Title Page

Data Linkage: Canadian and Australian Perspectives on a Valuable Methodology for Intellectual and Developmental Disability Research

Robert Balogh,1 Helen Leonard,2 Jenny Bourke,2 Kate Brameld,3 Jenny Downs,2 Michelle Hansen,2 Emma Glasson,2 Elizabeth Lin,4 Meghann Lloyd,1 Yona Lunsky,5 Melissa O'Donnell,2 Shahin Shooshtari6, Kingsley Wong,2 Gloria Krahn7

Author Affiliations:

1. Faculty of Health Sciences, Ontario Tech University, Oshawa, Ontario, L1H 7K4, Canada
2. Telethon Kids Institute, Centre for Child Health Research, The University of Western Australia, PO Box 855, West Perth, Western Australia 6872, Australia
3. Faculty of Health Sciences, Curtin University, Perth, Western Australia 6102, Australia
4. Evaluation and Data Management, Centre for Addiction and Mental Health, Toronto, Ontario, M5T 1R8, Canada
5. Azrieli Adult Neurodevelopmental Centre, Centre for Addiction and Mental Health, Toronto, Ontario, M6J 1H4, Canada
6. Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada, R3T 2N2
7. College of Public Health and Human Sciences, Oregon State University, Corvallis, Oregon, 97331, USA

Corresponding Author:
Robert Balogh, PhD, Associate Professor
Ontario Tech University
Faculty of Health Sciences
Oshawa, Ontario, Canada, L1H 7K4
Robert.balogh@uoit.ca

Acknowledgements:

The authors gratefully acknowledge staff at the WA Data Linkage Branch, Department of Health and the data custodians, including the Western Australian Register of Developmental Anomalies, for their assistance in obtaining the linked data used in this
study. The authors also acknowledge the Department of Communities, WA (previously the Disability Services Commission), the WA Department of Education, the Catholic Education Office, and the Association of Independent Schools of WA for assistance with data collection for the IDEA database. We acknowledge provision of data through (NHMRC) Program Grant #572742.

This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). This study also received funding from Special Olympics Canada. Parts of this material are based on data and information compiled and provided by the MOHLTC, Canadian Institute for Health Information. The analyses, conclusions, opinions and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

HL is supported by an Australian National Health and Medical Research Council Senior Research Fellowship (APP1117105)
RB and ML received funding from Special Olympics Canada

The authors wish to thank Marni Brownell who reviewed sections of the manuscript and James Noronha who helped prepare the data from Special Olympics Ontario.

Keywords: intellectual and developmental disabilities, data linkage, health surveillance
Health and social service researchers have long recognized the potential for new knowledge that could be derived from integrating information across multiple existing administrative datasets (e.g., Dunn, 1946). Administrative data originate with records generated from the administration of programs, are not originally intended for research purposes, and usually come from government sources such as health, education, and social services (Connelly, Playford, Gayle & Dibben, 2016). Data linkage using administrative data refers to the joining together of electronic records that belong to the same person or event. Yet developing the infrastructure to access this data for research purposes was initially accomplished in only a few locations. In 2008, Holman and colleagues (Holman, Bass, Rosman, Smith, Semmens, Glasson,…Stanley, 2008) listed only a handful of information-rich environments in the world where the necessary infrastructure was developed to allow for efficient and routine data-linking or records-linking. In the past decade, we have experienced an explosion in digital data on health in many countries. In the U.S., this has been prompted in part by the incentives and requirements of the Affordable Care Act of 2010 (see US DHHS, n.d.). During this same time, Big Data has captured the imagination of numerous innovators across multiple fields, including public health (e.g., Khoury & Ioannidis, 2014; National Academies of Sciences, Engineering and Medicine, 2013) and health care (e.g., Murdoch & Detsky, 2013).

As policymakers and researchers recognize the major contribution that social factors make to physical and mental health, there is an obvious desire to connect social variables with physical health data to investigate, understand and improve health outcomes. This “determinants of health” approach holds promise for expanding knowledge in a number of directions. Policy debates focus on whether public investments in housing or family supports can substantively improve health and reduce health care costs. Linked data across social, educational and medical
services has the potential to identify the relative value of services individually or collectively that predict better health outcomes. At an intervention level, linked data might better identify specific subgroups that experience health disparities and who are at particularly high risk for poor health and early mortality.

For disability researchers, the infrastructure for ongoing data linkages is highly promising given the array of service systems involved. To date, however, other than Canada and Western Australia, there are limited sites where the necessary infrastructure and data linkage capabilities have been applied to intellectual disability research. Some exceptions include research using linked educational data in Florida (Chapman, Scott, & Mason, 2002; Chapman, Scott, & Stanton-Chapman, 2008); Medicaid, educational and mental health data with disability in South Carolina (McDermott, Royer, Mann & Armour, 2018; Royer, Hardin, McDermott, Ouyang, Mann, Ozturk & Bolen, 2014); recent research in the United Kingdom (Cooper, Mackay, Wood, King, Clark, Smith, & Pell, 2016; Hatton, Emerson, Robertson, & Baines, 2018; Mackay, Smith, Cooper, Wood, King, Clark, & Pell, 2016; Mackay, Smith, Dobbie, Cooper, & Pell, 2013); and new efforts emerging in Ohio (Bonardi, 2019, this issue).

The purpose of the present paper is to provide examples from two well-established data environments that facilitate data linkage for population research on people with intellectual and developmental disabilities (IDD)—from Canada, the Manitoba Centre for Health Policy and related work in the neighboring province of Ontario at ICES; and from Australia, the Western Australia Data Linkage System. In both examples, source records from databases identify persons with intellectual disability (ID) who may have other developmental disabilities as well (e.g. cerebral palsy); but, neither would include people with a diagnosis of cerebral palsy who do not also have ID. Jurisdictions and studies do however vary in how terms are defined, and in
how administrative data and codes are used to identify their study group of interest. For example in Canada, a number of researchers have opted to include records of people with autism spectrum disorder regardless of the presence or absence of a co-occurring ID diagnosis. In this paper, the term IDD is used to refer to people with ID and/or developmental disabilities. When the term ID is used, sample members specifically had ID and this possibly co-occurred with other conditions.

The examples from Canada and Australia include descriptions of the databases and the linkage capabilities, how the linked databases are used to identify persons with IDD, and how the environments where the databases are kept have been leveraged to conduct population level research about persons with IDD. We conclude by describing some limitations and challenges of using linked administrative databases for research purposes and some commonalities between the approaches taken in the two countries.

