search within results](#)

Results fewer than expected? Check Search Preferences

[\[Download Data Elements to Prior Excel\]](#)
[\[Download Data Elements to Excel\]](#)
[\[Download Data Elements as XML\]](#)

[\[Download CDE Browser DTDs\]](#)

Sort order: (Default) Registration Status>>Workflow Status>>Long Name [Ascending]

[Add to CDE Cart](#)
[Add to CDE compare list](#)
[Compare CDEs](#)

1 - 30 of 30

<input type="checkbox"/>	Long Name	Preferred Question Text	Owned By	Used By Context	Registration Status	Workflow Status	Public ID	Version
<input type="checkbox"/>	Diastolic Blood Pressure Value	Diastolic BP	CCR	ABTC,AECC,CITN,COG,CTEP,DCP,ECOG-ACRIN,LCC,NICHD,NINDS,SPOREs,caBIG	Standard	RELEASED	2004291	1.0
<input type="checkbox"/>	Measurement Vital Signs Occurrence Date	Date of Measurement	caBIG	ABTC,AECC,CITN,COG,CTEP,LCC,NINDS	Standard	RELEASED	2829808	1.0
<input type="checkbox"/>	Measurement Vital Signs Occurrence Time	Time of Measurement	caBIG	CITN,LCC,NINDS	Standard	RELEASED	2829812	1.0
<input type="checkbox"/>	Person Height Unit of Measure Unified Code for Units of Measure Code		caBIG	ABTC,AECC,CCR,CITN,CTEP,DCP,NHC-NCI,NHLBI,NINDS,OHSU Knight,PBTC	Standard	RELEASED	2538920	1.0
<input type="checkbox"/>	Person Temperature Unified Code for Units of Measure Code	Temperature Unit of Measure UCUM Code	caBIG	ABTC,AECC,CTEP,DCP,LCC,NINDS	Standard	RELEASED	2956267	1.0
<input type="checkbox"/>	Person Vital Signs Respiratory Rate Physical Examination Value	Respiratory Rate	CTEP	ABTC,AECC,CCR,CITN,DCP,LCC,NICHD,NINDS,SPOREs,caBIG	Standard	RELEASED	2644399	1.0
<input type="checkbox"/>	Person Vital Signs Temperature Physical Examination Value	Body Temperature	CTEP	ABTC,AECC,CITN,DCP,ECOG-ACRIN,LCC,NICHD,NINDS,caBIG	Standard	RELEASED	2644401	1.0
<input type="checkbox"/>	Person Weight Unit of Measure Unified Code for Units of Measure Code	Weight unit of measure	caBIG	ABTC,AECC,CCR,CITN,CTEP,DCP,NHC-NCI,NHLBI,NICHD,NINDS,OHSU Knight,PBTC,SDC Pilot Project	Standard	RELEASED	2630200	1.0
<input type="checkbox"/>	Person Weight Value	Weight	DCP	ABTC,AECC,CCR,CITN,CTEP,ECOG-ACRIN,LCC,NHC-NCI,NIDA,NINDS,NRG,OHSU Knight,PBTC,SDC Pilot Project,Theradex,caBIG	Standard	RELEASED	2179689	4.0
<input type="checkbox"/>	Pulse Rate Measurement Value	Pulse	CCR	ABTC,AECC,CITN,CTEP,DCP,LCC,NICHD,NINDS,SPOREs,caBIG	Standard	RELEASED	2767073	1.0
<input type="checkbox"/>	Systolic Blood Pressure Value	Systolic BP	CCR	ABTC,AECC,CITN,COG,CTEP,DCP,ECOG-ACRIN,LCC,NICHD,NINDS,SPOREs,caBIG	Standard	RELEASED	2004289	1.0
<input type="checkbox"/>	Diagnostic Electrocardiogram Performed Date	Date of EKG	CTEP	ABTC,DCP,NINDS,NRG,PBTC,SWOG,caBIG	Qualified	RELEASED	2004069	3.0
<input type="checkbox"/>	Oxygen Saturation Percentage Level Number	Room air SaO2	DCP	NINDS	Qualified	RELEASED	2482251	1.0
<input type="checkbox"/>	Patient Height Measurement	Height	CTEP	Alliance,CTEP,NINDS,NRG,SWOG,caBIG	Qualified	RELEASED	649	4.1
<input type="checkbox"/>	Biospecimen Aliquot Number	Number of aliquots	NINDS			RELEASED	3192457	1.0
<input type="checkbox"/>	Biospecimen Centrifugation Indicator	Sample centrifuged?	NINDS	NRG		RELEASED	3192499	1.0
<input type="checkbox"/>	Biospecimen Clotted Procedure Indicator	Clotting procedure used?	NINDS			RELEASED	3192555	1.0
<input type="checkbox"/>	Biospecimen Collection Biospecimen Type Other Text	Type of sample other specify	NINDS			RELEASED	3192562	1.0
<input type="checkbox"/>	Biospecimen Collection Date	Date of sample collection	NINDS			RELEASED	3192128	1.0

