



**PHRN** Population  
Health  
Research  
Network

# **PHRN RESEARCHER CONSULTATIONS**

## **FINAL REPORT**

Dr Angela Rate

PHRN Program Office

Submitted to Management Council

October 2012

## Table of Contents

1	EXECUTIVE SUMMARY .....	4
1.1	Introduction.....	4
1.2	Consultation.....	4
1.3	Key Findings.....	4
1.4	Conclusions .....	6
2	INTRODUCTION.....	7
2.1	Background.....	7
2.2	Accessing and using linked data for cross-jurisdictional research projects.....	7
3	PHRN RESEARCHER CONSULTATIONS.....	7
3.1	Purpose of current consultations.....	7
3.2	Approach Adopted .....	8
3.3	Consultation Schedule .....	8
4	SUMMARY OF FINDINGS .....	9
4.1	Metadata.....	9
4.2	Other information beyond metadata .....	11
4.3	Online application system for cross-jurisdictional projects.....	12
4.4	Training.....	13
5	CONCLUSIONS.....	15
6	REFERENCES .....	16
	Appendix 1: Transcript of online survey and accompanying information sheet.....	17

Appendix 2: Face-to-face consultations dates and attendees..... 24

Appendix 3: Survey respondents – demographics..... 27

# 1 EXECUTIVE SUMMARY

## 1.1 Introduction

The Population Health Research Network (PHRN) has been established to build a nationwide data linkage infrastructure capable of securely and safely managing population information from around Australia. Part of PHRN's role is to facilitate and coordinate access by researchers to linked data.

This report describes a researcher consultation process undertaken by the PHRN Program Office to identify what information, tools and training researchers need to assist them to access cross-jurisdictional linked data through the PHRN. The consultations included an emphasis on three elements that are currently being investigated for the PHRN: metadata, web-based applications for cross-jurisdictional projects and training for users of linked data.

## 1.2 Consultation

Consultations with researchers who are current and future users of linked data occurred over a five month period between April and August 2012. They were conducted by Dr Angela Rate via face to face meetings involving round-table discussions with small groups or an online survey for those researchers who couldn't attend the meetings. A total of 100 researcher representatives were consulted.

## 1.3 Key Findings

### Metadata

- Currently there is variable use of metadata published by local data linkage units, custodian-published metadata, the Australian Institute of Health and Welfare's (AIHW) MeTEOR resource and other collection-specific metadata. Usage was dependent on availability, usefulness and knowledge of existence.
- General support was received for a centralised metadata resource in the form of a catalogue of collections from across Australia (including Commonwealth collections) and links back to source metadata where available.
- More detailed information about data quality was requested. This included the completeness and reliability/accuracy of the data such as whether the variable is recorded routinely; if a coding system or systematic classification is used; if the variable is self-reported; and changes in variables over time.
- Other metadata elements requested included publication of data collection instruments; a catalogue of projects using linked data and contact details of lead authors; and a catalogue of data validation studies for commonly used collections where available.
- There was limited support for annotatable metadata or a researcher forum for exchange of information about data as most participants said they would probably not have the time to contribute to such resources effectively.

## **Information beyond metadata**

- A range of other information was suggested for publication in a centralised information portal including how to access Commonwealth data, ethical/privacy/security aspects of linkage, geocoding, pricing for data linkage projects, researcher FAQs, statistical disclosure control and links to other useful resources including data linkage unit contacts, training opportunities and data linkage publications.

## **Web-based online application system for cross-jurisdictional data linkage projects**

- There was general support for an online application submission system for cross-jurisdictional projects.
- Tracking was thought to be a good idea. However concern was expressed as to how this may be effective for a cross-jurisdictional project involving multiple application forms to be reviewed by multiple custodians and ethics committees.
- All were supportive of the development of a single data application form that would be acceptable across the jurisdictions.
- There was variable use of currently available online National Ethics Application Form (NEAF) application submission websites but positive feedback from those who had used these services.

