

Process trumps potential public good: better vaccine safety through linked cross-jurisdictional immunisation data in Australia

Katherine M. Duszynski,^{1,2} Nicole L. Pratt,³ John W. Lynch,² Annette Braunack-Mayer,^{2,4} Lee K. Taylor,⁵ Jesia G. Berry,¹ Vicki Xafis,² Jim Buttery,^{6,7} Michael S. Gold¹ on behalf of the Vaccine Assessment Using Linked Data (VALiD) Working Group

The benefits of vaccines are globally acknowledged. Nevertheless, periodic concerns emerge regarding adverse events following immunisation (AEFI). Newly developed vaccines have incomplete safety profiles at the time of licensure because of limited participant enrolment in clinical trials and short duration of safety surveillance. Trials usually omit the vulnerable populations targeted by government vaccination programmes, including infants, pregnant women and the elderly. Common and acute reactions are readily identified, while rare and delayed AEFI may be missed without further assessment.

Without thorough assessment of all AEFI, appropriate government, regulatory and manufacturer action cannot be taken. Community confidence in immunisation benefits may waiver, resulting in reduced vaccine coverage, as became apparent for human papillomavirus (HPV) vaccination in Denmark^{1,2} and Japan³; and disease resurgence, as occurred with the measles-mumps-rubella vaccine in the United Kingdom.^{4,5}

To detect AEFI and mitigate the impact of any suspected concerns, the World Health Organization advocates all countries implement a post-licensure vaccine safety surveillance system.⁶ One recommended

Abstract

Objective: To provide insights into complexities of seeking access to state and federal cross-jurisdictional data for linkage with the Australian Childhood Immunisation Register (ACIR). We provide recommendations for improving access and receipt of linked datasets involving Australian Government-administered data.

Methods: We describe requirements for linking eleven federal and state data sources to establish a national linked dataset for safety evaluation of vaccines. The required data linkage methodology for integrating cross-jurisdictional data sources is also described.

Results: Extensive negotiation was required with 18 different agencies for 21 separate authorisations and 12 ethics approvals. Three variations of the 'best practice' linkage model were implemented. Australian Government approval requests spanned nearly four years from initial request for data, with a further year before ACIR data transfer to the linkage agency.

Conclusions: Integration of immunisation registers with other data collections is achievable in Australia but infeasible for routine and rapid identification of vaccine safety concerns. Lengthy authorisation requirements, convoluted disparate application processes and inconsistencies in data supplied all contribute to delayed data availability.

Implications for public health: Delayed data access for safety surveillance prevents timely epidemiological reviews. Poor responsiveness to safety concerns may erode public confidence, compromising effectiveness of vaccination programs through reduced participation.

Key words: vaccines, immunisation registry, record linkage, safety, surveillance

approach is a passive surveillance system (PSS), relying on reports submitted to regulatory agencies from health professionals, industry and community. One PSS aim is 'signal detection', and this is undertaken by the Therapeutic Goods Administration in Australia. Collated AEFI reports are examined for patterns involving

specific or groups of vaccines that may then need further investigation. However, as reports are generally non-mandatory, considerable under-reporting exists and is coupled with difficulty in determining numbers of administered vaccines. Signal evaluation is compromised, impeding assessment of a causal association between

1. Discipline of Paediatrics, Robinson Research Institute, School of Medicine, The University of Adelaide, South Australia

2. School of Public Health, The University of Adelaide, South Australia

3. Quality Use of Medicines and Pharmacy Research Centre, Sansom Institute, The University of South Australia

4. School of Health and Society, University of Wollongong, New South Wales

5. Centre for Epidemiology and Evidence, New South Wales Ministry of Health

6. SAEFVIC, Murdoch Children's Research Institute, Victoria

7. Department of Epidemiology and Preventive Medicine, Monash University, The Ritchie Centre, Hudson Institute, Victoria

Correspondence to: Ms Katherine Duszynski, Robinson Research Institute, The University of Adelaide, Ground Floor, Norwich Centre, 55 King William Road, North Adelaide, SA 5006; e-mail: katherine.duszynski@adelaide.edu.au

Submitted: September 2018; Revision Requested: March 2019; Accepted: June 2019

The authors have stated they have no conflict of interest.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

a vaccine and adverse event. Data linkage, or the matching and joining of records from administrative datasets, has been extensively used internationally, as a means of safety assessment. In the United States and Scandinavian countries, analysis of linked data has been used to both identify and refute associations between specific vaccines and adverse outcomes.⁷⁻¹⁰ A whole-of-population linked dataset enhances the scope, representativeness and population size for epidemiological assessments of vaccine safety.

Australia is well placed to employ data linkage for safety assessment of vaccines. Since 1996, routine childhood immunisations have been captured on the Australian Childhood Immunisation Register (ACIR), which was extended to all ages as the Australian Immunisation Register (AIR) in 2016. Adolescent human papillomavirus (HPV) immunisations have been recorded on the HPV Register since 2007, with data also integrated with AIR since late 2018. These registries have potential to be linked with other administrative data collections, such as jurisdictional hospital datasets and the National Death Index (NDI), to enable signal evaluation. However, application of linked administrative datasets for vaccine safety monitoring in Australia has been limited to a single investigation, the South Australian Vaccine Safety (SAVeS) study.¹¹

In late 2008, a team of investigators embarked on the Australian Research Council-funded study VALiD. The study objective was to investigate the acceptability and feasibility of linking Australian Government (hereafter AusGov) and jurisdictional data collections to evaluate the safety of vaccines. Two policy developments preceded project commencement. Firstly in 2006, the National Collaborative Research Infrastructure Strategy (NCRIS) identified population health and data linkage as a key priority for research investment. Enhanced data linkage capability was anticipated to expedite epidemiological research, leading to improvements in clinical practice and delivery of health and social services.¹² Secondly, two new vaccines were included on the NIP in 2007: a second-generation rotavirus vaccine (RV) protecting against diarrhoeal disease in children, and the HPV vaccine against cervical cancer. The nascent safety profiles of these vaccines warranted post-licensure surveillance, particularly since first-generation RVs were withdrawn in the US following

identification of the vaccine's increased risk of intussusception (bowel obstruction).¹³⁻¹⁵ Safety data for the HPV vaccine were also limited due to its recent introduction, with Australia the first country to include the vaccine in a funded national schedule.