Refresh tree

- caDSR Contexts
 - ABTC (Adult Brain Tumor Consortium)
 - AECC (Albert Einstein Cancer Center)
 - Alliance (Alliance)
 - BOLD (Breast Oncology Local Disease)
 - BRIDG (BRIDG Collaboration)
 - caBIG (NCI cancer Biomedical Informatics Grid)
 - caCORE (NCI Core Infrastructure)
 - CCR (NCI Center for Cancer Research)
 - CDC/PHIN (Centers for Disease Prevention and Control - Public Health N)
 - CDISC (Clinical Data Interchange Standards Consortium)
 - CIP (NCI Cancer Imaging Program)
 - CITN (Cancer Immunotherapy Trials Network)
 - COG (Children's Oncology Group)
 - CTEP (NCI Cancer Therapy Evaluation Program)
 - DCI (Duke Cancer Institute)
 - DCP (NCI Division of Cancer Prevention)
 - ECOG-ACRIN (ECOG-ACRIN)
 - EDRN (NCI Early Detection Research Program)
 - LCC (Lombardi Cancer Center)
 - NCIP CDE Data Standards (Shortcut)
 - NHC-NCI (Norton Cancer Institute)
 - NHLBI (National Heart, Lung and Blood Institute)
 - NICHD (National Institute of Child Health and Development)
 - NIDA (National Institute on Drug Abuse)
 - NIDCR (National Institute of Dental and Craniofacial Research)
 - NINDS (National Institute of Neurological Disorders and Stroke)
 - NRG (NRG Oncology Group)
 - OHSU Knight (Oregon Health & Science University Knight Cancer Institut
 - PBTC (Pediatric Brain Tumor Consortium)
 - PS&CC (NCI Population Sciences & Cancer Control)
 - SDC Pilot Project (SDC Pilot Project)

<https://cdebrowser.nci.nih.gov/CDEBrowser/>

Clinical Data Acquisition Standards Harmonization (CDASH)

- Standardize data collection fields intended to be used on CRFs
 - Divided into sixteen domains
 - Applicable to all clinical studies regardless of therapeutic area or phase of development
- Released in October 2008 by the Clinical Data Interchange Standards Consortium (CDISC)
- Link to site:
 - <http://www.cdisc.org/extranet/index.php?a=2531>

Elements of a CRF

Header

- Key identifying Information
- Study Number
- Site/Center Number
- Subject identification number

Safety Modules

- Demographic information
- Adverse Events
- Medical History/Cancer history (e.g., diagnosis, staging)
- Physical Exam, including Vital Signs
- Concomitant/Concurrent Medications/Measures
- Deaths
- Drop outs/off-study reasons
- Eligibility confirmation

Efficacy Modules

- Unique modules
- Can be more difficult to develop
- Protocol dictates the elements
- Repeated/battered of tests
- Define
 - Key efficacy endpoints of trial (primary and secondary)
 - Additional test to measure efficacy (e.g.: QOL)
 - Required diagnostics
- Includes baseline measurements

Poorly Designed CRF

- Poorly designed CRFs will result in data deficiencies including:
 - Data not collected as per protocol
 - Collecting unnecessary data (i.e.: data not required to be collected per protocol)
 - Impeding data entry process
 - Database requiring modifications throughout study

Electronic CRFs



- Use of remote data capture (RDC) is increasing
- Generally, concepts for the design of eCRFs/RDC screens are the same as for paper
- Advantages:
 - Faster data collection
 - Cleaner data collection due to system built “checks”
 - Easier monitoring
 - Central database for storage of all trial data
 - Near real-time data access to authorized personnel

CDMS v. CTMS

Clinical Data Management System (CDMS)

- Associated with traditional approach to clinical trial activities
- Heavy data management component
 - Collection of data on paper CRFs that is then transcribed onto electronic CRFs
 - Data is cleaned and transported to various reporting and care-associated destinations

Clinical Trial Management System (CTMS)

- Integrate data from many systems (e.g., labs, genomics, and adverse events), enter and clean the data in expedited steps, and store it in a repository that can serve multiple purposes over time
 - Key concept of 'services' or capabilities such as the ability to handle patient randomization at the same time as adverse event reporting and creating laboratory report alerts for the clinician.