**Data Linkage in Canada**

**The Canadian Context for Health Care**

The Canadian political and health care system is well suited to health research using linked administrative data. Canadians use the term ‘Medicare’ to describe its publically funded health care system. Under Medicare, Canadians have access to medically necessary health services without having to pay out of pocket (Government of Canada, 2018). However, Medicare is not a national plan. Rather, Canada’s 10 provincial and 3 territorial governments hold the authority to manage, organize, and deliver these services (Marchildon, 2013). Only approximately 25% of health services delivered to Canadians are funded by the federal government, and this occurs in the form of the ‘Canada Health Transfer’ (Government of Canada, 2019). In order for full transfer of these funds to occur, the provincial and territorial
governments must adhere to the principles set out in the Canada Health Act (Government of Canada, 2019). These principles are: (a) public administration, (b) universality, (c) comprehensiveness, (d) accessibility, and (e) portability. Public administration requires that the provincial and territorial health insurance plans be administered and operated on a non-profit basis by a public authority. In practice, this means that each province or territory has only one health insurance plan that covers all medically necessary hospital and physician services paid for through taxes. Universality requires that all ‘residents of a province or territory be entitled to the publically funded health services covered by the provincial/territorial plans’ (Standing Senate Committee on Social Affairs Science and Technology, 2002). Together, these two principles have led each province and territory to collect longitudinal data on the utilization of hospital-based and physician-provided services (i.e. those covered by the Canada Health Act) by their residents for administrative purposes. Under specified conditions these data are also available to conduct research.

**Partners and Data Sets of the Manitoba Centre for Health Policy**

The Manitoba Centre for Health Policy (MCHP) is a university-based research center founded in 1991 that conducts population-based research on the health and social determinants of health of Manitobans (Marchessault, 2011). MCHP is recognized internationally as a center of excellence for its leadership and innovations in the area of large administrative database linkage (Holman, Bass, Rouse, & Hobbs, 1999; Lewis, 2011). The pioneering work of Noralou and Les Roos in the area of linked databases led to the creation of MCHP. It was supported at the time by the perception of persons in government that putting some “science behind some of the decisions we are making” was of value (Fransoo, Marchessault, Black, & Decoster, 2011). Almost half of
the MCHP’s funding comes from the provincial government and the other half from research funding institutes (Manitoba Centre for Health Policy, 2019a).

The initial focus of MCHP was to conduct research on the patterns of illness and health care service use of Manitobans. Since then, MCHP has expanded to include research on social issues such as educational achievement (Marchessault, 2011; Roos, Roos, Brownell, & Fuller, 2010). The typical use of MCHP’s research is to inform health and social policy and to monitor policy performance (Manitoba Centre for Health Policy, 2019a). It accomplishes this by accessing and analyzing a number of linkable databases made available through the Manitoba Population Research Data Repository (henceforth the Repository). Originally, the Repository primarily consisted of linkable health databases, but it has grown over the years to include a comprehensive set of administrative, registry, survey and other data from education, social services and the justice system (see Figure 1) (Marchessault, 2011).

A key component of the Repository is the MCHP Registry which contains information on demographic, residential and provincial health insurance eligibility on virtually all Manitobans (2016 pop. 1.3 million) from 1970 to the present (Roos & Nicol, 1999; Statistics Canada, 2019). The MCHP Registry accomplishes this through regular data updates from the Government of Manitoba’s Health Insurance Registry; the completeness of the data at MCHP obtained from this source is made possible because of the ‘universalism’ and ‘public administration’ principles of the Canada Health Act (Roos & Roos, 2001). A central use of the MCHP Registry is to provide a link between all the other administrative databases held in the Repository. This is done using a scrambled identifier that is available in both the Registry and in some of the other databases in the Repository (e.g. physician visits and hospitalization data) (Manitoba Centre for Health Policy, 2016a). For other databases the scrambled identifier is used in a “multi-stage de-
Data Linkage for IDD Health Surveillance

identification” process that permits linkage and addresses privacy and confidentiality concerns (see Roos, Brownell, Lix, Roos, Walld, MacWilliam, 2008 for details). All together, the Repository databases make it possible to compile a comprehensive history for all individuals living in Manitoba while maintaining the confidentiality of personal information (Finlayson, Lix, & Roos, 2011).

Insert Figure 1 About Here

In 1992, a similar arms-length research center called ICES (formerly the Institute for Clinical Evaluative Sciences) was established in the province of Ontario (pop. 13.5 million, 2016) (ICES, 2019a; Statistic Canada, 2019). ICES has historically conducted population-based research using a number of health databases from Ontario including hospitalization and physician visits. It has been steadily expanding its holdings to include other kinds of data (e.g., census, mortality, disability income, and immigration information). At ICES, the Registered Persons Database (RPDB) plays the same key role for linking databases as the MCHP Registry (Government of Ontario, 2017). ICES has similar mission and vision statements as MCHP (ICES, 2019b). In addition to both centers excelling at conducting and producing rigorous population level research, the centers and researchers also invest substantial effort towards communicating and engaging with partners. These include government officials, health and social service administrators and others in an effort to increase the likelihood that findings are ‘not only useful, but usable’ (Brownell, Kozyrskyj, Fuchs, & Santos, 2011). Since their inception, both centers have not only published academic publications but also reports and ‘atlases’ that address specific policy questions (Wolfson, 2011). Later this paper will provide examples of how this ‘integrated knowledge translation’ has been applied in the context of IDD.
Methodology for Data Linkage

**Developing IDD identifiers.** To conduct and publish research for people with IDD that is similar to research for the general population of Manitobans and Ontarians, it was first necessary to identify persons with IDD within databases held at MCHP and ICES. While persons with IDD are entitled to the same publically financed health services as all Canadians and are included in the MCHP Registry and ICES RPDB, neither province maintains a population level IDD registry that can be used for linking. Further the diagnosis of IDD is not consistently and comprehensively flagged in any of the included databases held by either center. To identify persons with IDD, researchers created an appropriate algorithm tailored to the coding conventions of each database that potentially included a marker for IDD. Once a cohort of persons with IDD was created from the linked records, essentially any of the research conducted for the general population at MCHP and ICES can be duplicated in persons with IDD.

