Validity of autism diagnoses using administrative health data

L. Dodds, PhD (1); A. Spencer, MSc (1); S. Shea, MD (2); D. Fell, MSc (1); B. A. Armson, MD (3) A. C. Allen, MD (1); S. Bryson, PhD (2)

Abstract

It is necessary to monitor autism prevalence in order to plan education support and health services for affected children. This study was conducted to assess the accuracy of administrative health databases for autism diagnoses. Three administrative health databases from the province of Nova Scotia were used to identify diagnoses of autism spectrum disorders (ASD): the Hospital Discharge Abstract Database, the Medical Services Insurance Physician Billings Database and the Mental Health Outpatient Information System database. Seven algorithms were derived from combinations of requirements for single or multiple ASD claims from one or more of the three administrative databases. Diagnoses made by the Autism Team of the IWK Health Centre, using state-of-the-art autism diagnostic schedules, were compared with each algorithm, and the sensitivity, specificity and C-statistic (i.e. a measure of the discrimination ability of the model) were calculated. The algorithm with the best test characteristics was based on one ASD code in any of the three databases (sensitivity = 69.3%). Sensitivity based on an ASD code in either the hospital or the physician billing databases was 62.5%. Administrative health databases are potentially a cost efficient source for conducting autism surveillance, especially when compared to methods involving the collection of new data. However, additional data sources are needed to improve the sensitivity and accuracy of identifying autism in Canada.

Introduction

The prevalence of autism spectrum disorders (ASDs) and autism, specifically, is reported to have been increasing over time.[1-4] If this is in fact correct, there would be major implications for the education system and agencies that provide services for these children: the availability of support and services will not match the increasing demands on the education system and health service providers. To date, there have been isolated efforts in Canada to estimate the prevalence of ASDs in some jurisdictions, but there are currently no systems in place to routinely monitor and report autism incidence and prevalence. Active surveillance of autism, conducted by population screening, provides excellent prevalence information, but is expensive and generally limited to short-term investigations.[1] Passive surveillance using existing databases provides a relatively inexpensive method to derive ongoing, population-based prevalence estimates.

The broad continuum of associated cognitive and neurobehavioural disorders, of which autism is the most extreme, are called pervasive developmental disorders (PDDs) or autism spectrum disorders (ASDs).[1,5] According to the diagnostic criteria of the International Classification of Diseases (ICD-10) by the World Health Organization (WHO), PDDs include childhood autism, atypical autism, Rett syndrome, other childhood disintegrative disorders, overactive disorders associated with mental retardation and stereotyped movements, Asperger's syndrome, other pervasive developmental disorders and unspecified pervasive developmental disorders. Childhood autism, atypical autism and Asperger's syndrome represent the more common diagnoses. In this study, we use the term ASD, which is equivalent to PDD, except that ASD does not include Rett syndrome and childhood disintegrative disorder, both of which are extremely rare.

In 1985, Bryson et al. made the first effort to estimate autism prevalence in Canada by screening all children (i.e. n = 20 800) aged 6 to 14 years in a specific geographic area of Nova Scotia and conducting follow-up diagnostic assessments for children who screened positive (i.e. n = 46).[6] Of the 46 children who screened positive, 21 children fell within the relatively narrow autism spectrum that was defined at the time (i.e. most, if not all, of whom would meet the more stringent criteria for autistic disorder).[7]

More recently, researchers in Canada have used existing data to estimate ASD prevalence. Ouellette-Kuntz et al. reported estimates of the prevalence of PDDs among children 15 years or younger during 2002 in the provinces of Prince Edward Island (PEI) and Manitoba.[8] In PEI, cases were identified by the Department of Social Services and Seniors and the Department of Education; parental consent was required for the researchers

