Using administrative health data to identify individuals with intellectual and developmental disabilities: a comparison of algorithms

E. Lin,[1] R. Balogh,[2] V. Cobigo,[3] H. Ouellette-Kuntz,[4] A. S. Wilton[5] & Y. Lunsky[6]

1 Provincial System Support Program, Centre for Addiction and Mental Health, Toronto, ON, Canada

2 University of Ontario Institute of Technology, Oshawa, ON, Canada

3 School of Nursing Sciences, University of East Anglia, Norwich, UK

4 Community Health and Epidemiology, Queen's University, Kingston, ON, Canada

5 Institute for Clinical Evaluative Sciences, Toronto, ON, Canada

6 Dual Diagnosis Program, Centre for Addiction and Mental Health, Toronto, ON, Canada

Correspondence: Dr Elizabeth Lin, Provincial System Support Program, Centre for Addiction and Mental Health, Room T314, 33 Russell Street, Toronto, ON, Canada M5S 2S1 (e-mail: elizabethbetty.lin@camh.ca).

Abstract

Background Individuals with intellectual and developmental disabilities (IDD) experience high rates of physical and mental health problems; yet their health care is often inadequate. Information about their characteristics and health services needs is critical for planning efficient and equitable services. A logical source of such information is administrative health data; however, it can be difficult to identify cases with IDD in these data. The purpose of this study is to evaluate three algorithms for case finding of IDD in health administrative data.

Methods The three algorithms were created following existing approaches in the literature which ranged between maximising sensitivity versus balancing sensitivity and specificity. The broad algorithm required only one IDD service contact across all available data and time periods, the intermediate algorithm added the restriction of a minimum of two physician visits while the narrow algorithm added a further restriction that the time period be limited to 2006 onward. The resulting three cohorts were compared according to socio-demographic and clinical characteristics. Comparisons on different subgroups for a hypothetical population of 50 000 individuals with IDD were also carried out: this information may be relevant for planning specialised treatment or support programmes.

Results The prevalence rates of IDD per 100 were 0.80, 0.52 and 0.18 for the broad, intermediate and narrow algorithms, respectively. Except for 'percentage with psychiatric co-morbidity', the three cohorts had similar characteristics (standardised differences <0.1). More stringent thresholds increased the percentage of psychiatric co-morbidity and decreased the percentages of women and urban residents in the identified cohorts (standardised differences = 0.12 to 0.46). More concretely, using the narrow algorithm to indirectly estimate the number of individuals with IDD, a practice not uncommon in planning and policy development, classified nearly 7000 more individuals with psychiatric co-morbidities than using the intermediate algorithm.

Conclusions The prevalence rate produced by the intermediate algorithm most closely approximated the reported literature rate suggesting the value of imposing a two-physician visit minimum but not restricting the time period covered. While the statistical differences among the algorithms were generally minor, differences in the numbers of individuals in specific population subgroups may be important particularly if they have specific service needs. Health administrative data can be useful for broad- based service planning for individuals with IDD and for population level comparisons around their access and quality of care.

Keywords administrative data, cohort algorithms, intellectual disability, standardised differences, surveillance

Introduction and background

Although individuals with intellectual and developmental disabilities (IDD) are highly vulnerable to physical and mental health problems, their care is often inappropriate (Ouellette-Kuntz et al. 2005; MENCAP 2007; Emerson et al. 2011; World Health Organization 2011). Planning efficient and equitable services requires a better understanding of their healthcare needs, and there is substantial literature suggesting that administrative health data can provide important contributions towards such understanding (Adams et al. 2009; World Health Organization 2011).

Compared to data from household surveys or patient records, administrative data are less expensive, routinely collected, not impacted by self-report limitations and have a better potential of capturing more of the population of interest particularly if they are linked across multiple sectors (Westerinen et al. 2007; Balogh et al. 2010; Ouellette-Kuntz et al. 2010). Health administrative data sets also allow for comparisons with other groups or the general population, which are crucial in identifying healthcare disparities. Patterns of healthcare use and disparities in both care and outcomes have been described for a variety of population groups including individuals with depression (Lin et al. 2011), stroke (Tirschwell et al. 1999), HIV (McGinnis et al. 2003) and low- income status (Khan et al. 2011). Administrative health data are increasingly being used to study the population with IDD. However, it can be difficult to identify cases with IDD on a population-wide basis (Westerinen et al. 2007). This paper contributes to a better definition and identification of cases with IDD in health administrative data sets by comparing three IDD case-identifying algorithms using administrative health data sets in Ontario, Canada.

In the IDD literature, there are two common methods of case identification. One approach, used predominantly by prevalence studies, combines general population surveys with clinical assessment. Well-designed surveys first screen for potential cases and then confirm diagnoses using specialised health professionals (e.g. Christianson et al. 2002; Hosain et al. 2007; Xie et al. 2008; Brugha et al. 2011; also see Fombonne 2009) although few studies manage to clinically assess all respondents (e.g. Camp et al. 1998; Heikura et al. 2003; Gustavson 2005). While these studies are felt to yield the most reliable prevalence estimates for IDD – particularly those with a prospective design – their primary drawbacks are cost, labour-intensiveness and the consequent difficulty in repeating them on a regular basis.

A second method uses registries or recipients of social, health or educational supports specifically created for people with either disability in general or IDD in particular to identify individuals with IDD (e.g. Massey & McDermott 1996; Statistics Canada 2001; Madsen et al. 2002; Leonard et al. 2003; McConkey et al. 2006; Ouellette-Kuntz et al. 2006). Some countries, such as Australia and Ireland, have used this approach to create registries of persons with IDD (Sullivan et al. 2003; Kelly et al. 2007).

While these data sources have some level of specificity because of the evaluation process required to become a ‘registered’ case, definitions vary across evaluators and across organisational eligibility requirements. A more serious limitation is that the results produced using these specialised databases are likely to be underestimates. Westerinen et al. (2007) used eight national health and social benefits registries to estimate the prevalence of IDD in Finland of which only one, hospital discharge data, was a general rather than a disabilities population database. Their analyses showed that 8% of the identified cases were recorded only in the hospital data. van Schrojenstein Lantman-de Valk et al. (2006) combined cases with IDD identified through service providers and cases identified through general practitioner patient records to extrapolate a national prevalence rate for the Netherlands. Their results showed that

between 33% and 38% of the estimated number of individuals with IDD were identified only through general practitioner records. Consequently, developing reliable methods for the identification of persons with IDD in health administrative data sets is likely to help in providing a more accurate estimate of the prevalence of IDD.

Case identification in administrative health data- bases uses algorithms that are typically validated through chart reabstraction or similar clinical review methods (Hux et al. 2002; Tirschwell & Longstreth 2002; Spettell et al. 2003; Tu et al. 2007). They are generally based on a combination of the main diagnosis being treated, threshold numbers or types of service encounters and a specific, uniform time window. A typical example would be to include only individuals with at least two physician visits for condition X or who had had an inpatient stay for the condition within the past 2 years (Hux et al. 2002). The goal is to arrive at an algorithm that balances sensitivity (i.e. including all possible cases) with specificity (i.e. excluding all possible non-cases) and also minimises bias (e.g. using the same time window across all data sets).

