Clinical Data Management
-An Introduction
Clinical Data Management is involved in all aspects of processing the clinical data, working with a range of computer applications, database systems to support collection, cleaning and management of subject or trial data.
Clinical Data Management is the collection, integration and validation of clinical trial data.

During the clinical trial, the investigators collect data on the patients' health for a defined time period. This data is sent to the trial sponsor, who then analyzes the pooled data using statistical analysis.
Why CDM

• Review & approval of new drugs by Regulatory Agencies is dependent upon a trust that clinical trials data presented are of sufficient integrity to ensure confidence in results & conclusions presented by pharma company
• Important to obtaining that trust is adherence to quality standards & practices
• Hence companies must assure that all staff involved in the clinical research are trained & qualified to perform data management tasks
Key Members

The Key members involved in Data Management

• Project Manager / Data Manager
• Database Administrator
• Database Programmer / Developer
• Clinical Data Associate
Clinical Trials in a Nut Shell
Multidisciplinary Teams in Clinical Trials

1. Clinical Investigator
2. Site coordinator
3. Trial Pharmacists
4. Biostatistician
5. Lab Coordinator
6. Project manager
7. Clinical Research Manager/Associate
8. Monitor
9. Ethics committee
10. Regulatory affairs
11. Clinical Data Management
12. Pharmacovigilance
13. IT/IS personnel
14. Clinical supply
15. Auditor/Compliance
Responsibilities of CDM

**Study Setup**
- CRF design and development (paper/e-CRF)
- Database build and testing
- Edit Checks preparation and testing

**Study Conduct**
- Data Entry
- Discrepancy Management
- Data Coding (using MedDRA and WHODDE dictionaries)
- Data review (Ongoing QC)
- SAE Reconciliation
- Data Transfer

**Study Closeout**
- SAE Reconciliation
- Quality Control
- Database Lock
- Electronic Archival
- Database Transfer
CDM Process Overview

Startup phase
- Protocol design
- CRF Design
  - Quality check
- Database Design
  - Quality check
- Edit Checks
  - Quality check
- Database activated

Conduct phase
- Data entry
- Discrepancy management - Query to investigator
- Coding of medical terms
- Database updates

Close out phase
- Database QC
- Database Lock
Study Start Up Process Review

- Protocol
- Database design
- CRF design
- Validation/derivation Procedures
- Activated database ready to accept production data
CRF Design/Review

A representation of the study as outlined in the protocol is made (including CRF completion guidelines if necessary). Therefore a final protocol needs to be available before this activity can be initiated. CRF design usually takes about three rounds: First draft (rough without detail but correct content), second draft (as good as we can get it) and final version. We need input from our sponsor to correct draft versions and to approve the final version.

QADATA EDC
- Traditional Paper Based Case Report Forms
- e-CRF (Electronic Case Report Form)- Study information directly entered into computer.
LYME DISEASE CASE REPORT FORM

<table>
<thead>
<tr>
<th>Patient's last name:</th>
<th>First name:</th>
<th>Telephone:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address:</td>
<td>City:</td>
<td></td>
</tr>
<tr>
<td>State:</td>
<td>County:</td>
<td></td>
</tr>
<tr>
<td>Age (yrs):</td>
<td>Sex: M/F</td>
<td></td>
</tr>
<tr>
<td>Race: Asian/Pacific Island</td>
<td>Ethnicity:</td>
<td></td>
</tr>
<tr>
<td>Unspec.: Black/White</td>
<td>Unknown:</td>
<td></td>
</tr>
</tbody>
</table>

SYMPTOMS AND SIGNS OF CURRENT EPISODE (PLEASE MARK EACH QUESTION)

| DERMATOLOGIC: | Erythema migrans (physician diagnosed EM at least 5 cm in diameter) | [Y] [N] [?]
|---------------|---------------------------------------------------------------------|
| RHEUMATOLOGIC:| Arthritis characterized by brief attacks of swelling in one or a few joints | [Y] [N] [?]
| NEUROLOGIC:   | Bells palsy or other cranial neuralgia | [Y] [N] [?]
|               | Radiculoneuropathy | [Y] [N] [?]
|               | Lymphocytic meningitis | [Y] [N] [?]
|               | Encephalitis | [Y] [N] [?]
| CARBONIC:     | 2nd or 3rd degree atrioventricular block | [Y] [N] [?]

