Preliminary summary of analytical findings

Emerging lessons from NDDA pilot

National
Disability
Data Asset

1   Emerging lessons from the NDDA Pilot

As part of the 18-month pilot phase of the National Disability Data Asset (NDDA), participating jurisdictions have committed to undertaking five priority data linkage test cases. This report consolidates lessons discussed to date on data sharing and enduring linkage, and presents preliminary assessments of expected analytical outcomes from the five test cases.

The test cases were designed to demonstrate three things:

1. That jurisdictions will share more data relevant to people with disability's outcomes than previously.

2. That lessons on process, technology, capability and other insights on sharing, linkage and analysis should shape the design and implementation of an enduring asset which supports future policy and research initiatives to improve outcomes for people with disability.

3. That linked data across jurisdictions could produce insights that inform and drive policy development and service improvement.

The Committee has previously discussed the first category of lessons of the test cases, with all data committed being shared. The Committee, and the Senior Executive Steering Committee, have also previously discussed the high-level lessons around lessons for an enduring data asset, with further detail provided below.

This paper also set out early findings on the third category. While the success of the Pilot was not designed to be assessed only on the basis of the analytic outputs generated by these test cases, partners have also worked hard to produce compelling outputs for ministers. The Committee has discussed that it will be important to demonstrate this practical value, ideally through showing insights and a link to how jurisdictions are considering using insights to improve their policy settings or service delivery. The ultimate value of the NDDA will come from establishing the infrastructure to enable access to large linked datasets in weeks, rather than years; however, and test case insights show what that will deliver in future.

Most test case teams have only just begun receiving their linked datasets, and more insights will be provided towards the end of September (as planned). The Early Childhood and Justice test cases are the most advanced, and so have more developed initial insights.

1.1   Consolidated lessons for enduring sharing and linkage

The below early insights from the test cases focus on the insights that will input into the enduring design of the NDDA.

Key Takeaways

The current project-by-project approach to data integration cannot deliver the responsiveness and scale of insights needed for data-driven decisions across government or enable people to navigate services effectively.

The pilot has shown that even with funding for dedicated resources and high-level government agreement to share, time to get approvals, access data and get linked data to analysts is a minimum of 12 months between two governments and involves 30+ process/artefacts that need to be redone for every linkage/purpose.

These delays are primarily the result of the time taken across all players to develop data sharing agreements (adhering to relevant legislation and policy) rather than the mechanics of data linkage itself.

Streamlined governance is what will differentiate the NDDA from previous linked assets and take Australia to the next generation of using its data to inform decisions.

Scaling up the production of insights will require development of agreed approaches for identifying various groups of people with disability within data, and establishing reliable, high-quality indicators alongside pre-linked data. Indicator and cohort identification capability will not develop sufficiently through a project-by-project approach for high-quality insights.

The shift from the current research-focused approach (as used in the test cases) to a government and other platform use case focused asset will require investment. A roadmap for how this can be accomplished, including an outline of supporting governance arrangements, will be delivered at the completion of the pilot.

1.2 Notable milestones

1. A first-ever extension of national linkage mapping across the Commonwealth and States to an unprecedented number of datasets has been accomplished, with the capacity to identify people with disability over multiple years.

2. Barriers to national data sharing and collaboration have been overcome: 51 national and state/territory datasets have been linked across the five test cases. These are being securely accessed and analysed by researchers in partner agencies. More data than originally committed by jurisdictions has been provided and delays have not resulted from any unwillingness to share.

3. For the first time, insights will be able to be generated on mainstream service use by people with disability across the education, employment, health, justice and housing sectors.

1.3 Testing design options to inform the enduring asset

The test cases are assessing different and novel approaches to data integration, harmonisation and access to inform the design of the enduring asset.

1.3.1 Approaches to data integration

Key Learnings

The pilot facilitated unprecedented levels of collaboration and cooperation across custodians, linkage units and data users, with all three approaches resulting in high linkage rates (that is, proportions of successfully matched individuals in different data collections, of 95% or higher). Although spine-to-spine linkage was more efficient, it did not necessarily perform better than other approaches. An approach that maps all data to a single spine will produce the greatest gains.

