Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

CAHS Research Education Program

Research Skills Seminar

Data Collection and Management

28 July 2023

Presented by

Dr Jane Mugure Githae
CAHS REP Research Fellow

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

Child and Adolescent Health Service, Department of Research

Department of Health, Government of Western Australia

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

Data Collection and Management

PRESENTATION SLIDES

CAHS Research Education Program Research Skills Seminar Series
✉ ResearchEducationProgram@health.wa.gov.au
cahs.health.wa.gov.au/ResearchEducationProgram

Government of Western Australia
Department of Health

Perth Children's Hospital

CAHS Research Education Program

2023 Research Skills Seminar Series

Data Collection and Management

Presented by *28th July 2023*

Dr Jane Mugure Githae
CAHS REP Research Fellow

Perth Children's Hospital Foundation

1

Acknowledgement of Country

The Child and Adolescent Health Service acknowledge Aboriginal people of the many traditional lands and language groups of Western Australia.
We acknowledge the wisdom of Aboriginal Elders both past and present and pay respect to Aboriginal communities of today.

2

CAHS Research Education Program

Research Skills Seminar Series

A free, open-access resource designed to upskill busy clinical staff and students and improve research quality and impact.

- Over 20 topics across the research process
  - 1h overview
  - Handouts are provided
- Recorded and uploaded
- Feedback
  - Back of handout
  - Emailed link
- Please hold questions to the end
  - Use provided microphone

3

- ❑Definitions
- ❑Quality data
- ❑Data Collection
- ❑Data Management
- ❑Database Development

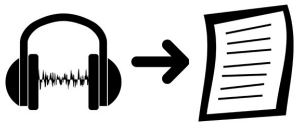
4

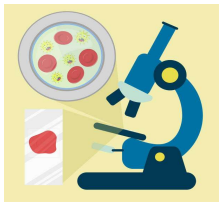
Definitions

5

Research Data

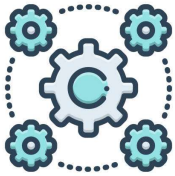
- Any information collected, observed or created during a study
- Can be:
  - Numerical
  - Descriptive
  - Visual
  - Tactile
- May be raw or processed


6

Quality Data: Data should be FAIR



FINDABLE ACCESIBLE INTEROPERABLE REUSABLE

https://www.nature.com/articles/sdata201618
https://ardc.edu.au/resources/aboutdata/fair-data/

7

Data Management Processes

Data collection

Data organisation

Data storage

Version control
Access & security
Backup

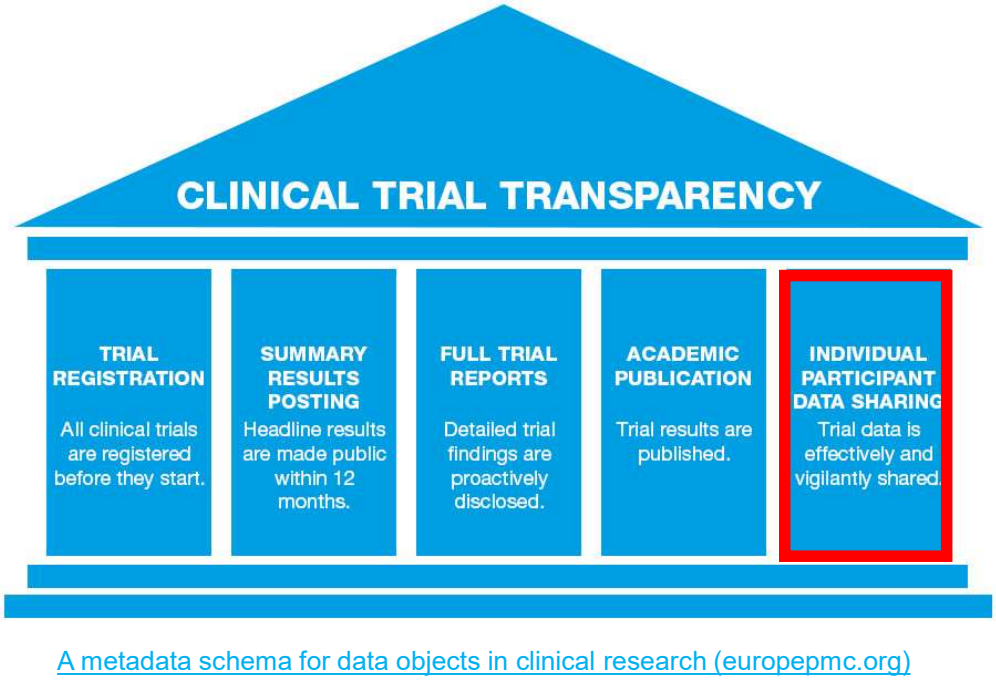
Archiving & Disposal

Data Sharing

8

Quality Data

9



A metadata schema for data objects in clinical research (europepmc.org)

10

Data Management Guidelines

- Australian Code for Responsible Conduct of Research(2018)
  - Management of Data and Information in Research NHMRC
- Institutional policy
  - Data and statistics (health.wa.gov.au)
- Funder policy
  - Funding Rules/Grant Guidelines | Australian Research Council
- Publisher policy
  - On acceptance of manuscript, data must be availed

11

Australia Policies and Codes

May 2020	Commonwealth Privacy Act (S95 and S95A)
Jul 2018	National Statement on Ethical Conduct in Human Research
Jun 2018	NHMRC: Australian Code for the Responsible Conduct of Research
May 2018	Competencies for Australian Academic Clinical Trials
1992	Freedom of Information Act (WA)

WA Health Research Governance Policy and Procedures

12

Institutional responsibilities

- ❑Training of researchers
- ❑Ownership, stewardship and control of research data and primary materials
- ❑Storage, retention and disposal
- ❑Safety, security and confidentiality
- ❑Access by interested parties

Management of Data and Information in Research NHMRC
Data and statistics (health.wa.gov.au)

13

Researcher responsibilities

Data management plan:

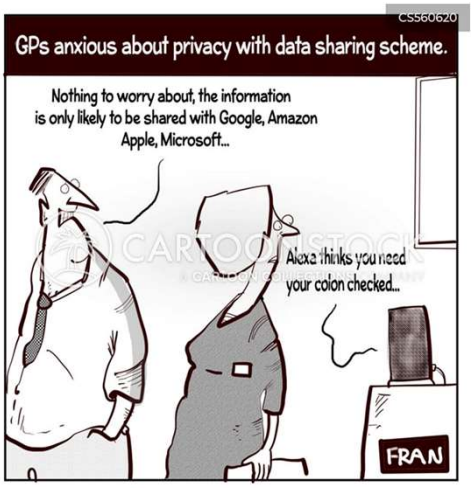
- ❑Security
- ❑Policies & procedures
- ❑Licensing agreements
- ❑Training team members
- ❑Data collection form
- ❑Retention & publication

Management of Data and Information in Research NHMRC

14

Benefits to Researchers

- Easily accessible data
- Better use of project resources by avoiding duplication
- Allows for meta-analysis & secondary analysis



15

Data Collection

16

I. Planning Data Management

Research Data Management Plan (RDMP)

- Versatile – evolves with project
- Useful for new members, audit & post-project review
- RDMP Checklist (Institutional)
  - Data Management Plans | Australian Research Data Commons | ARDC
  - Ethics & Governance - WSLHD (nsw.gov.au) - scroll down for Research Data I Plan /Checklist

17

Research Data Management Checklist

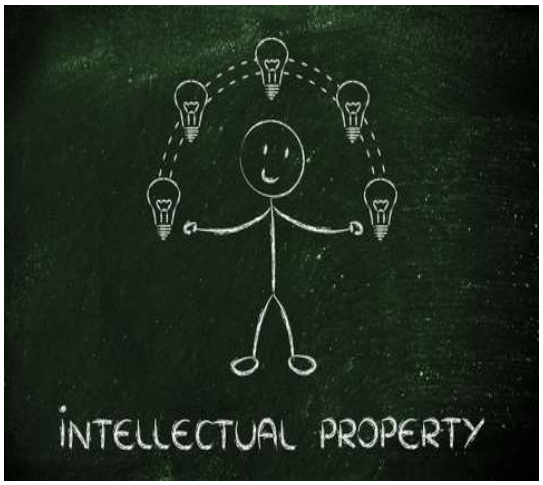
- ❑Type of data
- ❑Data stakeholders - IP
- ❑File organization
- ❑Metadata
- ❑Storage – backup, archiving, disposal
- ❑Security – access and confidentiality
- ❑Sharing
- ❑Budget

18

Planning Research Data Management (2)

Ownership:

- ❑Clarify from the start:
- ❑Who can publish findings?
- ❑ Who can take data with them when they leave



Intellectual property (health.wa.gov.au)

19

II. Data collection

1. Types of data

a) Quantitative
- Test a hypothesis
- Precise measurement of an occurrence
- Gain large-scale statistical insights

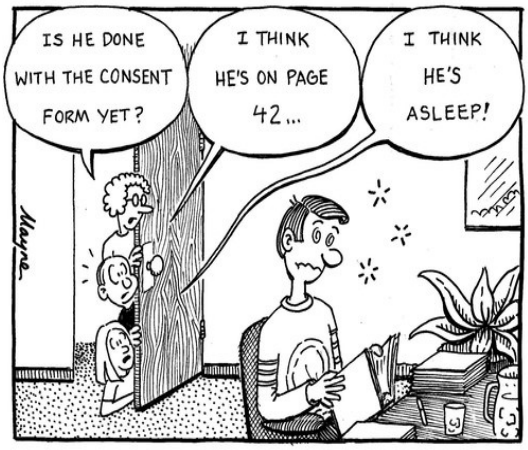
b) Qualitative
- Explore ideas
- Understand experiences
- Gain insights on a specific issue

20

Data Collection (2)

2. Consent forms

- ❑ Storage & security
- ❑ Confidentiality – Anonymisation
- ❑ Data usage
- ❑ Data sharing
- ❑ Future research

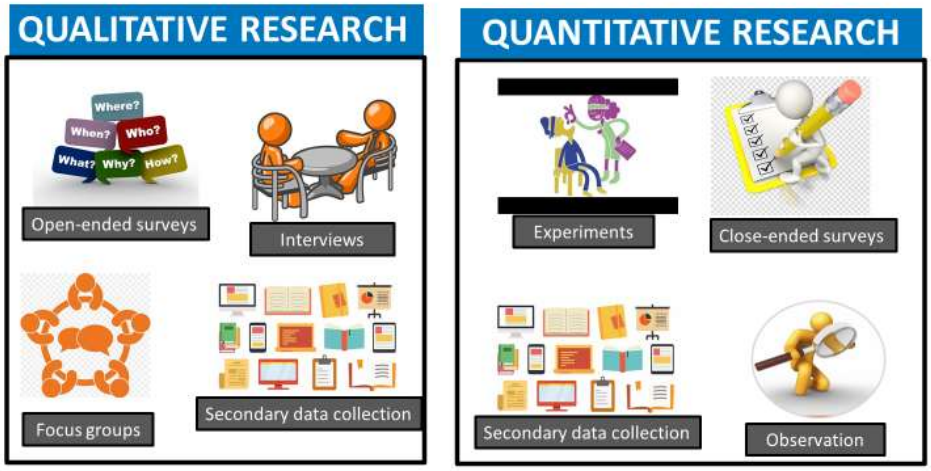


RGS - Document Templates (health.wa.gov.au)

