

Chapter 4: Ongoing and Emerging Issues in Homelessness Data Collection

Harnessing the Potential of Linked Administrative Data for Homelessness Research

Associate Professor Lisa Wood, Shannen Vallesi, Professor Paul Flatau, Centre for Social Impact, UWA Business School, University of Western Australia

Linked government and administrative data sets are increasingly recognised as a powerful resource for program evaluation and policy research in a range of sectors, including health care, housing, and social services,^{1,2} and have been used world-wide for conducting health and social science research.³

In Australia, the use of linked administrative data for policy relevant research and evaluation is well entrenched in public health. However, in large part reflecting low levels of investment in research and data infrastructure, the application in the housing and homelessness fields is still in its relative infancy. Relatively untapped also is the rich research potential of linking comprehensive homeless data collections (such as the Specialist Homelessness Service (SHS)), the Australian Institute of Health and Welfare (AIHW) collection and the Registry Week data (held by homelessness agencies) with government and administrative data collections from other sectors such as housing, health, justice and social services where people who are homeless are frequently over-represented. As articulated by Petrilá⁴:

Policy initiatives in one area — for instance, housing — typically can affect individual and community outcomes in other areas such as health or education. As a result, analysing data from only one system frequently results in a one dimensional perspective that misses myriad outcomes in other systems, and thus makes it more difficult to accurately diagnose a problem and develop a solution.

With an urgent fiscal imperative to build evidence for effective interventions that can reduce homelessness, administrative data sets provide opportunity to tap into high quality detailed information that is collected systematically and longitudinally. Administrative data sets from government agencies and departments provide rich sources of complementary data that when linked together capture the 'big picture' of an individual's experience and can be used to calculate the cost effectiveness of a program or initiative to determine the area of greatest savings.

Data linkage uses a process where data that has already been collected for other purposes is merged at the individual level using a unique identifier to create new data from existing sources.^{5,6} Illustratively, in our recent study;⁷ administrative data on public housing, National Partnership Agreement on Homelessness (NPAH) program participation and hospital and health service use was linked to examine the health and economic impacts of supporting people who are homeless to access public housing tenancies.

Among the 3,383 previously homeless people in the study, there was a marked reduction in emergency department presentations, hospital admissions, length of stay, psychiatric unit bed-days and intensive care units, within just one year of their entry into public housing. This equated to a combined cost saving of \$16.4 million for the WA health system in that single year, with the cost per person saving greatest among a cohort of 983 clients supported through NPAH programs. The study found a large cost saving of \$84,135 per person per year for

those receiving support from the NPAH Mental Health program, which was primarily related to a reduction in psychiatric inpatient admissions.

This study provides an example of the economic impact that can be measured by linking two datasets together by capturing the changes in health service use from the provision of public housing and support through the NPAH programs. Importantly, our study provided significant supporting evidence for the effectiveness of NPAH programs at a time when some were questioning the robustness of previous NPAH program evaluations.

Administrative data linkage is advantageous for homelessness research at differing levels, as illustrated with the examples of recent research below:

- *whole of government (where multi-agency data is linked):* as illustrated by the Australian Institute of Health and Welfare (AIHW) study⁸ which linked data from SHS agencies and public housing authorities in Western Australia (WA) and New South Wales (NSW);
- *whole of sector to evaluate the impact of a targeted government policy initiative:* our Centre for Social Impact, University of Western Australia NPAH study highlights the potential for multi-faceted interventions to simultaneously improve mental health and homelessness outcomes, and yield cost savings in both domains. This is important in the current policy climate as it strengthens evidence for the continuation of NPAH and similar programs that can yield fiscal and social outcomes across multiple sectors and government funding domains;



- *individual project level*: for example in our current evaluation of the 50 Lives 50 Homes project in WA, the evaluation dataset includes administrative data from hospitals and health services; the Ruah After Hours Support Service (providing support and access to primary health care for participants) and VI-SPDAT registry week data;
- *Agency level*: for example the use of linked administrative data to improve knowledge of client history and outcomes and to track progress of these (such as patterns of change in health service use or justice pre/post engagement with the agency).

A further benefit of administrative data is that there are standardised methods of data collection, recording and reporting, hence outcomes of different studies and interventions can be more readily compared. In homelessness research to date, comparison across studies is often difficult as there is considerable variation in the type, source and quality of data or evaluation tools and measures used. Administrative data can also be usefully triangulated with other data sources for richer understanding of homelessness trajectories and outcomes. In our NPAH study⁹ administrative health and housing data was also linked to data from a survey completed by a subsample of public housing tenants in WA, which provided additional insights into homelessness and housing experiences that valuably complemented the empirical findings from the linked hospital and housing data.

Whilst the appeal of linking inter-agency data for homelessness research is compelling, it's not without its limitations. Some of these relate to the data itself, and some to the data linkage and access process.

