The 2nd Clinical Data Management Training

Clinical Data Review

September, 2010 at SMMU, Shanghai
Purpose of Data Review

Data and Error Source

Types of Data Review

Perform Data Review

Key Messages
ICH GCP requires:

- **2.10** All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation and verification.

- **5.1.1** The sponsor is responsible …… and data are generated, documented (recorded), and reported in compliance with the protocol, GCP, and the applicable regulatory requirement(s).

- **5.1.3** Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly.
1. Purpose of Data Review (Individual Trial)

The ultimate goal is to provide complete, accurate and integrate data that serves as the base to draw correct conclusions.

- No matter how much care is taken into collecting and entering data, discrepancies and data errors will invariably find their way into a clinical database.
- The vast majority of these data inconsistency and errors can be alleviated with careful data review and data cleaning.
Agenda (Why, What, Where, When, How)

1. Purpose of Data Review
2. Data and Error Source
3. Types of Data Review
4. Perform Data Review
5. Key Messages
2.1 Data Source – What and Where

Data

- CRFs
- Lab
- ePRO
- Safety
- Admin

CRF – Case Report Form
Admin – Administration Data
ePRO – Electronic Patient Report Outcome
2.2.1 Error Type - Fraud

- Fraud
  - Sponsor
  - Investigator
  - Patient
2.2.2 Error Type – Systemic Error

- **Systemic error**
  - **Protocol deviation**
    - Oral vs. Axillary temperature
    - Fasting vs. non-fasting blood sample
    - Adverse event reporting
  - **Poor CRF design**
    - Ambiguous questions
    - Double negative questions
    - Limited options for close question
2.2.2 Error Type – Systemic Error

- **Systemic error**
  - **Measurement**
    - Un-calibrated sphygmomanometer
    - Different lab test method and reagent
  - **Evaluator/analyst**
    - Lack of experience
    - Un-qualified
  - **Data transfer between systems**
    - Missing record
    - Duplicate record
2.2.3 Error Type – Non-systemic Error

- Non-Systemic error
  - Transcription error
  - Other random errors
Agenda (Why, What, Where, When, How)

1. Purpose of Data Review
2. Data and Error Source
3. Types of Data Review
4. Perform Data Review
5. Key Messages
3.1 Data Review Type_Subject Set

- **Aggregate grouped data** across a study population (e.g. tabular summaries, bar charts, scatter plots)

- **Line listings of data** for a group of subjects (e.g. a list of all the AEs for a single subject, or a list of all subjects with a specific AE term)

- **Clinical data over time of individual subject** (e.g. all medically relevant data for a subject displayed time-aligned)
## Example 1

<table>
<thead>
<tr>
<th>Visit</th>
<th>Summary</th>
<th>Site 1 (N = 45)</th>
<th>Site 2 (N = 40)</th>
<th>Site 3 (N = 42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>N</td>
<td>45</td>
<td>40</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>36.2</td>
<td>35.5</td>
<td>36.8</td>
</tr>
<tr>
<td></td>
<td>Min - Max</td>
<td>35.3 - 37.6</td>
<td>34.8 - 37</td>
<td>35.4 – 38.6</td>
</tr>
<tr>
<td>Visit 1</td>
<td>N</td>
<td>45</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>36.5</td>
<td>35.3</td>
<td>36.4</td>
</tr>
<tr>
<td></td>
<td>Min - Max</td>
<td>35.5 - 37</td>
<td>35 - 36.8</td>
<td>36 - 37.2</td>
</tr>
</tbody>
</table>
### Listing of Adverse Events by Subject

<table>
<thead>
<tr>
<th>Site</th>
<th>Subject</th>
<th>Adverse Event</th>
<th>Date of Onset</th>
<th>Intensity</th>
<th>Outcome</th>
<th>Date Resolved</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>1001</td>
<td>FATIGUE</td>
<td>14-Apr-08</td>
<td>MODERATE</td>
<td>Resolved</td>
<td>19-Apr-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANEMIA</td>
<td>09-May-08</td>
<td>MODERATE</td>
<td>Resolved</td>
<td>13-Jun-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANOREXIA</td>
<td>10-Mar-08</td>
<td>MODERATE</td>
<td>Resolved</td>
<td>22-Mar-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CONFUSION</td>
<td>11-Sep-08</td>
<td>MODERATE</td>
<td>Resolved</td>
<td>16-Sep-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANEMIA</td>
<td>28-May-08</td>
<td>MILD</td>
<td>Resolved</td>
<td>19-Aug-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DRY MOUTH</td>
<td>19-Aug-08</td>
<td>MILD</td>
<td>Resolved</td>
<td>05-Feb-09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DRY SKIN</td>
<td>07-May-09</td>
<td>MODERATE</td>
<td>Resolved</td>
<td>21-Dec-09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LOST APPETITE</td>
<td>10-Mar-08</td>
<td>MODERATE</td>
<td>Ongoing</td>
<td></td>
</tr>
</tbody>
</table>
## Example 3