Ethical issues and resources | NHMRC

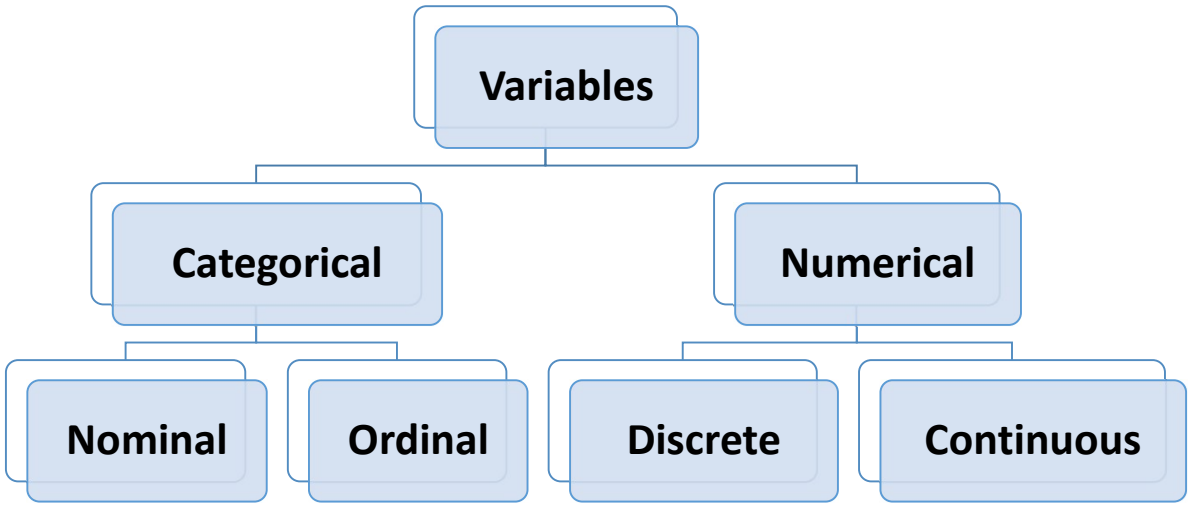
21

Methods of Data Collection



22

Data collection: Variables



23

Data Collection Procedures

- ❑Sampling & Recruitment strategies
- ❑Data collection tool (CRF)
- ❑Mode of data collection and storage
- ❑Training
- ❑Piloting
- ❑Quality assurance & monitoring

24

Data Collection Form

- ❑Only collect what you need!
- ❑Unique identifier
- ❑Collect data as continuous variables
  - ▪Categorize during analysis
- ❑Test and modify the tool as required

25

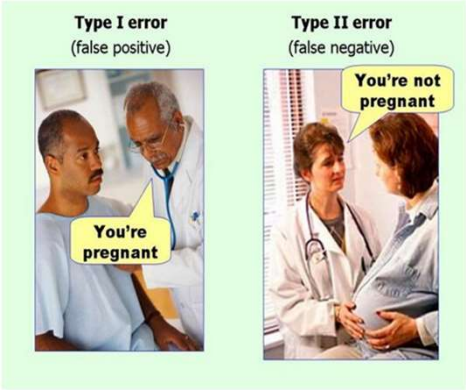
Mode of Data Collection

1. Paper based data capture
2. Electronic data capture (**REDCap**)
   - ✓Free & Secure
   - ✓User friendly
     - ▪ Data Library Department of Health
     - ▪ CAHS REDCap Resources
3. Hybrid
   - ✓ Manual or scanned

26

III. Data Cleaning

- Spotting and resolving data errors or inconsistencies
- Improves data quality
- Clean data:
  - Complete and Consistent
  - Valid and Accurate
  - Single entry
  - Uniform format



27

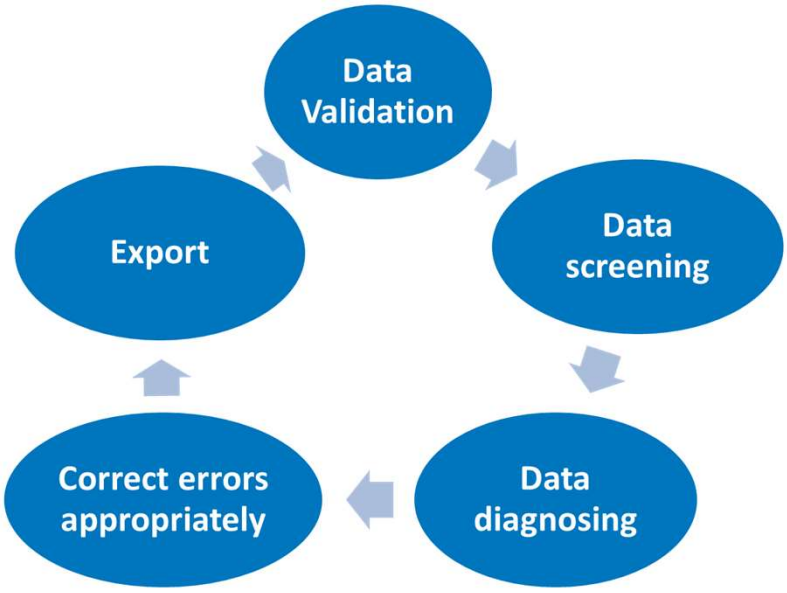
Data Cleaning (2)

- Software available
  - OpenRefine (Google Refine)
  - Statistical software
    - SPSS, R, Python

- Back up and preserve original dataset
- Save cleaned file with new name (command file)
- Document each step

28

Data Cleaning Process



- Inconsistent data
- Manually or statistical software

- Duplication
- Invalid data
- Missing Data
- Outliers

- Deduplication
- Data standardisation
- Deal with Missing Data
- Deal with Outliers

29

Data Management

30

Starting Point

https://retprogram.org/training/ich-good-clinical-practice-gcp-certificate/

31

File Naming

- CRC2023_draft.docx
- CRC2023_draft2.docx
- CRC2023_draftupdated.docx
- CRC2023_final.docx
- CRC2023_finaldraft.docx
- CRC2023_finalFinal.docx
- CRC2023_lastone.docx

32

File Naming Tips

- Standardisation:
  - Order of elements
  - Punctuation
  - Keywords
  - Hyphen or underscore between words for compatibility with analysis software
- Consider how files will be sorted

Date created YYYYMMDD — Creator Initials

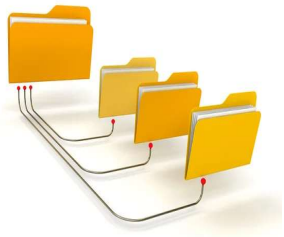
20230701_CRCdata_PN_JG.xlsx

Content — Project name

33

File Organisation Tips

- Always make a copy before edits
- Define and intuitive directory structure:
  - /projectA/data/masterfile
  - /projectA/data/analysis
  - /projectA/data/backups…
- Save in format that can be accessible in future:
  - .csv .xlsx .txt



34

Versioning

- Allows reverting to previous iteration if needed
- Helps understand evolution of data set
- Tips:
  - Decide how many versions to keep
  - Keep previous versions in same location
  - Versioning software:
    - Git (git-scm.com)
    - Mercurial SCM (mercurial-scm.org)

35

Versioning (2)

- Version control template
  - Records Management Advice - Version Control (www.wa.gov.au)
  - Archive digital content, University of Sydney Library

Version Control Table Template

Much of this information is better captured in a separate version control table, rather than as part of the file name. A version control table documents the edit history of your data.

Title				
Description				
Created By				
Date Created				
Maintained By				
Version Number	Modified By	Modifications Made	Date Modified	Status

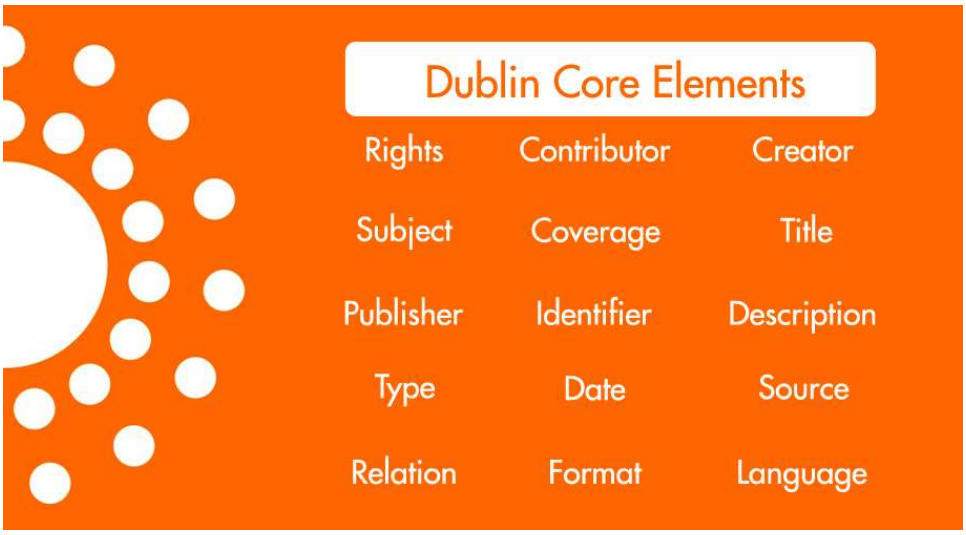
36

Metadata

- ✓Data about data
- ✓Varying structure
- ✓Footprint:
  - What would someone else need to find / re-use my data?
- ✓Store with dataset in a 'ReadMe' file or eNotebook

37

Metadata Schema



https://www.dublincore.org/resources/userguide/

38

Data Storage

- ❑Access
- ❑Storage and archival policy
- ❑Ease of Sharing
- ❑Actual storage location – cloud?
- ❑Documented tight backup system (automated)
- ❑Data repository

39

Retention period

- Standard:
  - Adult research: minimum 5 years
  - Paediatric: 5 years after last reference or until child turns 25 years
- Permanent:
  - Gene therapy research
  - Data of international or national significance
  - Studies that are impossible to repeat

Information Retention and Disposal Policy

40

Data Sharing & Publication

- ✓Included in research proposal & consent form
- ✓Data license
  - Outlines access to and use of data
- ✓Publish metadata along dataset
- ✓Persistent identifiers:
  - DOI (dataset)
  - ORCID (researchers)

creativecommons.org.au

http://www.dataarchive.ac.uk/media/2894/managingsharing.pdf

41

Database development

42

Database definitions

- ✓ Electronic filing systems for entry, storage and manipulation of data
- ✓ Data stored in tables:
  - Row - Record; Column – Variable
- ✓ Relational vs non-relational databases
- ✓ Maximize data integrity before analysis
- ✓ Allow multiple users – concurrent data entry

43

Database software

EpiData	Small, free portable program
EpiInfo	Small, free, does analysis, easy to use
REDCap*	Free, stored locally, web access, online tutorials, standard forms, reports, front end checks and balances, multiple users
Medrio	Clinical trial data entry, good security, cloud-based
Webspirit	Clinical trials – available through PTNA
SPSS Data Entry	Stand-alone product, allows validation
Qualtrics	Online tool for creating surveys, Australia-based, paid option
iApply	Online forms such as surveys
+ Oracle, MySQL, Access, lots of others	

44

Data Dictionary

- ❑Variable names & descriptions
- ❑Validation and coding rules
  - Coding: categorical variables, missing data
  - Range of continuous data and dates
- ❑Table names & relationships
- ❑Relationships with other databases
- ❑Dynamic

45

Data Dictionary - Report Example

CLIN_OBS Table

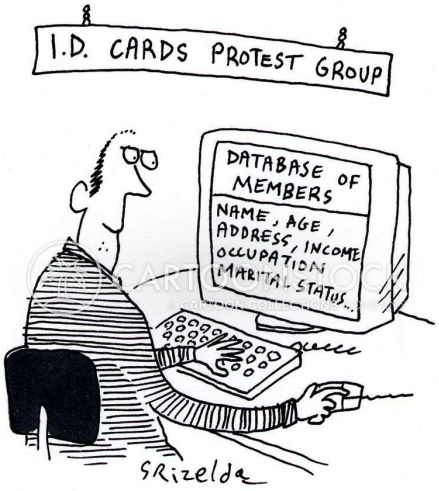
Clinical observations starting from when the child was admitted to ISOP (7B). Primary key: (study_id, obs_time_pt, obs_date). Relationships: Links to DEMOGRAPHIC table via study_id.

