Addressing the challenges of cross-jurisdictional data linkage between a national clinical quality registry and government-held health data

Nadine E. Andrew,[1] Vijaya Sundararajan,[2] Amanda G. Thrift,[1] Monique F. Kilkenny,[1,3] Judith Katzenellenbogen,[4,5] Felicity Flack,[5] Melina Gattellari,[6,7] James H. Boyd,[8] Phil Anderson,[9,10] Brenda Grabsch,[3] Natasha A. Lannin,[11] Trisha Johnston,[12] Ying Chen,[13] Dominique A.Cadilhac[1,3]

Data linkage brings together information belonging to an individual, family, place or single event from two or more independent data sources.[1] Advances in electronic recording of health information have increased the potential for the secondary use of health data collected primarily for clinical purposes. While many state jurisdictions in Australia now provide data linkage services to researchers, linkage of cross-jurisdictional data-sets is under-utilised.[2]

Data used for monitoring and evaluating health service performance and patient outcomes are collected using multiple systems such as: health information systems; routine audits; and clinical registries. Hospital administrative datasets provide general clinical information on all patient contacts, often record multiple diagnosis codes and procedures and differentiate between acute, non-acute and sub-acute types of care delivered and the number of admissions. On the other hand, clinical registries and audits collect data on processes of care which are not often routinely recorded in other systems. Some national registries, such as the Australian Stroke Clinical Registry (AuSCR)[3]

Abstract

Objective: To describe the challenges of obtaining state and nationally held data for linkage to a non-government national clinical registry.

Methods: We reviewed processes negotiated to achieve linkage between the Australian Stroke Clinical Registry (AuSCR), the National Death Index, and state held hospital data. Minutes from working group meetings, national workshop meetings, and documented communications with health department staff were reviewed and summarised.

Results: Time from first application to receipt of data was more than two years for most state data-sets. Several challenges were unique to linkages involving identifiable data from a non-government clinical registry. Concerns about consent, the re-identification of data, duality of data custodian roles and data ownership were raised. Requirements involved the development of data flow methods, separating roles and multiple governance and ethics approvals. Approval to link death data presented the fewest barriers.

Conclusion: To our knowledge, this is the first time in Australia that person-level data from a clinical quality registry has been linked to hospital and mortality data across multiple Australian jurisdictions.

Implications for Public Health: The administrative load of obtaining linked data makes projects such as this burdensome but not impossible. An improved national centralised strategy for data linkage in Australia is urgently needed.

Key words: data linkage, clinical registry, health data

or the Riksstroke registry, a national stroke registry in Sweden,[4] also include follow-up data such as health-related quality of life (HRQoL) or functional outcomes.

Data collected by each system have their own strengths and limitations. The use of data linkage techniques, through which person-based data from different sources

1. Stroke & Ageing Research, School of Clinical Sciences at Monash Health, Monash University, Victoria
2. Department of Medicine, St. Vincent's Hospital, Melbourne University, Victoria
3. Florey Institute of Neuroscience and Mental Health, Heidelberg, Victoria
4. Western Australian Centre for Rural Health, The University of Western Australia
5. Telethon Kids Institute, The University of Western Australia
6. South Western Sydney Clinical School, University of New South Wales
7. Ingham Institute for Applied Medical Research, New South Wales
8. Population Health Research Network Centre for Data Linkage, Centre for Population Health Research, Curtin University, Western Australia
9. Data Linkage Unit, Australian Institute of Health and Welfare, Australian Capital Territory
10. Faculty of Health, University of Canberra, Australian Capital Territory
11. School of Allied Health, College of Science, Health and Engineering, La Trobe University, Victoria
12. Health Statistics Branch, Queensland Department of Health
13. Victorian Data Linkages, Department of Health and Human Services, Victoria
Correspondence to: Dr Nadine Andrew, Translational Public Health & Evaluation Division, Stroke and Ageing Research, School of Clinical Sciences at Monash Health, Monash University, Level 1/43-51 Kanooka Grove, Clayton, VIC 3168; e-mail: nadine.andrew@monash.edu

are combined, is a logical and efficient method for maximising use of these existing data collections. Data linkage can also provide: (i) a cost effective means of collecting data, thereby reducing participant burden and avoiding duplication of effort; (ii) detailed longitudinal data on outcomes for patients not otherwise available; and (iii) the infrastructure to address clinical research questions that may not be readily investigated using a single dataset. Such data linkages have informed improvements in patient care, healthcare delivery and health service infrastructure.[4-8]

Linking stroke patient data is likely to yield important information about stroke care as between 86% and 96% of patients who suffer a stroke are admitted to hospital.[9-11] Further, these patients are likely to have a significant history of comorbidities and risk factors recorded from admissions prior to their stroke event. Post-stroke hospitalisations due to recurrent vascular events and other conditions either unrelated or related to the patient's stroke event are also common.[12] As the risk of recurrent vascular events, subsequent hospitalisations and post-discharge mortality is high, large amounts of information post-stroke can be captured using linked data. Stroke-based data linkage projects within Australia have already provided valuable information related to stroke care and health service delivery.[7,8,13-16] However, these studies have been restricted to within-state linkages.

