A comparison of content from across contemporary Australian population health surveys

Branislava Godic,[2]*[iD] Selin Akaraci,[3] Rajith Vidanaarachchi,[1,4][iD] Kerry Nice,[1][iD] Sachith Seneviratne,[1,4][iD] Suzanne Mavoa,[5,6] Ruth Hunter,[3] Leandro Garcia,[3][iD] Mark Stevenson,[1,4] Jasper Wijnands,[7][iD] Jason Thompson[2]

[1]Faculty of Architecture, Building, and Planning, University of Melbourne, Melbourne, Australia
[2]Phoenix Australia Centre for Posttraumatic Mental Health, Department of Psychiatry, University of Melbourne, Melbourne, Australia
[3]Centre for Public Health, Queen's University Belfast, Belfast, Northern Ireland, UK
[4]Faculty of Engineering and Information Technology, University of Melbourne, Melbourne, Australia
[5]School of Population and Global Health, University of Melbourne, Melbourne, Australia
[6]Murdoch Children's Research Institute (MCRI), Parkville, Victoria, Australia
[7]Royal Netherlands Meteorological Institute (KNMI), de Bilt, the Netherlands

Abstract

Objective: Associations between place and population health are of interest to researchers and policymakers. The objective of this paper is to explore, summarise and compare content across contemporary Australian geo-referenced population health survey data sets.

Methods: A search for recent (2015 or later) population health surveys from within Australia containing geographic information from participants was conducted. Survey response frames were analysed and categorised based on demographic, risk factor and disease-related characteristics. Analysis using interactive Sankey diagrams shows the extent of content overlap and differences between population health surveys in Australia.

Results: Thirteen Australian geo-referenced population health survey data sets were identified. Information captured across surveys was inconsistent as was the spatial granularity of respondent information. Health and demographic features most frequently captured were symptoms, signs and clinical findings from the International Statistical Classification of Diseases and Related Health Problems version 11, employment, housing, income, self-rated health and risk factors, including alcohol consumption, diet, medical treatments, physical activity and weight-related questions. Sankey diagrams were deployed online for use by public health researchers.

Conclusions: Identifying the relationship between place and health in Australia is made more difficult by inconsistencies in information collected across surveys deployed in different regions in Australia.

Implications for Public Health: Public health research investigating place and health involves a vast and inconsistent patchwork of information within and across states, which may impact broad-scale research questions. The tools developed here assist public health researchers to identify surveys suitable for their research queries related to place and health.

Keywords: public health, health data sets, health policy, public health research

Introduction

The availability of population health survey data sets is important for the conduct of high-quality public health research,[1] policy-making, planning,[2,3] and project and program evaluation.[3–6] The collection of public health information through population health surveys makes it possible for many important public health research questions to be posed and answered by researchers, health services and government, and even interested citizen scientists. Improving public health survey accessibility, transparency and openness is also useful for increasing the efficiency of researcher time and budgets, which are usually funded by the public purse. These considerations are key, given that both the custodians of public health data sets and their users have a moral and ethical responsibility to facilitate their utilisation for the greatest public good.

Whether in the case of a region-specific or country-wide public health research question, there will often be multiple data sets applicable to that region. Sifting through each data set to determine which one(s) might hold variables of interest can be time-consuming. Furthermore, individual surveys are often unlikely to contain all required information, meaning that combinations of data sets may be necessary. The content of various surveys must then be compared because while content across surveys may overlap (e.g. measures of consumption, activity, disease incidence, or exposure risk), differences in sample populations, response frames, and/or survey questions may make some more suitable for use than others. In addition, access to geo-referenced information, which refers to geospatial information, such as postcodes or other area-related measures, is necessary for understanding questions related to place and health, which is increasingly becoming recognised as a prominent research area.[7–9] Many Australian health surveys contain geo-referenced information at varying levels of granularity, such as the 45 and Up Study[10] and the Household, Income, and Labour Dynamics (HILDA) study.[11,12]

To assist in overcoming some of the difficulties associated with understanding and comparing content across Australian public health survey data sets, this research aimed to collect and analyse the content within current Australian population health surveys. In doing so, we hoped to make the process of population health data collection more efficient and increase the use of publicly available population health data sets for public good.