## **Researcher Training**

- Current levels of undergraduate training were not deemed sufficient for conducting research using linked data and much learning was done via on-the-job training and through supervisors.
- Many of the researchers had completed the *Introductory Analysis of Linked Data* course offered by Prof D'Arcy Holman from UWA or the adapted version of this offered by the University of Sydney and had found it to be extremely useful. Others said limited availability and costs associated with travel had so far prevented them from attending.
- A smaller number of researchers had completed Prof Holman's *Advanced Analysis of Linked Data* course.
- A growing number of researchers had or were intending to complete the SURE training offered by the Sax Institute.
- A small number of other training modules had been undertaken by a few of the researchers including a Biostatistics Collaboration of Australia (BCA) traineeship, a 5-day course on multilevel modelling, Prof D'Arcy Holman's *Data and Biospecimen Law for Epidemiologists* and vendor-run courses on the use of analytical software.
- There was general support for provision of basic information such as how linkage is conducted, how to merge data, limitations of administrative data, legal/ethical aspects of linkage and access to Commonwealth data but it was

felt by a majority that this could be provided through an online information portal rather than training *per se*.

## **1.4 Conclusions**

The PHRN researcher consultations found that researchers were supportive of a streamlined approach to the access and delivery of linkable data, including the development of an online national metadata resource, provision of general information for researchers and an online application and tracking system, as well as the increased availability of existing training programs around data analysis and the use of the SURE facility.

The consultation report will be tabled at the PHRN Management Council and the results of the consultation will be used to inform the development of the PHRN Metadata Framework and online application system.

## **2 INTRODUCTION**

### **2.1 Background**

The Population Health Research Network has been established to provide Australian researchers with access to linkable de-identified data from a diverse and rich range of population datasets, across jurisdictions and sectors. This will support research of national relevance which results in improved policy making and service delivery and demonstrates global best practice in maximising the benefits to the Australian community whilst preserving individual privacy. The Network is an initiative of the Australian Government being conducted as part of the National Collaborative Research Infrastructure Strategy (NCRIS) and the Education Investment Fund Super Science Initiative (EIF-SSI).

### **2.2 Accessing and using linked data for cross-jurisdictional research projects**

Researchers who wish to apply for access to linked data held within Australian states and territories are currently directed to the local data linkage units in each jurisdiction. Within each unit there is generally a dedicated client services team that can assist researchers with various aspects of projects involving linked data including design and cohort selection, provision of quotations, completion of data applications, advice on the Human Research Ethics Committee (HREC) process and applications, negotiation of contracts with custodians and in some instances, archiving of data extracts at the completion of the study.

Compared with single-jurisdiction projects, the process for access to linked data from multiple jurisdictions is more complex due to the different data custodian and HREC requirements in each state/territory. There may also be multiple linkage units involved in linking the data. In order to maximise the efficiency of the data linkage infrastructure developed under the PHRN NCRIS and EIF-SSI investments and hence support as many projects as possible, the PHRN can play a role in providing support services to the users of linked data.

## **3 PHRN RESEARCHER CONSULTATIONS**

### **3.1 Purpose of current consultations**

A series of consultations with researchers who are current or future users of linked data from across Australia was organised to meet a number of objectives:

- 1) To provide current and future users of linked data with an overview of the PHRN and its initiatives;
- 2) To obtain information about access and usage practices around current client services for data linkage research available around Australia;
- 3) To understand the gaps or areas for improvement with a particular emphasis on three priority elements of the national data linkage infrastructure that the

Program Office is funded to investigate for the PHRN through the PHRN NCRIS and EIF-SSI Investment Plans<sup>1,2</sup>: a) metadata; b) web-based applications for cross-jurisdictional projects; and c) training for users of linked data.

The findings and recommendations from these consultations will help to ensure that users of linked data can use the PHRN infrastructure to its maximum potential.

### **3.2 Approach Adopted**

Face-to-face consultations were arranged with researchers who currently or are planning to use linked data from each of the jurisdictions. An online anonymous survey was provided for interested users who could not attend the face-to-face consultations. The information sheet provided to all participants as well as a transcript of the online survey are provided at **Appendix 1**.

Representatives from each PHRN Participant as per the NCRIS and EIF-SSI Funding Agreements were contacted by Dr Angela Rate via email in April 2012. They were provided with information outlining the purpose of the consultations accompanied by a request for a list of potential users of linked data to invite to participate. Dates of the face-to-face consultations and related logistics including venues and catering were arranged in consultation with the PHRN Participant Representatives. Dr Rate then contacted the nominated linked data users via email to invite them to attend either the face-to-face consultations or alternatively to complete the online survey should they be interested in providing feedback but unavailable to attend the face-to-face consultations.