Variable name	Description	Data Type	Values/Rules
STUDY_ID	Study identification number issued to child from randomisation form. 1000's = Indigenous 26 wks & less; 2000's = non-Indig 26 wks & less; 3000's = Indig > 26wks; 4000's = non-Indig > 26 wks	number	1000-4000
OBS_TIME_PT	Clinical observations time point	number	1=Baseline hosp obs
			2=Enrolment obs
			3=12 hourly obs
OBS_DATE	date and time (24hr) of this clinical obs	date	
TEMP	Temperature (deg C)	number	25-45
PULSE	Pulse rate (beats per minute)	number	50-250
RESP	Respiratory rate (breaths per min)	number	20-120
OXY	Supplemental Oxygen (L/min)	number	0-10
RA_SAT	Oxygen saturation on room air (%)	number	60-100

46

Unique Identifiers

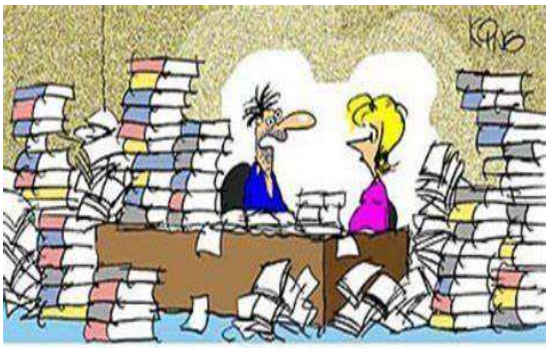
- ❑Record - specific
- ❑Automated or manual
- ❑Not hospital record number
- ❑Links data between database tables



47

Variable Naming

- ❑Unique
- ❑Short but informative
- ❑No ‘space’ or punctuation marks
- ❑Be compatible with data entry & statistical packages



"18 hours going through 350 baby name books, and you decide on the name Bob?"

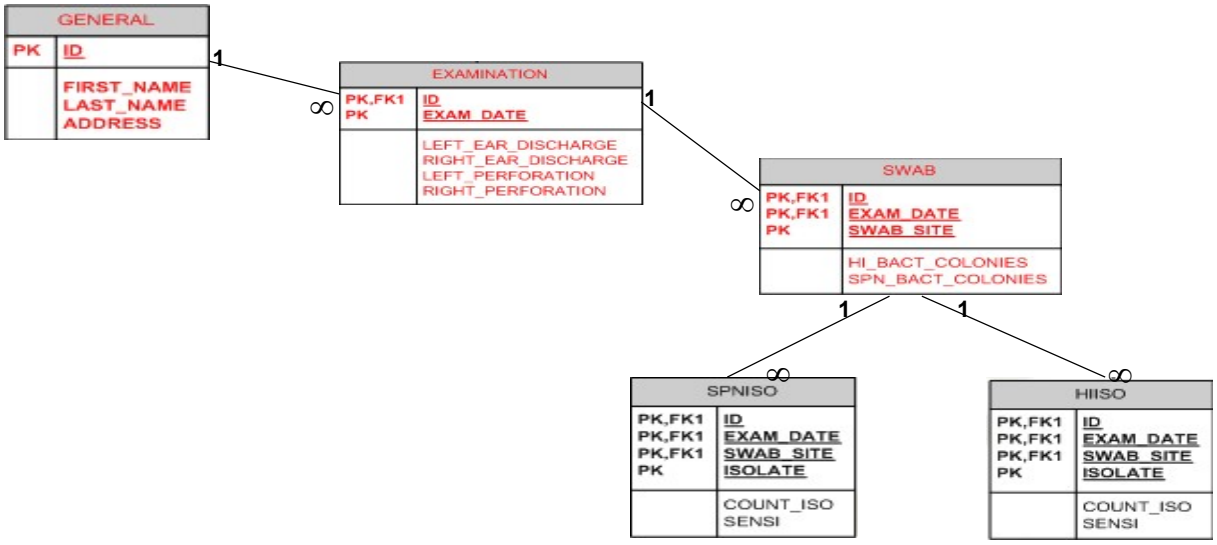
48

Selecting Data Types for the Database

- ❑Numerical vs Categorical
- ❑Validate formats where applicable
  - DD/MM/YYYY
- ❑Utilize calculated fields
  - Age; Time
- ❑Code for missing data

49

Types of Databases: Relational Database



50

Types of Databases: Non-relational Database

ID	NAME	DOB	AGE	EXAM_DATE	BP
1	Joe Bloggs	1/03/1987	20	12/05/2007	120/80
2	Smith, Jane	12/05/1998		12/05/2007	110/70
1	Joe Blogs	1/03/1978	21	1/06/2008	130/90

ID	EXAM_DATE	MEDICATION	DOSE
1	12/05/2007	Penecillin	400 mg daily for 7 days
2	12/05/2007	Amoxycillin	350
1	1/06/2008	Penicilin	200 mg twice a day for 7 days

51

Minimising Errors in the Database

- Checks:
  - Compulsory fields
  - Data validation
- Skips:
  - Branching logic

52

Test the Database

1. Ensure structure and integrity checks work as planned
2. Try to "crash" the database
3. Verify nothing unexpected
4. *ensure error messages/ prompts if incorrect data entered

53

Role of Database Manager

- ❑Fine tune data collection form
- ❑Design database
- ❑Create data dictionary
- ❑Database queries and reports
- ❑Manage user access rights
- ❑Advice on data cleaning

54

Acknowledgements

A/Prof Sue Skull

- Providing some PowerPoint slides

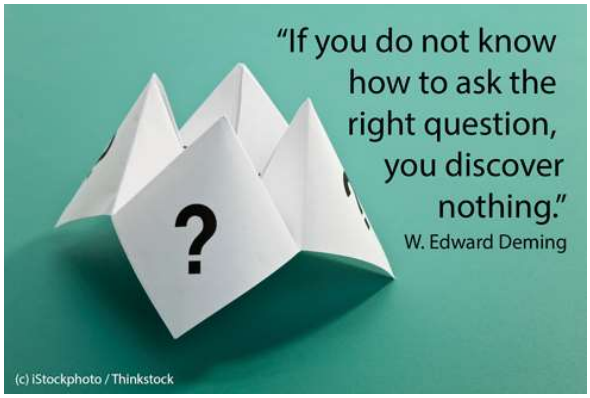
55

Questions?

Comments?

56

Take Home Message



57

Upcoming Research Skills Seminars

4 Aug Rapid Critical Appraisal of Scientific Literature
Dr Natalie Strobel, ECU

18 Aug Media & Communications in Research
Keryn McKinnon, Telethon Kids Institute

Register → **researcheducationprogram.eventbrite.com.au**

We love feedback

A survey is included in the back of your handout, or complete online
https://tinyurl.com/datacollectionevaluation

ResearchEducationProgram@health.wa.gov.au cahs.health.wa.gov.au/ResearchEducationProgram

58

Government of Western Australia
Child and Adolescent Health Service

TELETHON KIDS INSTITUTE
Discover. Prevent. Cure.

2023 Child Health Research Symposium

Celebrating Innovation, Collaboration and Translation

8 – 10 November 2023 Perth Children's Hospital

Abstract submissions close 13 August 5pm (AWST) https://tinyurl.com/CHRsymposiumabstracts

Neonatology | Community Health | Mental Health | Perth Children's Hospital

59

✉ ResearchEducationProgram@health.wa.gov.au
cahs.health.wa.gov.au/ResearchEducationProgram

60

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

Data Collection and Management

RESOURCE NOTES

CAHS Research Education Program Research Skills Seminar Series
✉ ResearchEducationProgram@health.wa.gov.au
cahs.health.wa.gov.au/ResearchEducationProgram

Table of Contents

1. Why we need good data management

The whole point of conducting research is to obtain high quality data that can have impact. This is an ethical issue, as poor quality, corrupted or lost data can mean not answering the research question, or falsely influencing policy and practice – in other words, potentially wasting participants and your time and resources, and affecting your reputation as well as that of your institution.

Good data management is a foundation of good research. It should be planned from the beginning of a research project and become part of standard research practice. Unfortunately, data management is often done at the last minute, using the first method that comes to mind. This approach is generally more time-consuming and error-prone. Taking time at the start of a research project to put in place robust, easy-to-use data management procedures will usually pay off several times over in the later stages of the project. If data are properly organised, preserved and well documented, and their accuracy, validity and integrity is controlled at all times, the result is high quality data, efficient research, outputs based on solid evidence and the saving of time and resources. By contrast, inadequate data management can lead to catastrophes like the loss of data or the violation of people's privacy. Some journals also require data to be entered to a data repository for open access, or availability of data on request.

In short, research data should be accurate, complete, identifiable, securely stored, retrievable and able to be made available to others. Good management of data results in high quality data able to provide a meaningful and interpretable result for a research question.

2. What are the steps to achieving good data management?

Data management includes all activities associated with data other than the direct use of the data. Steps to be considered are:

- Analysis planning – based on a clear, answerable question and objectives
- Data base design and development
- Database testing
- Developing a data management plan
- Data file management
- Coding sheets and Data dictionaries
- Data collection
- Data entry
- Data validation
- Data cleaning
- Security
- Data sharing and collaboration

- Archiving

2.1. Research Data Management - Good Sources of Open Access Learning Materials

- **Australian Research Data Commons (ARDC)**
  Webinar – Data Management in the revised National Statement on Ethical Conduct in Human Research - 5 September 2018 (189) Data management in the NHMRC's revised National Statement on Ethical Conduct in Human Research - YouTube

- **Curtin University Library - Research Data Management**
  http://libguides.library.curtin.edu.au/research-data-management

- **Research Data Services**
  Research Data Australia

- **Australian National Data Service (ANDS) Data Management**
  http://www.ands.org.au/working-with-data/data-management

- **Edith Cowan University**
  - Research Journey for Research Students – Data Management
    http://intranet.ecu.edu.au/research/for-research-students/research-journey/designing-and-undertaking-your-research/data-management
  - Library Guides – Manage Research Data
    http://ecu.au.libguides.com/research-data-management

- **Coursera online courses**
  https://www.coursera.org/courses?query=data%20management
  - Data Management for Clinical Research
    https://www.coursera.org/learn/clinical-data-management

- **Zenodo Research data management (RDM) open training materials**
  https://zenodo.org/communities/dcc-rdm-training-materials/?page=1&size=20

- **Digital Curation Centre**
  http://www.dcc.ac.uk/

- **UK Data Archive**
  http://www.data-archive.ac.uk

- **UWA Research Data Management Toolkit**
  Welcome - Research Data Management Toolkit - Guides at University of Western Australia (uwa.edu.au)

3. Data management responsibilities

Funding bodies and governments increasingly require sound data management. Researchers have a responsibility to make themselves aware of any relevant codes and to comply with them. Failure to comply with funding body requirements (eg In Australia: ARC or NHMRC – increasingly required in grant applications – so PLAN, get input from IT, data services etc) may jeopardise future research funding. Failure to comply with legal requirements, such as those that safeguard the privacy of participants in medical research, may lead to prosecution.

Good Clinical Practice training provides information on many aspects of standards required for data collection and management in clinical research.

"An introduction to GCP" one hour overview is available via the Research Education Program. Please submit the short access survey form to access the recording.

Training options include:

- **Global Health Trials**
  https://globalhealthtrials.tghn.org/elearning/

- **ARCS Australia**
  https://www.arcs.com.au/events/online-training/

- **Research Education & Training Program (RETProgram) WAHTN**
  Online Training - RETProgram

The state public sector in Western Australia does not currently have a legislative privacy regime. Various confidentiality provisions cover government agencies and some of the privacy principles are provided for in the Freedom of Information Act 1992 (WA).

http://www.austlii.edu.au/au/legis/wa/consol_act/foia1992222/

3.1 Important documents/sites

- WA Health Research. Governance Policy and Procedures Handbook.
  Research Governance Policy (health.wa.gov.au)

- WA Health Patient Information Retention and Disposal Schedule (PIRDS) 2016
  Information Management (health.wa.gov.au)

- Research Data Management Toolkit. University of Western Australia.
  Includes Planning, Intellectual Property, Documentation, Storage/Backup, Sharing/Reuse, Retention/Disposal, Support/Contacts/Useful Resources.
  https://guides.library.uwa.edu.au/RDMtoolkit

- The Telethon Kids Institute has policies for Confidentiality of Research Data, Information Retention and Disposal, Archiving, and Information Security and Handling Procedures. Code of Conduct for RPR Part A (telethonkids.org.au) Code of Conduct RPR Part B Handling of allegations of research misconduct (telethonkids.org.au)

- Souhami R. 2006. Governance of research that uses identifiable personal data. BMJ. http://www.bmj.com/content/333/7563/315

- Australian Code for the Responsible Conduct of Research (2018) https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018

- WA Health and Institutional policies – Research Governance Framework https://rgs.health.wa.gov.au/Pages/Research-Governance-Framework.aspx

- UWA Code of Conduct for the Responsible Practice of Research. University of Western Australia https://www.uwa.edu.au/policy/-/media/Project/UWA/UWA/Policy-Library/Policy/Code-of-Conduct/Integrity/Research-Integrity/Research-Integrity-Policy.doc

- Australian Clinical Trials Handbook https://www.tga.gov.au/publication/australian-clinical-trial-handbook

- Information Management Governance Policy 0152/21 https://ww2.health.wa.gov.au/~/media/Corp/Policy-Frameworks/Information-Management/Information-Management-Governance-Policy/Information-Management-Governance-Policy.pdf

- WA Health Information Storage and Disposal Policy 0144/20 https://ww2.health.wa.gov.au/~/media/Corp/Policy-Frameworks/Information-Management/Information-Retention-and-Disposal-Policy/Information-Retention-and-Disposal-Policy.pdf

- WA Health Portable Computer and Storage Devices Policy 0145/20 https://ww2.health.wa.gov.au/~/media/Corp/Policy-Frameworks/Information-Management/Information-Storage-Policy/Information-Storage-Policy.pdf

- WA Health Hospital Morbidity Data System. HMDS Reference Manual July 2014. https://ww2.health.wa.gov.au/-/media/Files/Corporate/general-documents/Clinical-Information-Assurance/Part-A-HMDS-Ref-Manual-2018-19.pdf

4. Data management plans

Documentation of project information in a Data Management Plan is an invaluable resource to later members of the project team as well as other researchers that come to investigate your project and its data after project completion, or audit personnel. Increasingly these are required for grant applications. Get input from IT, data services etc.