Good quality metadata is a critical input at various stages of the data integration process including approval, intake, linkage, and use, and is a dependency for automation and scalability. The current availability and quality of metadata varies widely across custodians and substantial investments will be required to establish a standard capable of supporting the NDDA.

The pilot has demonstrated the benefits of leveraging existing data integration infrastructure at the State/Territory and Commonwealth (AIHW/ABS) levels. By September, these lessons will inform a proposed data integration model for the NDDA that significantly advances Australia's national data linkage system.

Three different approaches to data integration were adopted across the five test cases:

- all data linked by AIHW

- multi-stage linkage leveraging existing State-level mappings; and

- multi-stage leveraging the ABS-AIHW interoperable spine.

This has allowed the development of an enduring linkage approach which builds on existing capabilities (such as the use of interoperable spines to activate the national network of data linkage units) and is more likely to be agreed by jurisdictions because it does not involve 'handing over' all their data to a central entity. This would be a first for Australia.

In the first approach, all identifiers or linkage keys (as applicable) were provided by custodians to AIHW for mapping to the NDDA Linkage Map. This process is closely aligned with historical approaches to data linkage projects in Australia.

Two of the test cases trialled a multi-stage linkage process that leveraged existing mappings of datasets within State Data Linkage Units (DLU). In this approach, a local population spine was derived by the DLU and subsequently mapped to the NDDA population spine by AIHW. This approach maximised existing linkage infrastructure and data sharing agreements within jurisdictions, and increased privacy by reducing the amount of identified data transferred to AIHW.

In a separate approach, two test cases assessed a multi-stage linkage process that leveraged the AIHW-ABS interoperable spine. Here, an initial mapping and cohort definition was produced by AIHW and the identifiers were then used by ABS to provide content data corresponding to the defined cohort. As above, this approach minimised the amount of identified data transferred to linkage authorities and benefited from existing data agreements and infrastructure.

All linkages were conducted in secure environments with processes that minimise disclosure risks, based on ethics approvals, data sharing agreements and complying with the Five Safes[1] framework.

Linkage methods adopted for the test cases included probabilistic and deterministic methods, which are based on linking records using full personal identifiers such as names, addresses and dates of birth. Whereas deterministic (or exact) linking uses a unique identifier to link datasets, probabilistic linking uses a number of identifiers, in combination, to identify and evaluate links. Statistical Linkage Keys (SLK) were also used in particular cases, which is a de-identified key that is highly likely to be unique and is based on the letters in a person's name, date of birth and sex. Separation principles were adhered to for each linkage process, which ensures that identifiers remain separated from their corresponding content data.

1.3.2   Analytic environment

> **Key Learnings**
>
> The test cases have highlighted the effort required to set up the AIHW SRAE and ABS DataLab to support the analysis and the preparation and training needed to prepare analysts to undertake analysis of these complex datasets.

---

[1] Desai, T., Ritchie, F., & Welpton, R. (2016). Five Safes: Designing data access for research.

> Future environments will need to appropriately configure the computing facilities to support the demands of the NDDA. This is an important insight into the ICT investment needed for the national impact intended through scaled sharing.

Both Commonwealth partner linkage agencies are enabling secure access to the test case data through their remote analytic environments. Four test cases are piloting the AIHW's Secure Remote Access Environment (SRAE), and one test case is leveraging the ABS DataLab environment. Notably, one of the test cases will test both analytic environments, apportioning the analytic work across the ABS DataLab and AIHW SRAE.

The volume and complexity of data being analysed by each of the test cases has posed unanticipated challenges related to computing capabilities within the environments, which has limited the analytic approaches that can be applied within each test case. Limited information available to the analysts about each dataset and familiarity with the analytics environments made it difficult for them to plan their analysis. Similarly, establishing the environments to meet the analyst's requirements was part of the try, test, learn approach.

Although the teams have worked collaboratively with the AIHW and ABS to find workarounds and progress the analyses within the constraints of the environments, this has highlighted the need to design future systems that can cope with large amounts of data and computationally-intensive analysis.

1.3.3   Disability indicators

> **Key Learnings**
>
> The test cases have highlighted the need for ongoing methodological development and collaboration to develop disability indicators of relevance to the disability community. A suitable platform to share learnings and promote comparability will be needed, as will continued engagement with the disability community via the new governance arrangements.

