The future of data

1

The future of data

Key messages

Data can be incredibly valuable. They can provide reliable information on health and wellbeing and on the impact of policies and programs. However, they do not speak for themselves – high-quality analysis is needed to bring data to life.

Much of the potential that data offer is yet to be realised as data are not always brought together in the way they could be. Data on government services have traditionally been analysed in isolation. This approach does not provide the sort of insights that can be gained by looking across service systems and focusing on how people engage with multiple services.

Data linkage can provide much more comprehensive insights than are possible by looking at individual services in isolation, but it can be slow and resource intensive.

The National Disability Data Asset, being led by the Department of Social Services, is based on a new, enduring approach to data linkage in Australia. To deliver this, the Australian Bureau of Statistics (ABS) and the AIHW are working with the states and territories to implement the Australian National Data Integration Infrastructure (ANDII). Importantly, the ANDII will bring together data from the states and territories and the Australian Government to provide more detailed insights than are possible by looking at data at one level of government in isolation.

While data linkage and integration have huge potential, they cannot overcome the adverse impact of the substantial data gaps that still exist. Some of these gaps have been highlighted by Royal Commissions and major inquiries. For example, the Royal Commission into Aged Care Quality and Safety highlighted several data gaps that limit an understanding of the quality of age care services and how these services affect wellbeing.

This chapter provides information on plans to improve data on aged care and on family, domestic and sexual violence. It also highlights the successful use of both single-touch payroll data by the ABS to provide very timely data on changes in employment, and linked data by the AIHW to better understand the health system.

Why is comprehensive and high-quality data important?

The word 'data' is the plural form of the Latin word *datum*. The ABS defines data as measurements or observations that are collected as a source of information; the *Cambridge Dictionary* defines data as information, especially facts or numbers, collected to be examined and considered and used to help decision-making, or information in an electronic form that can be stored and used by a computer.

On occasion, some people can be uncomfortable about quantitative data, feeling that they can obscure personal experience. This reservation can be overstated: data can, in fact, reflect the experience of thousands or millions of people.

Data can provide:

- representative information to inform policy and an understanding of societal trends
- important insights into how things are changing for different people. How do employment rates vary, for example, by age, gender, location and ethnic background?

Used well, data can provide reliable information on the quality of people's lives and how that is changing; importantly, they can show how this can vary for different groups. They can also provide information on the impact of policy and programs.

Data do not speak for themselves

> *'I have been in rooms with data and listened very carefully. They never said a word.'*
> (Wolpoff 1975)

While data can be very valuable, data do not speak for themselves. The value of data is obviously a function of how well they are used.

The quality of statistical analysis can affect the conclusions drawn from data. Advances in statistical techniques over the last 30–40 years have, in some cases, substantially improved the quality of insights that can be gleaned. To give an example, before the 1980s, economists often ran linear regression equations on time series data. As Engle and Granger (1987) pointed out, without care, this approach could lead to spurious correlations, as any apparent correlation between the data sets may simply reflect the fact that both series are increasing or decreasing over time. Engle and Granger (1987) developed a new approach to avoid such spurious conclusions, which transformed the way that economists use time series data.

In many cases, the data themselves will not answer the questions that people want them to. When assessing the impact of a policy or program, one should not just 'look up the data'. For example, one should not assess the impact of an employment program by simply observing the employment data, as employment is affected by many factors that may have nothing to do with that program, such as the state of the economy.

In evaluating the impact of policy or programs, one needs a sense of what the counterfactual is – that is, what would have happened in the absence of the policy or program? This sort of assessment requires considered analysis through, for example, modelling and the use of comparison or control groups.

Data are increasingly being used to gain new insights in many fields, including professional sports; but, again, the value of the data comes from the quality of analysis not just from the data themselves. A well-known example of this is evidence around the so-called 'hot hand' in basketball. The idea of a hot hand is that once a player shoots one goal, this can increase the likelihood that they will shoot more goals in a streak, and that the perceived hot hand does not simply reflect random variation. A 1985 paper by Gilovich and colleagues concluded that there is no hot hand in basketball; indeed, it was subsequently described as a myth. Since then, it is now generally accepted that the conclusion of these researchers that there is no hot hand in basketball was misplaced – and that there is now evidence that the so-called 'hot hand' is real (Gelman 2015). Importantly, this new conclusion did not come from new data – it came from better analysis. Data do not speak for themselves.

The use of artificial intelligence and machine learning can provide insights that may not come from more traditional statistical analysis but, in some cases, the approaches can be similar. For example, some machine learning approaches use logistic regression, which has been borrowed from statistics. In some cases, machine learning approaches will find connections in the data that a human would be unlikely to find. For some applications, it may not matter whether these relationships are causal or have a clear explanation. If a set of variables can be used to forecast what will happen in another variable, this may be valuable for forecasting, as long as this predictive relationship holds. However, if data are to be used to inform policy, one would normally want to understand the causal links.

It is difficult to predict exactly what machine learning and artificial intelligence will mean for data analysis, but it is clear that they will probably transform the amount of analysis that will be possible and the speed at which it can be done. These approaches are also likely to transform the way in which data are coded and structured.

