



Australian Government

Australian Institute of
Health and Welfare

*Authoritative information and statistics
to promote better health and wellbeing*

Office of the Chief Statistician and Performance
4N-06 St Andrews House
Regent Road
Edinburgh
EH1 3DG
Scotland
UK

A Scotland-wide Data Linkage Framework for Statistics and Research: Consultation Paper on the Aims and Guiding Principles

Introduction

The Australian Institute of Health and Welfare (AIHW) welcomes the opportunity to make a submission in response to the consultation paper on '*A Scotland-wide Data Linkage Framework for Statistics and Research: Aims and Guiding Principles*'.

The AIHW is a major national agency set up by the Australian Government under the Australian Institute of Health and Welfare Act to provide reliable, regular and relevant information and statistics on Australia's health and welfare. We are an independent statutory authority established in 1987, governed by a Management Board, and accountable to the Australian Parliament through the Health and Ageing portfolio.

Our aim is to improve the health and wellbeing of Australians through better health and welfare information and statistics. We collect and report information on a wide range of topics and issues, ranging from health and welfare expenditure, hospitals, disease and injury, and mental health, to ageing, homelessness, child protection and of course, disability.

Relevant AIHW capability

AIHW has the following capabilities that are relevant to the consultation and provide context to subsequent AIHW feedback:

1) Well established data collection and management infrastructure and expertise

Over time the AIHW has acquired considerable expertise and infrastructure allowing for the ongoing collection of administrative data across a range of

related human services sectors including health, community services, housing and disability. Key elements of this capability include technical infrastructure, staff expertise and well-developed and longstanding internal processes for data processing and database management.

2) Legislated authority to collect, manage and disseminate information.

As mentioned, the activities of AIHW are guided by the Australian Institute of Health and Welfare Act 1987 (AIHW Act). This act gives the AIHW legislated authority to perform the following functions as they related to health and welfare-related information: data standards development and maintenance, data collection and management, production of information and statistics, reporting to the Minister and the public, and dissemination of information and statistics to third parties.

3) A well-established and multi-faceted data supply chain.

AIHW has the experience of receiving and managing information from governments and non-government organisations, and has the flexibility to deal with a range of information collection types and processes, including both administrative and survey data. In order to efficiently gather the data we require in a timely and usable (i.e. standardised) manner, we collaborate closely and have effective data partnerships with many experts from around Australia, including the Australian Bureau of Statistics, governments at all levels, universities, research centres, and non-government organisations.

4) Leadership in data standards development and maintenance

The AIHW plays a significant leadership role in developing and maintaining national metadata standards that are essential to the development of robust statistical collections. This is achieved through active participation and leadership in national committees and our maintenance of METeOR (Metadata electronic Online Repository), Australia's repository for national metadata standards for health, community services and housing assistance information. The Institute works closely with governments and non-government organisations to achieve greater adherence to these standards in administrative data collections to promote national consistency and comparability of data and reporting.

5) Major supplier of information to a range of stakeholders, in multiple formats

AIHW plays a significant role in the dissemination of data and information for public policy, public reporting, research and to inform community debate. The Institute produces around 150 reports each year, and provides information that is made available to agencies for their own purposes, such as the COAG Reform Council (CRC) and the Report on Government Services (ROGS), as well as for broader research purposes.

Virtually everything that the Institute produces is publicly available and all our publications are available as free downloads from our website

(www.aihw.gov.au), which had more than 1.3 million visits in 2009-10. We have developed a more innovative presentation of hospitals information, the MyHospitals website (www.myhospitals.gov.au), and we have plans to release more of our products in an online format in the future.

6) Expertise in data linkage

The Institute possesses significant expertise and has extensive experience in undertaking data linkage for research into a wide range of health and welfare policy and research work.

