
Application of a national administrative case definition for the identification of pre-existing diabetes mellitus in pregnancy

V. M. Allen, MD (1,2); L. Dodds, PhD (1,2,3,4); A. Spencer, MSc (4); E. A. Cummings, MD (3); N. MacDonald, MD (3); G. Kephart, PhD (2)

This article has been peer reviewed.

Abstract

Introduction: Accurate ascertainment of pregnant women with pre-existing diabetes allows for the comprehensive surveillance of maternal and neonatal outcomes associated with this chronic disease.

Method: To determine the accuracy of case definitions for pre-existing diabetes mellitus when applied to a pregnant population, a cohort of women who were pregnant in Nova Scotia, Canada, between 1991 and 2003 was obtained from a population-based provincial perinatal database, the Nova Scotia Atlee Perinatal Database (NSAPD). Person-level data from administrative databases using hospital discharge abstract data and outpatient physician services data were linked to this cohort. Various algorithms for defining diabetes mellitus from the administrative data, including the algorithm suggested by the National Diabetes Surveillance System (NDSS), were compared to a reference standard definition from the NSAPD.

Results: Validation of the NDSS case definition applied to this pregnant population demonstrated a sensitivity of 87% and a positive predictive value (PPV) of 66.4%. Use of ICD-9 and ICD-10 diagnostic codes among hospitalizations with diabetes mellitus in pregnancy showed important increases in sensitivity and PPV, especially for those pregnancies delivered in tertiary centres. In this population, pregnancy-related administrative data from the hospitalization database alone appear to be a more accurate data source for identifying pre-existing diabetes than applying the NDSS case definition, particularly when pregnant women are delivered in a tertiary hospital.

Conclusion: Although the NDSS definition of diabetes performs reasonably well compared to a reference standard definition of diabetes, using this definition for evaluating maternal and perinatal outcomes associated with diabetes in pregnancy will result in a certain degree of misclassification and, therefore, biased estimates of outcomes.

Keywords: *diabetes mellitus, pregnancy, validation studies*

Introduction

Monitoring the prevalence and incidence of diabetes, estimating the burden of illness, and evaluating the impact of care on prevention and progression are

essential for planning and evaluating treatment and prevention programs for chronic disease.^{1,2} Increasing maternal age³ and changing maternal characteristics such as pre-pregnancy weight^{4,5} may contribute to increasing rates of pre-existing diabetes

in pregnant women,⁶ with associated increased costs related to diabetes care.⁷ Obstetrical complications associated with pre-existing diabetes also have important maternal and neonatal consequences.^{6,8-11} Accurate ascertainment of pregnant women with pre-existing diabetes allows for the comprehensive surveillance of maternal and neonatal outcomes associated with this complication.

The identification of diabetes cases in the population using administrative data began in Canada in 1991, followed by the development of a provincial diabetes database in Manitoba in 1998.¹² To be labelled as having diabetes, a person must have recorded two physician claims within a two-year period or one hospitalization with a diagnosis of diabetes. With further refinement related to age threshold, and clarification of the claim date, the National Diabetes Surveillance System (NDSS) established an algorithm for the collection of national data related to diabetes.¹³ The NDSS is a collaborative network of provincial and territorial surveillance systems. Supported by the Public Health Agency of Canada, it was developed in 2001 to improve the breadth of information about the burden of diabetes in Canada so that policy makers, public health and health care professionals and the general public can make better public and personal health decisions. At the provincial level, the NDSS compiles administrative health care data relating to individual diabetes cases and sends aggregate anonymous data to the Public Health Agency of Canada for national analyses.¹⁴

Author references:

1. Department of Obstetrics and Gynaecology, Dalhousie University, Halifax, Nova Scotia, Canada
2. Department of Community Health and Epidemiology, Dalhousie University, Halifax, Nova Scotia, Canada
3. Department of Pediatrics, Dalhousie University, Halifax, Nova Scotia, Canada
4. Perinatal Epidemiology Research Unit, Dalhousie University, Halifax, Nova Scotia, Canada

Correspondence: Victoria M. Allen, Department of Obstetrics and Gynaecology, IWK Health Centre, Room G2141, 5850/5980 University Avenue, Halifax, NS B3K 6R8; Tel.: (902) 470-6486; Fax: (902) 425-1125; Email: victoria.allen@dal.ca

Validation studies have evaluated the NDSS case definition for the detection of diabetes mellitus in the general population as well as a more liberal definition that requires only one physician visit or hospitalization. They have compared these definitions to reference standards such as provincial health surveys, diabetes registries, medical charts and laboratory data.^{2,15} The results of these validation studies have been inconsistent as measured by ascertainment rate¹² or sensitivity,¹⁵⁻¹⁹ which may be related to regional, temporal or reference standard definitions.^{12,15-19} In addition, the utility of the NDSS in sub-populations such as pregnancy becomes challenging when diagnoses of gestational diabetes are also considered. The NDSS is limited in its ability to distinguish between type 1, type 2 and gestational diabetes, and although gestational diabetes has its own ICD-9* and ICD-10† codes, increases in the prevalence of pre-existing diabetes among women of child-bearing years (20-49 years) may be the result of changing maternal characteristics, such as increasing pre-pregnancy weight,⁴ or the result of miscoding.^{14,20} To eliminate gestational diabetes cases that were miscoded with a diabetes mellitus code, the NDSS case definition excludes women first meeting the case definition for diabetes 120 days preceding or 90 days after any pregnancy-related visit.