Identifying individuals with IDD using these methods is challenging, however. The conditions comprising IDD do not usually require direct healthcare interventions and thus are likely to be recorded in health records only when individuals are first assessed (usually early in the lifespan). Furthermore, assessments of cognitive and adaptive functioning are typically completed by psychologists and thus would not be captured in administrative data sets that are limited to medical records. For databases which allow the recording of multiple diagnoses or conditions, IDD might be captured as a comorbidity, but it is unclear how consistently co-morbidities are recorded by health service providers (Iezzoni 2002; Balogh et al. 2005; Juurlink et al. 2006). The expectation therefore is that IDD is underdetected in administrative health data and that consequently maximising sensitivity is more pressing than avoiding false positives or potential bias.

Intellectual and developmental disability studies are increasingly using general population, administrative health data. Some studies have combined such data with information from more targeted sources such as registries or disability support recipients (van Schrojenstein Lantman-de Valk et al. 2006; Hall et al. 2007; Westerinen et al. 2007; Balogh et al. 2010) and others have relied solely on health data (Lunsky & Balogh 2010; Lunsky et al. 2011). These studies used strategies to maximise the detection of IDD cases including using multiple data sources, choosing the widest time window possible and including all of the available diagnostic data (i.e. not just the primary diagnostic field). To our knowledge, no study has attempted to compare case identification strategies to determine which one seems to be the most accurate and should be recommended for research and service planning purposes.

This paper creates and evaluates three case identification algorithms which vary in their degree of inclusiveness of individuals with IDD. These are applied to administrative health data from the province of Ontario, Canada, and the resulting cohorts are compared in terms of their size and socio- demographic and clinical characteristics. Because health administrative data are more robust when reporting broad versus narrow diagnostic categories (Kisely et al. 2009a,b), our investigation will focus on evaluating algorithms for the general category of IDD, rather than for the different conditions that comprise IDD. The implications and future directions for both research and service planning are discussed.

Methods

Data sources

The seven data sources used for this report are housed and managed at the Institute for Clinical Evaluative Sciences, Ontario and were made avail- able for analysis in the form of linked, anonymised data. They include five provincial administrative health data sets, a registry of persons eligible for provincial health coverage and the Canadian Census. The five administrative health data sets capture the vast majority of the formal medical services that all legal residents of Ontario are eligible for. The Ontario Mental Health Reporting System (OMHRS) and the Canadian Institute of Health Information Discharge Abstract Database (DAD) capture inpatient discharges for all acute care psychiatric and non-psychiatric hospital beds. Same Day Surgery (SDS) and the National Ambulatory Care Reporting System (NACRS) record ambulatory care visits for inpatient surgery or to the emergency department. The Ontario Health Insurance Plan (OHIP) contains all claims submit- ted to the province by fee-for-service physicians. Table 1 provides information about the content of these seven data sets, their inception date and the variables used for this report. Also included are the numbers of cases as defined by the broadest IDD algorithm (described below). Data linkage was accomplished through an encrypted unique identifier for the administrative health and registry data- bases and the individual's residential postal code for the Census variables. The entire project, including the linkage and anonymisation methods, was reviewed and approved through the research ethics processes at the Institute for Clinical Evaluative Sciences at Sunnybrook Hospital and at the Centre for Addiction and Mental Health, both in Toronto.

Intellectual and developmental disability

Our definition of IDD is based on the one used by our provincial government to determine eligibility for disability support services and is consistent with recent Ontario legislation (Government of Ontario 2010). The conditions covered under this legislation are those that are characterised by lifelong limitations in cognitive and adaptive functioning that originate before the age of 18 and impact on activities of daily living. The definition thus includes a broad range of conditions, not based on either aetiology or an overall IQ cut-off. IDD conditions consistent with the above definition were identified using ICD-9, ICD-10 or DSM codes or their data- base equivalents in consultation with both clinicians and policy makers. The list of diagnostic codes used is shown in Table 2.

Cohort algorithms

As described earlier, standard algorithms for administrative health data aim for a balance between sensitivity and specificity as well as a method that minimises bias (i.e. by using uniform time windows

Table 1. Datasources: Content, inception date, number of IDD cases , and relevant sociodemographic and clinical variables

Datasource	Content	Inception Year	Clinical variables	IDD Cases[1] Any datasource: n = 67,337 (single datasource: n = 53,881)	Sociodemographic variables
OMHRS (Ontario Mental Health Reporting System)	Inpatient discharges from psychiatric beds (2006 to present)	2006	Psychiatric DSM-4 (3 Axis I fields) Provisional diagnosis (16 categories) Intellectual disability variable Medical ICD-10 (6 fields) Specific illnesses (6 conditions)	4419 (2264)	Age Sex
DAD (Discharge Abstract Database)	Acute care inpatient discharges (prior to 2006) Acute care inpatient discharges from non-psychiatric beds (2006 on)	1988	ICD-9 (16 fields, 1988-2001) ICD-10 (25 fields, 2002-present)	17,301 (7,667)	
SDS (Same Day Surgery database)	Hospital visits for same-day surgery	1991	ICD-9 (16 fields, 1988-2001) ICD-10 (25 fields, 2002-present)	7,297 (1,628)	
NACRS (National Ambulatory Care Reporting System)	ED visit	2000	ICD-10 (10 fields)	3,814 (527)	
OHIP (Ontario Health Insurance Plan)	Fee-for-service physician visit	1991	ICD-9 equivalent (1 field)	53,192 (41,745)	none
RPDB (Ontario Registered Persons Database)	Ontarians covered under provincial health insurance plan	1990	None	n/a	Age Sex
Canadian Census 2008 intercensal population estimates		2009	none	n/a	Average neighborhood income Rurality index (Population denominators)

Table 2. Intellectual and developmental disability diagnostic codes: ICD, DSM, and database equivalent

