What is Cross-Jurisdictional Data Linkage?

Bringing together records from different systems across states/territories (or Commonwealth) related to the same individual (or ‘entity’)
Now I would like to propose a toast to our glorious Federation... the beautiful coming together in perfect harmony of all the states a man has ever got himself in during the course of his life...

Why do cross-jurisdictional data linkage?
Complete patient pathways, care and outcomes

Longitudinal studies & population mobility

Over time (5yrs), people move

Source: PHRN POC1
Cross-border service use

Source: Australian Bureau of Statistics

Rare Conditions & Outcomes

RD CATEGORY PREVALENCE

Source: Curtin University

Rare Disease prevalence data from [source].
Why do cross-jurisdictional data linkage?

- Increase statistical power for research on rare conditions or outcomes
- Accurate data for longitudinal studies
  - People move or die interstate, seasonal or out of state workers (FIFO)
- Ascertain complete patient pathways, care and outcomes (welcome to Federation)
- Assess cross-border service utilisation
- Evaluate state-based variations in hospital and health care

Examples of cross-jurisdictional studies

- Proof of Concept 1
- Proof of Concept 2
- Proof of Concept 3
- Proof of Concept 4

Proof of Concept 1

- In-hospital and post-discharge mortality: learning about quality of care using data linkages from four Australian states
- XJ linkage of hospital records & death registrations, WA, SA, NSW & Qld
- Outcomes: Refined SMRs, cross-border flows, completeness of patient pathways

### POC#1: Data cleaning by state

<table>
<thead>
<tr>
<th>Excluded episodes of care (hospital records)</th>
<th>NSW</th>
<th>WA</th>
<th>QLD</th>
<th>SA</th>
<th>Total excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential funding duplicate records – one kept</td>
<td>31,345</td>
<td>0</td>
<td>9,358</td>
<td>964</td>
<td>20,903</td>
</tr>
<tr>
<td>Missing age at admission</td>
<td>842</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>842</td>
</tr>
<tr>
<td>Missing sex or intersex</td>
<td>199</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>211</td>
</tr>
<tr>
<td>Missing principal diagnosis</td>
<td>28,738</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28,738</td>
</tr>
<tr>
<td>Newborn³, boarder, organ procurement care types</td>
<td>604,128</td>
<td>36,473</td>
<td>427,795</td>
<td>22,042</td>
<td>1,090,438</td>
</tr>
<tr>
<td>Non-hospital facility (rehab, nursing, hospices)</td>
<td>140,416</td>
<td>0</td>
<td>13,447</td>
<td>1,879</td>
<td>164,742</td>
</tr>
<tr>
<td>Non-medical hospital encounter (Z76)</td>
<td>1,899</td>
<td>4,137</td>
<td>1,496</td>
<td>760</td>
<td>8,292</td>
</tr>
<tr>
<td>Cancelled procedures (Z53)</td>
<td>1,609</td>
<td>0</td>
<td>35</td>
<td>0</td>
<td>1,644</td>
</tr>
<tr>
<td>Missing, unknown or overseas postcodes</td>
<td>84,888</td>
<td>8,412</td>
<td>30,788</td>
<td>3,267</td>
<td>127,355</td>
</tr>
<tr>
<td>Separation date &lt; admission date or missing</td>
<td>125</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>125</td>
</tr>
<tr>
<td><strong>Total excluded</strong></td>
<td><strong>903,189</strong></td>
<td><strong>49,025</strong></td>
<td><strong>482,925</strong></td>
<td><strong>28,915</strong></td>
<td><strong>1,443,290</strong></td>
</tr>
</tbody>
</table>

³ Inclusion of newborns requires special considerations due to unique data handling methods.
## POC#1: Coding variation between states

<table>
<thead>
<tr>
<th></th>
<th>NSW</th>
<th>WA</th>
<th>QLD</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Day case chemotherapy (Z51.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public hospital</td>
<td>26,784</td>
<td>11.5</td>
<td>150,826</td>
<td>53.5</td>
</tr>
<tr>
<td>Private hospital</td>
<td>206,693</td>
<td>88.5</td>
<td>131,011</td>
<td>46.5</td>
</tr>
<tr>
<td>Day case renal dialysis (Z49.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elective</td>
<td>1,606,258</td>
<td>91.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Emergency</td>
<td>12,581</td>
<td>0.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Not assigned</td>
<td>136,494</td>
<td>7.8</td>
<td>729,104</td>
<td>100</td>
</tr>
<tr>
<td>Single delivery (O80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elective</td>
<td>25,798</td>
<td>30.4</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Emergency</td>
<td>2,155</td>
<td>2.5</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Not assigned</td>
<td>57,029</td>
<td>67.1</td>
<td>23,983</td>
<td>100.0</td>
</tr>
</tbody>
</table>

ED deaths recorded as in-patient: Yes, No, No, No

---

### Other PoCs

- **Proof of Concept 2** = burden of injury, linkage of hospital, ED & death records, multi-state (NSW, SA, Qld)
- **Proof of Concept 3** = perinatal risk factors & developmental outcomes, multi-state linkage & AEDC
- **Proof of Concept 4** = vaccination update, effectiveness and burden of infection, linkage of NSW & WA (hospital, ED, perinatal) to ACIR.
Examples of other cross-jurisdictional studies

- Marfan disease (rare disease)
- Epilepsy study (registry)
- Continuity of Care (popn, $ NHMRC 2015)
- SHIP study (cohort, $ NHMRC 2016)