The WA Data Linkage System (WADDS) was the first comprehensive state-wide data linkage system developed in Australia in 1995 and provided within state linkages as well as linkages with some Commonwealth-held datasets. Subsequently, most Australian jurisdictions have developed data linkage capacity. However, prior to the Population Health Research Network (PHRN) there was limited capacity in Australia to link data collections across state jurisdictions e.g. between two or more States or between the multiple States and the Commonwealth.[17] The main focus of the PHRN has been to develop a range of facilities and capabilities to enable researchers to access linked data. This included support for state-based linkage units and national linkage facilities through the Centre for Data Linkage (CDL) and Australian Institute for Health and Welfare (AIHW). However, to date only a limited number of projects involving linkages with state held hospital morbidity

data, across multiple state jurisdictions, have been completed in Australia.[18,19] The feasibility of national cross-jurisdictional data linkage between government-held data and a researcher-generated national clinical registry, in which identifiable data are held, has yet to be tested.

In 2012, Cadilhac and colleagues were awarded a four-year National Health and Medical Research Council (NHMRC) Partnerships for Better Health grant (GNT1034415). One of the project objectives was to create, for the first time, a comprehensive integrated national dataset for stroke by obtaining person-level linkages with data from a national clinical registry and emergency, hospital and mortality data across multiple jurisdictions.[20] Here, we describe the methods, approval processes, challenges and solutions involved in achieving this aim.

Methods

Study Cohort

Our study cohort was made up of all AuSCR registrants (n=approximately 17,000 patients) admitted to participating hospitals (n=41) from Queensland, New South Wales, Victoria and Western Australia, between 2009 and 2013. The AuSCR cohort was used to define the index stroke event, this being the first ever stroke episode registered in the AuSCR.

Data Sources for linkage

Five data sources were used in the Stroke123 data linkage study as outlined below. Date ranges for state hospital datasets are outlined in Table 1.

1. **The AuSCR** – A collaborative national effort to monitor, promote, and improve the quality of acute stroke care in Australia (www.auscr.com.au).[3] The AuSCR collects data on all stroke/TIA cases (except sub-arachnoid haemorrhage) admitted to participating hospitals from New South Wales, Victoria, Western Australia, Tasmania and Queensland. Cases are defined by clinicians using a clinical diagnosis

Table 1: Date ranges for hospital datasets.

State	Admitted patient datasets Availability	Emergency department datasets Availability
Queensland	1995-2015	2008-2015
Victoria	1993-2015	2000-2015
New South Wales	2001-2015	2005-2015
Western Australia	1970-2015	2002-2015

of stroke (not the primary discharge International Classification of Disease, 10th Revision (ICD10) codes) and so are not reliant on interpretation by coders. The AuSCR is designed so that prospective data are collected systematically to identify variations in clinical processes of care and health outcomes. Identifying patient information, clinical information and processes of care received by patients with stroke or TIA are entered into an online web form by hospital staff or uploaded into a central database. A minimum core set of data, including four process of care indicators, are collected nationally along with a range of demographic variables and 90-180 day follow-up data (See Table 2).

2. **The National Death Index (Australian Institute of Health and Welfare [AIHW])** – is a listing of all deaths in Australia since 1980 set up specifically for data linkage research. Information on the fact/date of death and underlying and associated cause of death are available. Diagnostic codes as well as text fields can be provided.

3. **Admitted patients (New South Wales, Victoria, Western Australia and Queensland)** – Includes data on all inpatient separations (discharges, transfers and deaths) from all public, private, psychiatric and repatriation hospitals. Each separation ends when the patient is formally separated from the facility, i.e. the patient is either discharged, transferred, dies, or when the principal clinical intent changes within the same period of stay. This may result in multiple separations for a single health event per patient. While the depth of information collected varies between states, all states comply with requirements for the Admitted Patient National Minimum Data Set (NMDS), enabling cross-jurisdictional comparisons of hospitalisation data .[21] Coding of information by trained clinical coders using standardised International Statistical Classification of Diseases and Related Health Problems, Australian Modification (ICD-AM) and Australian Classification of Health Interventions (ACHI) codes and coding rules also ensures consistency between data (see Table 2).

