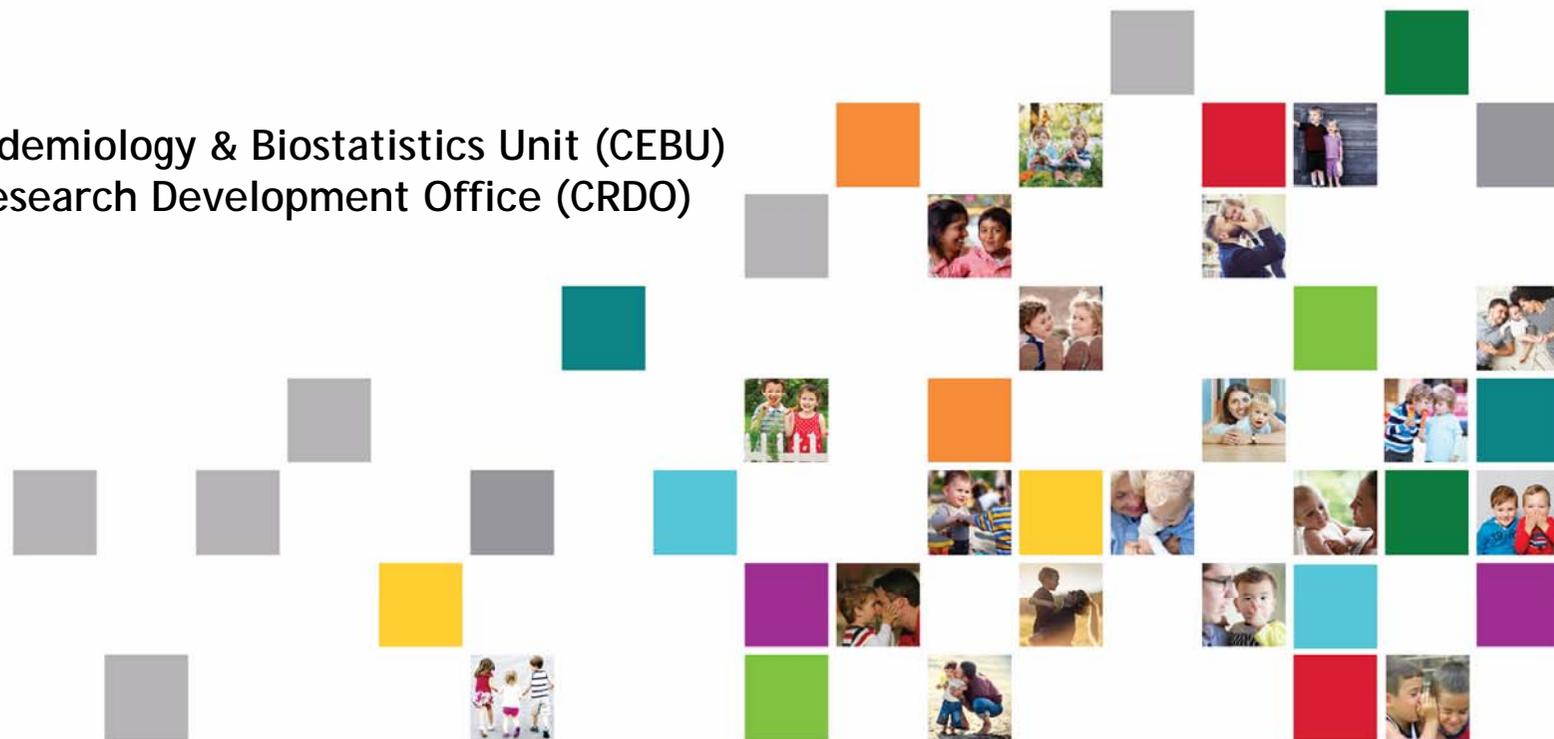


The evolving landscape for managing research data

Luke Stevens, Clinical Epidemiology & Biostatistics Unit (CEBU)
Fiona Williams, Clinical Research Development Office (CRDO)