This paper describes the complexities encountered in accessing cross-jurisdictional data for linkage with the ACIR to establish a national linked dataset for vaccine safety evaluation. We suggest a series of recommendations for improving access and timely delivery of linked datasets to enhance safety surveillance.

Methods

Data sources

Eleven data sources from two federal and five jurisdictional agencies were identified for linkage in the VALiD study (Table 1). Linkage occurred in stages due to delays in approval for release of jurisdictional data. Two project datasets were created with the ACIR as the primary data source: 1) a national linkage with death registration records from the NDI (1999–2010); and 2) a cross-jurisdictional linkage with hospital inpatient and emergency department (ED) attendance records (2003–2013) from four of five jurisdictions: South Australia, New South Wales, Victoria and Queensland. Western Australia (WA) withdrew in the final stages of approval (December 2015), citing legislative restrictions preventing release of hospital data to external agencies. Established in 1996, the ACIR maintains records of all immunisations (specified by the National Immunisation Schedule) administered to children up to 7 years of age. This includes

up to 21 separate vaccinations, protecting against 14 diseases, which are administered at six intervals during childhood.¹⁶ The ACIR records child name, address, demographic information, aboriginality, immunisation history (including vaccine administered, date of immunisation, dose and batch number) and provider contact information.

Linkage methodology – the model we proposed

Initially, we proposed linking datasets by replicating the methodology implemented for the SAVeS project,^{11,17} where linkage of immunisation and hospital data for South Australian (SA) children was conducted by the SA Department of Health (SADH). When meeting with AusGov agencies (November 2009), we presented two linkage models: a centralised model involving SADH taking responsibility for linking all required datasets and a distributed model, where ACIR identifiers would be sent to individual jurisdictions to perform state-based linkage with hospital data. However, a new policy introduced by the Commonwealth Department of Health required a different linkage methodology for AusGov data. Launched as 'High Level Principles for Data Integration involving Commonwealth Data' in February 2010,¹⁸ this policy stated linkage of AusGov datasets could only be undertaken by an approved agency or 'integrating authority'. At the time of application, two integrating authorities had received accreditation for linking AusGov data: the Australian Bureau of Statistics and the Australian Institute of Health and Welfare (AIHW), with the latter subsequently assigned as the linkage agency for the VALiD project.

Linkage methodology – the model prescribed by the Australian Government

The best practice protocol for data linkage stipulates that person identifiers (name, address, demographic) used for matching information between administrative datasets should be kept separate from those relating to an individual's health information (e.g. immunisations, deaths, hospitalisations).¹⁹ This 'separation principle' ensures individuals' confidentiality and privacy are protected, with linkage staff having no access to health information components. Similarly, researchers never receive access to person identifiers, preventing potential matching of health information to specific individuals.

Table 1: Data collections linked for the VALiD study.

Jurisdiction	Dataset name
Commonwealth	Australian Childhood Immunisation Register
AIHW	National Death Index
QLD	Hospital Admitted Patient Data Collection Emergency Department Information System Mater Hospital Emergency Department Information System
NSW	Admitted Patient Data Collection Emergency Department Data Collection
SA	Integrated South Australian Activity Collection Emergency Department Data Collection
WA	Hospital Morbidity Data Collection Emergency Department Data Collection
VIC	Victorian Admitted Episodes Dataset Victorian Emergency Minimum Dataset

Following matching of identifiers, a national linkage key (NLK) was generated and sent to data custodians to attach health information. The NLK comprised two identification numbers – a new Person ID and encrypted Local Record ID assigned by the data custodian with the latter removed following attachment of health data.

