



Linking data to improve health outcomes

Routinely collected data, when linked, are a rich source of sound evidence for making health care decisions

Ian N Olver
MBBS, MD, PhD,
Chief Executive Officer

Cancer Council Australia,
Sydney, NSW.
ian.olver@cancer.org.au

doi: 10.5694/mja14.00374

Linkage of health records in Australia was initiated in Western Australia in the late 1970s under the leadership of Michael Hobbs, who, a decade earlier, had been associated with the Oxford Record Linkage Study in the United Kingdom.^{1,2} The ability to link datasets is key to providing the evidence to direct health policy towards achieving better medical outcomes, and two studies published in this issue of the *MJA* are excellent examples of this. Both link data from the New South Wales Central Cancer Registry, the National Death Index and the NSW Admitted Patient Data Collection.^{3,4}

Smith and colleagues compared hospital caseload with outcomes of oesophago-gastric cancer surgery.³ Their data showed improved 5-year survival when the initial surgery was performed in hospitals with higher volumes. Similarly, a previously published meta-analysis from a systematic review of the literature between 2000 and 2011 found a benefit to centralising oesophageal surgery to high-volume institutions, although the end points of in-hospital mortality and 30-day mortality were more immediate.⁵ Based on such evidence, health service managers could restrict oesophagectomies and gastrectomies to high-volume hospitals rather than sites being determined by surgeons' or hospitals' preferences.

Jorgensen and colleagues further demonstrated the utility of data linkage by comparing variation in colorectal cancer care between hospitals through measuring patient outcomes and care processes.⁴ Outcome end points included 30-day mortality and complications, and the care process measures included time to referral for treatment, discussion at multidisciplinary team meetings and documentation of disease stage. In showing that much of the variation between hospital outcomes was due to differing patient characteristics, this study demonstrates that poorer outcomes in hospitals must be further analysed to include patient characteristics in risk-adjustment models. This finding reflects previous reports that social, hospital and clinical characteristics can all impact on colorectal cancer outcomes.⁶

Both studies confirm that linked datasets are a valuable tool for health services planning. Linked data can be used to evaluate the effectiveness and quality of cancer services, the uptake of treatments and the impact of screening and vaccination programs.

Both studies also demonstrate the usefulness of linkage to health services data. For example, cancer registries provide incidence and mortality data, which, if linked with data from structured pathology reports or clinical sources, such as multidisciplinary meetings, add parameters such as stage and biomarkers that enable better interpretation of survival data, exploration of the prognostic significance of biomarkers and estimation of adherence to treatment

guidelines. Further, the ability to link to Medicare data that include records of billing for privately funded treatments allows analysis of outcomes by chemotherapy and radiotherapy use. Linking registry data to cancer screening and human papillomavirus vaccination data enables analysis of the benefits of these programs. Registries don't collect data on population risk factors, but linkage with population cohort studies allows study of associations between risk factor trends and cancer rates. Linkage with other data sources such as workforce statistics and databases of social determinants of health can provide essential data for health planning.

There are multiple barriers across Australia to researchers accessing publicly funded data collections, let alone being able to link them. One concern is the protection of individuals' privacy; however, there are many techniques proposed for preserving privacy while linking multiple datasets.⁷ The Western Australian data linkage system, established in 1995, has developed very sophisticated systems for linking data, which ethics committees have deemed adequate for protecting privacy.¹ For linking the data in the studies described in this issue of the *Journal*, the Centre for Health Record Linkage used probabilistic privacy-preserving methods that gained ethics committee approval.^{3,4} In general, patients express surprise that their data are not used for the public good, rather than concerns about privacy.⁸

Researchers wishing to gain access to administrative datasets can face barriers from multiple jurisdictions and administrations. For example, data from state registries should be standardised before comparisons between them are made, but this is a statistical exercise that should not prevent access to these data. Further, the possibility of researchers finding inferior outcomes in particular health units of jurisdictions can be seen as a potential reason for denying access to datasets, but instead should be seen as an opportunity to drive improvement. While data custodians need to be assured that only sound projects are granted access, having to present ethics proposals to each state to gain approval to work with national data seems unnecessary in an era of centralised ethics committees for multi-centre research.⁹ Researchers, however, should be aware of the workforce and financial burden that providing data imposes on data custodians, and look to minimise this.

Digital health records are now essential for providing the large datasets necessary for gaining high-quality evidence through data linkage.

The public needs to understand the benefit they derive from linked datasets and add to the call for researchers, health policymakers and data custodians to jointly seek to remove the barriers to accessing and linking health datasets.

Competing interests: No relevant disclosures.

Provenance: Commissioned; externally peer reviewed.

- 1 Holman CDJ, Bass AJ, Rouse IL, Hobbs MST. Population-based linkage of health records in Western Australia: development of a health services research linked database. *Aust N Z J Public Health* 1999; 23: 453-459.
- 2 Acheson ED. Medical record linkage. London: Oxford University Press, 1967.
- 3 Smith RC, Creighton N, Lord RV, et al. Survival, mortality and morbidity outcomes after oesophagogastric cancer surgery in New South Wales, 2001–2008. *Med J Aust* 2014; 200: 408-413.
- 4 Jorgensen ML, Young JM, Dobbins TA, Solomon MJ. Predictors of variation in colorectal cancer care and outcomes in New South Wales: a population-based health data linkage study. *Med J Aust* 2014; 200: 403-407.
- 5 Markar SR, Karthikesalingam A, Thrumurthy S, Low DE. Volume-outcome relationship in surgery for esophageal malignancy: systematic review and meta-analysis 2000-2011. *J Gastrointest Surg* 2012; 16: 1055-1063.
- 6 Sacerdote C, Baldi I, Bertetto O, et al. Hospital factors and patient characteristics in the treatment of colorectal cancer: a population based study. *BMC Public Health* 2012; 12: 775. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3551723> (accessed Mar 2014).
- 7 Kuzu M, Kantarcioglu M, Inan A, et al. Efficient privacy-aware record integration. *Adv Database Technol* 2013; 167-178. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3772958/#!po=2.38095> (accessed Mar 2014).
- 8 Olver IN. Opting-in for opt-out consent. *Med J Aust* 2014; 200: 201-202.
- 9 Hicks SC, James RE, Wong N, et al; Australasian Gastro-Intestinal Trials Group. A case study evaluation of ethics review systems for multicentre clinical trials. *Med J Aust* 2009; 191: 280-282. □