Service systems and data linkage

A high proportion of data analysis on the performance of government services focuses on particular services in isolation. While this may, at times, be appropriate, in many cases it can be quite limited. By focusing on individual services, it is not possible to understand how service systems operate or to observe the way in which people use multiple services.

There is no agreed definition of the term 'service system'; however, people often talk about the health system or the justice system. When people use these terms, they are referring to a range of services not just one type of service. For example, the health system refers not only to hospitals but also to primary health care, while the justice system refers not only to prisons but also to the role of police and the courts.

In some federations, there are quite clear distinctions between the roles of different levels of government. In Germany, for example, education is primarily the responsibility of the various states (or *Länder*), with the federal government playing a relatively minor role. In Australia, on the other hand, responsibilities are often split, even within the one policy area. For example, the Australian Government plays the key role in funding primary health care while the states and territories operate the hospital system. In housing, the states and territories are responsible for social housing and homelessness services while the Australian Government is responsible for rent assistance.

This split in responsibility across levels of governments in Australia has made it challenging to bring data together to understand service systems; however, these challenges are increasingly being overcome through data sharing and data linkage. Through data linkage, it is possible to understand how people use various services; this, in turn, provides information to policy advisers on the interface between those services.

The AIHW created the National Integrated Health Services Information Analysis Asset (NIHSI AA) in 2018. This linked data asset brings together Australian Government data both on Medicare services covered by the Medicare Benefits Schedule (MBS) (such as general practitioner, specialist, pathology and diagnostic imaging services) and prescriptions supplied under the Pharmaceutical Benefits Scheme (PBS) with state and territory hospital and mortality data. This asset has been used to gain new insights on the operation of the health system (including analysis of service use in the last year of life) and on the extent to which people receive appropriate medications once they leave hospital. None of this analysis would have been possible with existing data on individual services.

The AIHW has also undertaken a large number of data linkage projects covering welfare services, including housing, child protection and the receipt of income support statements. The ABS has facilitated many linkage studies through the Multi-Agency Data Integration Project (MADIP), which combines information from data sets such as the ABS Census of Population and Housing (Census), the ABS National Health Surveys, the MBS, the PBS and mortality data.

As noted in *Australia's welfare 2021* (see AIHW 2021), several Australian states are leaders in data linkage. New South Wales and Victoria, for example, have both used it to gain a much better understanding of how people use various state services.

New national linkage system – the National Disability Data Asset and the ANDII

While notable progress has been made in recent years, Australia still lacks a genuine national data linkage system; rather, it has several state and territory systems, with Australian Government agencies (such as the ABS and the AIHW) playing an important role. In practice, it is time consuming to bring Australian, state and territory government data together for data linkage.

The National Disability Data Asset (disability data asset) is led by the Department of Social Services and overseen by the Disability Reform Ministers Council. The ABS and the AIHW are partnering to design, develop and deliver technical aspects of the disability data asset. The Pilot found that the key limiting factor when it comes to data linkage in Australia is data governance, not data availability or information technology (IT) infrastructure. (See http://www.ndda.dss.gov.au for information on the disability data asset.)

To deliver an enduring disability data asset, the ABS and the AIHW are working with the states and territories to implement the ANDII.

The ANDII is the national linkage and integration infrastructure. It includes a national spine and linkage model. It also includes data governance and streamlined data-sharing arrangements that enable the creation of data assets. ANDII also refers to an Information Communications Technology (ICT) solution, supported by linkage and analytical capabilities. The ANDII ICT solution is being built in the Cloud, which provides opportunities for enhanced security, scalability and controlled data sharing.

Importantly, the disability data asset will be co-governed by the Australian, state and territory governments and the disability sector, including people with lived experience. This will ensure that data are used appropriately, and that state and territory and Australian Government data are regularly linked together.

The ANDII will also be co-governed with the states and territories. It will entail more enduring data linkage than the slow, project-by-project approach that currently characterises much data linkage in Australia. The new system will also involve a 2-way data flow so that states and territories and the Australian Government have better access to data from other levels of government. Both the disability data asset and the ANDII will be subject to strict privacy and ethical oversight.

The ANDII has a huge amount of potential and can provide a genuine national approach to data linkage in Australia. Importantly, it will also facilitate the linkage of data across sectors to get a better understanding of how people use services across different policy areas. The value of this approach was shown by several of the disability data asset Pilot studies. For example, the Victorian study showed how National Disability Insurance Scheme (NDIS) clients with psychosocial disability used the

Victorian mental health system, while the New South Wales study showed how people with disability use the justice system. 'Chapter 5 Use of mainstream services and outcomes achieved for people with disability' includes Pilot findings from the disability data asset.

With access to linked data, policy advisers can move beyond looking at individual services in isolation. For example, how do people who use homelessness services use the health system? How can the health system better serve people who experience homelessness?

Linked longitudinal data can also be used to better understand how early events in someone's life can affect subsequent events. The AIHW demonstrated, for example, that a high proportion of young people in the juvenile justice system had earlier involvement with the child protection system (AIHW 2022c, see also 'Chapter 6 Children who have experienced child protection, youth justice and homelessness').