In the context of the new Commonwealth data linkage framework, AIHW has expanded its capabilities in data linkage by becoming one of the first organisations accredited to undertake complex data linkage projects involving Commonwealth data. This accreditation gives assurance to data custodians and the public that high levels of security and strong governance processes are in place to protect the privacy of individuals, while maximizing the potential of available data sets held by the Institute to provide new insights into complex policy issues, thereby supporting better policy decisions.

7) Robust governance and oversight of information

AIHW has a long history of providing high quality information and data within an environment that protects the privacy of individuals. The AIHW is bound by the Privacy Act of 1988, which incorporates a series of Information Privacy Principles (IPPs) that regulate how the AIHW obtains, stores, uses and discloses personal information. The AIHW Act also contains stringent provisions for the protection of personal information, including criminal penalties for unauthorised disclosure. In addition, the AIHW Ethics Committee, provides oversight to the provision of data for research purposes, as well as to our internal activities including data linkage. Finally, in producing our reports and other information products the Institute employs numerous quality control checks to ensure the accuracy and the appropriate application of statistical approaches in the reporting of information.

8) Experience developing performance measures

One of our more recent key roles has been to develop, build and improve performance indicators and targets for national agreements between the Commonwealth and State and Territory governments; particularly in the areas of health care and hospitals, housing and homelessness, disability services and Indigenous health, welfare reform and disability. In this work, the aim is to provide more outcome-based and “actionable” measures of impact and performance. This is taking place not only at the level of COAG and ROGS reporting, but also within individual sectors and programs the National Disability Agreement auspices, where providers and policymakers alike, along with service recipients, are increasingly interested in this type of information.

More information about the Institute, in terms of its activities, capabilities and performance can be found on our website (www.aihw.gov.au/about) or in our 2010-11 Annual Report.

Consultation questions

A detailed response to the consultation paper is attached in the respondent information form and draws on the AIHW's expertise in the data linkage/integration area. We hope this provides some useful information from an international perspective and would be pleased to discuss any of the matters raised in our submission in further detail, at your convenience.

Yours sincerely



Mr David Kalisch
Director
Australian Institute of Health and Welfare

6 June 2012

CONSULTATION QUESTIONS

**Are you responding *primarily* as a data custodian, data user or data subject?
(We recognise all people are data subjects and many organisations act as data guardians and data users, but please tick only one box)**

- Data Custodian
- Data User (e.g. researcher)
- Data Subject (e.g. member of the public or group representing citizens)

Australian Institute of Health and Welfare - Organisation Overview

The Australian Institute of Health and Welfare (AIHW) is the national agency set up to provide information and statistics on Australia's health and welfare. We are a statutory authority accountable to the Australian Parliament, and are governed by a Board comprising independent and government members. Our work is used to inform debate and decisions on policy and services. We have a strong commitment to providing impartial and objective information and advice in all our work, and this independence is valued by stakeholders and users of our statistics.

We publish more than 150 reports each year, ranging from flagship publications to substantive reports, 'in brief' publications, journal articles and short discussion papers. Our flagships, [Australia's health](#) and [Australia's welfare](#), are published in alternate years. We have an expanding range of innovative web products, including the [MyHospitals website](#) reporting individual hospital data, the [Closing the Gap Clearinghouse](#), and online interactive health and welfare data sets.

We collaborate with a range of experts from around Australia, including Australian, state and territory government health, housing and community services agencies; universities and research centres; and peak non-government organisations. We collaborate closely with the Australian Bureau of Statistics, particularly in our use of Census and survey data. Researchers in universities and similar organisations carry out work relevant to our mission under collaborative arrangements that have the legislative backing of the *Australian Institute of Health and Welfare Act*.

We enjoy national and international recognition for our expertise in analysis and reporting of health, housing and community services information; our leadership in data integration; our management of national data collections and health registers; and the development of data set specifications, data definitions and classification systems. An important part of our work is promoting consistency among national, state and territory statistics so that we can produce and compare comprehensive national data of the highest standard. Our work also contributes to international statistics and draws on international standards reported by bodies such as the World Health Organization and the Organisation for Economic Co-operation and Development (<http://www.aihw.gov.au/capability-statement/>).