Author References

1 Perinatal Epidemiology Research Unit, Departments of Obstetrics and Gynaecology and Pediatrics, Dalhousie University, Halifax, NS, Canada.
2 Department of Pediatrics, Dalhousie University, Halifax, NS, Canada.
3 Department of Obstetrics and Gynaecology, Dalhousie University, Halifax, NS, Canada.
Correspondence: Linda Dodds, PhD, Departments of Obstetrics and Gynaecology and Pediatrics, Dalhousie University, 5850/5980 University Avenue, P.O. Box 9700, Halifax, NS, Canada B3K 6R8, Tel.: 902-470-7191, Fax: 902-470-7190, Email: l.dodds@dal.ca

to collect the information. In Manitoba, cases were identified through referrals to the Children's Special Services program of the Department of Family Services and Housing. PDD prevalence rates among 1- to 15-year-olds in both provinces were similar (i.e. 2.84 per 1000 in Manitoba and 3.52 per 1000 in PEI). Fombonne et al. reported prevalence (of PDDs) based on a population of children registered at a large Anglophone school board in the Montreal area on October 1, 2003 (i.e. n = 27 749).[9] In Quebec, school boards submit information on children with PDDs and other disorders to the Ministry of Education in order to receive supplemental funding. In this 2003 survey, a total of 180 identified children had been diagnosed with a PDD (i.e. rate of 6.5 per 1000), 61 of whom were specifically diagnosed with autism.[9] In summary, surveillance and reports of autism prevalence in Canada are infrequent and variable rates have been reported.

To date, administrative health databases have not been used in Canada to estimate autism incidence or prevalence, although they have been used to estimate the incidence and prevalence of other conditions; e.g. algorithms have been developed and tested using administrative data for determining the incidence and prevalence of childhood asthma, osteoporosis, diabetes mellitus and diabetic macular edema.[10–13] In a study to evaluate the validity of ICD codes from administrative hospital discharge data, Quan et al.[14] compared ICD-9 and ICD-10 coding (i.e. the coding systems used in the administrative health databases) with medical chart data for 32 clinical conditions (ASD was excluded from the conditions assessed). They found that detection rates (e.g. sensitivity) varied by condition from 82% for renal failure to 9% for weight loss.[14]

Administrative health databases are a potential source for determining autism prevalence, but the validity of ASD diagnoses from administrative health data must be determined before these databases are used to measure the prevalence of autism in a population. Based on a cohort of children born in Nova Scotia between 1989 and 2002, we used administrative health databases linked to a "gold standard" clinical autism database to assess the accuracy of autism diagnoses ascertained from administrative health databases.

Methods

This study was based on data from a retrospective cohort study designed to examine prenatal, obstetrical and neonatal factors related to the development of autism. A cohort of all children born in Nova Scotia between 1989 and 2002 was identified from the Atlee Perinatal Database, i.e. a population-based database of all hospital births in Nova Scotia. The cohort of births was linked to the administrative health databases at the Population Health Research Unit at Dalhousie University. Data linkage was accomplished using encrypted health card numbers, common to all data sources. The cohort of children born between 1989 and 2002 were followed, by way of the administrative health databases, until December 2005.

For residents of Nova Scotia, as in the rest of Canada, access to hospital and physician services is universal within a system of publicly funded health care. For this study, three administrative health databases in Nova Scotia were used to identify diagnoses of autism spectrum disorders (ASD), i.e. the Hospital Discharge Abstract Database (available since 1989); the Medical Services Insurance (MSI) Physician Billings Database (available since 1989); and the Mental Health Outpatient Information System (MHOIS) Database (available since 1992). The Hospital Discharge Abstract Database includes diagnoses, which are noted in the medical chart and abstracted upon discharge. The MSI Physician Billing Database included a physician diagnostic code(s), which was sent to the provincial agency that handled payment for these insured services. The MHOIS Database was used for all outpatients seen in the mental health clinics and day patients in mental health day-treatment programs. Diagnoses were recorded by psychiatrists or psychologists, or both. An ASD diagnosis was defined from these administrative databases by an ICD-9 code 299 or an ICD-10 code F84 from any primary or secondary diagnostic field.