4. **Emergency departments (New South Wales, Victoria, Western Australia and Queensland)** – Include data on presentations to the majority of public emergency departments. Emergency Department data-sets between states

vary in terms of the number and type of variables available as well as the overall reliability and quality of the recorded data, including diagnostic information. Nevertheless a few variables are consistent across datasets such as triage category, type of ED visit and discharge destination (Table 2).

5. *Clinical costings (New South Wales, Victoria and Queensland)* – Clinical costing data can be used to describe the costs of treating stroke and compare changes in costs over time. Data are provided on the direct, indirect and total costs associated with a large number of services such as staffing, intensive care and other specialised care types, imaging, pathology and surgical costs.

Processes

1. A working group of experts was brought together to better understand the data linkage environment in Australia and to guide and facilitate the process. Group members included data linkage experts, database managers, epidemiologists and statisticians. Membership included representation from Victoria, Western Australia, New South Wales, South Australia and Queensland as well as organisations such as the PHRN and the AIHW. This group met monthly to discuss the challenges and processes of the Stroke123 data linkage project and share knowledge to develop effective and efficient methods for data linkage.

2. A detailed protocol was developed in consultation with the working group. This protocol was used to identify common and complementary information available across datasets and to develop important research questions that could be assessed from the integrated data. This enabled us to provide consistent applications to each data custodian.

3. We identified governance clearances and approvals that were required for each dataset. This included identifying the linkage units and data custodians responsible for each dataset, documenting (where possible) the processes involved for each jurisdiction and identifying which jurisdictions required ethics clearances in addition to the already-obtained project ethics approval.

4. An application plan was developed using a staged process. We first applied for linkage to the National Death Index as these data were considered to be most important and, because these are compiled at a national level by the AIHW, linkage was possible with the whole AuSCR cohort through a single application. Prioritisation for the state hospital data was based on Stroke123 project priorities, such as the need for linked data for a Queensland quality of care study, and to those states with greatest number of hospitals participating in the AuSCR.

5. Annual face-to-face workshops were undertaken to share knowledge between members of the expert working group and with researchers and data linkage experts external to the project to work through related issues.

Data collection and analysis

We undertook a review of the processes negotiated and timelines associated with obtaining data containing person level linkages between the AuSCR, the National Death Index, and state-held hospital data.[21] We reviewed minutes from 27 monthly expert working group meetings that occurred between February 2013 and September 2015, and three annual national workshops (2013 to 2015). We also examined email communications between the data linkage unit members of the sub-committee and study staff to confirm the contents of the minutes. Information obtained from these sources was reviewed by authors NA and DC and summarised chronologically to report on the processes, timelines and specific challenges identified during the project and how these were overcome.

Results

Challenges and solutions

Challenge 1: Separation of roles

Clinical registries, such as the AuSCR, are commonly managed within research institutes where the data custodian is also a leading researcher in their field. Unlike data linkage between administrative datasets, where the data are released to independent researchers, linkage involving a clinical registry often means that the data custodian will also be involved in the research. In Stroke123 the principal investigator (PI) is also the AuSCR data custodian. This duality of roles raised privacy concerns with some government data custodians about the potential to re-identify government records.

Variables	AuSCR	Hospital data	ED data
Table 2: AuSCR minimum dataset variables and commonly collected hospital variables.			
Patient name*	√	√	√
Address*	√	√	√
Country of birth*	√	√	√
Date of birth*	√	√	√
Sex*	√	√	√
Hospital code/name*	√	√	√
Indigenous status	√	√	√
Admission date	√	√	√
Discharge date	√	√	√
Discharge destination	√	√	√
ICD10 primary diagnosis codes	√	√	
Next of Kin contact details	√		
Type of stroke	√		
Date & time of stroke inset	√		
Ability to walk independently on admission	√		
First ever (incident) event status	√		
Use of intravenous thrombolysis (tPA)	√		
Access to a stroke unit	√		
Discharged on an antihypertensive agent	√		
Care plan provided at discharge	√		
Date of death	√		
Place of residence at 90-180 days	√		
Living alone status at 90-180 days	√		
Recurrent stroke event within 90-180 days discharge	√		
Readmission to hospital within 90-180 days	√		
Quality of life (EuroQoL5D adults PedsQL children at 90-180 days)	√		
Admission mode		√	
Admission urgency status		√	
Condition onset		√	
Diagnosis related group, code		√	
ICD10 additional diagnosis codes		√	
Procedure codes		√	
Number of days of hospital-in-the-home care		√	
Hospital geographic remoteness code		√	
Hospital region identifier		√	
Hospital sector code		√	
Transport mode to Ed			√
Triage category			√
Triage date & time			√
Type of ED visit			√
ED Waiting time			√

AuSCR: Australian Stroke Clinical Registry; ED; Emergency Department
* linkage variables

Solution: To overcome this we established three levels of separation.