An emerging lesson of the test cases, and the broader body of disability research using linked data, is that coordinated effort on data improvement and development is likely to be a key value of a national data asset.

There is no single marker of disability that exists across administrative datasets and such a marker is not feasible given the community rejection of a single definition of disability (as well as practical constraints). Disability can be inferred through various means such as provision of disability services, clinical diagnoses, identification of needs in education or other service systems, self report, or other such measures. Each source of information has strengths and limitations. Bringing these disparate pieces of data together provides a more comprehensive

and reliable picture of disability than any source on its own; however, this must be done carefully in order to account for underlying differences in data capture, collection purposes, and temporal relationships.

Each test case in the NDDA pilot is adopting a different approach to identify people with disability. This is because each test case is looking at a different population of interest, and therefore including only the relevant data required to address the defined policy questions. This provides a good opportunity to assess the suitability of various approaches and combinations of data to identify disability (including, where relevant, disability type or severity) and uncover the most appropriate methodologies.

The current "gold standard" for measuring disability prevalence in Australia is the ABS' Survey of Disability, Ageing and Carers (SDAC), which aligns closely with the WHO International Classification of Functioning, Disability and Health (ICF). Whereas four of the five test cases will assess the quality of their disability indicators against SDAC population estimates, the Outcomes test case will directly compare its derived indicators against individual SDAC records. This will provide a more robust assessment of the appropriateness and potential limitations of identifying people with disability in administrative data.

These learnings will pave the way for the enduring NDDA to establish a core of high quality, reliable disability indicators. A platform (or 'wiki') approach will be needed nationally to share capabilities in indicators and analytical methods for deriving indictors, to ensure that individual silos in governments and research institutions are not reinventing the wheel within each project. Contribution to this shared capability should be made a condition of accessing linked data in the enduring design so there is reciprocity and continual improvement built in.

2 Early analytical insights from the Test Cases

2.1 Early Childhood Supports in NSW

> **Key Learnings**
>
> For the first time, we can explore supports through mainstream and targeted services across state and national health, education and human services systems that promote optimal educational outcomes.

For the first time, this will enable answers to questions like:

1. What early childhood supports for developmental delay and disability are accessed by children in NSW?

2. What is the relationship between the early childhood supports for children with disability and developmental delay, participation in early childhood education, and a range of developmental and educational outcomes for those children as they progress through their schooling?

3. Outcomes will include school participation (enrolment, attendance and exclusion), educational attainment (literacy and numeracy assessment at various stages), and development.

4. What data gaps and further sources of data are required to better assess the impact of services/supports for children with developmental delay and disability in the future?

Highlights and accomplishments:

- This test case trialled a novel linkage approach which resulted in high linkage rates and has informed the design of the enduring asset.

- Using a NSW child population approach, we are able to identify the data of children who received a service for disability or developmental delay prior to school, and also the children who did not receive a service, but may have benefitted from it. This enables us to understand where and how services need to be directed to reach all children.

- Service information is available from 23 NSW and National datasets providing information on health, Early Childhood Interventions, education and social services.

- Around a third of children in NSW attend a government preschool. This is the first time the service journey of children who did not go through this system can be understood in terms of trajectories to school educational outcomes.

- This test case goes beyond service use for children with disability to those with developmental vulnerability who may or may not have been identified and supported prior to school.

- The dataset includes information on 2.35 million children born in the years between 2003-2019. These children were either born in NSW, received certain health or education services in NSW, or had a NSW address in the period 2003-2019.

  - Around 5.5% of the cohort were identified in the Disability Services National Minimum Dataset (NMDS) and/or the National Disability Insurance Scheme (NDIS) dataset.

  - Approximately 7.2% of the children represented in the Australian Early Development Census (AEDC) were identified as of interest to this test case, with either 'special needs', 'requiring further assessment, or 'developmentally vulnerable' on each domain of the AEDC.

- The results of this test case will be incorporated into a data visualisation tool to inform needs-based service planning and developing outcomes frameworks.

