

Productivity Commission Draft Report

Data availability and use

Comments from Associate Professor Rebecca Mitchell

Australian Institute of Health Innovation, Macquarie University

The investment by the Australian government to generate and enhance health record linkage infrastructure in Australia has been seen as world-leading. However, for academics, there are several areas that could be enhanced to expedite the approval process to conduct research that involves national health record linkage and to gain timely access to data for research purposes.

The examples provided of the length of research timeframes in Box 1 of the Draft Report on page 6 are not uncommon. One example from my own research is a linkage of national child injury-related hospitalisation data to mortality data. The Australian Institute of Health and Welfare (AIHW) is linking national hospital admission data from the National Hospital Morbidity Database to the National Death Index (NDI) for this research. To probabilistically link health records to other datasets, unique person identifiers are commonly used. Below are a couple of the access and timeliness issues for this child injury research project are described.

Unique person identifiers

A unique person identifier (e.g. first name, surname, address, date of birth, age, gender) currently do not exist on the National Hospital Morbidity Database. To link the National Hospital Morbidity Database to the NDI, each state and territory health department must provide unique person identifiers from their hospitalisation data collection to the AIHW, along with a state unique identifier. Before this can occur, the academic must get approval from all state and territory health data custodians, approval from each state and territory Human Research Ethic Committee (HREC), and approval from the AIHW HREC.

HREC approval

For national research that involves health record linkage, having one national HREC that has the jurisdiction to provide ethics approval, or if each jurisdiction-based HREC would **recognise an existing HREC approval**, would be ideal. Currently, using the National Hospital Morbidity Database to NDI linkage example, I was able to obtain ethics approval from all jurisdictions (except Western Australia) and the AIHW HREC in 9 months and this is considered to be a 'best case' scenario, as the AIHW HREC only meets 4 times a year.

The Queensland Health HREC was the only HREC that ratified an existing HREC approval, while the University of Tasmania HREC had an expedient approval process, if a researcher already had one HREC approval. All other states required a full HREC review and the South Australian HREC also required the completion of a Site Specific Assessment (SSA). Clinical trials are able to get HRECs to comply with the NHMRC's national approach to single ethics

review, so it would be helpful if the single ethics review approach was also identified as applicable to research that involved record linkage and was promoted to the HRECs.

Research timeliness

For the child injury research project described earlier that involves national data linkage by the AIHW Data Linkage Unit, each jurisdiction's health data application process, the HREC approval process in multiple jurisdictions, the waiting period for each jurisdiction to provide their unique person identifiers to the AIHW for linkage, the legal review of confidentiality agreements between the academic researcher's institution and the health departments in each jurisdiction, the conduct of the linkage and the provision of data extracts to the researcher for analysis all take extended periods of time to occur. This impacts on the timeliness of the research results. The example of linking the National Hospital Morbidity Database to the National Death Index to examine child injury in Australia is **still ongoing**, taking over **3 years**. All of the data extracts have still not been provided to the academics.

In another example, the approval process for a case-control study examining adult hospitalised injury in NSW, South Australia and Queensland that I am undertaking took **4 years to achieve**. I am just now publishing some of the results in academic peer-review journals (See: Mitchell, R. Cameron C. McClure R. Quantify the morbidity and mortality attributable to traumatic injury using a population-based matched cohort in Australia. [BMJ Open](https://doi.org/10.1136/bmjopen-2016-013266) 6 (12) e013266 doi:10.1136/bmjopen-2016-013266), however one comment from a peer-reviewer that I received was:

"1. Data is 7 years old. This should be discussed in the limitations section."

Currently, I spend more time filling in paperwork and getting approvals to access linked data extracts and conduct research that involves record linkage than I do in conducting the actual research. **The length of time for the approval process for research involving record linkage is impacting on the timeliness and the relevance of the research involving record linkage being conducted in Australia.**

Data use can improve social outcomes (page 69 in the Draft Report).

