

The Lowitja Institute Data Linkage Roundtable: Background Document

Ray Lovett

Introduction

Data linkage (sometimes called record linkage) refers to finding records in different data sets that refer to the same person. This can be across two or more different data sources, e.g. hospital data and another administrative data set. Linking of multiple data records for one person is one way of enhancing understanding of health outcomes, processes involved in health care delivery or to gain greater understanding at the population level of relationships between variables in data sets and health outcomes. Data linkage is becoming a common method in health research because of the opportunity to work across disciplines and because it is cost effective when compared to primary data collection (Doiron, Raina, & Fortier, 2013).

Who conducts record linkage in Australia?

There are Data Linkage Units in all Australian jurisdictions which link health and other related data at the person level using secure, privacy-preserving methods (<https://www.sure.org.au/resources/australian-health-data-linkage-units>). These include:

- Centre for Health Record Linkage (CheReL)
- Data Linkage Western Australia
- Centre for Data Linkage (Population Health Research Network)
- SA-NT DataLink
- Menzies Research Institute Tasmania Data Linkage Unit
- Health LinQ
- Victorian Data Linkages (VDL)
- AIHW Data Linkage Unit
- Population Health Research Network Telethon Institute for Child Health Research.

Data linkage genesis concerning Aboriginal and Torres Strait Islander people in Australia

Much of the past data linkage work involving Aboriginal and Torres Strait Islander people in Australia is concerned with under-reporting of Indigenous status in administrative data. Data linkage relating to Aboriginal and Torres Strait Islander people has also been used to:

- compare Indigenous status across data sets
- compare Indigenous and non-Indigenous status on outcome variables of interest
- exclusively select Aboriginal and Torres Strait Islander people from data sets to study particular variables and link this with other datasets.

A recent report from the Australian Institute of Health and Welfare (AIHW) reviews data linkage projects undertaken on Aboriginal and Torres Strait Islander people in Australia. (<http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129543445>).

Closing data gaps through data linkage strategies, including the linking of registry and administrative databases across jurisdictions by the Australian Institute of Health and Welfare (AIHW), 2013(Australian Institute of Health and Welfare, 2013).

Data linkage to improve under-recording of Indigenous status

Data linkage has been used to improve identification of Aboriginal and Torres Strait Islander people in relation to different datasets where this is often missing. In these projects registry data or national databases are linked to, for example, State and Territory hospital record data.

Improved identification of Indigenous deaths increased 10% by linking death records from the National Mortality Database, the National Death Index, the Residential Aged Care Dataset, National Hospital Morbidity data and the National Perinatal Data Collection. This AIHW project has improved state and territory registrars' population records for births, deaths and marriages for Aboriginals and Torres Strait Islanders; it has also contributed to estimate life expectancy.

Using data from Perinatal Data Collection (PDC) and Hospital Administration Patient Admission data to identify babies from non-Indigenous mothers and Indigenous fathers (Indigenous fathers are not included in the PDC) allowed to establish that 29.1% of Indigenous babies have an Indigenous father and non-Indigenous mother.

Data linkage to improve data

Data linkage has also been used to improve data of interest other than Indigenous under-reporting. In this context, variables of interest are collected and linked through different databases and Aboriginality is linked from the database that has the best data on this. For example, in an investigation into reasons for hospitalisation of children before two years of age-linked registered births data and admitted patients' data from the West Australian Data Linkage System (WADLS) with the Western Australian Midwives Notification System data, Indigenous status was derived from the latter, because it contained the best data for this.

Another example is a study in which child maltreatment and factors increasing the risk of child maltreatment were examined using a retrospective cohort study. In this study, child protection data, midwives notifications system, birth registrations, hospital morbidity database and mental health database were linked.

A further example included the Cerebral Palsy Register, Births Defects Register and Intellectual Disability Exploring Answers Dataset linked to identify disabled children. Aboriginal and Torres Strait Islander status was initially taken from the births registrations and midwives notifications, and missing records were filled in by data from hospital morbidity data or child protection data.

Data linkage where Indigenous status is a secondary variable

Data linkage is also used in studies where Aboriginality is defined as one of multiple risk factors for certain diseases and disabilities. In these circumstances data linkage is used to improve the identification of Aboriginal and Torres Strait Islander people or the clinical data to analyse the risk for Aboriginal and Torres Strait Islander people on certain health related outcomes. A study, extracting hospital morbidity data and deaths data from the WADSL, examined recurrent gastroenteritis among WA infants. This study concluded that Aboriginality was the only significant risk factor for re-admission.