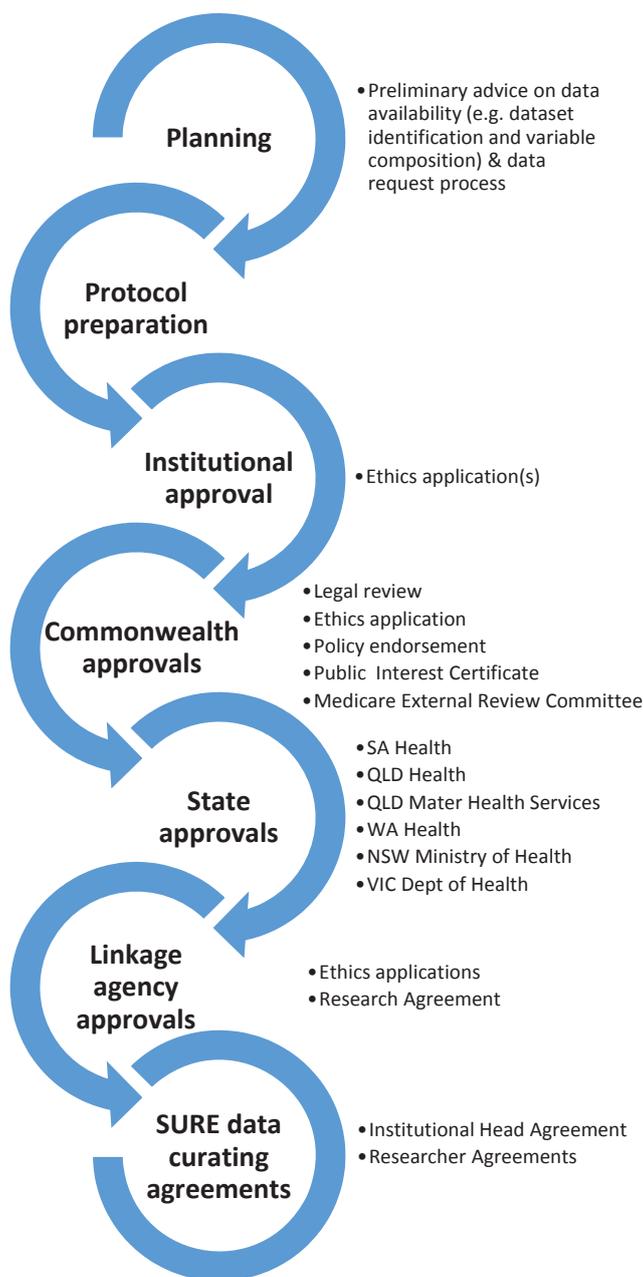
The AIHW implemented a variation to best practice linking methodology in handling AusGov data, receiving both identifier and health data extracts. Within the AIHW, partitioning of extracts was achieved by having separate data processing domains for receipt of identifier information and health data. Following linkage, integrated data were sent to a secure, remote access computing facility known as 'SURE' (Secure Unified Research Environment), hosted by the Sax Institute in NSW.

Results

Process and scope of approvals for data access

Figure 1 outlines the process, scope of approvals and agreements required before data linkage could proceed. Extensive negotiation was required between 18 different agencies to obtain 21 separate authorisations and 12 ethics approvals. Approvals were needed from four domains: Australian Government agencies, state-based authorities, ancillary organisations comprising linkage and data curating facilities, and human research ethics committees (HRECs). The process commenced with developing a study protocol and initial ethics submission to the University HREC. While the submission was under review, approval for data release was first sought from the AusGov Department of Health. Release of ACIR immunisation data for linkage and analysis required three tiers of approval: legislative, policy authority and administrative. Legislative approval involved review of the relevant Commonwealth statute (*Health Insurance Act 1973* (Cth)) regulating the function and management of ACIR. While disclosure of immunisation information was permitted in a restricted range of circumstances, including research, only non-identifiable information could be released. Non-identifiable data, however, are impracticable for data linkage activities that require person identifiers for dataset matching. Consequently, approval in the form of a *Public Interest Certificate* (PIC) signed by the data custodian, or their delegate, was

Figure 1: Overview of process and range of approvals/agreements required for securing release of cross-jurisdictional and Commonwealth data for linkage.



necessary to provide legislative exemption to data release protections.

Before delegate approval could be granted, policy authority was required from the relevant AusGov section with procedural responsibility for the ACIR dataset. During this process, overall project objectives and proposed methodology were reviewed by the Immunisation Branch together with an assessment of the risks of data release weighed against the potential public benefit of data provision. Procedural review for the VALiD project progressed through four levels of governance to achieve policy

authorisation. In-principle approval for release of ACIR data for the VALiD project was granted in February 2011, some 2.5 years after project commencement. A further 14 months ensued before the PIC, authorising release of the immunisation register data, was conferred in April 2012. Our ARC competitive grant ended in 2010.

The final AusGov authorisation before data release to AIHW involved approval from the Commonwealth Department of Human Services (DHS). This Department had administrative responsibility for the dataset and was ultimately responsible for

preparing and sending the data extract to the assigned integrating authority. This required a further application to the Medicare External Requests Review Committee and, while timely approval was secured, a further year ensued before data were extracted and transferred to the AIHW in May 2013. Additional requirements supplementary to the Commonwealth process included completion of *Confidentiality Deed Polls* for those investigators accessing data.

In summary, provision of AusGov approval took almost four years from the initial request for data access in August 2008, with a further year before ACIR data were transferred to the AIHW for the first linkage. An additional 14 months was required for linkage and provision of data, due to re-issue of ACIR data following identification of data integrity issues related to same-individual duplications in person identifier number. This process was completed in August 2014, six years after our initial consultation with the AusGov Department of Health.

Jurisdictional and ethics approvals

The jurisdictional request process was similarly multi-layered and protracted (Table 2). Broadly, this involved agency consultation, identifying required datasets, selecting relevant variables, completion of a data application (and additional forms where necessary), securing data custodian approval and, finally, ethics application. No two jurisdictions were the same in their approach to sanctioning data release. Time to securing data release ranged from nine months in NSW, to more than four years for Victorian data. In NSW, the application process required receipt of the AusGov PIC documentation before an ethics application could be submitted, delaying completion of state approval processes until secured.

As well as being non-uniform, the extent of documentation varied between jurisdictions, sometimes involving a single application form (such as Queensland's Public Health Application) to additional documents such as Privacy Forms (NSW) and technical feasibility assessments (NSW, WA, SA). In SA, QLD and WA, executive approvals were also required for trans-border flow of data to the AIHW and/or release of data from hospital area health services. Due to differences in hospital funding arrangements, two application processes for release of hospital data were undertaken in QLD involving the QLD Health Statistics Unit and Mater Health

Services (ED data only from former Mater Children's Hospital) with the latter requiring considerable documentation.