Previous validation studies used to develop a Nova Scotia diabetes repository demonstrated an unacceptably high number of false positive diagnoses of diabetes mellitus using the NDSS case definition in the general population.²¹ The purpose of our study was to evaluate the application of the NDSS case definition for diabetes mellitus using data derived from administrative databases to a population of pregnant women, and to compare this application to a clinical definition for the diagnosis of pre-existing diabetes in pregnancy using data derived from a reference standard perinatal database.

Methods

The province of Nova Scotia has a homogeneous, predominantly Caucasian population of about one million, with

approximately 10 000 live births each year.²² The population of Nova Scotia has universal health coverage with a single payer health system within Canada. Although nine hospitals offer intrapartum obstetrical care, 50% of deliveries occur at one of the tertiary maternity facilities.

Data sources and linkage

Information on all women who delivered in Nova Scotia between 1988 and 2003 is available from the Nova Scotia Atlee Perinatal Database (NSAPD), which is managed by the Reproductive Care Program (RCP) of Nova Scotia. The NSAPD is a high quality, provincial population-based database containing clinical information on all births born at a gestational age of at least 20 weeks or having a birth weight of at least 500 grams. It contains maternal and newborn information, such as demographic variables, procedures, interventions, maternal and newborn diagnoses and morbidity, and mortality information for every pregnancy and birth in Nova Scotia since 1988. Home births without hospital admission are currently not entered into the database. (However, there are few home births, approximately 0.2% per year.) Information in the database is abstracted by trained health records personnel using standardized forms and hospital medical records across Nova Scotia. Detailed information on several hundred variables is collected on specific lifestyle and other subject characteristics, medical conditions, labour and delivery events and neonatal outcomes. All information is entered into the database soon after the time of collection. In addition to the routine data checks and edits that are made at the time of collection, an ongoing data quality-assurance program, which carries out periodic rigorous abstraction studies, has shown that the information in the database continues to be reliable. In particular, the information collected on pre-existing diabetes was considered the reference standard for the diagnosis of diabetes for this study because cases were clinically confirmed and accurately coded.²¹ The database has been used previously for numerous studies,

including diabetes-related studies,^{9,23,24} and has been used to validate other sources of data.²⁵

The data from the NSAPD were linked to the two administrative health databases relevant to this study, the Canadian Institute for Health Information's Discharge Abstract Database (CIHI-DAD) and the Medical Services Insurance (physician visits) Database (MSID). The administrative databases are housed at the Population Health Research Unit (PHRU), Department of Community Health and Epidemiology, Dalhousie University, and include population-level administrative health data for the Province of Nova Scotia. These administrative data are obtained from provincial billing information (MSI) and recorded from hospital medical records abstracted by trained health records personnel. Nine facilities provide regional or tertiary level obstetrical services in Nova Scotia; in 6 of these facilities, the data abstractor who codes and abstracts information is the same coder for both CIHI-DAD and NSAPD data, while in 3 facilities, the data are collected for the CIHI system and the NSAPD system by two different individuals. Each data abstractor is registered with the Canadian Health Information Management Association and is qualified for and knowledgeable about data collection in either system for data collection.

These health databases capture all diagnosed cases of diabetes in both outpatient and hospital settings. The CIHI-DAD contains information on medical diagnoses and procedures from hospital discharge data. Discharges are coded using ICD-9 codes for 1987 to 2001 and ICD-10 codes since 2001; surgical and other procedures are coded using the Canadian Classification of Procedures for 1987 to 2001 and the Canadian Classification of Health Interventions for 2001 to the present day. The MSID records outpatient visits and diagnoses through physician billing, including information on physician specialty. In Nova Scotia, clinical fees for obstetrical services are coded separately for prenatal visits, admission to hospital, care for labour and delivery, and postpartum care.

* *International Classification of Diseases, 9th Revision.*

† *International Classification of Diseases, 10th Revision.*

NSAPD (reference-standard) criteria

Since inception, the NSAPD has defined pre-existing diabetes in pregnancy using the White classification, then ICD-10-CA (Canada) or CCI (Canadian Classification of Health Interactions) code, and finally the NSAPD code for diabetes mellitus during a pregnancy-related admission (Table 1a).²⁶ However, during this study, we used only the White classification, which considers duration of diabetes and the presence of vascular, retinal and renal complications of diabetes mellitus.²⁷ Pre-existing diabetes in pregnancy is coded in the NSAPD when it is identified anywhere in the patient record, regardless of whether the diagnosis was of an outpatient or an inpatient. The NSAPD case definition is able to distinguish gestational diabetes mellitus (White classification, Class A) from pre-existing type 1 or type 2 diabetes mellitus (White classification Class B-T).

NDSS criteria

The case definition used by the NDSS for the diagnosis of diabetes mellitus requires that an individual have either at least one hospitalization or at least two medical claims coded with a diagnosis of diabetes mellitus (250 in ICD-9 or E10–E14 in ICD-10, Table 1b) within two years (Algorithm A, Table 2). To meet the two MSI physician claims requirements, the claims could not occur on the same day. These case definitions are applied to all patient-level claims, irrespective of age or gestational status.²⁰ To eliminate miscoding of gestational diabetes cases as diabetes mellitus, and because birth date information is not available to NDSS, the NDSS case definition distinguishes pre-existing diabetes mellitus from gestational diabetes by removing any cases with a diagnostic code for diabetes mellitus (Table 1b) that occur 120 days before or 90 days after any pregnancy-related visit (relevant obstetrical claims codes summarized in Table 1b). The NDSS case definition includes type 1 and type 2 diabetes mellitus but is unable to distinguish between types.