Linked longitudinal data have considerable potential to increase the quality of policy and program evaluation in Australia. There is often a lack of good evidence on programs and policies due to a dearth of good evaluations (Gray and Bray 2019). This lack of high-quality evaluation reflects several factors: one is a lack of access to high-quality data. Under existing arrangements, it can be difficult to use linked data for evaluations because of the time required to obtain all the necessary approvals. The more enduring approach to linkage through the ANDII will help to resolve this issue.

Linked data can be very useful for evaluation as it can be used to construct control and comparison groups so that outcomes for people who do and do not use a particular program can be compared. The marginal cost of using linked data in this way can be low – it is not always realistic given the cost, to undertake new surveys to conduct evaluations.

Data linkage and data gaps

While data linkage has huge potential, it cannot overcome the challenges created by data gaps. Several Royal Commissions have highlighted critical data gaps – these cannot be resolved with data linkage. Key recommendations of the Royal Commission into Aged Care Quality and Safety (Royal Commission 2021) were to create an aged care National Minimum Data Set and improve data on the interaction between the health and aged care systems. Another Royal Commission – the Royal Commission into Violence, Abuse, Neglect and Exploitation of People with Disability – highlighted a lack of useful, and nationally consistent, data on the extent of violence against (or abuse, neglect and exploitation of) people with disability. The 2020 Productivity Commission Report on mental health also highlighted key data gaps, particularly for mental health services provided in the community (Productivity Commission 2020).

AIHW's role in creating and reporting on evidence

The AIHW is a leading health and welfare statistics agency. Its legislated role is to work with others to develop information standards and collections across health and welfare, and to publish statistics across these areas. It works closely with governments – including state and territory authorities with health and welfare responsibilities, the ABS, and other statistics agencies – and with the academic and non-government sector to make this happen.

The AIHW collects, accesses and uses data from a range of sources, including from administrative, survey, longitudinal and linked sources, and the Census to present information on:

• the characteristics of people and their health and welfare needs, and how these change during their lives

• how the health and welfare needs of people differ, depending on where they live

• the availability and accessibility of health and welfare services in those places

• how service use changes over time.

The AIHW provides regular information on the health and welfare of the Australian population to assess how outcomes are changing over time. It also provides information on the health and welfare of particular groups, and on how they fare relative to the rest of the population. These groups include Aboriginal and Torres Strait Islander people; Australians who live in remote areas; veterans; older Australians in the aged care system; and Australians with disability, mental illness, in the child protection and justice systems, or who are prisoners.

The AIHW validates and standardises the data to allow comparisons to be made between different population groups, different places and over time.

The AIHW strives to make its data and findings accessible to a range of audiences through the release of many products, including summary and detailed reports. It has improved accessibility to information through interactive data visualisation. Much of the AIHW's data are now presented in an interactive form to make it easier to gain insights from data – and to make the data interesting.

The AIHW has a range of websites, some of which provide detailed data on particular products, such as:

• Australia's Disability Strategy Outcomes Framework

• GEN aged care data

• housing data dashboard

• mental health

- My Hospitals

- National Mental Health Service Planning Framework

- suicide and self-harm monitoring

- Aboriginal and Torres Strait Islander Health Performance Framework

- Indigenous Mental Health and Suicide Prevention Clearinghouse

- Regional Insights for Indigenous Communities.

These websites can all be accessed at https://www.aihw.gov.au/.

Case studies

The following section of this article presents 7 case studies. They show how the AIHW is filling data gaps and using linked data to gain new insights. These studies provide information on data gaps and challenges, plans to improve data, the importance of high-quality and comprehensive data and the potential for data linkage and integration across systems to support reporting and decision-making. They also identify work being done to improve data outside the AIHW. One study describes how The Smith Family, though partnership with the South Australian Department for Education, is using data to improve the operation of their programs to support students and their families. Another highlights the successful use of single-touch payroll data by the ABS to provide very timely data on changes in employment during the COVID-19 pandemic.

The 7 case studies presented are:

1. National Disability Data Asset

2. Using data integration to improve health and welfare insights for the ex-serving Australian Defence Force population

3. Data developments in family, domestic and sexual violence

4. Aged care data: challenges and improvements

5. Using data linkage to explore patterns of health service use in the last year of life among people who died by suicide

6. Better data for better futures: The Smith Family *Learning for Life* program

7. Using single-touch payroll data to meet changing needs.

1. National Disability Data Asset

About 1 in 6 (18%, or 4.4 million) Australians live with disability. The Australian Disability Strategy 2021–2031 (the Strategy) outlines a vision for a more inclusive and accessible Australian society where all people with disability can fulfil their potential as equal members of the community. The Strategy also commits to reporting on progress achieved against the Australian Disability Strategy Outcomes Framework (ADSOF), which reports on a broad range of measures from areas including health, education and employment.

Many of these measures are reportable through national surveys and administrative data collected by specialist disability supports, such as the NDIS and the Disability Employment Scheme. However, some key measures cannot currently be reported due to gaps in the existing evidence base.

Most people with disability do not use specialist disability services (only around 1 in 4 people with disability under age 65 are active participants in the NDIS); however, all use mainstream services such as health care, welfare and/or education services. But the data systems for these services often do not identify people with disability. And where these data are collected, they are often inconsistently defined or collected only for a subset of disability types. As a result, there is little consistent information on the use of mainstream services by people with disability.