Seven algorithms were derived from combinations of requirements for single or multiple ASD claims from the three administrative databases. For example, in one algorithm, a child was considered to have an autism diagnosis if there was at least one autism code from the hospital discharge database; autism codes from the other databases were not required. The algorithm allowing for the most "hits" for an autism diagnosis was required for at least one ASD claim from any of the three aforementioned databases.

"Gold standard" diagnoses were obtained from a clinical database generated by the Autism Team of the IWK Health Centre. Referrals to the Autism Team were made largely by health care professionals and some teachers in the Halifax Regional Municipality to assess children with suspected autism. The IWK Autism Team consisted of pediatricians, psychologists, social workers, psychiatrists, speech-language pathologists, occupational therapists and nurses. Final determination of diagnoses was made by psychologists and/or pediatricians or psychiatrists, who led or co-led the diagnostic teams and was based on the Autism Diagnostic Interview – Revised, the Autism Diagnostic Observation Schedule and clinical judgment using DSM-IV-TR.[15–17] These instruments and criteria were consistent with recommended practice parameters for diagnosing ASDs.[18,19] Diagnoses made by the Autism Team, considered the "gold standard," were recorded in a database starting in 2001.

The linkage between the Atlee Perinatal Database, the administrative health databases and the "gold standard" data was accomplished using a multi-step procedure to ensure anonymity. The first step was the creation of a "cross-walk file," which included a unique number assigned to all individuals in each of the databases, along with their encrypted health card number. A third party used a sophisticated algorithm to encrypt health card numbers, assigned to every individual in the province and a common field in each data source). Finally, the requested variables from each file were linked back to the "cross-walk file," using the unique encrypted number

TABLE 1
Comparison of algorithms[1] using combinations of autism spectrum disorder (ASD) diagnoses from three administrative health databases compared to a "gold standard" diagnosis

Type of administrative data			Comparison of results to "gold standard"				Test characteristics of algorithms		
Hospital data (# of times ASD coded)	Physician billing data (# of times ASD coded)	Mental health outpatient data (# of times ASD coded)	# True positives	# True negatives	# False positives	# False negatives	Sensitivity	Specificity	C-statistic
≥ 1			21	86	2	155	11.9%	97.7%	0.55
	≥ 1		105	75	13	71	59.7%	85.2%	0.72
		≥ 1	29	81	7	147	16.5%	92.0%	0.54
≥ 1	≥ 1	≥ 1	122	68	20	54	69.3%	77.3%	0.76
≥ 1	≥ 1		110	73	15	66	62.5%	83.0%	0.74
≥ 1	≥ 2	≥ 2	75	78	10	101	42.6%	88.6%	0.67
≥ 1	≥ 2		65	82	6	111	36.9%	93.2%	0.65

[1] Algorithms based on autism code(s) from more than one database indicates that an autism diagnosis was assigned if an autism code was used in either of the databases indicated.

assigned to the individuals in each database, and a linked, anonymous analysis file containing data elements from each data source was generated.

Diagnoses of children assessed by the "gold standard data" (i.e. the IWK Autism Team) from 2001 to 2005 were compared to ASD diagnoses from each of the seven algorithms, based on the administrative health databases. The accuracy of each algorithm was evaluated by calculating the sensitivity, specificity and a C-statistic (i.e. a nonparametric estimate of the area under a receiver operating characteristic curve that provides a measure of a method's ability to predict an autism diagnosis). C-statistic scores range from 1.0 for a "perfect" test with a sensitivity and specificity of 100%, to 0.5 for a method that was unable to discriminate.[20]

For the true ASD cases that were missed by the administrative databases (i.e. false negatives [FN]), codes for other psychological conditions were examined. In addition, codes that occurred both before and after the date of the true (i.e. "gold standard") diagnosis were evaluated. Various factors were evaluated for those patients who had an autism code in one of the administrative databases, but who were not given an ASD diagnosis after assessment by the Autism Team (i.e. false positive [FP]). These included the number of incorrect claims; the years when these ASD claims occurred, whether the incorrect claims occurred after the IWK negative diagnosis date; and whether there had been other claims made in relation to psychological conditions. Sensitivity and specificity rates were compared for maternal and infant factors, such as low birth weight and maternal age (available from the Atlee Perinatal Database), to determine if certain characteristics were associated with the accuracy of autism diagnoses based on administrative health data.

