



2.5 Secondary use of health information

Every day, the clear majority of Australians log on to some form of digital technology, such as a smartphone or a tablet. In 2014–15, 86% of Australian households—including 97% of those with children under 15—had internet access (ABS 2016). This 'doorway' to the digital world has brought with it unprecedented access to vast volumes of data—including data on our health and welfare.

Health data relate to information gathered on people's health (and factors that affect it) and the health system. These data can be collected from a diverse range of sources, including the health care system itself, health practitioners, medical research, or individuals tracking their own health. For example, wearable trackers have been used to track and increase physical activity (Kaiser et al. 2016) and social media data have been used to provide early insight into disease outbreaks (Brown 2015).

The amount of health data generated is increasing rapidly. One report estimated that the global amount of health data being generated grew by 48% every year, and that the amount of health data is expected to exceed 2,000 exabytes (2 billion terabytes) by the year 2020 (EMC 2014).

The growth in digital technology and 'digital health'—the term used when this technology is used to collect and share a person's health information (see Chapter 2.4 'Digital health')—has increased the volume of data available for health research.

This, in turn, has increased opportunities for 'secondary use' of health data to improve the health system and the wellbeing of Australians.

What is meant by secondary use?

'Secondary use' of data refers to any application of data beyond the reason for which they were first collected (known as the primary use or purpose). For example, the primary use of data collected to treat a patient in a hospital is to provide the patient with the care they need in that hospital episode; a secondary use could be to aggregate patients' data to compare hospital performance across Australia.

Secondary use of data presents an enormous opportunity to improve the health treatment that Australians receive. Health data can be used in this way to monitor and improve hospital performance—this can include improving flagging facilities (where this can be done) and identifying top-performing hospitals. Processes and practices of top-performing health facilities can then be adopted by others, thereby improving quality of care overall.

However, use of an individual's health data to enable these improvements must be balanced against the risk to their privacy. Australia has many legal requirements and systems in place to ensure that privacy is protected.



This article focuses on two aspects of secondary use of health data: the insights that can be gained from secondary use of administrative data, and how data integration can be used to gain a deeper and broader view of the health system and the many factors that affect a person's health outcomes. Aspects of privacy protection with the processes that support these uses are also discussed.

Insights from administrative data

Health care in Australia is delivered by a variety of professionals in many different settings—from state-of-the-art operating rooms to care provided in a patient's home. Administrative data can be considered as data collected from running these services and programs. In the case of health, these cover data from hospitals, health professionals, pharmacies and allied health care.

Using these data for secondary purposes can help to glean a holistic overview of the health system and the pathways that individuals take when using health services. For example, access to the data allows a variety of analyses to be undertaken, including examining trends in health expenditure, examining patient outcomes from particular treatments, looking at potential risk factors and determinants of health and disease, and tracking hospital wait times. Another application of secondary use is identifying, and then inviting, patients for clinical trials and studies.

Using this vast amount of data has the potential for improvements in the quality and safety of health care—and in clinical and administrative processes—as well as for better targeting of resources. Ultimately, the potential is there for a healthier population.

Many of the chapters in this report show the wide application and benefits of secondary use of data for research purposes.

Data linkage

The value of administrative data increases when individual data sets are linked to provide a more robust picture of the people using the services and the system. Data linkage and data integration combine information from two or more data sets to create a new, more detailed data set that can tell a much more meaningful story than would be possible from a single data source. This may involve bringing together data from the same source over different time periods to see the effect of events that have occurred across a person's life, or bringing together two different data sets to get a broader understanding of factors that can affect a person's health. For example, researchers at the University of New South Wales are examining the factors that lead to poor outcomes following prescription opioid use by linking Medicare, pharmaceutical, mortality and cancer data sets (AIHW 2017a). Bringing these data together will provide researchers with a much more comprehensive view than if they were to examine each data source individually.

Today's unprecedented volume, diversity and speed of data generation present many new opportunities for data linkage. As more health data become available, academics, researchers and organisations such as the AIHW are looking at how data linkage can provide insights into the health of Australians and the performance of the health system.



Box 2.5.1 provides two examples of data linkage projects which have enabled better responses to health problems.