All face-to-face consultations were conducted by Dr Angela Rate (Policy and Client Services Coordinator). Dr Felicity Flack (Policy and Client Services Manager) was also in attendance for the Perth and Adelaide consultations. The meetings were conducted over a 1-1.5-hour period in a round-table format with a maximum of 13 attendees at any one place. Dr Rate provided a general background to the PHRN including its structure, funding model and key initiatives including an overview of the PHRN Proof of Collaboration program and more detailed information was then delivered on the three client services elements of particular interest: a) metadata; b) web-based applications for cross-jurisdictional projects; and c) training for users of linked data. Dr Rate supported her introduction with a Powerpoint presentation in Canberra, Hobart, Sydney and Brisbane. The floor was then opened for a general discussion around the PHRN and client services elements of specific interest.

### **3.3 Consultation Schedule**

Consultations with users of linked data were conducted between April and August 2012. A total of 60 current and future users of linked data attended the face-to-face consultations held in Perth, Sydney, Canberra, Adelaide, Hobart, Melbourne and Brisbane. A number of other interested stakeholders who were not necessarily current or future users of the linked data including data linkage unit staff and management, HREC members and representatives from the ABS were also in

attendance at a number of the meetings. A further 40 individuals completed the online survey. A list of dates and locations of the face-to-face consultations and all attendees can be found at **Appendix 2**. Demographic information for the online survey respondents is provided at **Appendix 3**.

## **4 SUMMARY OF FINDINGS**

Below is a summary of the feedback received from the consultations with researchers. Please note that the source of feedback has been removed to maintain confidentiality. Original minutes and online survey results have been retained as a record of the consultations.

### **4.1 Metadata**

Participants were provided with a definition of the term 'metadata' and were advised that the PHRN is considering ways to provide information about data collections across Australia that are potentially available to be linked. It was made clear that the intention would be that links will be provided back to the source metadata wherever it is available rather than the PHRN developing its own metadata per se.

#### ***Participant feedback from face to face consultations***

- Variable use of metadata published by local data linkage units, custodian-published metadata and the AIHW's MeTEOR resource which was dependent on availability, usefulness and knowledge of existence
- More experienced researchers do not use formal metadata publications as they just 'know' the collections that they use
- Some researchers had access to more detailed metadata through their affiliations with state health departments
- General support for a centralised metadata resource in the form of a catalogue of collections from across Australia (including Commonwealth collections) and links back to source metadata where available
- General acknowledgement of a lack of resourcing of custodians in collating and maintaining detailed metadata
- More detailed information about data quality was requested across all consultations, particularly the completeness and reliability/accuracy of the data such as whether the variable is recorded routinely, if a coding system or systematic classification is used, if the variable is self-reported, changes in variables over time
- Endorsement from a number of groups was received for the publication of data collection instruments that could be useful in determining underlying data quality
- Wide support for a list of data linkage projects with the data collections involved and contact details of the chief investigator to enable direct sharing of

information about quality and usefulness of particular collections rather than publication of annotatable metadata or creation of an online researcher forum

- Support for development of a catalogue of validation studies of commonly used collections and links to these was supported by a number of groups
- One researcher suggested that PHRN could encourage/assist in the development and/or publication of concept dictionaries similar to that produced by the Manitoba Centre for Health Policy Research
- Another researcher suggested that PHRN could provide guidelines or a framework for metadata collation to data custodians which reflects what information researchers are looking for.

### ***Participant feedback from online survey***

In relation to current use of existing metadata resources, 28 respondents (70%) use a variety of metadata resources including metadata on AIHW's MeTEOR (5 people), the CHeReL website (4 people), NSW Health HOIST or SAPHARI (4 people), the WADLB website (3 people), the SANT DataLink website (3 people), 45&Up metadata (2 people), custodian-supplied metadata (2 people), ABS (1 person), LSAC/HILDA metadata (1 person), National Coroner's Database (1 person) and ICD-10 metadata (1 person). Metadata quality was reported to be variable depending on the source.

Respondents were provided with a list of suggested metadata items and were asked to nominate whether they agreed with their inclusion in a PHRN metadata resource. Below is a summary of responses:

<b>Metadata Item</b>	<b>% agreement</b>
Which data collections are available	100%
Where the data collections are held	97.4%
Custodian contact details	95.0%
Legislation covering the collection, use and disclosure of each data collection	87.2%
Approvals required for access to the data collections	95%
Quality of the data (including reliability, accuracy, completeness)	95.0%
Detailed descriptions of the variables in the data collections (including mode and method of collection, changes over time)	95.0%
Links to any validation studies of the data available	92.5%
Glossary of commonly used data linkage terms	82.1%
Researcher forum for exchange of information/ experiences with using specific data collections	81.1%

Other items suggested for consideration included approval requirements, links to publications, links to other metadata sources including Research Data Australia,

contact details for local data linkage units, classification and re-coding methods applied to data and concept dictionaries.

