Title

A mathematical model for determining age-specific diabetes incidence and prevalence using body mass index

Journal

Annals of Epidemiology, 23(5)

ISSN

1047-2797

Authors

Appuhamy, JAD Ranga Niroshan
Kebreab, E
France, J

Publication Date

2013-05-01

DOI

10.1016/j.annepidem.2013.03.011

Peer reviewed

Annals of Epidemiology 23 (2013) 248–254

Contents lists available at SciVerse ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

A mathematical model for determining age-specific diabetes incidence and prevalence using body mass index

J.A.D. Ranga Niroshan Appuhamy PhD [a,b,*], E. Kebreab PhD [b], J. France DSC [a]

[a] Centre for Nutrition Modelling, University of Guelph, Canada
[b] Department of Animal Science, University of California, Davis

ARTICLE INFO

Article history:
Received 3 October 2012
Accepted 19 March 2013
Available online 19 April 2013

Keywords:
Mathematical model
Diabetes prevalence
Diabetes incidence
Body mass index

ABSTRACT

Purpose: Few models have been developed specifically for the epidemiology of diabetes. Diabetes incidence is critical in predicting diabetes prevalence. However, reliable estimates of disease incidence rates are difficult to obtain. The aim of this study was to propose a mathematical framework for predicting diabetes prevalence using incidence rates estimated within the model using body mass index (BMI) data.
Methods: A generic mechanistic model was proposed considering birth, death, migration, aging, and diabetes incidence dynamics. Diabetes incidence rates were determined within the model using their relationships with BMI represented by the Hill equation. The Hill equation parameters were estimated by fitting the model to National Health and Nutrition Examination Survey (NHANES) 1999–2010 data and used to predict diabetes prevalence pertaining to each NHANES survey year. The prevalences were also predicted using diabetes incidence rates calculated from the NHANES data themselves. The model was used to estimate death rate parameters and to quantify sensitivities of prevalence to each population dynamic.
Results: The model using incidence rate estimates from the Hill equations successfully predicted diabetes prevalence of younger, middle-aged, and older adults (prediction error, 20.0%, 9.64%, and 7.58% respectively). Diabetes prevalence was positively associated with diabetes incidence in every age group, but the associations among younger adults were stronger. In contrast, diabetes prevalence was more sensitive to death rates in older adults than younger adults. Both diabetes incidence and prevalence were strongly sensitive to BMI at younger ages, but sensitivity gradually declined as age progressed. Younger and middle aged adults diagnosed with diabetes had at least a two-fold greater risk of death than their nondiabetic counterparts. Nondiabetic older adults were found to be under slightly higher death risk (0.079) than those diagnosed with diabetes (0.073).
Conclusions: The proposed model predicts diagnosed diabetes incidence and prevalence reasonably well using the link between BMI and diabetes development risk. Ethnic group and gender-specific model parameter estimates could further improve predictions. Model prediction accuracy and applicability need to be comprehensively evaluated with independent data sets.

Published by Elsevier Inc.

Introduction

Diabetes prevalence is rising dramatically worldwide and is expected to rise from 366 million in 2011 to 552 million by 2030 [1]. More than 10% of world health care expenditure and about 14% of U.S. healthcare costs are attributable to diabetes [2]. Quantifying diabetes prevalence is important to allow rational planning of prevention programs and allocating resources for people affected by diabetes [2]. Mathematical models can be used effectively to estimate disease prevalence and help understand factors affecting disease development risk. The majority of diabetes-related mathematical models explain clinical aspects of glucose–insulin dynamics, whereas few models have been specific to the epidemiology of diabetes [3]. Diabetes prevalence varies significantly with age implying the mechanisms underlying risk of developing diabetes could be age specific. Boutayeb and Derouich [4] proposed a mathematical model for predicting the age-specific prevalence of diabetes and its complications. Accurate prevalence predictions from such a model require reliable estimates of diabetes population dynamics, such as incidence rates and death rates.

