



# Data Management

Tsitsi Bandason

11<sup>th</sup> March 2019

# Objective of the Session

- To enable you to recognize the role of a good data management system in research



# Introduction

- Research data management (RDM) is a process which includes
  - Case Report Form(CRF)/Questionnaire Design
  - Procedures for data collection
  - Quality control of data (cleaning/validation)
  - Tracking of participants
  - Databases design
  - Data Storage and Backup
  - Processing of data into reportable results (Analysis)
  - Data Storage and Transfer/Sharing

# What is DATA ?

- Facts and statistics collected together for reference or analysis (*dictionary*)
- A set of values of subjects with respect to qualitative or quantitative variables (*Wikipedia*)
- **“Research data, unlike other types of information, is collected, observed, or created, for purposes of analysis to produce original research results”** (*University of Edinburgh*)

# Data Management Plan

- Data management plan is set up before the beginning of the study
  - It is a document that outlines how the research data will be handled both during and after the end of the project (Wikipedia)

# Why We Need a Concise Data Management Plan

- Transparency & Integrity
  - Provide high quality reliable/traceable data
  - Keep errors and missing data low
- Compliance
  - Protocol Guidelines
  - Good Clinical Data Management Practices
  - General Data Protection (GDPR) standards.

# Who Uses The Data or Results

- Your Institution
  - Publications
- Policy Makers
- Other Researchers
  - Data Sharing

# Key Questions



- What data is required to be collected?
  - Only data specified in the protocol
  - Data required to answer the study questions
- When will this data be collected?
  - Enrolment / follow-up
- What CRFs will need to be designed.
  - Who is going to complete the CRFs.
- How is the data going to be analysed



# Other Considerations

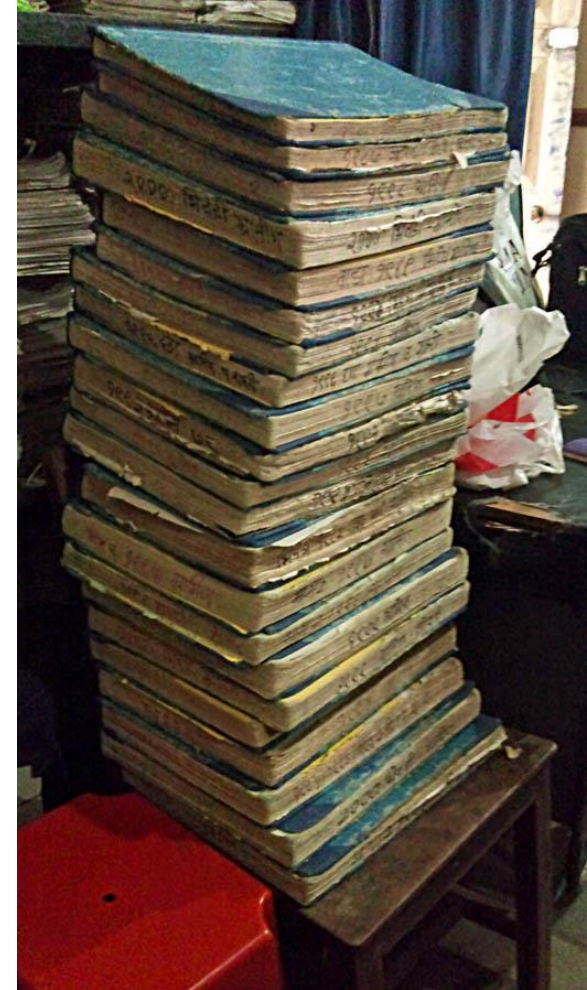
- Have a clear understanding of how much data will be collected to allow for
  - Selection of database management software
  - Procedures for data entry
  - Procedures for quality control
  - Procedures for data storage, and maintenance of hard copies and electronic data
- Need to obtain authorisations for data collection
  - Approvals
  - Consent

# Other Considerations

- Need to use appropriate and reliable methods of data collection
  - Data Collection and Management Software
  - Paper or Electronic
  - Show evidence of when data collected, changes made and reasons for change (audit trail)