Methods

Using online search engines, a search was undertaken for publicly available Australian population health survey data sets published from around 2015–2020 using existing published information.[13,14] The current paper focused on public health surveys and data sets and specifically on the analysis of their questionnaires with a focus on noncommunicable diseases. Public health data sources, including hospital data and surveys, that do not include these areas were excluded as a result. It is important to note that while every effort was made to identify appropriate data sets at the time of searching, the search methodology undertaken was not systematic. This is due to the original search occurring as a byproduct of another study that inspired the research question. Data sets were included if they contained demographic, health, disease, and risk-exposure information (e.g. risk factors and clinical diagnoses, etc.) for people over 18 years. Once individual data sets were identified, data dictionaries and/or survey questionnaires were sought from the survey custodian. If a data dictionary or survey was not available, it was excluded. Please note that one survey was publicly available at the time of analysis, but the analysed version is no longer available. A total of 13 population health survey data sets were included (Table 1).

Datasets analysis process

Analysis occurred via a two-step process. In the first step, items from each survey were categorised into sets of common themes to capture the breadth of content across surveys (see Supplementary Material for a complete list of themes). Broadly, self-reported survey items were categorised across demographic information (e.g. age, gender, background, marital status, occupation, income, housing, location, etc.), presence of illness as classified by the World Health Organisation International Statistical Classification of Diseases and Related Health Problems version 11 (ICD-11) codes,[59] Non-communicable disease (NCD) risk factors (e.g. smoking, level of physical activity, alcohol consumption, etc.), and remaining characteristics. Remaining characteristics covered a broad range of other factors assessed by the surveys that did not fall under the previous categories, including pregnancy, transport, health screening behaviours, health promotion, and many more. Note that all information in the surveys was self-reported. Some studies may contain links to other sources of information, such as hospital data; however, these data were not considered.

The second step quantified the extent to which the different themes were covered across surveys by counting the number of data points allocated to each subject category in each survey. Individual survey questions were assigned to categories to allow comparison between data sets, and only questions answered by and for adults were considered. The number of characteristics associated with individual questions was first determined. Answers to binary response frame questions, and single-select multiple-choice questions were recorded as a single data point. For multiple-choice questions, each response contributed to one data point. Where questions covered multiple domains (e.g. questions about whether someone has a family history of a condition would be considered a question about the risk factor of prior family history, along with asking about the condition), points were allocated to all relevant categories. Only direct questions (e.g. not derived characteristics) for adults were assessed.

The categorisation process was independently validated by two researchers (B.G. and S.A.), with any discrepancies between categorisations discussed until a mutual agreement was reached. Sankey diagrams were then generated, illustrating the number of items allocated to each category across surveys.[60] Interactive Sankey diagrams were also generated and imported into a new website created for the findings of the study (https://aushealthdatasets. w3spaces.com/).

Results

All surveys included characteristics on employment, housing, and income. All surveys asked questions related to self-rated health. The most frequently covered ICD-11[59] dimension was symptoms, signs and clinical findings, covered by all surveys. Symptoms, signs and clinical findings is a broad category in the ICD-11, which refers to indicators of illness that might not reach a threshold for a diagnosis. Twelve surveys asked about neoplasms and mental, behavioural or neurodevelopmental disorders. Eleven surveys covered diseases of the nervous system, endocrine, nutritional, or metabolic diseases, diseases of the respiratory system, and injury, poisoning, or other consequences of external causes. Ten surveys covered questions about nonspecific "other important illnesses," diseases of the musculoskeletal system or connective tissue, and diseases of the circulatory system. It should be noted that questions related to "other important illnesses" could apply to any of the categories in the ICD-11; however, not all surveys asked for the condition to be specifically named.

The "symptoms, signs or clinical findings not elsewhere specified" category covers a broad range of symptoms from across all of the body systems that have not been diagnosed as a disease due to transience or indeterminate causes.[59] This category was covered in greatest detail by the Raine Study followed by the Busselton Healthy

Table 1: Australian population health surveys included in the analyses, indicating the title of the data custodian, the survey collection period for the version analysed, the sample size within the survey analysed, whether the study is longitudinal overall, multigenerational, and/or provides biosample or clinical data.