Expertise in data linkage

The Institute possesses significant expertise and has extensive experience in undertaking data linkage for research into a wide range of health and welfare policy and research work.

In the context of the new Commonwealth data linkage framework, AIHW has expanded its capabilities in data linkage by becoming one of the first organisations accredited to undertake complex data linkage projects involving Commonwealth data. This accreditation gives assurance to data custodians and the public that high levels of security and strong governance processes are in place to protect the privacy of individuals, while maximizing the potential of available data sets held by the Institute to provide new insights into complex policy issues, thereby supporting better policy decisions.

1. Are there any benefits of data linkage for statistical and research purposes that are not sufficiently described here?

Yes, there are further benefits No, the benefits are described fully

If you ticked 'yes', please describe the further benefits of data linkage for statistical and research purposes.

The benefits outlined in the consultation paper provide a good high level summary with appropriate examples. These support development of the secure Data Linked Framework for Scotland (this includes a National Data Linkage Centre and an Analytical Privacy Advisory Service) which will enable research to improve services for the Scottish population. Utilising the data linkage technology outlined in the consultation paper would help maximise the value of administrative data which is an important resource for both government and university researchers.

Further benefits of such a programme include improved research skills and the development of analytical tools for complex linked data will enable new research that enhances the delivery of public services. Research studies using similar international infrastructure to that proposed have contributed significantly to both services and outcomes.

Research using data linkage in Scotland (and overseas) has already proven powerful in not only influencing policy and improving public services. Data linkage provides information on whole populations that generates a more complete picture of the community than is possible using other research methods. It is also a very cost-effective research tool. Once the linkage infrastructure is in place, the cost of accessing linkable data becomes more affordable.

Some examples of where data linkage has been used to great effect across a range of health and health related areas can be found on the PHRN website

(www.phrn.org.au/about-us/what-we-have-learnt).

The proposal has parallels with a number of national and international projects which should be considered to maximise benefits, these include:

Australian Institute of Health and Welfare (AIHW)

The Institute possesses significant expertise and has extensive experience in undertaking data linkage for research into a wide range of health and welfare policy, and research work.

In the context of the new Commonwealth data linkage framework, AIHW has expanded its capabilities in data linkage by becoming one of the first organisation accredited to undertake complex data linkage projects involving Commonwealth data. This accreditation gives assurance to data custodians and the public that high levels of security and strong governance processes are in place to protect the privacy of individuals, while maximizing the potential of available datasets held by the Institute to provide new insights into complex policy issues, thereby supporting better policy decisions.

There is substantial public benefit to be gained from research using integrated data. This research methodology can identify evidence of cause and effect, and the nature and strengths of relationships over time and across traditionally separated domains of data collection. The demand to create and provide access to integrated data is growing, and the number and breadth of projects involving data integration (both undertaken at AIHW and more generally) is expanding rapidly.

SAIL Databank – Wales

The Secure Anonymised Information Linkage (SAIL) Databank brings together the widest possible range of person-based data using robust anonymisation techniques for health related research. It is operated by the Health Information Research Unit (HIRU) based at Swansea University. HIRU aims to realise the potential of electronically-held, person-based, routinely-collected data to conduct and support health-related studies. The Databank holds over 500 million records and continual growth is in progress.

The Secure Anonymised Information Linkage (SAIL) system has been designed and developed to: 1) ensure data transportation is secure; 2) operate a reliable record matching technique to enable accurate record linkage across datasets; 3) anonymise and encrypt the data to prevent re-identification of individuals; 4) apply measures to address disclosure risk in data views created for researchers; 5) ensure data access is controlled and authorised; 6) establish methods for scrutinising proposals for data utilisation and approving output; and 7) gain external verification of compliance with Information Governance.

Population Health Research Network (PHRN) – Australia

The Population Health Research Network (PHRN) has been established to develop data linkage units in all States/Territories around Australia supporting nationally and internationally significant population-based research that will improve health and enhance the delivery of health care services in Australia.