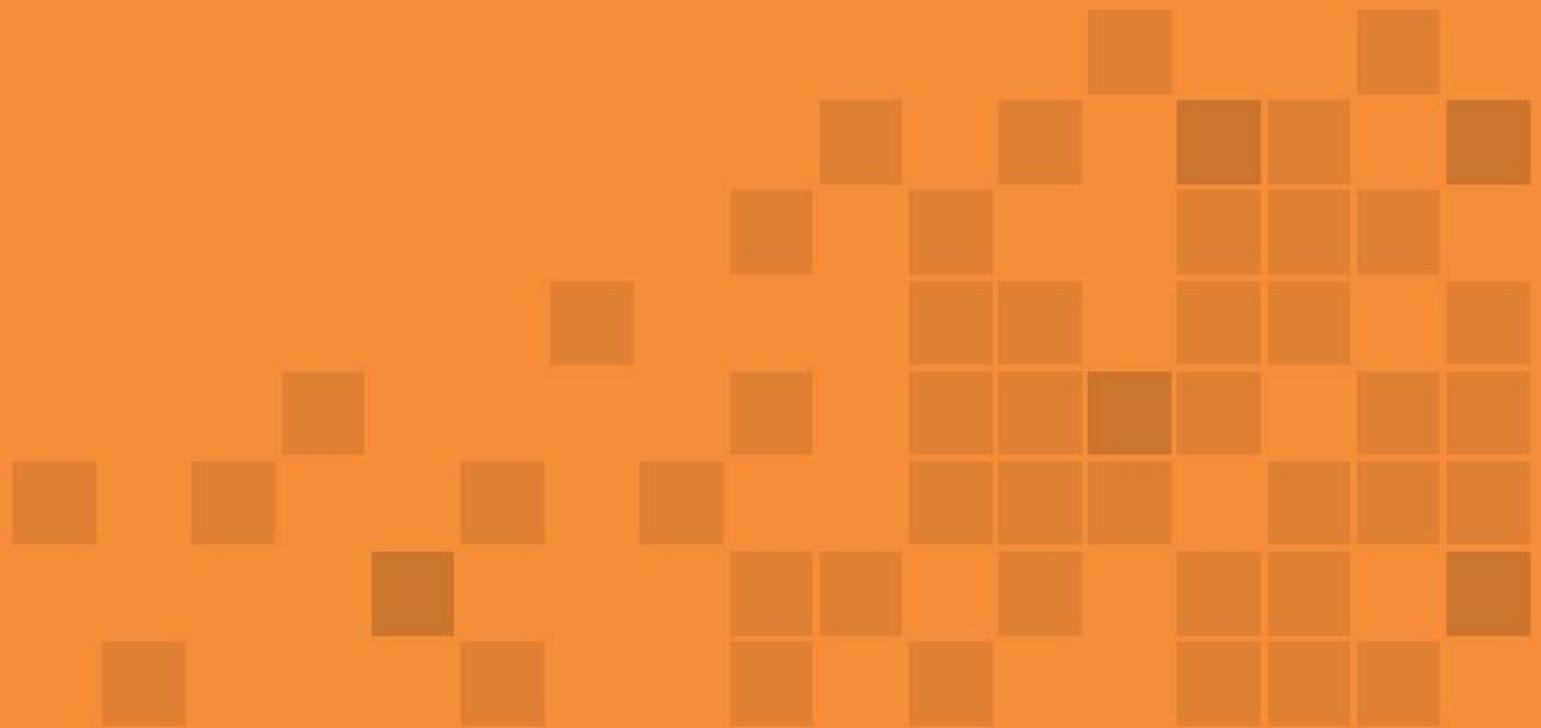
Research Coordinator Forum
10 April 2019



Outline

- What is data? And what is data management planning?
- What guidelines do we work to in managing data for research ? What has changed and how do we implement the changes?
- How are data management systems evolving?

What is data - and its management?



What is data? What is information?