NIH Systems and Databases

- C3D: Cancer Central Clinical Database
- CRIMSON
- Medidata RAVE
- BTRIS: Biomedical Translational Research Information System
 - Integrates clinical research data from the CC and ICs
 - Provides Investigators with access to identifiable data for the subjects on their own active protocols
 - Provides NIH Investigators with access to de-identified data across all protocols

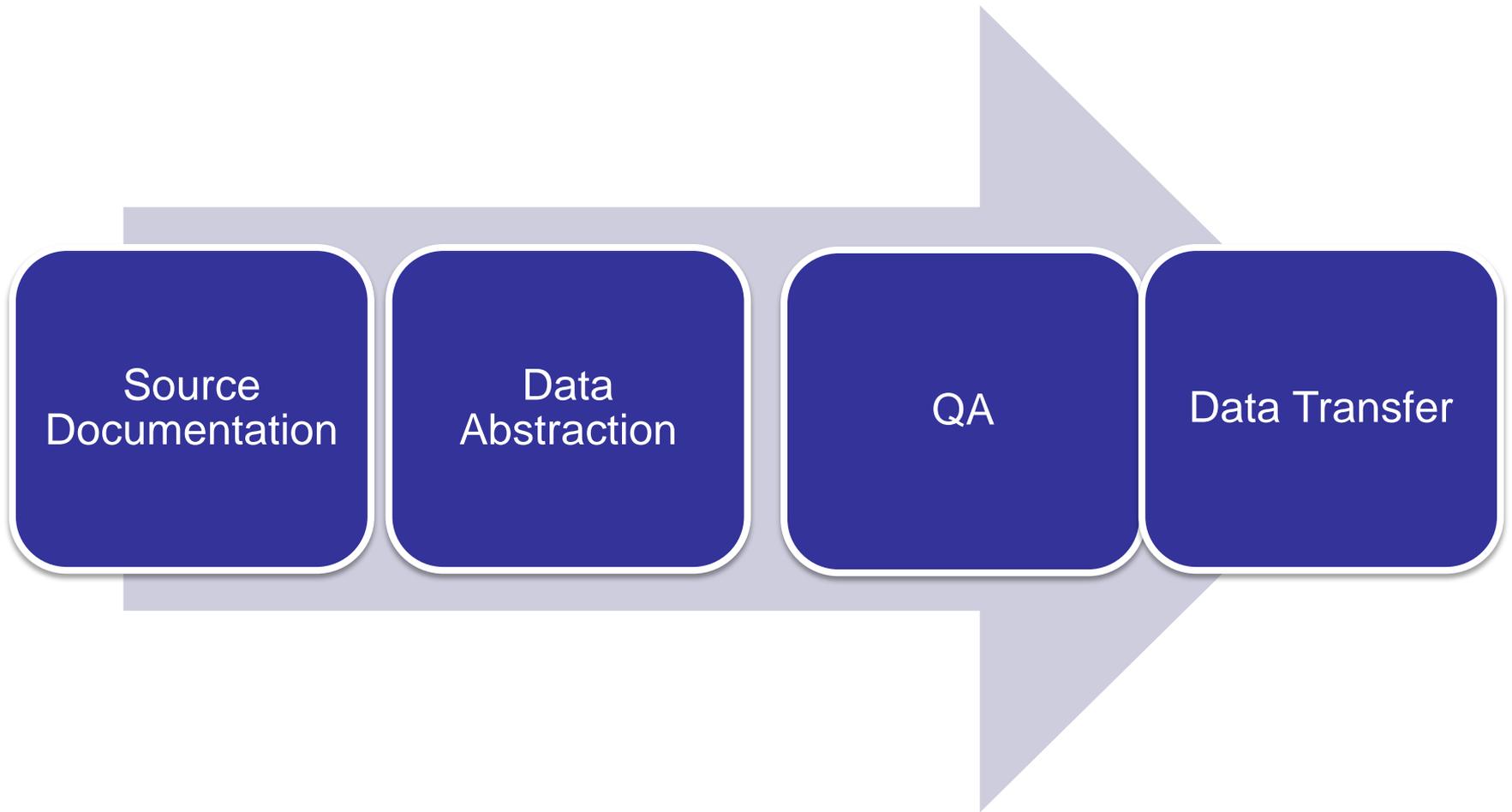
CRF Completion

- Per GCP Guidelines (4.9.1), investigator should ensure the accuracy, completeness, legibility, and timeliness of the data reported on the CRFs including:
 - all sections have been completed
 - all alterations have been properly made
 - all adverse events are fully recorded and that for all serious adverse events, any specific documentation has been completed

