Democratising data to address health system inequities in Australia

Stephanie M Topp, Fintan Thompson, Karen Johnston, Deborah Smith, Alexandra Edelman, Maxine Whittaker, Christopher Rouen, Nishila Moodley, Malcolm McDonald, Ruth Barker, Sarah Larkins

ABSTRACT

Understanding the health status of a population or community is crucial to equitable service planning. Among other uses, data on health status can help local and national planners and policy makers understand patterns and trends in current or emerging health and well-being, especially how disparities relating to geography, ethnicity, language and living with disability influence access to services. In this practice paper we draw attention to the nature of Australia’s health data challenges and call for greater ‘democratisation’ of health data to address health system inequities. Democratisation implies the need for greater quality and representativeness of health data as well as improved access and usability that enable health planners and researchers to respond to health and health service disparities efficiently and cost-effectively. We draw on learnings from two practice examples, marred by inaccessibility, reduced interoperability and limited representativeness. We call for renewed and urgent attention to, and investment in, improved data quality and usability for all levels of health, disability and related service delivery in Australia.

INTRODUCTION

Effective health management requires effective health measurement. Health data are crucial for understanding the scope and distribution of health needs, and for delineating health system inequities. Specifically, health data help us recognise how even the smallest health disparities manifest among different racial, ethnic and socioeconomic groups over time.1–3 They also emphasise disparities in access to healthcare services, the quality of care received by different populations and health outcomes. Health data are fundamental to monitoring and evaluating services and tracking progress towards programmatic and policy goals, including measuring barriers to care.4 Health data are critical for identifying areas where access to and quality of health services are limited, but also underscore informed decision-making regarding broader reforms or more targeted interventions to address these.5

Globally, ongoing efforts to strengthen health systems hinge on the ability to measure disparities in health and healthcare access. Notwithstanding often positive overall health system evaluations, significant inequities exist in Australia’s health system, compounded by limitations and weaknesses in the management of health information and service data.6 Practical experiences reported in this paper help illustrate overwhelming barriers to access and use of multiple, siloed and sometimes inadequately representative health data.

We call for urgent attention to, and investment in, efforts to improved representativeness, access and usability of Australia health data for all levels of health, disability and related service delivery.
Sharing health data and information across sectors and stakeholders remains the hallmark of a mature and high-functioning health system.\textsuperscript{12} Health service coverage, quality and acceptability can be greatly enhanced by combining data from multiple sources; increasing the potential to improve healthcare, health outcomes and health equity.\textsuperscript{13} In this practice paper we draw attention to the nature of Australia’s health data challenges and call for greater ‘democratisation’ of available health data to address health system inequities. We use democratisation to imply the need for greater quality and representativeness of health data as well as improved access and usability that enable health planners and researchers to respond to health and health service disparities efficiently and cost-effectively. We draw on learnings from two practice examples: (1) a data scoping exercise for reform of disability, rehabilitation and lifestyle services in northern Australia (hereafter Disability and Rehabilitation Services Reform) and (2) the compilation of a North Queensland Health Atlas (hereafter the NQ Health Atlas) designed to inform a multisectoral place-based planning project for better healthcare in prioritised communities in NQ.\textsuperscript{14} Both experiences were marred by multiple challenges in sourcing and compiling Australian health data to inform even basic needs assessment and accurate planning.

In the first practice example, we conducted a data scoping exercise to inform Disability and Rehabilitation Services Reform. We identified 20 discrete data collections with indicators relevant to the project including major government (e.g., Australian Bureau of Statistics (ABS) and Australian Institute for Health and Welfare (AIHW)) and non-government (e.g., Public Health Information Development Unit (PHIDU)) institutions. Of these, 4 (20\%) were unsuitable, 5 (35\%) required formal requests to access the data and 11 (55\%) were suitable and immediately accessible.

In the second practice example, a data review and compilation exercise was conducted to create the NQ Health Atlas.\textsuperscript{15} The idea of the atlas was to support and inform a hyperlocal place-based planning project by creating an intuitive and visual platform for local stakeholders that brought together data on population, health status, service availability and workforce (noting the absence of any similar tool locally or nationally). The atlas was designed to be used without the requirement for any specialty skills, with the aim of improving accessibility to data by local planners and service providers. To compile the atlas, which includes approximately 200 layers, 41 separate data collections with a mix of public and non-government sources were reviewed. Of these, 29 (70\%) were publicly accessible; 7 (17\%) were partially accessible (i.e., we were able to see some data, and/or we had to pay to access the data); and 5 (12\%) sources were not accessible at all to the public. The final atlas was used to inform the production of health and service profiles for initial place-based planning workshops in the project communities. Recognising its broader utility, the atlas was launched\textsuperscript{14} and made open to the public,\textsuperscript{15} with discussions ongoing with regional stakeholders about long-term hosting, maintenance and scale-up.