Title	Data custodian	Study aim(s)	Version analysed and collection period	Survey timespan	Intervals of data collection	Location/level of data collection and level of geospatial information released	Sample cohort and sample size of the version used	Longitudinal with same cohort?	Multigenerational	Biosamples collected?	Sampling method
National studies											
45 and Up [10]	The Sax Institute[10]	Six priority areas: (1) socioeconomic factors about healthy ageing, (2) weight and activity, (3) environmental effects, (4) disease risk factors, (5) health service usage and (6) health status of people aged >80 years [10]	Wave 2: first follow-up, collected between 2012 and 2015[15,16]	2006 onwards [10]	Every 5 years [10]	New South Wales (NSW) collection, geolocated data available (SA2, more granular data released if approved) [10,17]	142,548 men and women aged ≥45 [16]	Yes [17]	No[10]	Yes[17]	Random sampling from the Medicare Australia enrolment database. Individuals aged >80 years and rural residents are oversampled[10]
Ten to Men [18]	Australian Government: Australian Institute of Family Studies (previously the University of Melbourne)[18]	Six priority areas: (1) mental health and well-being, (2) health service usage, (3) behaviours, (4) health status, (5) health literacy, (6) socioenvironmental factors [18]	Wave 2 (Adult data only), collected between 2015 and 2016 [18,19]	2013 onwards [18]	Regular intervals not reported, currently between 2 and 5 years [18]	Australia-wide,[20] geolocated data available (SA1, SA2, SEIFA) [18]	10,339 men, aged from 10 years old [18]	Yes[18]	Yes [18]	No[18]	Stratified, multistage, cluster random sampling [18]
Australian Longitudinal Study on Women's Health [21,22]	Australian Government Department of Health, the University of Newcastle and the University of Queensland [21,23]	Surveys aim to provide a holistic investigation into factors related to women's health, encompassing sociodemographic factors, health-related behaviours, health issues and symptoms, general well-being and health status, and health service utilisation [23]	1946-51 survey wave 9, collected between 2018 and 2020 [23–25]	1996 onwards [23,25]	Varies depending on cohort, but between 6 months and every 4 years [26]	Australia-wide, derived variables, including ARIA, RRMA, SEIFA 1996 Indexes, Modified Monash Model [23,27]	7,844 women [23]	Yes, with the same group of participants in multiple cohorts[23]	No[28]	No[23]	Stratified random sampling, with oversampling in rural areas [22]
Household, Income and Labour Dynamics in Australia (2019) [11,12]	Australian Government Department of Social Services, the Melbourne Institute of Applied Economic and Social Research [11]	Provide broad-ranging information related to sociodemographic and health-related information, along with wider topics, such as attitudes, experiences, among broader topics covered less frequently (family dynamics, cognition and more) [11]	2019 survey, collected in 2019–2020 [12,29,30]	2001 onwards [11]	Annually [11]	Australia-wide, derived geospatial variables, including statistical local area, local government area, statistical subdivision, statistical division, section of state, major statistical region, SA2, SA3, SA4, Greater Capital City Statistical Area and Remoteness Area[11]	7,142 men and women [12]	Yes, with the same group of participants over time[11]	Yes[11]	Yes[12]	National probability sample of Australian households occupying private dwellings [12]
National Health Survey[31]	Australian Bureau of Statistics[31]	Aims to collect broad-ranging health information about Australians, including long-term diagnoses, risk factors, and sociodemographic information [31]	2017–2018 survey, collected across 2017–2018 [31]	1989–onward [32]	Collected every 3–6 years [32]	Australia-wide, limited results provided per SA2[31]	21,315 males and females [32]	No, the study is conducted on a new sample in each wave [32]	Yes[32]	Yes[32]	Random multistage area sample [32]