Code (comment)	Label
ICD-9 (* = include only if all 6 digits present)	
299-29999	Pervasive development disorders (e.g. autism)
317-31799	Mild mental retardation
318-31899	Other specified mental retardation
319-31999	Unspecified mental retardation
7580-75839	Chromosomal anomalies for which a developmental disability is typically present (e.g. Down syndrome, cri-du-chat syndrome)
7585	Other conditions due to autosomal anomalies
7588, 75889 (exclude 75881)	Other conditions due to chromosome anomalies
7589	Conditions due to anomaly of unspecified chromosome
7595	Tuberous sclerosis
75981	Other and unspecified congenital anomalies: prader-willi
759821*	Other and unspecified congenital anomalies: de Lange
759827*	Other and unspecified congenital anomalies: Seckel's
759828*	Other and unspecified congenital anomalies: Smith-Lemli-Opitz
75983	Other and unspecified congenital anomalies: fragile x
759874 *	Other and unspecified congenital anomalies: Beckwith-Wiedemann syndrome
759875*	Other and unspecified congenital anomalies: Zellweger's syndrome
75989	Other and unspecified congenital anomalies: other (e.g. menkes disease, Laurence-Moon-Biedl,rubinstein-taybi syndrome etc.)
76071	Fetal alcohol syndrome
76077	Fetal hydantoin syndrome
ICD-10 (= include only if all 5 digits present)**	
F700	Mild mental retardation with the statement of no, or minimal, impairment of behaviour
F701	Mild mental retardation, significant impairment of behaviour requiring attention or treatment
F708	Mild mental retardation, other impairments of behaviour
F709	Mild mental retardation without mention of impairment of behaviour
F710	Moderate mental retardation with the statement of no, or minimal, impairment of behaviour
F711	Moderate mental retardation, significant impairment of behaviour requiring attention or treatment
F718	Moderate mental retardation, other impairments of behaviour
F719	Moderate mental retardation without mention of impairment of behaviour
F720	Severe mental retardation with the statement of no, or minimal, impairment of behaviour
F721	Severe mental retardation, significant impairment of behaviour requiring attention or treatment
F728	Severe mental retardation, other impairments of behaviour

Code (comment)	Label
F729	Severe mental retardation without mention of impairment of behaviour
F730	Profound mental retardation with the statement of no, or minimal, impairment of behaviour
F731	Profound mental retardation, significant impairment of behaviour requiring attention or treatment
F738	Profound mental retardation, other impairments of behaviour
F739	Profound mental retardation without mention of impairment of behaviour
F780	Other mental retardation with the statement of no, or minimal, impairment of behaviour
F781	Other mental retardation, significant impairment of behaviour requiring attention or treatment
F788	Other mental retardation, other impairments of behaviour
F789	Other mental retardation without mention of impairment of behaviour
F790	Unspecified mental retardation with the statement of no, or minimal, impairment of behaviour
F791	Unspecified mental retardation, significant impairment of behaviour requiring attention or treatment
F798	Unspecified mental retardation, other impairments of behaviour
F799	Unspecified mental retardation without mention of impairment of behaviour
F840	Childhood autism
F841	Atypical autism
F843	Other childhood disintegrative disorder
F844	Overactive disorder associated with mental retardation and stereotyped movements
F845	Asperger's syndrome
F848	Other pervasive developmental disorders
F849	Pervasive developmental disorder, unspecified
Q851	Tuberous sclerosis
Q860	Fetal Alcohol Syndrome
Q861	Fetal hydantoin syndrome
Q871	Aarskog, Prader willi, DeLange, Seckel etc.
Q8723**	Rubinstein-Taybi
Q8731**	Sotos
Q878	Other
Q900-Q939 except: Q926 Q971 Q992 Q998	(i.e., all Down Syndrome Types, cri du chat, etc except Extra marker chromosomes)
DSM-IV (OMHRS database only)	
299, 299.00	Autistic disorder
299.1, 299.10	Childhood disintegrative disorder
299.8, 299.809	Asperger’s disorder

Code (comment)	Label
317	Mild mental retardation
318, 318.0	Moderate mental retardation
318.1	Severe mental retardation
318.2	Profound mental retardation
319	Mental retardation, severity unspecified
Database equivalent codes (comments)	
299 (OHIP database only)	Childhood psychoses (e.g., autism)
319 (OHIP database only)	Mental retardation
Q3 (OMHRS database)	Mandatory intellectual disability variable (yes/no)

across data sets). In contrast, methods to identify IDD cases in administrative health data appear to weigh sensitivity more heavily. In some instances [e.g. the method used by Balogh et al. (2010)], maximising the time window across different databases creates the possibility of bias if there were different database inception dates. To address the tension between these two approaches, we created three algorithms for evaluation.

The broad algorithm is based on the approach used by Balogh et al. (2010). The narrow algorithm mimics other standard cohort algorithms such as the one used in Canada for diabetes (Hux et al. 2002; Public Health Agency of Canada 2009) by imposing a uniform time window and by requiring either an inpatient or emergency department contact or a minimum of two physician visits for IDD. The argument for this last restriction is that the physician claims data set (OHIP) only permits a single diagnostic code per claim. The concern is that only requiring a single physician visit would capture individuals who are being assessed for IDD (and therefore receiving an IDD diagnostic code) but who do not end up qualifying for a diagnosis. We also created an intermediate algorithm.

The cohorts produced by the algorithms are nested – that is, the narrow IDD group is a subset of the intermediate IDD group which, in turn, is a subset of the broad group. The pool of individuals eligible for classification included any adult, aged 18–64, who was alive and living in Ontario on 1 April 2009 and who had at least one IDD record in any database. An IDD record was defined as an administrative health record with an IDD diagnosis in any diagnostic field. The specific algorithm definitions are: Broad: any inpatient/SDS IDD discharge or any ED IDD visit or any physician IDD visits since database inception Intermediate: any inpatient/SDS IDD discharge or any ED IDD visit or two or more physician IDD visits since database inception Narrow: any inpatient/SDS IDD discharge or any ED IDD visit

Other variables

Age and sex were drawn from either administrative health data or the Ontario Registered Persons Data- base (RPDB). Age was divided into five groups (18–24, 25–34, 35–44, 45–54 and 55–64). Average neighbourhood income and percentage rural were provided through Census intercensal estimated data. Ontario neighbourhoods were grouped in approximately equal-sized quintiles from poorest (Quintile 1) to wealthiest (Quintile 5) using 2006 census dissemination areas taking into account household size and community of residence. Urban– rural status was derived from census subdivisions using Statistics Canada's (2007) Statistical Area Classification of Statistics; rural represents the areas that are outside of the commuting zones of larger urban centres with a core population of 10 000 or more. Census intercensal estimates also provided the 2009 population denominators for calculating prevalence rates.

The presence of other clinical conditions was ascertained using the variables listed in Table 1. We used previously created Institute for Clinical Evaluative Sciences definitions to determine the prevalence of diabetes (Hux et al. 2002), hypertension (Tu et al. 2008), acute myocardial infarction (AMI; Tu et al. 1999), chronic obstructive pulmonary dis- order (COPD; Gershon et al. 2009a), asthma (Gershon et al. 2009b) and congestive heart failure (Schultz 2012). The psychiatric co-morbidity algorithm was developed as part of an Ontario study on mental health disorders and IDD (Lunsky et al. 2012) and included all mental and substance-related conditions except for mental retardation diagnoses, psychological development disorders, behavioural/ emotional disorders that onset in childhood or adolescence, and sleep disorders (the full list of psychiatric codes used is available from the authors).