## 4.2 Other information beyond metadata

The PHRN could provide researchers using linked data with a range of other information beyond metadata and so the consultations attempted to determine if there was support for additional information to be made available in a centralised information portal.

### ***Participant feedback from face to face consultations***

Beyond the general metadata information requested, researchers were also keen to have more information provided on:

- How to access Commonwealth data (common request from all face-to-face consultations)
- Links to local data linkage units
- Ethical/privacy aspects of linkage (although many of the researchers who had completed the SURE training said that this was covered in the training)
- Geocoding
- Pricing for data linkage projects

### ***Participant feedback from online survey***

Below is a list of suggested information elements that were put to survey respondents and the percentage of individuals who agreed with their publication:

<b>Information Portal Item</b>	<b>% Agreement</b>
Linked data application process and approximate timelines	85.0%
Approximate timelines for applications	82.5%
Information about pricing of data linkage projects	87.5%
Data linkage methodology – how data linkage works	81.6%
Design of data linkage projects	62.2%
Ethical issues including considerations for HREC applications involving linked data	94.9%
Standard FAQs for researchers	94.9%
Legal framework governing data linkage projects	82.5%
Security aspects of data linkage projects	90.0%
Statistical disclosure risk assessment and control	85.0%
Links to data linkage publications	75.7%
Other useful resources e.g. training opportunities	91.7%

### 4.3 Online application system for cross-jurisdictional projects

Access to data from multiple jurisdictions currently requires the completion of individual data applications for each data collection as well as multiple HREC applications. Face-to-face consultation and survey participants were advised that the PHRN is considering development of an online application management system that will serve as a central point for researchers to electronically complete and submit data applications and ethics applications for cross-jurisdictional projects.

#### ***Participant feedback from face to face consultations***

- General support received for an online application submission system for cross-jurisdictional projects.
- Tracking thought to be a good idea. However concern was expressed as to how this may be effective for a cross-jurisdictional project involving multiple application forms to be reviewed by multiple custodians and ethics committees.
- All supportive of the idea of a single data application form that would be acceptable across the jurisdictions.
- Most researchers outside WA had completed a NEAF form (in-house ethics forms preferred in WA).
- Of those who had completed the NEAF, there had been variable use of the Australian Online Research Forms website and/or the NH&MRC's NEAF website. The majority that had used them had found the sites easy to use/navigate and instead expressed more frustration around the content of the NEAF rather than the process of completing it.
- During a number of discussions around application submission, the issue of multiple ethical reviews was raised and a number of participants were keen for the PHRN to assist in the progression of the Harmonisation of Multi-Centre Ethical Review process.

#### ***Participant feedback from online survey***

The online survey provided respondents with an opportunity to submit feedback about the functions of an online application system should it be developed. The following items were provided as possible features and the percentage of the 40 respondents in agreement with the suggestions are shown:

<b>Online application feature</b>	<b>% Agreement</b>
Web interface to allow access from any computer	100.0%
Individual user accounts for each researcher that can hold multiple applications for different projects	100.0%
Electronic applications specific for each data collection/jurisdiction that can be updated and saved on a periodic basis	97.3%
Ability to import and export information in XML format	71.4%

<b>Online application feature</b>	<b>% Agreement</b>
Ability to invite other registered users to view/edit applications	86.5%
Automated project tracking to provide information on the progress of project applications	100.0%

From the survey, 16 of the 40 respondents (40%) had used the Australian Online Research Forms website and/or the NH&MRC's NEAF website to complete the NEAF. Their general comments reflected those made during the face-to-face consultations.

#### **4.4 Training**

Training for researchers around using linked data was identified as a priority for the PHRN to consider following a series of consultations with PHRN Participants by the PHRN National Training and Education Coordinator in mid-2011. A number of areas for focus for training for data users were suggested by the PHRN Participant staff as detailed in the PHRN Training Consultation Final Report. This current set of consultations with researchers provided an opportunity for the PHRN to conduct an audit of what training had been completed by this target group, whether they perceived benefit in the development of training in areas suggested by the PHRN Participant staff and if there were any other gaps in training that the PHRN could address.