Next steps for this test case:

- Quantify cross agency service use by NSW children with disability and/ or developmental delay prior to school.

- Characterise the children who have accessed services prior to school.

- Explore the relationships between service use prior to school and educational and developmental outcomes at school. This includes the potential to explore needs and service mapping methods in a practical context.

- Track journeys through service systems to various outcomes for particular cohorts. This will illustrate the power of person-centred cross-system linked data to optimise service delivery for better outcomes.

- The NDDA digital platforms workstream will partner with this test case to develop the data visualisation tool to interrogate aggregated data to a far greater extent than can be reported from this test case. This will ensure that this valuable dataset can be repeatedly and rapidly used to inform policy and program design.

2.2 Interaction of People with Disability and the Justice System: a Commonwealth / NSW Linkage Study

Key Achievement

First-ever opportunity to examine the victimisation and offending rates of a cohort of individuals with intellectual, physical, sensory and psychiatric disabilities.
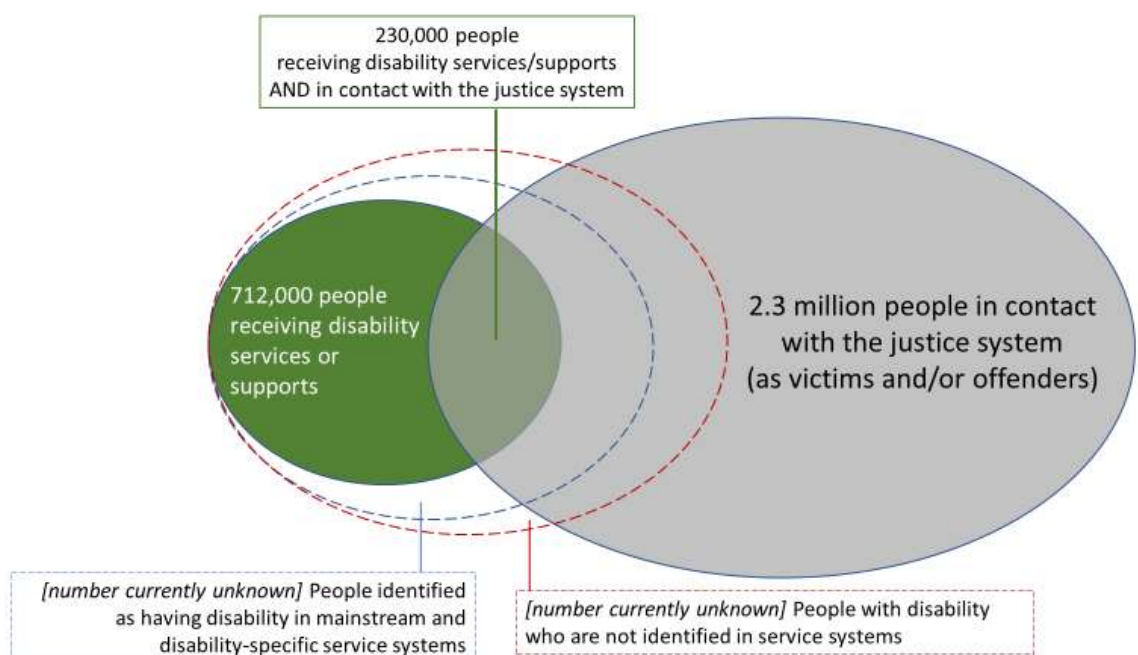
For the first time, this will enable us to explore:

1. How often victims identified in the criminal justice system are people with disability.

2. Factors associated with an increased risk of contact with the criminal justice system as victims or offenders, among people with disability.

3. How people with disability interact with the criminal justice system including whether their outcomes differ from the outcomes of those who are not identified as having a disability.

4. The use of diversionary options for offenders with mental health, psychosocial and /or cognitive disabilities.

5. How disability, health and other services are accessed by people with disability before, during and after contact with the criminal justice system.

Highlights and accomplishments:

- Over 99 million records linked across 11 State and Commonwealth datasets, with linkage rates of 78-99%, and an average linkage rate of 95%.

- The study cohort contains information on 2.8 million individuals aged 10 years and over who had contact with the NSW criminal justice system or received disability services or supports between 2009 and 2018. This includes 2.3 million individuals who had contact with the NSW criminal justice system, and 712,000 individuals who had contact with disability services and supports.