Approval for this study was obtained by the Research Ethics Board of the IWK Health Centre.

Results

The IWK Autism Team evaluated 270 patients linked to the overall study cohort of children born in Nova Scotia. According to the team's assessment, there were 176 confirmed ASD cases and 88 non-cases (i.e. 6 had undetermined diagnoses and were dropped from further analysis). All remaining 264 children had at least 2 years of administrative data available following the date of their birth. When seen by the Autism Team, 58% of the children were 4 years or younger; only 12% of the children were 10 years or older when the team saw them. The majority of confirmed cases were coded with a general diagnosis of ASD, without any specific autism diagnosis noted.

Table 1 shows the definition of each of the seven algorithms tested, along with the sensitivity, specificity and C-statistic associated with each algorithm. The algorithm with the highest C-statistic (i.e. 0.76), the highest sensitivity (i.e. 69.3%) and a specificity of 77.3% was the algorithm that defined an ASD diagnosis by at least one claim in any of the three administrative databases. Using this algorithm, 190 of the 264 children were correctly diagnosed. There were 20 FPs and 54 FNs, which were examined in more detail to help explain the inaccuracies in the administrative databases.

An examination was made of the 54 FN children diagnosed with ASD by the Autism Team, but who did not have an ASD claim in any of the three databases, to see if other claims might have been systematically recorded instead of ASD. Of the 54 FNs, 46 children had at least one MSI physician billing claim for neurotic disorders, personality disorders and other non-psychotic mental disorders (i.e. ICD-9 codes 300-316). Of these 46 children, 35 (i.e. 76%) had an ICD-9 code of 315 (i.e. "specific delays in development") coded at least once. This code occurred in 22 children before the Autism Team diagnosis date and in 26 children after; some children had an ICD-9 code of 315 before and after the Autism Team diagnosis date.

TABLE 2
Comparison of the number of autism spectrum disorder (ASD) claims per child among false positives and true positives

		False positives	True positives
Database	Frequency of ASD claims per child	# children (%) (n = 20)	# children (%) (n = 122)
Hospital	1 or more	2 (10%)	21 (17%)
	2 or more	1 (5%)	7 (6%)
MSI	1 or more	13 (65%)	104 (85%)
	2 or more	4 (20%)	55 (45%)
MHOIS	1 or more	7 (35%)	29 (24%)
	2 or more	7 (35%)	23 (19%)
Any of 3 databases	1 or more	20 (100%)	122 (100%)
	2 or more	11 (55%)	74 (61%)

The number of ASD claims from each of the three databases was compared between the 20 FP children and the 122 TP children (see Table 2). For the 20 FPs, 2 children (i.e. 10%) had ASD coded from Hospital Discharge Data, 13 (i.e. 65%) from MSI Physician Billing Data and 7 (i.e. 35%) from MHOIS Data (see Table 2). Among the 13 subjects from the FP group with one or more ASD claims from the Physician Billing Database, 4 of 13 (i.e. 31%) had more than one ASD claim in the Physician Billing Database, compared to 55 of 104 (i.e. 53%) of the true positives. Among the MHOIS claims for the FP group, all had more than one MHOIS claim for ASD. Of the 122 TPs, 21 (i.e. 17%) had hospital claim(s), 104 (i.e. 85%) had MSI claim(s) and 29 (i.e. 24%) had MHOIS claim(s); 27 (i.e. 22%) had claims from 2 databases and 5 (i.e. 4%) had claims from all 3 databases (data not shown). While most ASD claims from the hospitalization and MHOIS databases occurred after the Autism Team diagnosed TPs, 55 of 104 (i.e. 53%) of children had MSI claim(s) before this date. Other than ASD codes, the most common code used was ICD-9-CM 315 ("specific delays in development"), which was recorded equally before and after the Autism Team diagnosis date.

