Linked data – Handle with Care
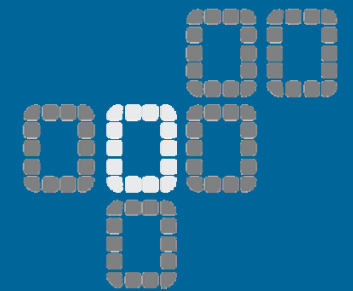
Brooke Baxter [1]

Diana Rosman [1]

Geoff Davis [1]

1 Department of Health, Perth, Western Australia.

Linked data – Handle with Care

Objective:

- Describe how the WA Data Linkage Branch has been structured to protect the privacy of individuals & service providers data, where linked data are provided, for ethically approved research in WA & elsewhere.

Background:

- Dunn first described record linkage in 1946. Dunn likened *record linkage* to the creation of a Book of Life[1]

- Records of significant events over the course of ones life are like the pages in a book[1]

- First page represents birth and the last page is death[1]

- Today, data linkage described as the bringing together of data from two or more sources that relate to the same individual, family, place or event.

WA Data Linkage System

- WA Data Linkage System (WADLS) established in 1995 in collaboration with:
  - Department of Health - Western Australia
  - University of Western Australia
  - Curtin University of Technology
  - Telethon Institute for Child Health Research.

- WADLS managed by the Data Linkage Branch (DLB) at the Department of Health – Western Australia

- DLB is led by a Program Manager & employs 18 staff predominantly from WA State Government & some University staff, who work in system & liaison roles.

Two distinct groups within Data Linkage Branch

1. Linkage Services

- o Linkage Officers

- o Systems Officers & Programmers

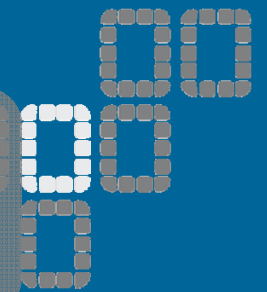**See demographic data**

↓

Creation of linkage keys

2. Client Services

- o Project Manager

- o Project Officer

- o Quality Assurance

- o Analysts

NO access to demographic data

↓

Provision of data extracts for ethically approved research

Data Linkage Branch

Linkage Services Group:

- Led by a Manager, Data Linkage Systems
- Responsible for creating links for the WA population across many datasets derived from health data collections & data collections from other jurisdictions
- Links created using special software & tools
- Access to identifying data (i.e. names, addresses & dates of birth) is fundamental to creating links
- WADLS is dynamic in nature, links are created as a "chain of links", allowing for updates every moment of everyday
- No access to clinical or service data.

Data Linkage Branch
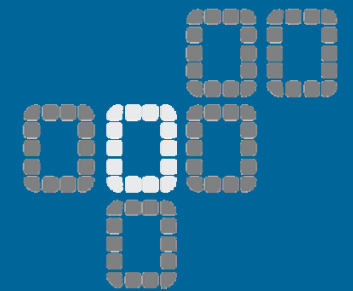
Client Services Group:

- Since 1995 more than 500 projects have been approved for access to linked data in WA
- With increased demand, Client services group established in 2004
- Linked data requests now more complex
- Inclusion of recent cross-jurisdiction linkages to Australian Govt data & other state agencies
- No access to identifying data
- Work with multiple data collections to facilitate the extraction of clinical/service data for ethically approved research projects.
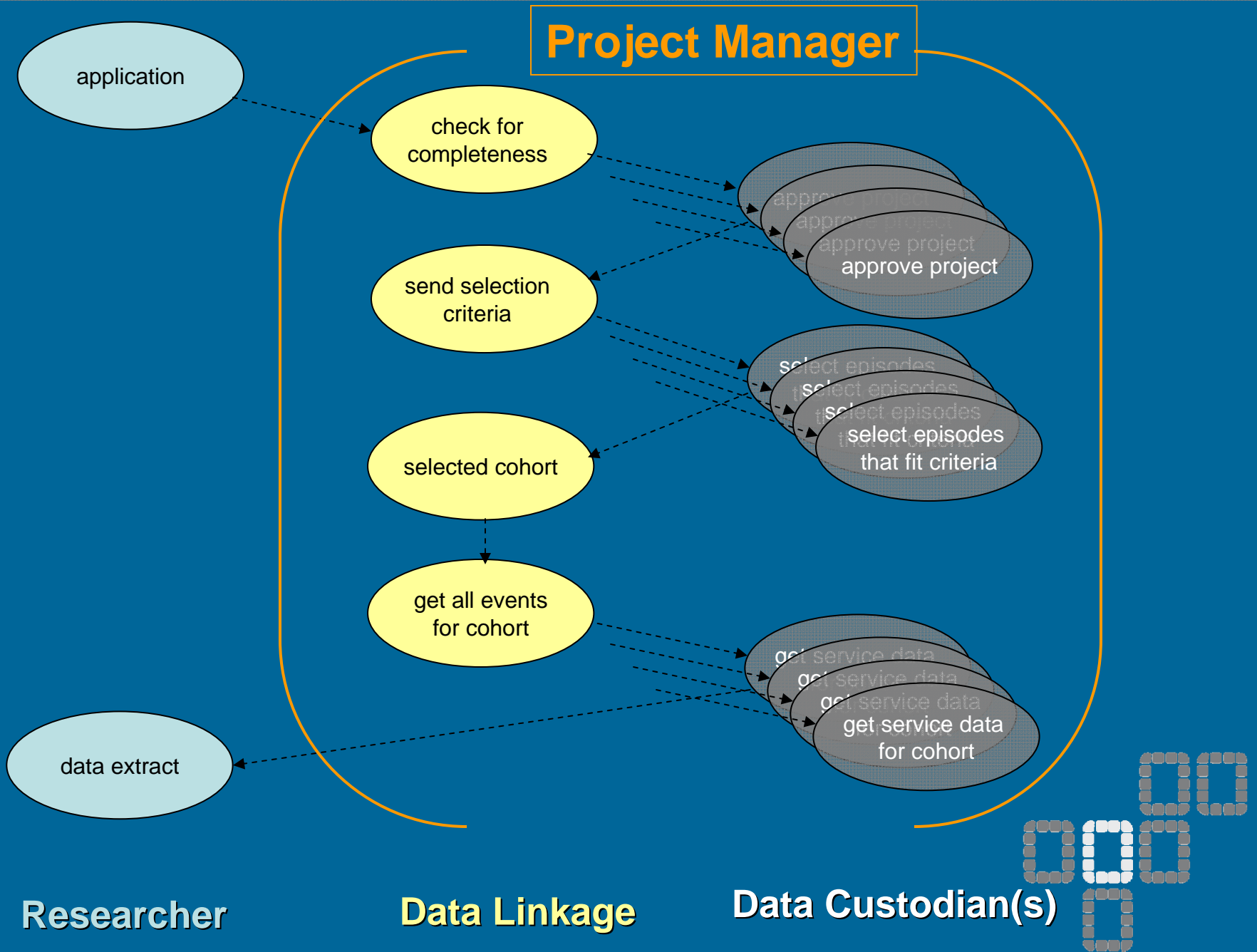
Client Services Group