A data management plan needs to cover:

1. Survey of existing data: What existing data will need to be managed?
2. Data to be created: What new data will your project create?
3. Data owners & stakeholders: Who will own the data created, and who would be interested in it?
4. File formats: What file formats will you use for your data?
5. Metadata: What metadata will you keep? What format or standard will you follow?
6. Access & security: Who will have access to your data? If the data is sensitive, how will you protect it from unauthorised access?
7. Data organisation: How will you name your data files? How will you organise your data into folders? How will you manage transfers and synchronisation of data between different machines? How will you manage collaborative writing with your colleagues? How will you keep track of the different versions of your data files and documents?
8. Storage: Where will your data be stored? Who will pay? Who will manage it?
9. Backups: Hard drives on desktop and laptop computers fail regularly. You must follow a credible backup strategy of regular backups. Consider including an off-site backup so that your data will not be lost in a “worst case scenario” eg your building burns down. Rather than relying on memory, consider automated backup.
10. Bibliography management: What bibliography management tools will you use? How will you share references with the other members of your group?
11. Data sharing, publishing and archiving: What data will you share with others? Who will be allowed to have future access or re-use data? How will you organise this?
12. Retention and disposal procedures and provisions: What data will you destroy? When? How?
13. Responsibilities: Who will be responsible for each of the items in this plan?
14. Budget: What will this plan cost? Possible costs include those for backups, research assistant time for data curation, metadata creation, archiving, data manager time etc.
15. Ownership and protection of intellectual property
16. Anything else: Don't restrict yourself to the items above. Stop and think. Is there anything missing?

5. Good file management practices

Standardisation creates uniformity within/across projects and allows for easy retrieval of files because everyone knows where to find them, as well as version control. Covered in more detail by Good Clinical Practice (GCP) training.

A few basics:

- Maintain a single master copy of your data file.
- Always make a copy of the data file with a new name before any data entry or edits are made.
- Name your copies using the current date e.g. MyDataFile_2018-09-28.rec. This y-m-d format helps you sort files in date order.
- Define an appropriate directory structure that makes clear what the purpose is of any files therein: e.g.
  /myproject
  /myproject/data
  /myproject/data/backups

6. Confidentiality

- Keep the database and/or your computer password protected.
- **Use codes to de-identify the data**, i.e. id numbers (never an "identifiable" number such as a hospital record number)
- Keep the participant's name and contact details in a separate file to their survey results. For longitudinal studies, details such as contact details of family/friends, withdrawals should also be "stripped" from the main database and kept separately
- Ideally use a data capture program like REDCap which enables
  - de-identification of identifying variables (e.g. Name, date of birth, hospital record number, contact details), and enforces allocation of a de-identied "unique identifier – that can be used to link back to identifying details if required
  - database users to be ascribed user rights at different levels – i.e. only certain people will have access to data, and the level of access can be controlled
- Ensure hard copy data is kept in locked filing cabinet (or equivalent).
- In WA Health, de-identified data may be moved on a Dept of Health encrypted USB but NOT stored.
- Have a standard operating procedure to ensure all project staff understand their responsibilities around keeping data confidentially.
- Ensure you are familiar with your institution/state/national policies on data confidentiality – this includes movement of and access to data
- See websites earlier in this handout for further information and policy.

7. Data sharing / collaborative data entry work

- REDCap is a good example of data capture software enabling multiple users, with the ability to keep certain areas accessible only to certain people, and multiple data entry persons.
- If using a server-based database for multiple users, ensure only the latest version can be accessed
- If you don't have access to a joint access database, concurrent data entry may be best managed by always entering data into a blank dataset, then having one person responsible for merging the results
- If a single-user database, each user can be given a separate copy of the database with its own name into which they enter all their data. The data in these separate copies can later be appended. Note that this remains less optimal to a multi-user data entry system like REDCAP as appending will not prevent a duplicate record from being entered – e.g. one by a user into their copy of the database and another by a second user into their copy.
- UK Data Archives - excellent Guide To Managing and Sharing Data May 2011 http://www.data-archive.ac.uk/media/2894/managingsharing.pdf
- UWA Research Data Sharing/Re-use (Part of the Research Data Management Tool Kit) Sharing/collaborating - Research Data Management Toolkit - Guides at University of Western Australia (uwa.edu.au)

8. Data archiving / storage and data destruction

At the end of the project create an archive of:

- All data, electronic and paper
- All cleaning and analysis command files
- All related documentation
- Date when the archive was created and know when or if it can be destroyed
- In WA, minimum storage times are 5 years after all reference to the documents has ceased (data or related research documents) AND for interventional studies involving children, until the child reaches age 25 years. Clinical trials data are generally kept for 15 years.
- Hard copies can be archived off site as long as they are retrievable for the required period and are not required frequently (e.g. not more than once per year)

No destruction of records, information or data can be conducted unless it is in accordance with an approved disposal authority;

The approved disposal authorities which should be used for WA Health staff are contained within

- The Patient Information Schedule RD 2007079;
- Administrative and Functional Records RD 2005001 (available from the Department of Health) and
- The General Disposal Authority for State Information (available from our website – www.sro.wa.gov.au);

All information and data able to be legally destroyed now, must be documented – that is all records, information and data which can legally be destroyed must be listed with proof of the authorised disposal authority under which it is being destroyed.

At a minimum, this information must include:

- file or document group or subject;
- date range of same;
- identification of the disposal authority – name and number;
- evidence of sign-off by CEO;
- date of destruction; and
- method of destruction.

For the destruction of data and electronic records, the metadata must be kept which identifies the items listed above.

As the State Archive is full, the Department of Health advises suitable alternative storage via :

WA Dept of Health: Data Storage and Disposal Policy 2020
https://ww2.health.wa.gov.au/~/media/Corp/Policy-Frameworks/Information-Management/Information-Storage-Policy/Information-Storage-Policy.pdf

Note there are two main types of data to be considered: type 1 below should be retained, type 2 can eventually be destroyed after following the appropriate timing recommendations. If in doubt, keep it.

8.1. Research administrative and functional records (approval, monitoring, publications etc)

Retain

- Records associated with the activities involved in investigating or enquiring into a subject or area of interest in order to discover facts, principles etc. e.g. Ethics and governance documentation

8.2. Patient information (data, consent etc)

Eventual Destruction

- Patient or Subject Research Records - Refers to records (including laboratory results, reports, questionnaires and surveys) obtained from consenting patients or subjects for the specific purpose of researching a project, theory or trial.
- Records of Consent or Authorisation - Refers to records of consent or authorisation for the use of patient's or subject's results in research.
- Research Requests - Refers to documented requests to perform research.

8.3. Preparation for storage - courtesy of State Records Office

- Group the records into their proper series (SRO can assist here if you're unsure).
- Box the records sequentially by their numeric, alphabetic or alpha-numeric numbering system.
- List the records in full on consignment/manifest lists in the same order in which they are boxed (i.e. sequentially).
- Include a copy of the consignment list in the box and record details in your organization's Recordkeeping system.
- Ensure boxes are numbered sequentially and the range of items for each box is noted on the end of the box (i.e. the first and last record number for each box).
- De-metalling (removal of staples, paper clips, metal fasteners) is not recommended when preparing archives for transfer

Dept of Health WA recommends archival records awaiting transfer to the SRO/recommended facility be boxed by the custodial organisation in approved archive boxes. Standard archive boxes are made from acid free board (neutral pH and buffered).

The standard box, suitable for most files, measures 385mm x 250mm x 168mm (internal measurements when box is made up). These boxes are commonly referred to as Type 1 archive boxes.

To determine how many boxes will be needed, measure the length of the files as they would stand on a shelf and divide by 0.16m. Other size boxes are also available to accommodate smaller or larger items. If not available at the Health Dept., archival quality boxes can be ordered from suppliers.

Some agencies have needed to utilise commercial storage providers on the current Common Use Arrangement:

https://www.contractswa.finance.wa.gov.au/group.jsp?groupID=OF&STMP=140520150248112#17

9. Data collection and analysis planning

9.1. Basics of setting up databases

What is a database?

- A computer software program that facilitates the entry, storage and manipulation of data
- The data is stored in a table and the file is called a data file
- Each row represents a case/record
- Each column represents a field/variable
- One or more of these tables stored together create a database.
- A relational database has multiple tables that are related to each other. Usually they are linked together by unique identifiers.

9.2. Database software

REDCap – the preferred data capture program for WA Health and many leading institutions

- Data stored locally, accessed via web
- FREE – acknowledge only
- Online tutorials available
- Standard forms provide assistance with set up
- Front end checks and balances available to improve accuracy of data entry
- See extensive information on the Research Education Program open access website
- Seminar 1h overview
  CAHS - Seminar schedule (health.wa.gov.au)
- REDCap Workshops and associated handouts
  CAHS - REDCap resources (health.wa.gov.au)
- General access and help information including other instructional videos:
  CAHS - Additional Resources (health.wa.gov.au)

EpiData

- Small free portable program (www.epidata.dk)
- Data entry forms can be set up to resemble paper questionnaire

SPSS Data Entry

Stand-alone product that will allow validation checks

Medrio

- for clinical trial data entry – cloud-based, secure

Webspirit

- For clinical trial data – available through Paediatric Trial Network Australia

Qualtrics

- Paid online tool for creating surveys (www.qualtrics.com)
- used by a number of institutions in WA eg Curtin, UWA, the Raine Study

Microsoft Access

- Has limitations – security issues, potential to overwrite data etc
- Can deal with complex relational databases
- Data entry forms can be set up to look like your paper questionnaire
- Able to support multi-user access to database
- Customised reporting available

Microsoft Excel - users BEWARE

- NOT recommended. VERY EASY TO DESTROY YOUR DATA IRREVOCABLY
- Comprehensive data checking will be required once all the data have been entered
- Unable to enforce uniqueness for an identifier
- Not a relational database
- Only one person can access a file at any one time.
- Need to be careful with dates

Survey Monkey - users BEWARE

- NOT recommended. Preferably don't use, or use for non-sensitive data with caution
- You don't own the data
- Unclear where data are
- You may be breaching Dept of Health policy
- Never use for sensitive, identifiable or re-identifiable health data
- **Not recommended by many sites** including Telethon Kids Institute and Perth Children's Hospital for research data.
- **Evolving policy to watch for at WA Health** – not recommended for patient or staff data, particularly if potentially sensitive

9.3. Variables, coding sheets and data dictionaries

Before creating a database, you need to design its structure. This is done by creating a data dictionary or coding manual in conjunction with the data collection forms.

A data dictionary includes:

- Table names
- Variable names
- Variable descriptions
  - meaning, data type, units of measure

- Validation/coding rules
  - code for categorical variables
  - ranges for continuous data and dates
  - codes for missing data

- Relationships between tables

Naming Variables

- Each variable must have a unique name
- Choose an informative name – something anyone could understand
- Do not use spaces, special characters or punctuation marks
- Begin variable names with a letter

Select variable names that are compatible with both your data entry and statistical packages.

Check:

- The maximum number of characters allowed for a variable name
- The characters that can be used to make up a variable name
- Is the package case sensitive?
- Does the package have any special names that can't be used for variable names?