Data linkage methodologies have the potential to fill many of these data gaps by sharing data across systems to:

- enable people with disability to be identified in mainstream systems that do not include a disability status 'flag'
- study pathways taken by people with disability through specialist and mainstream services
- study outcomes achieved for people with disability using particular services.

In recent years, governments in Australia have been working together to develop the National Disability Data Asset to enable such data sharing and linkage. During 2020 and 2021, Pilot studies integrated over 50 data sets from the Australian Government; the National Disability Insurance Scheme (NDIA); and the New South Wales, Victorian, Queensland and South Australian governments. This work provided valuable insights for 5 high-priority research projects focused on:

- early childhood supports
- interaction of people with disability with the justice system
- pathways from education to employment
- services and supports for people with disability and mental health issues
- outcomes measurement, focusing on housing-related supports.

(continued)

These studies highlighted the tremendous potential for such data linkage approaches to support outcome reporting and inform policy development. For example, the investigation into specialist homelessness services (SHS) found that the current administrative data collections may be seriously underestimating the number of clients with disability.

The study found that 37% of all SHS clients in 2019–20 in the 4 participating jurisdictions were identified as having disability, compared with 2.3% of clients identified using the current disability flag. Psychosocial disability, in particular, had not been well captured in the existing homelessness services data collection (73% of clients had psychosocial disability according to linked data compared with 62% having a current mental health issue as described by existing data). The linked data also showed that 45% of SHS clients with disability had multiple disabilities and that clients with disability were more likely to be older than clients without disability (Figure 1.1).

Figure 1.1 SHS clients with disability are older than SHS clients without disability

Age distribution of SHS clients with and without disability in NSW, Vic, Qld and SA, 2019–20
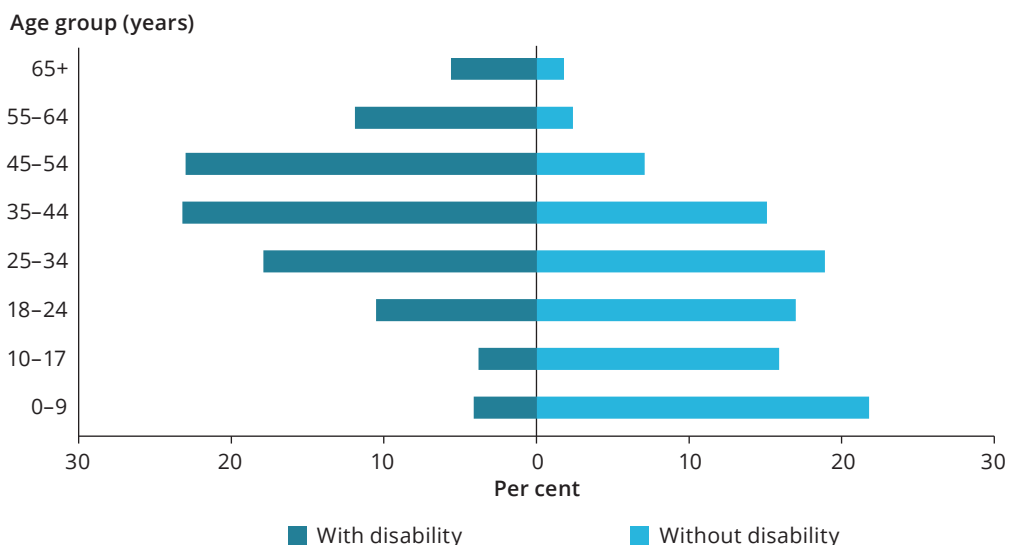
Notes
1. There are limitations with identifying people with disability under the age of 25 and over the age of 64 years. Caution should be used for ages outside this range.
2. Percentages may not sum to 100 due to rounding.
3. Per cent calculations are based on totals less 'Not stated' (unless 'Not stated' has been combined with another category for confidentiality reasons).
Source: AIHW (2022) NDDA Outcomes Reporting pilot, unpublished AIHW analysis of data asset.

'Chapter 5 Use of mainstream services and outcomes achieved for people with disability' further explores the findings of the 4 National Disability Data Asset Pilot studies that relate to community services provided for people with disability.

The National Disability Data Asset partners are continuing to work on developing the asset, based on the learnings and consultations undertaken during the Pilot phase, and building towards the evidence required by the ADSOF and other emerging policy priorities.

2. Using data integration to improve health and welfare insights for the ex-serving Australian Defence Force population

The AIHW is a leader in providing high-quality health and welfare information. Its work program has built a strong evidence base for better decisions that deliver improved health and welfare outcomes. The evolution of its data integration program has exemplified innovation in identifying and rectifying key information gaps, as well as in responding to opportunities to develop and capture the data required to inform national priorities.

The AIHW conducts data integration in partnership with data custodians and specialists in integration and analysis. Linkage projects requiring the integration of Australian Government data are undertaken by an accredited integrating authority (see https:// statisticaldataintegration.abs.gov.au/roles-and-responsibilities/integrating-authorities).

The AIHW has met stringent criteria covering project governance, capability and data management to gain this accreditation. In this capacity, it is trusted to integrate Australian Government data for high-risk research projects. To date, its integration projects have generated improved research outcomes that have identified vulnerable population groups, improved the understanding of health risk factors, and contributed to the development of targeted interventions. These projects have fostered new insights into dementia, disability, health service use, patient experiences of health care, and suicide.