##### ***Participant feedback from face to face consultations***

- Current levels of undergraduate training were not deemed sufficient for conducting research using linked data and a lot of learning was done via on-the-job training and through supervisors
- Many of the researchers had completed the *Introductory Analysis of Linked Data* course offered by Prof D'Arcy Holman from UWA or the adapted version of this offered by the University of Sydney and had found it to be extremely useful
- A smaller number had completed Prof Holman's *Advanced Analysis of Linked Data* course
- Some researchers in states outside of WA and NSW said that they would be interested in completing the *Introductory Analysis of Linked Data* course but limited availability and costs associated with travel had so far prevented them from attending
- A growing number of researchers had or were intending to complete the SURE training offered by the Sax Institute
- A small number of other training modules had been undertaken by a few of the data users including a Biostatistics Collaboration of Australia (BCA) traineeship, a 5-day course on multilevel modelling, Prof D'Arcy Holman's *Data*

and *Biospecimen Law for Epidemiologists* and vendor-run courses on the use of analytical software

- There was general support for provision of basic information such as how linkage is conducted, how to merge data, limitations of administrative data, legal/ethical aspects of linkage and access to Commonwealth data but it was felt by a number of data users that this could be provided through an online information portal rather than training *per se*.
- Some data users felt that the PHRN should be less focused on delivering linkage-specific training but should instead be supporting training aimed at data analysis in general such as the units and traineeships offered by the Biostatistics Collaboration of Australia

### **Participant feedback from online survey**

- 34 respondents (85%) had completed a post-graduate degree (Combination of MPH, PhD, Masters of Biostatistics, Masters of Health Information Management)
- 97% of these individuals felt that their formal qualifications had not equipped them with the complete skill set required to use linked data for research, rather these had been gained through additional formal training or on-the-job training
- 28 respondents (70%) had completed additional formal training in the area of linked data:
  - D’Arcy Holman’s Introductory Analysis of Linked Data – 16 people
  - D’Arcy Holman’s Advanced Analysis of Linked Data – 8 people
  - USyd Introductory Analysis of Linked Data – 7 people
  - SURE training – 6 people
  - D’Arcy Holman’s Data and Biospecimen Law for Epidemiologists – 1 person
  - IBMSAS/COBOL – 1 person
  - Data sharing seminars at Uni of Qld – 1 person
  - HealthLinQ workshop – 1 person
  - Internal ABS training – 1 person
- The following suggestions of additional training came from the original PHRN training consultations with PHRN Participant staff. Researchers were asked whether they would be interested in undertaking training in these areas and the following is a summary of their responses:

<b>Training Topic</b>	<b>% Agreement</b>
Understanding administrative data and its constraints	28.6%
Introduction to the legal aspects of using administrative data	28.6%

<b>Training Topic</b>	<b>% Agreement</b>
Privacy and data linkage	28.6%
Ethics and data linkage including how to write ethics applications	31.4%
Data security – physical security, hardware security, software security, safe handling of output and access control	36.1%
Researchers' roles and responsibilities	20.6%
Statistical disclosure risk assessment and control	45.7%
Data application and approval processes for state/territory and Commonwealth data	54.3%
Methods for linking administrative data	42.9%
Merging datasets	32.4%
Interpreting administrative data	48.6%
Analysis of linked data	40.0%
How to use the Secure Unified Research Environment (SURE; <a href="http://www.sure.org.au">www.sure.org.au</a> )	51.4%

- The general comments in this section highlighted the fact that many of the respondents had either completed relevant training already or would prefer to receive much of this information by way of online information packages rather than formal training. This was consistent with the findings from the face-to-face consultations.

## **5 CONCLUSIONS**

The PHRN researcher consultations found that researchers were supportive of a streamlined approach to the access and delivery of linkable data, including the development of an online national metadata resource, provision of general information for researchers and an online application and tracking system, as well as the increased availability of existing training programs around data analysis and the use of the SURE facility.