# Other Considerations

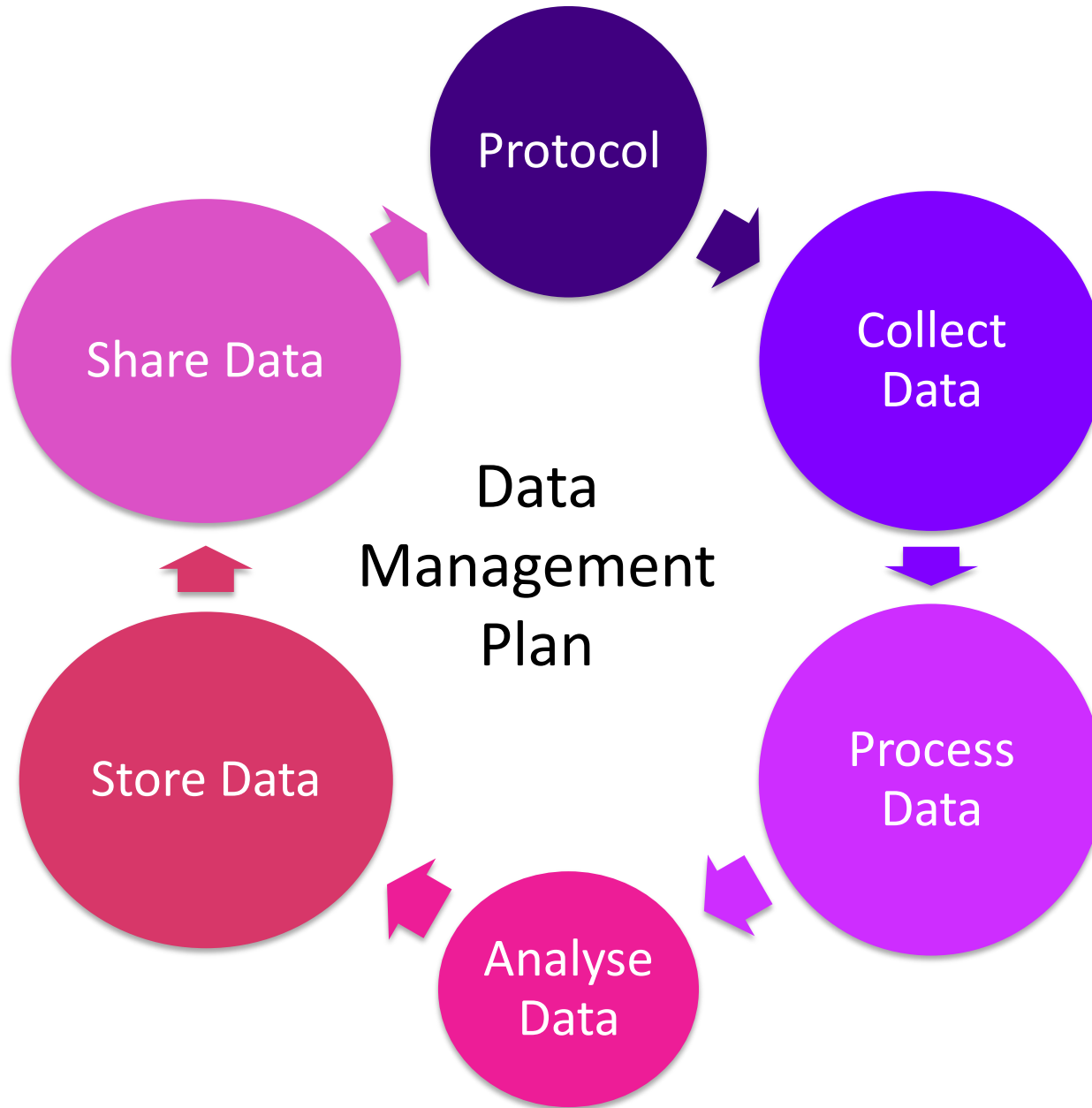
- Need to pay attention to detail
  - Photographic memory
  - Good record keeping
- Keep a record of everything
  - What is done
  - What is observed
  - What is achieved



# Other Considerations

- Data protection is priority
  - Data should be stored in a safe place
  - Data should be backed up if on electronic devices
  - Stored in locked cabinets if on paper
- Confidentiality agreements should be honoured
- Data should be preserved for a prescribed period
  - Based on institutional policies (10+ years)





# Data Collection

- The goal of data collection is to have an exact electronic copy of the data from the field/laboratory
- Choose an appropriate and convenient software
  - Allows for correction of entry errors
  - Allows QC checks during entry
- Data should be entered on real time basis as far as possible to allow for correction of errors
- All hard copies of forms should be correctly filed

# Data Entry/Collection

- Data Entry from CRF
  - Excel
  - Epi Info
  - Epi data



- Data Entry on Mobile Phone/Tablet
  - Epi Info
  - ODK
  - Survey CTO



# Why I now choose Mobile instead of Paper

- Several challenges with paper data collection
  - High human error rate
  - Delays in data entry
  - Lack in flexibility



# Data Checks

- Data Checks (Manual and Electronic)
  - Limit/Accuracy Checks
  - Related fields
  - Required field /Missing
  - Pattern

# Database Software

- A **Database** is a collection of information that is organized so that it can be easily accessed, managed and updated (WhatIs.com)
- Database
  - Ms Excel, Ms Access, SQL, Oracle,....

# Data Storage and Security

- Back-up data
  - CDs, Tapes (Label appropriately)
  - Store securely on/off site
  - Document location of storage media
- Filing of questionnaires – ensure ease of access
  - Label appropriately
  - Store securely

# Data Analysis

- *“Science of statistics cannot be ignored if you are in the business of collecting data”*
- It is a process of systematically collecting, describing, detecting trends, testing hypothesis and presenting and interpreting data
- This information is documented in Data analysis Plan



**That's work for another day**

# Data Management Disasters

- Data Collection
  - Missing Forms or Data
  - Coding errors
  - Lost Data
  - Duplicates records that cannot be identified
  - Inaccurate data
- Other
  - Forms do not link (linked data/relational data)

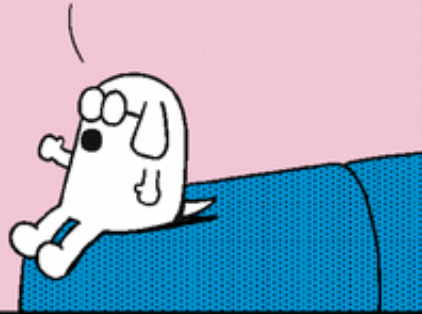
# Data Management Disasters

I NEED SOME DATA FROM AN UNREACHABLE GUY NAMED ED. WHAT SHOULD I DO?



www.dilbert.com scottadams@aol.com

JUST MAKE UP A BUNCH OF DATA LIKE EVERYONE ELSE DOES.



11-11-06 © 2006 Scott Adams, Inc./Dist. by UFS, Inc.

EVERY— ONE ELSE DOES THAT?



ARE YOU DOUBTING MY DATA?

<C:\Users\zenith\Downloads\Data Sharing and Management Snafu in 3 Short Acts.mp4>

# Essential Elements

- Have and Implement a Data Management Plan
- Make sure you have essential staff with appropriate expertise and willing to learn
  - Train & re-train & re-train your staff

Thank You