In one recent study assessing both under-recording of Indigenous status and algorithms on estimated disparities in cardiovascular and injury outcomes, it was found that 60 per cent of the variation in recording of Aboriginal status was related to the hospital of admission. The poorest recording of

Aboriginal and Torres Strait islander status was found among private and major city hospitals (Randall, Lujic, Leyland, & Jorm, 2013).

Data linkage in studies of Aboriginal and Torres Strait Islander people exclusively

Other studies in which data linkage is useful include those that use a dataset with an Indigenous focus. This Indigenous focused dataset can be linked to other datasets to add data missing from the other datasets. This is especially useful, for example, in studies using Indigenous focused surveys. Surveys act as the register for Aboriginal and Torres Strait Islander people and can serve as a gold standard.

In a study examining the influence of different lifestyle factors on coronary heart disease, survey data were linked to WA hospital, cancer and deaths data, which served to identify coronary heart disease risk factors. This study found that certain lifestyle factors were a predictor of coronary heart disease.

Establishing better communication/collaboration across the different health and information systems to improve data linkage capacity and work

The National Advisory Group on Aboriginal and Torres Strait Islander Health Information and Data (NAGATSIHID) was established, as a result of a decision by the Australian Health Ministers' Advisory Council (AHMAC) in October 2000, to improve reporting on the status of health of Indigenous Australians.

The group's main focus is to:

- strategically align the work of this roundtable to particular areas of the NAGATSIHID Strategic plan. Areas such as providing advice on projects and seeking sponsorship of projects outlined from the roundtable would be a focus
- ensure greater involvement of the community controlled health sector which might benefit from data linkage approaches to identifying local and regional needs in terms of planning health service delivery or in identifying service delivery gaps.

Ethical considerations in data linkage including data ownership, custodianship and use

Data linkage and other health research can lead to harmful conclusions about Aboriginal and Torres Strait Islander people. An example of this is a study mentioned earlier in this paper that concluded Aboriginality was identified as the only significant factor associated with hospital admission among Western Australian infants with gastroenteritis. This research implies that Aboriginality rather than environmental and other demographic factors were related to the health outcome of gastroenteritis. Occasionally, Indigeneity is used as an explanatory variable in health outcomes. This is incorrect. It is however, often the case that a substantial number of Aboriginal and Torres Strait Islander people are adversely affected by a range of complex societal, political and environmental factors that contribute to poor health outcomes and it is these factors that require attention to improve health outcomes.

The outcomes of data linkage studies can directly or indirectly influence the wellbeing of Aboriginal and Torres Strait Islander people. Policies and programs may be developed based on data linkage studies outcomes (Australian Institute of Health and Welfare and Australian Bureau of Statistics, 2012). Studies that use data linkage must therefore adhere to the Values and Ethics guidelines for health research involving Aboriginal and Torres Strait Islander people (National Health and Medical Research Centre, 2003). This means that, additional to ethics and data custodian approval, community approval must be sought as well. There has been limited examination of this in Australia; however, there are international

approaches that can be called upon for guidance in this area
(<http://cahr.uvic.ca/nearbc/documents/2009/FNC-OCAP.pdf>) (First Nations Centre, 2007).

References

- Australian Institute of Health and Welfare. (2013). Report on the use of linked data relating to Aboriginal and Torres Strait Islander people. Cat. no. IHW 92: Canberra: AIHW.
- Australian Institute of Health and Welfare and Australian Bureau of Statistics. (2012). National best practice guidelines for data linkage activities relating to Aboriginal and Torres Strait Islander people (Vol. 2013). Canberra: AIHW.
- Doiron, D., Raina, P., & Fortier, I. (2013). Linking Canadian population health data: maximizing the potential of cohort and administrative data. *Can J Public Health, 104*(3), e258-261.
- First Nations Centre. (2007). OCAP: Ownership, Control, Access and Possession. Sanctioned by the First Nations Information Governance Committee, Assembly of First Nations. Ottawa, ON:: National Aboriginal Health Organization.
- National Health and Medical Research Centre. (2003). Values and Ethics - Guidelines for Ethical Conduct in Aboriginal and Torres Strait Islander Health Research. Canberra: Commonwealth of Australia.
- Randall, D. A., Lujic, S., Leyland, A. H., & Jorm, L. R. (2013). Statistical methods to enhance reporting of Aboriginal Australians in routine hospital records using data linkage affect estimates of health disparities. *Aust N Z J Public Health, 37*(5), 442-449. doi: 10.1111/1753-6405.12114.