Examples of statistical package variable naming issues

- Variable names that start with a number will be prefixed by an "n" in EpiData.
- The underscore character "_" is OK in Stata but not in EpiData.
- Lowercase is recommended if you will be analysing your data in Stata.
- In EpiData a variable name can be up to 10 characters long, contain letters or numbers, and must begin with a letter.
- REDCap variables may not start with a number, use underscores not spaces, and can contain lowercase letters or numbers
- If your variable names contain numbers only, Stata will name the variables v#.
- If your variable names start with a number, the number will be stripped in Stata.
- MySQL and Oracle databases will not accept variable names starting with a number.

Selecting data types

- Data are generally numeric, text or date
- Use numeric variables for continuous data such as height and weight or for variables on which you wish to perform mathematical operations.
- Use numeric codes to hold categorical data wherever possible – easier for data entry plus many statistical procedures will only work with data stored in numeric format.
- If using string codes (free text), be consistent with spelling and upper vs. lower case text.
- For date variables select a date format rather than entering dates as strings or entering day, month and year as separate variables (otherwise very hard to work with).

Numeric Data

<u>Categorical</u>

- Nominal

  - no ordering implied
  - binary if only 2 categories
  - no inherent meaning e.g.: 1=single; 2=married; 3=defacto; 4=divorced/separated

- Ordinal

  - categories assume a natural ordering
  - codes convey the order e.g: 1=poor; 2=fair; 3=good; 4=very good; 5=excellent

- Continuous

  - e.g. height, weight, waist
  - – can only take specific values (visit #: 1 2 3 4…)

String/text

- can be categorical, e.g. never; sometimes; always
- you will need to convert string variables to numeric formats in your statistical package prior to doing most statistical procedures, e.g. never=1, sometimes = 2, always = 3
- generally avoid wherever possible

Date

- Use of a date format in database packages such as REDCap, EpiData and Access will ensure that only valid dates will be accepted.

Continuous and discrete variables

Continuous and discrete variables contain numbers only, for example height or weight (continuous) or number of people in a household (discrete). You will often have limits or ranges for continuous variable. For example, an Apgar score must be an integer between 0 and 10. When setting up a database you can specify these limits to reduce the chance of data entry error. Many measurements come in units such as days, grams or millilitres. Decide on the most appropriate unit (grams or kilograms? minutes or hours?) and specify this unit on the coding sheet, questionnaire and data entry form. All entries for one variable must use the same unit. Don't mix up (for example) grams and kilograms or months and years. Setting ranges for variables helps prevent this.

String variables

String variables can contain text and numbers. As it is difficult to perform mathematical or statistical operations on strings, they should not be used if a numeric variable can be used instead. You might, for instance, have a question that asks for community. The paper questionnaire can have a space for the community to be filled in, but before the questionnaires are entered in the database, you can give each community a numeric code which is entered in the database. String variables usually contain data that is too complicated to categorise, such as long comments. Many programs have limitations on the length of string variables.

Date variables

Date formats vary around the world. Both day-month-year and year-month-day are logical date formats that are easily understood. Use one date format that will be understood by everybody working with your data, and use it consistently throughout the database.

Unique identifiers

- Each record or case must have a unique identifier
- Unique identifiers are used to refer to specific records in the database *without using identifying information* such as names. They are also used to link data between tables in the database. The unique identifier for a participant is usually a number assigned by the researcher.
- The unique identifier can be a single field, such as studyid, or a combination of two or more fields, e.g. studyid and visitdate.
- Hospital record number should not be used as the unique identifier.
- If using REDCap, the first variable will automatically be set up as the unique identifier, and will generally be consecutive numbers.

Missing data

- Include a code for missing data
- For numeric variables, missing data is conventionally represented by “9” for one-digit variables, “99” for two-digit fields, and so on. The missing code must never be a valid response for that variable. For dates, 9/9/9999 can often be used as the missing code.
- Some fields, such as id numbers and eligibility criteria, should never be missing.
- Change missing codes to missing values in your statistical package before you begin your analysis.

Using special codes for missing data allows data entry personnel to indicate a blank on the data collection form as opposed to data that has been accidentally skipped during data entry. Fields with missing data can be collated later on and efforts made to retrieve missing data where possible. All fields should have data in them unless they are skipped because of a conditional jump. For example, if a participant says no to “Have you been diagnosed as diabetic?”, then they shouldn’t have data for the question “Date of diabetes diagnosis”.

*The missing code must never be a valid response for that variable, so if someone’s weight could be 99, then “999”, and not “99”, should be used as a missing code. Alternatively to maintain consistency you may prefer to nominate a large number, say 9999, as the missing code for all your numeric variables.

Data Entry Queries

- Include a code for data queries encountered during data entry
  For example, if a response written on a data collection form is illegible, unlikely, unclear or inconsistent with other responses then the data entry operator can enter a query code. This indicates something that needs to be followed up.
- Use a value that is not a valid response for the field to indicate a query (e.g. 77 or 7/7/7777 for dates).
- Unique identifier fields must have a valid value and therefore should not have a query code.
- All queries should be dealt with before beginning analysis.

Examples:
- Date of birth is not feasible
- Date field is outside the date range for the project.
- Male respondent answers yes to "Have you been pregnant?"

Consistent dates and times

Example of data downloaded from hospital patient admission system. The date of admission and date of discharge were exported from the hospital system formatted as mm/dd/yyyy and emailed to the project coordinator. Not realising the formatting issue, the data was imported in to the project database which used dates formatted as dd/mm/yyyy. The problem arose when only the dates with days greater than 12 were transposed to dd/mm/yyyy format and the rest were unchanged. This of course meant that many of the dates were incorrect. In this case the data had to be requested again specifying the dates to be formatted as dd/mm/yyyy.

Enter raw data rather than summary data

e.g. when measuring heights of 2-year olds you may take three measurements and use the average of the three measurements in analysis. Still enter all three measurements and allow your stats package to calculate the average. Doing this also allows you to check how discrepant the three measurements are. If one is very different from the other two you may want to calculate the average using the other two measurements only. Entering average height: you're more likely to make an error calculating the average than your stats package will, and it's more difficult to check the computerized records against the paper records.

Do not confuse data entry and calculations.

Calculations can be done later in your statistical package. The exception would be where the results of a calculation would inform your next steps, e.g. how to treat a patient, determining what other data needs to be collected, etc. Sometimes calculated fields are included on the data entry form for data validation but these fields are not saved to the database.

Identify sections within a data collection form with repeating information collected on different subjects or at different time points.

For instance, assume we're collecting data on pregnant women and the babies they deliver. A pregnancy can result in the birth of more than one baby. In the database how many babies should we allow for? Singletons, twins, triplets, quads? We could create one file with enough space to allow for quads but this would result in a very wide file with many empty variables. An alternative is to enter the data in two separate data files, one for pregnancy and another for babies. The baby file should include the pregnancy id number so that each baby can be linked to the pregnancy. There would be a separate table for each mother and a mother may have several pregnancies over the course of the data collection phase.

Badly Designed Table – Example 1

ID	NAME	DOB	AGE	EXAM_DATE	BP
1	Joe Bloggs	1/03/1987	20	12/05/2007	120/80
2	Smith, Jane	12/05/1998		12/05/2007	110/70
1	Joe Blogs	1/03/1978	21	1/06/2008	130/90

- Table holds data on more than one subject - demographics and examinations
- NAME and BP fields are not atomic – break into FIRST_NAME and SURNAME fields, SYS_BP and DIAS_BP
- AGE field has not been calculated by the computer program and is more likely to result in errors.

Calculated fields are not stored in the database because they can lead to inconsistencies, for example if the DOB field was updated then the researcher must remember to also update the AGE field. The AGE field could be added on the data entry form and calculated from the DOB and EXAM_DATE fields. This would allow the data entry person to check that the participant is eligible for the study or highlight data entry errors with the date fields.

Badly Designed Table Example 2

ID	EXAM_DATE	MEDICATION	DOSE
1	12/05/2007	Penecillin	400 mg daily for 7 days
2	12/05/2007	Amoxycillin	350
1	1/06/2008	Penicilin	200 mg twice a day for 7 days

- MEDICATION has repetition of text – a drop down list or code set is indicated here
- DOSE field is not atomic – free text field impossible to analyse

An example of how a medication table might be designed is given below.

Variable Name	Description	Data Type	Values	
ID	Unique identification number (Risk Factor Study ID)	number		
DATE_PRES	Date drug prescribed	date		
DRUG_NAME	Antibiotic name	number	1	PROCAINE PENICILLIN
			2	AZITHROMYCIN
			3	AMOXYCILLIN 250MG/5ML
			4	METRONIDAZOLE
			5	BENZATHIINE PENICILIN
			7	LA BICILLIN
			8	SOFRADEX
			9	TRIMETHOPRIM
DOSE	Dose of prescribed antibiotics, amount/dose	number		
DOSE_UNIT	Unit of dose (mls/mgs/units/drops)	number	1	mls
			2	mgs
			3	units
			4	drops
			5	topical
			9	not determined/missing
FREQ	Frequency of dose	number		
FREQ_UNIT	Unit of frequency of dose (daily/weekly/ongoing)	number	1	Daily
			2	Weekly
			3	Ongoing
			9	not determined/missing
TOT_DAYS	Total number of days medication given	number		

Variable Definition Examples

Variable Definition Example 1

Studyid

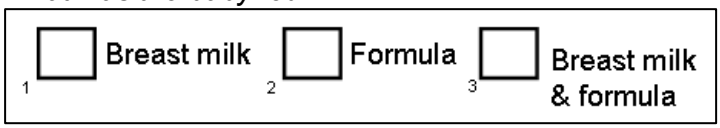
- Essential to have a unique identifier for each record
- Id numbers should be written on every form
- Specify the range of id numbers on the coding sheet
- There should be no missing values

Variable name	Description	Data type	Values/Rules
Studyid	Participant's unique study id number	Number	Must be unique 1001-2000

Variable Definition Example 2

Oralfeed

What was the baby fed?

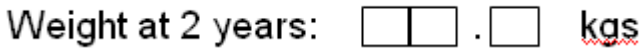


= nominal, categorical variable

Variable name	Description	Data type	Values/Rules
Oralfeed	Oral feeds	Number	1 = breast milk 2 = formula 3 = breast milk & formula 7 = query 9 = missing

Variable Definition Example 3

weight2



- Continuous
- Include unit of measure
- Specify plausible range
- For missing and query codes, use numbers that are outside the possible range of values for this variable

Variable name	**Description**	**Data type**	**Values/Rules**
weight2	Weight at 2 years (kg)	number	< 20 77 = query 99 = missing

Data Dictionary Variable Example - Hospital observation data

ADMISSION OBSERVATIONS

	DATE	Time (hrs)	Temp °C	Heart rate	Resp rate	O_2 Flow (L/min)	SaO_2 on room air
Baseline hospital observations							
Enrolment observations							

WARD OBSERVATIONS UNTIL DISCHARGE

TIME 0 IS THE TIME CLOSEST TO ENROLMENT
PLEASE RECORD OBSERVATIONS EVERY 12 HOURS FROM ENROLMENT ONWARDS UNTIL THE RESPIRATORY EPISODE ENDPOINT (i.e. 16 HOURS OFF OXYGEN, FEEDING ADEQUATELY ETC)

12 HOUR PERIOD	DATE	Time (hrs)	Temp °C	Heart rate	Resp rate	O_2 Flow (L/min)	SaO_2 on room air
1							
2							
3							
4							

This is an example of a data collection form for recording vital observations in hospital. Creating the data dictionary for the form will involve documenting each of the variables including any ranges and code sets.

Creating the data dictionary can be done using Microsoft Word, Excel or some other text file. Each record must relate to a participant who should be allocated a unique identification number. In this case we have created a variable called STUDY_ID for this purpose. The observation time point can be defined as a categorical numeric variable with a code for each observation type (1 = Baseline hospital observation, 2 = Enrolment observation, and 3 = 12 hour period observation).

The rest of the variables have ranges and unit measures as detailed overleaf.