Analyses

We compared the socio-demographic and clinical characteristics of the three cohorts to determine whether the observed differences were statistically significant. Standard tests could not be applied for two reasons. First, the three cohorts were not mutually exclusive and thus violated the assumption of independent samples. Second, even when independent groups could be created (e.g. by comparing the narrow cohort with the new individuals added by the intermediate algorithm), the large sample sizes meant that a majority of comparisons would be statistically significant even with a conservative P-level (Cohen 1994).

Consequently, we used the standardised difference, also known as Cohen's effect size index (Cohen 1992), for the statistical comparisons. This statistic is used to represent the magnitude of difference between two populations (where 0.2, 0.5 and 0.8 are interpreted as reflecting small, medium and large effect sizes, respectively). It is also used when doing propensity score matching to determine whether the test and resulting 'control' groups are 'balanced' – that is, acceptably similar (Normand et al. 2001; Austin 2009). A commonly used threshold for supporting balance is <0.1 (i.e. less than a small effect size). Standardised differences have the advantage that they do not require mutually exclusive samples (relevant since our cohorts were nested) and also are independent of sample size. We used the standardised difference in two ways. First, we compared the cohorts with each other. Second, we wanted to evaluate the new individuals added by the next broadest definition. Thus, individuals in the narrow cohort were compared to the new individuals added by the intermediate algorithm, and all of the individuals in the intermediate cohort were compared to the new individuals added by the broad algorithm.

We were also interested in examining the potential impact of the three IDD algorithms on service planning. For this analysis, we used an indirect estimation approach, one method used in planning and forecasting when exact numbers for a given jurisdiction are not known (Schaible 1996). For each of the three IDD cohorts, we took the percent- ages for the socio-demographic and clinical characteristics we examined and applied them to a hypothetical 'population' of 50 000 individuals with IDD. The figure of 50 000 was calculated by applying a 0.5% prevalence rate [in keeping with the adult rate for ID reported in a recent meta-analysis by Maulik et al. (2011) as well as adult rates for IDD found in a national postcensal Canadian survey by Human Resources and Skills Development Canada (2006)] to an adult population of 10 million people (roughly the size of the 18- to 64-year-old Ontario population; Statistics Canada 2012). This method allowed us to compare the numbers of individuals, estimated by each algorithm, who might require targeted outreach or intervention services because of their socio-demographic or clinical characteristics.

SAS software was used to calculate the descriptive statistics and Excel to calculate the standardised differences and estimated numbers of individuals needing targeted care.

Results

As would be expected, the number of individuals identified as having IDD decreased as the restrictiveness of the definition increased (Table 3). The broad algorithm identified over 67 000 individuals with the number dropping by 34% for the intermediate and again by 77% for the narrow algorithms. These translate into IDD prevalence rates per 100 of 0.80, 0.52 and 0.18, respectively. Table 3 also shows the socio-demographic and clinical characteristics of the three IDD cohorts. Of note is the clear relationship between age group and prevalence with the rates for the youngest age group being 2 to 2.5 times the rates for the oldest across all three algorithms (broad: 1.42 vs. 0.56; intermediate: 0.91 vs. 0.36; narrow: 0.30 vs. 0.14, respectively).

Figure 1 shows the standardised differences for the three cohorts compared to each other (left side) and for each cohort compared to the new individuals added by the next broadest algorithm (right side). For the cohort versus cohort comparisons, all three could be considered reasonably similar in that all the standardised differences except one are below the commonly used threshold of <0.1 (solid vertical line). The one exception is the percentage with psychiatric co-morbidity where the narrow cohort differs from the intermediate (0.29) and broad (0.36) cohorts. These standardised differences fall within the small effect size range.

The cohort versus added individuals comparisons (Fig. 1, right side) show greater differences although still below the 0.2 threshold for a small effect size (dashed vertical line), again with the exception of percentage with psychiatric co-morbidity. Compared to the intermediate cohort, the new individuals added by the broad algorithm (white diamonds) have a higher proportion of women and urban residents, a considerably lower rate of psychiatric comorbidity and lower rates of diabetes and congestive heart failure. Compared to the narrow cohort, the new individuals added by the intermediate algorithm (grey diamonds) have considerably lower rates of psychiatric comorbidity and lower rates of congestive heart failure.

Table 4 shows the results of the comparisons on a hypothetical population of 50 000 individuals with IDD. The estimated numbers based on the broad and narrow cohort percentages are compared with the intermediate cohort. This was chosen as the reference because the intermediate algorithm produced a prevalence rate closest to the 0.5 reported by both Maulik et al. (2011) for adult and adult- plus-child/adolescent ID populations and a national Canadian survey (Human Resources and Skills Development Canada 2006) for adult Canadians with IDD. For ease of reading, absolute differences of 1000 or more (equivalent to a misclassification proportion of 2% or greater of the hypothetical IDD population) are bolded while absolute differences of 5000 or more (equivalent to a 10% or greater misclassification) are both bolded and italicised.

Compared to the intermediate algorithm, the narrow algorithm produces a more skewed age distribution (fewer individuals under age 35, more individuals 54 years and older) and identifies more rural individuals and a higher proportion with psychiatric co-morbidity and diabetes, but a lower proportion with asthma. The broad algorithm identifies more women, more urban residents and more individuals without psychiatric co-morbidity. To give a specific example, the intermediate algorithm would ‘miss’ 1050 men when compared to the broad estimate but ‘overshoot’ the narrow estimate by 300 men.

Table 3. Treated prevalence rates and sociodemographic and clinical characteristics of individuals with IDD using three administrative data algorithms

Characteristic	IDD Algorithm		
	broad (n = 67,337)	**intermediate** (n = 44,161)	**narrow** (n = 15,487)
Prevalence (overall)	0.80	0.52	0.18
By age groups			
18-24 years	1.42	0.91	0.30
25-34 years	0.79	0.53	0.17
35-44 years	0.68	0.45	0.16
45-54 years	0.66	0.44	0.17
55-64 years	0.56	0.36	0.14
Sociodemographic characteristics			
% Age groups			
18-24 years	26.9	26.4	24.8
25-34 years	20.6	20.9	19.3
35-44 years	19.5	19.4	19.5
45-54 years	20.3	20.6	22.6
55-64 years	12.7	12.6	13.8
% Male	55.7	57.8	58.4
% Rural	13.4	14.5	15.5
% Income quintile			
Quintile 1 - Low	26.3	27.0	26.7
Quintile 2	21.0	21.1	21.0
Quintile 3	18.0	17.8	17.9
Quintile 4	17.6	17.4	17.7
Quintile 5 - High	15.9	15.6	16.0
Clinical characteristics			
% other psychiatric disorder	53.8	56.9	70.8
% diabetes	9.1	9.9	11.0
% hypertension	13.5	13.6	13.8
% acute MI	0.5	0.5	0.5
% COPD	5.1	5.6	6.5
% asthma	18.1	17.6	16.0
% congestive heart failure	1.0	1.3	1.8

Figure 1. Standardized differences: Cohort vs. Cohort and Cohort vs Added-Individuals