**Inclusion in the database.** In Canada, the procedures for identifying persons with IDD within administrative databases have evolved over the years. One of the first Canadian studies used a single, unlinked database of hospitalization data as a source for identifying persons with IDD (Balogh, Hunter, & Ouellette-Kuntz, 2005). Because only one dataset was used and the coding algorithm for identifying IDD was rudimentary compared to current approaches, the authors emphasized that a number of people with IDD were probably missing from the analysis and that these were likely persons with a mild form of the disability. The first study where record linkage was used to identify persons with IDD within multiple health and non-health databases held at MCHP was a prevalence study by Ouellette-Kuntz and colleagues (Ouellette-Kuntz et al., 2009). The Venn diagram in Figure 2 shows the proportions of persons with IDD
identified by the various Manitoba sources in that study. Venn diagrams (figures 2, 3 and 6) are used in this paper to emphasize the importance of including databases from multiple sources in order to maximize case ascertainment. For example a substantial proportion of people with IDD living in Manitoba would have been missing from analyses had data from education not been accessible. The Venn diagrams also provide a sense for the relative contribution of each data source to the identification of IDD. MCHP maintains and updates an online “Concept Dictionary and Glossary” that provides definitions for variables on about 300 different concepts relevant to researchers using Repository data (Manitoba Centre for Health Policy, 2017). The codes used for identifying persons with IDD from the Repository are described in detail on the website (Manitoba Centre for Health Policy, 2017, 2019b).

Building on this methodology, in 2010 Ontario-based researchers were awarded funding from the Canadian Institutes of Health Research to explore data linkage options for identifying persons with IDD using Ontario databases. At that time, unlike MCHP, ICES did not include linked data from the provincial government’s social services, which also captures information about disability diagnoses eligible for public disability income support. The work of Yona Lunsky and the Health Care Access Research and Developmental Disabilities team (H-CARDD) led to the first transfer of social services data to ICES to create a representative cohort of people with IDD in Ontario (Health Care Access Research and Developmental Disabilities, 2019; Lin et al., 2014). Taking advantage of the longitudinal nature of the data held at ICES and the fact that IDD is a lifelong diagnosis, H-CARDD also developed a coding algorithm that searched for IDD
Data Linkage for IDD Health Surveillance

diagnoses back to the inception of the primary linked ICES health care databases as well as IDD diagnoses within social services disability income support data (Lin, Balogh, Cobigo, Ouellette-Kuntz, Wilton & Lunsky, 2013; Lunsky, Klein-Geltink, & Yates, 2013). The Venn diagram in Figure 3 shows the proportions of persons with IDD identified from the data sources used to create the H-CARDD Ontario cohort. This represents one of the largest population-based cohorts of persons with IDD in the world.

Illustrating the Value of Data Linkage

The successful identification of Manitobans and Ontarians with IDD using linked databases has resulted in a number of valuable studies as documented in peer reviewed publications. These publications commonly report on the epidemiology of various diseases (e.g. osteoporosis, diabetes, dementia, mental illness, addictions) and on health service outcomes (e.g. preventable hospitalizations, repeated emergency department use, cancer screening). The large study samples and longitudinal nature of the linked databases, lends itself to the use of retrospective cohort designs, making it possible to report disease incidence as well as risk ratios. Another common study design featured in the publications is use of a comparison group (usually the general population without IDD) to produce study outcome benchmarks against which the results for people with IDD can be compared. Some of the peer reviewed publications from MCHP and ICES are described here.

**Hospitalizations for ambulatory care sensitive conditions.** Using the linkable databases at MCHP, researchers found that after adjusting for relevant covariates, Manitobans
with IDD were 6 times more likely to experience a preventable hospitalization compared to Manitobans without IDD (Balogh, Brownell, Ouellette-Kuntz, & Colantonio, 2010). Higher than normal preventable hospitalization rates for chronic conditions (e.g. asthma and diabetes, sometimes referred to as ambulatory care sensitive conditions) in a sub-population is indicative of a problem with primary care access in that group (Ansari, 2007). Improvements in primary care management of chronic conditions in people with IDD could decrease the need to hospitalize by slowing disease progression or by preventing complications. The issue of preventable hospitalizations for diabetes was explored further in a study using the linked databases at ICES (Balogh, Lake, Lin, Wilton, & Lunsky, 2015). After controlling for differences in diabetes prevalence between groups, the study found that persons with IDD were 2.6 times more likely to be hospitalized for diabetes and diabetes related complications. These studies have helped stress the need to improve the knowledge and training of primary care providers and make resources for people with IDD and their families freely available to help make primary care health service encounters more successful. Organizations like the Developmental Disabilities Primary Care Program and H-CARDD are now working to advance these initiatives (Developmental Disabilities Primary Care Program, 2018; Health Care Access Research and Developmental Disabilities, 2019).

**Chronic disease, screenings, and injuries.** Chronic disease management and secondary prevention strategies have been a focus of some other research from linked databases in Canada. For instance researchers reported on disparities in uptake of cancer screening (e.g. cervical, breast, colon cancer) between persons with and without IDD in Ontario (Cobigo et al., 2013; Ouellette-Kuntz, Cobigo, Balogh, Wilton, & Lunsky, 2015). As noted, research identified consistently higher prevalence and incidence of diabetes in persons with IDD compared to
Ontarians without IDD (Balogh et al., 2015). Researchers from ICES reported that persons with IDD were approximately 3 times more likely to experience a low trauma fracture compared to those without IDD (Balogh et al., 2017).

Mental health in children. Researchers from Manitoba led by Shahin Shooshtari studied Manitobans with IDD at different life stages using the record linkage methodology and coding algorithm for IDD at MCHP. The availability of education data at MCHP has meant that researchers were able to create and study a more complete cohort of children with IDD (Shooshtari et al., 2017). One study identified an alarmingly high prevalence of depression among children 5-9 years of age, leading to recommendations that those providing support to this group be aware of symptoms and that screening should be considered by health providers in order to detect early signs of depression (Shooshtari et al., 2014).

Expanding possibilities. As new databases are established and linked to the existing databases held at MCHP and ICES, new types of investigations have become possible. For example, a dataset of maternal-infant data for all in-hospital births in Ontario was used to determine fertility rates in women with IDD compared to those without IDD. Findings indicated general fertility was lower in women with IDD, but the age-specific fertility was similar in young women with and without IDD (Brown, Lunsky, Wilton, Cobigo, & Vigod, 2016). Another study soon to be published used data from the Immigration, Refugees and Citizenship Canada (IRCC) database to investigate the prevalence of IDD in newcomers vs. non-newcomers (Durbin et al., 2018).

Canadian research in the area of IDD using linked databases continues to evolve. Thus far, databases from education (for MCHP) and social services (for ICES) have been primarily used to identify persons with IDD in order to create useful samples of persons with IDD.
Data Linkage for IDD Health Surveillance

Canada, less focus has been put on how these same databases –and others– can be used to study the impact of social determinants on the health of persons with IDD. There are still many gaps in our understanding of what role social determinants play in the lives of people with IDD. Evidence from England suggests that up to 31% of the higher risk of poorer health in people with IDD can been attributed to differences in social factors (Emerson & Hatton, 2007; Graham, 2005). Other opportunities are emerging, including what we believe is one of the first linkages with a charitable organization. Recently, we linked a registry of Special Olympics Ontario participants with the databases held at ICES (These datasets were linked using unique encoded identifiers and analyzed at ICES). Through this process, it was determined that of the 70,000 adults with IDD known to service entities in 2015/16, almost 19,000 of them were current or past Special Olympics Ontario participants; Close to 8,000 adults with IDD were identified in the Special Olympics registry, but not other sources. It was found that among children and youths (0-19 years, N=58,000) close to 9,000 had participated in Special Olympics; Close to 3700 young people were identified with IDD solely using the Special Olympics registry. Going forward it will be possible to investigate if participation vs. non-participation in Special Olympics (a marker for social networking and participation in a health promotion activity) is associated with health outcomes such as obesity or diabetes.