OTHER CLINICAL

<table>
<thead>
<tr>
<th>Date of onset of first symptoms:</th>
<th>Date of diagnosis:</th>
<th>Date of report to health agency:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]</td>
<td>[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]</td>
<td>[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]</td>
</tr>
</tbody>
</table>

OTHER HISTORY

| Was the patient hospitalized for the current episode? | [Y] [N] [?]
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of antibiotic(s) used during episode? Use index:</td>
</tr>
</tbody>
</table>
| Was the patient pregnant at the time of the illness? | [Y] [N] [?]
| Where was the patient most likely exposed? County: | [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] |

LABORATORY RESULTS

<table>
<thead>
<tr>
<th>Seminotic test results:</th>
<th>Positive</th>
<th>Negative</th>
<th>Equivocal</th>
<th>Not done</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture results:</td>
<td>[ ] [ ]</td>
<td>[ ] [ ]</td>
<td>[ ] [ ]</td>
<td>[ ] [ ]</td>
<td></td>
</tr>
<tr>
<td>Other (specify):</td>
<td>[ ] [ ]</td>
<td>[ ] [ ]</td>
<td>[ ] [ ]</td>
<td>[ ] [ ]</td>
<td></td>
</tr>
</tbody>
</table>

VISIT 1

- Required consent:
  - Data informed consent signed: [ ] [ ] [ ]

- Demographics:
  - Gender: M/F
  - Race: White/Black/Asian/Other (specify)

- EOC Vaccination:
  - Yes: [ ]
  - No: [ ]
  - Unknown: [ ]

<table>
<thead>
<tr>
<th>Visit</th>
<th>Visit Date</th>
<th>Date of VTEC 0012</th>
<th>VTEC 0012 Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]</td>
<td>[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]</td>
<td>[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]</td>
</tr>
</tbody>
</table>

[Image of e-CRF interface]
How many CRFs do you need?

- Eligibility or Screening
- Randomisation
- Physical Exam / Vitals
- Medical History
- Follow-up Visit
- AE form/ SAE form
- Concomitant therapy form
- Laboratory test form
- Status Evaluation
Data Base Design

Data from a clinical trial will be collected and stored in the CDMS. A database is simply a structured set of data. A collection of rows and columns.
--QADa
ta CDMS
<table>
<thead>
<tr>
<th>Subject Id</th>
<th>Name of patient</th>
<th>Age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>A23691</td>
<td>XYZ</td>
<td>23</td>
<td>M</td>
</tr>
<tr>
<td>A23692</td>
<td>XYA</td>
<td>24</td>
<td>M</td>
</tr>
<tr>
<td>A23693</td>
<td>XYB</td>
<td>25</td>
<td>F</td>
</tr>
<tr>
<td>A23695</td>
<td>ABX</td>
<td>26</td>
<td>M</td>
</tr>
</tbody>
</table>

**DBMS:**
- MS Access, MS Excel
- Oracle Clinical
- Clintrial
- Phaseforward InForm
- medidata Rave
CRF Annotation

• An annotated CRF is generally defined as a blank CRF with markings, or annotations, that coordinate each data point in the form with its corresponding dataset name.
• Essentially, an annotated CRF communicates where the data collected for each question is stored in the database.
• CRF Annotation is the first step in translating the CRFs into a database application.
• CDM annotates the CRFs by establishing variable names for each item to be entered.
• Reviewed by CDM and Statistician
### Informed Consent

NO study related activities may take place before the patient has signed the Informed Consent form.