Areas of expertise:

- Consultation

- Project administration & governance

- Help desk facility via phone & website

- Project Management

- Coordinate modes of data transfer between WADLS & disparate data sources

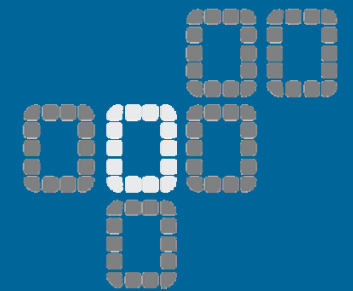- Quality assurance of data extracts.

Project Manager

application

check for completeness

approve project
approve project
approve project
approve project

send selection criteria

select episodes that fit criteria
select episodes that fit criteria
select episodes that fit criteria
select episodes that fit criteria

selected cohort

get all events for cohort

get service data for cohort
get service data for cohort
get service data for cohort
get service data for cohort

data extract

Researcher   **Data Linkage**   **Data Custodian(s)**

Data Linkage Branch

Adoption of the best practice protocol [2]

- DLB physically separated into two groups
- Separation of the personal identifiers from the clinical or service data
- Coordinated effort to manage the supply of service data to research clients through provision of encrypted project specific IDs
- Follow principles of Kelman et al (2002) best practice protocol.
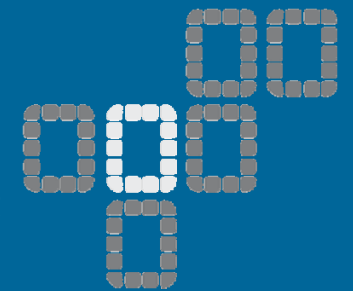
Data delivery

Future directions:

- Currently tailored extracts are provided to approved researchers

- Consideration is being given to the creation of an extract server to house snapshots of standardized data

- Where projects are approved, a linkage key file would be used to join the snapshot of various data extracts

- No personal identifiers to be stored on proposed extract server.

References

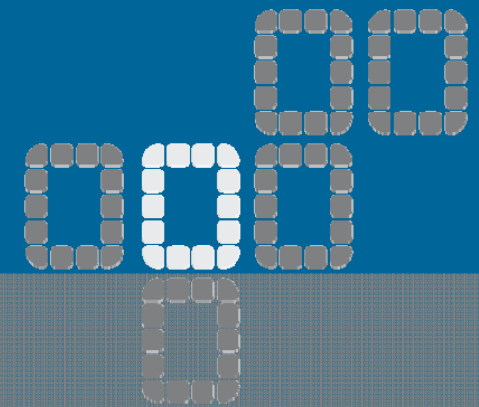[1]Dunn HL, Record Linkage. *Am J Public Health* 1946; 36:1412-1416.

[2] Kelman CW, Bass AJ, Holman CDJ, Research use of linked health data – a best practice protocol. *Aust and NZ J Public Health* 2002; 26 (3): 251-255.

Data Linkage Branch
Public Health Division
Department of Health - Western Australia

www.datalinkage-wa.org.au

Brooke Baxter
Project Manager, Data Linkage Branch
Brooke.Baxter@health.wa.gov.au
phone 08 9222 2370