The PHRN comprises a Program Office located in Perth, Western Australia, a Centre for Data Linkage (CDL) at Curtin University in Western Australia, a Remote Access Laboratory at the Sax Institute in New South Wales and a network of Project Participants and Data Linkage Units located in each Australian state/territory (including the Data Linkage Branch in the Western Australian Department of Health and SA/NT DataLink).

The PHRN has the objective of building a national infrastructure for the development and promotion of data linkage for population and clinical health datasets for research purposes. This network includes national elements including the National Linkage System (Centre for Data Linkage, Curtin University) and the Secure Unified Research Environment (Sax Institute, Sydney).

The Secure Unified Research Environment (SURE) is a remote-access computing environment that allows researchers to access and analyse linked health-related data files for approved studies in Australia. SURE has been developed by the Sax Institute as part of the Population Health Research Network (PHRN) initiative.

Data Linkage Australia (DLA) – Australia

The Data Linkage Australia (DLA) draws on the unique competitive advantages possessed by Western Australia in its ability to attract external science investment, due to its possession of comprehensive, population-based health databases, and recent local achievements in community leadership and medical informatics that have created individual identity-relationships between millions of database records. The collaboration brings together a multi-disciplinary group of science and technology leaders who have developed infrastructure that supports population based research (and research into cross portfolio development pathways).

2. Are there challenges or barriers preventing more effective and efficient data linkages for statistical and research purposes taking place that are not sufficiently described here?

Yes, there are further challenges No, the challenges have been identified

If you ticked 'yes', please describe the challenges or barriers.

The framework should build on existing national and international collaborations, infrastructure and skills – it should leverage these developments. The success of the programme will depend on effective collaboration between the various partner organisation and stakeholders.

Typical inhibitors and barriers to data linkages for research and statistical purposes include:

- The usability of the systems - it is very often time-consuming to establish data linkage projects in terms of legal and governance arrangements. A process that manages/streamlines all the processes involved in executing a linkage project (from start to end) would ensure effective and efficient data linkages (these processes need to address any uncertainty around legal, privacy and data custodian constraints);
- Efficient and effective technology - the infrastructure involved in undertaking data linkages need to be scalable, fast and efficient to ensure timeous response to important policy and research questions (this includes application, workflow, linkage, metadata and analysis);
- Skills and/or culture – there is often a lack of understanding of what's available, how it can be accessed and the complexity of analysing linked information (analytical skills for complex linked data may need strengthened);
- Security and privacy – the design of any developments need to facilitate safe and secure access to linked data (i.e. that they apply the appropriate governance process, protocols and standards);
- Cost effective national infrastructure – development and maintenance of the data linkage infrastructure should be relatively inexpensive (and can leverage national and international experience).

The challenges for the framework include:

- Creating transparent assessment and approval processes for data linkages;
- A development/design which provides comprehensive security features and makes collaborations between researchers easier and more efficient;
- The facility needs to address privacy and confidentiality concerns associated with the use of data for research projects and facilitating the use of this data within the constraints of privacy, confidentiality and the law;

- Independent assessment of the balance of public good against the privacy imposition and risks to confidentiality. Examples include community representation on steering committees, the use of an ethics committee, or an advisory committee with community representation and the ability to report independently to the agencies involved in the proposal.
- Providing researcher with skills, analytical software and tools to deliver high impact studies.

3. Are the guiding principles sufficient and appropriate? Please explain your answer fully and make suggestions for improvement.

Yes, they are sufficient and appropriate No, they are not

Please explain your answer fully and make suggestions for improvement.

The statistical and research value of administrative data should be maximised, within legal and practical constraints, by granting broad access for research purposes to data that is not likely to enable identification. Design mechanisms to achieve adequate control and manage risk including the use of specified standards and audits against those standards, and the potential application of sanctions should be considered.

Developing a strong Information Governance framework throughout the infrastructure will be necessary to ensure the required safeguards for, and appropriate use of, patient and personal information.