The AIHW has extensively used the ABS's MADIP data asset to gain health and wellbeing data insights. The MADIP data asset was established in 2015. It is a secure data asset, combining data from various government agencies on health, education, government payments, income and taxation, employment and population demographics (including data from the Census) over time. It provides whole-of-life insights on various population groups in Australia, such as the interactions between their characteristics; their use of services like health care and education; and their outcomes, like improved health, employment and income circumstances. Integrated data assets, such as the MADIP data asset, allow complex questions and changes over time to be analysed with new and expanded insights to be gained that are not available from single data sources (ABS 2022b).

The ABS is trusted as the accredited integrating authority for the MADIP data asset. It collects and combines the data and provides access to authorised researchers while protecting individual privacy and keeping the information available in the MADIP data asset secure at all times.

Two AIHW projects have used the MADIP data asset to better understand the ex-serving Australian Defence Force (ADF) population following their separation from service. These projects were 2 of the first data integration projects to use Linkage Spine Interoperability

(continued)

(von Sanden 2020); this enabled Department of Defence (Defence) personnel data held at the AIHW to be combined with data held at the ABS in the MADIP data asset without the need for these organisations to share personal identifying information. Linkage Spine Interoperability increases efficiency in creating integrated data products between 2 agencies; it also reduces the risks of sharing personal information, and facilitates better use of person-centred data, enabling research efforts to be expanded.

Through data integration, the 'Understanding the wellbeing characteristics of ex-serving ADF members' project (Project 1) was, for the first time, able to include:

- comprehensive insights on female veterans not typically available through national surveys

- exploratory analysis of veterans' families, including spouses and children

- the release of 'Wellbeing outcomes on education, employment, income and housing circumstances' – a key highlight being the release of employment status for ex-serving ADF members following their transition from the ADF

- modelling analysis that provided insights on the ADF service characteristics that had the strongest statistical association with wellbeing outcomes, such as being employed, unemployed, attaining a bachelor's degree and earning a higher income (AIHW 2022b).

Data insights from this project paint a largely positive picture of wellbeing outcomes following separation from the ADF.

At the time of the 2016 Census, many ex-serving ADF members had attained higher education qualifications, were employed, earned higher incomes than the Australian population, owned their own homes (including people paying mortgages) and were socially connected by living in a family type household.

However, this is not the case for everyone. Those people who separated from the ADF involuntarily for medical reasons or served fewer years had an increased risk of facing wellbeing challenges (including higher unemployment rates and lower levels of education qualifications) and were receiving lower incomes in 2016 than people who separated for any other reason. People who served fewer years experienced similar wellbeing challenges when compared with people who served longer or served in the army or air force.

This project highlights the varied wellbeing outcomes following transition from ADF service. Through the strategic partnership between the AIHW and the Department of Veterans' Affairs, the work program will continue to ensure that the research captures the variation of outcomes of the ex-serving ADF population and supports the needs of all Australia's veterans and their families. Future work will include analysis of the 2021 Census ADF service status question which asks all Australians participating in the Census whether they had served in the ADF (see https://www.abs.gov.au/articles/australian-defence-force-service for more information).

(continued)

Data integration with the MADIP data asset was also used to explore socioeconomic characteristics of ex-serving ADF members who died by suicide (AIHW 2022a) (Project 2). For the first time, the AIHW undertook analysis to identify potential demographic and socioeconomic factors related to death by suicide among ex-serving ADF males. The analysis compares the cohort with the total ex-serving ADF male population, and with the general Australian male population who died by suicide.

Findings from this project show that ex-serving ADF males who died by suicide were more likely to be younger, to live alone, to have lower incomes or to be widowed/divorced/separated/never married.

This project complements the existing AIHW veterans' health and welfare program that aims to build a comprehensive profile of the health and welfare of Australia's serving and ex-serving ADF population. It also builds on the broader suicide and self-harm monitoring work conducted by the AIHW on the whole Australian population. It was used to inform the Royal Commission into the Defence and Veteran Suicide interim report (see https://www.aihw.gov.au/suicide-self-harm-monitoring and https://defenceveteransuicide.royalcommission.gov.au/publications/interim-report).

3. Data developments in family, domestic and sexual violence

Family, domestic and sexual violence (FDSV) is a major health, welfare and social issue in Australia and can have a lasting impact on individuals, families and the community. It can affect people of all ages and backgrounds, but predominantly affects women and children. There is currently no national definition of what constitutes FDSV. Violence is a broad term, referring to behaviours (or patterns of behaviour) that cause harm. Violence can occur in the form of assault, threat, abuse, neglect or harassment and is often used by a person, or people, to intimidate, harm or control others. Examples of forms of violence include sexual assault, emotional abuse, financial abuse, and stalking. FDSV can occur within a range of relationships and settings (see 'Family, domestic and sexual violence' at https://www.aihw.gov.au/reports/domestic-violence/family-domestic-and-sexual-violence).

The AIHW adopts a broad approach to reporting on FDSV, which includes violence against all people, and supports understanding of gender-based violence. Gender-based violence refers to violence against someone because of their gender, and is the focus of the National Plan to End Violence against Women and Children 2022–2032.