Multiple ethics approvals were also required. The significant delay in acquiring approval for data release from the ACIR delayed submissions for release of state hospital data, effectively separating our research objectives into two projects and two separate linkage processes. Submissions included two Commonwealth, six state, two institutional and two linkage agency applications. This resulted in 12 (rather than nine) separate ethics submissions seeking approval of ACIR linkage with the NDI, and then hospital datasets. Three jurisdictions also required additional submissions to state-based Aboriginal Health Ethics Committees for approval relating to the release of an Indigenous identifier.

Other approvals and agreements

Other approvals and agreements included a (data) Risk Assessment completed by the AIHW and submitted to Commonwealth DHS. Researcher and institutional agreements were also needed for the linkage agency's data storage facility (SURE). Confidentiality agreements with states were also required.

Data linkage methodology

Three variations of the data linkage model were implemented across agencies contributing data. Model 1 involving NSW, QLD Health and SA complied strictly with the best practice protocol.¹⁹ Each jurisdiction only supplied AIHW with hospital-derived person identifiers for linkage with ACIR identifiers. The resulting NLK was then returned to jurisdictions for appending health data extracts, then uploaded to SURE. Model 2 involved supply of AusGov ACIR data and QLD Mater ED data. These agencies sent both person identifier and health data extracts to separate domains within AIHW. Following processing and linkage, health data extracts were similarly attached to the NLK before transfer to SURE. Model 3 applied to linkage of Victorian hospital data. Provisional enquiries seeking access to Victorian hospital data commenced in November 2009. However, these negotiations were suspended until September 2012, when AusGov approval for release of immunisation data was assured. Initially, person identifier variables from the Victorian hospital datasets and approved health data extracts were to be supplied to the AIHW as described for Model 1. Identifier

variables were restricted to date of birth, gender, postcode and residential suburb. During consultation, the 'Better Patient Data' project, designed to improve linkage specificity, was undertaken by the Victorian Department of Health in 2014. This initiative expanded the range of identifiers available from the Patient Master Index of Victorian hospitals and involved adding names and addresses to retrospective data holdings from all Victorian hospitals. The expanded person identifier set was made available to the AIHW in late 2016, while health data from Victorian hospitals were derived from AIHW hospital data collections, to be added to the NLK, creating the third linkage model.

Datasets

Variation in data characteristics and variable formats were also noted across jurisdictional data sources. For example, state hospital datasets varied in availability of named data for linkage, limiting observation periods. In WA, for example, named separations data from public hospitals were from 1970; in QLD, only from 1995. In SA, named separations data were available from July 2003 onwards and for Victoria from mid-2010, although work is ongoing to expand named data back to mid-2005. Another inconsistency related to data variables released. Specific justifications were required for release of date of birth so that researchers could apply the most basic of epidemiological analyses – age adjustments – to ensure comparability of risks relative to age for AEFI. Most states prevented release of the postcode variable, instead providing a coded measure of geographic location and socioeconomic status for health data. However, coded measures were inconsistent across jurisdictions, with some states providing broader geographic locations such as statistical local areas.

Cost

Various costs were associated with creating the two linkage datasets. These included costs (GST exclusive) for ACIR (linkage 1: \$9,580; linkage 2: \$6,588) and state data extracts (\$0 in SA, QLD & VIC and \$34,000 in NSW), linkage services (linkage 1: \$18,900; linkage 2: \$59,600) and access to the SURE facility over four years (\$31,233). Excluding indirect personnel time in administrating approvals, direct costs for the VALiD study totalled almost \$126,000, although the true cost was higher due to all jurisdictions absorbing fees for data extracts. Final

personnel costs during the application stage were difficult to estimate, as research personnel prepared other study components in addition to the data linkage applications and negotiation of approvals. Conservatively, we estimate personnel time expended on approval activities across the six state and Commonwealth jurisdictions was one year's duration – equivalent to \$95,000 base salary with on-costs for a mid-level research person.

Discussion

This paper describes the complex and convoluted application, approval and operational processes to link cross-jurisdictional and AusGov data to establish a national linked dataset for vaccine safety signal evaluation. The cost (direct and personnel) and time to secure approvals for data linkage in Australia is clearly impractical for routine or even periodic safety monitoring of vaccines. This is concerning, given the importance of the public regard for safe vaccines.²⁰⁻²³ In earlier studies, the investigators identified considerable parental support (94%) for linking their child's vaccination and hospital records for vaccine safety surveillance, with high confidence (84%) in identity protections.²³ Parents also emphasised the public benefits of generating knowledge on potential harms, which prevailed over concerns regarding permissions for data access.²⁴ Our experience demonstrates the disconnect between public attitudes and the reality of what can be achieved with administrative data in Australia. While data protection is important, it needs to be balanced against the significant public health advantages arising from linking data and adding value to an existing and government-funded resource – the Australian immunisation registers.

Key findings

Investment in Australia's data linkage infrastructure has expanded availability of administrative data for health research.²⁵ However, practical access to these data is limited by considerable administrative burdens and delays for researchers. The experience of the VALiD investigators in establishing a national integrated dataset for safety assessment of vaccines identified three key areas of complexity in securing data for cross-jurisdictional projects involving Commonwealth and state data. These relate to: 1) distributed dataset access; 2) variable

Table 2: Stages of jurisdictional application process and associated application documentation.