The NDSS case definition was applied to this study population using all coding fields in the hospital discharge data for any hospital admission in pregnancy and

TABLE 1A
Diagnostic codes used to define pregnancies with pre-existing diabetes in the NSAPD

Diagnostic codes	Year of use
White Classification	1988–2003
Class A: Gestational diabetes	
Class B: Less than 10 years duration; no vascular disease; onset after age 20 years	
Class C: Duration 10–19 years; minimal vascular disease; onset after age 10 years	
Class D: Duration 20 years or more; benign retinopathy; onset before age 10 years	
Class F: Patient with Class D and nephropathy	
Class R: Patient with proliferative retinopathy	
Class T: Diagnosis made by level of glucose challenge test equal to or greater than 10.3 mmol/l	
ICD-10-CA or CCI	2003–2006
NSAPD code	2006–present

Abbreviations: CCI, Canadian Classification of Health Interactions; ICD-10-CA, *International Classification of Diseases, 10th Revision, Canadian version*; NSAPD, Nova Scotia Atlee Perinatal Database.

TABLE 1B
Diagnostic codes used in the application of the National Diabetes Surveillance System case definition for pre-existing diabetes to pregnancies in the NSAPD

	Diabetes mellitus codes	Obstetrical codes	Diabetes mellitus in pregnancy codes
ICD-9 ^a	250	640–669	648.0
ICD-10 ^b	E10–E14	O265, O290–O30, O318, O320–O369, O40–O439, O60–O669, O680–O849, O890–O899, O904, O908, O95–O97, Z354–Z356	O24.0, O24.1, O24.2, O24.3, O24.9

Abbreviations: ICD-9, *International Classification of Diseases, 9th Revision*; ICD-10, *International Classification of Diseases, 10th Revision*; NSAPD, Nova Scotia Atlee Perinatal Database.

^a In use 1987–2001.

^b In use 2001–present day.

TABLE 2
Algorithms based on the application of the NDSS case definition for diabetes mellitus to pregnancies in the NSAPD, or existing diagnostic codes for diabetes in pregnancy using CIHI-DAD and MSI

Algorithm	Definition
A (NDSS)	removes cases with at least one hospitalization or at least two MSI claims with a diagnostic code for diabetes mellitus (ICD-9 250 or ICD-10 E10–14) that is followed within 120 days, or 90 days after, by an obstetrics claims code
B	removes cases with at least one hospitalization or at least two MSI claims with a diagnostic code for diabetes mellitus (ICD-9 250 or ICD-10 E10–14) that is followed within 150 days, or 90 days after, by an obstetrics claims code
C	removes cases with at least one hospitalization or at least three MSI claims with a diagnostic code for diabetes mellitus (ICD-9 250 or ICD-10 E10–14) that is followed within 120 days, or 90 days after, by an obstetrics claims code
D	includes cases with at least one hospitalization with a diagnostic code for diabetes mellitus during pregnancy (ICD-9 648.0 or ICD-10 O24.0, O24.1, O24.2, O24.3, O24.9)
E	algorithm A or algorithm D

Abbreviations: CIHI-DAD, Canadian Institute for Health Information's Discharge Abstract Database; ICD-9, *International Classification of Diseases, 9th Revision*; ICD-10, *International Classification of Diseases, 10th Revision*; MSI, Medical Services Insurance; NDSS, National Diabetes Surveillance System; NSAPD, Nova Scotia Atlee Perinatal Database.

physician claims data for the study period. The NSAPD began coordinating data collection in Nova Scotia in 1988; however, since the databases at the PHRU only included data beginning April 1, 1989, and to allow for a two-year period before a pregnancy (NDSS case definition includes two medical claims codes within two years), the study period included all pregnancies between April 1, 1991, and December 2003. In Nova Scotia, ICD-9 was replaced by ICD-10 in the CIHI-DAD, beginning April 1, 1997; however, MSI coding with respect to billing continued in ICD-9.

Statistical analysis

All pregnancies in the NSAPD (i.e. ≥ 20 weeks gestation and births ≥ 500 g) between April 1, 1991, and April 1, 2003, that resulted in live-birth singletons were considered for analysis. Because diabetes status in pregnancy may change over time, only the diabetes status for a nulliparous pregnancy recorded in the NSAPD was considered. A pregnant woman must also have been eligible to receive MSI (i.e. did not move out of province or die) during a period at least two years before the start of pregnancy to 90 days after the delivery date to ensure sufficient time to meet the NDSS case definition. Since the administrative databases began in April 1, 1989, the earliest delivery date was April 1, 1991. The delivery date, which was determined from the NSAPD, was not used in the administrative definitions, but was used only to place the cases in the appropriate time intervals. Only patient obstetrical and diabetes records were retained.