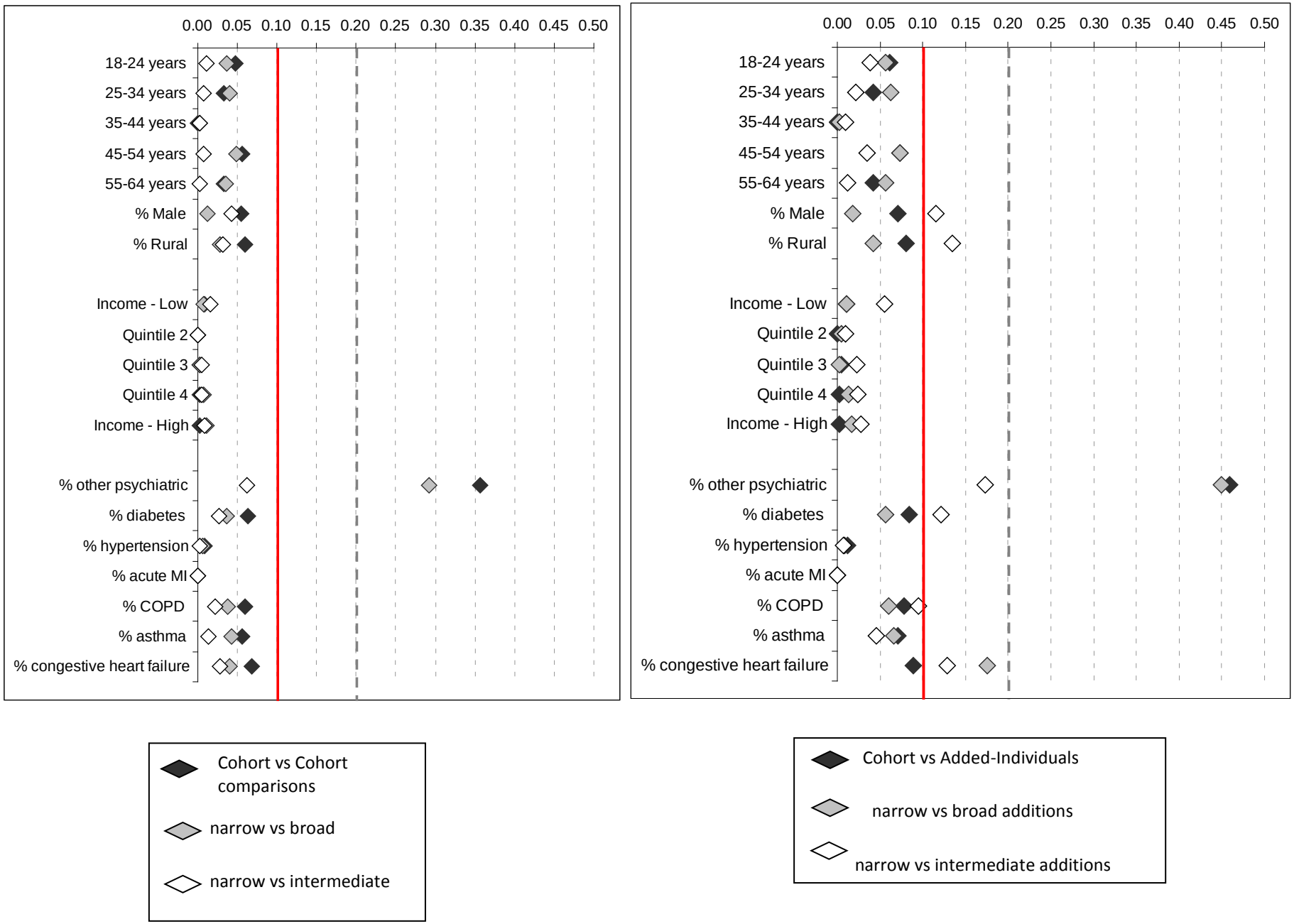

Table 4. Variance of estimated numbers of individuals with particular sociodemographic or clinical characteristics among a hypothetical population of 50,000 individuals with IDD

Characteristic	**Intermediate** Estimate	Difference from **Intermediate** Estimate	
		Broad Estimate	**Narrow** Estimate
% Age groups			
18-24 years	13,200	+250	-800
25-34 years	10,450	-150	-800
35-44 years	9,700	+50	+50
45-54 years	10,300	-150	**+1,000**
55-64 years	6,300	+50	+600
% Male	28,900	**-1,050**	+300
% Rural	7,250	-550	+500
% Income quintile			
Quintile 1(Low)	13,500	-350	-150
Quintile 2	10,500	0	0
Quintile 3	8,900	+100	+50
Quintile 4	8,700	+100	+150
Quintile 5 (High)	7,800	+150	+200
Clinical Characteristics			
# psychiatric comorbidity	28,450	**-1,500**	**+6,950**
# diabetes	4,950	-400	+550
# hypertension	6,800	-50	+100
# acute MI	250	0	0
# COPD	2,800	-250	+450
# asthma	8,800	+250	-800
# congestive heart failure	650	-150	250

Discussion

Population-based administrative health data sets are the best source of information on healthcare use and are particularly valuable in comparative analyses between IDD and non-IDD populations. As suggested by the literature, there may be important numbers of individuals with IDD who are missed when specialised data sets such as case registries or support recipients are used. We created three algorithms based on two different case identification approaches described in the literature in which the competing goals were balancing sensitivity and specificity versus emphasising sensitivity. In general, the three IDD algorithms create cohorts that are ‘acceptably similar’ (based on their standardised differences) in their socio-demographic and clinical characteristics. The one exception is that the narrow algorithm identifies a group with more psychiatric co-morbidities albeit the difference is at the small effect size level.

The major difference among the three algorithms is the resulting prevalence rate. The broad algorithm yielded a rate that was four times that of the narrow algorithm. Assuming that a prevalence rate of 0.5 (Human Resources and Skills Development Canada 2006; Maulik et al. 2011) is a legitimate criterion, the

broad definition appeared to be overly sensitive, the narrow one overly specific. Thus, the algorithm of choice would be the intermediate definition.

The higher prevalence rates in the youngest age group is consistent with findings reported elsewhere (Human Resources and Skills Development Canada 2006; Ouellette-Kuntz et al. 2010; Maulik et al. 2011). One explanation for our results is clearly methodological. Because the earliest database was created in 1988, older cohorts have fewer chances compared to younger cohorts of being identified since their IDD service contacts may have occurred when they were younger – that is, before database inception. A second explanation explored by researchers such as Hertz-Picciotto & Delwiche (2009) who examined the reasons for rising rates of autism (Centers for Disease Control and Prevention 2012) is that this is partly due to changes in diagnostic criteria for the autism spectrum disorders (ASD). It will be important to monitor these types of age-related patterns longitudinally to determine whether this is a cohort effect or whether, through processes such as higher mortality or some other type of attrition, it is an age-related phenomenon.