Use for Policy and Program Planning

A number of studies using linked databases from Canadian sources provided important background information and supported recommendations in the recently published Canadian consensus guidelines on the primary care of adults with intellectual and developmental disabilities (Sullivan et al., 2018). For example, referencing Balogh and colleagues (2015; 2017),
the guidelines now recommend that persons with IDD be screened for diabetes at an earlier age, and that males and females with IDD should be screened for osteoporosis in early adulthood.

Both MCHP and ICES share a mandate to generate information relevant to public policy (ICES, 2019a; Manitoba Centre for Health Policy, 2019c). In keeping with this mandate, scientists have generated a number of reports using the linkable data to support decision making, policy development, service provision, and research relevant to the health and wellbeing of Canadians with IDD (for example, see https://www.porticonetwork.ca/web/hceardd/kte/h-cardd-reports). Clear communication with the intended audience is critically important in order for these reports to maximize their impact.

The need for knowledge translation and exchange expertise is well recognized by both MHCP and ICES (Manitoba Centre for Health Policy, 2019d). In Ontario, this combination of report creation and communication has had some impact on provincial policy for individuals with IDD over the past few years. In particular, findings have contributed to a discussion about broadening the provincial government’s definition of dual diagnosis (currently defined as IDD and mental health need) to include addictions as well as mental illness (Ministry of Health and Long-Term Care and Ministry of Community and Social Services, 2008). Findings have also been cited in a 2014 report submitted to the Ontario legislature addressing an urgent need to improve the IDD services and in a 2016 Ombudsman investigation on crisis situations involving adults with developmental disabilities (Dube, 2016; Select Committee on Developmental Services, 2014). The Ombudsman investigation generated 60 recommendations that the government was required to implement and report back on regularly.

The most recent Ontario example of this report-knowledge-exchange approach is the February 2019 release of a report examining five health and health care outcomes -- 30-day
repeat emergency room visits, 30-day repeat hospitalization, alternate level of care (known in other jurisdictions as delayed discharge or bed blocking), admissions to long-term care, and premature mortality (Lin et al., 2019). One of the goals of this report was to bridge silos of care as the portfolios for these five outcomes were held by different parts of the provincial government. Analyses, reporting formats, and dissemination strategies were designed throughout the project via discussion with a multi-ministry advisory committee. The main finding was that adults with IDD compared to adults without IDD had worse results across all five outcomes regardless of age, sex, and the wealth or poverty of their neighborhood of residence. The consistency of this disparity, together with the fact that there was considerable overlap in the recommended interventions for each outcome, supported the necessity of taking a systems-wide rather than an outcome-specific approach. The intent is to build on the cross-government discussion established by the Advisory toward a multi-stakeholder creation of more systemic solutions.

**Uptake in other areas.** Other Canadian provinces (e.g. British Columbia) have adopted the linking methods described here in order to study the health and health service use of people with IDD in their own jurisdictions (Marquis, McGrail, Hayes, & Tasker, 2018). There is, however, a concern for Canadians with and without IDD that there is no integrated pan-Canadian capacity to conduct research similar to what is occurring in the provinces separately. As expressed in an article by Stephen Lewis “Canada’s constitutional structure means that developments in each province and territory will follow their own paths” (Lewis, 2011).

**Data Linkage in Australia**
The Australian Context for Health and Disability Services

Somewhat like Canada, the Commonwealth of Australia is a Federation of States so that there are different tiers of government—Commonwealth (federal), State and even Local Government. This means that the funding and provision of services can come from different government entities as well as from Non-Governmental Organizations. For health care, General Practitioner (GP) services are funded by the Commonwealth and hospitals by the state. Similarly, with disability, carer allowances and pensions are funded by the Commonwealth while support services have been traditionally funded by states but differently from state to state.

In order to obtain reliable, consistent data about disability services with minimal load on the field, the Australian Institute of Health & Welfare (AIHW) has, since 1991, facilitated an annual collection and collation of nationally comparable data about disability support services under a series of Commonwealth/State agreements. This data collection has had multiple iterations in an attempt to improve the quality of the data. Data are provided by numerous heterogeneous government and non-government service providers across the nation, with the latest name for the collection being the Disability Services National Minimum Data Set (DSNMDS)(Australian Institute for Health and Welfare, 2019). One of the first reports generated by AIHW in 1997 provided estimates of the incidence and prevalence of ID in Australia using the Australian Bureau of Statistics 1993 Survey of Disability, Ageing and Carers as a data source (Wen, 1997). Disappointingly, none of the many reports subsequently generated by this unit have focussed specifically on ID.

Western Australia (WA) represents about 10% of the total Australian population of 26 million people and over three quarters live in the major metropolis Perth. In WA, disability services have generally been funded by the state government and until recently many of the
programs such as early intervention, accommodation, employment services, recreation and respite were both funded and provided by the state Disability Services Commission (DSC; recently referred to as Department of Communities but ‘DSC’ will be retained for the purposes of this paper). In recent years, DSC has continued to fund services but with their provision increasingly outsourced to Non-Government Organizations. DSC is quite unique in Australia in that, since the 1950s, it and its predecessors have maintained records of those of its clientele who meet the criteria for a diagnosis of ID.

A new federally funded Disability Support Scheme, known as the National Disability Insurance Scheme, is currently being implemented across Australia (National Disability Insurance Agency, 2019). It aims to give consumers more choice in obtaining the services they require and, in WA, in replacing the state-funded services provided by the DSC. This new national scheme, estimated to cost $22 billion annually, commenced its roll-out in 2016 and is anticipated to make a complete transition by 2021. Currently there is some uncertainty as to how ID will be assessed in this new system and whether the services provided will be an improvement on the previous system.