CLIN_OBS Table Clinical observations starting from when the child was admitted to ISOP (7B). Primary key: (study_id, obs_time_pt, obs_date). Relationships: Links to DEMOGRAPHIC table via study_id.			
Variable name	**Description**	**Data Type**	**Values/Rules**
STUDY_ID	Study identification number issued to child from randomisation form. 1000's = Indigenous 26 weeks and less; 2000's = non-Indig 26 wks and less; 3000's = Indig > 26wks; 4000's = non-Indig > 26 weeks	number	1000-4000
OBS_TIME_PT	Clinical observations time point	number	1=Baseline hospital obs
			2=Enrolment obs
			3=12 hourly obs
OBS_DATE	date and time (24hr) of this clinical observation	date	
TEMP	Temperature (deg C)	number	25-45
PULSE	Pulse rate (beats per minute)	number	50-250
RESP	Respiratory rate (breaths per min)	number	20-120
OXY	Supplemental Oxygen (L/min)	number	0-10
RA_SAT	Oxygen saturation on room air (%)	number	60-100

Linking Tables

Tables which need to be linked must contain unique identifiers. It's good practice to give the linking variables the same name in all tables.

Relational Databases

Relational Database Rules (not covered in any detail in the seminar)

Relational database design should adhere to the following rules:

- Each database file should deal with only one project
- Each table should contain data relating to one topic/theme e.g. demographics, lab results, contacts
- Each table must have a primary key/unique identifier, field or fields that are unique for each record
- Each field must relate to the unique topic/theme of the table
- Fields should be atomic, hold just one piece of information
- No derived (calculated) fields in the table
- Repeating groups or fields in a table indicates the need for another table, a one to many relationship (e.g. Med1, med2, med3)

A database design which takes these rules into account allows the data to be easily extracted and manipulated later e.g. REDCap. Excel is NOT a relational database.

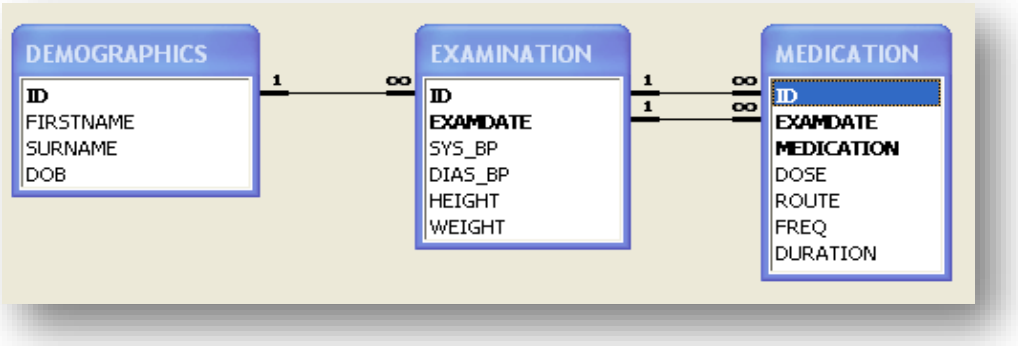
Relational Database rules explained

- Each table contains information about one subject. e.g. a “patient” table would contain data at the patient level (dob, gender etc) and a “visit” table would contain data collected at each visit.
- Add a new variable to one table only. Adding “DateOfBirth” to the visit table would result in “DateOfBirth” appearing as many times as there are visits which leads to redundancy and potential data discrepancies.
- Have one piece of information per cell e.g. split patient names into two variables, firstname and lastname.
- Set up all dates and times with the same format, e.g. dd/mm/yyyy, 12 or 24 hour clock.
- Enter raw data rather than summary data
- Repeating information collected on different subjects or at different time points should be stored in a separate table.

Relationship Diagram – Example 1

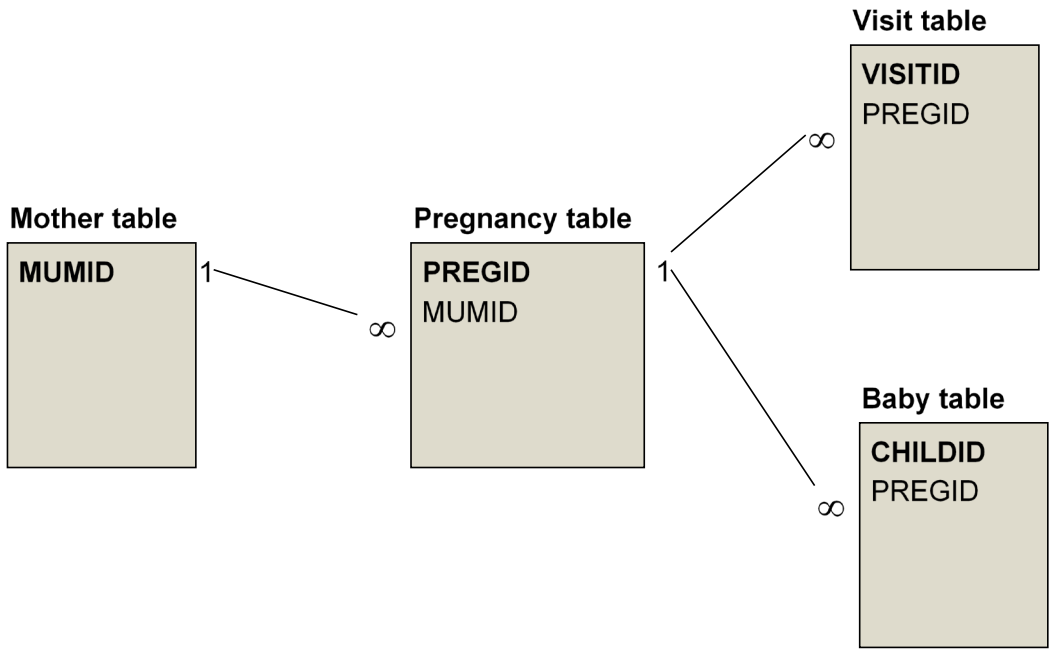
For complex studies, produce a relationship diagram prior to setting up the database and:

- include all the main tables
- indicate the linking fields
- indicate the type of relationship, i.e. one-to-one(1) or one-to-many(∞)



This relationship diagram shows how the previous badly designed table examples (1 and 2 above) could be designed and linked. Both relationships are one-to-many, the person can have more than one examination on different dates and on these dates the person can be given more than one medication.

Relationship Diagram – Example 2



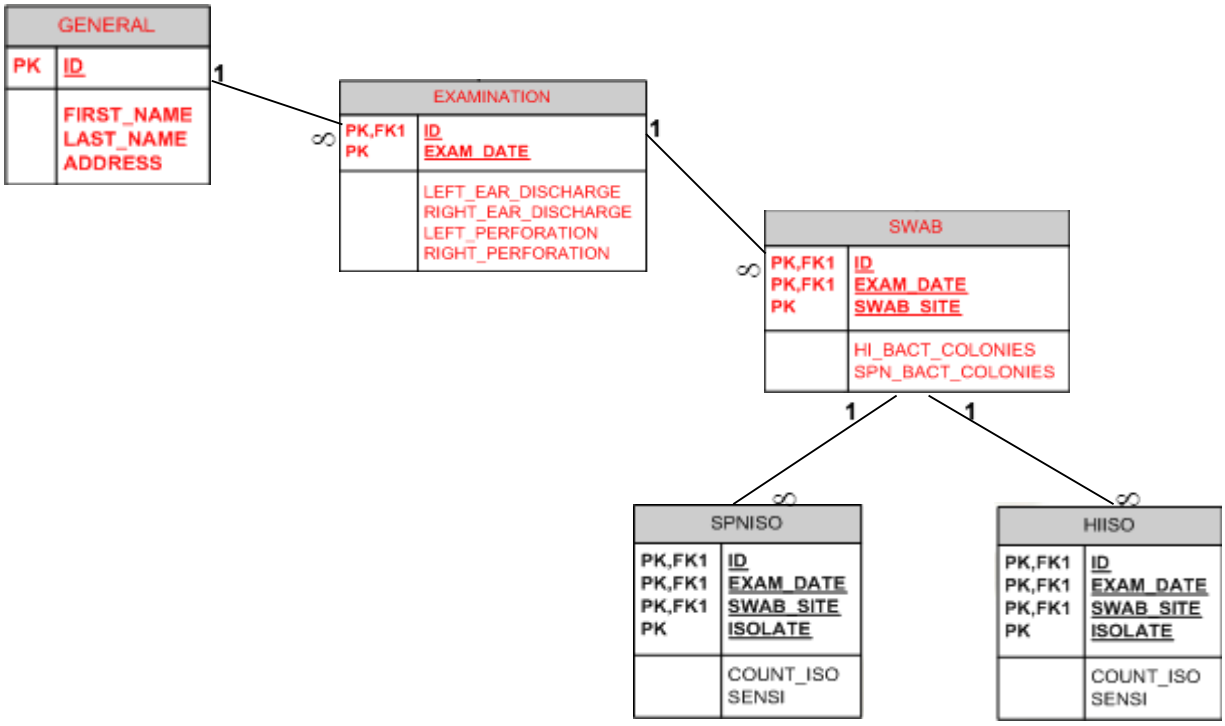
This is an example of how a many-to-many relationship could be represented. If we were collecting data on pregnant women and the babies they deliver, then the mother could have many pregnancies and each pregnancy could result in the birth of more than one baby.

In the database how many babies should we allow for? Singletons, twins, triplets, quads? We could create one table with enough variables to allow for quads but this would result in a very wide file with many empty variables. An alternative is to enter the data in two separate tables, one for pregnancy and another for babies. The baby table should include the pregnancy id number so that each baby can be linked to the pregnancy. There would be a separate table for the mothers and a mother could have several pregnancies over the course of the data collection phase. Therefore we create a junction table with the pregnancy id and the mother id which allows us to link a mother with a pregnancy and a child or children.

This example also shows a visit table where information could be stored on the hospital visits the mother had for each pregnancy.

The unique identifiers for these tables would be MUMID for the mother table, PREG_ID for the pregnancy table, CHILDID for the baby table and VISITID for the visit table. By adding the MUMID into the pregnancy table and PREGID into the visit and baby tables we can link all the information from all four tables.

Relationship Diagram – Example 3



In this example we have a series of one-to-many relationships. The unique identifiers or primary keys (PK) involve more than one field. A participant can have more than one examination, more than one swab can be taken at each examination and these swabs can have more than one isolate extracted from them.

10. Testing a database

The purpose of testing your database is to ensure that it has the structure and integrity checks that you expect. Try to "crash" the database to make sure these checks are working. Verify that the database does everything that you expect it to do, and nothing unexpected.

INSTRUCTIONS

- **Continuous fields**: Test the boundaries of each continuous field by entering minimum and maximum values (should succeed). Then try to enter values just outside the valid range (should fail).
- **Categorical fields**: Check that only valid responses can be entered for categorical fields. Typically a web interface will display only valid values, but EpiData and (sometimes) Access can allow entry of the raw numeric value code. E.g. gender:

1=Male, 2=Female, 8=Query, 9=Missing - ensure these are the only allowed values by testing outside the boundaries: 1, 2, 8, 9 are accepted - 0, 3, 4, 5, 6 and 7 are not.

- **Try entering a duplicate record - and ensure the attempt fails!** Typically this is with a single field, such as a unique study ID number per record in a participant dataset. However in some datasets you may use a combination of fields such as StudyID + Visit Date as a unique "key".
- **Test that skips are working properly**. Consider also what should happen if data is later amended, e.g. what happens to data recorded for pregnancy information when you edit gender from female to male?
- **Check that required fields cannot be left blank**. Note that in EpiData data entry should be done using the Enter, Tab or arrow keys to move from field to field. If using the mouse, jumps and "must enter" rules will be ignored.
- **Test warnings for values that are out of sequence**. e.g. dob should be < visit 1 date, visit 1 date should be < visit 2 date, but can also apply to numeric fields, e.g. systolic bp should be > diastolic bp. For most date fields a date after the current date should not be allowed. Note that "current date" can mean different things: date of questionnaire completion is rarely the same as the date of data entry.
- **Check the time difference between dates is valid**, e.g. if all visits occur between 1 and 2 years of age, check the minimum allowed difference between visit date and dob is 1 year and the maximum allowed difference is 2 years.