Beyond establishing prevalence rates, developing reliable case identification methods is important to IDD-related policy development and service planning (Adams et al. 2009; Mercadante et al. 2009). As demonstrated by our analyses, variations in the chosen algorithm may impact the proportion of individuals who have both IDD and a psychiatric co-morbidity. Our narrow algorithm identified a meaningfully larger proportion of individuals with both conditions compared to the intermediate algorithm. Statistically, the difference was consistent to a small effect size. However, it represents a difference of 6950 persons on a hypothetical population of 50 000 individuals with IDD, which is equivalent to a 14% disagreement between the two algorithms (i.e. a difference of 6950 individuals who were not detected by the intermediate algorithm). An important question is whether our narrow algorithm may have overdetected psychiatric co-morbidity because of the mandatory yes/no question in the OMHRS data (Martin et al. 2007). However, regardless of the source of the algorithm disagreement, either over- or underidentification of this group would have important real-life implications given the costs and complexity of the care and supports these individuals need. Thus, a keen awareness of the potential biases and limitations in the data on which the extrapolation is based is essential.

Statistical comparisons may not always be helpful for practical applications. For example, the two types of comparisons we made converged in identifying some meaningful differences (e.g. ‘% with psychiatric co-morbidity’), but they were not always consistent. For example, for the broad and inter- mediate algorithms, the standardised differences for ‘% male’ and ‘% rural’ were less than 0.1 (for the cohort vs. cohort comparison) and between 0.1 and 0.15 (for the cohort vs. added individuals comparison) – both less than a small effect size. However, the number of men is nearly twice the number of rural residents that were ‘misclassified’ (1050 vs. 550) by the broad algorithm. Whether these differences are important depends entirely on the context. If men or rural residents with IDD are heavy users of resource-intensive services and require specialised outreach or interventions to decrease such use, then even small discrepancies in planning estimates may have significant economic and social consequences. On the other hand, if the impact of ‘misclassification’ is minor, then these differences would be inconsequential. As noted earlier, a conventional threshold of <0.1 is used for standardised differences. However, Austin (2009) cautions that there is no clear consensus on what threshold to use.

The primary limitation of our study is the lack of external validation of our IDD cases. Other studies have dealt with this limitation in a variety of ways including accepting eligibility evaluations for IDD- related services at face value, incorporating multi- stage survey designs or employing multistage and labour intensive strategies to establish valid and reliable case finding (e.g. van Schrojenstein Lantman-de Valk et al. 2006). We could find only two published validation studies involving administrative health data, both limited to ASD. One by Dodds et al. (2009) used existing clinical information to validate the administrative data while the other (Van Naarden Braun et al. 2007) used more standard reabstraction methods but only on a sample of cases. This is clearly an issue that awaits further work and perhaps innovative thinking.

This limitation has some important implications. First, as noted earlier, health administrative data are invaluable for population level comparisons using broad diagnostic categories (e.g. IDD) but less effective for evaluating narrow diagnostic groupings (e.g. different types of IDD) (cf. Kisely et al. 2009a,b). A more appropriate vehicle would be a combination of administrative and clinical data, such as is used in standard validity studies. An important issue for this approach would include evaluating and making recommendations about the type and quality of clinical documentation needed to support the identification and care of individuals with IDD.

The second implication of the lack of external validation relates to our preference for the intermediate algorithm, particularly over the broad algorithm. This was based on two assumptions: that the prevalence of IDD is approximately 0.5 and that the broad definition is overly sensitive. As noted above, the rate of 0.5 was reported in a meta- analysis which did not include ASD (Maulik et al. 2011) and in a Canadian survey which did (Statistics Canada 2001). The literature provides mixed results for estimating what the prevalence of IDD might be. These differences are due to both variations in the definitions used and a lack of clarity about what the overlap is between ID and ASD (Fombonne 2009; Brugha et al. 2011; Pinborough- Zimmerman et al. 2011). Based on these mixed results, we conclude that we have chosen a conservative threshold under conditions where the true prevalence of IDD in adults is still uncertain.

The assumption that the broad definition is overly sensitive cannot be evaluated using administrative data alone. Minimally, both validation data and a clear understanding of the purpose of the analysis are required to establish the acceptable balance between sensitivity and specificity. The sole difference between the broad and intermediate algorithms is the inclusion of individuals who have only a single physician visit where IDD has been coded. Our physician billings data have features which increase the likelihood of false negatives: there are only two broad codes available for IDD (mental retardation; childhood psychoses, e.g. autism) and only one diagnostic field allowed per claim. In the absence of validation data, however, we have no way of assessing the magnitude of the false positives versus the false negatives. The study by Dodds et al. (2009) reported that autism algorithms which included single physician visits showed the expected pattern of a higher sensitivity but lower specificity than those which raised the threshold to two or more visits. Extrapolating these findings would suggest that our broad definition includes higher rates of false positives. However, comparable validity studies are not available for other IDD diagnoses or for the specific database we used. Future research as well as policy discussions would allow more definitive comment on this point.

A second limitation is that, despite the value of using health administrative data, they still do not identify all individuals with IDD and thus underestimate the true prevalence. Linkages with other sources of data may provide more accurate figures.

There are important benefits of using administrative health data in IDD research. They provide useful information for broad-based service planning and for population level comparisons around access and quality of care, particularly if the comparison groups are thoughtfully chosen. In terms of what algorithmic approach to use, our comparisons suggest an important value in going back to data- base inception, or at the very least, having a lengthy time window for identifying IDD cases as well as requiring a two-visit threshold for physician visits. However, it is important either to recognise that the proportion of persons with IDD and concomitant mental health problems may vary depending on the chosen algorithm or to try to supplement the administrative health data with more specialised information sources, a point that is also made by other authors (Dodds et al. 2009).

Acknowledgements

This study was funded by a Canadian Institutes of Health Research Partnerships in Health Systems Improvement grant (PHE# 103973) in partnership with the Ontario Ministry of Community and Social Services and Surrey Place Centre and is part of the Health Care Access Research in Developmental Disabilities (H-CARDD) Program. The work was also supported by the Centre for Addiction and Mental Health (CAMH) and the Institute for Clinical Evaluative Sciences (ICES), which receive annual grants from the Ontario Ministry of Health and Long Term Care (MOHLTC), as well as by CAMH and Fonds de Recherche en Sante du Quebec postdoctoral support for R. Balogh and V. Cobigo. The opinions, results and conclusions reported in this paper are those of the authors and are not necessarily those of the funding sources. No endorsement by CIHR, CAMH, ICES, the Ontario MOHLTC, or the Fonds de Recherche en Sante du Quebec is intended or should be inferred. The authors gratefully acknowledge the assistance and support of Jennifer Bennie at the Institute for Clinical Evaluative Sciences and Thivia Jegathesan for assistance with the literature search. In addition, the feedback from the anonymous reviewers was invaluable. Of particular note were their comments regarding age-related prevalence rates.