### Listing of Blood Pressure (mmHg) Over Time by Subject

<table>
<thead>
<tr>
<th>Visit</th>
<th>Subject 1001</th>
<th>Subject 1002</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DBP*</td>
<td>SBP**</td>
</tr>
<tr>
<td>Baseline</td>
<td>79</td>
<td>128</td>
</tr>
<tr>
<td>Visit 1</td>
<td>78</td>
<td>125</td>
</tr>
<tr>
<td>Visit 2</td>
<td>83</td>
<td>121</td>
</tr>
<tr>
<td>Visit 3</td>
<td>91</td>
<td>150</td>
</tr>
<tr>
<td>Visit 4</td>
<td>89</td>
<td>142</td>
</tr>
<tr>
<td>Visit 5</td>
<td>80</td>
<td>124</td>
</tr>
<tr>
<td>Visit 6</td>
<td>78</td>
<td>120</td>
</tr>
</tbody>
</table>

*Mean of 3 Diastolic Blood Pressure readings

**Mean of 3 Systolic Blood Pressure readings
3.2.1 Data Review Type_Method

❖ Data Quality Review

- Synonym: electronic data review / edit check review / logic check review

- Edit checks are programmed and discrepancies triggered by system

- Mainly apply to individual patient data

- Errors can be detected by computer and an algorithm can be defined

- Carefully designed edit checks can greatly increase efficiency and data quality by automating many data review processes

- Examples (see next slide)
Was a blood sample drawn at this visit? [ ] Yes [ ] No
If yes, date of blood sample collection: __/__/____

If Yes is checked, date should be present. If No is checked, date should be blank.

Blood Pressure _____/_____ (SBP/DBP)
Systolic BP is expected to be 100 – 200
Diastolic BP is expected to be 60 – 110
SBP should be greater than DBP

Weights at Month 1, Month 2, and Month 3
Weight changes more than 20% compared with last month
Lab test performed at Visit 1, Visit 2, and Visit 3

Lab test data of performed, lab test results are same for Visit 1 and Visit 3.

Drug A 20mg  
Start: 01Jan09  
Stop: 01Jun09

Drug A 20mg  
Start: 01May09  
Stop: 01Oct09

Same medication, sort by start date, and start date is prior to the stop date of previous record

Date of birth: ___/___/_____

Date of birth must be provided.
Sex: [ ] Male  [ ] Female

Was a pregnancy test performed? [ ] Yes  [ ] No

If Female, pregnancy test question should be answered.
If Male, the question should not be answered.

Reconcile with external lab data

Inconsistent birth data, date of sample collection, mismatched record, duplicate records, etc

Reconcile with safety DB (serious adverse event)

Inconsistent birth data, onset date, intensity, outcome, mismatched record, etc
3.2.2 Data Review Type_Method

- Medical Review
  - Synonym: clinical review / science review / manual review
  - Apply to all kinds of subject data (individual, group, aggregate)
  - Medical judgment is needed and an algorithm cannot be defined
  - Examples (see next slide)
**Example**

Anti-hypertension drug is recorded

No hypertension is recorded in medical history or reported as adverse event

**Example**

Tylenol is recorded being used for diabetes

Tylenol is not indicated for diabetes

**Example**

ECG result: [ ] Normal  [X] Abnormal

Specify abnormality: [Complete atrioventricular block]

No relevant medical history or adverse event reported
Out of normal range value

Need to review medical history and adverse event or other data to determine if the value is valid or not

Weight changed from baseline or previous visit

Need to review the variance in weight is due to medical history or result from adverse event (e.g. worsening HF)
Agenda (Why, What, Where, When, How)