TABLE 1.   Continued

4

Full Length Article

Title	Data custodian	Study aim(s)	Version analysed and collection period	Survey timespan	Intervals of data collection	Location/level of data collection and level of geospatial information released	Sample cohort and sample size of the version used	Longitudinal with same cohort?	Multigenerational	Biosamples collected?	Sampling method
Jurisdiction-specific studies											
Busselton Healthy Ageing Study (Busselton Baby Boomer Study)[33]	Busselton Population Medical Research Institute[33]	The identification of the diseases and their related and underlying factors that constitute the disease burden impacting upon healthy ageing[33]	Version 4[a], collected between 2010 and 2015 (34)	2010 onwards[34]	Every 5 years[34]	Busselton, Western Australia, geospatial information not mentioned[33,34]	5,107 men and women born between 1946 and 1964 (34)	Yes[33]	No[33]	Yes[33]	Sampling from the Busselton, Western Australia electoral roll with randomised order of invitation[33]
North West Adelaide Health Study[35,36]	The University of Adelaide, Adelaide Medical School[35,36]	Understanding chronic diseases, including the illness itself, related risk factors, and individual impacts, with the overarching aim of improving the health status of those in north-west Adelaide[35]	15-year follow-up, collected in 2015 (37)	1999—ongoing[37]	Follow-ups every ~5 years[37]	Adelaide, South Australia, geospatial information not mentioned[36]	~4,000 adults[37]	Yes[35,36]	Yes[37]	Yes[36]	Random sampling from the Northern and Western suburbs of Adelaide, Australia[36]
The Raine Study[38]	The University of Western Australia[38]	Investigation into the contributing factors for health and disease over the life course, considering genetic, physiological, behavioural, environmental, and sociodemographic elements[38]	Gen2, 27-year follow-up, current version collected between 2016 and 2019[39–41]	1989 onwards[38,42]	Regularly every 1—4 years[38]	Western Australia, latitude and longitude data are collected, and SEIFA variables are derived[38,43]	Sample size of this version N/A for this follow-up. This sample of male and female participants recruited at birth between 1989 and 1991[38,44]	Yes[38,44]	Yes[38]	Yes[38]	Enrolment in the study at birth, children born to women recruited from the King Edward Memorial Hospital in Perth, Australia[38]
South Australian Population Health Survey[45,46]	Wellbeing SA, Government of South Australia, SA Health[45,46]	Observation of the health status of the South Australian population[46]	2020 version, collected in 2020[b]	2018 onwards[46]	Monthly, with annual reports released[46]	South Australia, SA4 on website[47] and SEIFA in report[45,48]	~7,000 people of all ages (n=11,477 in 2020)—a proxy is used if person is under 16 years of age[45,46,48]	No, the study is conducted on a new sample of participants annually[45]	No[45]	No[45]	Random sampling through random digit dialling[45]
Geelong Osteoporosis Study[49]	Barwon Health[49]	The study aimed to (1) collect evidence related to characteristics of bone mineral density (BMD) and bone loss and how these may differ among the population, (2) utilise BMD data for fracture prediction, and (3) collect genetic and environmental data to outline osteoporotic and fracture-related risks[49]	15-year follow-up for men[a], collected between 2016 and 2019[50]	Collection for women began in 1993 onwards, and for men from 2001 onwards[49]	Between every 1 and 5 years[49]	Barwon Statistical Division, Victoria, geospatial information not mentioned[49]	624 men[50]	Yes, with the same group of participants[49]	No[49]	Yes[49]	Age-stratified random sampling from the electoral roll[49]
Western Australia Health and Wellbeing Surveillance System (2019)[51]	Western Australia Department of Health[51]	Observation of the health status of the Western Australian population and aid with policy evaluation and development[51]	2019 questionnaire, collected in 2019[52,53]	2002 onwards[51]	Monthly, with annual reports released[51]	Western Australia, geography is mentioned but not defined[51]	6,907[53]	No, the study is conducted on a new sample each month[51]	No[51]	No[51]	Stratified random sampling, with oversampling in rural areas[51]

(continued)

TABLE 1. Continued

Title	Data custodian	Study aim(s)	Version analysed and collection period	Survey timespan	Intervals of data collection	Location/level of data collection and level of geospatial information released	Sample cohort and sample size of the version used	Longitudinal with same cohort?	Multigenerational	Biosamples collected?	Sampling method
Australian Capital Territory General Health Survey [54]	Australian Capital Territory Health [54]	Aims to obtain health, lifestyle, and well-being-related data about individuals living in the Australian Capital Territory [54]	2019 survey, collected in 2019 [54,55]	2007 onwards [54]	Annual [54]	Australian Capital Territory, geospatial information not mentioned [54]	~1,200 adults [54]	No, the study is conducted on a new sample of participants [54]	No [54]	No [54]	Randomly sampled participants [54]
New South Wales Adult Health Survey [56]	New South Wales Government [57]	Aims to collect health-related information, including health status, health-related behaviours and mediating factors, of individuals living in NSW [56]	2019 questionnaire, collected in 2019 [57,58]	1997 to present [57]	Annual [57]	New South Wales, geospatial information not provided [57]	~13,000(56)	No, the study is conducted on a new sample of participants [57]	No [57]	No [57]	Random sampling [57]

NSW = New South Wales; SA2 = Statistical Area Level 2; SA1 = Statistical Area Level 1; SEIFA = Socio-Economic Indexes for Areas; ARIA = Accessibility/Remoteness Index for Australia; RRMA = Rural, Remote, and Metropolitan Area; N/A = Not applicable.
aProvided through personal correspondence.
bVersion no longer publicly available.