## 6 REFERENCES

- (1) Australian Government 2009 *Funding Agreement between the Commonwealth of Australia as represented by the Department of Innovation, Industry, Science and Research and The University of Western Australia regarding funding for the National Collaborative Research Infrastructure Strategy's Research Capability known as the Population Health Research Network.*
- (2) Australian Government 2011 *Funding Agreement between the Commonwealth of Australia as represented by the Department of Innovation, Industry, Science and Research and The University of Western Australia regarding funding for Implementing an Investment Plan for the Population Health Research Network Education Investment Fund Super Science Initiative.*

## Appendix 1: Transcript of online survey and accompanying information sheet



### PHRN LINKED DATA RESEARCHER CONSULTATION SERIES 2012

#### What is the Population Health Research Network?

The Population Health Research Network (PHRN) is an Australian initiative responsible for building national data linkage infrastructure. The PHRN is overseeing the development of new and expanded data linkage capacity in each Australian state/territory and supporting the establishment of a national Centre for Data Linkage, a Commonwealth Integrating Authority and a state-of-the-art remote access laboratory. This will enable researchers to access a wider range of data collections from across Australia and have the ability to link data from across jurisdictions including the Commonwealth and other sectors including education and justice. Further information is available at [www.phrn.org.au](http://www.phrn.org.au).

#### Accessing data for cross-jurisdictional data linkage projects

Researchers who wish to apply for access to linked data held within Australian states and territories are currently directed to the local data linkage units in each jurisdiction. Within each unit there is generally a dedicated client services team who can assist researchers with various aspects of projects involving linked data; including design and cohort selection, provision of quotations, completion of data applications, advice on the Human Research Ethics Committee (HREC) process and applications, negotiation of contracts with custodians and in some instances, archiving of data extracts at the completion of the study.

The process for access to linked data from multiple jurisdictions will be more complex due to the different data custodian and HREC requirements in each state/territory. In an effort to streamline this activity, the PHRN is proposing to have a centrally administered client services hub where researchers will have access to information to help them plan their study, know what approvals are required and be able to complete and submit their applications in one place rather than having to approach each jurisdiction separately.

As part of the development of the central client services hub for cross-jurisdictional linkage projects, the functional requirements for a number of elements are being drafted. These include:

- A centralised information portal for researchers including access to metadata;

- An online application and tracking system for researchers wanting access to linked data across jurisdictions; and
- A training program for researchers.

In order to design these elements to work efficiently and meet researcher needs, a series of consultations with users of linked data from across Australia has been planned to learn from users experiences and to understand their needs and requirements. The findings and recommendations from these consultations will guide development of the central client services hub and will help to ensure that researchers can use the PHRN infrastructure to its maximum potential.

A short survey has been created in Survey Monkey about the three elements – please go to <http://www.surveymonkey.com/s/FHM62VJ> to complete it.

## ONLINE SURVEY

### 1. Information Portal for Researchers

#### 1.1. Metadata

Metadata is often called ‘data about data’. More precisely, it is the underlying definition or structured description of the content, quality, condition or other characteristics of data. In reality, data collections are held by a range of organisations and the collation and publication of metadata is generally completed by individual data custodians.

The PHRN is proposing the creation of a database of data collections across Australia that are available to be linked and that researchers may apply for access for approved projects. The intention is that links will be provided back to the source metadata wherever it is available.

<b>Q1: Which items do you think should be included in the metadata database?</b>		
<b>Item</b>	<b>Please tick box if you would find useful</b>	<b>Additional comments</b>
Which data collections are available	<input type="checkbox"/>	
Where the data collections are held	<input type="checkbox"/>	
Custodian contact details	<input type="checkbox"/>	
Legislation covering the collection, use and disclosure of each data collection	<input type="checkbox"/>	
Approvals required for access to the data collections	<input type="checkbox"/>	

Item	Please tick box if you would find useful	Additional comments
Quality of the data (including reliability, accuracy, completeness)	<input type="checkbox"/>	
Detailed descriptions of the variables in the data collections (including mode and method of collection, changes over time)	<input type="checkbox"/>	
Links to any validation studies of the data available.	<input type="checkbox"/>	
Glossary of commonly used data linkage terms	<input type="checkbox"/>	
Researcher forum for exchange of information/ experiences with using specific data collections	<input type="checkbox"/>	
<b>Others? Please list here:</b>		
<p><b>Q2: Do you currently use existing metadata resources relating to data you use/intend to use for linkage-based research? If yes, which metadata sources do you use and do they meet your needs in terms of content, ease of use and accessibility?</b></p>		

## 1.2. Other information relevant to researchers

In addition to metadata about individual data collections, it is proposed that a comprehensive suite of information is provided to researchers within the information portal.