Significant challenges were encountered in the process of both the data scoping and data compilation exercises outlined above. We use the vignettes below relating to data sourcing processes for three foundational health indicators—health status, health service coverage and health workforce—to illustrate these challenges in more detail.

**CHALLENGES IN SOURCING AND ANALYSING HEALTH DATA IN AUSTRALIA**

**Vignette 1: sourcing health status data**

Understanding the health status of a population or community is crucial to equitable service planning. Among other uses, data on health status can help local and national planners and policy makers understand patterns and trends in current or emerging health and well-being, especially how disparities relating to geography, ethnicity, language and those living with disability influence access to services.

Data on the incidence or prevalence of specific health conditions or disability are limited across Australia, particularly in rural and remote settings. The ABS has two key national surveys for health and disability data, the National Health Survey (NHS)\textsuperscript{16}—a recommended source for both national and state/territory-level prevalence data for all long-term health conditions—and the Survey of Disability, Ageing and Carers (SDAC).\textsuperscript{17} However, both these surveys (and other surveys with relevant data such as the General Social Survey\textsuperscript{18} and Household Expenditure Survey\textsuperscript{19}) were not conducted in Modified Monash 7 areas, very remote areas and discrete Aboriginal and Torres Strait Islander communities.\textsuperscript{20} This means the data do not represent people living in most of the far northern Australian localities (cf NHS Methodology\textsuperscript{21} and SDAC Methodology\textsuperscript{22}). While Small Area Estimates (based on the NHS) are theoretically possible, they either cannot be generated for rural and remote areas or are considered unreliable due to small population size. The ABS website itself notes that ‘Small Area Estimation involves producing estimates for small geographical regions for which direct survey estimates are statistically unreliable’.\textsuperscript{23} The separate National Aboriginal and Torres Strait Islander Health Survey (NATSIHS)\textsuperscript{24} does provide some aggregated health measures in Small Area Estimates for regions not covered by the NHS, but the work required to combine and curate data from these different surveys represents a barrier to accessing in its own right.

To inform the NQ Health Atlas, we found additional data sources covering the incidence or prevalence of specific health conditions, such as the Australian Cancer Atlas\textsuperscript{25} which provides modelled estimates of the incidence of various cancers for small geographical units (Statistical Area Level 2); National Diabetes Services Scheme\textsuperscript{26}; the National Study of Mental Health and Well-being\textsuperscript{27}; and the Medicare Benefits Schedule General
Practitioner (GP) Management of Chronic Conditions data. Although data about prevalence of some long-term health conditions are now collected through the ABS Census of Population and Housing, and are available at small geographical levels, limitations include the method of data collection (self-reported via a completion of a form, usually unassisted, or via phone, meaning household members often complete the Census survey for each other) and the 5-year data collection intervals.

Our work experience highlighted the labour-intensive nature of integrating data sets due to differing definitions and measures. For example, the national ABS Census of Population and Housing reports disability as severe or profound and necessitating assistance with a core activity; while other surveys, such as the NATSIHS, measure disability in broader terms, where a long-term condition exists that restricts everyday activities.

Vignette 2: sourcing service coverage data
Indicators of physical availability of health services—service coverage—are key to understanding health system function and progress on universal health coverage (UHC). Australia is praised for having a universal Medicare Benefit Scheme (MBS) and Pharmaceutical Benefits Scheme (PBS). Yet inequity in physical, financial and linguistic access to health services remains a major challenge for people living with disability, in rural and remote areas, Aboriginal and Torres Strait Islander peoples and many ethnic and non-English-speaking populations. Health planners must be provided with a baseline, number and type of health services, planned and delivered. Location is critical when planners assess availability and practicability of service delivery. They need to understand the service gaps and information to address compounding time, transport and accommodation costs that individuals and families must pay to access services outside the community. Such data can also help identify where there are overlapping or duplicated services by different providers, particularly when reliant on non-government providers.