The framework could consider the following areas to be included in the principles:

- The risks of indirect as well as direct identification should be carefully managed when data are disseminated. This management must take account of the potential increase in identifiability of one set of data when combined with another set. It might involve strict data use licensing conditions, reducing detail, perturbing data, or seeking the consent of the individual or business involved to release potentially identifiable data, the last of these being most likely in the case of business data;
- Process to agree the nature of valid uses that can be made of the linked datasets and the approval mechanism to be applied to applications using the datasets, as well as any control mechanisms to be applied to such use;
- Removing personal identifiers as soon as they are no longer required to meet the approved purposes of the statistical data integration. Where identifiers need to be retained, for example for longitudinal studies, they could be kept separate from the integrated dataset;

- The type of matching used (deterministic, probabilistic or statistical) should be chosen as the minimum needed to support the approved purposes, and the range of attributes used to establish a common identity should be the minimum necessary for the linking operation to succeed;
- Developing minimum standards for the secure management of information and Information Governance including ISO/IEC Information Security standards (covering: Information technology/Security techniques/Information security management systems);
- Management of data access (and risk of privacy and/or confidentiality breaches) for research projects ensuring/considering:
 - o Confidentiality and privacy
 - o Data Protection
 - o Information security
 - o Record Management (Data retention/destruction policies are applied)
 - o The international information security standard e.g. ISO/IEC 27002: 2005
 - o Data access agreements

4a. Are the objectives set out for a Privacy Advisory Service in Section 3c the right ones?

Yes, the objectives are right

No, they are not

Please explain your answer fully and make suggestions for improvement.

The creation of a Privacy Advisory Service addresses some of the challenges associated with facilitating data linkage research and provides a focal point for ethical, legal and social issues associated with data linkage projects. The objectives outlined for the Privacy Advisory Service includes Client Services functions which will be crucial in making data linkage service availability and quality data linkage projects possible (establishing a greater level of transparency throughout the process).

4b. Do you wish to be consulted on firmer proposals for a Privacy Advisory service as and when they are developed?

Yes No

5a. Are the functions that will be led by the National Data Linkage Centre set out in section 3d the right ones?

Yes, they are the right functions

No, they are not

Please explain your answer fully and make suggestions for improvement.

The proposal outlines a similar infrastructure, training and support model to many national and international examples. As a comparison, the AIHW Data Linkage Unit (which is a National Data Linkage Centre) facilitates the development and analysis of person-centred data to support whole-of-government and whole-of-life approaches to policy. The Unit achieves this by investigating data linkage and analytical methods, by undertaking data linkage and analyses of linked datasets, and by providing leadership and assistance to analyses undertaken elsewhere within the AIHW in close collaboration with subject matter staff. The Unit is the main point of contact with the wider national and international data linkage communities.

Objectives of the AIHW Data Linkage Unit:

- Develop the AIHW's capacity to fill information gaps through the continued enhancement of data linkage and analytical methodologies;
- Increase confidence in, and acceptance of, data linkage by explaining its benefits and by publicising the AIHW's strong privacy protections;
- Develop better ways of presenting key findings from linked data and integrating the findings with other statistical analyses in the AIHW's reports;
- Enhance methods and systems to create datasets required for the AIHW's expanded program of linkage-based statistical analysis and research;
- Support the production of those Council of Australian Governments (COAG) performance indicators that will be derived from linked data views of Australians' health and welfare, by undertaking and supporting analyses of linked data.

The National Data Linkage Centre in Scotland needs to provide a high-performance data linkage environment with enhanced speed, storage capacity. Although the concept is not new, designing and building an infrastructure with comprehensive security features will be a challenge.

For more information on Data Linkage/Integration at AIHW please contact:

Phil Anderson (andersop2@aihw.gov.au) or
James Boyd (james.boyd@aihw.gov.au).

5b. Do you wish to be consulted on firmer proposals for a National Data Linkage Centre as and when they are developed?

Yes

No