Sensitivity and specificity values were compared according to maternal and neonatal characteristics (see Table 3). The sensitivity of the administrative data in identifying an ASD diagnosis was similar across most factors, including for males (i.e. 69.7%) and females (i.e. 66.7%). The sensitivity of the administrative data in identifying an ASD diagnosis was not significantly lower for children with a major congenital anomaly (i.e. 55.6%) compared to children without an anomaly (i.e. 69.9%). The sensitivity was not significantly higher among children outside of Halifax County compared to residents of Halifax County (i.e. 75.0% versus 68.1%), although specificity was lower (i.e. 66.7% versus 80.0%, respectively).

Discussion

In the current study, we used codes from three administrative health databases to evaluate multiple algorithms for their accuracy in identifying autism among children in Nova Scotia. Although the overall study cohort included all children born in Nova Scotia between 1989 and 2002, only children seen by the Autism Team (between 2001 and 2005) who linked to the study cohort were included in this validation study. Based on the algorithm defining autism by at least a single claim in any one of the hospitalizations, the physician billing or the outpatient mental health databases, the ability of administrative health databases in Nova Scotia to correctly identify children with autism was moderately successful (i.e. sensitivity of 69%). Most of the true ASD cases who were incorrectly identified within the administrative data (i.e. FNs) had codes indicating some other non-psychotic psychological disorder or developmental delay, suggesting that physicians may have been reluctant to use an autism code before an autism diagnosis was verified.

A strength of this study was the quality of the autism diagnosis in the "gold standard" population. However, the "gold standard" diagnosis was limited to children who were referred to the Autism Team. It should be noted that children in this validation study without an ASD diagnosis when assessed by the Autism Team would have had some behavioural and/or developmental feature that warranted referral to the Autism Team. Therefore, the false positive rate observed in this study is likely higher, and the specificity lower, than it would have been had we been able to establish a "gold standard" diagnosis for all children in the administrative databases. Nevertheless, the specificity we observed was reasonably high (i.e. 77%), an estimate which is likely below the true specificity. Other algorithms tested in this study had more stringent requirements for defining autism (e.g. two physician claims required), and therefore had better specificity than the one-claim algorithm, albeit at the expense of reduced sensitivity.

In order to improve the detection rate observed in this study, other data sources would be required. In Canada, information on ASD diagnoses is available from regional school boards in some areas or from some provincial Departments of Social Services or Family Services, as previously discussed. The use of education data sources (i.e. alone or in conjunction with clinical data) and data from other government-administered programs have been used to identify autism cases in the United States. The Centers for Disease Control and Prevention have established a multi-source surveillance network for ASD and other developmental disabilities.[21]

Children 8 years of age with ASD who reside within one of the 16 states comprising part of the network area were identified in a two-phase process. First, children suspected of having an ASD were identified through screening and abstraction of records from multiple sources within clinical and education records. In phase two, the abstracted behavioural data were scored

TABLE 3
Comparison of sensitivity and specificity of autism spectrum disorder (ASD) diagnoses using administrative data compared to "gold standard" diagnoses, according to maternal and neonatal factors

Factor	Number	Sensitivity (95% CI)	Specificity (95% CI)
Maternal age			
< 35	218	71.2% (63.2 to 78.1)	77.2% (66.8 to 85.2)
≥ 35	46	62.2% (46.1 to 76.0)	77.8% (44.3 to 94.7)
County of residence			
Halifax County	214	68.1% (60.0 to 75.1)	80.0% (69.1 to 87.8)
Outside Halifax	50	75.0% (57.7 to 87.0)	66.7% (43.6 to 83.9)
Birth weight			
< 2500 g	15	72.7% (42.9 to 90.8)	75.0% (28.9 to 96.6)
≥ 2500 g	234	69.1% (61.7 to 75.7)	77.4% (67.3 to 85.1)
Major congenital anomaly*			
Yes	13	55.6% (26.6 to 81.2)	100% (45.4 to 100)
No	247	69.9% (62.5 to 76.4)	76.5% (66.2 to 84.5)
Sex			
Male	231	69.7% (62.0 to 76.4)	80.3% (69.8 to 87.8)
Female	33	66.7% (45.2 to 83.0)	58.3% (31.9 to 80.7)
Birth order			
First born	145	69.8% (60.0 to 78.1)	77.6% (64.0 to 87.1)
Second or higher	119	68.8% (57.9 to 77.9)	76.9% (61.5 to 87.6)