Our experiences highlighted that Australia has no single publicly accessible state, federal or other repository of all health services, as planned or delivered, down to the community level. For example, as part of the curation of data for the NQ Health Atlas, identification of service location required manual and tedious search, extraction, collation and revalidation of data from multiple websites and documents. We found a list of hospitals, noting location required manual and tedious search, extraction, collation and revalidation of data from multiple websites and documents. We found a list of hospitals, noting that some hospitals in Queensland are known as Multi-Purpose Health Facilities under the Queensland Rural and Remote Health Service Framework, and aged care services, with addresses and coordinates at the AIHW MyHospitals and GEN website, respectively. To identify gaps, cross-reference and build on those initial lists, we conducted manual searches of individual Hospital and Health Services (HHS—similar to health districts or local health authorities elsewhere) websites. The searches identified general and dental practice services through individual non-government Aboriginal and Torres Strait Islander Community Controlled Health Organisations websites (cf ref 31); the National Health Service Directory; HealthMap, and the Yellow Pages. Websites for health organisations known to the authors were also accessed since several did not appear on any other database.

In rural and remote areas, delivery of outreach services is heavily relied on and we used the CheckUp database to locate data for the number, frequency and type of outreach services delivered and associated workforce. Funding programmes for outreach services managed by CheckUp were the Rural Health Outreach Fund, Medical Outreach Indigenous Chronic Disease Program, Healthy Ears, Visiting Optometry Scheme and Eye and Ear Surgical Support initiatives in Queensland. However, outreach services delivered under other funding schemes or by community organisations (such as the Royal Flying Doctor Service) are not captured in this data set. Similarly, detailed data for outreach services delivered and funded by the government sector are often effectively inaccessible, requiring specialist knowledge of coordination of outreach services used by a particular organisation as well as the resources and expertise to pursue laborious approval processes (eg, identifying and contacting organisational data custodians and pursuing requisite research ethics and site-specific approvals to access the data). Data for emergency patient retrieval between public health services (services coordinated by Retrieval Services Queensland) were publicly available from the Queensland Government Open Data Portal. Other sources of data not fully investigated as part of our practice experiences but relevant to the broader points being made here would include data from private sector services such as private allied health and support services.

Vignette 3: sourcing health workforce data
Health workforce data are critical to understanding the capacity of health services to meet local health needs. Such data are essential for helping mid to high-level policy makers and planners identify where skills gaps exist and can inform local through sector-wide strategies that align role development and recruitment and service model design with priority health needs. Both our practice experiences highlighted significant challenges regarding access to and accuracy of health workforce data in Australia.

At the time of writing, access to several important workforce data sets remained restricted. These included Queensland Health, Hospital and Health Service (HHS) workforce data which are restricted even internally; and the relatively new federal government-funded Health Demand and Supply Utilisation Patterns Planning tool which has been rolled out only to restricted users in specific organisations. Among data sets that were available, currency and completeness were a challenge. For example, while ABS workforce data provide acceptable detail for specific health occupations, it is based on the
five-yearly Census of Population and Housing, and thus dates quickly. The CheckUp database provides an understanding, though incomplete, of workforce involved in delivering outreach services.

In our combined practice experiences, we found the main and most regularly updated source of accessible data on health workforce was the National Health Workforce Data Set (NHWDS),22 which draws on the Australian Health Practitioner Regulation Agency (AHPRA)43 registration information supplemented by an annual workforce survey. However, AHPRA data do not include non-registered health professions (eg, speech pathologists, Aboriginal and Torres Strait Islander Health Workers,44 data entry clerks) despite their critical roles within and contributions to health and disability services. Moreover, the NHWDS only captures the main practice location of health professionals (though limited data about practice location in the week before completing registration are captured), meaning workforce that services rural and remote areas in locum-style service models is missed. A further limitation of NHWDS data is that estimates of full-time equivalency are based on self-reported hours of work.

Clear and accessible information on using and interpreting the NHWDS is lacking, including poor descriptions on data suppression. Data suppression refers to the use of an indicator such as ‘0 to 3’ in place of the actual health professional count. Use of this type of indicator in rural and remote areas renders workforce data near useless, since many services have fewer than three full-time registered health professionals, and the difference between zero and one is the difference between a service being available or completely absent.

**IMPLICATIONS FOR POLICY AND PLANNING**

The practice experiences and data sourcing vignettes described in this practice paper identify some major (but non-exclusive) challenges and limitations of Australian health data. We are not the first to identify these45 and initiatives such as the federally funded PHIDU have helped improve access to small-area statistics and support planning and research over several decades. Nonetheless, as described in more detail below, Australia’s health data remain beset by issues of accessibility, compatibility, quality and representativeness that require coordinated, large-scale, cross-jurisdictional efforts to remedy.