1. Purpose of Data Review
2. Data and Error Source
3. Types of Data Review
4. Perform Data Review
5. Key Messages
4.1 Perform data review – Start Up

- **Start Up**
  - No data review activities
  - Develop Data Review Plan

- **Conduct**
  - Data collection starts
  - Majority data review work is conducted

- **Close Out**
  - Minimum data review work may be needed
  - Database is locked and frozen
4.1.1 Develop Data Review Plan

- DRP is intended to
  - describes the objectives of data review and
  - document the key data review activities (frequency, tools and reports used, etc) and
  - define the roles and responsibilities of each study team member in the data review process

- For each data review activity, DRP will define:
  - Type of Data
  - Purpose
  - Roles and Responsibilities
  - Frequency
  - Tools
<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Purpose</th>
<th>Process/ Role &amp; Responsibility</th>
<th>Frequency</th>
<th>Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRF data Central lab data Other data defined in DVC</td>
<td>To check data are complete, logical and correct</td>
<td>DM:  ◆ Continuous review discrepancies generated by the validation specifications  ◆ Monitoring the quality of discrepancy management;</td>
<td>Ongoing</td>
<td>CRF DVC Study Reports</td>
</tr>
<tr>
<td>Key data for efficacy data and signal detection</td>
<td>Focused medical data review and signal detection</td>
<td>Science: Performs medical data review and detection of potential signals and risk management.</td>
<td>Monthly Additional timepoint/ongoing</td>
<td>CRF I-Review Study Reports</td>
</tr>
</tbody>
</table>
4.1.1 Develop Data Review Plan (DRP)

- Fully understand
  - Protocol
  - CRF and other data source
  - Working scope and contract (CRO)

- Define critical data

- Two main components:
  - Data Validation Check
  - Medical Review Checklist
4.1.1 Develop DRP – Define Critical Data

Data used to make decisions in clinical trials must be accurate. Decisions about dosages, risk of adverse event and risk-benefit profiles of the treatment are made using these data.

- Study objectives
- Primary and secondary efficacy endpoint
- Primary and secondary safety endpoint

Critical data are defined by statistician, clinical science, and reviewed and agreed by the study team.
4.1.1 Develop DRP – Determine data review type

Use the appropriate, efficient method and tools to perform data review

- Study complexity / Study population
- Clinical database limitations
- Nature of data error
  - Some edit checks may be less feasible or efficient than a manual review (e.g., free text data fields);
  - Some data irregularities may be more appropriately identified by biostatisticians than through edit checks or manual reviews;
  - Some unexpected data trends may be indicative of systemic problems with data collection or processing and may not be easily identified by an edit check or manual review.
4.1.1 Develop DRP - Data Validation Checks (DVC)

- Synonym: Edit Check Specifications, Data Validation Guidelines, etc

- The DVC is a document containing the specifications of all automatic validation checks programmed to ensure data is accurate and complete within a study.
4.1.1 Develop DRP - Data Validation Checks (DVC)

DVC usually include:

- Edit Check No.
- CRF Section/Domain
- Applicable Visit / Page
- Question
- Description
- Error Message Text
- Validation Procedure Name
- Testing Information (status, dummy data location, tester initial and date, reason for failure, action)

*Additional columns may be needed per company’s standard*
<table>
<thead>
<tr>
<th>VITAL SIGNS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Rate (Seated):</td>
<td>bpm</td>
</tr>
<tr>
<td>Respiration:</td>
<td>/min</td>
</tr>
<tr>
<td>Height:</td>
<td>in</td>
</tr>
<tr>
<td>Weight:</td>
<td>lb</td>
</tr>
<tr>
<td>Arm Circumference:</td>
<td>in</td>
</tr>
<tr>
<td>Oral Temperature:</td>
<td>°F</td>
</tr>
</tbody>
</table>
# DVC Example