Partners in Development of the Western Australia Intellectual Disability Database

**Initial collaborations.** DSC and its forerunners have been the principal service providers for children with ID in WA and, since 1964, have maintained an internal register of their client base. Originally, their clientele was restricted to people with ID but more recently has included all disabilities. The impetus for the establishment of the WA Intellectual Disability Database

---

1 The name assigned to the organization providing services to people with ID in WA, which is now DSC, has changed over time. Originally it was the WA Health Department's Division for Mental Deficiency, subsequently “Irrabeena”, the Department of Intellectual Handicap, the Authority for Intellectually Handicapped Persons and then DSC.
Database was provided in part in 1996 by Louisa Alessandri, who had a disability herself, and her colleagues who recognized the need for improved capability in counting the number of persons with different disabilities in WA (Alessandri, Leonard, Blum, & Bower, 1996). This prompted one of the current authors (HL), as a medical doctor working at DSC in 1999, to investigate whether their administrative data could be used as an epidemiological database for ID. Earlier in that decade, these data had been used to investigate the cause and prevalence of ID (Wellesley, Hockey, & Stanley, 1991; Wellesley, Hockey, Montgomery, & Stanley, 1992). It was clear that the estimates obtained were biased towards people who were registered with DSC and the likelihood that a significant proportion of the population was missing, especially those from the disadvantaged or indigenous families who were not receiving services from DSC. This prompted our group to approach the Department of Education to investigate the possibility of linking Education data to DSC data. Following fruitful discussions, we were able to establish such a database of de-identified data on people with ID that combined DSC and Education data, and that could be linked as an entity to other administrative databases for research purposes (Petterson, Leonard, Bourke, Sanders, Chalmers, Jacoby, & Bower, 2005). Using this combined disability data source, now known as the Western Australian Intellectual Disability Exploring Answers (IDEA) Database, we were able to demonstrate that the prevalence of intellectual disability was actually twice (14.3 vs 7.6/1000 births) that previously estimated (Leonard, Petterson, Bower, & Sanders, 2003; Wellesley, Hockey, Montgomery & Stanley, 1992). Along with ourselves (HL, JB) at the Telethon Kids Institute, our core partners are DSC and the Department of Education. We are hoping that in the future data from the National Disability Insurance Scheme will also be included.
Inclusion in the database. Eligibility for services is determined by psychometric and adaptive behavior assessments and any child or adult who is registered and meets the DSC’s criteria for ID is considered eligible for the IDEA Database. Children identified through the Department of Education for educational support are assessed similarly. After linkage of the datasets from DSC and Education and reconciling of all information, individuals are deemed eligible for the IDEA Database if they clearly meet criteria for ID (Schalock, Bothwick-Duffy, Bradley, Buntinx, Coulter, Craig,… Yeager, 2010) from both or either source. Where there is conflicting information on level of ID from the two sources the most recent assessment will generally be used to determine eligibility. Children registered with DSC who are too young to undertake psychometric tests may be deemed “vulnerable to ID” based on developmental assessments (such as Griffiths Mental Development Scales) and be eligible for the database. Many of these children will be reassessed at school age and their eligibility for the database confirmed.

Data linkage capacity. Data Linkage WA (the name given to the Western Australian Data Linkage system) undertakes the data linkage of records from children born in WA since 1980, identified through the Midwives Notification System, to DSC and Education Department data. DSC has provided the funding for Telethon Kids Institute to employ a part-time research officer to manage the database and assist with analyses. The strength of the system is that through Data Linkage WA there is capacity to link to other databases such as Midwives Notifications, Birth and Mortality Registers, Hospitalizations, Mental Health System Contacts, Child Protection and Justice - the latter through an infrastructure known as the Developmental Pathways Project, although funding is required to cover the costs of extraction on each occasion it is requested. The Developmental Pathways Project represents a collaboration between
Data Linkage for IDD Health Surveillance

researchers at Telethon Kids Institute and a large number of government departments who provide their data for approved linkages. It also includes a community reference group who provide community input into the research that is conducted through utilization of the linked data.

Figure 4 about here- Data Linkage Structure in WA

Infrastructure and Methodology for Data Linkage in WA

Active since 1985, Data Linkage WA is the longest running and most extensive linkage system in Australia in terms of the number and duration of its linked data collections. The linkage system initially involved a small number of discrete core statutory datasets, but has since expanded to include numerous other health, education, welfare, administrative and disability datasets (Figure 4). Best practice protocols are used to link and extract data that are accessible from the multiple data custodians and these protocols are guided by principles to protect private and confidential information (Kelman, Bass, & Holman, 2002). Importantly, during the data linkage processes, personal identifiers are separated from records to be linked and the linkage keys that connect data to personal identifiers are confined to the linkage process. With unique project-specific individual identification numbers given to every new data extraction, individuals cannot be matched across research projects. Additionally, linked data are only provided for approved research with specific governance, data extracts only include the variables essential to the research, and strict commitments to security must be upheld (Kelman et al., 2002). This method has resulted in a substantial reduction in the use of name-identified data being used in data linkages and extractions over time (Trutwein, Holman, & Rosman, 2006). For simple linkages involving small numbers of cases or data collections, data encryption and unique
identifiers are used in extractions for approved research projects. However, for complex population-based research using several data collections, specialized procedures to link and extract data are applied to ensure individual privacy is maintained (Eitelhuber, Davis, Rosman, & Glauert, 2014).

After unique identifiers are created by the Data Linkage Unit for individuals identified with ID by DSC and Department of Education, the information related to their level or cause of ID is securely transferred directly from these government departments to the Telethon Kids Institute using this unique identifier (Figure 5). This process is repeated every two years. The provision of the de-identified data to the IDEA Database is covered under a contract between the Telethon Kids Institute and DSC, and a Memorandum of Understanding between Telethon Kids Institute and the Department of Education.

Value of linked data for researching ID in Western Australia

Aside from developing the best method for counting and capturing people, the overarching aim and premise for the establishment of the IDEA Database was to further understand the determinants and outcomes associated with ID in order to: (a) prevent ID by reducing risk factors and increasing protective factors; (b) target specific communities such as migrants or other disadvantaged groups where prevalence is higher and there is need for a greater focus; and (c) reduce the physical and mental health burden associated with ID on affected individuals and their families.
We have already used these data to investigate associations with ID of sociodemographic (H. Leonard et al., 2011; H. Leonard et al., 2005) and pre and perinatal (Langridge et al., 2013; H. Leonard, de Klerk, Bourke, & Bower, 2006; H. Leonard et al., 2008) factors including maternal alcohol use (O’Leary et al., 2013) and maternal physical (H. Leonard et al., 2006) and mental health (Fairthorne, Hammond, Bourke, de Klerk, & Leonard, 2015). We have also examined outcomes such as hospitalization patterns for children with ID (Bebbington, Glasson, Bourke, de Klerk, & Leonard, 2013; Fitzgerald, Leonard, Pikora, Bourke, & Hammond, 2013), dual diagnosis of psychiatric morbidity in the entire ID population (Morgan, Leonard, Bourke, & Jablensky, 2008), and increased mortality risk at different child ages (Bourke, Nembhard, Wong, & Leonard, 2017; Glasson, Jacques, Wong, Bourke et al., 2016).