* A major anomaly is defined as a defect of structure or function that is present at birth and affects length of life, impacts quality of life or requires surgery.

by clinicians to determine whether they met the ASD case definition. The rates varied somewhat between sites, with an overall mean prevalence rate of 6.6 per 1000 eight-year-old children.[22] Extensive quality assurance activities were incorporated into the network to maximize data quality and consistency.

Newschaffer et al. used a national source of administrative data (i.e. the United States Department of Education, Office of Special Education Programs) to examine trends in ASD between 1992 and 2001. However, limitations of these data were noted, in particular with the specific classification of impairment and the likelihood of underestimating autism prevalence based on special education data alone.[23] In California, individuals with autism (and other conditions) are eligible to receive services through the Department of Developmental Services. Eligibility is based on diagnoses provided by qualified health care professionals. Croen et al.[24] used these data to estimate autism prevalence. They suspected that their observed prevalence of 12.3 per 10 000 children for the years 1987 to 1994 was an underestimation, since approximately 20% to 25% of the children who were eligible to receive services were not enrolled in the program.[24]

In Canada, all provinces and territories have administrative data that include hospitalizations and physician visits. In Nova Scotia, the addition of an outpatient mental health database increased the sensitivity of ASD diagnoses by about 7%, compared to the sensitivity using only hospitalization and claim data regarding physician visits. On the other hand, the specificity increased by about 6% when the mental health outpatient data were excluded. Since relatively few children were hospitalized for (or with) autism (i.e. 12% of the true cases had an autism code from the hospitalization data), this source, by itself, was inadequate to determine autism diagnoses in a population. However, an autism diagnosis in the hospitalization database was very likely correct. Although we explored ICD codes that were used other than ASD codes, their use was too inconsistent to suggest an algorithm that would improve the false positive or false negative rates.

Research or surveillance of health conditions using administrative health databases has advantages over other data collection methods. Administrative health databases are available in all Canadian provinces and territories and provide a source for a large number of population-based cases, likely at a lower cost than would be possible with newly collected data. In addition, diagnoses are entered into the databases without knowledge of underlying exposure-outcome hypotheses. However, there are limitations to using administrative data, particularly with respect to the accuracy of diagnoses that are being used for billing purposes (as is the case with the Physician Billing Database).

Given that we measured maximum sensitivity at 69%, it is likely that administrative health data alone would underestimate the true incidence and prevalence, as observed in this study. This would suggest that additional data sources are necessary to enhance the detection rate of ASD diagnoses from existing databases, since it is unlikely that a single source of administrative data will provide a complete accounting of all autism cases in Canada. Although challenging, the jurisdictions should work together toward acquiring standard data from multiple sources to enable ongoing, passive surveillance of ASDs in Canada.

Acknowledgements

This project was funded by Cure Autism Now (i.e. now Autism Speaks). Linda Dodds was supported by the Clinical Research Scholar Award from Dalhousie University and the New Investigator Award from the Canadian Institutes of Health Research. The authors thank the Reproductive Care Program of Nova Scotia and the Population Health Research Unit of Dalhousie University for data access. Although this research is based partially on data obtained from the Population Health Research Unit, the observations and opinions expressed are those of the authors and do not represent those of the Population Health Research Unit.

References

1. Bryson SE, Smith IM. Epidemiology of autism: prevalence, associated characteristics, and implications for research and service delivery. Ment Retard Dev Disabil Res Rev. 1998;4:97–103.