Data are essential to understand the extent, nature and impact of FDSV. Data also provide insight into peoples' understanding and attitudes towards violence and how people engage with health and welfare services after experiences of violence. The knowledge gained from these data can then be used to inform decision-making, service planning and resource allocation to improve outcomes for people who are, or may be, affected by FDSV. Another important role for data is in monitoring and evaluating programs provided specifically to assist people affected by FDSV and gender-based violence.

Through the National Plan to End Violence against Women and Children 2022–2032, the Australian, state and territory governments are committed to growing the evidence base and strengthening data collection systems. This builds on the work completed by governments and national information and research agencies to improve FDSV data and reporting under the previous national plan (National Plan to Reduce Violence against Women and their Children 2010–2022).

FDSV-related data and reporting have been substantially improved over the past decade but several national gaps remain. Currently, available data, collated from a range of sources, provide an incomplete national picture. These data sources include hospitals, child protection services, homelessness services, police, courts and population surveys. National gaps include having limited FDSV-related data on:

• services, including specialist FDSV services (services that respond specifically and mostly to people experiencing FDSV), primary health care, ambulance/paramedic care, emergency department care, drug and alcohol services and mental health services

(continued)

- service pathways, impacts and outcomes for victim-survivors, perpetrators and families
- select population groups.

In some cases, data are available at the state and territory level; however, comparability across collections is limited as different definitions and/or methods are used to support different requirements (which can be related to legislation and/or service scope). Data availability and/or comparability can also vary across public and private sectors.

The Australian Government funds the AIHW to develop a prototype specialist FDV services data collection and a national FDSV integrated data system. The prototype collection aims to draw together data from providers of specialist FDV crisis services. A large component of the project is devoted to understanding what data are collected, the systems used to collect and store the data, and current reporting by states/ territories and/or services. The project will culminate in the development of a prototype collection that will support the collection of consistent and comparable data, regardless of the jurisdiction where the service is provided. The data to be included are still to be agreed; however, as well as basic demographics of people seeking and receiving specialist services (and information about the services), the collection of information on unmet demand, risk assessments, pathways through services (including referrals), outcomes and case complexity will be explored. A key principle of the prototype collection is that it uses a flexible and staged approach to data collection, which can be expanded and built on over time to include other types of services.

The integrated data system project aims to develop the foundations for an enduring national integrated data system related to FDSV victim-survivors and people who use violence. The long-term aim is to provide a more complete picture and a better understanding of the life experiences and outcomes of people experiencing FDSV by analysing 'joined up' de-identified national data collections. There are some challenges in developing this system due to the limited national data available to identify victim-survivors and people who use violence, but the collection of data from specialist FDV services will go some way to rectifying this in the future. The scoping stage of the project has involved consultation with stakeholders on potential research and policy questions that could be dealt with by a FDSV integrated data system.

4. Aged care data: challenges and improvements

The Royal Commission into Aged Care Quality and Safety concluded in 2021. Its final report identified shortcomings in aged care, including in data about aged care. The AIHW and the Department of Health and Aged Care are partnering to deliver:

• an aged care data strategy to drive system-wide improvement

• an Aged Care National Minimum Data Set (NMDS) to improve aged care data quality

• a National Aged Care Data Asset to better understand interactions and outcomes.

The data strategy will guide the approach for developing and continually improving a comprehensive aged care data system. Many people use aged care information and data, and for many different reasons. The AIHW is using what was heard through stakeholder consultations to develop the draft data strategy: the aged care data system must respond to the needs of different users and provide them with timely and comprehensive information.

The Aged Care NMDS will focus on a core set of aged care data variables. Common data standards and specifications are key to improving the quality of data about aged care.

The National Aged Care Data Asset will reveal how Australians access aged care and other related systems like health care, and what their needs and outcomes are. Data systems traditionally focus on individual programs, areas or settings and specific time periods. But, as people age, their care and support needs often become more complex and they rely on different services.

The data asset will be modular, including aged care, health and other services. It will allow for a broad system-based view, and ongoing study of questions related to evaluation, policy and research.

Benefit	What does this look like?
Improved reporting	Producing indicators of aged care system performance, such as how frequently or intensively people who use aged care also use other services, or how these patterns vary between different groups of people using aged care.
Monitoring effects	Assessing how particular population groups are affected as policies are implemented, the environment changes or other transitions occur. The data will facilitate follow-up analyses on the same cohort over time and increased visibility of cross-sector outcomes.
Planning at the local level	Supporting flexible use of in-scope information where data cross sector or region boundaries. For example, sometimes information is better centred around the geographical region where events take place, rather than on individual sectors, to present a meaningful picture of activity at the local level.

(continued)

Benefit	What does this look like?
Resource for research	Providing access to relevant data for research on the interface between aged care and health care, other high-priority topic areas and questions relating to particular populations, such as: • younger people in aged care • people with dementia • people using palliative care • veterans.
Setting the stage for the future	Adjusting to new or expanded data, systems, platforms and environments as they come into being. The data asset will be part of the scaffolding to build a better aged care data system.