In this section we first explain the value of capturing ID data through a database such as ours in comparison to use of hospital morbidity data as an ascertainment source; secondly demonstrate our findings investigating assisted reproduction as a possible risk factor for ID; thirdly describe how health service utilization in the last year of life may differ between people with ID and the general population; fourthly consider how risk of maltreatment varies according to the type of ID; and finally discuss our findings in relation to the use of gastrostomy in children with ID over a period of 35 years.

**Investigating the use of hospitalization data to ascertain ID.** In the absence of other sources, the coding of hospitalizations has been used to identify people with ID in research studies (e.g., Balogh, Hunter, & Ouellette-Kuntz, 2005; Akobirshoev, Parish, Mitra, & Rosenthal, 2017). However, it is not known how well people with ID who are admitted to hospital are captured by the appropriate diagnostic codes. In WA we were able to investigate this issue by linking the hospitalization episodes of people with ID ascertained by the IDEA Database
with the ICD codes they had been assigned for each admission record (Bourke, Wong, & Leonard, 2018). If we accept that an individual had an ID if they were identified either through our database or coded as having an ID using ICD codes similar to elsewhere (Lin, Balogh, Cobigo, Ouellette-Kuntz, Wilton & Lunsky, 2013), we found that 10,218 of persons with ID born between 1983-2010 had been hospitalized at least once. Of those, almost all (95%) had been identified through the IDEA Database while less than 20% could be identified using hospital codes (Figure 6). Those who were identified in both sources tended to be younger, more likely to be female, non-Aboriginal and have severe ID compared with those only identified in the IDEA Database. A quarter of the 478 identified only in the hospital system had died before one year of age, most at under one month. Of children surviving to one year of age and identified in both sources, syndromic or monogenic etiologies of ID such as Down syndrome, neurofibromatosis and fragile X syndrome were most likely to be identified and ID of unknown cause, autism or fetal alcohol syndrome were least likely to be identified in hospital sources. Overall, this indicates the limitation of using hospital coding alone as a source of identification of individuals with ID.

Association of assisted reproduction with presence of ID in the offspring. Since its inception 40 years ago, the use of Assisted Reproductive Technology (ART) has increased dramatically so that in Australia 4.3% of births are now conceived using these treatments (Fitzgerald, Harris, Paul, Chambers, 2017; Australian Institute for Health and Welfare, 2017). Since the results of current literature are conflicting and often confounded by methodological
issues, the second example that we provide relates to whether or not assisted reproduction is associated with any increased risk of ID (Hansen, Greenop, Bourke, Baynam, Hart, & Leonard, 2018). For our study the WA IDEA Database (Petterson et al., 2005) was linked to Midwives Notifications for all live births from 1994-2002 (Hutchinson, Joyce, 2016), the WA Register of Developmental Anomalies (Bower, Baynam, Rudy, Quick, Rowley, Watson, Cosgrove, 2015) and the Reproductive Technology Register which has collected data from all WA fertility clinics since 1993 (Western Australia Reproductive Technology Council, 2003). We found that children conceived with ART (n=2876) had a slight increased risk (RR 1.58, 95% CI1.19-2.11) of ID. However, the risk more than doubled for those born very preterm, for severe ID and if intracytoplasmic sperm injection (ICSI) had been used. Those conceived by ICSI (n=937) also had a greater likelihood of having a known genetic cause of ID including occasional cases of Angelman, Prader Willi and Russell Silver syndromes. This study was limited to children born from 1994 to 2002 and ID data were available up until 2010, allowing eight years of follow-up for any potential diagnoses of ID. Major changes are occurring over time in ART practice with single embryo transfer replacing the previously common method of multiple embryo transfer which was associated with a higher prevalence of multiple births and perinatal complications (Fitzgerald et al., 2017). Therefore, some of our results may be more pertinent to countries like North America, Latin America and the Middle East where multiple embryo transfer is still common and ICSI use rates are high (Boulet, Mehta, Kissin, Warner, Kawwass, Jamieson, 2015; Dyer, Chambers, de Mouzon, Nygren, Zegers-Hochschild, Mansour, . . . & Adamson, 2016; Zagadailov, Hsu, Seifer, Stern, 2017).

Comparing patterns of health service use in people with ID and the rest of the population in the last year of life. We know from previous research that health care utilization
in people with ID is much greater than in the general population (Bebbington et al., 2013),
generally because of the complex medical conditions that often accompany ID, particularly when
it has a genetic etiology (Fitzgerald et al., 2013; Leonard et al., 2016). However, there is also
concern that people with ID do not have the same access to adequate health care as the general
population (Krahn, Hammond, & Turner, 2006) and that lack of adequate health care may have
contributed to some premature mortality (Bourke et al., 2017; Heslop, Blair, Fleming, Hoghton,
Marriott, & Russ, 2014). To investigate this, we used the linked data of WA IDEA Database or
relevant ICD codes from hospitalizations or death certificates to identify 591 Western
Australians with an ID aged 20 years and over who had died during 2009-2013. We used
“coarsened exact matching” to identify a matched cohort of 29,713 people without ID who had
died during the same period such that there was a similar balance in age-groups, sex, indigenous
status, socioeconomic status and country of birth between the two cohorts. We then examined the
health service utilization in the last year of life of people with ID compared with people in the
general population without ID (Brameld, Spilsbury, Rosenwax, Leonard, & Semmens, 2018).
Decedents with ID tended to be younger, lived in areas of more social disadvantage, did not have
a partner, and were Australian born compared with all other decedents. Decedents with ID
attended emergency departments more frequently (3.2 vs 2.5 visits) and were admitted to
hospital less frequently (4.1 vs 6.1), but, once admitted, stayed longer (5.2 days vs 4.3 days).
Decedents with ID had increased odds of presentation, admission, or death from conditions
identified as ambulatory care sensitive and potentially preventable, including vaccine-
preventable respiratory disease, asthma, cellulitis, and convulsions and epilepsy. These findings
suggest that, at least in Western Australia, primary health care for people with ID needs to be
improved.
Risk of maltreatment among children with ID. Previous research has indicated that children with disabilities have an increased risk of maltreatment but has not examined variation of risk among children with different disabilities (O’Donnell, Nassar, Leonard, Jacoby, Mathews, Patterson, & Stanley, 2010), (Sullivan & Knutson, 1998). To investigate this, we conducted a population cohort study of children who were born from 1990-2010 and followed them until 2010 (Maclean, , Sims, Bower, Leonard, Stanley, & O’Donnell, 2017). We utilised the Western Australian Midwives and Birth registration data (births); the IDEA database, WA Register of Developmental Anomalies and Hospital Morbidity Data System (HMDS); and Department of Communities Child Protection and Family Support data (maltreatment). Of the 524,534 children in the cohort, 8551 had ID. Of children with ID, 18.7% had a child protection allegation, 10.6% had a substantiated allegation, and 6.2% entered out-of-home care. Controlling for socioeconomic status and parental mental health or substance-related admissions, we identified a 2-fold increase in risk for child maltreatment allegations, substantiated allegations, and entry into out-of-home care. Less severe disability was related to an increased risk of maltreatment allegation; children with borderline-mild ID had an almost three-fold increased risk, with mild-moderate ID having a two-fold increased risk. Children with Down syndrome and children with autism had much lower risks of maltreatment allegations while the risks were greater for some types of familial ID. Policy implications are that more support is needed for families of children with disabilities to meet their health and developmental needs and to manage the more complex parenting environments that come from having a child with additional needs. Further, there was a large attenuation of risk after adjustment of demographic and psychosocial risk factors, indicating that these factors also need to be considered when assessing the supports required for families to ameliorate risks for child maltreatment.
Use of gastrostomy in children with ID over a period of 35 years. WA data from the IDEA database extrapolated to the entire country suggest that approximately 750 children are born each year in Australia with a moderate to severe ID (Bourke, de Klerk, Smith, & Leonard, 2016). Many of these children experience difficulties with poor feeding (Romano, Dipasquale, Gottrand, & Sullivan, 2018) and comorbidities such as epilepsy (Robertson, Hatton, Emerson, & Baines, 2015). Children with severe ID require up to 10 times as many hospitalizations compared with the general population (Bebbington et al., 2013), often for lower respiratory tract infections or exacerbation of epilepsy. This group of children are frequently on the brink of ill health and require substantial care and support.