- Of the individuals who had contact with disability services:

  o 72% (510,000) had received the Disability Support Pension.

  o 38% (273,000) had contact with disability services captured by the Disability Services National Minimum Dataset.

  o 27% (189,000) had contact with the National Disability Insurance Scheme.

- Approximately 1 in 3 individuals with disability services contact had at least one criminal justice system contact, either as an offender or a victim.

- Approximately 1 in 10 individuals with at least one criminal justice system contact (as an offender or victim) had contact with disability services.

230,000 people
receiving disability services/supports
AND in contact with the justice system

712,000 people
receiving disability
services or
supports

2.3 million people in contact
with the justice system
(as victims and/or offenders)

[number currently unknown] People identified as having disability in mainstream and disability-specific service systems

[number currently unknown] People with disability who are not identified in service systems

- Data remediation work was undertaken by the NSW Bureau of Crime Statistics and Research (BOCSAR) to enable victimisation data to be used in this test case. Analysis of the linkage of victim records shows:

    o the within-dataset linkage is of a high quality, and similar to that of the NSW BOCSAR Re-offending Database.

    o a strong external linkage rate of 90% with the Medicare Enrolment File, a 15% higher linkage rate than the dataset with the lowest linkage rate in the sample.

- Data for inmates accessing Statewide Disability Services run by NSW Corrective Services were linked to the NSW Re-offending database, enabling broader identification of people with disability and the analysis of service usage by people with disability in correctional facilities.

- Methods to identify disability across a range of service system data, beyond disability services, have been developed. This will result in a more comprehensive disability indicator and help to identify service delivery gaps.

- This test case has an embedded Aboriginal Perspectives Advisory panel. The panel works with the researchers and analysts to shape the research, support interpreting the data and results, identify data gaps and help with the communication of these sensitive results, where there is an over-representation of First Nations people.

- The Panel has provided insights into data gaps in this test case, as well as mechanisms which lead to gaps in accessing disability support services for First Nations people in contact with the justice system. Selected quotes from Panel members are highlighted below:

" For good reasons many First Nations people **won't disclose** they have a disability. It may not be to your benefit to tell anybody too much about yourself. It might make you **more vulnerable.**"

""Very common for a child to be noted in their education data as having special needs because **school is a place where it becomes very obvious**, even if they have not been identified anywhere else and may not continue to be."

Next steps for this test case:

- Apply the cross-service system disability indicators and compare these indicators with the identification of disability through disability-specific support services.

- Examine the outcomes of charges involving people with disability as victims and/or offenders.

- Examine the service usage pathways experienced by people with disability before, during and after their interaction with the justice system.

- Identify further data sources required to better assess the impact of services/supports and associated risk factors for people with disability who have interactions with the NSW criminal justice system.

2.3 Identification of People with Disability in Linked Administrative Data for Service Use and Outcomes Reporting

Key achievement

First-ever holistic examination of services accessed by people with disability across multiple service systems (Commonwealth, NDIS, and States).

This test case has two overarching aims:

- Testing whether existing administrative data collections can support creation of a comprehensive indicator for people with disability, including those within and outside of the scope of the NDIS and DSP.

- Assessment of the accuracy of the derived disability indicator and whether it can be used to reliably report on NDS/NDIS outcomes for people with disability, as well as services and supports accessed by people with disability.

Housing was chosen as a test system for this project. Stable, accessible and affordable housing is fundamental to improving outcomes for people with disability. It is currently not possible to fully measure the degree of interaction people with disability have with the housing system:

- Housing supports are provided via multiple systems (e.g. Commonwealth, state and territory, NDIS). Each system uses its own definition of disability, whereas some systems do not collect disability data at all.

- People with disability may use more than one type of housing support. However, as reporting is system-centric, it is not known how people with disability interact with multiple systems or move between systems.

- It is not known how many people with disability do not access housing supports.

Thus, for the first time, this test case will be able to answer questions like:

1. Can cross-system data matching provide an accurate and complete measure to identify people with disability? Can sub-groups of people with disability be comprehensively defined on the basis of information available in the linked datasets? Which cohorts of people with disability have reduced/missing coverage in administrative data?