<b>Q3: Do you think the information portal should include any of the following items?</b>		
Item	Please tick box if you would find useful	Additional comments
Linked data application process and approximate timelines	<input type="checkbox"/>	
Approximate timelines for applications	<input type="checkbox"/>	

Item	Please tick box if you would find useful	Additional comments
Information about pricing of data linkage projects	<input type="checkbox"/>	
Data linkage methodology – how data linkage works	<input type="checkbox"/>	
Design of data linkage projects	<input type="checkbox"/>	
Ethical issues including considerations for HREC applications involving linked data	<input type="checkbox"/>	
Standard FAQs for researchers	<input type="checkbox"/>	
Legal framework governing data linkage projects	<input type="checkbox"/>	
Security aspects of data linkage projects	<input type="checkbox"/>	
Statistical disclosure risk assessment and control	<input type="checkbox"/>	
Links to data linkage publications	<input type="checkbox"/>	
Other useful resources e.g. training opportunities	<input type="checkbox"/>	
<b>Others? Please list here:</b>		

## 2. Online Application/Tracking of Cross-Jurisdictional Projects

Access to data from multiple jurisdictions currently requires the completion of individual data applications as well as multiple HREC applications. To streamline application processes for cross-jurisdictional linkage projects, the PHRN is proposing to develop an online application management system (AMS) that will serve as a central point for researchers to electronically complete and submit data applications required by individual data custodians and in some instances, complete and submit HREC applications.

<b>Q4: Which features should be included in the AMS?</b>		
Feature	Please tick box if you would find useful	Additional comments
Web interface to allow access from any computer	<input type="checkbox"/>	

Feature	Please tick box if you would find useful	Additional comments
Individual user accounts for each researcher that can hold multiple applications for different projects	<input type="checkbox"/>	
Electronic applications specific for each data collection/jurisdiction that can be updated and saved on a periodic basis*	<input type="checkbox"/>	
Ability to import and export information in XML format	<input type="checkbox"/>	
Ability to invite other registered users to view/edit applications	<input type="checkbox"/>	
Automated project tracking to provide information on the progress of project applications	<input type="checkbox"/>	
<b>Others? Please list here:</b>		

**Q5: Do you currently use the Australia Online Forms for Research ([www.ethicsform.org/au/](http://www.ethicsform.org/au/)) or the NEAF website ([www.neaf.gov.au](http://www.neaf.gov.au)) to complete and submit NEAF? If so, have you found the site(s) easy to use and navigate?**

### 3. Researcher Training

As researchers, you are one of the three key target groups of the PHRNs training strategy, along with HREC members and staff of the PHRN Participant Organisations. Consultations have occurred with both PHRN staff and HREC members and we are now interested in hearing from you about your training needs and priorities.

**Q6. Do you have a tertiary qualification and if so, did your tertiary education provide you with the necessary skills to use linked data?**

**Q7. Have you attended training relevant to data linkage before? If yes, what training have you attended? If no, how have you learnt how to use linked data?**

**Q8. Would you like to receive training on any of the following topics?**

Topic	Please tick box if you would like to receive training	Additional comments
Understanding administrative data and its constraints	<input type="checkbox"/>	
Introduction to the legal aspects of using administrative data	<input type="checkbox"/>	
Privacy and data linkage	<input type="checkbox"/>	
Ethics and data linkage including how to write ethics applications	<input type="checkbox"/>	
Data security – physical security, hardware security, software security, safe handling of output and access control	<input type="checkbox"/>	
Researchers' roles and responsibilities	<input type="checkbox"/>	
Statistical disclosure control	<input type="checkbox"/>	
Data application and approval processes for state/territory and Commonwealth data	<input type="checkbox"/>	
Methods for linking administrative data	<input type="checkbox"/>	
Merging datasets	<input type="checkbox"/>	
Interpreting administrative data	<input type="checkbox"/>	
Analysis of linked data	<input type="checkbox"/>	
How to use the Secure Unified Research Environment (SURE; <a href="http://www.sure.org.au">www.sure.org.au</a> )	<input type="checkbox"/>	
<b>Others? Please list here:</b>		

#### 4. Information about you

<b>Q9. How long have you been conducting research using linked data?</b>	
<1 year	<input type="checkbox"/>
1-3 years	<input type="checkbox"/>
3-5 years	<input type="checkbox"/>
5-10 years	<input type="checkbox"/>

>10 years	<input type="checkbox"/>
-----------	--------------------------

<b>Q10. What type of institution are you affiliated with?</b>	
University	<input type="checkbox"/>
Research Institute	<input type="checkbox"/>
Government Agency	<input type="checkbox"/>
Other: please specify	<input type="checkbox"/>

## **Appendix 2: Face-to-face consultations dates and attendees**

A total of 77 people attended the consultations including 60 current and future users of linked data. Below is a list of attendees by consultation.