Enteral feeding via a gastrostomy tube is often used (Sleigh & Brocklehurst, 2004), with documented weight gain (Viktorsdottir, Oskarsson, Gunnarsdottir & Sigurdsson, 2015), but the evidence base supporting gastrostomy is otherwise limited. Randomized clinical trials would not be feasible or ethical, but studies using administrative hospital datasets and linked data methodologies allow evaluation of the epidemiology of gastrostomy and any inequities in its use. Using the IDEA Database and linking to hospitalization records, we investigated the epidemiology of gastrostomy insertion in WA children with ID and whether accessibility was similar across demographics of socioeconomic status, geographical remoteness and Aboriginality (Wong, Leonard, Pearson, Glasson, Forbes, Ravikumara, . . & Downs, J., 2018). Gastrostomy was increasingly used in recent years, including in children from the lowest socioeconomic status quintile, those who lived in regional/remote areas or who were Aboriginal. Accessibility to gastrostomy did not vary by these indicators of disadvantage.

There are further important research questions related to health outcomes that data linkage methodologies could usefully inform. For example, what is the hospitalization pattern
before and after gastrostomy and does this vary by the cause of the hospital admission? What is
the life expectancy of children who undergo gastrostomy and does gastrostomy improve
survival? Does gastrostomy reduce the cost of long term health care utilization? Does
gastrostomy confer advantages to children, maternal and family quality of life? Data linkage can
have an important role in identifying the evidence-base for some clinical procedures and care
management that otherwise would be difficult to address.

Use in Policy and Program Planning

Our findings have been used to inform submissions to Senate inquiries in relation to the
importance of recognizing the special needs of people with ID in the development of the
National Disability Insurance Scheme (Parliament of Australia, 2019). More recently Australia’s
Royal Commission into the institutional response to the sexual abuse of children requested
evidence from us about the burden of abuse and neglect in this group and the increased risk for
children with different types of disability. Currently, our findings are being used by advocacy
groups lobbying politicians for better health and medical services for Australians with ID
(NSWCID, 2019).

Uptake in other areas. While WA has the only dedicated ID database in Australia, some
ID linkage work has been undertaken in another Australian state (New South Wales), originating
from a program focusing on mental health in people with ID (Florio & Trollor, 2015). This work
has been achieved by linking data from separate government agencies and using disability data
generated from the Department of Family & Community Services, Ageing, Disability & Home
Care (Reppermund, Srasuebkul, Heintze, Reeve, Dean, Emerson, . . . & Trollor, 2017). However,
this linkage is not a routine process and acts as a snapshot of data applicable only to New South Wales.

**Concluding Observations, Commonalities and Future Directions**

The examples on data-linkage from both Canada and Western Australia illustrated several common themes. A first commonality is that both examples cite the visionary leadership and early work of individuals and their colleagues who recognized the potential of data linkage, and who persisted in advancing its possibilities. In Canada, Noralou and Les Roos are credited with helping realize the value of linked databases to help create MCHP. In Western Australia, Fiona Stanley is credited with advancing the vision of data linkage across datasets, and Louisa Alessandri and her colleagues for raising awareness of the need to be able to count the number of persons with different disabilities. Additionally, both teams included experts in epidemiology, a key consideration credited for seeing the potential range of other uses for administrative data, primarily collected for purposes other than research (Wolfson, 2011). A second observation is that in both examples, linkage of IDD related datasets was part of a larger data-linkage environment. The presence of centres of excellence with existing linking capabilities for the general population made it feasible to pursue similar research for people with IDD. In Canada, work began with an established data repository for the general population, developing capacity to identify people with IDD in that environment. In Australia, an active data linkage system was investigating health outcomes in the general population (Holman, Bass, Rouse, & Hobbs, 1999). The data registry of people with ID was first established in 1999 with the first linkage of data between DSC and Education (Leonard, Petterson, Bower & Sanders, 2003), and could then be used to link with other datasets. A third similarity is that in both examples, the work on IDD began modestly and has grown and evolved over time. The evolution saw increasing accuracy of
Data Linkage for IDD Health Surveillance

identification algorithms, and expansion of databases to access. A fourth observation is that in both examples, the principal persons involved were attentive to their audiences of policymakers and dedicated knowledge translation efforts to ensure findings were usable and useful to them. Work also included persistence in formalizing agreements between individual departments (Petterson et al., 2005) and promoting a shared sense of value from maintaining the data collections. Both Australia and Canada noted the importance of involvement of advocates to drive research on behalf of people with IDD (Leonard, Glasson, Bebbington, Hammond, Croft, Pikora, . . . & Glauert, 2013; Lin, Balogh, Durbin, Holder, Gupta, Volpe, … & Lunsky, 2019). Finally, both examples imply that a high degree of rigor was used in conducting their studies, as evidenced by reports and testimony to governmental entities and by peer-reviewed scientific publications. Taken as whole these observations suggest that the following conditions should be in place for a jurisdiction to feasibly and successfully conduct this type of research for people with IDD: 1) a number of population level administrative databases need to exist with capacity to link records at an individual level based on identifier variables; 2) the legislative environment needs to allow for the use of linked administrative data for research purposes; 3) existing expertise in database linking in a secure environment should already be in place, 4) consensual commitment to privacy protections, scientific integrity and adherence to data-sharing protocols; and 5) commitment of resources to establish and maintain the linkage capability.

