Can J Diabetes 37 (2013) 249–253

Contents lists available at SciVerse ScienceDirect

Canadian Journal of Diabetes

journal homepage: www.canadianjournalofdiabetes.com

Original Research

Concordance Between Self-Report and a Survey-Based Algorithm for Classification of Type 1 and Type 2 Diabetes Using the 2011 Population-Based Survey on Living With Chronic Diseases in Canada (SLCDC)-Diabetes Component

Edward Ng PhD [a,*], Saskia E. Vanderloo MSc [b], Linda Geiss MA [c], Jeffrey A. Johnson PhD [d]

[a] *Health Analysis Division, Statistics Canada, Ottawa, Ontario, Canada*
[b] *Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada*
[c] *Division of Diabetes Translation, US Centre for Disease Control, Atlanta, Georgia, United States of America*
[d] *School of Public Health, University of Alberta, Edmonton, Alberta, Canada*

ARTICLE INFO

Article history:
Received 12 October 2012
Received in revised form
24 May 2013
Accepted 25 May 2013

Keywords:
Surveillance
survey
type 1 diabetes
type 2 diabetes

ABSTRACT

Objective: There are 2 major forms of diabetes mellitus: types 1 and 2. A major limitation of most current population-based diabetes surveillance systems is the classification of diabetes types. Our objective was to examine the concordance of self-reported diabetes type with a previously developed classification algorithm, using a nationally representative survey sample.
Methods: Self-reported data were available from 2 544 adults with self-reported diabetes, aged ≥20 years and older, who responded to the diabetes component of the 2011 Survey of Living with Chronic Diseases in Canada. We examined the concordance of self-reported diabetes type with an algorithm based on self-reported, but objective, respondent characteristics, such as age of diagnosis and treatment patterns. Concordance was measured using kappa coefficients. Sensitivity, specificity and positive and negative predictive values were calculated using the algorithm as the reference "standard."
Results: Approximately 11% of the estimated population did not self-report diabetes type; almost all of these respondents would be classified as having type 2 diabetes by the algorithm. Of those self-reporting diabetes type, we found moderate overall agreement between the algorithm and self-reported type (kappa, 0.52; 95% confidence interval [CI], 0.52 to 0.53). Perfect agreement was noted in the youngest age group (kappa, 1.0; 95% CI, 1.0-1.0) but agreement was poor for the oldest age group (kappa, 0.20; 95% CI, 0.19 to 0.20).
Conclusions: An algorithm based on self-reported, objective characteristics related to diabetes diagnosis and treatment patterns may have the potential to overcome limitations of simple self-report diabetes type for the classification of diabetes type in older adults.

Mots clés :
Surveillance
enquête
diabète de type 1
diabète de type 2

RÉSUMÉ

Objectif : Il existe 2 formes principales de diabète sucré : le type 1 et le type 2. Un obstacle important à la plupart des systèmes de surveillance habituels du diabète dans la population est la classification des types de diabète. Notre objectif était d'examiner la concordance du type de diabète déclaré à un algorithme de classification élaboré antérieurement au moyen d'un échantillon représentatif sur le plan national.
Méthodes : Les données déclarées disponibles provenaient de 2 544 adultes âgés de 20 ans et plus ayant un diabète déclaré, qui avaient répondu au volet sur le diabète de l'Enquête sur les personnes ayant une maladie chronique au Canada de 2011. Nous avons examiné la concordance du type de diabète déclaré à l'aide d'un algorithme basé sur les caractéristiques déclarées, mais objectives, des répondants, telles que l'âge au diagnostic et les profils thérapeutiques. La concordance a été mesurée à l'aide des coefficients Kappa. La sensibilité, la spécificité et les valeurs prédictives positives et négatives ont été calculées en utilisant l'algorithme comme étalon de référence.

* Address for correspondence: Edward Ng, PhD, 24R, R.H. Coats Building, Statistics Canada, 100 Tunney's Pasture Drway, Ottawa, Ontario K1A 0T6 Statistics Canada.
E-mail address: edward.ng@statcan.gc.ca.

1499-2671/$ – see front matter © 2013 Canadian Diabetes Association
http://dx.doi.org/10.1016/j.jcjd.2013.05.007

Résultats : Approximativement 11 % de la population estimée n'a pas déclaré le type de diabète; presque tous ces répondants pourraient être classifiés comme ayant le diabète de type 2 par l'algorithme. Parmi ceux ayant déclaré le type de diabète, nous avons observé une concordance globale modérée entre l'algorithme et le type déclaré (Kappa, 0,52; intervalle de confiance [IC] à 95 %, 0,52-0,53). Une concordance parfaite a été notée dans le groupe d'âge le plus jeune (Kappa, 1,0; IC à 95 %, 1,0-1,0), mais la concordance a été médiocre dans le groupe d'âge le plus âgé (Kappa, 0,20; IC à 95 %, 0,19-0,20).
Conclusions : Un algorithme basé sur les caractéristiques objectives déclarées en lien avec le diagnostic du diabète et les profils thérapeutiques peut avoir le potentiel de surmonter les obstacles de la simple déclaration du type de diabète par la classification du type de diabète chez les adultes plus âgés.