	NSW	QLD	QLD Mater	SA	VIC	WA
Planning						
Consultation	◆	◆	◆	◆	◆	◆
Preparation of application documentation						
Site Specific Application Form			◆			
Public Health Application form		◆				
Data application form	◆			◆	◆*	◆
Data services form(s)						◆
Variable list form(s)	◆		◆	◆		◆
List of ICD codes						◆
Privacy form	◆		◆			
Documentation of funding notification	◆		◆			
Research protocol	◆		◆			◆
Ethics application	◆			◆	◆*	◆
Review of draft application	◆	◆	◆			◆
Supporting documentation			◆			◆
In-principle data custodian support						◆
Technical feasibility assessment form	◆					◆
Data custodian approval	◆	◆	◆	◆	◆	
Declaration of confidentiality						◆
Indemnity			◆			
Investigator CVs			◆			
Ethical review						
Ethics approval	◆	MA	MA	◆	◆	◆
Indigenous ethics approval	◆			◆		◆
Allied approvals						
Executive approval	◆	◆		◆		◆ (4)
Final data custodian approval						◆
Confidentiality Agreement	◆			◆	◆	
Research Agreement			◆			
Total application/approval documents	8	1	8	4	2	16

Notes:

*Data request negotiated directly with data custodian with ethics approval secured through ethics application.

MA=Mutual acceptance of ethics approval

and frequently non-transparent application processes; and 3) lengthy approval times involving multiple tiers of authorisation (Table 3). While detailed as separate challenges, these issues are all entwined, and arise from the distributed operational responsibility for healthcare in Australia.

Hospitalisations, for example, are administered by jurisdictions, while federal programs such as the National Immunisation Program and the Pharmaceutical Benefits Schedule are administered by AusGov agencies. Researchers need to negotiate approvals with individual jurisdictional and national agencies to obtain authorisation for data release and/or use. Complex and lengthy application processes ensue, progressing through multiple tiers of authorisation for data release approvals. Further obstacles arise through the non-uniform application process, with no two jurisdictions alike in their requirements, leading to differences

in application stages (1–4), application documents (1–16) and time to approval for data release – ranging from nine months to six years. Complicating the negotiations were variations in willingness to release specific health data variables, particularly those deemed sensitive and with potential for re-identification of individuals, e.g. postcode. Transparency in the application processes also varied. At the Commonwealth level, the application process was opaque and lacked coordination between AusGov agencies, with limited information provided by agencies on sequence and timing of government approvals. The delay in securing AusGov authorisations also delayed approvals from state agencies. Applications could not be commenced or progressed in one jurisdiction without receipt of the PIC approving release of data from ACIR, the primary data source. One further concerning feature was the decision not to release hospital data in

one jurisdiction (WA) at the final approval stage involving WA formal data custodian endorsement. This happened despite our investigators completing extensive application documentation and securing six individual (ethics and governance) approvals. Collectively, these challenges create substantial time and resource commitments for researchers brokering approvals to achieve data release. Furthermore, as data were no longer current by the time of data release, revised submissions became necessary to obtain more recent data. The challenges encountered in seeking to link cross-jurisdictional data are not isolated to the current study, with parallels also seen in a two-state immunisation effectiveness study and a national injury surveillance study.²⁶⁻²⁸ The difficulties encountered raise significant ethical concerns about wasting public research funds and serious concerns regarding the feasibility of undertaking research of significant public benefit due to uncoordinated and disparate governance structures. The 2017 Productivity Commission Data Availability and Use Inquiry Report²⁹ ('PC Report') further emphasise that delays in data access impact on researchers' ability to provide real-world evidence on topics of concern.

Recommendations

A simpler, more transparent application and approvals model for release and linkage of Commonwealth and jurisdictional data is urgently needed to ensure researcher efforts and resources are directed to investigating study objectives rather than negotiating the approval pathway. A uniform model of managing approvals would also reduce the burden experienced by data custodians when assessing and approving dataset requests. Four principal recommendations are proposed to simplify data access and reduce time to data approval and release. These are to:

1. Prioritise access to high-utility data collections, e.g. jurisdictional hospital

separations/ED presentations, national registries, Pharmaceutical Benefits Scheme (PBS), Medicare Benefits Schedule (MBS).

2. Establish multiple portals (state-specific or institutional organisations) for managing approvals and data release of high-utility data.
3. Simplify data and ethics applications to a single application process supported by a uniform template for assessing risks of data release.
4. Implement a web-based electronic system with facility for monitoring progress of data requests, approvals, data extract preparation, linkage and data release.

Implementation of these recommendations would benefit researchers and assist data custodians in streamlining processes for data access. Situating Australian Government high-value datasets across multiple approved agencies with a uniform application template would: 1) clearly identify a nominated agency(ies) for data release approvals, thereby reducing burden (and waiting times) related to repeated requests for AusGov datasets with Commonwealth Departments; 2) resolve recursive legislative and policy reviews through a standing agreement with AusGov agencies pre-specifying conditions of data release; 3) reduce the cost of data extracts; 4) establish data dictionaries consistent with minimum dataset templates; and thereby 5) provide a uniform format for release of data variables from cross-jurisdictional and national data holdings.

Analytic environments

Allied with our recommendations is an appeal for expanding the options for secure analytic environments used to manage and access integrated data. Researchers are currently constrained by computing and analysis environments pre-determined by data custodial agencies. These environments are inflexible to meet the evolving analytical requirements of researchers. As research questions become more complex and the scope of datasets expands, computing

infrastructure in the existing environment will be insufficient for enhanced processing of multiple extremely large datasets and running complex statistical models. As suggested in one submission³⁰ to the PC Inquiry, institutions could provide a more flexible analysis facility for trusted users working with linked data. Through an accreditation mechanism, institutions would need to demonstrate appropriate security access measures, auditing of file access and transfer are in place. Allowing accredited institutions to establish secure computing environments employing cloud-based storage software, rather than requiring physical servers,³¹ would also defray some of the significant (and increasing) costs related to accessing linked data.