Analyses compared the reference standard to two modifications of the application of the NDSS case definition to this pregnant population (Algorithm B and C, Table 2) and two alternate definitions using administrative databases (Algorithm D and E, Table 2). Algorithm B removed diabetes mellitus claims followed within 150 days by an obstetrical claim (instead of 120 days used in the NDSS definition) to identify a case of pre-existing diabetes mellitus. Algorithm C required three (not two) MSI physician claims (within two years) or one hospital claim to identify a case of pre-existing diabetes mellitus. Algorithm D

was defined by using only hospitalizations with ICD-9 and ICD-10 diagnostic codes specific for diabetes mellitus in pregnancy (Table 1b) for the duration of the study (1991–2003) from the CIHI-DAD, because the fourth digit for the ICD-9 code 648, which distinguishes pre-existing diabetes in pregnancy (ICD-9 648.0) from gestational diabetes (ICD-9 648.8), only became available for MSI physician claims in Nova Scotia on March 31, 1996. Algorithm E required either the application of the original NDSS case definition to this pregnant population or at least one hospitalization with ICD-9 and ICD-10 diagnostic codes specific for diabetes mellitus in pregnancy (based on the model with either algorithm A or D). The analysis was also done for each algorithm separating the cohort into those who delivered in tertiary hospitals from those that delivered in non-tertiary (regional or community) hospitals, and also into two time periods, before and after April 1, 1997.

The development and maintenance of study databases such as those used in this study is consistent with the Tri-Council's guidelines pertaining to database linkages under their Code of Ethical Conduct for Research Involving Humans. This research project received approval from the IWK Health Centre Research Ethics Board and from the Joint Data Access Committee of the RCP of Nova Scotia.

Results

Linkage of the NSAPD and the administrative databases housed by the PHRU, which included the CIHI-DAD and MSI, yielded 41 533 nulliparous pregnancies in the NSAPD with corresponding hospitalization and outpatient physician visit administrative codes. There were 8.4% less women in the PHRU eligibility file compared to the data file derived from the NSAPD.

Table 3 summarizes the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the evaluation of each algorithm of the NDSS case definition used in comparison with the reference standard diagnosis. During the study period, 200 women with pre-existing diabetes mellitus were identified using

the NSAPD, while 262 women who met inclusion criteria were identified as having pre-existing diabetes mellitus by applying the NDSS case definition (Algorithm A). The estimated prevalence in the study population was 0.48% (95% CI 0.42–0.55) using data from the NSAPD and 0.63% (95% CI 0.56–0.71) using administrative data based on the NDSS case definition (31% higher prevalence in the NDSS). Based on the reference standard, the NDSS case definition of diabetes had a sensitivity of 87.0%, specificity of 99.8%, PPV of 66.4% and NPV of 99.9%. There were 88 pregnancies falsely identified as having pre-existing diabetes when applying the NDSS case definition compared to the reference standard. Of these 88 pregnancies, 22 were diagnosed by the NSAPD as having gestational diabetes. The NSAPD does not collect information on glucose intolerance in pregnancy.

The consequences of variations in components of the NDSS case definition are summarized in Table 3. Modification of the NDSS case definition to remove women diagnosed with diabetes 150 days preceding any pregnancy-related visit (Algorithm B) slightly reduced the number of false positive cases of diabetes mellitus to 84, but also decreased the sensitivity to 84.5% compared to the NSAPD. Modification of the NDSS case definition removing women with diabetes codes 120 days preceding or 90 days after any pregnancy-related visit using three MSI codes (Algorithm C) reduced the number of pregnancies falsely identified as having diabetes mellitus to 50, with a concurrent slight decrease in sensitivity to 82.5% compared to the NSAPD.

When testing the algorithm that used diagnostic codes specifically for diabetes mellitus in pregnancy among hospitalizations (Algorithm D), we found that 228 pregnancies had at least one hospitalization with a diagnostic code for diabetes mellitus in pregnancy during the study period, while 200 pregnancies were identified by the NSAPD as having pre-existing diabetes mellitus. Compared to the reference standard, 51 pregnancies were falsely identified with diabetes mellitus, with a sensitivity of 88.5%, specificity of 99.9%, PPV of 77.6% and NPV of 99.9% (Table 3).

When testing Algorithm E (based on the model with either algorithm A or D), sensitivity was increased to 92.0% but PPV was decreased to 60.1% (Table 3).

Categorization by type of delivery hospital (tertiary versus non-tertiary) showed higher sensitivity and PPV for deliveries within a tertiary hospital (n = 26 165) for all of the algorithms compared to all pregnancies combined or to deliveries in non-tertiary hospitals (n = 15 368; Table 3). The best performance was for Algorithm D for

deliveries in a tertiary hospital, where the sensitivity was 98.0% and PPV was 82.0%. Poorer performance was seen when applied to deliveries in non-tertiary hospitals. The prevalence was lower than the entire study population when deliveries occurred in non-tertiary hospitals (0.31%) and higher when deliveries occurred in tertiary hospitals (0.59%).

Categorization by period (before April 1, 1997, n = 20 993, or equal to or later than April 1, 1997, n = 20 540) showed a very

slightly higher sensitivity in the later time period, but poorer PPV (71.3% in the earlier time period versus 63.4% in the later time period), with application of the NDSS case definition (algorithm A) (Table 4). For all the other algorithms, only small differences were observed in the test characteristics between the two time periods (Table 4). The prevalence was lower than the entire study population when deliveries occurred in the first period (0.40%) and higher when deliveries occurred in the second (0.57%).