Aging Study and the Geelong Osteoporosis Study. Questions about this category often related to questions about mental or behavioural symptoms, with all surveys asking about mood or affect, motivation or energy, and personality-related symptoms. Other symptoms commonly covered were those related to the respiratory system (11 surveys) and general symptoms, which includes questions about pain (10 surveys).

All surveys asked questions about risk factors for ill health, including alcohol consumption, diet, medical treatments, physical activity, and weight and related factors. A broad range of additional topics were covered by the population health surveys to varying degrees, including transport, providing or receiving care from others, general opinions related to health conditions, leisure activities, community values and attendance at health screenings, and more. An interactive diagram of these findings can be seen at https://aushealthdatasets.w3spaces.com/.

Discussion

The current paper demonstrates that Australian population health surveys cover a broad range of topics, but their inconsistencies produce a "patchwork" of information, which hampers their ability to provide broad insight. The current paper demonstrates that Australian population health surveys cover a broad range of topics, but their differences produce a "patchwork" of information. In a practical sense, these differences are a likely and understandable result of differing funding sources and resources, initial aims and goals, scopes, recruitment methodologies and more. These differences hamper researchers' ability to collate and integrate these data sets to answer larger-scale questions related to place and health. Our study aimed to assist researchers with this process by providing a tool and initial analyses to assist with the identification of suitable data sets and those that may be comparable to one another. It is important to note that the implication of our study is to further suggest that comparisons between data sets would not be required if greater consistency existed between data sets in Australia. Consistency of geo-referenced health information benefits researchers wanting to investigate the effects of place on health.

Examples of such variances can be seen in Table 1, which contains information about the level of geospatial information released publicly in some form (including website presentations). For example, studies such as 45 and Up[10] and Ten to Men[18] may possess SA2-level data, they could not be combined with studies that contain derived variables only, such as SEIFA, as seen with the Australian Longitudinal Studies for Women's Health.[21,22] In addition, researchers will often encounter obstacles related to the release of the data, as some studies may hold the data but will not release it at a lower level of granularity depending on the licensing and privacy regulations of the data custodian. Similarly, other obstacles may be the release of data at a level (e.g. SA2), which will be dummy coded and no longer geographically accurate, defeating the purpose of the collecting geographic data (e.g. the Ten to Men Study[61]) as the understanding of where data has been collected from is lost. The need for geographical granularity to explore questions related to place and health is somewhat self-evident. Jia et al.[8] emphasised the need to incorporate geocoded information into cohort data sets to allow exposures to various environmental elements to be explored, particularly (where possible) in a longitudinal context. Gao et al.[9]

describe the broad range of questions that can be explored with relation to population health, ranging from the potential direct effects of the environment, such as neighbourhood features, on health outcomes, but can also be expanded to consider broader questions that have indirect impacts on health, such as the relationships between place and inequality, that in turn influence well-being. Encouraging advancements in processes and procedures to permit, advance, and standardise geocoding and sharing of geocoded information, while balancing ethical and privacy concerns, allows such research questions to be explored for the greatest benefit of the population.

In general, Australian population health surveys and associated data sets show consistent coverage of basic demographic information, common risk factors (alcohol consumption, diet, medical treatments, physical activity and weight-related factors) and self-rated health. The main ICD-11 category asked about in the surveys was related to symptoms, signs, and clinical findings, primarily focusing on mental and behavioural symptoms, and often related to mental ill-health symptoms assessed via scales such as the Kessler-10 Psychological Distress Scale.[62] A broad range of life- and lifestyle-related items also featured with questions about transport and providing and/or receiving care being the most common. Studies also often provided geo-referenced information; however, this occurred to varying degrees of granularity. This may be an important consideration for research requiring granular location-related information.

Our study has several limitations. First, it is important to note that some longitudinal surveys, such as HILDA, have "rotating content," which explores a different set of characteristics in each study wave,[12] and the findings in this study should be considered as a general guide. Additionally, some longitudinal surveys contain characteristics such as age in the first wave of the survey, so collection in subsequent waves is unnecessary. Similarly, characteristics such as male or female gender were part of the participant screening Ten to Men[19,20] and the Australian Longitudinal Study on Women's Health.[24] Questions surrounding location characteristics may also have been asked during an earlier phase of the study, such as during recruitment, which may not appear in later versions of the survey. Location data may also not be released depending on the data custodian's policies, despite the need for such data to investigate the relationship between place and health.[7] Finally, although every effort was taken to identify all potential large-scale health surveys, some may have been missed during the search.

