What is data linkage

- Unique person identifiers are not often available within or across siloed data collections in Queensland and Australia.
- Instead data linkage is used.
- Data linkage is a process that uses person-level identifying information (such as name, date of birth, address) to determine which records within a data source, or between multiple data sources, pertain to a particular individual.
Probabilistic linkage

• Uses the concept of probability to produce a composite score based on weighted values for each element used for linkage

• Relies on completeness and quality of identifiers (names, date of birth, address) – but not all have to be present to enable a match.
Linkage in Queensland Health

- Funded by ongoing National Collaborative Research Infrastructure Strategy (NCRIS) funding, one-off Queensland Government Innovation (DEEDI) funding and Queensland Health funding since 2011
- Queensland Department of Health recurrently funds positions and infrastructure in addition to NCRIS funds to provide an ongoing data linkage service
Linkage in Queensland Health

- Production Linkage – creation of a Master Linkage File (MLF) with enduring linkages between core health data collections which is maintained in near-real time. The MLF currently includes approx. 60M records.

- Request linkage – provision of data from the MLF and custom linkage for ad hoc research and government requests
Key achievements in the past 12 months

- Additional data in MLF – data collections and years
- Enduring linkage keys
- Moving to a new server and version of linkage software
- National work to improve access for cross-jurisdictional projects – working with PHRN/AIHW on processes for approvals and use of enduring linkages
- Client Services team added
- Community of practice website started
## Data Collections linked - MLF

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queensland Hospital Admitted Patient (public and private)</td>
<td>01/2000-current (07/2007 private)</td>
</tr>
<tr>
<td>Emergency Department (public)</td>
<td>07/2008-current</td>
</tr>
<tr>
<td>Perinatal (public and private)</td>
<td>07/2007-current</td>
</tr>
<tr>
<td>Death Registration</td>
<td>01/2000-current</td>
</tr>
<tr>
<td>Birth Registration</td>
<td>07/2007-current</td>
</tr>
<tr>
<td>Elective Surgery Waiting List</td>
<td>07/2014-current</td>
</tr>
<tr>
<td>Outpatient waiting list (public)</td>
<td>07/2015-current</td>
</tr>
</tbody>
</table>
### Data Collections linked - MLF

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery Connect</td>
<td>Up to 05/2018</td>
</tr>
<tr>
<td>Qld Ambulance Service</td>
<td>11/2016 - 03/2018</td>
</tr>
<tr>
<td>Air Retrievals</td>
<td>2017/18</td>
</tr>
<tr>
<td>Non-Admitted Patient (public)</td>
<td>2016/17</td>
</tr>
<tr>
<td>Notifiable Conditions System</td>
<td>01/2017 - Current</td>
</tr>
<tr>
<td>Vaccination Information</td>
<td>Approved</td>
</tr>
<tr>
<td>Community Mental Health</td>
<td>In progress</td>
</tr>
<tr>
<td>Education (NAPLAN, AEDC)</td>
<td>Negotiating access</td>
</tr>
</tbody>
</table>

**Statistical Services Branch, Queensland Health**
Linkage requests received by year and type
Completed requests

Statistical Services Branch, Queensland Health
Research use of linked data

Annual count of peer reviewed publications based on linked data

- 2017-18
- 2016-17
- 2015-16
- 2014-15
- 2013-14
- 2012-13

Count of publications

Queensland
Australia

Statistical Services Branch, Queensland Health
Use of linked data in Queensland Health

- Person-level data rather than event-level data enables:
  - Insight into post-discharge outcomes (e.g., readmissions, deaths), patient journeys, scope of activity for an individual
  - Accurate cost estimates for models of care and care optimisation. Enables value-based service funding.

- QH projects include
  - Chronic disease reports and dashboard (costs, comorb, services)
  - Predictive modelling to target services
  - Disease registers across collections – improved case ascertainment and quality (e.g., congenital anomalies, trauma)
More information

Queensland Health data linkage unit

- email: [HSBresearch@health.qld.gov.au](mailto:HSBresearch@health.qld.gov.au)

Queensland Health research ethics and governance unit

- email: [HIIRO_REG@health.qld.gov.au](mailto:HIIRO_REG@health.qld.gov.au)
AIHW update

There is now a clear process for obtaining access to MBS and PBS data for research projects:

1. AIHW Ethics Committee approval (includes technical assessment application)

2. AIHW seek data custodian approval from Department of Health via a Public Interest Certificate (PIC)

3. AIHW extract and provide data for MBS (from 1984) and PBS (from July 2002)
DOMINO (Data Over Multiple Individual Occurrences) can now be accessed through AIHW. AIHW can also link DOMINO to other datasets.


- AIHW will work with researchers to obtain approval from DSS, again in the form of a PIC
AIHW update

- MBS, PBS and DOMINO data must all be accessed through a secure research environment, currently SURE.
- AIHW now has a technical assessment form which must be signed off by the data linkage unit before submitting to the AIHW and then is part of the ethics submission.
- The AIHW online ethics application system EthOS is being upgraded and this will be completed by the end of the year.
- AIHW data linkage web pages are currently being updated. In the interim, please contact the AIHW linkage unit directly.
National Integrated Health Service Information (NIHSI) – enduring national linkage

• AIHW is creating an enduring linked dataset with national (MBS, PBS, NDI, RAC) and state (hospital admissions, ED, NAP) data collections to enable person-level analysis across services and over time
• QH has provided 5 years of hospital data – annual updates
• Deidentified data will be available for approved projects (health service policy, planning, evaluation) via a remote secure access environment - initially to govt. departments providing data but will be available for researchers
<table>
<thead>
<tr>
<th>TOPIC</th>
<th>PRESENTER</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data linkage output tips and online resources available to help analysts working with linked data</td>
<td>Dr Ronald Webster and Dr Dale Steinhardt, Statistical Analysis and Linkage Unit, Queensland Health</td>
<td>9:10</td>
</tr>
<tr>
<td>Partners in Prevention: Linking police, ambulance and health data to understand and enhance first responses to suicide crisis situations</td>
<td>Carla Meurk, Principal Researcher, Queensland Centre for Mental Health Research and Project Manager, Queensland Forensic Mental Health Service, Queensland Health</td>
<td>9:30</td>
</tr>
<tr>
<td>End of life hospital costs for patients with dementia</td>
<td>Megan McStea, Centre for Health Services Research and Centre for Research in Geriatric Medicine, The University of Queensland</td>
<td>10:00</td>
</tr>
<tr>
<td>Education data holdings and processes for requesting data for research and linkage projects</td>
<td>Dr Michelle Nesic, Principal Statistical Officer, Performance and Reporting, Queensland Department of Education</td>
<td>10:20</td>
</tr>
<tr>
<td>Born in Queensland Study</td>
<td>Macarena San Martin, Institute for Social Science Research, The University of Queensland</td>
<td>10:40</td>
</tr>
<tr>
<td>Data linkage for identifying costs and inequalities: two case studies from maternity and cancer care</td>
<td>Associate Professor Emily Callander, Principal Research Fellow – Health Economics, Australian Institute of Tropical Health and Medicine, James Cook University</td>
<td>11:00</td>
</tr>
<tr>
<td>Accelerated Chest pain Risk Evaluation (ACRE) project</td>
<td>Sarah Ashover, Senior Project Officer, ACRE Project, Queensland Health</td>
<td>11:30</td>
</tr>
</tbody>
</table>