### **University of Western Australia School of Population Health, Perth – 26.04.12**

- Asst/Prof Louise Stewart
- Asst/Prof Qun (Bella) Mai
- Ms Sylvie Price
- Assoc/Prof Frank Sanfillipo
- Prof David Preen
- Asst/Prof Eva Malacova

### **Telethon Institute for Child Health Research, Perth – 26.04.12**

- Dr Hannah Moore
- Dr Lyn Colvin
- Mr Peter Cosgrove

### **University of New South Wales, Sydney - 24.05.12**

- Dr Rebecca Mitchell
- Mr Han Xu
- Dr Anne Kricker (USyd staff member)
- Assoc/Prof Claire Vajdic
- Dr Heather Gidding
- Dr Efty Stavrou

### **Kolling Institute, University of Sydney, Sydney - 24.05.12**

- Prof Christine Roberts
- Dr Natasha Nassar
- Ms Samantha Lain
- Mr Charles Algert
- Ms Mandy Ampt
- Dr Anthony Ashton
- Dr Amina Khambalia
- Ms Megan Macfarlane
- Dr Sharon McCracken
- Ms Jillian Patterson

### **University of Sydney, Sydney – 24.05.12**

- Assoc/Prof Sallie Pearson
- Dr Tim Dobbins
- Ms Mikaela Jorgensen
- Prof Judy Simpson
- Dr Tim Driscoll

### **National Centre for Epidemiology and Population Health, Australian National University, Canberra - 14.06.12**

- Prof Emily Banks
- Ms Ellie Paige

- Dr Grace Joshy
- Ms Megan Shipley
- Ms Tanya Mather
- Ms Rosemary Macfarlane
- Mr Ivan Hanigan

#### **SA-based Researchers, Adelaide - 23.07.12**

- Prof John Lynch
- Ms Jesia Berry
- Dr Katherine Duszynski
- Ms Alyssa Sawyer
- Dr Catherine Chittleborough
- Ms Louisa Santucci
- Dr Michael Davies
- Prof Jon Karnon
- Dr James Doidge
- Ms Almond Sparrow (SANT DataLink)
- Ms Stacy Vasquez (SANT DataLink)
- Mr Chris Radbone (SANT DataLink)
- Ms Mon Lim (SANT DataLink)

#### **VIC-based Researchers, Melbourne - 27.07.12**

- Dr Vijaya Sundararajan

Note: VDL requested that Dr Rate refrain from consulting with other Victorian-based researchers at this time.

#### **QLD-based Researchers, Brisbane - 24.08.12**

- Dr Angela White
- Dr Suhail Doi
- Dr Abdullah Mamum
- Dr Deidre McLaughlin
- Mr Richard Hockey
- Assoc/Prof Peter Baade
- Mr Shannon Dias
- Ms Lyn McPherson
- Ms Karen Cosgrove (QLD Dept of Treasury)
- Ms Kit Loke (QLD Record Linkage Group)
- Mr Dinu Corbu (QLD Record Linkage Group)
- Ms Nyree Divitini (HealthlinQ)
- Ms Anita Paydar (HealthlinQ)

#### **TAS- based Researchers, Hobart - 28.08.12**

- Dr Kristen Hinds
- Dr Julia Walters
- Dr Jim Chalman
- Mr Tim Albion
- Dr Kristy Sanderson
- Prof Don Chalmers (TDLU Management Committee member/ Dean of UTas Law School)

- Ms Jo Dickinson (UTas HREC member)
- Mr Glenn Sward (ABS)
- Ms Julia Curtis (ABS)
- Ms Alison Kreft (ABS)
- Ms Ruth McArdle (ABS)
- Dr Kelly Shaw (TDLU Management Committee member)
- Mr Brian Stokes (TDLU Manager)

### **Appendix 3: Survey respondents – demographics**

The online survey was designed as an anonymous survey however respondents were asked to complete a section regarding their research affiliations and experience. The following is a summary of information provided by the 40 respondents:

#### ***Length of time conducting data linkage research***

<1 year	30.5%
1-3 years	16.5%
3-5 years	25%
5-10 years	16.5%
>10 years	11.5%

#### ***Research affiliation***

University	80%
Research Institute	12.5%
Government agency	5%
Other	2.5% (non-profit organisation)