- *Data collected for purposes other than homelessness*: Administrative data sets created for purposes unrelated for homelessness for example (such as hospital or corrections data) often do not have a variable that suffices as a robust measure of homelessness or housing status, and the use of 'no fixed address' variables is not an ideal proxy, and misses the nuances of different types of homelessness.¹⁰ The way in which

demographic or other variables are collected can also vary from collection to collection and make matching more difficult, and people who are homeless are more likely to have missing data (for example service history records may be erratic, or missing if they have moved around the country);

- *Data access and approval processes:* Timeliness of access to valuable linked data is a challenge for research that seeks to have traction in current policy discourse and intervention development. Administrative data from government agencies is, by its very nature, often confidential and of a sensitive nature, and state-wide data linkage systems have a strong imperative to ensure that concerns about confidentiality, privacy and use for intended purpose only can be rigorously addressed.¹¹ As a consequence, however, the application and approval processes for use of linked data can take months, and the linking process itself is labour intensive particularly when multiple data sets are sought from different jurisdictions or sectors. This can unfortunately hinder the realisation of the vast opportunities that linked data can yield for public policy and the forging of greater links between research, policy and practice. However, as data linkage gains further momentum around Australia, the need for timely access to data is being addressed, and this should not deter researchers from considering its use. There is also a cost involved typically in obtaining linked data, but this is often far less than the cost of undertaking direct data collection on the ground, and any cost is usually well outweighed by the benefits of comprehensive longitudinal data;

- *Scope of the administrative data:* There is a misconception that administrative data sets, by their very nature, capture the records of the whole population. In spite of very broad coverage, administrative data sets typically have quite specific geographic and organisational restrictions. For example, the Australian homelessness administrative data only refer to clients of 'specialist homelessness services' (that is, services receiving specific forms of government housing/homelessness

funding) not all services. Health and social service administrative data sets are invariably state/territory based at present and have not been linked across jurisdictions;

- *Scale of data and multi-agency data:* Whilst administrative data can be linked for small client samples, it is the accessing of large scale data collections from multiple service providers and/or that capture the majority of a population of interest that have the greatest potential. For example, jurisdiction level hospital record or justice data or nationally collated Centrelink data if linked to SHS collection data would provide a rich mine of data for homelessness research. However, this is more easily said than done with multiple data custodians approvals and separate ethics applications among considerable privacy concerns around potentially re-identifiable data.

A low hanging fruit for data linkage research in homelessness lies in linking SHS data to administrative data sets. The recent AIHW was the first national foray into this,¹² and used linked data from SHS agencies and public housing authorities in Western Australia and New South Wales. Two key results highlight the importance of linking housing and homelessness data. Firstly, of those adults assisted by specialist homelessness agencies into public housing who then exited public housing, just under one-half returned to a specialist homelessness agency for support.¹³ Secondly, of those who only accessed support from SHS agencies after losing their public housing tenancy, almost half were identified as homeless.¹⁴ A critical piece of future research is to extend the current work so that it combines the homelessness, health and housing data.

Conclusion

With other sectors such as health, housing and justice bearing much of the cost and consequence of recurring homelessness, we need to look beyond homelessness data collections alone if our research is to drive substantial policy and funding change. There is currently sub-optimal use of big data for policy-making and collaborative social impact¹⁵ and if more effectively harnessed, the homelessness field has much to gain. Such gains can be amplified through a collaborative

research agenda around the use of data linkage in homelessness research and policy evaluation, as there are immense synergies in leveraging shared learnings and data sources. Complex social issues such as homelessness therefore require cross-sectoral approaches that cut across government and non-government silos. Research that harnesses linked administrative data can assist in guiding and evaluating the impact of more integrated solutions to ending homelessness.

Endnotes

1. Kamateri E, Panopoulou E, Tambouris E, Tarabanis K, Ojo A, Lee D, Price D 2014, *A comparative analysis of tools and technologies for policy making*, in Janssen M, Wimmer MA, Deljo A, (eds) *Policy practice and digital science — integrating complex systems, social simulation and public administration in policy research*, Springer Science, Berlin, pp. 125–56.
2. Oliver I 2014, Linking data to improve health outcomes, 2014, *Medical Journal of Australia*, vol. 200, no. 7, pp.368–9.
3. Jutte DP, Roos LL, Brownell MD 2011, Administrative record linkage as a tool for public health research, *Annual Review of Public Health*, vol.32,, No.1, pp.91–108.
4. Petrla J 2014, *Data Integration For Social Policy: Challenges And Opportunities*, Available from: <http://www.whatcountsforamerica.org/portfolio/data-integration-for-social-policy-challenges-and-opportunities/>.
5. Winglee M, Valliant R, Scheuren F 2005, A case study in record linkage, *Survey Methodology*, vol.31, no.1, pp.3–11.
6. Jutte DP, Roos LL, Brownell MD 2011 op cit.
7. Wood L, Flatau P, Zaretsky K, Foster S, Vallesi S, Miscenko D 2016, *What are the health, social and economic benefits of providing public housing and support to formerly homeless people*, Australian Housing and Urban Institute, Melbourne.
8. Australian Institute of Health and Welfare 2015, *Exploring transitions between homelessness and public housing*, 2015, Australian Institute of Health and Welfare, Canberra.
9. Wood L, Flatau P, Zaretsky K, Foster S, Vallesi S, Miscenko D 2016 op cit.
10. Chamberlain C, MacKenzie D 1992, Understanding contemporary homelessness: Issues of definition and meaning, *The Australian Journal of Social Issues*, vol.27, no.4, p.274.
11. Oliver I 2014, op cit, pp.368–9.
12. Australian Institute of Health and Welfare 2015 op cit.
13. *ibid.*
14. *ibid.*
15. Misuraca G, Mureddu F, Osimo D 2014, *Policy-Making 2.0: Unleashing the Power of Big Data for Public Governance*. http://link.springer.com/chapter/10.1007/978-1-4614-9563-5_11