Towards sharing health-services data

Data sharing would speed progress in improving the Australian health system

Michael D Coory
MB BS, FAFPHM, PhD, Public Health Physician[1]

Craig L Fry
BSc(Hons), DrPH, NHMRC Career Development Fellow (Population Health),[1] and Associate Professor/Principal Fellow[2]

1 Murdoch Childrens Research Institute, Melbourne, VIC.
2 Centre for Applied Philosophy and Public Ethics, University of Melbourne, Melbourne, VIC.

Michael.Coory@mcri.edu.au

doi: 10.5694/mja11.11238

The benefits of data sharing have been demonstrated in genomics, where it has accelerated progress and provided an understanding of complex issues sooner.[1] Hoping to emulate this success, 17 funders of research from around the world (including the National Health and Medical Research Council [NHMRC]) recently signed a joint statement committing to do all they can to encourage data sharing from the research that they fund.[1]

The joint statement clarifies that this noteworthy commitment to data sharing does not cover routinely maintained databases to which the signatories contribute no funding.[1] In Australia, these include a multitude of health-services databases that are maintained by various organisations (eg, federal and state health departments, private health funds, private hospital groups).

For the type of research the NHMRC typically funds, the path to data sharing has been comprehensively discussed, particularly overseas.[2] Informed consent provided by research participants must allow for data sharing; there should be no risk of breaching confidentiality; and appropriate attribution should be given to the research team that originally established and continues to maintain the database. These changes imply a fundamental change to the culture of research.

However, for health-services data, the path to data sharing will be different, and has not been comprehensively discussed. For example, consent is not routinely obtained; the data are often collected under federal or state legislation; and there is often a formally appointed data custodian who decides whether or not to release the data.

In this article, we aim to provide an understanding of the benefits of and barriers to sharing health-services data in Australia. Our main argument is that maintaining current barriers to data sharing will mean unnecessary delays and increased costs in producing the research needed to improve the Australian health care system.

Benefits

Data sharing is the process of facilitating access to data by research groups and other stakeholders (eg, policymakers or budget-holders) who were not involved in collecting the original data.[1] This includes sharing data both within and between countries.[2] In this article, we concentrate on sharing Australian data within Australia.

Australian health-services databases include the National hospital morbidity database (NHMD), National elective surgery waiting times data collection and National non-admitted patient emergency department care database.[3] Their data could be put to many uses. For example, the NHMD and other large health-service databases will be used to establish the *efficient price* for different types of public hospital admissions through analysis of actual activity and costs. This benchmarking is the cornerstone of activity-based funding and the hospital agreement between the federal and state governments.[3] If all stakeholders agreed to wide sharing of these databases, this funding mechanism could be optimised.

More generally, the benefits of sharing health-services data include:

- better understanding of how the health system is performing;
- faster production of information to improve the health system;
- increased transparency about the underlying assumptions of statistical methods;
- faster development of innovative statistical methods; and
- increased opportunities for training the next generation of health-policy researchers.

One way to increase the utility of routine health-services databases is to link information about the same patient across different databases. This requires sharing among the relevant data custodians (eg, federal and state governments). There is well established linkage infrastructure in Western Australia[4] and New South Wales,[5] and similar infrastructure is being developed in other states and nationally.[6,7] Although the increased utility of linked versus unlinked data is usually listed as a benefit of data sharing,[2] such linkage does not guarantee that the data will be shared beyond the sharing required for linking. That is, all the barriers to sharing unlinked data (discussed in the next section) also apply to linked data.

Barriers

The main barriers to sharing health-services data in Australia are issues related to waiving consent, a limited understanding of patients' attitudes, difficulties interpreting the legislation, and the vagaries of some data-custodian approvals. Perhaps the least problematic of these is waiving consent, with the NHMRC's *National statement on ethical conduct in human research* providing clear guidelines.[8] Ethics committees can waive consent if: there is sufficient protection of patients' privacy; there is no patient contact; there is little or no risk to the patients' physical or psychological health; and any risk (if present) is outweighed by the benefits of the research.[8]

In relation to waiving consent, it must be noted that health-services databases are by-products of direct patient care, and only collect information that is already in the patient's clinical notes. Thus, biobanks are not health-service databases because they collect tissue specimens, and consent is required for initial collection and any subsequent sharing of data. Similarly some clinical registries collect quality-of-life data through contact with patients after discharge, and this would require consent, or at least arrangements that allow patients to opt out and have their data removed or not made

available. Also not considered as health-services data are personally controlled electronic health records, whereby patients consent to having parts of their medical history shared among service providers. Use of such data in research would require specific additional consent.

Compared with waiving consent, there is currently less clarity about patients' attitudes to data sharing.[9] More work is needed, although patients' concerns about data security might be allayed by better software algorithms (which are now available) to render data non-identifiable but still useful for research.[5,10]

More work is also needed on the intent of the legislative and regulatory arrangements that cover the collection of health-services data, because there is very little case law to guide interpretation. Some work has been done by the Australian Law Reform Commission, which found that all the various pieces of federal and state legislation and regulations permit data sharing for bone fide research.[11]

Perhaps the most problematic barrier to data sharing is data-custodian approval. Despite the state-specific examples of successful data linkage and sharing in WA[4] and NSW,[5] the limited Australian research available suggests that there is a degree of arbitrariness about approval process and about what can subsequently be published.[12,13] Also, when data are linked across jurisdictions or health care settings (eg, primary and hospital care), it can be difficult to identify the relevant data custodian, and there is often no clear pathway for appealing refusal of access or unnecessary delays in access to data by data custodians.