One of the often stated limitations of using administrative data for conducting research on persons with IDD is the possibility that a number of persons with IDD may not have been identified with this diagnosis in any of the administrative databases, resulting in those people being misclassified as not having IDD. This type of misclassification means that prevalence
estimates are undercounts, and differences in outcomes reported in studies investigating disparities between those with and without IDD are likely biased towards the null.

While Western Australia maintains a disability registry, no regions of Canada benefit from this type of linkable population level registry of persons with IDD. A few Canadian and American studies have investigated the validity of using administrative data for IDD diagnoses, typically focusing on single conditions like Down syndrome or autism (Coo, Ouellette-Kuntz, Brownell, Shooshtari, & Hanlon-Dearman, 2018; Jensen, Cooke, & Davis, 2014). There are publications that point to methods to modify algorithms for identifying cases in administrative databases that balance the need for sensitivity while not sacrificing specificity (Lin et al., 2013, 2014). These studies show that the cohorts of persons with IDD created from multiple linked databases are representative of the population, and that including databases from multiple domains (e.g. health, education, social services) can make significant improvements to case ascertainment. Urbano and colleagues have published an example of the necessary syntax required to link administrative datasets using open source software for disability researchers (Urbano, Beck, & Stephens, 2013).

Maintaining data-linkage resources requires ongoing effort and funding. Databases need regular updates, requiring ongoing efforts at maintaining linkage capabilities. Larger systems changes intended to benefit the IDD population can create disruptions in established data collection methods. One such example is the change in delivery of services to Australians with ID through the implementation of the NDIS. Therefore the challenge for the future will be again to harness the same attributes employed in the formative years of the WA IDEA Database to ensure its sustainability into the future by making new arrangements and agreements with the federal NDIS.
Despite those limitations and challenges, the findings from these two countries illustrate the incredible value of linked datasets in addressing questions about the health of persons with IDD. These questions cannot, efficiently, be answered in other ways while including total populations. We are optimistic that other entities, including in the U.S., will expand this methodology and commitment to advance knowledge to inform policy and practice to support health and well-being of people with IDD.
References


Data Linkage for IDD Health Surveillance


Developmental Disabilities Primary Care Program. (2018). Developmental Disabilities Primary Care Program.


researching down syndrome. *Medical Care, 52*(8), e52-57. https://doi.org/10.1097/MLR.Ob013e31827631d2


Data Linkage for IDD Health Surveillance


Figure captions

Figure 1:
The de-identified linkable databases called the Manitoba Population Research Data Repository housed at the Manitoba Centre for Health Policy
(Manitoba Centre for Health Policy, 2016b).

Figure 2:
Venn diagram showing approximate proportion of distinct individuals with intellectual and developmental disabilities identified according to database source, all ages, Manitoba, 1998/99-2002/03 (Total N=5384, Intellectual and developmental disability administrative prevalence=4.7 per 1000 population) (Ouellette-Kuntz et al., 2009)
Figure notes: Not shown are the approximately 1% identified from the Social Assistance database. The darker grey area represents individuals who were found in multiple databases (~15%).

Figure 3.
Venn diagram showing approximate proportion of distinct individuals with intellectual and developmental disabilities identified according to database source, 18-64, Ontario, 2009/10 (Total N=66,484, prevalence=7.8 per 1000 population) (Lunsky, Klein-Gelting, & Yates, 2013)
Note: Health databases include hospital visits (same day and overnight stays), emergency room visits, and physician visits.

Figure 4:
Data Collections available to access through the Western Australia Data Linkage System for approved research projects.

Figure 5:
Protocols for the transfer of identifying information used by the WA Data Linkage Unit to generate unique identifiers and the subsequent provision of information on individuals with intellectual disability from the government bodies to the IDEA Database.

Figure 6:
Venn diagram showing counts of individuals with intellectual disability as identified by the Hospital Morbidity Data System (HMDS) and the IDEA Database in Western Australia.
Manitoba Population-Based Registry

Hospital
Physician Services
Nursing Home
Home Care
Immunization
Vital Statistics
Pharmaceuticals
Emergency Department
Clinical (e.g. intensive care unit)
Health Surveys
Medical Laboratory
Cancer Care
Education
Family Services
Healthy Child Manitoba
Income Assistance
Social Housing
Justice
Education
35%

Hospital
30%

Physician claims
20%
Social Services 37%

Multiple sources 30%

Health 33%
<table>
<thead>
<tr>
<th>CORE DATA COLLECTIONS</th>
<th>EXAMPLES OF SUPPLEMENTARY DATA COLLECTIONS</th>
<th>ADDITIONAL FEATURES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Births</td>
<td>Home care</td>
<td>Geocoding</td>
</tr>
<tr>
<td>Deaths</td>
<td>Notifiable diseases</td>
<td>Family links</td>
</tr>
<tr>
<td>Midwives</td>
<td>Birth defects</td>
<td>SES coding</td>
</tr>
<tr>
<td>Hospitalisations</td>
<td>Reproductive technology</td>
<td>Ability to link</td>
</tr>
<tr>
<td>Emergency hospital</td>
<td>Education</td>
<td>external data</td>
</tr>
<tr>
<td>Mental health</td>
<td>Intellectual Disability (IDEA)</td>
<td>Ability to request</td>
</tr>
<tr>
<td>Cancer</td>
<td>Corrective services</td>
<td>control groups</td>
</tr>
<tr>
<td>Electoral roll</td>
<td>Housing</td>
<td>Aboriginal</td>
</tr>
<tr>
<td></td>
<td>Transport</td>
<td>identifiers</td>
</tr>
</tbody>
</table>
Disability Services Commission

Identifying information on individuals with ID sent to DLU

Unique Root number generated by DLU for each individual

Data on individual (medical cause and ID level) using Root number

IDEA Database

Unique Root number generated by DLU for each individual in DSC and/or Education datasets

Data on individual (ID level) using Root number

Department of Education

Identifying information on individuals with ID sent to DLU

Unique Root number generated by DLU for each individual

WA Data Linkage Unit
HMDS n=478
IDEA Database n=8305
IDEA and HMDS n=1435