Prevalence is the proportion of a population affected by a disease at a particular time point, whereas the incidence rate is the rate of occurrence of new cases of the disease. Incidence rates, indicative of risk of contracting or developing the disease, can be also used to measure the efficacy of disease prevention strategies. Nonetheless, obtaining reliable diabetes incidence rate estimates is often

* Corresponding author. Department of Animal Science, One Shield Avenue, University of California, Davis, CA 95616. Tel.: 530-752-2401; fax: 540-752-0175.
E-mail address: jappuhamy@ucdavis.edu (J.A.D.R.N. Appuhamy).

1047-2797/$ – see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.annepidem.2013.03.011

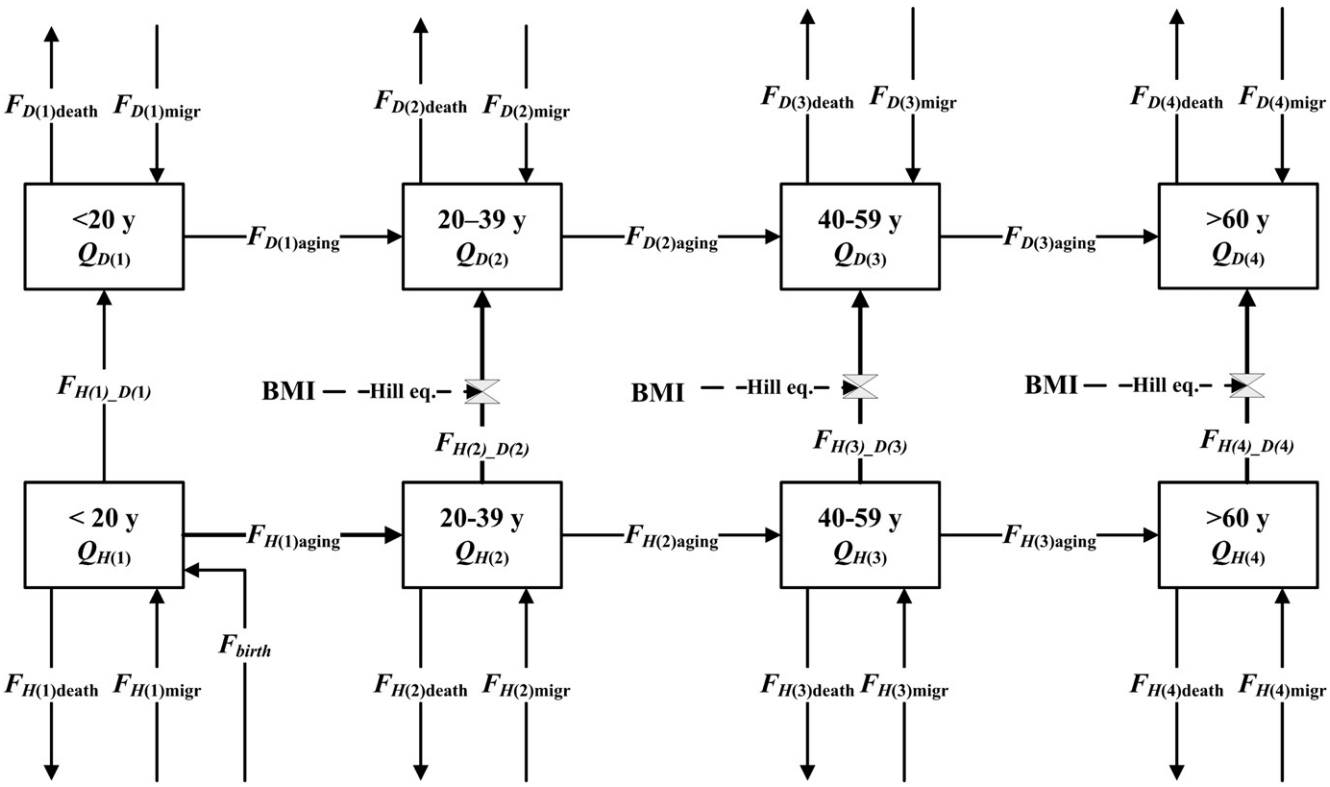

Fig. 1. Schematic representation of the model. Boxes, solid arrows and dashed arrows represent pools (Q), flows (F), and effects of body mass index (BMI) on diabetes incidence, respectively. Letter '*D*' and '*H*' denote diagnosed diabetic and nondiabetic individuals respectively. Time unit for the model is a year (y).