2. What proportion of people with disability receive housing-related supports? How does contact with housing supports differ for people with disability compared to those without disability? How does contact with housing supports vary across sub-cohorts of people with disability? What trends (and event shifts) were observed during 2010-2020?

3. How do housing systems perform for people with disability, by sub-cohort, and relative to people without disability?

4. How suitable is the NDDA as a data source to report against the new National Disability Strategy Outcomes Framework, and inform the development of its indicators and measures?

Highlights and accomplishments:

- This test case is the only one of the NDDA Pilot test cases to analyse data across multiple states. It uses data from Commonwealth data sources and the NDIS, as well as data from all states participating in the NDDA Pilot (by using components of the state-owned and nationally-collated data in AIHW collections) and will provide key insights on how a future coordinated national linkage system could work.

- The test case does not have an age restriction and will attempt to identify people with disability in all age groups.

- The linked dataset for this test case contains information on 15.7 million people across seven data sources over years 2010-2020. The linked records include:

   o Around 5.7 million people identified as part of the *initial* disability cohort.

   o About 10 million people who received housing and homelessness services and were not identified as having a disability.

- The initial disability cohort selection criteria for this test case were kept intentionally broad and included receipt of disability-related services through NDIS, MBS, or under NDA: an array of social security payments related to disability, sickness or medical conditions, and restricted capacity to work. For practical reasons, the final disability cohort will likely be refined to exclude people who only experienced transient medical conditions.

- Linkage rates were very high, ranging from 96.0% to 99.4% for name-based linkage, and from 88.5% to 97.5% for statistical linkage key (SLK)-based linkage.

- Preliminary disability indicators have been developed (including indicators of disability type and severity), drawing on information from across the linked data sets, for testing against individual SDAC and Census records. Initial comparisons are underway to assess the size and demographic profile of the preliminary disability cohort against published SDAC data.

- Data quality activities have been performed on the source data. These will inform further development and use of the derived indicators, as well as contributing to the NDDA Disability Data Improvement Plan.

Next steps for this test case:

- Finalise derivation of disability indicators from administrative data.

- Analyse housing supports and services received by people with disability (including public housing, homelessness services, Commonwealth Rent Assistance, and NDIS housing-related supports).

- Assess quality and performance of the derived disability indicator (based on linked administrative data) against SDAC 2018 and Census 2016 data. Summarise learnings and outline final approach in a methodology paper.

2.4 Services and Supports used by People Living with Disability and Mental Health Issues in Victoria

Key Achievement

First holistic picture of the disability and mental health system in Victoria than has otherwise been possible from state-based data alone

For the first time, this test case will generate insights to answer the following types of policy questions:

1. Who are the cohorts of people living with disability and mental health issues (i.e. psychosocial disability only, psychosocial and other co-existing disability, disabilities with mental issues) and their characteristics?

2. What supports are currently regularly accessed by the different cohorts of people living with disability (including psychosocial disability) and mental health issues?

3. What service gaps exist for people with a disability (including psychosocial disability) with mental health issues?

4. Which supports are effective in improving health outcomes for people living with disability (including psychosocial disability) and mental health issues?

5. What data gaps and further sources of data are required to better assess whether the needs of people living with disability and mental health issues are being met?

Highlights and accomplishments:

- This is the only test case where an existing population-wide linkage map of an entire state (Victorian Linkage Map [VLM]) was linked to the entire AIHW National Linkage Map (NLM).

- This test case brought together more than 164 million records across Victorian and Commonwealth datasets. Linkage rates ranged from 79% to 98%. This was slightly lower than linkage rates in other test cases and was due to varying quality in the identifiers contained in the Victorian Linkage Map.

- 45% of linked individuals for this test case were associated with more than one anonymous identifier in the Victorian Linkage Map. In future, this type of linkage approach could be used to improve the quality of jurisdiction-specific linkage maps.

- The spine-to-spine linkage approach was cost-effective and flexible. Future linkages could be refreshed routinely and re-used for multiple collaborative projects (with relevant approvals in place and pending improvements in the underlying quality of identifiers), increasing scalability of data linkage and access.