Use of linked data is also able to contribute towards informing public health policy. Linkage of police-reported road crash data to mortality data, emergency department presentation and hospital admissions data in NSW was used to examine several areas of road safety, including:

Bambach M. Mitchell R. Grzebieta R. (2013) The protective effect of roadside barriers for motorcyclists. [Traffic Injury Prevention](https://doi.org/10.1186/1745-2875-14-756) 14, 756-765.

This research examined single motorcyclist collisions and found that motorcyclist collisions into various roadside hazards (i.e. trees, posts and utility poles) compared to guardrails resulted in a greater number of serious injuries to motorcyclists. However, due to the ability to now review the injuries experienced by motorcyclists colliding with guardrails (by linking the police-reported crashes to the health data collections), this research was able to inform

the Australian Roadside Design Guides (Austroads, 2010) of the need to increase the severity indices for barriers by around 25% for motorcyclists compared to the provisions for motor vehicle collisions.

Mitchell R. Bambach M. Toson B. (2015) Injury risk for matched front and rear seat car passengers by injury severity: An exploratory study. Accident Analysis and Prevention 82, 171-179.

This research examined injury severity risk for motor vehicle occupants located in the front versus the rear seat of the same vehicle in a collision and found that occupants in the rear seat had a higher likelihood of being severely injured compared to occupants in the front seat, suggesting that rear seat occupant protective mechanisms in vehicles should be examined. This research has since been replicated in the United States with similar findings (i.e. Atkinson et al. Paired vehicle occupant analysis indicates age and crash severity moderate likelihood of higher severity injury in second row seated adults in frontal crashes, 2016, 89, 88-94).

Page 135 – Draft Recommendation 3.1 – suggest making this more explicit – ie, “All Australian Government agencies should create comprehensive, easy to access data registers (listing both data collections and data variables that are available....”).

Page 140 - Draft Recommendation 3.2 – Linked datasets held by researchers are not currently released as researchers generally are required to sign confidentiality agreements ensuring the security and confidentiality of the linked data extract(s). In NSW, the maximum penalty for disclosure of data under the Health Administration Act 1982 No 135 is six months gaol.

Page 200 - Draft Recommendation 5.1 – Recommend that applying the same method that the Australian Bureau of Statistics (ABS) uses to ‘confidentialise’ some of their data that requires researchers to end up doing multiple statistical iterations of a data extract is not conducted. Researchers still need access to **unit record data**.

Page 201 - Draft Finding 5.2 – In relation to Finding 5.2, there needs to be consistency in the data variables that state and territory health departments will release to researchers. For example, the Victorian hospitalisation data custodian will not release hospital admission or separation dates from the Victorian hospital separation data collection to researchers. This makes calculation of certain information, such as 30-day mortality from the date of hospital admission or 28-day hospital readmission impractical in a national study.

Page 206 - Draft Recommendation 5.3 – In relation to enabling the ongoing use of existing linked datasets, the Centre for Health Record Linkage (CHeReL) provides a list of current research projects that use linked NSW data on their website as do the WA Data Linkage

Unit. It also might be worthwhile reviewing the Australian Data Archive, where previous (unlinked) data sets have been archived - <https://www.ada.edu.au/>

Retention of linked datasets

In terms of practical ways to ensure that improvements to linked datasets are available to subsequent data users: I assume that this relates to multiple data extracts that have been linked with a linkage key added to each data extract, each data extract has been cleaned, organised and re-linked into one dataset for analysis and new researchers would like access to the 'one cleaned dataset'. I would suggest that it would depend on what the research question was and on how the linked data extracts had been cleaned and organised.

It would probably be more practical for a 'new researcher' to work with the existing research team, rather than to simply hand over a clean data extract. In order for the new researcher to understand what 'cleaning' and 'organisation' of a data collection had been conducted, the existing research team would need to provide copies of all their data cleaning, organisation and data analysis syntax. This can sometimes be weeks of work, depending on the number of data extracts that were linked. You would also be reliant on the new researcher having knowledge of the statistical package that had been used to clean and organise the data extracts to be able to review and understand what data cleaning and organisation had been undertaken.