challenging and requires larger survey samples than those required for prevalence estimates. Therefore, a mathematical representation calculating diabetes incidence rates within the model itself can provide a better option for an efficient and more accurate prediction of diabetes prevalence. Mathematical models also allow for estimating parameters and determining sensitivities [3]. A model representing all major population mechanisms such as births, deaths, aging, migration, and diabetes incidence will help to assess relative sensitivities or strength of associations of each of these mechanisms to diabetes prevalence. Moreover, such a mathematical model also allows for estimating parameters of some critical mechanisms, for example, death rates [5]. The death rates associated with diseases are often estimated based on the information reported on death certificates. However, the reliability of death certificate-oriented death rate estimates appears to be doubtful [6].

Obesity has been a major factor in the recent increase in diagnosed diabetes incidence in the United States [7]. Therefore, body mass index (BMI) can potentially be a leading diabetes risk predictor. Huang et al. [8] constructed a comprehensive Markov chain model for predicting diabetes incidence and prevalence across different BMI categories in the total U.S. population. However, Narayan et al. [9] demonstrated that the link between BMI and diabetes development risk can vary significantly with age, suggesting a need for separate mathematical representations of age-specific associations between BMI and diabetes incidence rates. An appropriate mathematical representation quantifying the age-specific associations between BMI and diabetes incidence can be postulated to predict diabetes prevalence accurately. The main objective of this study was to propose a mathematical model to predict diabetes prevalence in different adult age groups. The specific objectives were to (1) develop a mathematical representation for quantifying the effect of BMI on diabetes development risk in adult age groups commonly defined in epidemiology, (2) assess sensitivities of diabetes prevalence to incidence, death and migration rates, and (3) estimate death rate constants and other parameters for diabetic and nondiabetic adults by fitting the model to National Health and Nutrition Examination Survey (NHANES) 1999–2010 data.

Materials and methods

Model development

The time unit for the model is a year (y). Total population size was arbitrarily set at 10,000, held constant, and divided into four age groups (x): (1) younger than 20, (2) 20 to 39, (3) 40 to 59, and (4) 60 years or older (Fig. 1). Individuals in each age group were allocated to two pools: Diabetic ($Q_{D(x)}$) and nondiabetic ($Q_{H(x)}$), which also includes undiagnosed cases. The diabetes incidence rate of each age group ($F_{H(x)\_D(x)}$) was taken to be a linear function of $Q_{H(x)}$ with corresponding diabetes fractional incidence rate $k_{H(x)\_D(x)}$:

$$F_{H(x)\_D(x)} = k_{H(x)\_D(x)} Q_{H(x)}.$$

Death rates of nondiabetic ($F_{H(x)\text{death}}$) and the diabetic ($F_{D(x)\text{death}}$) individuals in each age group were also taken as linear functions of $Q_{H(x)}$ and $Q_{D(x)}$, respectively, with corresponding fractional death rates $k_{H(x)\text{death}}$ and $k_{D(x)\text{death}}$:

$$F_{H(x)\text{death}} = k_{H(x)\text{death}} Q_{H(x)};$$

$$F_{D(x)\text{death}} = k_{D(x)\text{death}} Q_{D(x)}.$$

A fractional rate expresses an absolute rate or a flux (i.e., $F_{H(x)death}$) as a proportion of the pool of interest (i.e., $Q_{H(x)}$). Because units of the rates and pools are individuals per year and individuals respectively, the unit of the fractional rates is y^{-1}. For example, a fractional death rate of 0.01 y^{-1} means that 1% of the population dies annually.