The AIHW is one of a small number of accredited integrating authorities in Australia—the others include the Australian Bureau of Statistics and the Australian Institute of Family Studies (NSS 2014). These organisations can undertake high-risk data integration projects using Commonwealth data as they have strict privacy and confidentiality controls in place. There are also state-based data linkage units that link state and territory data for research purposes. These units, along with the AIHW, are members of the Population Health Research Network, a collaboration to build data linkage infrastructure at the national level (Productivity Commission 2017).

Box 2.5.1: Data linkage—two case studies

In 2017, the AIHW first published a report providing new insight into suicide among contemporary Australian Defence Force (ADF) personnel (with at least 1 day of service since 1 January 2001). The study constructed a data set by linking ADF personnel administrative data and the National Death Index. This linked data set was then supplemented by information on confirmed suicide deaths from the Defence Suicide Database and cause of death information from the National Mortality Database. This linkage allowed the AIHW to identify characteristics of the contemporary ex-serving population associated with suicide risk (see Chapter 5.6 'Veterans'). Following from this, the AIHW and the Department of Veterans' Affairs have now begun a 3-year partnership to monitor veteran suicide rates and develop a comprehensive profile of the health and welfare of Australian veterans and their families to understand their current situation and future needs (AIHW 2017b).

Researchers from the University of South Australia, the University of Adelaide and the Flinders Centre for Innovation used data linkage to uncover important insights into colorectal cancer treatment patterns. The researchers linked data from the South Australia Cancer Registry to hospital and radiotherapy data. The Population Health Research Network carried out the linkage, ensuring that the researchers themselves did not access any information that could identify individuals. The study found that recommended treatments were less likely to be performed on people who were older, who lived in rural areas or who had severe or multiple diseases. This information promises to improve outcomes by highlighting a segment of the population that can be targeted to ensure that they receive the recommended treatment for colorectal cancer (Beckmann et al. 2014).

Programs such as the Data Integration Partnership for Australia are developing ways to efficiently derive the most societal benefit from new linkage opportunities, while maintaining the consideration and protection given to privacy. The program brings together agencies from across the Australian Government to work together on data integration projects that will better indicate how factors that influence health and welfare affect outcomes for Australians.



Taking a person-centred approach to data

The increasing availability of digital data comes with a community expectation that public data will be used to improve outcomes for Australian people. There is now growing interest in person-centred data—a model that focuses on the individual and recognises that the various aspects of a person's life interact, and are influenced by, their own specific circumstances. For example, a person's health outcomes are affected by multiple influences, including their genetic make-up; where they live; their housing and employment; how much they earn; and a multitude of other biological, lifestyle, socioeconomic, societal and environmental factors.

Privacy, security and current protections

Governments, organisations and individuals now have unprecedented access to a growing volume of digital data. The ever-increasing delivery of faster, high-quality, useable information comes with the obligation that public data are held securely, and that individual privacy is protected. The Productivity Commission's *Inquiry on data availability and use* suggests that, on the whole, people are willing to make their information available, but this willingness is conditional. The public needs to trust how data are handled; they also need to feel that they have some control over how they will be used and who will use them, and to see (and ideally directly benefit from) the potential value of this use (Productivity Commission 2017).

A 2017 survey showed that 93% of Australians support using medical records for research. As well, more than 95% of survey respondents had moderate, high or very high trust that researchers would use health data responsibly (Research Australia 2017).

Australia has many legislative arrangements to protect individual privacy. These protections ensure that the privacy of individuals is respected, while ensuring that data are available to use for public benefit, particularly in health research. The *Privacy Act 1988* (Cwlth) recognises that it is often impractical or impossible for researchers to obtain people's consent for the use of their data in specific research projects. As such, a set of guidelines produced by the Privacy Commissioner must be followed by any researcher approved to use health data without patient consent. These guidelines also assist human research ethics committees in deciding whether research projects should be approved (OAIC n.d.).

Commonwealth Integrating Authorities adhere to strict protections and data standards when creating linked data sets. At the AIHW, an Ethics Committee reviews each linkage project to ensure that it complies with applicable legislation and individual privacy guidelines. As well, many research data sets are adjusted before release to researchers to reduce the risk of disclosure of personal information. These adjustment processes can include de-identification, where information that would identify individuals is removed to protect an individual's privacy.