This is the first paper to directly compare the content of health surveys available to researchers in Australia. Our analysis enables researchers to readily assess the content of population health surveys' characteristics of interest and more efficiently determine how and whether the content of these surveys will be of use for them in exploring their own research questions.

To allow researchers to interact with these findings and efficiently identify the data sets that could assist them, we have developed a website for this purpose: https://aushealthdatasets.w3spaces.com/. This website contains interactive Sankey diagrams containing the same data contained in this paper, along with interactive diagrams for the various symptoms, signs, and clinical findings, risk factors, and remaining variables covered by included surveys. This will allow researchers to highlight and directly compare areas of interest between different surveys with greater ease, particularly for

categories that are covered by many surveys. Future work may incorporate new surveys as they are developed and update this analysis for rolling time windows (e.g. released in the last 5 years) and focus on standardising response frames across surveys to address the lack of comparability between data sets.[5]

Acknowledgments

Author ORCIDs

Branislava Godic 🔟 https://orcid.org/0000-0003-2976-7689
Rajith Vidanaarachchi 🔟 https://orcid.org/0000-0003-1645-0084
Kerry Nice 🔟 https://orcid.org/0000-0001-6102-1292
Sachith Seneviratne 🔟 https://orcid.org/0000-0001-9094-2736
Leandro Garcia 🔟 https://orcid.org/0000-0001-5947-2617
Jasper Wijnands 🔟 https://orcid.org/0000-0002-5832-4134

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst AJ, et al. A systematic review of barriers to data sharing in public health. *BMC Publ Health* 2014;**14**(1):1144.
2. World Health Organization. *Framework and standards for country health information systems*. 2nd ed. World Health Organization; 2012.
3. Chan M, Kazatchkine M, Lob-Levyt J, Obaid T, Schweizer J, Sidibe M, et al. Meeting the demand for results and accountability: a call for action on health data from eight global health agencies. *PLoS Med* 2010;**7**(1):e1000223.
4. AbouZahr C, Adjei S, Kanchanachitra C. From data to policy: good practices and cautionary tales. *Lancet* 2007;**369**(9566):1039–46.
5. Murray CJ. Towards good practice for health statistics: lessons from the Millennium Development Goal health indicators. *Lancet* 2007;**369**(9564):862–73.
6. United Nations. *Sustainable development goals* [Available from: https://sdgs.un.org/goals.
7. Macintyre S, Ellaway A, Cummins S. Place effects on health: how can we conceptualise, operationalise and measure them? *Soc Sci Med* 2002;**55**(1):125–39.
8. Jia P, Lakerveld J, Wu J, Stein A, Root ED, Sabel CE, et al. Top 10 research priorities in spatial lifecourse epidemiology. *Environmental health perspectives* 2019;**127**(7):074501.
9. Gao X, Berkowitz RL, Michaels EK, Mujahid MS. Traveling together: a road map for researching neighborhood effects on population health and health inequities. *Am J Epidemiol* 2023;**192**(10):1731–42.
10. 45 and Up Study Collaborators BE, Redman S, Jorm L, Armstrong B, Bauman A, Beard J, Beral V, et al. Cohort profile: the 45 and up study. *Int J Epidemiol* 2008;**37**(5):941–7.
11. Wilkins R, Botha F, Vera-Toscano E, Wooden M. *The Household, income and Labour Dynamics in Australia survey: selected findings from waves 1 to 18*. Melbourne Institute: Applied Economic & Social Research, University of Melbourne; 2020.
12. Summerfield M, Garrard B, Hahn M, Jin Y, Kamath R, Macalalad N, et al. *HILDA user manual - release 19*. Melbourne Institute: Applied Economic & Social Research, University of Melbourne; 2020.
13. Grunseit A, McGill B, Barnard D, Indig D, Bauman A. *Australian health cohorts. The Australian prevention partnership centre*. 2017.
14. Australian Institute of Health and Welfare. AIHW Data Collections List October 2020 [Available from: https://www.aihw.gov.au/getmedia/6520cd48-04ab-40d5-b5a6-847a2d5c5c1e/Our-data-collections-October-2020.pdf.aspx].
15. Sax Institute. *The 45 and up study: wave 2 data dictionary (first follow-up 2012-2015)*. 2021.