There is a larger body of research on data-custodian approval from overseas, which suggests that data custodians in government tend to be overly risk averse and, if in doubt, tend not to share data.[14] More specifically, some overseas commentators have suggested that government officials are reluctant to share data because they do not want to expose the data to interpretations that compete with their own interpretation.[2]

In Australia, the mission of the Council of Australian Governments Reform Council is "strengthening public accountability of the performance of governments through independent and evidence-based monitoring, assessment and reporting".[3] This process would be greatly enhanced if the data were shared, thereby harnessing the intellectual capital of our universities, research institutes, primary health care system and hospitals. However, it is unclear how a research team would obtain data to participate in this process.

Another barrier to sharing, identified in the international literature, is that some data custodians consider interpretation of the data they manage to be so dependent on understanding the local circumstances of data collection that the data are worthless for other research. A related concern is that secondary users might spot the errors that are an inevitable by-product of data collection in the real world.[2] Ideally, sharing of data should, over time, improve its relevance and quality, with researchers collaboratively providing feedback about data quality to those who collect the data.

Conclusion

Providing affordable, high-quality health care for all patients is a complex challenge facing all health care systems around the world.[15] Solutions applicable in Australia will be context specific; that is, we will not be able to rely on research done overseas.

Given that research capacity is finite, sharing of health-services data would mean that faster progress could be made in improving the accessibility, quality and cost-effectiveness of the Australian health care system. One guiding principle is that analyses of data, which are invariably expensive to collect, should not be constrained by the capacity of the group that collected or linked the data.[2]

Sharing health-services data requires transparent policies and procedures. Such governance arrangements need not add to the administrative burden of research. For example, the Heritage Provider Network in California (of 2100 primary care physicians and 30 000 specialists) is running a competition to develop a statistical model to identify patients most at risk for admission to hospital (and then provide them with the ambulatory care they need to avoid admission).[16] Unit-record, non-identifiable data are available to research groups on the internet, which shows that health-services data can be shared with a minimum of administrative overhead.

The data that governments collect, and do not fully use, do not come free. The collection is paid for with taxpayers' money, and impediments to data sharing have an opportunity cost. Specifically, they delay understanding of how the health system is working and how to improve it. As has previously been argued, it is irrational for governments to invest so much in collecting data, but so little in facilitating the best use of it.[2]

Acknowledgements: We are supported by the Murdoch Childrens Research Institute (Victorian Government Operational Infrastructure Support Program).

Competing interests: No relevant disclosures.

Provenance: Not commissioned; externally peer reviewed.

1 Walport M, Brest P. Sharing research data to improve public health. *Lancet* 2011; 377: 537-539.
2 Pisani E, AbouZahr C. Sharing health data: good intentions are not enough. *Bull World Health Organ* 2010; 88: 462-466.
3 COAG Reform Council. National Healthcare Agreement: Performance report for 2009–10. Canberra: CRC, 2011. http://www.coagreformcouncil.gov.au/reports/healthcare.cfm (accessed Nov 2011).
4 Stanley F. Privacy or public good? Why not obtaining consent may be best practice. *Significance* 2010; 7: 72-75. doi: 10.1111/j.1740-9713.2010.00423.x.
5 The Centre for Health Record Linkage. Centre for Health Record Linkage: the first three years, 2006–07 to 2008–09. Sydney: CheReL, 2009, http://www.cherel.org.au/media/13571/cherel-the-first-three-years-2006-07-to-2008-09.pdf (accessed Nov 2011).
6 National Collaborative Research Infrastructure Strategy. Population Health Research Network. http://ncris.innovation.gov.au/Capabilities/Pages/PopHealth.aspx (accessed Mar 2012).
7 National Statistical Service. Vision statement: http://www.nss.gov.au/nss/home.nsf/pages/Vision+Statement?OpenDocument (accessed Nov 2011).
8 National Health and Medical Research Council. National statement on ethical conduct in human research 2007 [updated 2009]. Canberra: NHMCR, 2007. http://www.nhmrc.gov.au/guidelines/publications/e72 (accessed Mar 2012).
9 O'Keefe CM, Connolly CJ. Privacy and the use of health data for research. *Med J Aust* 2010; 193: 537-541.
10 El Emam K, Dankar FK, Issa R, et al. A globally optimal k-anonymity method for the de-identification of health data. *J Am Med Inform Assoc* 2009; 16: 670-682.
11 Australian Law Reform Commission. The Privacy Act and health information. In: Australian Privacy Law & Practice Report (ALRC Report 108). Canberra: ALRC, 2008. http://www.alrc.gov.au/publications/report-108 (accessed Nov 2011).
12 Holman CD. An end to suppressing public health information. *Med J Aust* 2008; 188: 435-436.
13 Yazahmeidi B, Holman CD. A survey of suppression of public health information by Australian governments. *Aust N Z J Public Health* 2007; 31: 551-557.
14 Pullman D, Buehler SK, Felt L, et al. Sorry, you can't have that information: data holder confusion regarding privacy requirements for personal health information and the potential chilling effect on health research. *Healthc Policy* 2009; 4: 61-76.
15 Straus SE, Tetroe J, Graham I. Defining knowledge translation. *CMAJ* 2009; 181: 165-168.
16 Heritage Provider Network. Health Prize Competition, 2011. http://www.heritagehealthprize.com (accessed Nov 2011).