Further privacy protection is ensured by adhering to the 'separation principle'. This principle, which is embedded in the processes of all integrating authorities, requires that directly identifiable information (for example, names, addresses) and content data (for example, instances of services used) are kept separately. This means that the people able to identify subjects of the data are not able to learn any information about them, further minimising any risk to people's privacy.

Other codes and guidelines ensure research integrity and the protection of privacy. The Australian Code for the Responsible Conduct of Research provides a framework for institutions engaged in research to develop appropriate procedures and codes of conduct (NHMRC 2016). The National Statement on Ethical Conduct in Human Research contains guidelines used by researchers and ethics committees. These guidelines cover areas such as ethical review, risk management, patient consent and lines of responsibility (NHMRC 2013).

Where to from here?

Technology—in particular, digital technology—has had a huge impact on data generation, access and availability. The increasing opportunities for secondary use of data promise that Australians' health can be improved through the appropriate use of high-quality, useable information from new and existing sources. Two of the biggest influences on the future of secondary use of health data will be the growth in data availability and the trend towards 'open data'. These trends come with increased risks to privacy and steps need to be taken to ensure privacy protections keep pace with technological advances.

More health data available

The increased use of technology in providing health care is one of the biggest drivers behind the growth in health data. Electronic patient medical records, such as the Australian Government's My Health Record, are 'one of the most visible aspects of the increasing proliferation of data in health care' (Stanford Medicine 2017:14). The My Health Record is a secure online summary that combines health data from multiple sources (see Chapter 2.4 'Digital health'). By the end of 2018, every Australian who has not opted out of the program will receive their own My Health Record (ADHA 2018).

While the primary use of My Health Record is to store and make available individuals' medical information, it also provides a rich opportunity for secondary use of data. At the time of writing, there have been no secondary uses of these data. However, the Framework to guide the secondary use of My Health Record system data was released and the AIHW was appointed to manage and release My Health Record data for secondary purposes (Department of Health 2018).

Further to this, the all-encompassing presence of technology in our daily lives—wearing pedometers and heart rate monitors on our wrists, and tracking our exercise from smartphones—means that the potential for data provision and use is expanding rapidly. The share of global health data taken up by connected devices is expected to more than double between 2016 and 2020 (EMC 2014). Patient monitors and video cameras are examples of these connected devices—they are components of the so-called 'Internet of Things'. The data generated by these devices offer a rich opportunity for secondary uses.



Open data

As more data become available, greater access—particularly to person-centred data—will continue to unlock added value from these data assets. Open data are used to describe data sources that have little to no restrictions on their access and use. However, due to the sensitivity and personal nature of many types of health data, most individual-level health data will not be open and must be treated with appropriate respect for an individual's privacy. In practice, there are still many open data sources relevant to health research; for example, data bases on the weather are relevant to researching seasonal illnesses such as influenza.

There has been a push for increasing access to government data and the availability of open data resources. Technological developments can enable personal health information, such as that collected by mobile devices, to be combined with open data sources to provide more relevant real-time data to users, practitioners and researchers. Where health data sources are not classified as open data, de-identified views and extracts may be prepared to be compatible with a wider variety of future uses.

Challenges ahead

The opportunities presented by the increasing availability and scope of health data are enormous, but there are also challenges. As this article has stated, one of the main challenges is how to ensure that potential gains of data sharing and secondary use do not impinge on the privacy of individuals.

The variety of different types of information poses a challenge when it comes to the use of the data. Secondary use of administrative and other data sets can be very insightful, given their reach and scope, but these data also have limitations. Examples of limitations can include poor documentation, small numbers of data items, and uncertainty about the quality of information (Jorm 2015). Assuring that health data collections are inter-operable ensures that the technical challenges inherent in using research data are minimised, and that the promise of secondary data use can be realised.

What is missing from the picture?