1. *Separation of data.* Within the AuSCR database, de-identified clinical data are stored in a file separate to the identifying data required for performing the 90-180 day follow-up interviews. Separation of data is common practice in data linkage as a method of preserving privacy[17] and was maintained throughout the project. According to best practice data linkage methods, only the personal and clinical identifiers, such as date of admission, that were necessary for the linkage process were submitted to the data linkage units. Only de-identified clinical data with a project-specific person ID number were submitted to the researchers responsible for the analyses.[17]

2. *Separation of roles within organisations.* It was identified that staff that had access to identifying data should not have access to the linked content data to be used for analysis and vice versa. To overcome this, the PI for Stroke123 will not be directly involved in analysing the linked data but instead will provide advice on the analysis and interpretation of the data. For some states such as Western Australia it was necessary to nominate another person to act as data custodian for the project.

3. *Physical separation of staff.* The Stroke123 project is a collaborative project involving a number of institutions. This meant that, for this project, we were able to provide an extra level of separation. The research staff responsible for analysing the linked data will not be provided access to the AuSCR web tool and are located at a different institution to the staff involved in the day-to-day running of the registry.

Challenge 2: Consent

The AuSCR has ethics approval for an opt-out consent approach to recruit registrants consistent with the recommendations in the "Operating Principles and Technical Standards for Australian Clinical Quality Registries".[22] In the AuSCR participant information sheet potential registrants are advised that data collected for the registry may be linked to other health data for research and monitoring purposes, subject to approval from an ethics committee and the AuSCR Research Task Group. The text was based on advice from the New South Wales Population & Health Services Research Ethics Committee and thought to adequately, and sensitively,

cover linkage with a range of relevant health information. The data custodians and ethics committees from New South Wales and Western Australia considered this approach to be too broad to cover the data linkage activities associated with Stroke123.

Solution: In New South Wales and Western Australia we applied for a waiver of consent to allow the linkages between AuSCR and hospital data. The conditions for waiver of consent were met based on the size of the study, the public health importance of this particular research project, and the impracticality of obtaining informed consent.

Challenge 3: Release of identifiable data to a third party

Variations exist between states with regards to their ability to release identifiable data to linkage facilities outside their state. Legislation concerns were raised, by health department data custodians from Victoria and Queensland about submitting identifiable linkage variables to a third party such as the CDL. Further concerns were raised about the potential for re-identification of data because the AuSCR holds identifiable patient information.

Solution: Our team, with assistance from CDL staff, developed a novel, albeit complex, cross-jurisdictional data flow model that will enable linkage of datasets between different states using linkage keys that do not involve the submission of personal identifiers to a third party. This data flow method was acceptable to most States and will provide additional levels of data security to models used previously in that the linkage map will be held by a third party that does not hold identifiable or content data[7] (Figure 1).[17]

Challenge 4: Ownership of data held by the AuSCR

In some jurisdictions the agency collecting the data retains ownership of the data as opposed to the data custodian. For example, in Queensland the hospitals collecting the data that was submitted to AuSCR retained ownership. This meant that AuSCR data collected from this jurisdiction could not be sent, in an identified form, to a third party for additional linkage (or any other purpose) without approval from the collecting agency.

Solution: Additional approvals were sought and granted to allow data from the AuSCR, collected by Queensland hospital staff, to be submitted to the AIHW for linkage with the National Death Index based on a

public interest exemption (section 160 of the Hospital and Health Boards Act 2011). However, data from patients registered in AuSCR by Queensland hospital staff were not able to be submitted to other state data linkage units or the CDL. This means that only within-state linkages will be available for Queensland registrants.