Whether the cleaned dataset is suitable to answer the new research question would also need to be reviewed. For example, a hospital separation data extract consists of multiple patient episodes of care. Each episode of care can be a patient move within (e.g. from a ward to ICU to rehabilitation – a statistical discharge) or a transfer between hospitals. When cleaning a hospital separation data extract, there are multiple ways that statistical discharges within a hospital and transfers between hospitals can be organised to form a 'period of care' (i.e. all episodes of care related to an index hospitalisation until discharge from the health system). Some health conditions or specific types of injuries are associated with multiple statistical discharges and transfers, and some are not. The original researcher may have included (or excluded) patient transfers between hospitals within an index admission period of care, if the patient presented to the second hospital with the same principal diagnosis classification within 1-2 days of the index admission. This type of data organisation decision would impact on the ability of the linked data extract to answer new research questions where counting the number of patient transfers between hospitals was critical. This is why it would be critical for the new researcher to fully understand how the data extracts had been cleaned and organised. It may be worthwhile considering providing access to each of the data extracts with their linkage key and the new researcher conducts their own data cleaning, organisation and re-linkage as per what information they need to answer their research question.

Page 215 - Draft Recommendation 5.5 – It would be useful if the Commonwealth data collections were not restricted to being accessed and analysed within Secure Unified Research Environment (SURE). Australian universities already have secure computer infrastructure to store and analyse data and to also transfer data securely between universities (i.e. Cloudstor). Access to SURE can get quite costly for a research project budget. Project set up fees in SURE are around \$5,000, annual project operation fees are

\$4,620, \$110 each for user training, and annual user fees are \$4,620 for each user, and there is a \$500 fee for study archiving. If you have multiple staff that need to work on and analyse the data, this can get quite costly. I tend to limit the number of staff who work on datasets within SURE to try to keep costs down, this then has the knock-on effect of 'de-skilling' some staff. Whereas, if the data extracts could be retained within the university secure computer infrastructure this problem would be resolved.

Page 237 - Draft finding 6.2 – As well as data being available in an electronic format, data dictionaries/ metadata also needs to be made available.

Page 323 – Different models for interoperability with release authorities

A more coordinated approach to data release to researchers would improve timeliness of data access. Currently, researchers are reliant on a data custodian providing (in a timely manner) the unique record identifiers for their data set securely to the data linkage centre for linkage and, after the linkage is complete, providing the data extract with the linkage key to the researcher. In some instances, data custodians are inexperienced in performing both functions, which results in delays. In addition, data custodians also have competing priorities for their time, which can also result in delays. Having one centre conduct the data linkage of the unique identifiers and a second centre provide the data extract and the linkage key would still provide a separation between personal identifiers and health (or other) information and would seem to be a good alternative, and would be likely to reduce some of the delay experienced by researchers waiting for linked data extracts to be provided.

Page 352 - Draft recommendation 9.3 – Examples of the data collections that are available for linkage in Canada, the United Kingdom, and New Zealand are listed in Table 1. As a researcher, who examines injury and health outcomes across the lifespan, the data collections that I believe are in the national interest to be routinely linked to enable rigorous health services and public policy research and policy development are: ambulance records, emergency department presentations, hospital separations, mortality, births, trauma registries, police-report road crash data, pharmaceutical benefits scheme, Medicare benefits scheme, census data, cancer registry, residential aged care, and aged care assessment team data collections.

There are a couple of issues that have not been discussed in the Draft Report, which are:

Health data release procedures by the AIHW

The AIHW is applying the '1,000 denominator rule' to the release of linked health data from the National Hospital Morbidity Database (i.e. *patient level data should not be provided with a combination of demographic data items that distinguishes groups with an estimated population of less than 1,000*). In the majority of cases, the release of unit record linked health data contravenes this data release policy. This is not a data release policy of the jurisdictions, that regularly provide their jurisdiction-based health data for analysis.

Application of this rule by the AIHW, has meant that the AIHW requires each data custodian to **re-confirm their approval** to the AIHW to provide their jurisdictions' data to a researcher, even after they have provided approval for the research, HREC approval has all been received, and confidentiality agreements have been signed between the jurisdiction health department and the researcher's institution.