<table>
<thead>
<tr>
<th>Date informed consent was obtained:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Date informed consent was obtained:</td>
<td>DD</td>
<td>MON</td>
<td>YYYY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Time informed consent was obtained:

<table>
<thead>
<tr>
<th>Time informed consent was obtained:</th>
</tr>
</thead>
<tbody>
<tr>
<td>dd:mm</td>
</tr>
</tbody>
</table>

### Demographics

Date of Birth:

<table>
<thead>
<tr>
<th>Date of Birth:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of Birth:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Validation Checklist:

Edit specifications list describes in detail which data shall be checked and queried if necessary. The programming of the checks occurs according to this list. Before the programming starts, the sponsor will be asked to give approval of this list.

Test subjects are entered in the database to test the entry screens and the programming. The exact number of test subjects is not standard, but every check has to pass and fail (negative and positive proof) at least once.
Database setup and testing are always performed in a secure, non study data environment (test site). Only when a database has been reviewed and fully tested, will it be set in ‘production’, a separate environment where only study data will be entered. Changes in structure or programming will always first be performed and tested in the non study data environment before they are made effective in the ‘production’ database.
<table>
<thead>
<tr>
<th>table_name</th>
<th>crf_section_name</th>
<th>column_name</th>
<th>CRF Page</th>
<th>DCF Type</th>
<th>Query text</th>
<th>Test Condition</th>
<th>Test Plan</th>
<th>Pass/Fail, Date, Patient number</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHAND_CIV</td>
<td></td>
<td>invddat</td>
<td>Vis 1 page 10/Vis 2 page 14/Vis 3 page 18/Vis 4 page 23/Vis 5 page 28</td>
<td>Blank</td>
<td>Investigator date has not been completed, please provide.</td>
<td>Visit 1</td>
<td>Q: Investigator date = blank</td>
<td>P, 99-010, 25SEP201</td>
</tr>
<tr>
<td>EHAND_DS</td>
<td>informed consent</td>
<td>dvsdtat</td>
<td>Vis 1 Day 1</td>
<td>Blank</td>
<td>Date of informed consent has not been completed, please check and provide</td>
<td>Date of visit = 01 JUN 2012</td>
<td>Q: Consent date = blank</td>
<td>P, 99-010, 12SEP201</td>
</tr>
<tr>
<td>EHAND_DS</td>
<td>informed consent</td>
<td>dvsdtat</td>
<td>Vis 1 Day 1</td>
<td>invalid</td>
<td>Date of informed consent may not be after Vis 1 date, please check</td>
<td>Date of visit = 01 JUN 2012</td>
<td>Q: Consent date = 02 JUN 2012</td>
<td>P, 99-010, 12SEP201</td>
</tr>
<tr>
<td>EHAND_DS</td>
<td>informed consent</td>
<td>dssdtim</td>
<td>Vis 1 Day 1</td>
<td>Blank</td>
<td>Time of informed consent has not been completed, please check and provide</td>
<td>Q: Consent time = blank</td>
<td>Q: Consent time = 08:00</td>
<td>P, 99-010, 12SEP201</td>
</tr>
<tr>
<td>EHAND_DS</td>
<td>inclusion / Exclusion criteria review</td>
<td>evyn</td>
<td>Vis 1 Day 1 page 7</td>
<td>Blank</td>
<td>Did the subject meet all eligibility criteria is not completed, please check and provide details</td>
<td>Q: Subject met all eligibility criteria = Blank</td>
<td>Q: Subject met all eligibility criteria = Yes</td>
<td>P, 99-010, 12SEP201</td>
</tr>
<tr>
<td>EHAND_DS</td>
<td>inclusion / Exclusion criteria review</td>
<td>eincl</td>
<td>Vis 1 Day 1 page 7</td>
<td>Blank</td>
<td>Is the subject eligible to continue is not completed, please check and provide details</td>
<td>Q: Patient included in the study = blank</td>
<td>Q: Patient included in the study = Yes</td>
<td>P, 99-010, 21SEP201</td>
</tr>
<tr>
<td>EHAND_DS</td>
<td>inclusion / Exclusion criteria review</td>
<td>ictestcd</td>
<td>Vis 1 Day 1 page 7</td>
<td>invalid</td>
<td>Reason has not been specified, but ‘No’ answered above, please check</td>
<td>Patient included in the study = No</td>
<td>Q: Reason = blank</td>
<td>P, 99-010, 21SEP201</td>
</tr>
<tr>
<td>EHAND_DS</td>
<td>inclusion / Exclusion criteria review</td>
<td>ictestcd</td>
<td>Vis 1 Day 1 page 7</td>
<td>Questionable</td>
<td>Reason has been specified, but ‘Yes’ answered above, please check</td>
<td>Reason for non-inclusion = ABC</td>
<td>Q: Patient included in the study = Yes</td>
<td>F, 99-010, 21SEP201</td>
</tr>
</tbody>
</table>
CRF Tracking