**Access** The first challenge is to access useful Australian health and disability and respective service data. No single public source holds sufficient health and health service data necessary to inform broad and local policy and planning processes for health and disability services. Currently, an individual’s ability to access these multiple data sources is contingent on time and the confidence to negotiate multiple highly technical platforms, data dictionaries and registration processes. Data are ‘protected’ by often complex ethics and governance (including site-specific) approval processes. Among the data sources that are publicly accessible, several still require institutional approvals, specialised training and even out-of-pocket payment. Despite being publicly funded, and without obvious sensitivities relating to client privacy and confidentiality, several sources remained blocked for public use.

In the National Academy of Medicine Sharing Health Data report the authors note the importance of ‘resolving from a mindset of data guarding to data sharing’.13 The 2022 Australian government-commissioned Strengthening Medicare Taskforce Report similarly emphasised the need for more connected health data under the section titled ‘Modernisation’.16 While limited in geographical scope, the NQ Health Atlas provides an elegant proof of concept that despite the manifest challenges it is possible to curate and collate Australian, state and more localised data into a format and platform that is accessible to planners and policy makers at different levels for both urban and remote regions. Achieving a similar outcome at a larger state or national scale is possible, requiring adequate resourcing to identify or develop a technical platform with a user-friendly interface, and identification of a trusted institutional home to manage the ongoing governance, technical and compatibility requirements to ensure its currency.

**Compatibility and interoperability** of the multiple health, disability and service data sources in Australia remains a challenge. Integration of large data sets (eg, MBS and PBS with state-managed hospital data sets) currently requires time-consuming and expensive application procedures that are cost prohibitive. Variability in the way indicators are measured and interpreted inhibits even integration of many sources. While it is theoretically possible to interrogate data from many ABS surveys including via machine learning and artificial intelligence, for example, differences in the definitions of measurement of certain concepts are a barrier to true and effective interoperability. Using Confidentialised Unit Record Files (CURFs) from ABS surveys (eg, SDAC and NATSIHS) to find comparable indicators across the data collections is another option but requires accessing the ABS CURFs through the online DataLab platform to resolve differences in how certain concepts are measured between surveys. As of 2023, access to the ABS DataLab requires travelling to a capital city, completing a training course, passing a quiz and paying a fee.17 To address these multiple barriers, institutional incentives should be introduced for custodians of national and health data sets to catalyse investment in improved data interoperability across foundational, structural, semantic and organisational levels.88

**Quality** of and gaps in data both emerged as significant concerns. Inconsistencies and gaps are conspicuously exposed when there are multiple data collections with overlapping but inconsistently defined indicators. For service coverage, very few data sources had consistent information about the mode of service delivery, permanent versus visiting services, or the availability or
frequency of satellite services. Gaps between the reported and actual service delivery and health workforce availability were another concern. Health status data frequently displayed a large number of missing responses for certain geographical regions due in part to the modes of contacting people and/or the small number of respondents. For example, the 2016 ABS Census covered the entire NQ region and had an indicator on Care Activity Need for Assistance. Yet, for the entire region of Far North Queensland (FNQ), the proportion of missing information on this indicator was 10.9%, and even higher at 13.5% in the most remote areas (Statistical Area 3, on a scale ranging from Statistical Area 1 to Statistical Area 4) of the Far North. This is compared with 7.6% and 7.1% in Queensland and Australia, respectively (cf Appendix 4, 2016 ABS Census). Missing data increase dependency on proxy indicators. Indeed, various measures of health service utilisation are often used as proxy measures for health status—for example, admissions in children <5 years for dental extractions as a measure of oral health; preventable hospitalisations as proxies for chronic disease rates and estimates of the number of people with disability based on historical supply (most often undersupply in NQ). Yet, to be useful for local-level planning and analysis, quality and deidentified data must be available, with appropriate ethical considerations met at an individual level.

Representativeness of health data (especially health status data) is often poor and remains a consistent concern in the context of the underserved populations of rural and remote northern Australia. Issues of representativeness (see Vignette 1) included the exclusion of very remote areas from data collection in several large ABS surveys (SDAC and the NHS), which although not affecting national estimates have a profound impact on local representation; 17 of 21 local government areas comprising the region of Far North Queensland (FNQ), for example, have some or all of their population in ‘Very Remote’ areas and are thus partially or not represented within the SDAC and NHS.

