PHRN Infrastructure Planning Summary

September 2019

Contents

Executive Summary

The Population Health Research Network (PHRN) has been granted additional funding through the Australian Government's NCRIS 2018 Program with funding committed through to June 2023 to maintain and enhance the existing PHRN national data linkage infrastructure. A consultation process was conducted to determine the infrastructure priorities for this investment.

The consultations took place between February and June 2019. A wide range of stakeholders participated in the extensive consultation process including researchers, representatives from the therapeutic development sector and PHRN Participants. The opinions of the different stakeholder groups were gathered using online surveys, face to face workshops and individual interviews. The results of all the consultations were analysed thematically to determine the infrastructure priorities for the PHRN from 2020 to 2023.

The infrastructure priority areas that emerged from the consultations were:

- Simple and timely application and approval processes
- Therapeutic Development
- Secure Access and Cybersecurity
- The Governance and Authorising Environment
- Capital Upgrades for Existing Infrastructure

Introduction

In mid-2018 the PHRN was granted additional funding through the Australian Government's NCRIS 2018 Program with funding committed through to June 2023 to deliver national data linkage infrastructure.

In 2019 the PHRN Program Office coordinated a planning process to assess future infrastructure needs and to inform funding decisions for a three year period from July 2020.

As guided by NCRIS guidelines, the future investments are to maintain and enhance the existing PHRN national data linkage network and to deliver the maximum benefit to researchers and to the community. Investments must also align with the 2016 National Research Infrastructure Roadmap. The PHRN is included in the chapter on Therapeutic Development in the 2016 Roadmap.

Methods

Researcher Consultation

Survey

The online researcher survey included a mix of closed and open ended questions (13 total) which covered demographic, research interests, and needs for data and data linkage services and facilities over the next five years.

The survey was distributed to 364 researchers on the PHRN mailing list. All PHRN Participants were also asked to forward the survey to their researcher mailing lists.

The survey was open for two weeks with a reminder sent after one week.

Interviews

A series of semi-structured interviews with nine senior researchers were conducted either face to face or by telephone.

PHRN Participant Consultation

The Participant Consultation consisted of three parts:

- Participant Survey
- Participant Workshops
- Technical Forum discussions

Participant Survey

The online Participant survey included a mix of closed and open ended questions (13 total) which covered demographic, and needs for data and data linkage services and facilities over the next five years.

The survey was distributed to 84 people on the PHRN Participant mailing list.

The survey was open for two weeks with a reminder sent after one week.

PHRN Participant Workshops

A series of workshops were held within the PHRN and included:

- Participant Council, 22 February 2019.
- PHRN managers, 1 March 2019
- Program Office, 26 March 2019

After hearing a summary of the results of the researcher and Participant surveys, the participants were asked to compile a list of all the possible services, facilities or infrastructure that will be required to meet researchers' needs over the next five years. The participants then considered:

- Which items on the list should not be considered for PHRN funding?
- Which items on the list provided opportunities for national standardisation, interoperability, coordination or centralisation?
- Which infrastructure would make the greatest improvement in services for researchers?

PHRN Technical Forum

Approximately 25 staff from PHRN-related facilities attended the PHRN Technical Forum held in Hobart on 27-28 February 2019. The final session of the Forum included a facilitated discussion on technical priorities. Unit managers each identified their top technical priorities and the group then identified and discussed common themes.

Industry Consultation

Consulting firm Biointelect facilitated a discussion on 26 April 2019 between the PHRN, technical experts and industry experts on data needs from each of three target segments (market access for medicines, market access for devices and data needs for clinical trials) and representatives of the Medical Services Advisory Committee (MSAC) and the Pharmaceutical Benefits Advisory Committee (PBAC).

Results

Researcher Consultation

The results from both the researcher survey and interviews are summarised below.

Types of linked data

The highest priority for researchers who participated in the survey or interviews was access to a broad range of linked administrative data. Their next priorities were clinical data, longitudinal cohort data and clinical registries.