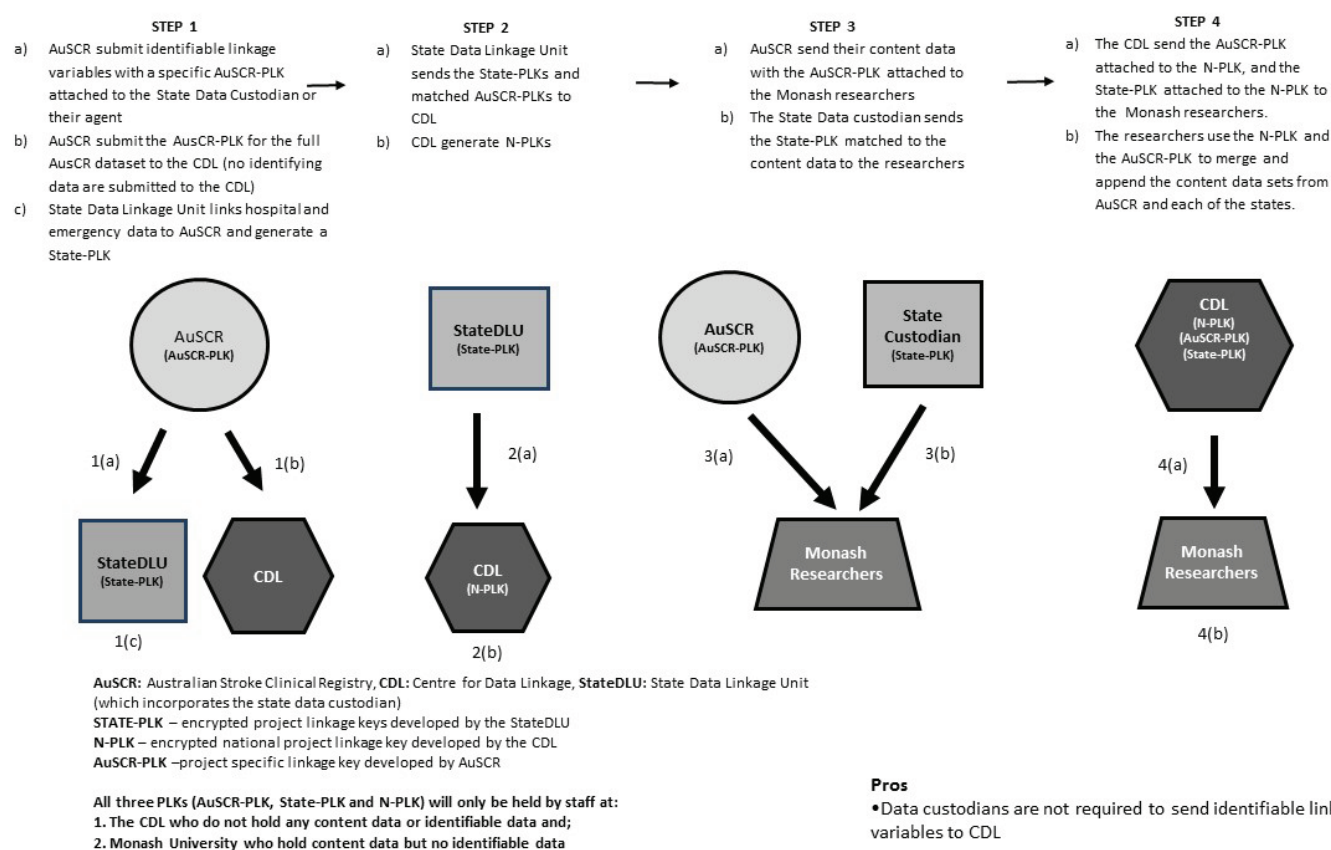### Challenge 5: Timelines and duplication of effort

Each jurisdiction required the completion of separate application forms which were often dissimilar. Various elements of the project were negotiated separately for each state. For example, Western Australia required a detailed research plan for each research question that was likely to be answered using the final data set, whereas New South Wales required a separate ethics amendment for each research question detailing the specific aims, methods and additional ethical issues related to the study, while Victoria required all manuscripts be approved by the Health Department prior to publication. In addition, four separate ethics approvals were required for the project and between two to five documents (median of three) were required to be signed by staff from different divisions of Health Departments for each jurisdiction. Waiting times for document signing were responsible for the greatest delays in our project (Figure 2). Obtaining National Death Index data was the quickest (1.5 years). These data would have been obtained within approximately six months in the absence of delays imposed by the additional approvals required for Queensland.

Solution: Members of the expert working group who were familiar with the workings of health departments were able to provide advice on how best to progress our applications. Having support from the State Stroke Clinical Networks in Queensland and Victoria was also crucial to obtaining sign-off from heads of health departments of these states.

Other potential challenges

Victoria has legislation governing the collection of clinical registry into the Department of Health for availability to third party research, subject to approval. This requirement may create additional challenges depending on whether or not this complies with the operating principles of the registry involved and is acceptable to clinical registry data custodian. Variations also exist

Figure 1: Proposed data flow for national cross jurisdictional data linkage using only linkage keys.

STEP 1
a) AuSCR submit identifiable linkage variables with a specific AuSCR-PLK attached to the State Data Custodian or their agent
b) AuSCR submit the AusCR-PLK for the full AuSCR dataset to the CDL (no identifying data are submitted to the CDL)
c) State Data Linkage Unit links hospital and emergency data to AuSCR and generate a State-PLK

STEP 2
a) State Data Linkage Unit sends the State-PLKs and matched AuSCR-PLKs to CDL
b) CDL generate N-PLKs

STEP 3
a) AuSCR send their content data with the AuSCR-PLK attached to the Monash researchers
b) The State Data custodian sends the State-PLK matched to the content data to the researchers

STEP 4
a) The CDL send the AuSCR-PLK attached to the N-PLK, and the State-PLK attached to the N-PLK to the Monash researchers.
b) The researchers use the N-PLK and the AuSCR-PLK to merge and append the content data sets from AuSCR and each of the states.