TABLE 3
Test characteristics of the NDSS case definition compared to the NSAPD (reference standard) for nulliparous pregnancies, by type of delivery hospital, Nova Scotia, 1991–2003

Algorithm		Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)
A (NDSS)	All pregnancies	87.0 (81.4–91.1)	99.8 (99.7–99.8)	66.4 (60.3–72.0)	99.9 (99.9–99.9)
	Delivery in a non-tertiary hospital	72.3 (57.1–84.0)	99.8 (99.7–99.8)	46.6 (35.0–58.6)	99.9 (99.9–100)
	Delivery in a tertiary hospital	91.5 (85.6–95.2)	99.8 (99.6–99.8)	74.1 (67.1–80.0)	99.9 (99.9–100)
B	All pregnancies	84.5 (78.6–89.1)	99.8 (99.8–99.9)	66.8 (60.6–72.5)	99.9 (99.8–100)
	Delivery in a non-tertiary hospital	70.2 (54.9–82.2)	99.8 (99.8–99.8)	46.5 (34.7–58.6)	99.9 (99.9–100)
	Delivery in a tertiary hospital	88.9 (82.6–93.2)	99.8 (99.7–99.8)	74.7 (67.7–80.7)	99.9 (99.9–99.9)
C	All pregnancies	82.5 (76.4–87.4)	99.9 (99.8–99.9)	76.7 (70.4–82.1)	99.9 (99.8–99.9)
	Delivery in a non-tertiary hospital	68.1 (52.8–80.1)	99.9 (99.8–99.9)	68.1 (52.8–80.5)	99.9 (99.9–100)
	Delivery in a tertiary hospital	86.9 (80.3–91.6)	99.9 (99.8–99.9)	82.6 (75.7–88.0)	99.9 (99.9–100)
D	All pregnancies	88.5 (83.1–92.4)	99.9 (99.8–99.9)	77.6 (71.6–82.8)	99.9 (99.9–100)
	Delivery in a non-tertiary hospital	57.4 (42.3–71.4)	99.9 (99.8–99.9)	60.0 (44.4–73.9)	99.9 (99.8–99.9)
	Delivery in a tertiary hospital	98.0 (93.9–99.5)	99.9 (99.8–99.9)	82.0 (75.5–87.1)	100.0 (100–100)
E	All pregnancies	92.0 (87.1–95.2)	99.7 (99.7–99.8)	60.1 (54.4–65.6)	100.0 (99.9–100)
	Delivery in a non-tertiary hospital	72.3 (57.1–83.9)	99.6 (99.5–99.7)	38.2 (28.3–49.2)	99.9 (99.9–100)
	Delivery in a tertiary hospital	98.0 (93.9–99.5)	99.7 (99.7–99.8)	69.1 (62.5–75.1)	100.0 (100–100)

Abbreviations: CI, confidence interval; NDSS, National Diabetes Surveillance System; NPV, negative predictive value; NSAPD, Nova Scotia Atlee Perinatal Database; PPV, positive predictive value.

TABLE 4
Test characteristics of the NDSS case definition compared to the NSAPD (reference standard) for nulliparous pregnancies, by period, Nova Scotia, 1991–2003

Algorithm		Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)
A (NDSS)	Delivery before April 1, 1997	85.7 (76.0–92.1)	99.9 (99.8–99.9)	71.3 (61.3–79.6)	99.9 (99.9–100)
	Delivery April 1, 1997, or later	87.9 (80.3–93.0)	99.7 (99.7–99.8)	63.4 (55.4–70.7)	99.9 (99.9–100)
B	Delivery before April 1, 1997	81.0 (70.6–88.4)	99.9 (99.5–99.6)	71.6 (61.3–80.1)	99.9 (99.9–100)
	Delivery April 1, 1997, or later	87.1 (79.3–92.3)	99.7 (99.6–99.8)	63.9 (55.9–71.3)	99.9 (99.9–100)
C	Delivery before April 1, 1997	79.8 (69.3–87.4)	99.9 (99.8–99.9)	79.8 (69.3–87.4)	99.9 (99.9–100)
	Delivery April 1, 1997, or later	84.5 (76.3–90.3)	100.0 (99.9–100)	74.8 (66.3–81.8)	99.9 (99.9–100)
D	Delivery before April 1, 1997	82.1 (71.9–89.3)	99.9 (99.8–99.9)	72.6 (62.4–81.0)	99.9 (99.9–100)
	Delivery April 1, 1997, or later	93.1 (86.4–96.8)	99.9 (99.8–99.9)	81.2 (73.3–87.3)	99.9 (99.9–100)
E	Delivery before April 1, 1997	89.3 (80.2–94.7)	99.8 (99.7–99.8)	61.5 (52.2–70.0)	100.0 (99.9–100)
	Delivery April 1, 1997, or later	94.0 (87.5–97.3)	99.6 (99.5–99.7)	59.2 (51.8–66.3)	100.0 (99.9–100)

Abbreviations: CI, confidence interval; NDSS, National Diabetes Surveillance System; NPV, negative predictive value; NSAPD, Nova Scotia Atlee Perinatal Database; PPV, positive predictive value.