<table>
<thead>
<tr>
<th>Check No.</th>
<th>CRF Section</th>
<th>Question</th>
<th>Applicable visit</th>
<th>Description</th>
<th>Error Message Text</th>
<th>VP Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS001</td>
<td>Vital Signs</td>
<td>Heart Rate</td>
<td>All</td>
<td>Heart rate should be between 50 – 100 beats per min inclusive.</td>
<td>Heart Rate is not within the expected range. Please review and amend or verify as correct.</td>
<td>V_VS_HR_01</td>
</tr>
<tr>
<td>VS002</td>
<td>Vital signs</td>
<td>Weight</td>
<td>All</td>
<td>Weight should be between 40 and 150 inclusive.</td>
<td>Weight is not within the expected range. Please review and amend or verify as correct.</td>
<td>V_VS_WT_01</td>
</tr>
<tr>
<td>VS003</td>
<td>Vital signs</td>
<td>Height Screening</td>
<td></td>
<td>Height should be between 140 and 200 inclusive.</td>
<td>Height is not within the expected range. Please review and amend or verify as correct.</td>
<td>V_VS_HT_01</td>
</tr>
</tbody>
</table>
4.1.1 Develop DRP – Medical Review Checklist

Medical Review Checklist usually include:

- CRF Section/Domain
- Applicable Visit / Page
- Question
- Description (what should be reviewed and checked)
- Tools or reports

To facilitate manual review, some study-specific data listings, reports and tools are needed.
**DEMOGRAPHICS**

**Date of Birth:**

- \( dd / mmm / yyyy \)

**Race:**

- [ ] Caucasian
- [ ] Black
- [ ] Asian
- [ ] Hispanic/Latin
- [ ] Other, please specify: ______________________

**Gender:**

- [ ] Male
- [ ] Female

If "Female", please complete below:

Pregnancy Potential:

- [ ] Post-menopausal for at least 2 years
- [ ] Surgically sterile
- [ ] Able to bear children
## Screening (Visit 1)

<table>
<thead>
<tr>
<th>PHYSICAL EXAMINATION</th>
<th>Body System</th>
<th>Not Done</th>
<th>Assessment</th>
<th>If Abnormal, please specify</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. Skin</td>
<td>□</td>
<td>□ Normal</td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td>2. HEENT</td>
<td>□</td>
<td>□ Normal</td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td>3. Neck</td>
<td>□</td>
<td>□ Normal</td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td>4. Lymph nodes</td>
<td>□</td>
<td>□ Normal</td>
<td>Abnormal</td>
</tr>
<tr>
<td></td>
<td>5. Thorax</td>
<td>□</td>
<td>□ Normal</td>
<td>Abnormal</td>
</tr>
</tbody>
</table>
## Medical Review Checklist - Example

<table>
<thead>
<tr>
<th>CRF Section / Domain</th>
<th>Applicable Visit / Page</th>
<th>Question</th>
<th>Description</th>
<th>Tools or Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demography</td>
<td>Screening / Page 1</td>
<td>Race, Other</td>
<td>Check the specification is valid and not included in any of the pre-defined Race</td>
<td>Listing of Demography by Subj.</td>
</tr>
<tr>
<td>Physical Examination</td>
<td>Screening, Visit 10 / Page 1, 38</td>
<td>Abnormalities</td>
<td>Check the specification of abnormalities is valid and belongs to the correct body system</td>
<td>Listing of Abnormalities of Physical Examination</td>
</tr>
</tbody>
</table>
4.2 Perform data review – Conduct

❖ **Start Up**
  - No data review activities
  - Develop Data Review Plan

❖ **Conduct**
  - Data collection starts
  - Majority data review work is conducted

❖ **Close Out**
  - Minimum data review work may be needed
  - Database is locked and frozen
Discrepancy Management

Site → Monitor → Data Entry

Data Manager, Science, Statistician, Monitor

Data Review

Any Query?

Yes → Dirty DB

No → Clean DB

Send queries to the site

Apply for paper study
4.3 Perform data review – Close Out

- **Start Up**
  - No data review activities
  - Develop Data Review Plan

- **Conduct**
  - Data collection starts
  - Majority data review work is conducted

- **Close Out**
  - Minimum data review work may be needed
  - Database is locked and frozen
Agenda (Why, What, Where, When, How)

1. Purpose of Data Review
2. Data and Error Source
3. Types of Data Review
4. Perform Data Review
5. Key Messages
5. Take-home Key Messages

- Purpose – deliver data that fits the quality level and serve as the basis for correct conclusion and interpretation
- Understand the data and error source and concentrate on the critical data
- Develop Data Review Plan before the review activities start and the efficient and appropriate data review method
- Data quality is a shared responsibility across functions and departments which needs teamwork and joint efforts
Question?