2. Fombonne E. The prevalence of autism. JAMA. 2003;289:87–9.

3. Fombonne E. Epidemiology of autistic disorder and other pervasive developmental disorders. J Clin Psychiatry. 2005;66 Suppl 10:3–8.

4. Williams JG, Higgins JP, Brayne CE. Systematic review of prevalence studies of autism spectrum disorders. Arch Dis Child. 2006;91:8–15.

5. Filipek PA, Accardo PJ, Baranek GT, et al. The screening and diagnosis of autistic spectrum disorders. J Autism Dev Disord. 1999;29:439–84.

6. Bryson SE, Clark BS, Smith IM. First report of a Canadian epidemiological study of autistic syndromes. J Child Psychol Psychiatry. 1988;29:433–45.

7. American Psychiatric Association. DSM IV diagnostic and statistical – manual. 4th ed. Washington (D.C.): American Psychiatric Association; 1994.

8. Ouellette-Kuntz H, Coo H, Yu CT, Chudley AE, Noonan A, Breitenbach M, et al. Prevalence of pervasive developmental disorders in two Canadian provinces. J Appl Res Intellect Disabil. 2006;3:164–72.

9. Fombonne E, Zakarian R, Bennett A, Meng L, McLean-Heywood D. Pervasive developmental disorders in Montreal, Quebec, Canada: prevalence and links with immunizations. Pediatrics. 2006;118: e139–50.

10. To T, Dell S, Dick P, Cicutto L, Harris J, Tassoudji M, Duong-Hua M. Burden of childhood asthma. Toronto, Ontario: ICES, 2004.

11. Lix LM, Yogendran MS, Leslie WD, Shaw SY, Baumgartner R, Bowman C, et al. Using multiple data features improved the validity of osteoporosis case ascertainment from administrative databases. J Clin Epidemiol. 2008;61:1250-60.

12. Hux JE, Ivis F, Flintoft V, Bica A. Determination of prevalence and incidence using a validated administrative data algorithm. Diab Care. 2002;25:512–6.

13. Bearelly S, Mruthyunjaya P, Tzeng JP, Suner IJ, Shea AM, Lee JT, et al. Identification of patients with diabetic macular edema from claims data. Arch Ophthalmol. 2008;126:986–9.

14. Quan H, Li B, Saunders D, Parsons GA, Nilsson CI, Alibhai A, et al. Assessing validity of ICD-9CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. Health Serv Res. 2008;43:1424–41.

15. Lord C, Rutter M, Le Couteur A. Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord. 1994;24:659–85.

16. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal EL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord. 2000;30:205–23.

17. American Psychiatric Association. Diagnostic and statistical manual. 4th ed. Washington (D.C.): American Psychiatric Association; 2000.

18. Filipek PA, Accardo PJ, Baranek GT, Cook EH, Dawson G, Gordon B, et al. The screening and diagnosis of autistic spectrum disorders. J Autism Dev Disord. 1999;29:439–84.

19. Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EH, Dawson G, et al. Practice parameter: screening and diagnosis of autism: report of the quality standards subcommittee of the American academy of neurology and the child neurology society. Neurology. 2000;55:468–79.

20. Ash A, Shwartz M. R^2: A useful measure of model performance when predicting a dichotomous outcome. Stat Med. 1999; 18:375–84.

21. Rice CE, Baio J, Van Naarden Braun K, Doernberg N, Meaney FJ, Kirby RS. A public health collaboration for the surveillance of autism spectrum disorders. Paediatr Perinat Epidemiol. 2007;21:179–90.

22. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2002 Principal Investigators; Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders – autism and developmental disabilities monitoring network, 14 sites, United States, 2002. MMWR Surveill Summ. 2007 Feb 9; 56(1):12–28.

23. Newschaffer CJ, Falb MD, Gurney JG. National autism prevalence trends from United States special education data. Pediatrics. 2005;115:e277–82.

24. Croen LA, Grether JK, Selvin S. Descriptive epidemiology of autism in a California population: who is at risk? J Autism Dev Disord. 2002;32:217–24.