Application processes in each jurisdiction

While jurisdiction-based legislation around data accessibility for research needs to be adhered to, there still needs to be a consistent health data access across jurisdictions. The Western Australia (WA) Health Department's 2012 *Information Access and Disclosure policy* requires researchers who need information at the hospital level (i.e. where a hospital is named) to write to each WA hospital Chief Executive (CE) for their permission to access data recorded by their hospital. It is not reasonable to require this of academics, especially as no contact details for the hospitals or CEs are provided by the WA Health Department and up to 109 hospitals (public and private) can need to be contacted. The WA Health Department does not provide any identifying person identifiers outside of their state, so currently does not participate in national linkage studies.

The additional requirement that all research that uses linked data in WA should also have a WA researcher on the team is an unnecessary restriction. If researchers do have any specific data queries regarding WA data, then the relevant WA data custodian could provide advice.

Confidentiality agreements

Prior to the academic receiving any health data, confidentiality agreements are required to be signed between the state or territory health department, the academic and the academic's university. It would be helpful if there was a one-size-fits-all confidentiality agreement between the jurisdictions. Despite jurisdiction-specific privacy legislation, this should be possible to achieve. NSW has the most streamlined approach, with the chief investigator receiving information regarding any data restrictions and signing a confidentiality agreement on behalf of all investigators. South Australia, WA and Victoria all require each investigator to sign a confidentiality agreement. In South Australia, and the language in Section 7.2 requires researchers to be personally and legally responsible (instead of the employer) for data security, if anything untoward happens to the data. This is an unreasonable clause, however, researchers are told that the South Australian Confidentiality Agreement is 'non-negotiable'.

Table 1: Data collections that are available for record linkage in Canada, New Zealand and the United Kingdom

Canada	New Zealand	United Kingdom
<p>Admission, Discharge, and Transfer (ADT) and E-Triage</p> <p>ADT: Adult data covers from Apr 1, 1999 to Dec 31, 2010. Children's data covers from Jun 1, 2006 to Mar 31, 2010.</p> <p>E-Triage: Adult and children's coverage from 2004/05 to 2011/12.</p>	<p>ACC injury claims data</p> <p>Source: Accident Compensation Corporation</p> <p>Time: From 1994</p>	<p>Annual District Birth Extract (ADBE)</p>
<p>Bone Mineral Density</p> <p>1990 to 2014/15 (more complete information beginning in 2000)</p> <p>Winnipeg Regional Health Authority St. Boniface Hospital</p>	<p>B4 School Checks</p> <p>Source: Ministry of Health</p> <p>Time: From 2011</p>	<p>Annual District Death Extract (ADDE)</p>
<p>Burns Registry</p> <p>1975 to 2014</p> <p>Winnipeg Regional Health Authority</p>	<p>Cancer registrations</p> <p>Source: Ministry of Health</p> <p>Time: From 1995</p>	<p>Bowel Screening Wales (BSW)</p>
<p>Cadham Provincial Laboratory</p> <p>1992/93 to July 2010</p> <p>Manitoba Health</p>	<p>Chronic condition/significant health event cohort</p> <p>Source: Ministry of Health</p> <p>Time: From 2007</p>	<p>Breast Test Wales (BTW)</p>
<p>Canadian Primary Care Sentinel Surveillance Network</p> <p>December 1998 to December 2015</p> <p>Manitoba Primary Care Research Network</p>	<p>General Medical Services (GMS) claims data</p> <p>Source: Ministry of Health</p> <p>Time: From 2002</p>	<p>Cervical Screening Wales (CSW)</p>
<p>Cancer Screening Databases</p> <p>Breast Screening: July 1995 to 2015</p> <p>Cervix Screening: January 2001 to 2015;</p> <p>Colon Screening: August 2007 to 2015</p> <p>CancerCare Manitoba</p>	<p>Health Tracker</p> <p>Source: Ministry of Health</p> <p>Time: 2006–13</p>	<p>Congenital Anomaly Register and Information Service (CARIS)</p>
<p>Cardiac Surgical Database</p> <p>April 1995 to March 2015</p> <p>Winnipeg Regional Health Authority</p>	<p>Laboratory claims data</p> <p>Source: Ministry of Health</p> <p>Time: From 2003</p>	<p>Emergency department Data Set (EDDS)</p>
<p>Critical Care / Intensive Care</p> <p>1999/2000 to 2012 for all six hospitals.</p> <p>Health Sciences Centre from 1988/89 onward.</p> <p>Grace Hospital from 1998/99 onward.</p> <p>St. Boniface General Hospital, Victoria, Concordia, and Seven Oaks from 1999/2000 onward.</p> <p>Winnipeg Regional Health Authority</p>	<p>Mortality data</p> <p>Source: Ministry of Health</p> <p>Time: From 1988</p>	<p>National Community Child Health Database (NCCHD)</p>

