

The Centre for Health Record Linkage: fostering population health research in NSW

Katie A. Irvine^{A,C} and Lee K. Taylor^B

^ACentre for Health Record Linkage, Cancer Institute NSW

^BCentre for Epidemiology and Research,
NSW Department of Health

^CCorresponding author. Email: Katie.Irvine@cancerinstitute.org.au

The Centre for Health Record Linkage (CHeReL) was established in July 2006 with the mission:

To create and sustain a record linkage infrastructure for the health and human services sector, and provide access to these resources to bona fide researchers and health planners and policy makers.

The benefits of researcher access to record linkage infrastructure are well recognised.^{1,2} Linkage transforms data that are collected on a routine basis as part of health care into a powerful resource for research. Linked data can be used to investigate the safety and quality of health care, the effectiveness of prevention and screening programs, and the patterns, costs and outcomes of health care for people with specific conditions such as diabetes, cancer and heart failure. Linkage of health data with data from other agencies – such as education, aged care and community services – can be used for research in the social sciences and to study the broader outcomes of ill health and disability.

Use of the CHeReL's linkage services has increased steadily over time. To date more than 120 linkage projects using data from the health, education, human services, justice and transport sectors have been completed. These include:

- follow-up of cohorts of people with rare conditions or outcomes reported through population health datasets (e.g. infective endocarditis,³ childhood cancer⁴)
- morbidity and mortality associated with infectious diseases (e.g. Hepatitis C,⁵ influenza)
- follow-up of researcher-supplied cohorts to obtain information on service utilisation or health-related outcomes (e.g. *45 and Up Study*,⁶ the Australian Longitudinal Study on Women's Health and the Australian HIV Observational Database⁷)
- case control studies of cancer screening behaviour and outcomes
- validation of the accuracy of screening tests by linkage with outcome information (e.g. antenatal serum screening and pregnancy outcomes⁸)

- incidence of diseases and conditions by the identification of first-time events (e.g. first admissions for stroke and heart failure)
- reporting of outcomes of health care adjusted for co-morbidity using historically linked data (e.g. outcomes for stroke care in New South Wales (NSW) hospitals)
- studies of health care safety, utilisation and costs (e.g. adverse events in hospital, cancer patterns of care⁹)
- validation studies of the accuracy of information held in population health datasets
- improved ascertainment of health information (e.g. Aboriginality, diagnoses) using multiple data sources.

Using probabilistic linkage software,¹⁰ the CHeReL has created a Master Linkage Key of records from population health datasets commonly used by researchers (Table 1). The Master Linkage Key currently includes over 36 million records relating to about 8 million people. Large amounts of historical data can be accessed for research while avoiding the prohibitive cost and time of creating one-off links for individual projects or establishing *de novo* longitudinal studies. To the extent that these datasets provide coverage of complete populations the outputs of record linkage studies avoid some of the potential biases associated with unrepresentative or incomplete samples compared with traditional study designs.

The CHeReL also fosters high quality research by:

- participating in the Population Health Research Network¹¹ which has been established under the National Collaborative Research Infrastructure Strategy¹² to provide Australian researchers with access to linkable non-identified data from a diverse range of health datasets, across jurisdictions and sectors
- providing a formal introduction to the CHeReL's linkage methods in a short course on the analysis of linked data through the University of Sydney
- providing support for the NSW Health Data Linkage Special Interest Group, which meets 3–4 times per year
- providing advice to researchers on the design, feasibility, cost and process of linkage studies.

Robust data governance has been critical to the CHeReL's success. Data custodian approval and approval of a human research ethics committee is required for all research projects. The CHeReL also complies with best practice in privacy preserving record linkage, which involves the separation of the linkage of personally identifying information from the analysis of

Table 1. The Centre for Health Record Linkage Master Linkage Key, 30 November 2010