5. Using data linkage to explore patterns of health service use in the last year of life among people who died by suicide

Using the NIHSI AA version 0.5, the AIHW explored patterns of health service use in the last year of life for people who died by suicide. This analysis showed that between 1 July 2010 and 31 December 2017, people aged 15–64 who died by suicide accessed fewer health services in their last year of life than people who died from other causes. About half (49%) of the people who died by suicide did not have any contact with hospitals during their last year of life, compared with 24% of people who died by other causes. About 1 in 10 (11%) people who died by suicide did not access any of the analysed health services in their last year of life.

This analysis also revealed some differences in health service use by sex and age group. Females (59%) who died by suicide, for example, were more likely to attend hospital for any reason in their last year of life than men (48%). Females aged 34–44 were the most likely to attend hospital (61%) and males aged 55–64 were the least likely to attend (45%). Females (30%) who died by suicide were more likely than males (19%) to have had a mental health hospitalisation (this excludes Emergency Department presentations) in their last year of life. Only 13% of women and 6% of men who died by suicide had hospital admissions for intentional self-harm in their last year of life.

An advantage of using the NIHSI AA, compared with previous studies, is that patterns of service use can be explored in both the MBS and the PBS. As a result, this analysis showed that people who die by suicide are considerably more likely to access MBS and PBS services in their last year of life than hospital services. This analysis also highlighted differences in MBS and PBS use between males and females. For instance, females who died by suicide had a higher use of MBS mental health services (57%) or PBS mental health prescriptions (71%) than males (37% and 50%, respectively) in their last year of life.

This analysis reinforces the importance of ensuring that suicide prevention activities also focus on people accessing health and community services (or no services at all), and not just on people who access hospital, MBS or PBS services (See 'Patterns of health service use in the last year of life among those who died by suicide' at https://www.aihw.gov.au/suicide-self-harm-monitoring/data/deaths-by-suicide-in-australia/health-service-use-in-the-last-year-of-life).

6. Better data for better futures: The Smith Family *Learning for Life* program

The Smith Family, an Australian not-for-profit organisation helping young Australians to overcome educational inequality, run a number of learning programs. The largest is its *Learning for Life* scholarship program, which provides extra support for children experiencing disadvantage to engage with school. It uses an early intervention approach to help children overcome the obstacles caused by disadvantage that often make it difficult to move through school and onto employment or further study.

The *Learning for Life* program has 3 integrated components that provide financial, relational and programmatic support, with The Smith Family team members providing support directly to students and families when they need it. The program, which supports around 60,000 children and young people each year (The Smith Family 2022), seeks to improve young people's educational, employment and life outcomes. Each student has a Unique Student Identifier; this enables the program to track school attendance, achievement and completion, as well as post-school engagement in employment, education and training. For example, 84% of students in year 12 and on the program in 2020 were in education, work and/or training 18 months after leaving school (The Smith Family 2022).

The effectiveness of a program like this – one that provides targeted support when it is needed – is highly influenced by the quality and timeliness of available data; in this case, the student-level data that the program team members have access to, and how quickly they have that access. Until recently, the educational information (like school attendance data and literacy and numeracy grades) that The Smith Family team members need to help them determine when and what type of support is needed, was available only through school reports once a year – usually after the school year was concluded. This timing presented a barrier to providing timely and targeted frontline service.

An ongoing partnership between The Smith Family and the South Australian Department for Education has seen this data access dramatically improve for *Learning for Life* students in South Australia since 2021. Through this data exchange partnership, The Smith Family has gained access to key educational data in real time for students on their program via a Department-hosted dashboard, as well as to data for similar students not on the program (for outcome comparisons). Data include attendance, the reasons for absences, achievement data (for example, meeting/not meeting the literacy/numeracy standard), student demographics, school enrolment, and behaviour management incidents – for example, suspensions. The data in the dashboard are the latest available; for example, the student-level attendance data are updated daily. With such timely data, *Learning for Life* team members can support students and families as issues emerge.

(continued)

For example, they can:

- quickly identify falls in attendance, understanding if multiple children in a family are not attending regularly and focusing on absences that are unexplained

- identify students struggling in literacy or numeracy, offering them access to The Smith Family programs focused on these areas

- reconnect with families who have to change schools, sometimes in times of crisis, ensuring support for the family continues

- advocate for families with their school, helping to establish a more positive relationship between them

- intervene early when a senior secondary student is at risk of disengaging from education

- identify and celebrate students' educational success.

Providing timely, targeted support has the potential to positively alter students' pathways through school for better overall educational outcomes. Currently, the dashboard provides data for over 4,700 *Learning for Life* students in South Australia. The potential for this to expand to around 10,000 in the state over the next 5 years as The Smith Family seeks to offer the program to more students.

The data-sharing arrangements were enabled by changes to the *Public Sector (Data Sharing Act) 2016* and Regulations that allowed non-government organisations to be included in such arrangements. Consent from the involved families was of central importance; the proposed arrangement was discussed in detail with *Learning for Life* families and 95% of them provided consent for 2-way sharing of their child's data between The Smith Family and the South Australian Department for Education.

This project demonstrates the power of data-sharing partnerships between government and non-government organisations for service-based improvements, and the benefits to students and families of live data for timely, targeted interventions.

Given the value already being seen by team members in South Australia from the data exchange, The Smith Family is developing similar arrangements with other jurisdictions, including Western Australia and Queensland.

See https://www.thesmithfamily.com.au/programs/learning-for-life for more information on *Learning for Life*.