While the potential benefit from secondary use of health information is great, it is dependent on an understanding of the health information landscape. This requires knowledge of what health data are currently collected, where gaps and overlaps exist and the priority areas for researchers and policy makers. There is currently no holistic approach to implementing a national health information strategy. A coordinated, strategic approach to identify gaps, overlaps and priorities would further enhance the use of national health data assets.

Where do I go for more information?

More information on AIHW, what we do, our privacy policy and our data governance is available at <www.aihw.gov.au/about-us>.

The Framework to guide the secondary use My Health Record system data is available at <www.health.gov.au/internet/main/publishing.nsf/Content/eHealth-framework>.



References

ABS (Australian Bureau of Statistics) 2016. Household use of information technology, Australia, 2014–15. ABS cat. no. 8146.0. Canberra: ABS. Viewed 29 January 2018, <<http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/8146.0Main+Features12014-15?OpenDocument>>.

ADHA (Australian Digital Health Agency) 2018. Welcome to My Health Record. Canberra: ADHA. Viewed 29 January 2018, <<https://myhealthrecord.gov.au/internet/mhr/publishing.nsf/content/home>>.

AIHW (Australian Institute of Health and Welfare) 2017a. Approved AIHW linkage projects. Canberra: AIHW. Viewed 29 January 2018, <<https://www.aihw.gov.au/our-services/data-linkage/approved-aihw-linkage-projects>>.

AIHW 2017b. Incidence of suicide among serving and ex-serving Australian Defence Force personnel 2001–2015: in brief summary report. Cat. no. PHE 213. Canberra: AIHW.

Beckmann KR, Bennett A, Young GP & Roder DM 2014. Treatment patterns among colorectal cancer patients in South Australia: a demonstration of the utility of population-based data linkage. *Journal of Evaluation in Clinical Practice* 20:467–77.

Brown J 2015. Using social media data to identify outbreaks and control disease. Folsom, CA: Government Technology. Viewed 29 January 2018, <<http://www.govtech.com/data/Using-Social-Media-Data-to-Identify-Outbreaks-and-Control-Disease.html>>.

Department of Health 2018. Framework to guide the secondary use of My Health Record system data, May 2018. Canberra: Department of Health.

EMC Corporation 2014. The digital universe: driving data growth in healthcare. EMC Corporation. Viewed 29 January 2018, <<https://www.emc.com/analyst-report/digital-universe-healthcare-vertical-report-ar.pdf>>.

Jorm L 2015. Routinely collected data as a strategic resource for research: priorities for methods and workforce. *Public Health Research and Practice* 25(4).

Kaiser DW, Harrington RA & Turakhia MP 2016. Wearable fitness trackers and heart disease. *JAMA Cardiology* 1(2):239.

NHMRC (National Health and Medical Research Council) 2013. Summary of national statement content. Canberra: NHMRC. Viewed 28 January 2018, <<https://www.nhmrc.gov.au/research/responsible-conduct-research/summary-national-statement-content>>.

NHMRC 2016. Australian code for the responsible conduct of research. Canberra: NHMRC. Viewed 28 January 2018, <<https://www.nhmrc.gov.au/guidelines-publications/r39>>.

NSS (National Statistical Service) 2014. Accredited integrating authorities. Canberra: NSS. Viewed 28 January 2018, <<http://www.nss.gov.au/nss/home.nsf/pages/Data%20Integration:%20Accredited%20Integrating%20Authorities>>.

OAIC (Office of the Australian Information Commissioner) n.d. Health information and medical research. Canberra: OAIC. Viewed 28 January 2018, <<https://www.oaic.gov.au/privacy-law/privacy-act/health-and-medical-research>>.

Productivity Commission 2017. Data availability and use. Report no. 82, Canberra: Productivity Commission.

Research Australia 2017. Australia speaks! Research Australia 2017 public opinion polling. Darlinghurst: Research Australia. Viewed 29 January 2018, <<https://researchaustralia.org/wp-content/uploads/2017/08/2017-Opinion-Poll-Digital.pdf>>.

Stanford Medicine 2017. Harnessing the power of data in health. Viewed 28 January 2018, <<https://med.stanford.edu/content/dam/sm/sm-news/documents/StanfordMedicineHealthTrendsWhitePaper2017.pdf>>.