AuSCR: Australian Stroke Clinical Registry, **CDL:** Centre for Data Linkage, **StateDLU:** State Data Linkage Unit (which incorporates the state data custodian)
STATE-PLK – encrypted project linkage keys developed by the StateDLU
N-PLK – encrypted national project linkage key developed by the CDL
AuSCR-PLK –project specific linkage key developed by AuSCR

All three PLKs (AuSCR-PLK, State-PLK and N-PLK) will only be held by staff at:
1. The CDL who do not hold any content data or identifiable data and;
2. Monash University who hold content data but no identifiable data

Pros
• Data custodians are not required to send identifiable linkage variables to CDL

---

between states with regard to the identifying, demographic variables available for linkage. For example, Victoria is the only state to routinely collect patient medicare numbers in addition to names and addresses. The methods used to link the data and the level of data provided to researchers about the accuracy of the linkages also varies between jurisdictions. For example, Health Department staff members from Victoria use stepwise deterministic matching whereas in New South Wales they use probabilistic matching.

Discussion

Achieving cross-jurisdictional data-linkage between a non-government national clinical registry holding identifiable data and state and commonwealth held data was challenging, yet achievable. An iterative process was required to respond to the different legislative and data custodian requirements that emerged as applications progressed through the system. We found the greatest challenge to be negotiating the ever-changing political and legislative framework associated with accessing data from multiple jurisdictions. Since commencing Stroke123,

governments in most states have changed. This has resulted in new interpretations of the same privacy laws as well as changes to the willingness of states to share data, usually to a more conservative approach. These changes meant that methods applied in prior work on

which the feasibility of our project was based, were no longer relevant. Uncertainties about access to data and the prolonged timelines required for approval make it very difficult for researchers to plan for projects that involve cross-jurisdictional data-linkage. Having the

Figure 2: Stroke123 data linkage time line.
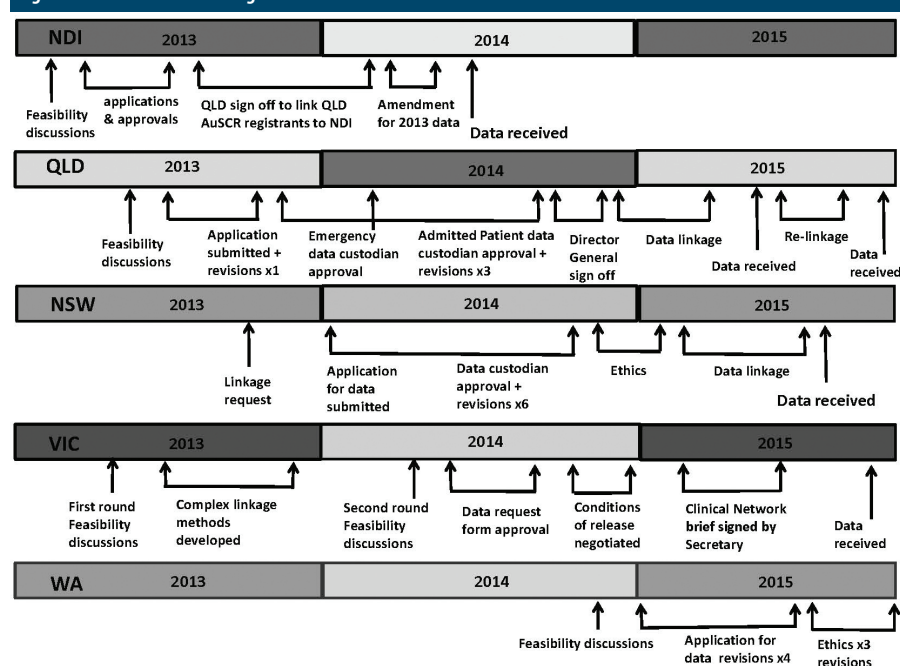
support of an expert working group and developing a collaborative approach with data custodians and health departments was a critical component in negotiating this landscape and advancing the data linkage work of Stroke123.