Introduction

Diabetes is currently the sixth leading cause of death in Canada (1) and, in 2010, approximately 1.8 million Canadians aged 12 years or older reported that they had been diagnosed with the disease (2). There are several forms of diabetes: type 1 and type 2 are the most common. Although all types of diabetes are characterized by the body's inability to maintain appropriate blood sugar levels, they differ in their causes, treatments and complications. Differences in etiology, associated risk factors, management and prevention strategies for type 1 and type 2 diabetes make it important for public health surveillance to distinguish diabetes type when monitoring incidence and prevalence.

In Canada, both administrative and survey data sets have been used to monitor diabetes trends. The Canadian Chronic Disease Surveillance System (CCDSS) uses administrative data to monitor trends of diabetes. Built on the model of the former National Diabetes Surveillance System, the CCDSS is a national network of provincial and territorial surveillance systems coordinated by the Public Health Agency of Canada. The CCDSS identifies chronic disease cases using disease-specific case definitions based on International Classification of Diseases (ICD) 9th and 10th revisions, Canadian version, diagnostic codes from physician billing claims and hospital discharge abstracts. The CCDSS captures diabetes cases using either one hospitalization with a discharge ICD-10-CA code of E10 to E14, or 2 physician claims with an ICD-9 code of 250 within a 2-year period (3,4). This case definition has shown a high sensitivity and specificity (3,5). Although ICD-10 codes from hospital discharge abstracts provide more specific coding for type 1 (E10) and type 2 (E11) diabetes, the majority of diabetes cases in the CCDSS are identified by the physician claims data and, therefore, are based on ICD-9 codes. A major limitation of the CCDSS case definition is that physician claims capture only the first 3 digits of the ICD-9 codes, and thus the CCDSS does not specify diabetes type.

Surveillance based on self-report also has been used to estimate the prevalence of diabetes in Canada, using survey data routinely collected by Statistics Canada's Canadian Community Health Survey (CCHS). The CCHS is a nationally representative population-based survey that contains questions about a host of chronic conditions—including diabetes—and a comprehensive array of demographic, socioeconomic, health and lifestyle variables. Since 2007, the CCHS has been conducted on an annual basis (6). Thus, the CCHS can be used to estimate and track the prevalence of diabetes in Canada among household populations aged 12 years and older, and to study associations with its risk factors; however, the CCHS does not include a direct question about diabetes type. Because the knowledge of diabetes type is important for public health surveillance, there is a need to evaluate effective ways to collect that information in national health surveys.

Ng, Dasgupta and Johnson (NDJ) (7) previously reported an algorithm to classify type of diabetes among respondents of the CCHS cycle 1.1 (2000/01) who reported that they had been diagnosed with diabetes by a healthcare professional (8). The NDJ algorithm was based on an earlier attempt (9), and classified the respondents' type of diabetes based on their answers to CCHS questions about diabetes during pregnancy, use of oral medications to control diabetes, use of insulin, timing of initiation of insulin treatment and age at diagnosis (Figure 1).

In this study, we used data from the recent CCHS survey supplement, Survey on Living with Chronic Diseases in Canada (SLCDC)—Diabetes component to evaluate the feasibility of self-reported diabetes type. The uniqueness of the SLCDC survey is that it contained all of the questions needed for the NDJ algorithm as well as a question asking the respondent about his/her diabetes type, allowing for a direct comparison. We compared self-reported type with the NDJ algorithm, which is based on self-reported, albeit objective, respondent characteristics and pattern of practice.

Methods

The SLCDC is a cross-sectional survey that collects information related to Canadians' experiences with chronic health conditions. Two cycles of the SLCDC have taken place (2009 and 2011), with 2 chronic diseases covered in each survey cycle. In 2011, the surveys focused on diabetes and respiratory conditions. Individuals aged 20 years and older who reported having diabetes diagnosed by a healthcare professional as part of the 2008 CCHS were eligible for participation in the SLCDC—Diabetes. The SLCDC coverage includes approximately 98% of the Canadian population residing in the provinces (10). People living in the territories, on Indian Reserves or on Crown land were not surveyed, nor were members of the Canadian Forces or people living in institutions (10). Women who reported having diabetes only during pregnancy (i.e. gestational diabetes) were excluded (10).

We applied the NDJ algorithm (Figure 1) to the SLCDC—Diabetes responses and compared the diabetes type classification results with the results from the self-reported diabetes type question ("What kind of diabetes do you have?") (11), based on the weighted sample to evaluate the concordance between the results of the NDJ algorithm and the self-reported data. Agreement between the 2 classification methods was quantified using kappa coefficients. Strength of agreement was assessed according to criteria from Landis and Koch (12): less than 0.40 represents poor to fair, 0.41 to 0.60 indicates moderate, and more than 0.60 represents substantial or near-perfect agreement. Sensitivity and specificity for classifying type 2 diabetes, as well as the positive predictive value (PPV) and negative predictive value (NPV), were calculated, with the NDJ algorithm-assigned diabetes type used as the reference standard. If we consider the classification of type 2 diabetes, for example, PPV is the proportion of respondents who had type 2 diabetes based on the NDJ algorithm among those who self-reported having type 2 diabetes, and NPV is the proportion of respondents who had type 1 diabetes based on the NDJ algorithm, among those who self-reported having type 1 diabetes. All comparisons were stratified by age group (i.e. 20 to 29, 30 to 64, $\geq$65 years). Calculations were performed using SAS software version 9.1 (SAS Institute, Inc., Cary, NC). Bootstrap weights have been incorporated into the estimation to account for complex survey design.