- Per the National Statement* section 3.2 , Element 4:
 - ✓ the term 'data' is intended to refer to **bits of information in their raw form**
 - ✓ the term 'information' generally refers to **data that have been interpreted, analysed/contextualised.**

Data and information may include:

- ✓ What is stated in interviews, focus groups, questionnaires/surveys, personal histories
- ✓ Images, audio recordings, other AV material
- ✓ Records for administrative purposes (e.g. billing) or required by legislation (e.g. disease notification)
- ✓ Digital information (via use of mobile devices and internet)
- ✓ Physical specimens/artefacts
- ✓ Information generated by analysis of existing personal information (e.g. clinical, organisational, social, observational, sources)
- ✓ Observations
- ✓ Results from experimental testing and investigations
- ✓ Information derived from human bio-specimens

* [National Statement on Ethical Conduct in Human Research](#) (NHMRC) (2007 and all updates)

What is data management planning?



High quality data and results when:

- ✓ Project data are properly organised, preserved and documented
- ✓ Your data have integrity, accuracy, validity
- ✓ Processes are reproducible

Good data management is the foundation for good research - planning is essential

- ✓ "Failing to plan is planning to fail."



- ✓ "In preparing for battle I have always found that plans are useless, but planning is indispensable."



What does data management planning involve?

Keep in mind the project life stage - it's not just the database and the analysis datasets...

- **Concept**
 - ✓ Documents, pre-existing and/or external data
- **Development**
 - ✓ Documents (protocol, planning, management, data documentation)
 - ✓ Systems development and piloting
- **Data Collection**
 - ✓ Documents (SOPs & administration)
 - ✓ Data capture tools (hardcopy/ electronic)
 - ✓ Database
 - ✓ QA/QC - data cleaning, interim analyses (SOPs, tools)
 - ✓ (SOPs) for Identifiability Storage, Access, disclosure
- **Output Generation**
 - ✓ Analyses, publications
- **Archive**
 - ✓ (data, metadata), sharing and re-use
 - ✓ Disposal
- **Long-term custodianship** (post-archive period)



What legislation/ guidelines do we work to in managing data for research?

What has changed? How do we implement the changes?

Data & Information requirements: Australian legislation (not research-specific)

The Health Records Act Victoria (2001)	Privacy and Data Protection Act (2014)	Privacy Amendment (Notifiable Data Breaches) Act 2017
<p>Intended to ensure fair & responsible health information by:</p> <ul style="list-style-type: none"> • protecting privacy • right of access for individuals • framework for resolving complaints <p>Requires appropriate ethics approval and informed consent before hospital patient data can be used for research</p>	<p>Covers responsible collection and handling of personal information - Information Privacy Principles</p> <ul style="list-style-type: none"> • relevancy, security, accuracy, integrity, limited use and disclosure <p>Requires data security regime</p>	<p>Covers unauthorised access to or loss of data on individuals.</p> <ul style="list-style-type: none"> • Requirement to assess potential harm - may require notification to the Information Commissioner. <p>In response: a <u>new MCRI data breach policy</u> was released in FEB 2018.</p>

Data & Information requirements: Privacy Amendment 2017 - MCRI response Data Breach Policy 2018

- **Policy owner** - MCRI Legal
 - **Definition**
 - A data breach is any incident in which data is compromised, disclosed, copied, communicated, accessed, removed, destroyed, stolen, lost or used by unauthorised individuals, whether by accident or intentional.
 - **Action required**
 - Containment and mandatory notification - follow the process in the MCRI [Data Breach Response Plan](#) includes notifying Head of Legal *or* CIO *or* COO
 - **Early detection and response essential**
 - To protect affected individuals and MCRI
 - To meet MCRI obligations under Privacy Act and limiting liability.
 1. MCRI must take reasonable steps to protect the personal information that it holds, including having a data response plan and notifying the affected individuals or the Information Commissioner in case of a data breach.
 2. Breaches of Privacy Act can incur fines up to AUD \$1,800,000.
-