To our knowledge, Australia is the only federated country that has national cross-jurisdictional capability for data linkage. However, national data linkage projects in Australia can be significantly hampered by the different interpretations of privacy laws and multiple applications required by the different jurisdictions. National data linkage is more frequently used in countries such as New Zealand or Scandinavian countries, that do not have a federated system, have a number of population-based registries, have well developed integrated data infrastructure to allow streamlining of data linkages and have a single unique identifier that can be used to link person-level data across datasets.[23-25] For example, unlike Australia, New Zealand has a unique National Health Index number (NHI number) that is applied to each person that utilises health and disability services within New Zealand. This number is used to easily and routinely link a number of health system databases. In some situations these routine national databases are annually linked to clinical cohorts, not dissimilar to our AuSCR cohort.[26]

Our experience has highlighted that a more centralised approach, similar to that used in New Zealand, is needed to enhance national data linkage in Australia (linkage between the AuSCR and the National Death Index was the most efficient). State approval processes are different to that used by the NDI, State and Territory Registrars of Births, Deaths and Marriages delegate approval for specific NDI linkage projects to the AIHW Ethics Committee under agreed protocols which includes AIHW reporting back on all projects, (and having a representative of the Registrars sit on the Ethics Committee). This resulted in a more centralised and efficient process and meant that these data were immediately available for analysis and reporting at a national level.[27] Hospitalisation datasets are different in that they are primarily used by state departments for the purposes of funding and resource allocation. The extent to which the NDI model can be extrapolated to hospital datasets is unknown. However, it may be possible to create a centralised resource for epidemiological and research purposes.

Longer-standing data linkage units had more formalised and well-documented application procedures, resulting in more specific guidance that facilitated an efficient approval process despite there being a greater amount of paperwork. Our results indicate that a centralised, consistent and formalised approach to national data linkage is likely to improve the efficiency of national data linkage in Australia.

It has taken three years to obtain approval for our applications and at the time of writing we are yet to receive all the requested data. These timelines are similar to those described by Spilsbury et al for their cross-jurisdictional data linkage project in four states (i.e. four years).[18] We staggered our state-by-state submissions sequentially in the hope that experience gained from an earlier submission would inform and streamline the next submissions. In hindsight, the lack of consistency between application processes was of little benefit since different issues were encountered for each jurisdiction. Often the lessons learned from the initial applications did not apply to subsequent applications, and this further highlights the need for a centralised national approach.

In response to delays experienced by our cross-jurisdictional data linkage work and the work of others,[18,19] the PHRN has developed, with assistance from state and commonwealth data linkage staff across Australia, a uniform on-line cross jurisdictional application form (www.phrn.org.au/). Advice on the feasibility and cost of potential projects and the data collections held by each state are also available. While individual approvals will still need to be negotiated separately with each state, this should facilitate a more streamlined approach. To enhance opportunities for data linkage we require consistency in state legislation and increased investment in data linkage infrastructure. Increases in funding for Data Linkage Units at both a state and Commonwealth level are also required, especially as the feasibility, quality and demand for linked data continues to increase. Without such initiatives project delays are also likely to increase.

Conclusion

Data obtained through linkage of person level data between a national clinical quality registry and health department administrative datasets can be used to address important research questions. For stroke, use of linked data to answer these questions will help guide policy, improve the quality of stroke care received by patients, identify missed stroke prevention opportunities and ultimately improve patient outcomes. Achieving timely routine linkages of health data is imperative if we are to maximise the use of routinely collected data to guide and improve health service delivery in Australia. Continued streamlining of processes and greater collaboration between jurisdictions, with increased infrastructure and funding support from governments are urgently needed.