Services and Facilities

The main services and facilities requested by researchers over the next five years were:

- Data linkage units
- Cross-jurisdictional linkage
- Secure data transfer
- Online application system (OAS)
- Secure remote access facilities
- Data repositories

There was very little support or suggestions for new facilities and services. The focus of most comments was on improving existing facilities, services and processes.

Barriers and Challenges

The major challenge reported by researchers was the time taken to access linked data. They identified a number of factors which they perceived to impact timeliness of access including:

- Different application forms required by every jurisdiction, even if the OAS is used.
- Processes changing all the time
- Complicated and duplicative approval processes (both data custodian and ethics).
- Lack of coordination
- Under-resourcing of data linkage units

Researchers were also concerned about pricing. There were concerns about the lack of transparency of pricing structures as well as lack of consistency in pricing between facilities. Some facilities were considered too expensive and in some cases costs were prohibitive.

Solutions

Researchers suggested a number of solutions to the barriers and challenges they identified. There was strong support for training and information about how to access and analyse linked data. They also suggested data repositories or data platforms as more efficient, and possibly cheaper, options for data access.

PHRN Participant Consultation

The results from the participant survey, workshops and Technical Forum are summarised below.

Types of linked data

Given that administrative data is the main type of data currently linked by PHRN data linkage units, Participants were asked what other types of data should be routinely linked. In addition to

administrative data the highest priority types of data were clinical registries, clinical data and longitudinal cohort data.

Services and Facilities

The Participants reported that a wide range of services and facilities would be required over the next five years including:

- Jurisdictional data linkage
- Cross-jurisdictional linkage
- Online application system
- Familial linkages
- Geocoding
- Secure data transfer
- Secure remote access facilities
- Data repositories
- Data storage
- Data archiving
- Specialist analytic software
- Supercomputing/high performance computing/flexible computing
- Data management platforms
- Assistance with study design and approvals
- Metadata/concept dictionaries
- Researcher training
- Project tracking and project management systems

There was very little support or suggestions for new facilities and services. The focus of most comments was on improving existing facilities, services and processes.

Investments which would make the greatest difference to researchers

When asked which infrastructure the PHRN should invest in to make the greatest difference to researchers there were a variety of answers from different groups of PHRN participants. However two main themes emerged:

- Streamlining the application and approval processes
- More efficient access to content data through linked and harmonised data repositories or data integration units.

Two other themes also had some support:

- Information security including the use of secure remote access facilities
- High quality and efficient linkage

Industry Consultation

All of the industry representatives wanted simplified information on the variables available in specific datasets, more information on who and how to contact with an idea, and some case studies that illustrate some typical enquiries that might be useful to the sectors. They were also interested in some form of aggregated data that could be made available 'routinely' so that private sector users could easily look up specific data as needed.

In addition, the needs of the therapeutic development sector have been canvassed through the PHRN collaboration with the Australian Clinical Trials Alliance. An extremely high priority for this sector is also simple and fast access to linked data.

Discussion

Types of data

Both researchers and PHRN Participants agreed that linking population level administrative data remained of high importance and core business for the PHRN. In addition to administrative data both groups also agreed that the following types of data are a high priority for linkage over the next five years:

- clinical data
- clinical registries
- longitudinal cohorts

A number of researchers stated that if they could just get easy access to a few linked core health data collections (hospital, emergency, death, PBS and MBS) they would have sufficient work to do for the remainder of their careers.

Researchers' requirements

Overwhelmingly the two main barriers to using linked data reported by researchers were pricing and the time taken to access linked data.

Comments about pricing were divided into the following categories:

- Transparency
- Consistency
- Free/minimal pricing

At a minimum researchers wanted pricing to be transparent and pricing information to be easy to find. Consistency in pricing across units was also important. All researchers wanted costs to be minimal and some did not think linkage costs should be paid for by grant funding. They thought these costs should be covered by governments through their support for data linkage units or by research institutions at the institutional level.

Researchers' solutions to reducing the time to access linked data focused heavily on streamlining and standardising application and approval processes. Many thought there should be a single application form used across all jurisdictions. They also suggested data repositories or data platforms as more efficient, and possibly cheaper, options for data access. Implied in their support for data repositories and explicitly described by one respondent was the assumption that the holder of the repository would be the data custodian for all the data in the repository and that this would remove the requirement for multiple data custodian approvals.