Marfan syndrome

- Describe epidemiology of this rare condition
- Estimate incidence & prevalence
- Examine hospital use, comorbidities, mortality
- Multi-state: WA, NSW, SA
- Cohort = ICD9 759.82; ICD10 Q87.4
- Linkage: Hospital admissions & Death registrations
- CIs = Brameld et al (Curtin University)
Epilepsy study

- Follow-up study of epilepsy patients in 2 state-based Epilepsy Registers
- Examine risk for mortality
- Explore patterns of AED use, co-morbidities
- Examine context of mortality (qualitative)
- CIs = D’Souza et al (UniMelb)
SHIP study

- Follow-up study of “SHIP” participants
- SHIP = large, national survey of people with psychotic illness
- Examine risk for mortality & morbidity
- Estimate economic & social costs
- NHMRC, 3 year project
- CIs = Morgan et al (UWA)
Continuity of Care

- Evaluate effects of coordinated health services
- Role of primary healthcare
- Assess patterns of care (primary & secondary)
- Outcomes for chronic / complex conditions
- Potentially preventable hospitalisations
- System costs
- NHMRC, 4 year project (2015-2018)
- Cis = Moorin et al (Curtin University)

Project Design & Data:
- Design: Whole of population, longitudinal
- Cohort: identified using Medicare enrolments
- Outcomes: from State (WA) & Commonwealth data

Datasets collections required:
- Medicare Benefit Scheme (MBS)
- Hospital Morbidity (WA)
- Emergency Department (WA)
- Deaths (WA)
Progress so far…

- Approval processes:
  - Curtin ethics - approved
  - AIHW Ethics - approved
  - WA Data Application - completed
  - Commonwealth Risk assessment - nearly there

- Data linkage requirements:
  - Combine State & Commonwealth data
  - Navigate Legal and Regulatory Barriers
  - State/AIHW data disclosure requirements
What are the challenges of cross-jurisdictional DL studies?

Challenges

- **Assessing feasibility** – data availability, quality, what’s possible & what’s not?
- **Approvals & application processes** – how many, how long?
- Wait times and cost
- Analysis – larger datasets, coding variations, comparability
Where do you start?

Resources

- PHRN Website (phrn.org.au)
  - General information
  - Processes – Ethics & Data Custodian approvals
  - Available datasets
  - Metadata links
- PHRN On-line Application system
  (https://oas.phrn.org.au)
- DLU Websites
- Human help (email or call with queries)
  cdli@curtin.edu.au  vdl@dhhs.vic.gov.au  phrn@uwa.edu.au
- SUFEX  - transferring data securely
- SURE  – analysing data securely
What datasets are available?

Check the PHRN Website
phrn.org.au & DLU websites

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>Avail Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Commonwealth</strong></td>
<td></td>
</tr>
<tr>
<td>ACD</td>
<td>1982+</td>
</tr>
<tr>
<td>MBS</td>
<td>2012</td>
</tr>
<tr>
<td>PBS</td>
<td>2012</td>
</tr>
<tr>
<td>NDI</td>
<td>1980</td>
</tr>
<tr>
<td>ACIR</td>
<td>1996</td>
</tr>
<tr>
<td><strong>Other National</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Victoria</strong></td>
<td></td>
</tr>
<tr>
<td>ED</td>
<td>2000</td>
</tr>
<tr>
<td>Admitted patients</td>
<td>1993</td>
</tr>
<tr>
<td>Cancer registry</td>
<td>2003</td>
</tr>
<tr>
<td>Deaths</td>
<td>1993</td>
</tr>
<tr>
<td><strong>WA</strong></td>
<td></td>
</tr>
<tr>
<td>Births</td>
<td>1982</td>
</tr>
<tr>
<td>Perinatal</td>
<td>1980</td>
</tr>
<tr>
<td>Hospital</td>
<td>1970</td>
</tr>
<tr>
<td>ED</td>
<td>2002</td>
</tr>
<tr>
<td>Cancer registry</td>
<td>1982</td>
</tr>
<tr>
<td>Death</td>
<td>1969</td>
</tr>
<tr>
<td>Mental Health</td>
<td>1969</td>
</tr>
</tbody>
</table>
Approvals processes

- At a minimum, institutional HREC
- If State-level: Data Custodian (DoH++) & HREC
- If Commonwealth, then AIHW EC & Commonwealth custodian approval(s)
- OAS Stages: RFQ, EoI, Full data application, HREC

**Advice:**
For HREC, use NEAF if possible (reusable). Check reciprocal arrangements.
Consent involved? Check wording with Cwth DHS
Complexity = longer timeframes & higher cost
Complexity – new linkage? Cases & controls? No & type of datasets?

---

**FAQs**

Q. Can I get access to Commonwealth data?
A. Which dataset? Time period? Linkage involved? Consented or unconsented? High risk project?

Q. What is SURE and do I have to use it?

Q. How do I use SUFEX?
A. Register via support@sufex.org.au. Free (for now).

*Any other questions? Please ask 😊*
What is SUFEX?

- A secure file transfer service for the PHRN and its stakeholders
- Uses secure online software to send and receive files from anywhere at anytime
- Easy to use, web-based
- It is not a file storage solution
- The service is provided (free!) but registration required
- Hosted and maintained by the CDL
- A secure remote computing environment for research using linked health data
- Operated by the Sax Institute
- SURE replaces current researcher computing environment only
access to SURE

- remote access to SURE is strongly authenticated
- additional means of authentication in addition to a username and secret password
  - one-time-use access code provided by a hardware token (Yubikey) or smartphone

Challenges aside...

Notwithstanding the challenges, working with linked data can be...

sweet, rich and delicious...

...just like chocolate!