Discussion

Accurate identification of a population with diabetes quantifies the burden of disease, but also contributes to the evaluation of disease management and outcomes associated with diabetes. Studies employing the NDSS case definition for the diagnosis of diabetes mellitus in the general population have demonstrated high ascertainment rates with the original case definition,¹² but have improved at estimating incidence by adding clearance periods to minimize the inclusion of prevalent cases,¹⁸ by modifying the number of hospitalizations or physician visits in the NDSS criteria,^{2,15} or by adding clinical data to the original case definition.^{2,16,17,19} The application of the NDSS case definition to a subpopulation such as pregnancy is challenging. We demonstrated that applying the NDSS case definition to a pregnant population underestimated true cases of pre-existing diabetes mellitus (sensitivity 87%) and a high number of false positive cases (PPV 66%). The prevalence of pre-existing diabetes mellitus among pregnant women in Nova Scotia was 0.5% using the reference standard, lower than the general female population in Canada as identified by the NDSS (0.7%–2.5% in 2006–2007 in women of child-bearing age).¹⁴ Grouping by type of delivery hospital increased both the sensitivity and the PPV for those delivering in a tertiary hospital, but usually resulted in poorer results for non-tertiary hospitals. The algorithm employing only hospitalization diagnostic codes for pre-existing diabetes in pregnancy (Algorithm D) among women delivering in tertiary centres performed the best, with sensitivity 98%, specificity 99.9%, PPV 82%, and NPV 100% compared to the reference standard (NSAPD). However, using an algorithm which excludes both outpatients and non-tertiary hospitals would limit the province-wide assessment of diabetes in pregnancy needed for programming and making policy decisions.

The false positive cases identified by applying the NDSS case definition to the studied pregnant population in Nova Scotia may reflect coding errors or misdiagnosis, such as coding glucose intolerance as diabetes in administrative data. The low PPV suggests a high potential for

misclassifying non-diabetic individuals as having diabetes mellitus. Implications of this misclassification become apparent when potential uses of the administrative data are considered. For an outcome study on the effect of pre-existing diabetes on birth outcomes, this degree of misclassification would be a major source of bias; if an administrative definition of pre-existing diabetes was used as part of a risk adjustment in a study in pregnancy examining an additional risk factor, then the misclassification would result in residual confounding. However, since the prevalence of pre-existing diabetes mellitus is small (0.5%–0.6%), the residual confounding resulting from misclassification would be small from the perspective of absolute numbers of misclassified women. In addition, if the administrative definition was used descriptively to measure the prevalence of pre-existing diabetes in pregnancy, the degree to which the misclassification biases the prevalence estimates should be taken into account. The influence of misclassification bias in understanding results using large administrative databases was recently highlighted in a cohort evaluation of the identification of diabetes mellitus in Ontario.²⁸ The authors emphasized the need for verifying the accuracy following the mass application of identification criteria to minimize misclassification bias, compared to regularly validated data collection employed by electronic databases such as the NSAPD.

Evaluation of the NDSS case definition applied to a population of pregnant women using the NSAPD demonstrated higher sensitivity and PPV for women who delivered at tertiary centres compared to those delivered at non-tertiary hospitals. Pregnancies complicated by severe diabetes mellitus may be preferentially delivered at a tertiary maternity facility, introducing severity bias into the assessment of the NDSS case definition. This difference in level of hospital for delivery may also represent variability in coding practice among centres. Additionally, specialists and subspecialists involved in the care of pregnant women with diabetes may be more likely to accurately code for pre-existing diabetes mellitus than general practitioners, as has been demonstrated

with other systemic diseases outside of pregnancy and where medical care unrelated to the disease is required.^{29,30}

A switch from the ICD-9 to ICD-10 coding system occurred in April 1997 in Nova Scotia, but despite this factor, the study interval did not affect the operating characteristics. This observation may be a result of improved coding as coders gained experience with new coding systems,³¹ balanced with increasing reimbursement for medical services using alternate funding programs established by the government of Nova Scotia. This change in the environment of funding, with a variable requirement for shadow billing to document clinical care, may have resulted in decreased accuracy in coding. In particular, there was a growth in alternate funding programs in the tertiary care centres in Nova Scotia in later years of the study. Alshammari and Hux demonstrated that detection of chronic disease is more likely with hospitalization, but that these diseases are less likely to be detected in surveillance programs dependent on administrative data algorithms in non-fee-for-service settings.¹⁶ Chronic diseases such as diabetes mellitus are treated largely on an outpatient basis, and surveillance efforts are heavily dependent on outpatient physician services claims. In the NDSS, nearly 75% of cases are detected by physician claims alone.¹⁶ For this population of women who delivered in Nova Scotia hospitals, the hospitalization code for pregnancies complicated by pre-existing diabetes is a more accurate method for identifying diabetes mellitus than the NDSS case definition. The addition of this hospitalization code (for pregnancies complicated by pre-existing diabetes) to the NDSS definition for the general population should increase sensitivity; however PPV may decrease.