Data collection	Years	No. records
NSW Admitted Patient Data Collection	July 2000–June 2009	19 874 083
NSW Registry of Births, Deaths & Marriages Birth Registrations	1994–2008	1 333 539
NSW Registry of Births, Deaths & Marriages Death Registrations	Jan 1994–June 2010	802 739
Australian Bureau of Statistics Mortality Data (NSW)	1985–2007	1 020 798
Australian Bureau of Statistics Perinatal Mortality Data (NSW)	1994–2005	9445
NSW Midwives Data Collection (mothers)	1994–2008	1 331 115
NSW Midwives Data Collection (babies)	1994–2008	1 331 115
NSW Central Cancer Registry	1994–2008	504 894
Australian Capital Territory Cancer Registry	1994–2006	14 821
The 45 and Up Study	2010 update	267 235
NSW Emergency Department Data Collection	2005–2009	9 526 946
NSW Notifiable Conditions Information Management System	1993–2008	421 870
NSW Perinatal Death Review Database	2000–2006	4657
Total		36 443 257

de-identified linked health records.¹³ This approach to data governance has been strongly supported by organisations that are custodians of health records, human research ethics committees, and the community. By providing a mechanism for researchers to access non-identified linked data, the CHeReL enables ethically approved research in the public interest to be carried out without consent, minimising bias and allowing researchers to access data on whole populations.

Conclusion

The CHeReL has become core infrastructure for health and health services research in NSW. Further information on the CHeReL is available from: www.cherel.org.au.

Acknowledgments

We gratefully acknowledge the funding and in-kind contributions of the NSW Department of Health and the Cancer Institute NSW as owners of the CHeReL; the Cancer Institute NSW as host organisation for the CHeReL; and the funding support of past and present CHeReL Member Organisations: ACT Health, Clinical Excellence Commission, the Sax Institute, the University of Newcastle, the University of New South Wales, the University of Sydney and the University of Western Sydney.

References

1. Roos LL, Menec V, Currie RJ. Policy analysis in an information-rich environment. *Soc Sci Med* 2004; 58(11): 2231–41. doi:10.1016/j.socscimed.2003.08.008
2. Holman CDJ, Bass AJ, Rosman DL, Smith MB, Semmens JB, Glasson EJ et al. A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system. *Aust Health Rev* 2008; 32(4): 766–77. doi:10.1071/AH080766
3. Sy RW, Kritharides L. Health care exposure and age in infective endocarditis: results of a contemporary population-based profile of 1,536 patients in Australia. *Eur Heart J* 2010; 31(15): 1890–7. doi:10.1093/eurheartj/ehq110
4. Stavrou EP, Baker DF, Bishop JF. Maternal smoking during pregnancy and childhood cancer in New South Wales: a record linkage investigation. *Cancer Causes Control* 2009; 20(9): 1551–8. doi:10.1007/s10552-009-9400-5
5. Gidding H, Amin J, Dore GJ, Ward K, Law M. Hospital related morbidity in people first notified with Hepatitis C: a population-based record linkage study in New South Wales, Australia. *J Hepatol* 2010; 53(1): 43–9. doi:10.1016/j.jhep.2010.01.036
6. Jorm LR, Walter SR, Lujic S, Byles JE, Kendig HL. Home and community care services. A major opportunity for preventive health care. *BMC Geriatr* 2010; 10: 26. doi:10.1186/1471-2318-10-26
7. Falster K, Wand H, Donovan B, Anderson J, Nolan D, Watson K et al. Hospitalisations in a cohort of HIV patients in Australia, 1999–2007. *AIDS* 2010; 24(9): 1329–39. doi:10.1097/QAD.0b013e328339e245
8. Lain SJ, Algert CS, Tasevski V, Morris JM, Roberts CL. Record linkage to obtain birth outcomes for the evaluation of screening biomarkers in pregnancy: a feasibility study. *BMC Med Res Methodol* 2009; 9: 48. doi:10.1186/1471-2288-9-48
9. Stavrou EP, Smith GS, Baker DF. Surgical Outcomes Associated with Oesophagectomy in New South Wales: An Investigation of Hospital Volume. *J Gastrointest Surg* 2010; 14(6): 951–7. doi:10.1007/s11605-010-1198-7
10. Open Source ChoiceMaker Technology. Available from: <http://oscm.sourceforge.net> (Cited 14 December 2010.)
11. Population Health Research Network. Available from: www.phrn.org.au (Cited 14 December 2010.)
12. Australian Research Council and Sargent M. National Collaborative Research Infrastructure Strategy. Canberra: Australian Research Council; 2006. Available from: <http://ncris.innovation.gov.au/Pages/SRARI.aspx#Roadmap> (Cited 1 August 2010.)
13. Kelman CW, Bass AJ, Holman CDJ. Research use of health data—a best practice protocol. *Aust N Z J Public Health* 2002; 26(3): 251–5. doi:10.1111/j.1467-842X.2002.tb00682.x