11. Data entry

11.1. Strategies for minimising errors

- Use a well designed questionnaire - clear, well presented (see previous Seminar: Survey Design and Techniques)
- Include codes on questionnaire
- Check questionnaires when returned
- Ensure the database fields follow the same sequence as the paper questionnaire
- Set up database to accept only valid responses
- Double enter data
- After data entry clean the entered data (10% check)

11.2. Validation (Database Design)

- Use of validation rules during data entry reduces time spent data cleaning. The following checks are common:
- Allow only certain values to be entered into a field
- Specify legal values for categorical data, e.g. 0, 1, 7 or 9.
- Specify a range for continuous data and dates. Remember the codes for missing/query.
- Program more complex checking procedures, e.g. consistency checks.
- Specify that some fields are compulsory
- Skip fields if particular values are entered
- Specify that id number must be unique

11.3. Double Data Entry

- The idea of double data entry is to identify discrepancies for correction. Options include:
  a) Entering the data twice into two separate files and then comparing the two files for differences.
  b) Prepare for duplicate entry. After the first file is completed, the second file is prepared based on a key field (the unique identifier) for the first file. While entering the second file, the value is checked for each field in each record against the same record of the first file. You are warned of any discordance so that you can ensure proper recording during the second entry process. This feature is available in EpiData.
- Double entry won't identify an error if the same error is made twice. The chances of this occurring may be reduced by a second data entry operator entering the duplicate data.

11.4. Data cleaning after database closure

- **Garbage in = Garbage out**
  Remember that the quality of the results you produce is directly related to the quality of your data.
- **Before analysing a set of data, it is important to check as far as possible that the data are correct.** Errors can be made at many points in the data collection process: when measurements are taken, when the data are originally recorded, when they are transcribed from the original source, or when being entered into a computer. We can't be sure about what is correct, but we can check if recorded values are plausible. This process is called data cleaning.
- **Fixing errors requires a data analysis program such as Stata or running queries in the database program**. After problems have been identified you should go back and compare them to the written collection forms. Some typing errors are obvious and just need to be corrected in the database. Other errors will require some interpretation. If it is not possible to decide on a meaningful response, these variables should be recorded as missing. Data can be corrected in the original database and re-exported for use in the statistical package. Any corrections made in the statistical package should be documented so that any repeated analysis will achieve the same result.
- **Checking and cleaning your data takes longer than you think**. Allow sufficient time for this stage when planning a study. Once the database has been closed, as decided by the principal investigator, any data corrections must be made in the statistical package.
- **DO NOT do any data cleaning interactively in your statistical package**. Cleaning must be documented and reproducible. Data must be cleaned via a command file (eg: a do-file in Stata, a syntax file in SPSS etc). Include comments in your data cleaning command file.
- **DO NOT overwrite original variables**. It is often necessary to re-code or modify original variables. It is good practice to assign the modified values to new variables and keep the original variables unchanged. The exception to this recommendation is replacing missing codes with missing values.

Prior to Database closure, cleaning might involve:

- Data checking throughout the data entry process – can minimize data entry errors due to interpretation differences between data entry personnel
- Double data entry - any discrepancies between the two copies are checked and corrected
- 10% hard copy checks – randomly select 10% of the records and two people then compare a print out of the selected electronic data with the paper collection forms

After database closure, cleaning might involve:

- Checking there are no query codes remaining
- Identifying blanks where there should be none
- Identifying implausible values
- Inconsistency checks - "Logic checks" – run queries on the data to pull out records that don't match certain rules defined by the project e.g. age criteria
- Replacing missing codes with missing values
- Attaching variable and value labels

Missing Data

- **Check for blanks**. It is advisable that codes for missing data are created prior to data entry. Therefore when data is entered, the only fields that should be left blank are those with a jump, e.g. If "No" skip next question. This makes it easy to spot fields that have been skipped in the data entry process.
- **If an error is found**, ideally the value should be changed to the correct value. However, if these is no record of what the value should be, the missing value code should be used, e.g. 9=missing.
- **Tell the package which values indicate missing data**. Usually this means converting the numeric missing code to the program's official 'missing value'.

Logical checks

- **Check for duplicate records**. Each record in a file must have a unique key. Usually this is a single variable (e.g. idno.), but it may be a combination of two or more variables (e.g. idno and visit date).
- **Check for consistency between variables**. Data values can depend on the value of another variable. For instance, in a study of survival after a kidney transplant, information on the number of previous pregnancies is relevant only for women, who should all have a non-missing value, whilst men should all have a missing value.
- If you have a set of criteria for selecting subjects for your study, check that all participants were eligible.
- **If a measure is recorded more than once at the same time point,** the repeated values should be within a reasonable range of each other. For instance, you may decide to newborn head circumference more than once to get a more accurate value. Discrepancies between the measurements should be small.

Checking categorical data (e.g. yes/no or mild/moderate/severe)

It is quite a simple task to check that all data values are plausible because there are a fixed number of pre-specified values. For each of the categorical variables produce frequency tables showing all the recorded values. Alternatively, if the package allows, include statements that make explicit checks on values. For instance, if gender has been coded as: 0=female, 1=male, 9=missing, then the statement would assert that gender contains only the values 0, 1 or 9. If the statement fails, you know that the variable contains at least one dubious value.

Checking continuous data (e.g. height or age)

For continuous data we can specify lower and upper limits on what is reasonable. Values that fall outside this range may not necessarily be wrong. All suspicious values should be checked and any errors corrected. If a value is felt to be impossible rather than just unlikely, it should be recorded as 'missing'. *Be aware that sometimes an apparently extreme value may be valid.*

- Produce summaries showing the mean, median, variance and minimum and maximum values for each continuous variable.
- As with categorical variables, you can include statements to check for values below the expected minimum and above the expected maximum values.
- Produce a dotplot to easily spot any possible errors.

Checking Dates

- Check all dates are within a reasonable time span. In a study where year 7 students are surveyed, the date of birth should be about 12 years prior to the survey date.
- Check dates are in the right order, eg dob < date of 1st visit< date of 2nd visit
- Ages and time intervals can be calculated via a statistical package using the relevant dates. Check that ages and time intervals lie within the expected range. eg: negative ages indicate data error

Longitudinal Studies

- Where the same variable is measured at several time points for each subject, it is valuable to plot each person's sequence of recorded values to ensure that they behave reasonably.
- Check that variables that shouldn't change over time are consistent.
- Only id numbers/subjects with a baseline record should have data at later time points.

11.5. SUMMARY: Steps to good data management

- Use well designed data collection forms
- Create a data dictionary to document the project's metadata (information about the data such as its meaning, relationships to other data, origin, usage, and format)
- Use a relational database where possible
- Ensure the database design will give the required outcomes with the greatest accuracy
- Carry out data cleaning before analysing the data
- Keep a record of all analyses (e.g. STATA do files)

- Archive all data at the completion of the project (both paper based and electronic)
- And budget appropriately in your project from the start for these activities, including input from a data manager
- Data quality is achieved with a meticulous, systematic and logical approach to data management.
- If not enough care, thought and time are given, problems can occur at the analysis stage. Your analysis will then be based on invalid data, leading to false results. *You need to be obsessive*.

12. Key resources

12.1. REDCap access and support

See extensive information on the Research Education Program open access website:

- **REP Seminar: "Using REDCap for data capture and management"**
  This 1hr overview reviews the basic functionality of REDCap and introduces some of its features.
  WATCH the seminar recording and DOWNLOAD the accompanying handout

- **REDCap Workshops and resources**
  View recordings of the REDCap Basics, REDCap Intermediate and Advanced REDCap workshops hosted by the Research Education Program and download the accompanying resources from our "REDCap Resources" page on the website.
  Tab #3 REDCap Recorded Seminars and Workshops
  CAHS - REDCap resources (health.wa.gov.au)

- **General access and help information including other instructional videos**:
  CAHS - Additional Resources (health.wa.gov.au)

12.2. Important REDCap information for CAHS staff

- Updates to the REDCap licensing terms and conditions mean a Telethon Kids Institute employee must be actively engaged in all projects using the Telethon Kids instance of REDCap.
- Projects where the entire team are external (Dept of Health employees without a Telethon Kids appointment) cannot reside on the Telethon Kids instance of REDCap.
- In the short-mid term, all existing projects set-up on the Telethon Kids instance of REDCap can remain.
- Access to a Dept of Health instance of REDCap is now available for all WA Health employees (with an active HE number and WA Health email address). See attachment below.
- REDCap support is still available to all CAHS Dept of Health based researchers through the Telethon Kids Biometrics team.
- For projects utilising the Dept of Health instance of REDCap, workshops and support/advice is still available, but there are limitations on the 'hands on' account related activities able to be performed.
- For projects with active Telethon Kids collaborations, there are no changes to the ability to use the Telethon Kids Instance or the processes by which you do this.

12.3. More useful websites

Other resources are embedded above through this document

- NHMRC: Australian Code for the Responsible Conduct of Research (2018) https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018
- NHMRC Competencies for Australia Academic Clinical Trialists (May 2018) https://www.nhmrc.gov.au/about-us/publications/competencies-australian-academic-clinical-trialists
- National Statement on Ethical Conduct in Human Research (2007) https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018
- EpiData Software link http://www.epidata.dk/
- UK Data Archives http://www.data-archive.ac.uk/
- University of Western Australia - Research Data Management Toolkit Welcome - Research Data Management Toolkit - Guides at University of Western Australia (uwa.edu.au)
- University of Western Australia - Code of Conduct for the Responsible Practice of Research. https://www.uwa.edu.au/policy/-/media/Project/UWA/UWA/Policy-Library/Policy/Code-of-Conduct/Integrity/Research-Integrity/Research-Integrity-Policy.doc

12.4. Data Linkage Branch Training for linked data

- Free workshops for researchers and other applicants interested in applying for linked data.
- Contact: DataServices@health.wa.gov.au
- Workshops generally cover core essentials:

  - The data linkage process
  - The preparation of data
  - The datasets available to researchers
  - Ethical considerations
  - The application process

12.5. Data Manager Support

Data managers can provide advice and/or assistance across a wide range of issues such as data base set up, data entry and cleaning. It is wise to allow in your project budget sufficient funds to seek assistance from a data manager from the earliest possible stages. Collaboration may be key to achieving access to what is sometimes a limited resource. Many research institutes, universities and some hospital departments have a data manager on site available for support.

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital **Foundation**

Perth Children's Hospital

CAHS Research Education Program

Research Skills Seminar Series

A free, open-access resource designed to upskill busy clinical staff and students and improve research quality and impact.

2023 Seminar Schedule

Interactive in pdf format
Last updated 7/7/23

	DATE	TOPIC	PRESENTER	ENROL	WATCH
1	3 Mar	**Research Fundamentals**	Dr Kenneth Lee, UWA	-	2023
2	17 Mar	**Introductory Biostatistics**	Michael Dymock, TKI	-	2023
3	28 Apr	**Scientific Writing**	A/Prof Tony Kemp, UWA	-	2023
4	5 May	**REDCap for Data Capture and Management**	Dr Jane Mugure Githae, CAHS	-	2023
5	12 May	**Using Social Media in Research**	Dr Kenneth Lee, UWA	-	2023
6	19 May	**Getting the Most out of Research Supervision**	A/Prof Sunalene Devadason, UWA/CAHS	-	2022
7	26 May	**Research Impact**	Dr Tamika Heiden, Vic	-	2023
8	2 Jun	**Survey Design & Techniques**	Dr Jane Mugure Githae, CAHS	-	2023
9	9 Jun	**Conducting Systematic Reviews**	Prof Sonya Girdler, Curtin Uni	-	2023
10	16 Jun	**Consumer & Community Involvement in Research**	Belinda Frank, TKI	-	2023
11	23 Jun	**Project Management**	Melanie Wright, SMHS	-	2023
12	30 Jun	**Sample Size Calculations**	Michael Dymock, TKI	-	2023
13	21 Jul	**Introduction to Good Clinical Practice**	Alexandra Robertson, CAHS	REGISTER	2021
14	28 Jul	**Data Collection and Management**	Dr Jane Mugure Githae, CAHS	REGISTER	2022
15	4 Aug	**Rapid Critical Appraisal of Scientific Literature**	Dr Natalie Strobel, ECU	REGISTER	2022
16	18 Aug	**Media and Communications in Research**	Keryn McKinnon, TKI	REGISTER	2022
17	25 Aug	**Oral Presentation of Research Results**	Dr Jane Mugure Githae, CAHS	REGISTER	2022
18	1 Sep	**Involving Aboriginal Communities in Research**	Cheryl Bridge, Mara West and Mel Robinson – TKI and CAHS	REGISTER	2022
19	8 Sep	**Knowledge Translation**	A/Prof Fenella Gill, Curtin Uni/CAHS	REGISTER	2021
20	13 Oct	**Research Governance**	Dr Natalie Giles, CAHS	REGISTER	2022
21	20 Oct	**Grant Applications and Finding Funding**	Dr Tegan McNab, TKI	REGISTER	2022
22	27 Oct	**Statistical Tips for Interpreting Scientific Claims**	Michael Dymock, TKI	REGISTER	2022
23	17 Nov	**Ethics Processes for Clinical Research in WA**	Natalie Giles, CAHS	REGISTER	2020
24	24 Nov	**Qualitative Research Methods**	Dr Shirley McGough, Curtin Uni	REGISTER	2022
25	1 Dec	**Innovation and Commercialisation**	Dr Helga Mikkelsen, Brandon BioCatalyst + Ashley Schoof	REGISTER	2022

Register via Eventbrite

View recorded seminars online

Subscribe to our mailing list

Contact Us

☎ (08) 6456 0514

✉ researcheducationprogram@health.wa.gov.au

cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program

Seminars are held from 12:30-1:30pm at Perth Children's Hospital Auditorium and are broadcast live online through Teams and Avaya.
Seminars are recorded and uploaded to our website within a week of presentation. Topics are subject to change with appropriate email notice provided.
Handouts are revised and updated regularly. A light lunch is provided for attendees at our PCH auditorium. Attendance certificates are available on request.