MCRI data breach policy 2018: types of data and risk of harm

Type of data	Non-identifiable data	Coded/re-identifiable data	Fully identifiable data	
			<i>Personal information</i>	<i>Sensitive information</i>
What does it cover?	Data that does not identify an individual.	A code is needed to link identifiable information to an individual.	Data that identifies an individual.	Special subset of personal information.
Examples	<ul style="list-style-type: none"> ■ Epic slicer/dicer 	<ul style="list-style-type: none"> ■ De-identified data set sent to an external collaborator for research or analysis purposes; ■ de-identified information about sex disorders or sexual activities of teenagers stored in a database and used for the purpose of publication. 	<ul style="list-style-type: none"> ■ Identification data and personal characteristics collected for a survey or for booking an appointment; ■ Professional data for HR purposes or invoicing management. 	<ul style="list-style-type: none"> ■ Health data, ■ racial or ethnic origin, ■ sexual orientation or practice, ■ criminal records <p>collected for research, clinical trial, medical care, diagnostic or treatment.</p>
Risk of harm in case of data breach	<p>Low to Medium</p> <p><i>depending on whether the non-personal information is sensitive or commercial.</i></p>	<p>Medium to high</p> <p><i>depending on whether the code has been secured and is strong or not.</i></p>	<p>Medium to high</p> <p><i>depending on the type of information.</i></p>	<p>Generally high</p>

Data & Information requirements: Australian guidelines for research

<p>Australian Code for the Responsible Conduct of Research (NHMRC, updated 2018)</p> <p>Guide to Managing and Investigating Potential Breaches of the Australian Code for the Responsible Conduct of Research (NHMRC, new 2018)</p>	<p>National Statement on Ethical Conduct in Human Research (NHMRC, 2007 last updated 2018)</p>	<p>Integrated Addendum to ICH E6(R1): Guideline for <u>Good Clinical Practice</u> ICH E6(R2) (annotated with TGA comments)</p>
<p>Applicable to all research:</p> <ul style="list-style-type: none"> • compliance required for NHMRC-funded research • other research agencies encouraged to adopt the Code as a mandatory requirement <p>2018 - Revamped format</p>	<p>Applicable to all human research</p> <p>2018 - Revamped format</p>	<p>Mandatory for drug/device clinical trials for submission to regulatory authorities <u>BUT</u> recommended for all research</p> <p>1st revision of the original guideline (1997) expanded section on data management</p>

Data & Information requirement updates: The Australian Code for the Responsible Conduct of Research 2018

- The 2018 Code provides broad principles and responsibilities:
 - Principles of research conduct =(8)
 - Responsibilities of institutions (n=13) & researchers (n=16)

 - The 2018 Code is supported by:
 - NHMRC's new guideline "Guide to Managing and Investigating Potential Breaches of the Australian Code for the Responsible Conduct of Research"
 - Further NHMRC guidance documents (*currently in draft - still to be released*)
 - Authorship Guide
 - Management of Data and Information in Research

 - No evident changes in data/data management requirements in documents released to date

 - Campus action required - to update documents
 - Melbourne Children's Governance Framework (RCH) /Framework for Administrating Allegations of Breaches of the Australian Code for the Responsible Conduct of Research (MCRI)
-

Data & Information requirement updates: National Statement (NS)

Chapter 3 “Ethical considerations in design, development, review and conduct of research”

- See 3.1 Element 4 for expanded section on collection, use & management of data & information

1. New discussion re identifiability of information

- Identifiability of information should be seen as existing on a *continuum*
 - NS no longer uses categorical terms ‘identifiable’, ‘potentially re-identifiable’, ‘anonymised’
 - Technological changes make possible identifying anonymised data (e.g. predictive analytics, machine learning, proliferation of data sets, degradation of privacy protections, data breaches)