References

Adams E., Krahn G., Horner-Johnson W. & Leman R. (2009) Fundamentals of disability epidemiology, ch. 7. In: Disability and Public Health (eds C. E. Drum, G. L. Krahn & H. Bersani Jr), pp. 105–24. American Public Health Association & American Association on Intellectual and Developmental Disabilities, Washington, DC.

Austin P. C. (2009) Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Statistics in Medicine 28, 3083–107.

Balogh R., Brownell M., Ouellette-Kuntz H. & Colantonio H. (2010) Hospitalisation rates for ambulatory care sensitive conditions for persons with and without an intellectual disability—a population perspec- tive. Journal of Intellectual Disability Research 54, 820–32.

Balogh R. S., Hunter D. & Ouellette-Kuntz H. (2005) Hospital utilization among persons with an intellectual disability, Ontario, Canada, 1995–2001. Journal of Applied Research in Intellectual Disabilities 18, 181–90.

Brugha T. S., McManus S., Bankart J., Scott F., Purdon S., Smith J. et al. (2011) Epidemiology of autism spectrum disorders in adults in the community in England. Archives of General Psychiatry 68, 459–65.

Camp B. W., Broman S. H., Nichols P. L. & Leff M. (1998) Maternal and neonatal risk factors for mental retardation: defining the 'at-risk' child. Early Human Development 50, 159–73.

Centers for Disease Control and Prevention (2012) Prevalence of Autism Spectrum Disorders – Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. MMWR. Surveillance Summaries 61, 1–19.

Christianson A. L., Zwane M. E., Manga P., Rosen E., Venter A., Downs D. et al. (2002) Children with intellectual disability in rural South Africa: prevalence and associated disability. Journal of Intellectual Disability Research 46, 179–86.

Cohen J. (1992) A power primer. Psychological Bulletin 112, 155–9.

Cohen J. (1994) The earth is round (p [illegible] 1003). The American Psychologist 49, 99

Dodds L., Spencer A., Shea S., Fell D., Armson B. A., Allen A. C. et al. (2009) Validity of autism diagnoses using administrative health data. Chronic Diseases in Canada 29, 102–7.

Emerson E., Baines S., Allerton L. & Welch V. (2011) Health Inequalities & People with Learning Disabilities in the UK: 2011. Learning Disabilities Observatory, Manchester.

Fombonne E. (2009) Epidemiology of pervasive developmental disorders. Pediatric Research 65, 591–8.

Gershon A. S., Wang C., Guan J., Vasilevska-Ristovska J., Cicutto L. & To T. (2009a) Identifying individuals with physician diagnosed COPD in health administrative databases. COPD. Journal of Chronic Obstructive Pulmo- nary Diseases 6, 388–94.

Gershon A. S., Wang C., Guan J., Vasilevska-Ristovska J., Cicutto L. & To T. (2009b) Identifying patients with physician-diagnosed asthma in health administrative databases. Canadian Respiratory Journal 16, 183–8.

Government of Ontario (2010) Services and Supports to Promote the Social Inclusion of Persons with Developmental Disabilities Act, 2008. Government of Ontario, Toronto, Ontario.

Gustavson K.-H. (2005) Prevalence and aetiology of congenital birth defects, infant mortality and mental retardation in Lahore, Pakistan: a prospective cohort study. Acta Paediatrica 94, 769–74.

Hall A., Wood D., Hou T. & Zhang J. (2007) Patterns in primary health care utilization among individuals with intellectual and developmental disabilities in Florida. Intellectual and Developmental Disabilities 45, 310–22.

Heikura U., Taanila P., Olsen P., Hartikainen A.-L., von Wendt L. & Järvelin M.-R. (2003) Temporal changes in incidence and prevalence of intellectual disability between two birth cohorts in Northern Finland. American Journal on Mental Retardation 108, 19–31.

Hertz-Picciotto I. & Delwiche L. (2009) The rise in autism and the role of age at diagnosis. Epidemiology 121, 84–90.

Hosain G. M. M., Chatterjee N., Ara N. & Islam T. (2007) Prevalence, pattern and determinants of mental disorders in rural Bangladesh. Public Health 121, 18–24.

Human Resources and Skills Development Canada (2006) Disability in Canada: A 2006 Profile. Human Resources and Skills Development Canada, Gatineau, Quebec.

Hux J. E., Flintoft V., Ivis F. & Bica A. (2002) Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. Diabetes Care 25, 512–16.

Iezzoni L. I. (2002) Using administrative data to study persons with disabilities. The Milbank Quarterly 80, 347–79.

Juurlink D., Preyra C., Croxford R., Chong A., Austin P., Tu J. et al. (2006) Canadian Institute for Health Informa- tion Discharge Abstract Database: A Validation Study. Insti- tute for Clinical Evaluative Sciences, Toronto.

Kelly F., Kelly C. & Craig S. (2007) HRB Statistics Series: Annual Report of the National Intellectual Disability Data- base Committee 2007. National Intellectual Disability Database Committee, Dublin.

Khan Y., Glazier R., Moineddin R. & Schull M. J. (2011) A population-based study of the association between socioeconomic status and emergency department utiliza- tion in Ontario, Canada. Academic Emergency Medicine 18, 836–43.

Kisely S., Lin E., Lesage A., Gilbert C., Smith M., Campbell L. A. et al. (2009a) Use of administrative data for the surveillance of mental disorders in 5 provinces. Canadian Journal of Psychiatry 54, 571–5.

Kisely S., Lin E., Gilbert C., Smith M., Campbell L. A. & Vasilladis H. M. (2009b) Using administrative data for the surveillance of mood and anxiety disorders. Austral- ian and New Zealand Journal of Psychiatry 43, 1118–25.

Leonard H., Petterson B., Bower C. & Sanders R. (2003) Prevalence of intellectual disability in Western Australia. Paediatric and Perinatal Epidemiology 17, 58–67.

Lin E., Diaz-Granados N., Stewart D. E. & Bierman A. S. (2011) Postdischarge care for depression in Ontario. Canadian Journal of Psychiatry 56, 481–9.

Lunsky Y. & Balogh R. (2010) Dual diagnosis: a national study of psychiatric hospitalization patterns of people with developmental disability. Canadian Journal of Psychiatry 55, 721–8.

Lunsky Y., Lin E., Balogh R. & Klein-Geltink J. (2011) Diabetes prevalence among persons with serious mental illness and developmental disability. Psychiatric Services 62, 830.

Lunsky Y., Lin E., Balogh R., Klein-Geltink J., Wilton A. & Kurdyak P. (2012) Emergency department visits and use of outpatient physician services by adults with devel- opmental disability and psychiatric disorder.Canadian Journal of Psychiatry 57, 601–7.

Madsen K. M., Hviid A., Vestergaard M., Schendel D., Wohlfahrt J., Thorsen P. et al. (2002) A population based study of measles, mumps, and rubella vaccination and autism. The New England Journal of Medicine 347, 1477–82.