Canada	New Zealand	United Kingdom
<p>Diabetes Education Resource for Children and Adolescents</p> <p>January 1986 to December 2015</p> <p>Winnipeg Regional Health Authority</p>	<p>National Immunisation Register (NIR)</p> <p>Source: Ministry of Health</p> <p>Time: From 2006</p>	<p>Outpatient Dataset (OPD)</p>
<p>Diagnostic Services Manitoba</p> <p>Winnipeg Lab - 2006 to 2014</p> <p>Westman Labs - 2002 to 2014</p> <p>Diagnostic Services Manitoba</p>	<p>National Needs Assessment and Service Coordination Information (SOCRATES)</p> <p>Source: Ministry of Health</p>	<p>Patient Episode Database for Wales (PEDW)</p>
<p>Drug Program Information Network</p> <p>Fiscal 1995/96 to 2015/16</p> <p>Pre-DPIN pharmacare 1973 to 1995 (aggregate data only, not individual-level)</p> <p>Nursing home drugs from 1985/86 to 2015/16</p> <p>Manitoba Health</p>	<p>National Non-Admitted Patient Collection (NNAPAC) data</p> <p>Source: Ministry of Health</p> <p>Time: From 2007</p>	<p>Primary Care GP dataset</p>
<p>Emergency Department Information System</p> <p>2007/08 to 2012/13</p> <p>Winnipeg Regional Health Authority</p>	<p>Pharmaceutical data</p> <p>Source: Ministry of Health</p> <p>Time: From 2005</p>	<p>UK Health Dimensions</p>
<p>Health Links - Info Sante</p> <p>November 18, 2003 to March 31, 2015</p> <p>Winnipeg Regional Health Authority</p>	<p>PHO enrolment data</p> <p>Source: Ministry of Health</p> <p>Time: From 2003</p>	<p>Welsh Cancer Intelligence and Surveillance Unit (WCISU)</p>
<p>Home Care MDS Assessment</p> <p>2002 to 2010/11</p> <p>Winnipeg Regional Health Authority</p>	<p>Population cohort demographics</p> <p>Source: Ministry of Health</p> <p>Time: From 2004</p>	<p>Welsh Demographic Service (WDS)</p>
<p>Home Care Utilization (MH MSSP)</p> <p>MSSP and VON 1990/1991 to June 27, 2015 (cumulative).</p> <p>The VON database contains data from 1993 to October 1999; it is no longer updated since this information is contained in the MSSP data.</p> <p>The last task date in the MSSP data is June 27, 2015; this data will no longer be updated.</p> <p>Manitoba Health</p>	<p>Population cohort addresses</p> <p>Source: Ministry of Health</p> <p>Time: 2005 to 2013</p>	
<p>Hospital Abstracts</p> <p>HAUM ICD-8/9-CM 1970/71 to March 31/04</p> <p>DAD/MADE:ICD-10-CA/CCI including non-MB residents served in MB facilities 2004/05 to 2014/15</p> <p>Reciprocal: MB registrants served in other (reciprocating) provinces 2004/05 to 2014/15</p> <p>Manitoba Health</p>	<p>Programme for the Integration of Mental Health Data (PRIMHD)</p> <p>Source: Ministry of Health</p> <p>Time: From 2008</p>	