Logistic way if it is paper based study.
EDC-electronic data capture if it is e-CRF.
Data Entry

Data Entry

Data entry is a process of entering/transferring data from case report form to the Clinical Data Management System (CDMS).

Data Entry: 1) Single data Entry
2) Double Data Entry
Discrepancy management is a process of cleaning subject data in the Clinical Data Management System (CDMS), it includes manual checks and programmed checks. Trivial discrepancies are closed as per self evident correction method or Internal rulings and discrepancies which require response from the site are queried by raising Data Clarification Forms (DCF).
The medical coding for a study is done as per the project specific protocol requirement. The dictionaries used for a study are:

Adverse Events: MedDRA (Medical Dictionary for Regulatory Activities)
Medications: WHODD (World Health Organization – Drug Dictionary)
SAE Reconciliation

• Serious Adverse Event (SAE) data reconciliation is the comparison of key safety data variables between Clinical Data Management System (CDMS) and Sponsor PV. Reconciliation is performed to ensure that events residing in both systems are consistent.
Study Close out Process Review

- Discrepancy management
  - Safety data reconciliation
  - Query generation
  - Resolution and/or update of database
  - Manual checks/QC/CRF tracking
  - Database lock & freeze
- Coding terms
Quality Control

• Quality Should be maintained for overall study by performing Quality checks at intervals for all data points (Critical & Non-Critical) prior to database lock.
• QC helps to ensure that all the data processed is accurate, clean and Correct.
but I eated it.
The database lock for a study is done to ensure no manipulation of study data during the final analysis.

Database lock for a study is done once all data management activities are completed. This includes the database lock checklist which ensures the same. Some of the activities included in database lock checklist are All discrepancies closed, DCFs received and updated, coding complete, SAE Reconciliation process complete etc.
Analysis & Reporting Process Review

Database release → Data extraction/Mapping → Statistical report generation → E-publishing → Published tables, figures and listings → Creation of Clinical study report (CSR) → Submission of CSR
CDM is a vital vehicle in Clinical Trials to ensure:

- The Integrity & quality of data being transferred from trial subjects to a database system
- That the collected data is complete and accurate so that results are correct
- That trial database is complete and accurate, and a true representation of what took place in trial
- That trial database is sufficiently clean to support statistical analysis, and its subsequent presentation and interpretation
Importance of CMD

CDM has evolved from a mere data entry process to a much diverse process today

- It provides data and database in a usable format in a timely manner
- It ensures clean data and a ‘ready to lock’ database
CDM Professionals

- ICH.E6.5.5.1: Utilize qualified individuals to:
  - Supervise overall conduct of trial (Project Manager)
  - To handle and verify the data (Data Manager)
  - To conduct the statistical analysis (Biostatistician)
  - To prepare study reports (Medical Writer)
The data management function provides all data collection and data validation for a clinical trial program.

Data management is essential to the overall clinical research function, as its key deliverable is the data to support the submission.

Assuring the overall accuracy and integrity of the clinical trial data is the core business of the data management function.
DM Role in Clinical Research

Data management starts with the creation of the study protocol.

At the study level, data management ends when the database is locked and the Clinical Study Report is final.

At the compound level (of the drug), data management ends when the submission package is assembled and complete.
Mission of CDM

Consistency
Accuracy
Validity
Archiving