Reforming data sharing and release

The propositions outlined above align with recommendations detailed in the PC Report²⁹ and proposed Australian Government reforms³² responding to the PC Report. Reforms would be legislated through the new Data Sharing and Release Bill.³³ The Bill aims to increase authorised sharing and release of Australian Government-held data while improving data safeguards and risk management tools to create a more transparent environment for data sharing.³³ Planned reforms for improving data sharing and release arrangements include identifying high-value data collections, creating Accredited Data Authorities to facilitate data provision, and implementing a trusted user framework for assessing data requests.³² High-value data collections with potential for delivering population benefits would be designated as National Interest Datasets and given priority access.

Accredited Data Authorities (ADAs) building on the current Integrating Authorities model would expand the network of agencies with facility for linking, sharing or releasing datasets. These Accredited Authorities would act as intermediaries between data custodians and users to facilitate data availability,^{29,32} including national datasets identified as high-value. Feasibly, ADAs could be assigned to existing state-based linkage units but also serve as an opportunity for other non-government agencies to expand their role to data provision. Increasing the number of agencies available to distribute high-value data on behalf of data custodians would considerably reduce the bottleneck associated with obtaining approvals, access

Table 3: Summary of complexities and recommendations for improving data release.

Issue	Recommendations
Fragmented dataset access leading to separate administrative responsibilities for national and cross-national data collections	Locate high-utility data collections in a designated repositories through a memorandum of understanding or standing agreement involving state and Australian Government departments
Protracted approval for data release arising from multiple tiers of authorisation with transparency lacking on time to approval and data release	Electronic system integrating data application process and monitoring progress of approvals, data preparation, linkage and release.
Variable and sometimes non-transparent application processes across jurisdictions	Single application process situated at repository with one uniform application form

and release of data currently experienced with the AIHW.

To increase transparency of approvals and streamline application processes, data access would be assessed by ADAs applying the 'trusted user' model, based on the Five Safes Framework.³² Originally intended for identified data with scope broadened to include deidentified data, the Framework comprises five principles or dimensions of data access to inform a process of safe data release.³⁴ These are: Safe projects (Can the researchers be trusted?); Safe people (Is the purpose of use appropriate? What analysis is being done?); Safe data (Can the data disclose identity?); Safe settings (Does the access environment prevent unauthorised use?); and Safe outputs (Are the statistical results non-disclosive?).^{34,35} The Five Safes provide a clearly articulated approach for assessing data requests; noting, however, that the five principles already underpin specifications outlined in current data and ethics applications. If uniformly applied by data custodians as a singular 'template', the framework would provide a transparent and simplified mechanism for managing data applications and also ethical review of linked data requests. In turn, the template would reduce burden of application review for data custodians and ethics committees as well as duplication of applications submitted by researchers.

The trusted user data model has been adopted by the Australian Bureau of Statistics³⁶ and the AIHW, with some states (Victoria and SA) also implementing this model for sharing data across government agencies.³⁵ Trusted users would be aligned with institutions (e.g. government agencies, universities) with existing arrangements for managing improper data use. Arrangements include agreeing to legal undertakings specifying protections for data use and having necessary computing infrastructure for storing data securely.³⁷

Monitoring data requests, approvals and release

Regarding our recommendation for a web-based monitoring system, some elements such as progress in data requests and approvals are considered in the online application established by the Population Health Research Network (PHRN).³⁸ The online application aims to harmonise application and approval processes for cross-jurisdictional and multi-jurisdictional research projects.

Useful for complex linkages, the online process provides an opportunity for dataset representatives to collectively raise issues about the application rather than requiring separate discussions with individual linkage agencies. However, there remains concerns by these researchers that the PHRN adds yet another layer of application to an already convoluted application process. Despite the central application point, there still remains a requirement to complete ancillary approval documentation for all jurisdictions (with the exception of SA and TAS), rather than consolidating data access requests into a single uniform application.³⁸

Programmatic linkages

Further considerations related to data access and data timeliness include establishing enduring (rather than ad hoc) linkages. The now-named Australian Immunisation Register (AIR) offers clear population benefits for evaluating the impact of immunisation. Linked with other datasets, the AIR could be established as a routine and enduring programmatic linkage for monitoring the effectiveness and safety of vaccines. An immunisation programmatic linkage could contribute to a national system of other programmatic linkages involving National Interest Datasets. Programmatic linkages could be managed by ADAs located across and between government, non-government and academic sectors.^{29,32} Inter-sectorial programmatic linkages, would also remove redundancy of repeated linkages of datasets. More critically, they would enable rapid and real-time safety surveillance to be conducted for vaccines, particularly seasonal or new vaccines introduced on the NIP.