16. Sax Institute. The 45 and Up Study: Wave 2 Data Book (First Follow-Up, 2012-2015) [Available from: https://www.saxinstitute.org.au/wp-content/uploads/W2-databook-May2021.pdf].
17. Sax Institute. Use the 45 and Up Study [Available from: https://www.saxinstitute.org.au/our-work/45-up-study/for-researchers/#FAQ].
18. Bandara D, Howell L, Silbert M, Mohal J, Garrard B, Daraganova G. *Ten to men: the Australian longitudinal study on male health - data user guide*. Melbourne: Australian Institute of Family Studies; September 2019., Version 3.0.
19. Australian Institute of Family Studies. Ten to Men Data Dictionary [Available from: https://tentomen.org.au/data-access-and-usage/data-documentation/data-dictionary/download].
20. Australian Institute of Family Studies. About the study [Available from: https://tentomen.org.au/about-study].
21. Women's Health Australia. Australian Longitudinal Study on Women's Health [Available from: https://alswh.org.au/].
22. Lee C, Dobson AJ, Brown WJ, Bryson L, Byles J, Warner-Smith P, et al. Cohort profile: the Australian longitudinal study on women's health. *Int J Epidemiol* 2005;**34**(5):987–91.
23. Australian Longitudinal Study on Women's Health. *Technical report #43*. 2020.
24. Australian Longitudinal Study on Women's Health. Women's Health Australia, Ninth survey for women of the 1946-51 cohort [Available from: https://alswh.org.au/wp-content/uploads/2021/01/1946-51-survey-9-2019.pdf].
25. Australian Longitudinal Study on Women's Health. *Technical report #42*. 2019.
26. Women's Health Australia. Cohort participation and retention [Available from: https://alswh.org.au/for-data-users/survey-data-overview/cohort-participation-and-retention/].
27. Fitzgerald D. Geo-Coding Variables in the ALSWH 2019 [Available from: https://alswh.org.au/wp-content/uploads/2020/08/DDS_Section5_GeoCoding2016Census.pdf].
28. Women's Health Australia. Recruitment methods [Available from: https://alswh.org.au/for-data-users/survey-data-overview/recruitment-methods/].
29. Melbourne Institute. New Person Questionnaire (NPQ) W19M [Available from: https://melbourneinstitute.unimelb.edu.au/__data/assets/pdf_file/0003/3115596/NewPersonQuestionnaireW19M.pdf].
30. Roy Morgan. HILDA: Living in Australia [Available from: https://melbourneinstitute.unimelb.edu.au/__data/assets/pdf_file/0004/3115597/SelfCompleteQuestionnaireW19M.pdf].
31. Australian Bureau of Statistics. National Health Survey: First results 2018 [Available from: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey-first-results/latest-release#survey-material].
32. Australian Bureau of Statistics. National Health Survey: First results methodology 2018 [Available from: https://www.abs.gov.au/methodologies/national-health-survey-first-results-methodology/2017-18].
33. James A, Hunter M, Straker L, Beilby J, Bucks R, Davis T, et al. Rationale, design and methods for a community-based study of clustering and cumulative effects of chronic disease processes and their effects on ageing: the Busselton healthy ageing study. *BMC Publ Health* 2013;**13**:936.
34. Busselton Population Medical Research Institute. Busselton Baby Boomer Study [Available from: https://bpmri.org.au/research/key-projects-studies/busselton-baby-boomer-study.html].
35. University of Adelaide: Adelaide Medical School. North West Adelaide Health Study [Available from: https://health.adelaide.edu.au/medicine/research/north-west-adelaide-health-study].
36. Grant JF, Taylor AW, Ruffin RE, Wilson DH, Phillips PJ, Adams RJ, et al. Cohort profile: the north west adelaide health study (NWAHS). *Int J Epidemiol* 2008;**38**(6):1479–86.
37. Adams R, Gill T. North West Adelaide Health Study Information Guide [Available from: https://universityofadelaide.app.box.com/s/aevb76iv2ut4gv39vta42tctmxjwo7rz/file/569941547948].
38. Straker L, Mountain J, Jacques A, White S, Smith A, Landau L, et al. Cohort profile: the western Australian pregnancy cohort (raine) study-generation 2. *Int J Epidemiol* 2017;**46**(5). 1384-5j.
39. The Raine Study. The Raine Study Gen2_27 year follow up [Available from: https://rainestudy.org.au/wp-content/uploads/2021/05/Gen2-27_Main_Questionnaire_CodingVersion_20201209.pdf].