- The linked dataset for this test case includes information on around 610,000 Victorians who received disability supports between 1 July 2008 and 31 December 2019.

  - Information is available on Victorian community care, mental health, hospital, emergency and outpatient services, as well as NDIS and disability support services.

  - For the first time, we will be able to understand how to best meet the needs of Victorians with disabilities and mental health issues.

- The results of this test case will provide baseline data to inform service delivery, planning and reporting on mental health and disability care in Victoria. In particular, these insights will inform key policy initiatives:

  o Victoria's 10-year Mental Health Plan – 'Victorians have good mental health and wellbeing' Focus Area in particular.

  o Victorian autism plan – Increase understanding of health and wellbeing needs, including mental health needs, for autistic Victorians.

  o The new Victorian State Disability Plan 2021-25 – Health domain: people with a disability achieve their optimal mental and physical health; and the Victorian Government's commitment to optimising the opportunities the NDIS presents for eligible consumers of specialist clinical mental health services.

  o Royal Commission into Victoria's Mental Health System Report recommendations, including on investment into developing improved collaborative research and knowledge sharing capacity to support mental health services.

Next steps for this test case:

- Data cleaning and preparation of key Commonwealth data holdings.

- Refinement and application of the methodology for generating indicators of disability across Victorian and Commonwealth service data.

- Analysis of mental health supports and services and the interaction with disability supports, with a particular focus on identifying gaps that indicate potential unmet need.

- Exploring options to trial machine learning methodologies to conduct the analyses in a try, test, learn approach.

2.5 Education to Employment Pathways for People with Disability: a SA Linkage Study

Key Achievement

First-ever examination of pathways into employment or further education for young adults with disabilities, through school education, senior secondary education, vocational education and training participation.

For the first time, this test case will generate insights to answer the following types of policy questions:

1. What are the educational and employment outcomes for young adults with disability? How do the outcomes for people with disability differ from those who do not have a disability?

2. What are the educational pathways for young adults with disability and which pathways and supports are most effective in achieving positive employment outcomes? How do the educational to employment pathways for people with disability differ from those who do not have a disability and how do these affect future employment?

3. What data gaps and further sources of data are required to better assess impact of services/supports and associated educational and employment outcomes for people with disability?

Highlights and accomplishments to date:

- This test case has leveraged existing AIHW – ABW interoperability arrangements:

    o AIHW has undertaken the linkage component, mapping identifiers from South Australia (SA) Education Enrolments, SA Certificate of Education, NDIS, disability services data (DS NMDS) and disability support payments (DOMINO) to the AIHW National Linkage Map, and creating the study cohort.

    o ABS is assembling all of the content data to be used in analysis (including MADIP data) and hosting the resulting linked dataset in the ABS DataLab.

- Linkage rates for the AIHW component of this test case ranged between 93.9% and 97.4% and include a combination of statistical linkage key (SLK) and full identifier linkages.

- The study cohort consists of around 190,000 individuals as captured by Year 10 (or equivalent) SA public school enrolment between 2005 and 2019.

- Indicators of disability will be available through a range of data included in the linkage including through interactions with the NDIS or receiving services under the National Disability Agreement as captured in the Disability Services National Minimum Dataset. Around 14,500 individuals have been identified within the study cohort under the NDIS and DS NMDS.

- Disability related information will be available in other data included in test cases including SA Education, VET and DOMINO but we do not yet have access to the

relevant data to identify the number individuals in the study cohort captured through these sources.

- A strong collaboration and commitment has been established between agencies who do not traditionally work together. Feedback from education stakeholders has highlighted the benefits of being able to "break down silos".

- The data agreement and transfer process has uncovered areas for improvement regarding data gaps, data quality and agreements to share data in a time-effective manner.

- The results of this test case will provide a deeper understanding of the extent to which people with disability interact with the education and employment systems and the impact of services and supports on their outcomes. This will enable SA government agencies and service providers to better plan and provide supports and services and, for governments, enable evidence-informed policy development. In particular, these insights will inform existing initiatives, including:

  o Inclusive SA 2019–2023 (South Australia's first disability inclusion plan). This sets out whole-of-government actions and measures how State Government agencies and local councils will implement and improve access and inclusion for South Australians living with disability.

  o The development of the next Inclusive SA (post-2023).

  o Implementation of Australia's Disability Strategy 2021-2031.