2. Expansion of data management requirements

- Data management plan (DMP) needed for all research projects - to address data/information generation, collection, access, use, analysis, disclosure, storage, retention, disposal, sharing and re-use
 - DMP should also detail the physical/electronic security, license & confidentiality agreements, training (CEBU DMP template being updated)
- Data & information should be stored for use in future research - reiterated in 2018 version with requirement for documenting valid reason(s) and researcher justification for non-banking

Data & Information requirement updates: National Statement (NS) *cont'd*

3. Secondary use of data & information

- Commonly understood as “Use of data/information collected for previous research or non-research purposes”
- 2018 version further addresses main issues (scope of initial consent & impracticability of re-consenting)
 - Suggested strategies to address:
 - Ensure research translated into practice
 - Publicise research results in a location & language suitable for general community
 - Acknowledge source of data/information in publications

4. Sharing of data & information

- Emphasised in 2018 version
 - “Data or information may collected... for an initial purpose, it is common for researchers to ‘bank’ data/information for possible use in future research projects or to otherwise share with other researchers...Also common for funding agencies to require the sharing of research data either via open access...or...mediated access.”
-

Data & Information requirement updates: ICH-GCP E6(R2) TGA-annotated

Section 5.0 “Quality Management” is new section - highlights importance of quality management to protect trial participants and ensure reliable trial results (i.e. the aims of GCP).

1. Identify the data that is critical to trial results points and processes

- Identify risks to collecting this data - assess likelihood of occurrence and impact of errors
- Mitigate /manage risks - review periodically

2. For electronic data handling - additional requirements

- Validate electronic systems
 - Document procedures - SOPs (i.e. set up, security, change control, audit trail, authorised users, back up)
 - Maintain data integrity (e.g. during software upgrades, migration of data)
-

MCRI resources & training

Some resources & training already in place:

1. Resources

- MCRI institutional policy (being updated)
- Data Management Plan template (see CEBU website)

2. Training (CEBU)

- 'Introduction to research data management' (mandatory on-line learning module - new staff)
- Research methods training (e.g. Intro to biostatistics, Research data essentials', My data rules)
- Data capture/entry - REDCap
- Data analysis - Stata

Next steps...

CEBU & CRDO are working on how to document data activities for your research project...

- Updating Data Management Plan template
- Creating new procedure templates to document the who, how and where of data activities
 - Testing your database (functionality)
 - Data collection and handling (includes data cleaning, data export/transfer, analysis, archiving)
 - System use (creating users, assigning permissions)
 - Change control process (making and approving changes to database structure & data)
 - Data back up (where data will be held and - if not on shared network - process for back up)

Also hoped that an outcome of the recent Data Review will be an increased strategic focus on data management at an Institute-wide level.

How are data management systems evolving?





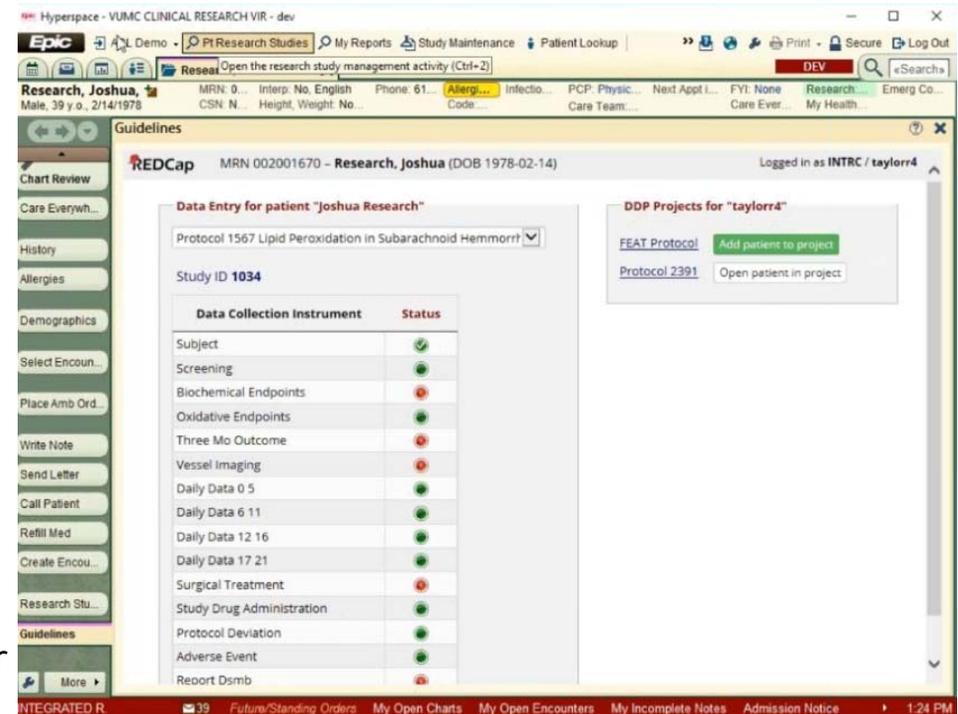
- Sync directories from your computer with your Home drive (H:\) on the MCRI network.
- Access via web browser from anywhere, on any device.
- Share files and directories with other ownCloud users (e.g. MCRI staff)*
- Enable download of files by anyone via a unique URL.
- Share files and directories with anyone (e.g. external collaborators)*
- Accessible backups (daily for past week - older can be retrieved from tape storage)