References

1. Hobbs M, McCall M. Health Statistics and Record Linkage in Australia. *J Chronic Dis*. 1970;23:375-81.
2. Olver IN. Linking data to improve health outcomes. *Med J Aust*. 2014;200:368-9.
3. Cadilhac DA, Lannin N, Anderson C, Levi C, Faux S, Price C, Middleton S, Lim J, Thrift AG, Donnan GA. Protocol and pilot data for establishing the Australian Stroke Clinical Registry (AuSCR). *Int J Stroke*. 2010; 5:217-26.
4. Asplund K, Hulter Asberg K, Appelros P, Bjarne D, Eriksson M, Johansson A, et al. The Riks-Stroke story: Building a sustainable national register for quality assessment of stroke care. *Int J Stroke*. 2011;6:99-108
5. Grau AJ, Eicke M, Biegler MK, Faldum A, Bamberg C, Haass A, et al. Quality monitoring of acute stroke care in Rhineland-Palatinate, Germany, 2001-2006. *Stroke*. 2010;41:1495-500.
6. Murphy P, Smith D, Vislosky M, Tripputi M, Cumbler E, Blakie J, et al. The impact of a statewide stroke quality improvement program: Colorado Stroke Alliance and Colorado Stroke Registry. *J Neurosci Nurs*. 2011;43: 246-52.
7. Gattellari M, Goumas C, Aitken R, Worthington J. Outcomes for patients with ischaemic stroke and atrial fibrillation: The PRISM study (A Program of Research Informing Stroke Management). *Cerebrovasc Dis*. 2011;32:370-82.
8. Sanders LM, Srikanth VK, Jolley DJ, Sundararajan V, Psihogios H, Wong K, et al. Monash transient ischemic attack triaging treatment: Safety of a transient ischemic attack mechanism-based outpatient model of care. *Stroke*. 2012;3:2936-41.
9. Thrift AG, Dewey HM, Macdonell RA, McNeil JJ, Donnan GA. Stroke incidence on the east coast of Australia: The North East Melbourne Stroke Incidence Study (NEMESIS). *Stroke*. 2003;1:2087-92.
10. Leyden JM, Kleinig TJ, Newbury J, Castle S, Cranefield J, Anderson CS, et l. Adelaide stroke incidence study: Declining stroke rates but many preventable cardioembolic strokes. *Stroke*. 2013;44:1226-31.
11. Islam MS, Anderson CS, Hankey GJ, Hardie K, Carter K, Broadhurst R, et al. Trends in incidence and outcome of stroke in Perth, Western Australia during 1989 to 2001: The Perth Community Stroke Study. *Stroke*. 2008;39:776-82.
12. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): A case-control study. *Lancet*. 2010;376:112-23.
13. Gattellari M, Goumas C, Garden F, Worthington JM. Relative survival after transient ischaemic attack: Results from the Program of Research Informing Stroke Management (PRISM) study. *Stroke*. 2012;43:79-85.
14. Katzenellenbogen JM, Begg S, Somerford P, Anderson CS, Semmens JB, Codde JP, et al. Disability burden due to stroke in Western Australia: New insights from linked data sources. *Int J Stroke*. 2010;5:269-77.

15. Katzenellenbogen JM, Vos T, Somerford P, Begg S, Semmens JB, Codde JP. Burden of stroke in indigenous Western Australians: A study using data linkage. *Stroke*. 2011;42:1515-21.

16. Kilkenny MK, Dewey H, Sundarrarajan V, Andrew NE, Lannin N, Anderson C, et al. Readmissions after stroke: Linked data from the Australian Clinical Stroke Registry (AuSCR) and hospital administrative databases. *Med J Aust*. 2015;203:102-7.

17. Boyd J, Ferrante A, O'Keefe C, Bass A, Randall S, Semmens J. Data linkage infrastructure for cross-jurisdictional health-related research in Australia. *BMC Health Serv Res*. 2012;12:480.

18. Spilsbury K, Rosman D, Alan J, Boyd JH, Ferrante AM, Semmens JB. Cross border hospital use: Analysis using data linkage across four Australian states. *Med J Aust*. 2015;202:582-6.

19. Mitchell RJ, Cameron CM, McClure RJ, Williamson AM. Data linkage capabilities in Australia: Practical issues identified by a Population Health Research Network 'Proof of Concept project'. *Aust NZ J Public Health*. 2015;39(4):319-25.

20. Cadilhac D, Sundararajan V, Andrew N, Kilkenny M, Flack F, Anderson P, et al. Using linked data to more comprehensively measure the quality of care for stroke – understanding the issues. *Australas Epidemiol*. 2013;20:15-19.

21. Australian Institute of Health and Welfare. *National Minimum Data Sets*. Canberra (ASUT): AIHW; 2015.

22. Australian Commission on Safety and Quality in Health Care. *Operating Principles and Technical Standards for Australian Clinical Quality Registries*. Melbourne (AUST): Monash University NHMRC Centre for Research Excellence in Patient Safety (CRE PS); 2008. p. 131.

23. Meretoja A, Roine RO, Kaste M, Linna M, Juntunen M, Erila T, et al. Stroke monitoring on a national level: PERFECT Stroke, a comprehensive, registry-linkage stroke database in Finland. *Stroke*. 2007;41:2239-46.

24. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: Possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol*. 2009;24:659-67.

25. Statistics New Zealand. *Integrated Data Infrastructure* [Internet]. Wellington (NZ): Statistics NZ; 2016 [cited 2016 May 4]. Available from: http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx

26. Wells S, Riddell T, Kerr A, Pylypchuk R, Chelimo C, Marshall R, et al. Cohort Profile: The PREDICT Cardiovascular Disease Cohort in New Zealand Primary Care (PREDICT-CVD 19). *Int J Epidemiol*. 2015 Dec 20. pii: dyv312. DOI: 10.1093/ije/dyv312

27. Cadilhac D, Lannin N, Anderson C, Andrew N, Kim J, Kilkenny M, et al. on Behalf of the AuSCR Consortium. *The Australian Stroke Clinical Registry Annual Report 2013*. Report No.: 5. Melbourne (AUST): The Florey Institute of Neuroscience and Mental Health; 2014.