The validation of administrative databases typically occurs with medical record audits, and results of validation studies have varied depending on the type of administrative data (inpatient versus outpatient and diagnostic versus procedural), specific disease area and codes used for case identification, and disease severity.³² Some provinces continue to use three-digit coding, which may decrease the PPV. Employing administrative databases for

the ascertainment of diagnoses is challenging in light of varying coding practices and the accuracy and comprehensiveness of data sources.^{33,34} The NSAPD is a validated database²⁵ and has been used to validate perinatal data in the CIHI-DAD.³⁵ It is considered a reference standard component of the Nova Scotia Diabetes Repository.²¹ The population-based nature of the administrative databases and the NSAPD in this study limits the selection bias that may occur with single-centre validation studies. In addition, health surveillance increases in pregnancy, reducing rates of undiagnosed diabetes.^{9,23} It would be important to validate the NDSS case definition using other perinatal database sources to rule out a regional bias in the comparison population,²⁸ and to assess coding quality and the coding environments in different provinces and regions.²⁹

A limitation to this validation study includes the introduction of new provincial health card numbers after 1995. Before this time, women had their father's (if under 18) or their husband's (if married) social insurance number plus a suffix, for a health card number (HCN), while after 1995 they were assigned their own HCN. Both the PHRU and the RCP have optimized mapping of the old to the new HCN; however, there may be occurrences where mapping is incomplete. This would lead to women appearing in the data as left censored or lost to follow-up when the old MSI number is changed, and would underestimate the prevalence of pre-existing diabetes.

The PPV of case definitions derived from administrative data is highly dependent on the prevalence of diabetes mellitus in the population to which they are applied. Accordingly, as prevalence decreases, more stringent case definitions are required in order to have an acceptable PPV.²⁷ In this study, the hospitalization code for pregnancies complicated by pre-existing diabetes performed the best. Other modifications could include the addition of clinical or laboratory data to improve detection.¹⁷

Conclusion

Validation of the NDSS case definition using the NSAPD as the reference standard diagnosis demonstrated adequate sensitivity but low positive predictive values. In the Nova Scotia pregnant population, administrative data using the ICD-9 and ICD-10 codes for diabetes mellitus in pregnancy from the CIHI-DAD (hospitalization database) alone appear to be a more accurate data source for the identification of pre-existing diabetes than the application of the NDSS case definition, particularly when pregnant women are delivered in a tertiary hospital. Although the NDSS definition of diabetes performs reasonably well compared to a reference standard definition of diabetes, using this definition for evaluating maternal and perinatal outcomes associated with diabetes in pregnancy will result in a certain degree of misclassification and, therefore, biased estimates of outcomes.

Acknowledgements

This study was funded by the Canadian Institutes of Health Research (CIHR). VMA and LD were supported by a Clinical Research Scholar Award from Dalhousie University and the New Investigator Award from CIHR. 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. The authors thank the Diabetes Care Program of Nova Scotia (DCPNS) for assistance with the methodology.

References

1. Saydah SH, Geiss LS, Tierney E, Benjamin SM, Engelgau M, Brancati F. Review of the performance of methods to identify diabetes cases among vital statistics, administrative, and survey data. *Ann Epidemiol.* 2004;14:507-16.

2. Southern DA, Roberts B, Edwards A, Dean S, Norton P, Svenson LW, et al. Validity of administrative data claim-based methods for identifying individuals with diabetes at a population level. *Can J Public Health.* 2010;101:61-4.
3. Joseph KS, Allen AC, Dodds L, Turner LA, Scott H, Liston R. The perinatal effects of delayed childbearing. *Obstet Gynecol.* 2005;105:1410-8.
4. Joseph KS, Young DC, Dodds L, O'Connell CM, Allen VM, Chandra S, et al. Changes in maternal characteristics and obstetric practice and recent increases in primary cesarean delivery. *Obstet Gynecol.* 2003;102:791-800.
5. Robinson HE, O'Connell CM, Joseph KS, McLeod NL. Maternal outcomes in pregnancies complicated by obesity. *Obstet Gynecol.* 2005;106:1357-64.
6. Pridjian G. Pregestational diabetes. *Obstet Gynecol Clin North Am.* 2010;37:143-58.
7. Johnson JA, Pohar SL, Majumdar SR. Health care use and costs in the decade after identification of type 1 and type 2 diabetes: a population-based study. *Diabetes Care.* 2006;29:2403-8.
8. Feig DS, Razzaq A, Sykora K, Hux JE, Anderson GM. Trends in deliveries, prenatal care, and obstetrical complications in women with pregestational diabetes: a population-based study in Ontario, Canada, 1996-2001. *Diabetes Care.* 2006;29:232-5.
9. Yang J, Cummings EA, O'Connell C, Jangaard K. Fetal and neonatal outcomes of diabetic pregnancies. *Obstet Gynecol.* 2006;108:644-50.
10. Rosenberg TJ, Garbers S, Lipkind H, Chiasson MA. Maternal obesity and diabetes as risk factors for adverse pregnancy outcomes: differences among 4 racial/ethnic groups. *Am J Public Health.* 2005;95:1545-51.
11. Jensen DM, Damm P, Moelsted-Pedersen L, Ovesen P, Westergaard JG, Moeller M, Beck-Nielsen H. Outcomes in type 1 diabetic pregnancies: a nationwide, population-based study. *Diabetes Care.* 2004;27:2819-23.