7. Using single-touch payroll data to meet changing needs

Since the onset of the COVID-19 pandemic, the ABS has played an active role in compiling and releasing timely data to support informed decision-making. This has involved modernising existing data collection activities and using big data from both public and private providers.

An important example of this effort involves the use of single-touch payroll (STP) data secured from the Australian Taxation Office (ATO) (ABS 2022a). During the pandemic, the ABS and the ATO expedited work to use STP data to provide close to real-time information on changes in jobs and wages as restrictions to reduce the spread of COVID-19 (which had flow-on effects for income and work) were implemented.

The ABS used its evolving suite of Cloud-based information and ICT services to acquire, store and analyse STP data. Cloud-based ICT services are scalable to handle these big data sets, support modern programming languages and analytical techniques, and provide robust data protection controls.

The ABS received the first STP file, containing 351 million transactions, on 2 April 2020, and published the first 'Weekly payroll jobs and wages in Australia' release 19 days later. This release not only provided a national and state/territory picture of changes in jobs and wages during the pandemic – with breakdowns for sex, age group, industry, and employment size – but also included sub-state regional data (from July 2021).

The STP data released by the ABS have been invaluable in providing a timely understanding of the impact of the COVID-19 pandemic on businesses and households.

STP data are compiled into weekly indexes of payroll jobs and wages and published monthly (ABS 2022a). For more information, see https://www.transparency.gov.au/annual-reports/australian-bureau-statistics/reporting-year/2020-21-7 and https://www.transparency.gov.au/annual-reports/australian-bureau-statistics/reporting-year/2019-20-53.

References

ABS (Australian Bureau of Statistics) (2022a) *Single touch payroll (STP)* [online document], ABS, accessed 2 May 2023. https://www.abs.gov.au/statistics/detailed-methodology-information/concepts-sources-methods/labour-statistics-concepts-sources-and-methods/2021/methods-four-pillars-labour-statistics/administrative-data/single-touch-payroll-stp

—— (2022b) *Use and benefits* [online document], ABS, accessed 29 March 2023. https://www.abs.gov.au/about/data-services/data-integration/use-and-benefits

AIHW (2021) *Australia's welfare 2021: data insights* [PDF 10.5 MB] AIHW, accessed 27 March 2023. https://www.aihw.gov.au/reports/australias-welfare/australias-welfare-2021-data-insights/report-editions

—— (2022a) *Socioeconomic characteristics of ex-serving ADF members who died by suicide* [PDF 2.1 MB], AIHW, accessed 29 March 2023. https://www.aihw.gov.au/reports/veterans/socioeconomic-characteristics-of-ex-serving-adf/contents/about

—— (2022b) *Understanding the wellbeing characteristics of ex-serving ADF members* [PDF 1.1 MB] AIHW, accessed 29 March 2023. https://www.aihw.gov.au/reports/veterans/wellbeing-characteristics-of-ex-serving-adf/contents/about

—— (2022c) *Young people under youth justice supervision and their interaction with the child protection system* 2020–21 [PDF 787 KB], AIHW, accessed 29 March 2023. https://www.aihw.gov.au/reports/youth-justice/young-people-under-youth-justice-supervision/summary

Engle RF and Granger CWJ (1987) 'Cointegration and error correction: representation, estimation and testing', *Econometrica*, 55(2):251–276, doi:10.2307/1913236.

Gelman A (9 July 2015) 'Hey – guess what? There really is a hot hand!' [blog], *Statistical Modelling, Causal Inference, and Social Science*, accessed 29 March 2023. https://statmodeling.stat.columbia.edu/2015/07/09/hey-guess-what-there-really-is-a-hot-hand/

Gilovich T, Vallone R and Tversky A (1985) 'The hot hand in basketball: on the misperception of random sequences', *Cognitive Psychology*, 17(3):295–314, doi:10.1016/0010-0285(85)90010-6.

Gray M and Bray RJ (2019) *Appendix B: Evaluation in the Australian Public Service: current state of play, some issues and future directions, Australia and New Zealand School of Government (ANZSOG)*, Melbourne, accessed 29 March 2023. https://www.apsreview.gov.au/sites/default/files/resources/appendix-b-evaluation-aps.pdf

Productivity Commission (2020) *Mental health* [online document], Productivity Commission, accessed 29 March 2023. https://www.pc.gov.au/inquiries/completed/mental-health/report

Royal Commission (Royal Commission into Aged Care Quality and Safety) (2021) Final report, Parliament of Australia, accessed 26 June 2023. https://agedcare.royalcommission.gov.au/publications/final-report

The Smith Family (2022) *Annual report 2021–22*, Department of Finance, accessed 29 March 2023. https://www.finance.gov.au/publications/annual-report/annual-report-2021-22

von Sanden N (2020) 'Improving inter-agency data sharing through linkage spine interoperability', *International Journal of Population Data Science*, 5(5), doi:10.23889/ijpds.v5i5.1577.

Wolpoff MH (1975) 'Comment on "Hypothesis testing in palaeoanthropology" by D Pilbeam and JR Vaisnys', in Russell H and Tuttle RH (eds) *Palaeoanthropology: morphology and paleoecology*, Mouton, The Hague:15–18.