## Options for developing in-house data management capacity

#### **Turning data into information...**







Lesley Workman Data Management, GHT January 2013

# "No study is better than the quality of its data"

#### Fundamentals of Clinical Trials, Friedman LM, Furberg CD, DeMets DL, 1998





## **Background: Computerised DM**

- 1990's PC's, databases and spreadsheets (DBase, Quatro)
- 1996 Commercial databases (Oracle Clinical)
- 2006 Open source commercial (OpenClinica)

Clinical Data Management is now a recognised profession, increasingly becoming recognised as an essential component of clinical research





## What is Data Management?

- Process of defining, gathering, capturing, cleaning, monitoring, analysing and reporting study data
- Starts with protocol development
- Ends with....
  - Once the final results have been published?
  - Once the study has been archived?
  - Once the data has been integrated into a larger data sharing repository to be used to answer important global health or related issues?





## **Objectives of a CDMS**

#### Produce data

- of provable quality
- timeously
- while adhering to the regulatory and legal requirements regarding:
  - Privacy
  - Auditability
  - Electronic signatures
- from which information can be generated to answer the research question









#### Database

- Database engine including the programs, meta-data and configuration tool
  - Distinct differences between packages
    - General purpose end-user packages such as MS Access
    - Industry-specific packages such as Oracle Clinical
- System security (controlled & secure access to participant data)
  - Is built into commercial packages
  - Must be programmed into MS Access (but with significant limitations)
- Validation encompasses
  - all aspects of the package installation
  - appropriate hardware and software versions
  - <u>correct installation of additional components</u>



#### Configuration

- Define the data and rules about the data
  - Commercial packages have a built in structure which can be readily configured by a trained researcher
  - The data structure and rules must be created in MS Access by a skilled MS Access resource
- System assurance
  - Configuration is fit for purpose
  - Consists of standard tests that ensure all rules are triggered and data is correctly stored
  - Assurance must be performed whenever there is a change to the configuration





### Management

- The management component of the CDMS comprises a Data Management Plan, SOPs and CRFs
  - Suitably trained staff
    - Role definitions, training plans and materials
  - Data Management Plan
    - Includes a quality management plan, CRF design, database design, validation and build, audit trails, CRF tracking and storage etc etc etc
  - SOPs
    - For every aspect of data handling process from CRF completion through to archiving of paper and electronic records
  - Audit procedures
    - Standard audit procedures to ensure that regulatory compliance requirements are met.



## **Types of CDM systems**

#### In-house built

- MS Access, Filemaker
- Open source packages
  - Openclinica, Redcap, EpiInfo
- Commercial packages
  - Cmed, Oracle Clinical, DataFax





#### Different types of studies require different type of CDMS

Phase 1, 2 and 3.

**FDA 21 CFR** – to deter record/signature falsification Part 11.10(e) stipulates an "independently-recorded" audit trail 21 CFR Part 11.10(j) states its intent is to ensure an irrefutable link between the electronic record and the electronic signature

- Phase 4
- Epidemiologic, Observational/surveys
- Record collection/folder review

#### But they all need to comply with ICH/GCP DM principles





#### **Industry vs Academic**

#### Pharma Industry

- Minimising time to the final study report
- Assured of compliance to regulations governing registration of a new product
- Academic/research Institutions
  - Scope of the work not well defined
  - Limited budget
  - Lack of suitable skilled staff



## Study requirements to consider when selecting a CDMS

- Will the data need to be submitted to regulatory authorities such as FDA for possible registration?
- Coding dictionaries e.g. MEDdra (licensing)
- National and other regulatory requirements/guidelines such as
  - CFR 21 Part 11 for the United States
  - Directive 2001/20/EC for the EU countries
  - ICH GCP and GCDM
  - POPIA (Protection of Personal Information Act) RSA 2013



## **Quality triangle**

#### Cost

 Reducing cost will either increase time, reduce scope or compromise quality

Time

- Shortening the time to database lock will comprise quality and/or increase cost
- Quality
  - Compromising on quality makes us less confident of the data



## **Out-source CDMS**

#### Issues relating to an out-sourced CDMS

- Sufficient budget? (hidden costs)
- Validation of the system to ensure the necessary compliance to regulations – who is responsible?
- Version control, back-up and server hosting
- Staff to do the data capture and query resolution (and could still need an in-house data manager)
- Clear roles and responsibilities
- How to handle external data and who is responsible for mapping
- Who owns the data at the end of the study?
- Issues if data is held offshore EU privacy laws impact 'export' of your own data



### **In-house CDMS**

- Data design issues (DB structures, specification)
- Security & audit trail issues
- Validation (IQ, OQ, PQ)
- Management (staff, SOP's and DM plan)
- Version control
- Back-up and recovery



## **Personal Experiences**

- Very high costs of out-sourcing
  - Compromise on the scope of the study (less patients/visits)
- Lack of trained Data Designers/Managers
  - Excel doesn't cut it!
  - spreadsheets' in Access
  - massive number of queries/large chunks of missing data
  - free text fields are not data
- Vendor management
  - Commercial agreements
  - Ownership of data
  - Additional cost for changes or omissions/misunderstandings



#### **Laboratory and Other External Data**

- Essential that the data integrity, quality and confidentiality are maintained
- QC on each stage of the data handling process to ensure all data are reliable and accurate
- Maintain a document and system audit trail
- Test data transmission before a live upload
- Early communication with external vendors to
  - Establish key and mandatory variables-specifications
  - Editing and verification procedures
  - Format and mapping of the transfer data
  - Query resolution
  - Maintain the study blind if applicable



#### How can we move towards in-house CDMS that meet appropriate regulatory requirements?

- Create a centre of excellence
  - Standardise on Open Source Database e.g. OpenClinica
  - Standardise Configuration
    - Standard Data Structure
    - External dictionaries
  - Standardise Management
    - DM plan/quality management plan
    - SOP's
    - Training courses



## And this will lead to.....

- Capacity building
- Strengthening of the Health Science Faculty research capacity
- A reference centre for other research institutions
- Retention of skilled expertise in the clinical research environment





#### "Experience is the name everyone gives to their mistakes"

(Oscar Wilde)





Lesley Workman Data Management, GHT January 2013