12. Blanchard JF, Ludwig S, Wajda A, Dean H, Anderson K, Kendall O, et al. Incidence and prevalence of diabetes in Manitoba, 1986-1991. *Diabetes Care*. 1996;19:807-11.
13. Clotley C, Mo F, LeBrun B, Mickelson P, Niles J, Robbins G. The development of the National Diabetes Surveillance System (NDSS) in Canada. *Chron Dis Can*. 2001(2);22:67-9.
14. Report from the National Diabetes Surveillance System: Diabetes in Canada, 2009 [Internet]. Ottawa (ON): Public Health Agency of Canada; 2009. Available from: <http://www.phac-aspc.gc.ca/publicat/2009/ndssdic-snsddac-09/pdf/report-2009-eng.pdf>
15. Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care*. 2002;25:512-6.
16. Alshammari AM, Hux JE. The impact of non-fee-for-service reimbursement on chronic disease surveillance using administrative data. *Can J Public Health*. 2009;100:472-4.
17. Chen G, Khan N, Walker R, Quan H. Validating ICD coding algorithms for diabetes mellitus from administrative data. *Diabetes Res Clin Pract*. 2010;89:189-95.
18. Asghari S, Couteau J, Carpentier AC, Vanasse A. Optimal strategy to identify incidence of diagnostic of diabetes using administrative data. *BMC Med Res Methodol*. 2009;9:62.
19. Ho ML, Lawrence N, van Walraven C, Manuel D, Keely E, Malcolm J, et al. The accuracy of using integrated electronic health care data to identify patients with undiagnosed diabetes mellitus. *J Eval Clin Pract*. 2011 Feb 17. doi: 10.1111/j.1365-2753.2011.01633.x.
20. National Diabetes Surveillance System. Responding to the challenge of diabetes in Canada: first report of the NDSS, 2003. Ottawa (ON): Health Canada; 2003 [cited 2010 Nov 22]. Available from: http://www.phac-aspc.gc.ca/ccdpc-cpcmc/ndss-snsd/english/pubs_reports/pdf/WEB_NDSS_English_Report-nocover.pdf
21. Diabetes Care Program of Nova Scotia. Development of a Nova Scotia Diabetes Repository: Provincial Report. Halifax (NS): Diabetes Care Program of Nova Scotia; Aug 2009.
22. Statistics Canada. 2006 Community profiles [Internet]. Ottawa (ON): Statistics Canada; 2006 [modified 2011 Jun 06; cited 2010 Nov 25]. Available at: <http://www12.statcan.ca/census-recensement/2006/dp-pd/index-eng.cfm>
23. McMahon MJ, Ananth CV, Liston RM. Gestational diabetes mellitus. Risk factors, obstetric complications and infant outcomes. *J Reprod Med*. 1998;43:372-8.
24. Russell C, Dodds L, Armson BA, Kephart G, Joseph KS. Diabetes mellitus following gestational diabetes: role of subsequent pregnancy. *BJOG*. 2008;115:253-60.
25. Fair M, Cyr M, Allen AC, Wen SW, Guyon G, MacDonald RC. An assessment of the validity of a computer system probabilistic record linkage of birth and infant death records in Canada. *Chronic Dis Can*. 2000(1);21:8-13.
26. Reproductive Care Program of Nova Scotia. The Nova Scotia Atlee Perinatal Database [Internet]. Halifax (NS): RCP; [cited 2011 Jul 17]. Available at: <http://rcp.nshealth.ca/atlee-database>
27. Gabbe SG, Niebyl JR, Simpson JL, editors. *Obstetrics: normal and problem pregnancies*. 5th ed. Philadelphia (PA): Churchill Livingstone; 2007.
28. Manuel DG, Rosella LC, Stukel TA. The importance of accurately identifying disease in studies using electronic health records. *BMJ*. 2010;341:c4226. doi: 10.1136/bmj.c4226.
29. Myers RP, Shaheen AA, Fong A, Wan AF, Swain MG, Hilsden RJ, et al. Validation of coding algorithms for the identification of patients with primary biliary cirrhosis using administrative data. *Can J Gastroenterol*. 2010;24:175-82.
30. Farrokhyar F, McHugh K, Irvine EJ. Self-reported awareness and use of the International Classification of Diseases coding of inflammatory bowel disease services by Ontario physicians. *Can J Gastroenterol*. 2002;16:519-26.
31. Quan H, Li Bing, Saunders LD, Parsons GA, Nilsson CI, Alibhai A, et al.; IMECCHI Investigators. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. *Health Serv Res*. 2008;43:1424-41.
32. Myers RP, Leung Y, Shaheen AAM, Li B. Validation of ICD-9-CM/ICD-10 coding algorithms for the identification of patients with acetaminophen overdose and hepatotoxicity using administrative data. *BMC Health Serv Res*. 2007;7:159.
33. Leslie WD, Lix LM, Yogendran MS. Validation of a case definition for osteoporosis disease surveillance. *Osteoporos Int*. 2011;22:37-46. doi: 10.1007/s00198-010-1225-2.
34. Juurlink D, Preyra C, Croxford R, Chong A, Austin P, Tu J, et al. Canadian Institute for Health Information Discharge Abstract Database: a validation study [Internet]. Toronto (ON): Institute for Clinical Evaluative Sciences; 2006 [cited 2011 Jul 17]. Available at: http://www.ices.on.ca/file/CIHI_DAD_Reabstractors_study.pdf
35. Joseph KS, Fahey J. Validation of perinatal data in the Discharge Abstract Database of the Canadian Institute for Health Information. *Chronic Dis Can*. 2009(3);29:96-100.