In a different example, the 2017/2018 Queensland Health Preventative Health Survey (n=2560 in FNQ) does cover the entire geography of the FNQ region but (in its most recent iteration) was conducted via telephone. This data collection method limits representativeness in remote areas due to lower telephone ownership and network coverage. In the ABS Census, the largest statistical collection of key characteristics of residents and network coverage. In the ABS Census, the largest statistical collection of key characteristics of residents and an important source of disability data, the 2016 Census missed 17.5%, and the 2021 Census missed 17.4%, of the estimated Aboriginal and Torres Strait Islander population. Both also undercounted people residing in remote localities.

Achieving comprehensive and accurate methods of data collection remains a long-standing challenge in Australia, complicated by the logistics of a large geography and dispersed population. However, the urban biases in the largely public-financed and operational arrangements, and methodological choices made by state and federal data collection agencies are also implicated, particularly in relation to quality and representation. English language, telephonic or online (internet) surveys can be used to capture most of the coastal-dwelling Australian population, but limit access to those in rural and remote geographies and populations with lower socioeconomic status or non-English-speaking backgrounds. While some gaps are mitigated through the work of special interest groups to collect, collate and release data about specific populations or health conditions (eg, Australian Heart Foundation; special interest groups of the Consumers Health Forum of Australia), reliance on such efforts adds to the complexity of the data landscape and effective decision-making.

Representational weaknesses in Australia’s current health data sets are anticipated consequences of operational and methodological choices embedded in publicly funded data collection exercises that limit the participation, and thus representation, of particular groups. Funding decisions that force data collectors to prioritise survey methods that capture ‘the majority’ at the expense of ‘the whole’ are exclusionary by design, making it easier to overlook disparities in health status and service access of geographically, ethnically or linguistically marginalised populations. Lack of representative data in turn undermines the accurate mapping of health profiles and service needs (preventive, curative and rehabilitative) in what are often the most vulnerable groups, derailing any effective planning and programming for health.

CONCLUSION

Accurate data are a fundamental pillar of UHC—including place-based planning for health service strengthening. Just like other public goods, it should be accessible to, and useable by, different types of stakeholders—from state and federal policy makers down to regional and local planners, providers and private citizens. All the data identified in the two practice experiences reported in this paper, including that which was ultimately inaccessible, were collected as part of publicly funded services and programmes. It should be more adequately representative, more easily accessed, and better curated to enable its use and interpretation by different stakeholders.

While further and more systematic documentation of the above problems is possible and likely warranted, the challenges described in this paper signal the need for urgent attention to be paid to, and investment made in efforts to strengthen data quality and to democratise (in the sense of improved representativeness, access and usability) health data for all levels of health, disability and related service delivery in Australia.

Twitter Stephanie M Topp @globalstopp and Sarah Larkins @SarahLarkins2

Contributors SMT conceived the paper and led writing of the first draft. FT and KJ curated data from the two practice experiences and contributed to writing the first draft. DS, AE, MW, RB, MM, CR, NM and SL contributed to the planning and delivery...
of the work including through critical reflection on one or both practice experiences and made critical edits to the manuscript.

**Funding**
The NQ Health Atlas was created as part of the Integrating Health Care Planning for Health and Prosperity in North Queensland. Funding for this project was received from the Cooperative Research Centre for Developing Northern Australia (CRCNA), which is part of the Australian Government’s Cooperative Research Centre Program (CRCP), with a financial contribution from Tropical Australian Academic Health Centre (TAAHC) and in-kind contributions from project partners. SMT holds an NHMRC Investigators Award (2020–24) GNT1173004.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository. The NQ Health Atlas and the Data Sources and Information for the NQ Health Atlas may be accessed at: https://www.taahc.org.au/research/integrating-health-care-planning-for-health-and-prosperity-in-north-queensland/northern-queensland-health-atlas/

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commericially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**
Stephanie M Topp http://orcid.org/0000-0002-3448-7983
Fintan Thompson http://orcid.org/0000-0002-0798-2671
Karen Johnston http://orcid.org/0000-0002-0477-4666
Alexandra Edelman http://orcid.org/0000-0002-2021-5266
Maxine Whittaker http://orcid.org/0000-0002-1677-2991
Christopher Rouen http://orcid.org/0000-0001-6471-1643
Nishila Moodley http://orcid.org/0000-0003-0375-2438
Ruth Barker http://orcid.org/0000-0002-2546-2581
Sarah Larkins http://orcid.org/0000-0002-7561-3202

**REFERENCES**