Implications for public health

The delay and complexity in accessing linked data for vaccine safety surveillance has both local and global public health implications. Firstly, it results in an inability to undertake timely epidemiological reviews to examine associations between reported serious events and a vaccine. This is particularly critical for newly licensed vaccines. Surveillance of AEFI and response to safety concerns provide the community with assurances that vaccines are appropriately regulated by the Australian Government. This, in turn, engenders public confidence and contributes to the success of vaccination programs through ongoing participation. Other work by the project

researchers has established community preference for employing linked data for public benefit, including vaccine safety monitoring, over and above preferences for individual consent and privacy concerns relating to the use of identifiable data.²²⁻²⁴ Secondly, countries differ in the selection and timing of vaccine administration. As vaccine trials are usually conducted outside Australia, information on possible safety issues derived from these trials may lack comparability with the Australian population. Routine surveillance following vaccine administration is therefore important for providing local knowledge on AEFI, including their incidence, occurrence and type. Thirdly, the World Health Organization advocates surveillance for AEFI as part of global efforts to enhance safety information across diverse populations, particularly for rare outcomes. Well-resourced countries like Australia, with strict regulatory controls on therapeutic products and a reputation for early adoption of new vaccines, have a role to play in these efforts by providing safety data on routinely administered vaccines.

Conclusion

Integration of immunisation registers with other data collections is achievable in Australia but remains infeasible for routine and rapid identification of vaccine safety concerns. Multiple lengthy authorisation requirements, convoluted application processes and inconsistencies in data supplied all contribute to delayed data availability. Prioritising access to national and jurisdictional datasets of high value, a single application process with transparent assessment of data requests and an electronic system for monitoring progress of approvals and data release would expedite data access. This would lead to a surveillance system that is rapid and responsive to monitoring vaccine safety concerns. Furthermore, data would be available for external parties to provide a measure of accountability for policy decisions, independent of government assessment.

Acknowledgements

We acknowledge the support of the Vaccine Assessment Using Linked Data (VALiD) Working Group throughout the VALiD project. The members are (from), University of Adelaide, SA : Michael S Gold, Annette J Braunack-Mayer, Philip Ryan, Katherine M

Duszynski, Jesia G Berry, Vicki Xafis, Jillian Carlson and Bernadette Richards. New South Wales Ministry of Health, NSW: Lee K Taylor. Royal Children's Hospital, Vic: Jim P Buttery. University of Melbourne, Vic: Cecily J Freemantle. Monash University, Vic: John J McNeil and Lisa L Demos. University of Wollongong, NSW: Colin Thomson. University of New South Wales, NSW: Glenda L Lawrence. The University of Sydney, NSW: Elizabeth J Elliot and Julie Leask. John Hunter Hospital, NSW: Rod Givney. Therapeutic Goods Administration, ACT: Gary Lacey. SA Health, SA: Tony Woollacott, Chris Gascoigne, Ann Koehler, Maureen Watson, Rebecca Horgan and Sarah Lawson. University of Western Australia, WA: Peter C Richmond. Medicare Australia, ACT: Sean Tarrant. University of South Australia, SA: Elizabeth E Roughead. We further recognise the contribution of the Commonwealth Department of Health in enabling access to Australian Childhood Immunisation Register data and the Australian Institute of Health and Welfare (AIHW) for access to the National Mortality Index. We are also grateful for the work of the Data Linkage Service Unit at the AIHW particularly, Tenniel Guiver and Dr Phil Anderson who coordinated data acquisition with the Commonwealth Departments of Health and Department of Human Services as well as their oversight of the dataset linkage. KMD also acknowledges the considerable support of Professor Elizabeth Roughead and A/Professor Nicole Pratt from the Quality Use of Medicines and Pharmacy Research Centre, Sansom Institute, University of South Australia during manuscript write-up.

Funding

This project work was funded by an Australian Research Council (ARC) Linkage Project Grant LP0882394. Michael Gold, Philip Ryan, John McNeil, Jane Freemantle, Colin Thomson, Elizabeth Roughead, Lee Taylor and Jim Buttery are Chief Investigators on this grant. Lee Taylor, David Filby, Jim Buttery and Elizabeth Elliot are Partner Investigators affiliated with organisations who provided financial and/or in-kind support for this work as part of the Vaccine Assessment Using Linked Data project. The organisations include the South Australian Department of Health & Ageing, the NSW Ministry of Health, Surveillance of Adverse Events Following Vaccination in Victoria (SAEFVic) and the Australian Paediatric Surveillance Unit (APSU). The study funders had no role in

study design and the collection, analysis, and interpretation of data, writing of the article or the decision to submit it for publication. KMD was supported by a University of Adelaide Division of Health Sciences Postgraduate Scholarship.