Canada	New Zealand	United Kingdom
Long Term Care MDS Assessment (WRHA) 2000/01 to 2015/16 Winnipeg Regional Health Authority	Publicly funded hospital discharges – event and diagnosis/procedure information Source: Ministry of Health Time: From 1988	
Long Term Care Utilization June 1973 to 2015/16 (cumulative) Manitoba Health		
Management Information System 1995/1996 to 2014/15 Manitoba Health		
Manitoba Adolescent Treatment Centre September 1985 to September 2015 Manitoba Adolescent Treatment Centre		
Manitoba Cancer Registry Cancer Registry Data: 1984 to 2013 Cancer Treatment Data: 1984 to 2015 Screening Data: Breast: 1996 to 2015; Cervix: 2001 to 2015; Colon: 2007 to 2015 CancerCare Manitoba		
Manitoba Fetal Alcohol Spectrum Disorder 1999 to 2013 Winnipeg Regional Health Authority		
Manitoba Heart Health Survey October 1, 1989 to February 28, 1990 Manitoba Health		
Manitoba Immunization Monitoring System 1986 to June 2016 (coverage is less complete prior to 1990) Manitoba Health		
Manitoba Maternal Serum Screening Program HSC: 1990/91 to 2002/03 CPL: 2002/03 to December 31, 2015 Winnipeg Regional Health Authority (1990/91 to 2002/03) Manitoba Health (2002/03 to present)		
Manitoba Telehealth 2003 to 2005/06 Winnipeg Regional Health Authority		
Medical Services Permanent Medical Statistical file: MB residents served in		

Canada	New Zealand	United Kingdom
<p>MB, PQ, or out-of-country 1970/71 to 2015/16</p> <p>Reciprocal: MB users in out-of-province facilities 1993/94 to 2015/16</p> <p>Nonstatistical file: Out-of-province users in MB and some non-fee-for-services for MB residents in MB 1993/94 to 2015/16 Manitoba Health</p>		
<p>Medicine</p> <p>2006/07 to 2012 for all four hospitals. Winnipeg Regional Health Authority</p>		
<p>Mental Health Management Information System</p> <p>1990/1991 to 2015/16 (Note 2008/09 data appears incomplete) Manitoba Health</p>		
<p>Midwifery Summary Reports</p> <p>2001/02 to 2014/15 Manitoba Health</p>		
<p>National Ambulatory Care Reporting System (NACRS)</p> <p>April 1, 2014 to March 31, 2015 Manitoba Health</p>		
<p>National Rehabilitation Reporting System</p> <p>2007 to September 2015 Manitoba Health</p>		
<p>Provincial Public Health Statistic System</p> <p>1997 to 1999 and 1999/2000 to 2007/2008. No data is available for the period January to March 1999.</p> <p>Manitoba Health</p>		
<p>Supportive Housing</p> <p>2004 to 2010/11 Winnipeg Regional Health Authority</p>		
<p>Trauma Registry</p> <p>2000 to 2012/13 Winnipeg Regional Health Authority</p>		
<p>Urgent Care (NACRS)</p> <p>2001/2002 to 2004/05</p> <p>2010/11 to 2012/13 Winnipeg Regional Health Authority</p>		
<p>Registries</p> <ul style="list-style-type: none"> • Manitoba Health Insurance Registry 1970/71 to June 2016 Manitoba Health • Metis Population Database 2009 Manitoba Metis Federation • Provider Registry 1978, 1991 to 1992 and quarterly March 1993 to June 2016 Manitoba Health • Vital Statistics Mortality 1970-2015 Vital Statistics 		

Data collection lists from the following websites:

<http://www.saildatabank.com/data-dictionary/sail-datasets>

http://umanitoba.ca/faculties/health_sciences/medicine/units/chs/departmental_units/mchp/resources/repository/datalist.html

http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/idi-data.aspx#health