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

CAHS Research Education Program

Research Skills Seminar Series 2023

A free, open-access resource designed to upskill busy clinical staff and students and improve research quality and impact

Rapid Critical Appraisal of Scientific Literature

4th August 2023

12.30-1.30pm

Given the sheer volume and variable quality of published papers even in high impact journals, it is essential to have skills to target and rapidly appraise relevant literature to answer current clinical questions. This seminar provides simple strategies to help focus your reading, examine validity of results, and decide whether to accept and apply them in your setting.

Perth Children's Hospital Auditorium

Level 5, 15 Hospital Ave Nedlands
Accessible via pink or yellow lifts

- OR -

Access online via Teams or Avaya

- OR -

Watch live
from a hosted video-conferencing site

- Bunbury Hospital
- Fiona Stanley Hospital
- Lions Eye Institute
- Royal Perth Hospital

Register via Eventbrite

View recorded seminars online

Subscribe to our mailing list

Meet the presenter

Dr Natalie Strobel
Senior Research Fellow

Centre for improving health services for Aboriginal and Torres Strait Islander children and families,Kurongkurl Katitjin, Edith Cowan University

Natalie is a Senior Research Fellow employed as the team leader on the evidence synthesis stream for the Centre for Improving Health Services for Aboriginal and Torres Strait Islander Children and Families (ISAC) at the Edith Cowan University. She has been working in health services research and epidemiology to improve service delivery to children, in particular Aboriginal and Torres Strait Islander children.

Dr Strobel has been consulting with WHO on various neonatal guidelines including for preterm and low birth weight infants. Her work has had a strong focus on ensuring projects delivered are needs-based and inform policy and practice.

researcheducationprogram@health.wa.gov.au (08) 6456 0514
cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

CAHS Research Education Program

Research Skills Seminar Series 2023

A free, open-access resource designed to upskill busy clinical staff and students and improve research quality and impact

Media and Communications in Research

18th August 2023

12.30-1.30pm

Understanding how to work with the media is essential and a critical responsibility for all researchers, whether it's the newspaper, TV, radio, or social media. This seminar will provide practical techniques on working with the media and ensuring your bottom line is delivered in an engaging, accurate, and responsible way.

Perth Children's Hospital Auditorium

Level 5, 15 Hospital Ave Nedlands
Accessible via pink or yellow lifts

- OR -

Access online via Teams or Avaya

- OR -

Watch live
from a hosted video-conferencing site

- Bunbury Hospital
- Fiona Stanley Hospital
- Lions Eye Institute
- Royal Perth Hospital

Register via Eventbrite

View recorded seminars online

Subscribe to our mailing list

Meet the presenter

Keryn McKinnon
Head, Strategic Communications
Telethon Kids Institute

Keryn is a corporate communications specialist with more than three decades experience as a senior journalist, newspaper editor, government media adviser and public relations executive.

With expertise in crisis communications, stakeholder engagement and media management, Keryn was Agency Director at Perth PR firm Hunter Communications prior to taking on the role as Head, Strategic Communications at Telethon Kids Institute in April this year.

A highly respected leader of communications teams, Keryn is an expert storyteller who thrives on delivering compelling, relevant and engaging content to targeted audiences.

TELETHON KIDS INSTITUTE
Discover. Prevent. Cure.

researcheducationprogram@health.wa.gov.au (08) 6456 0514
cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

CAHS Research Education Program

2023 Research Skills Workshop Series

Perth Children's Hospital Foundation

The Research Education Program (REP) Research Skills Workshop Series, supported by the Perth Children's Hospital Foundation, offers a series of interactive workshops that focus on building the most fundamental research skills required to undertake clinical research projects.
Workshops aim to directly build user skills and knowledge in a guided environment, with time to ask questions specific to your own project.

Presented by: CAHS Research Department **Location:** PCH, TKI Seminar Room, Level 5 (West)

Topic	Day	Date	Time	Max *(in-person)*
Workshop 1 - Setting up Clinical Trials Clinical trials are the benchmark for testing interventions in healthcare. This workshop aims to provide practical advice to clinical researchers who want to gain insight on how to develop and complete their clinical trial on time and within budget. Come learn practical aspects of the steps involved in developing a clinical trial from the research idea to protocol development and execution.	Tue	14 Mar	2.30pm - 4:00pm	WATCH the recording
Workshop 2 - Manuscript Writing Journal publications are an integral part of dissemination of research findings. However, it can be overwhelming to convert several months of research into a succinct manuscript that will be loved by peer-reviewers and attract readers. This workshop is designed to give those who have completed their research projects, practical skills to transform their research data into publishable peer-reviewed literature.	Tue	15 Aug	1.00pm - 3:30pm	**40** Register
Workshop 3 - Oral Presentation of Research Results Dissemination of research findings is integral in knowledge translation and clinical practice change. Oral presentations provide rapid dissemination of research findings to a target audience. We invite you to a practical session that will provide useful tips, practice sessions and personalised feedback to help deliver an adequate depth of your research findings to various research stakeholders.	Tue	5 Sep	1.00pm - 3:30pm	**40** Register
Workshop 4 - Navigating Research Ethics and Governance in WA If you are undertaking a research project or are thinking about becoming involved in research, understanding the review and approval requirements for your research project may appear intimidating. This workshop aims to help you understand the process of ethical and governance review for research approvals at CAHS - includes PCH, CACHS, CAHMS and Neonatology.	Tue	21 Nov	1.00pm - 3:30pm	**40** Register

IMPORTANT
Places are strictly limited and offered on a first-come, first-serve, basis. If you are not able to attend a workshop for which you have registered, please contact Research Education Program support via phone or email to cancel your reservation and/or be placed on the waitlist.

Register via Eventbrite

View recorded seminars

Subscribe to our mailing list

Contact Us

☎ (08) 6456 0514
✉ *researcheducationprogram@health.wa.gov.au*
cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

CAHS Research Education Program

Research Skills Workshop Series

Workshop 2: Manuscript Writing

15th August 2023 • 1.00 - 3.30pm • PCH, TKI Level 5 Seminar Room

The Research Education Program (REP) offers a series of interactive workshops that focus on building the most fundamental research skills required to undertake clinical research projects. REP Research Skills Workshops aim to directly build user skills and knowledge in a guided environment, with time to ask questions specific to your own project.

Journal publications are an integral part of dissemination of research findings. However, it can be overwhelming to convert several months of research into a succinct manuscript that will be loved by peer-reviewers and attract readers.

This workshop is designed to give those who have completed their research projects, practical skills to transform their research data into publishable peer-reviewed literature.

Places are strictly limited and offered on a first-come, first-serve basis. If you are unable to attend a workshop for which you have registered, please contact Research Education Program support via email to cancel your booking. Laptops are available if required

Register via Eventbrite

View our online recorded resources

Subscribe to our mailing list

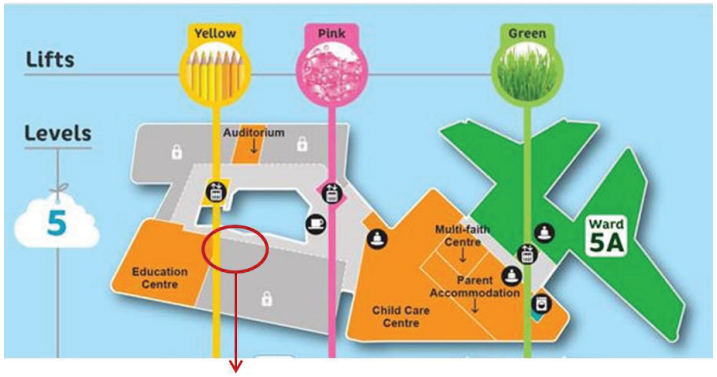
About the Presenter

Dr Kenneth Lee
Senior Lecturer
Pharmacy Practice, UWA

THE UNIVERSITY OF WESTERN AUSTRALIA

As a clinician, Dr Lee works in a large medical centre as the in-house pharmacist, as well as in private consultancy, where he primarily delivers comprehensive medication management review services to domiciliary patients.

Dr Lee is passionate about integrating practice, research, and teaching.



Location of the TKI Seminar Room
Accessible via yellow or pink lifts

Contact Us

(08) 6456 0514 researcheducationprogram@health.wa.gov.au
cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program

Perth Children's Hospital Foundation

REDCap Workshops are presented by the Research Education Program in partnership and with support from the PCH Foundation and Telethon Kids Institute as part of the Research Education Program REDCap Workshop Series, presented by the WA Department of Health CAHS Department of Research and invited speakers.

TELETHON KIDS INSTITUTE
Discover. Prevent. Cure.

Government of **Western Australia**
Child and Adolescent Health Service

TELETHON KIDS INSTITUTE
Discover. Prevent. Cure.

Child Health Research Symposium

Celebrating Innovation, Collaboration and Translation

cahs.health.wa.gov.au/Child-Health-Research-Symposium

8 - 10 November 2023

Perth Children's Hospital Auditorium
and Telethon Kids Institute Manda

Abstract submission deadline - 13 August 2023 5pm AWST

Incorporating the CAHS Nursing and CAMHS Symposiums

Neonatology | Community Health | Mental Health | Perth Children's Hospital

CAHS Research Education Program

Research Skills Seminar Series

A free, open-access resource designed to upskill busy clinical staff and students and improve research quality and impact.

Data Collection and Management

Thank you for your interest in this seminar

Please complete this 1-minute evaluation.
Your feedback will help guide future presentations and educational activities.

How did you attend the seminar?

- ○ Live seminar at Perth Children's Hospital
- ○ Hosted video-conference on-site (e.g. FSH, Lions Eye, RPH etc.)
- ○ Online via Avaya or Teams
- ○ Viewed online recording

Please rate your agreement with the following statements:

	N/A	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
The aims and objectives were clear	○	○	○	○	○	○
The session was well structured	○	○	○	○	○	○
Presentation style retained my interest	○	○	○	○	○	○
The speaker communicated clearly	○	○	○	○	○	○
The material extended my knowledge	○	○	○	○	○	○
The additional resources were helpful	○	○	○	○	○	○

What were the best aspects of the seminar?

What changes or improvements would you suggest?

How did you hear about the seminar?
(you can select multiple answer)

- ☐ Email invitation from Research Education Program
- ☐ CAHS Newsletters e.g. The Headlines, The View, CAHS Research Newsletter
- ☐ "Health Happenings" E-News
- ☐ Healthpoint Intranet Upcoming Events
- ☐ Collegiate lounge screen or other posted promotional material
- ☐ Telethon Kids Institute screen or other posted promotional material
- ☐ Telethon Kids Institute Newsletter
- ☐ Other

Thank you!

Government of **Western Australia**
Child and Adolescent Health Service

Perth Children's Hospital

cahs.health.wa.gov.au/ResearchEducationProgram

CAHS Research Education Program Research Skills Seminar Series

✉ ResearchEducationProgram@health.wa.gov.au
cahs.health.wa.gov.au/ResearchEducationProgram