References

- Brinth L, Theibel AC, Pors K, Mehlsen J. Suspected side effects to the quadrivalent human papilloma vaccine. *Dan Med J*. 2015;62(4):A5064.
- European Medicines Agency. *European Medicines Agency Assessment Report: Review under Article 20 of Regulation (EC) No 726/2004*. London (UK): EMA; 2015.
- Hanley SJ, Yoshioka E, Ito Y, Kishi R. HPV vaccination crisis in Japan. *Lancet*. 2015;385(9987):2571.
- The National Health Service Information Centre Workforce and Facilities. *NHS Immunisation Statistics England 2007–08*. Leeds (UK): Government of United Kingdom; 2009.
- Elliman D, Bedford H. MMR: Where are we now? *Arch Dis Child*. 2007;92(12):1055-7.
- World Health Organization. *Global Vaccine Safety Blueprint*. Geneva (CHE): WHO; 2012.
- Yih WK, Lieu TA, Kulldorff M, et al. Intussusception risk after rotavirus vaccination in U.S. infants. *N Engl J Med*. 2014;370(6):503-12.
- Barlow WE, Davis RL, Glasser JW, et al. The risk of seizures after receipt of whole-cell pertussis or measles, mumps, and rubella vaccine. *N Engl J Med*. 2001;345(9):656-61.
- Arnhem-Dahlstrom L, Pasternak B, Svanstrom H, Sparen P, Hviid A. Autoimmune, neurological, and venous thromboembolic adverse events after immunisation of adolescent girls with quadrivalent human papillomavirus vaccine in Denmark and Sweden: Cohort study. *BMJ*. 2013;347:f5906.
- Pasternak B, Svanstrom H, Molgaard-Nielsen D, et al. Risk of adverse fetal outcomes following administration of a pandemic influenza A(H1N1) vaccine during pregnancy. *JAMA*. 2012;308(2):165-74.
- Gold M, Dugdale S, Woodman RJ, McCaul KA. Use of the Australian Childhood Immunisation Register for vaccine safety data linkage. *Vaccine*. 2010;28(26):4308-11.
- Australian Department of Education and Training. *National Collaborative Research Infrastructure Strategy Strategic Roadmap 2006*. Canberra (AUST): Government of Australia; 2006.
- Murphy TV, Gargiullo PM, Massoudi MS, et al. Intussusception among infants given an oral rotavirus vaccine. *N Engl J Med*. 2001;344(8):564-72.
- Murphy BR, Morens DM, Simonsen L, Chanock RM, La Montagne JR, Kapikian AZ. Reappraisal of the association of intussusception with the licensed live rotavirus vaccine challenges initial conclusions. *J Infect Dis*. 2003;187(8):1301-8.
- Advisory Committee on Immunization Practices. Rotavirus vaccine for the prevention of rotavirus gastroenteritis among children. Recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Morb Mortal Wkly*. 1999;48(RR-2):1-20.
- Hull BP, Deeks SL, McIntyre PB. The Australian Childhood Immunisation Register-A model for universal immunisation registers? *Vaccine*. 2009;27(37):5054-60.
- Gold M, Dugdale S, and the SAVeS Project Advisory Group. *SA Vaccine Safety Data Linkage Project (SAVES): A New Way to Monitor Vaccine Safety in South Australia (Final Report to HSRIP: September 2004 - June 2006)*. Adelaide (AUST): State Government of South Australia; 2006.
- Cross Portfolio Statistical Integration Committee. *High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes*. Canberra (AUST): Government of Australia; 2010.
- Kelman CW, Bass AJ, Holman CD. Research use of linked health data - a best practice protocol. *Aust N Z J Public Health*. 2002;26(3):251-5.
- Berry JG, Gold MS, Ryan P, Duszynski KM, Braunack-Mayer AJ. Public perspectives on consent for the linkage of data to evaluate vaccine safety. *Vaccine*. 2012;30(28):4167-74.
- Berry JG, Ryan P, Duszynski KM, et al. Parent perspectives on consent for the linkage of data to evaluate vaccine safety: A randomised trial of opt-in and opt-out consent. *Clin Trials*. 2013;10(3):483-94.
- Berry JG, Gold MS, Ryan P, Duszynski KM, Braunack-Mayer AJ. Public perspectives on consent for the linkage of data to evaluate vaccine safety. *Vaccine*. 2012;30(28):4167-74.
- Berry JG, Ryan P, Duszynski KM, et al. Parent perspectives on consent for the linkage of data to evaluate vaccine safety: A randomised trial of opt-in and opt-out consent. *Clin Trials*. 2013;10(3):483-94.
- Xafis V. The acceptability of conducting data linkage research without obtaining consent: Lay people's views and justifications. *BMC Med Ethics*. 2015;16(1):79.
- Population Health Research Network. *Population Health Research Network Annual Review 2015–16*. Perth (AUST): PHRN; 2016.
- Moore HC, Guiver T, Woollacott A, de Klerk N, Gidding HF. Establishing a process for conducting cross-jurisdictional record linkage in Australia. *Aust N Z J Public Health*. 2016;40(2):159-64.
- 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 N Z J Public Health*. 2015;39(4):319-25.
- Mitchell RJ, Cameron CM, Bambach MR. Data linkage for injury surveillance and research in Australia: Perils, pitfalls and potential. *Aust N Z J Public Health*. 2014;38(3):275-80.
- Productivity Commission. *Data Availability and Use, Report No. 82*. Canberra (AUST): Government of Australia; 2017.
- University of NSW Centre for Big Data Research in Health. *Submission to the Productivity Commission Inquiry into Data Availability and Use (Submission 21)*. Canberra (AUST): Productivity Commission; 2017.
- Productivity Commission. *Storage and Cloud Computing*. Canberra (AUST): Government of Australia; 2017.
- Department of the Prime Minister and Cabinet. *The Australian Government's Response to the Productivity Commission Data Availability and Use Inquiry*. Canberra (AUST): Government of Australia; 2018.
- Department of the Prime Minister and Cabinet. *New Australian Government Data Sharing and Release Legislation: Issues Paper for Consultation*. Canberra (AUST): Government of Australia; 2018.
- Desai T, Ritchie F, Welpton R. *Five Safes: Designing Data Access for Research*. Bristol (UK): University of the West of England; 2016.
- Productivity Commission. *Data Availability and Use - Draft Report*. Canberra (AUST): Government of Australia; 2016.
- Australian Bureau of Statistics. *ABS Submission to the Productivity Commission Inquiry into Data Availability and Use (Submission 94)*. Canberra (AUST): Productivity Commission; 2017.
- Productivity Commission. *Overview: Data Availability and Use - Draft Report*. Canberra (AUST): Government of Australia; 2016.
- Population Health Research Network. *Online Application System* [Internet]. Perth (AUST): PHRN; 2017 [cited 2018 Jul 1]. Available from: <http://www.phrn.org.au/for-researchers/data-access/online-application-system/>