** Be careful if multiple people are editing the same document*

REDCap / Epic



- Map REDCap project fields to Epic fields (demographics, problem list, allergies, lab values).
- Launch REDCap within Epic (login credentials will be remembered for next time).
- Create REDCap project record for your Epic patient record, data for mapped fields copied across.
- From REDCap, look up Epic patient record by UR (accounting for permissions, BTG).
- REDCap adjudication workflow - see values and accept before they are saved to REDCap.
- Subsequent updates to Epic data automatically appear for adjudication.



APRIL 2019 - Following the Epic upgrade, the REDCap/Epic interface is not functioning. For resolution by REDCap in association with Epic.

- <https://mcri.figshare.com>
- Public archive of MCRI research projects and output.
- Publish details of project (who, what, when, where).
- Publish documents, datasets, files.
- Published items allocated a Digital Object Identifier (DOI)
- Items are findable, downloadable, citable.
- Can publish confidentially.



DataCite

Stevens, Luke (2018): MCRI figshare User Guide. figshare. Paper.

<https://doi.org/10.25374/MCRI.6729197.v1>

MCRI figshare User Guide

Paper posted on 03.07.2018, 11:41 by [Luke Stevens](#)

- MCRI's preferred cloud-based project planning and tracking tool.
- "Middle-ground" between Excel and more specialised productivity tools like Microsoft Project, basecamp (for document collaboration, discussion), Trello (for task lists)
- Features: dashboards, automated alerts and reminders, basic forms, attachments, audit trail.
- Use for: project planning/scheduling (Gantt charts), mapping milestones and task dependencies (critical path), issue tracking, resource management.
- Do NOT use for: research participant data (as the application is hosted in the USA).
- Free for users to access existing Smartsheets, licence required to build and make alterations.
- Training available.
- Look for announcement from IT next week...

Participant Relationship Management

- Customer Relationship Management (CRM) system e.g. Salesforce.
- Manage your contacts with research participants.
- Capturing and storing consents.
- Logging communications.
- Organising visits - reminder messages.
- Integrate with REDCap e.g. for creating project record, survey links.

- GenV currently evaluating candidates - for GenV, but also potential for wider availability.
- Work in progress, but important area of need identified (let us know of others!).

Contacts

Luke Stevens
Data Management Coordinator
Clinical Epidemiology and Biostatistics Unit (CEBU)
Murdoch Children's Research Institute
South Building, Level 4
T: (03) 9345 6552
E: luke.stevens@mcri.edu.au

The Clinical Research Development Office (CRDO) team
Murdoch Children's Research Institute
South Building, Level 4
E: crdo.info@mcri.edu.au



THANK YOU

@ Murdoch Children's Research Institute, 2017