Martin L., Hirdes J. P. & Fries B. E. (2007) Examining the characteristics of persons with intellectual disability receiving hospital services: Part 1 – psychiatric hospitals/ units. Journal on Developmental Disabilities 13, 89–103.

Massey P. S. & McDermott S. (1996) State-specific rates of mental retardation – Unites States 1993. MMWR. Morbidity and Mortality Weekly Report 45, 61–5.

Maulik P. K., Mascarenhas M. N., Mathers C. D., Dua T. & Saxena S. (2011) Prevalence of intellectual disability: a meta-analysis of population-based studies. Research in Developmental Disabilities 32, 419–36.

McConkey R., Mulvany F. & Barron S. (2006) Adult persons with intellectual disabilities on the island of Ireland. Journal of Intellectual Disability Research 50, 227–36.

McGinnis K. A., Fine M. J., Sharma R. K., Skanderson M., Wagner J. H., Rodriguez-Barradas M. C. et al. (2003) Understanding racial disparities in HIV using data from the veterans aging cohort 3-site study and VA administrative data. American Journal of Public Health 93, 1728–33.

MENCAP (2007) Death by Indifference. MENCAP, London. Available at: http://www.mencap.org.uk/ campaigns/take-action/death-indifference (retrieved 27 February 2012).

Mercadante M. T., Evans-Lacko S. & Paula C. S. (2009) Perspectives of intellectual disability in Latin American countries: epidemiology, policy, and services for children and adults. Current Opinion in Psychiatry 22, 469–74.

Normand S.-L. T., Landrum M. B., Guadagnoli E., Ayanian J. Z., Ryan T. J., Clearly P. D. et al. (2001) Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. Journal of Clinical Epidemiology 54, 387–98.

Ouellette-Kuntz H., Garcin N., Lewis M. E. S., Minnes P., Martin C. & Holden J. J. A. (2005) Addressing health disparities through promoting equity for individuals with intellectual disability. Canadian Journal of Public Health 96, S8–S22.

Ouellette-Kuntz H., Coo H., Yu C. T., Chudley A. E., Noonan A., Breitenbach M. et al. (2006) Prevalence of pervasive developmental disorders in two Canadian provinces. Journal of Policy and Practice in Intellectual Disabilities 3, 164–72.

Ouellette-Kuntz H., Shooshtari S., Temple B., Brownell M., Burchill C., Yu C. T. et al. (2010) Estimating administrative prevalence of intellectual disabilities in Manitoba. Journal on Developmental Disabilities 15, 69–80.

Pinborough-Zimmerman J., Bilder D., Bakian A., Satter- field R., Carbone P. S., Nangle B. E. et al. (2011) Socio-demographic risk factors associated with autism spectrum disorders and intellectual disability. Autism Research 4, 438–48.

Public Health Agency of Canada (2009) Report from the National Diabetes Surveillance System: Diabetes in Canada, 2009. Public Health Agency of Canada, Ottawa. Avail- able at: http://www.phac-aspc.gc.ca/publicat/2009/ ndssdic-snsddac-09/pdf/report-2009-eng.pdf (retrieved 25 January 2012).

Schaible W. L. (1996) Indirect Estimators in U.S. Federal Programs, Lecture Notes in Statistics 108 (ed. W. L. Schaible). Springer, New York. Available at: https://catalyst.library.jhu.edu/catalog/bib_1940588 (retrieved 26 January 2012).

van Schrojenstein Lantman-de Valk H. M. J., Wullink M., van der Akker M., van Heurn-Nijsten E. W. A., Metse- makers J. F. M. & Dinant G. J. (2006) The prevalence of intellectual disability in Limburg, the Netherlands. Journal of Intellectual Disability Research 50, 61–8.

Schultz S. (2012) Discussion on Ontario Congestive Heart Failure Database [Conversation].

Spettell C. M., Wall T. C., Allison J., Calhoun J., Kobylinski R., Fargason R. et al. (2003) Identifying physician-recognized depression from administrative data: consequences for quality measurement. Health Services Research 38, 1081–102.

Statistics Canada (2001) Participation and Activity Limitation Survey (PALS).

Statistics Canada (2007) Standard Geographical Classification (SGC). Volume I. The Classification [Internet].

Statistics Canada, Ottawa, ON. Available at: http:// www.statcan.gc.ca/pub/12-571-x/12-571-x2006001-eng.pdf (retrieved 25 January 2012).

Statistics Canada (2012) CANSIM Table 051-0001 (Ontario Population: 18–64, 2009). Available at: http:// www5.statcan.gc.ca/cansim/pick-choisir (retrieved 29 February 2012).

Sullivan S. G., Glasson E. J., Hussain R., Petterson B. A., Slack-Smith L. M., Montgomery P. D. et al. (2003) Breast cancer and the uptake of mammography screening services by women with intellectual disabilities. Preventive Medicine 37, 507–12.

Tirschwell D. L. & Longstreth W. T. J. (2002) Validating administrative data in stroke research. Stroke 33, 2465– 70.

Tirschwell D. L., Kukull W. A. & Longstreth W. T. (1999) Medical complications of ischemic stroke and length of hospital stay: experience in Seattle, Washington. Journal of Stroke and Cerebrovascular Diseases 8, 336–43.

Tu J. V., Naylor D. C. & Austin P. (1999) Temporal changes in the outcomes of acute myocardial infarction in Ontario, 1992–1996. Canadian Medical Association Journal 161, 1257–61.

Tu K., Campbell N. R. C., Chen Z.-L., Cauch-Dudek K.

J. & McAlister F. A. (2007) Accuracy of administrative databases in identifying patients with hypertension. Open Medicine 1, E18–E26.

Tu K., Chen Z. & Lipscombe L. L.; for the Canadian Hypertension Education Program Outcomes Research Taskforce (2008) Prevalence and incidence of hypertension from 1995 to 2005: a population-based study Canadian Medical Association Journal 178, 1429–35.

Van Naarden Braun K., Pettygrove S., Daniels J., Miller L., Nicholas J., Baio J. et al. (2007) Evaluation of a methodology for a collaborative multiple source surveil- lance network for autism spectrum disorders – autism and developmental disabilities monitoring network, 14 sites, United States, 2002. MMWR. Surveillance Summa- ries: Morbidity and Mortality Weekly Report 56, 29–40.

Westerinen H., Kaski M., Virta L., Almqvist F. & Iivanainen M. (2007) Prevalence of intellectual disability: a comprehensive study based on national registers. Journal of Intellectual Disability Research 51, 715–25.

World Health Organization (2011) World Report on Disability. World Health Organization, Geneva.

Xie Z.-H., Bo S.-Y., Zhang X., Liu M., Zhang Z.-X., Ji S. R. et al. (2008) Sampling survey on intellectual disability in 0-6-year-old children in China. Journal of Intellectual Disability Research 52, 1029–38.