Associated with researchers' support for data repositories or data platforms was an acknowledgement that in the future access to linked data will be through secure remote data access facilities.

Services and facilities required over the next 5 years

The major themes that emerged for services and facilities required over the next 5 years were:

- Data repositories and data harmonisation
- Streamlined governance and approval processes
- Jurisdictional and cross-jurisdictional linkage
- Secure access and data transfer
- Training/information about application and approval systems
- Transparent Pricing

Both researchers and PHRN Participants agreed that the PHRN should focus on improving and expanding the existing PHRN services and facilities over the next 5 years. There was clear support for the continued linkage of existing administrative data collections and for expansion of linked data collections to focus on clinical data, clinical registries and longitudinal cohorts. Cross-jurisdictional linkage, in particular, was seen as important and still challenging. There were very few suggestions for completely new services and facilities.

Whilst researchers focused on application and approval processes and data repositories as ways to improve timeliness of access, the Participants also suggested other improvements through linkage efficiency and quality and expansion of system capacity.

Many respondents supported data repositories. There were likely many different interpretations of the meaning of the term. All were talking about a shift from distributed management of content data to more central or coordinated management of content data. Descriptions of data repositories included:

- A single research centre which was controlled by researchers who had responsibility for linkage, content data management and making decisions about what the data was used for i.e. acted as the data custodian, data linker and data analyst.
- Data integration units e.g. CARES, NSW Data Integration Unit, SA NT DataLink Custodian Controlled Data Repository

The Participant Council identified information security and cyber security as a high priority. Other groups didn't explicitly identify this as a priority but it could be implied from the support for secure remote access facilities, data repositories and secure data transfer.

PHRN Priorities 2020 to 2023

The consultation resulted in the PHRN adopting the following priority areas and related challenges as its focus for the next two years.

Simple and timely application and approval processes

The overwhelming feedback from the research community was that they needed simpler application and approval processes and that the time from submission of an application to receipt of linked data should be significantly shorter.

Challenge 1

Implement a data application system in all jurisdictions and for cross-jurisdictional projects where the researcher is only required to complete one application form.

Challenge 2

Reduce the time it takes from when a project is approved to delivery of linked data to the researcher.

Therapeutic Development

Supporting and enabling the use of linked data for therapeutic development remains a high priority for the PHRN as set out in the 2016 National Research Infrastructure Roadmap.

Challenge 3

Routinely link clinically relevant data collections for use at several stages along the therapeutic development pipeline including long term follow-up of clinical trials participants and for regulatory approval and reimbursement submissions e.g. cost-effectiveness studies.

Challenge 4

Provide a service which provides simple descriptive analysis of linked data required to meet the planning needs of the therapeutic development sector.

Secure Access and Cybersecurity

The careful management of personal information through rigorous information governance and cybersecurity is very important to PHRN Participants and data custodians. Researchers indicated that they are willing to access data in secure remote facilities as long as the facilities provide the necessary software, compute capacity and storage and are not cost prohibitive.

Challenge 5

Provide secure remote access facilities which are flexible, scalable and meet researcher needs.

The Governance and Authorising Environment

Researchers requested information and training to assist them to navigate the complex application and approval processes. Although not always explicit in the consultation, it is clear that a major barrier to timely access to linked data is this complex environment. This is likely to continue to be a barrier to improvement for all PHRN Participants. In addition the establishment and maintenance of a social licence for the use of linked data for research will remain essential for all Participants.

Challenge 6

Support better understanding of the authorising environment, provide education and training for all stakeholders and involve the community in the work of the PHRN.

Capital Upgrades for Existing Infrastructure

There is an ongoing need to maintain and enhance existing data linkage infrastructure as provided for through NCRIS program funding and reflected in continued growth in demand for linked data in Australia.

Challenge 7

Maintain or improve the coordinated delivery of linked data to researchers through further investment in existing capital infrastructure.