Next steps for this test case:

- Data cleaning and preparation of disability indicators, including sharing the final methodology used in this test case to inform the enduring asset.

- Establishment of pathways into employment or further education for young adults with disability, through school education, senior secondary education, vocational education and training participation.

- Data quality assessments (and, where feasible, remediation) to inform the Disability Data Improvement Plan.

- Analysis of outcomes for school leavers with disability, and comparison to young people without disability.

APPENDICES

APPENDIX 1: Data sources for the NDDA pilot

Source	Dataset	Test cases				
		1	2	3	4	5
Commonwealth	Medicare Consumer Directory (MCD)	■	■	■	■	■
	National Disability Insurance Scheme (NDIS)	■	■	■	■	■
	Medicare Benefits Scheme (MBS)	■	■	■	■	
	Data Over Multiple Individual Occurrences (DOMINO)	■	■	■		■
	Child Care Subsidy (CCS)/Child Care Benefit (CCB)	■				
	Australian Early Development Census (AEDC)	■				
	Pharmaceutical Benefits Scheme (PBS)		■	■		
MADIP	Personal Income Tax (PIT) – ATO					■
	Higher Education Information Management System (HEIMS)					■
	Vocational Education and Training (VET)					■
	Survey of Disability, Ageing and Carers (SDAC)			■		
	Census		■			
National (AIHW)	Disability Services National Minimum Dataset (DS NMDS) and Commonwealth State/Territory Disability Agreement National Minimum Data Set (CSTDA NMDS)	■	■	■	■	■
	National Death Index (NDI)	■	■	■	■	■
	Specialist Homelessness Services Collection (SHSC)		■			
	National non-admitted patient emergency department care database (NNAPEDCD)		■			
	National Hospital Morbidity Database (NHMD)		■			
	Public Housing and State Owned and Managed Indigenous Housing (PH & SOMIH) Data Set			■		
SA	SA School Enrolment Census					■
	SA Certificate of Education					■
VIC	Victorian Linkage Map (VLM 1906)				■	
	Victorian Integrated Data Resource (IDR), derivations and tabulations				■	
	Clinical public mental health services (CMI/ODS)				■	
	Mental health Community Services (MHCSS)				■	
	Victorian Admitted Episode Database (VAED)				■	
	Victorian Emergency Minimum Dataset (VEMD)				■	
	Victorian Integrated Non-Admitted Health (VINAH)				■	
	Community Health Minimum Dataset (CHMDS)				■	
	Home and Community Care (HACC)				■	

Test cases:
Test Case 1: Early Childhood (NSW)
Test Case 2: Justice (NSW, Cth DSS)
Test Case 3: Outcomes data (Cth DSS)
Test Case 4: Mental Health and Psychosocial (VIC)
Test Case 5: Education to Employment (SA)

Source	Dataset	Test cases				
		1	2	3	4	5
NSW	NSW-AIHW Linkage Map	■				
	NSW Education disability	■				
	NSW school enrolment	■				
	NSW community preschool census data	■				
	NSW Government Preschools and Early Intervention Census	■				
	NSW Preschool Disability Support Program (PDSP) – administered by Northcott	■				
	NSW Disability and Inclusion Program (DIP) - Higher Learning Support Needs (HLSN)	■				
	NSW Best Start	■				
	Nationally Consistent Collection of Data on School Students with Disability (NCCD)	■				
	Suspension data	■				
	Attendance data	■				
	NAPLAN data	■				
	NSW Birth Registration Data Collection	■				
	NSW Deaths Registration Data Collection	■				
	NSW Perinatal Data Collection (PDC)	■				
	NSW Emergency Department Data Collection (EDDC)	■				
	NSW Admitted Patient Data Collection (APDC)	■				
	NSW ChildStory	■				
	NSW Re-offending Database (ROD)		■			
	NSW Police Victims' records		■			
	NSW Social housing data		■			
	NSW Child Protection		■			
	NSW Out of home care (OOHC)		■			