Timeliness of CRF Completion

- Ideally CRFs should be completed as soon after the subject's visit as possible
- Ensures that information can be retrieved or followed-up on while the visit is still fresh in the healthcare provider's mind, and while the subject and/or the information is still easily accessible

CRF Completion Process



Common Errors ...

- Logical
 - date of the second visit is earlier than the first visit
- Inaccurate information
 - source document says one thing, the CRF says another
- Omissions
 - AE is recorded on the CRF but not on the source document
- Transcription errors
 - date errors, 11-2-59 instead of 2-11-59

...Common Errors

- Abbreviations
 - unless an approved list of abbreviations is distributed and utilized, data entry personnel often misinterpret abbreviations
- Spelling errors
- Illegible entries/"write-overs"
- Writing in margins

Making Corrections

- Paper
 - Draw one horizontal line through the error;
 - Insert the correct data;
 - Initial and date the change;
 - **DO NOT ERASE, SCRIBBLE OUT, OR USE CORRECTION FLUID OR ANY OTHER MEANS WHICH COULD OBSCURE THE ORIGINAL ENTRY**
- Electronic
 - Make correction and the system does the rest
 - May need to include justification for change

Tips ...

1. Use the CRF completion/instruction manual
2. Make sure appropriate protocol, investigator and subject identifying information is included in the Header
 - Pre-populated for RDC
3. Ensure data is entered in the correct location or data field
4. Use the appropriate units of measurement (UOM), and be consistent
5. Use only the abbreviations authorized per manual

...Tips

4. Double check spelling
5. Watch for transcription errors
 - E.g.: sodium level should be “135” and entered as “153”
6. Use “comments” section to elaborate on any information, ***but keep to a minimum***
7. Perform quality control and logic checks
 - Is data consistent across data fields and across CRFs
 - Make sure visit dates match dates on the laboratory or other procedure reports
 - Make sure the birth date matches the subject’s age
8. Transfer data per sponsor

REMINDERS

- Data cannot be entered onto a CRF if it is not in the medical record or for some documents, in the research record
- If the individual completing the CRF, finds missing or discrepant source data he/she should:
 - Notify the health care provider who then will provide the data
 - If applicable, contact outside source (i.e.: outside lab or doctor's office)

PII v. PHI

Personally identifiable information (PII)

- Full name
- Mailing and Home Address
- Email address
- SSN/MRN
- IP address (in some cases)
- Vehicle registration plate number
- Driver's license number
- Face, fingerprints, or handwriting
- Credit card numbers
- Digital identity
- Date of birth
- Birthplace
- Genetic information
- Telephone number
- Login name, screen name, nickname

Protected Health Information (PHI)/(ePHI)

- Individually identifiable health information.
- Can be linked to a particular person.
- Specifically, this information can relate to:
 - The individual's past, present or future physical or mental health or condition,
 - The provision of health care to the individual, or,
 - The past, present, or future payment for the provision of health care to the individual.
- Common identifiers of health information include: names, social security numbers, addresses, and birth dates.

Data Security & Confidentiality

- **Do not:**
 - Include unnecessary protected health information (PHI)
 - Store patient information or PII on your personal computer/device.
- **Only:**
 - Store patient information or PII on government owned computers/devices that have a compliant encryption solution per HHS Cybersecurity Program Standard for Encryption
 - Use NIH-supported shared and personal NIH network drives. (These drives are secure and backed up nightly.)
- E-mailing of data files should be limited and done thru a secure system and encrypted if using PHI

Process of Data Transfer to Sponsor

**Traditional (Paper)
Electronic**

Database Example

CRF Development Activity

Quality Control

GCP Guidelines 5.1.3

QC should be applied to each stage of data handling to ensure that all data:

- Are reliable
- Have been processed correctly

Sponsor Queries

- Sponsor generates:
 - During/End of a monitoring visit
 - After data sent to sponsor and reviewed/entered in database
- Site corrects CRF:
 - During/between monitoring visit
 - May need to sign-off on query

Database Discrepancies

- Failure of entered data to pass a validation check as applied by a database
- Univariate discrepancy – single data element errors (e.g., not using provided pick-list, missing data in a field)
- Multivariate discrepancy – multiple data element errors (e.g., male patient with + beta HCG)

Assessing the QC/QA Process

- Are staff checking their own work?
- Are staff relying on others to check their work?
- Does the organization have a QA plan for monitoring protocol adherence and data collection?
- Are there SOPs related to data management?
- How soon after a visit is a CRF completed?
- Is all data, as defined in the protocol, captured from the source document to the CRF?

Terminology

- Quality Control
- Quality Assurance
- Quality Improvement



Quality Control (QC)

- Ongoing and concurrent review of subject data
 - Typically 100%
 - Checking your own work and work of others
- Verify that data collected and abstracted:
 - Correctly entered onto CRF
 - Able to be found in source document
 - Follows regulations and guidelines
- Individual team member level



Quality Assurance (QA)

- Planned, systematic check done at the branch or organizational level
- Verifies:
 - Trial is performed as per the approved plan
 - Data generated is accurate
- Identifies problems and trends:
 - Retrospective and involves sampling of subjects and data
 - Pulls all the pieces together to gain a picture (measurement) of compliance
 - Ensures staff is compliant with internal and external regulations/guidelines

QA Activities

- Internal monitoring/audits
 - Compile all data components and gain a measurement of compliance
- Clarification monitoring
 - Assess for trends
 - Review clarifications responses before they are submitted to sponsor
- Measure data inconsistencies and trends using a sampling of the data prior to audits/monitoring visits
- Summarize QA findings and report to management
- Identify learning needs

Quality Improvement (QI)

- Result of QC and QA
- Developing a plan includes:
 - Identifying root causes of problems
 - Intervening to reduce or eliminate these problems
 - Taking steps to correct the process(es)
 - Identifying trends and areas for improvement
 - Identifying solutions:
 - Assess work flow and time management activities
 - Develop tools for source documentation
 - Assess training needs
 - Involve appropriate staff in resolution
 - Implementing new/updated solution

Responsibilities

- Research Team responsibilities
- Research Nurse responsibilities
- Data Manager responsibilities

Research Team

- PI ultimately accountable
- Team should:
 - Ensure that all source data is documented in the Medical Record/Research Chart with accuracy, completeness, and consistency
 - Ensure the overall quality of the research data is verifiable and acceptable for sponsor submissions, publications, etc.
 - Review data discrepancy/clarification resolutions for accuracy, consistency and timely response

Research Nurse

- Provide accurate and complete source documentation
- Develop, implement, and maintain a team QC plan:
 - Establish a schedule of QC activities
 - Quality check source documentation, data abstraction, CRFs completion
 - Quality check of database
 - Verify function in database
- Review query/clarification
- Assist in developing team quality improvement plan

Data Management (er)

- Abstract data per CRF Instruction Manual
- Abstract data in a timely fashion, this includes entry into database
- Code accurately per protocol
- Apply quality control checks at each stage of data handling
 - Contact Research Nurse/PI for missing source data
- Resolve discrepant data – ongoing
- Utilize database report tools to assist with QC activities

Guiding Principles

- Source documents need to be accurate and complete
- Data abstraction should occur in real time
- QC/QI is the responsibility of every research team member
- QC/QI should be completed on all protocol data for all protocols
- QC/QI should be proactive and ongoing
- Each team member should know and understand the roles and responsibility of each team member

Data Management in Clinical Research

- 6 week webcourse developed by Vanderbilt University
- Available free of charge via [Coursera](#)
- Teaching strategies:
 - Online video presentations
 - Week 1: course information and data collection strategies (80 minutes)
 - Week 2: standardizing processes and fundamentals of electronic data capture (EDC) (70 minutes)
 - Week 3: planning a prospective study and setting up EDC (127 minutes)
 - Week 4: mid-study activities, data quality and monitoring, data sharing, de-identifying data, regulatory compliance (120 minutes)
 - Week 5: data collection using surveys, data management for multi-site studies, resource limited settings and global health, using data standards, course wrap-up (260 minutes)
 - Short weekly quizzes
 - Hands-on case report form and survey development

Questions???