40. The Raine Study. The Raine Study Gen2_27 year follow up, Drinks and Caffeine Questionnaire [Available from: https://rainestudy.org.au/wp-content/uploads/2021/05/Gen2-27_DrinksCaffeineQuestionnaire_CodingVersion_20201209.pdf].
41. The Raine Study. The Raine Study Gen2_27 year follow up: TiBS Study [Available from: https://rainestudy.org.au/wp-content/uploads/2021/05/Gen2-27_TiBs_Questionnaire_CodingVersion_20201209.pdf].
42. The Raine Study. The Raine Study Annual Activity Report 2019 [Available from: https://rainestudy.org.au/wp-content/uploads/2020/06/Raine-Annual-Activity-Report-2019-website-version-PDF.pdf].
43. The Raine Study. Available Data [Available from: https://rainestudy.org.au/information-for-researchers/available-data/].
44. The Raine Study. About The Study [Available from: https://rainestudy.org.au/information-for-participants/about-the-study/].
45. Government of South Australia: SA Health. Quick Facts: Methodology, South Australian Population Health Survey [Available from: https://das7nagdq54z0.cloudfront.net/downloads/SAPHS/SAPHS-Methodology-new-branding.pdf].
46. Wellbeing SA. SA population health survey [Available from: wellbeingsa.gov.au/evidence-data/sa-population-health-survey].
47. Wellbeing SA. South Australian population health data [Available from: https://www.wellbeingsa.gov.au/evidence-data/explore-and-request-data/open-data-portal/south-australian-population-health-data].
48. Government of South Australia. *South Australian population health survey 2020 annual report adults*. 2020.
49. Pasco JA, Nicholson GC, Kotowicz MA. Cohort profile: Geelong Osteoporosis study. *Int J Epidemiol* 2011;**41**(6):1565–75.
50. Tembo MC, Holloway-Kew KL, Mohebbi M, Sui SX, Hosking SM, Brennan-Olsen SL, et al. The association between a fracture risk tool and frailty: Geelong Osteoporosis Study. *BMC Geriatr* 2020;**20**(1):196.
51. Epidemiology Branch: WA Department of Health. WA Health and Wellbeing Surveillance System, Technical Paper Series No 1: Design and Methodology 2018 [Available from: https://ww2.health.wa.gov.au/~/media/Files/Corporate/Reports-and-publications/Population-surveys/Technical-paper-no1-Design-and-Methodology.pdf].
52. Government of Western Australia: Department of Health. Health and Wellbeing Surveillance System, 2019 Questionnaire [Available from: https://ww2.health.wa.gov.au/~/media/Files/Corporate/Reports-and-publications/Population-surveys/WA-Health-and-Wellbeing-Surveillance-System-Survey.pdf].
53. Dombrovskaya M, Landrigan T. *Health and wellbeing of adults in western Australia, 2019, overview and trends*. Western Australia: Department of Health; 2020.
54. ACT Government. Epidemiology [Available from: https://www.health.act.gov.au/about-our-health-system/population-health/epidemiology].
55. ACT Government. 2019 ACT General Health Survey [Available from: https://www.health.act.gov.au/sites/default/files/2020-01/ACTGHS%20Questionnaire%20final%20%28programming%20removed%29.pdf].
56. NSW Government: NSW Health. New South Wales population health surveys [Available from: https://www.health.nsw.gov.au/surveys/Pages/default.aspx].
57. Barr M. NSW Population Health Survey Methods - 2012 update 2012 [Available from: https://www.health.nsw.gov.au/surveys/other/Documents/population-Health-surveys-methods-2012.pdf].
58. NSW Government. NSW Population Health Survey, 2019 Collection Year (Version: 30/06/2020) [Available from: https://www.health.nsw.gov.au/surveys/adult/Documents/questionnaire-2019.pdf].
59. World Health Organization. ICD-11 for Mortality and Morbidity Statistics 2020 [Available from: https://icd.who.int/browse11/l-m/en#/?view=G0].
60. Pavlopoulos GA, Kontou PI, Pavlopoulou A, Bouyioukos C, Markou E, Bagos PG. Bipartite graphs in systems biology and medicine: a survey of methods and applications. *GigaScience* 2018;**7**(4).
61. Bandara D, Howell L, Silbert M, Daraganova G. Ten to men: the Australian longitudinal study on male health, release 3 (waves 1-3). *ADA Dataverse* 2022. https://doi.org/10.26193/JDE1TD.
62. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002;**32**(6):959–76.

Appendix A Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.anzjph.2024.100152.