Table 1
Mean values and standard deviations (SD, n = 6) for diagnosed diabetes prevalence, number of nondiabetics ($Q_{H(x)}$), number of diabetics ($Q_{D(x)}$), fractional diabetes incidence rate ($k_{H(x)\_D(x)}$), net migration rate of individuals diagnosed with diabetes ($F_{D(x)\text{migr}}$), and net migration rate of nondiabetic individuals ($F_{H(x)\text{migr}}$) in each age group across the period 1999–2010

Age group (yrs)	BMI		Prevalence		$Q_{H(x)}$		$Q_{D(x)}$		$k_{H(x)\_D(x)}$		$F_{D(x)\text{migr}}$		$F_{H(x)\text{migr}}$	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<20	ND	ND	0.003	0.001	2681	48	9	3	0.0005	0.0004	0.04	0.003	8.77	0.70
20–39	27.9	0.32	0.019	0.005	2668	79	52	13	0.0027	0.0009	0.25	0.020	16.4	1.32
40–59	28.9	0.25	0.102	0.014	2492	33	284	43	0.0086	0.0022	0.34	0.022	8.09	0.54
≥60	27.9	0.31	0.208	0.021	1435	22	378	53	0.0108	0.0017	0.48	0.037	2.52	0.2

ND = not determined.

Aging rates (F_{aging}; Fig. 1) by which nondiabetic and diagnosed diabetic individuals move to the next age group (from x to $x + 1$) were also represented as linear functions of preceding age group pool size and relevant fractional rates, set at 0.05 y^{-1} (because the age groups were 20 years). Related to the U.S. population, net migration flows were assumed to be positive and represented as immigration flows ($F_{D(x)\text{migr}}$ and $F_{H(x)\text{migr}}$ in Fig. 1). All births (F_{birth} in Fig. 1) were assumed to be nondiabetic. Rates of change in $Q_{H(x)}$ and $Q_{D(x)}$ were represented by ordinary dynamic differential equations. For example, rate of change of the nondiabetic younger than 20 pool was represented by the following equation which describes the balance between birth and migration inflows, and diabetes incidence, death, and aging outflows (Fig. 1):

$$\frac{dQ_{H(1)}}{dt} = \left(F_{\text{birth}} + F_{H(1)\text{migr}}\right) - \left(F_{H(1)\_D(1)} + F_{H(1)\text{death}} + F_{H(1)\text{aging}}\right).$$

Similarly, rate of change in the diabetic pool was calculated as a balance between incidence and migration inflows, and aging and death outflows (Fig. 1):

$$\frac{dQ_{D(1)}}{dt} = \left(F_{H(1)\_D(1)} + F_{D(1)\text{migr}}\right) - \left(F_{D(1)\text{death}} + F_{D(1)\text{aging}}\right).$$

$Q_{H(x)}$ and $Q_{D(x)}$ were then determined by numerical integration of the corresponding differential equations. Diabetes prevalence in each age group ($PRV_{(x)}$) was finally calculated using the corresponding $Q_{H(x)}$ and $Q_{D(x)}$:

$$PRV_{(x)} = \frac{Q_{D(x)}}{\left(Q_{D(x)} + Q_{H(x)}\right)}.$$

The $k_{H(x)\_D(x)}$ of adult age groups (≥20 years) were estimated using average BMI of the nondiabetic pools and the Hill equation from allosteric enzyme kinetics [10]. The Hill equation has three parameters ($Y_{\max}$, K, and n) and gives a nonlinear (sigmoidal) relationship between two variables X (substrate concentration) and Y (reaction rate):

$$Y = \frac{Y_{\max}}{1 + \left(\frac{K}{X}\right)^{n}}.$$

The fractional diabetes incidence rate in each healthy adult group ($Y = k_{H(x)\_D(x)}$) as a function of BMI ($X = BMI_{(x)}$) was therefore represented by:

$$k_{H(x)\_D(x)} = \frac{k_{H(x)\_D(x)\max}}{1 + \left(\frac{K_{(x)}}{BMI_{(x)}}\right)^{n_{(x)}}},$$

where $k_{H(x)\_D(x)\max}$ is the maximum fractional diabetic incidence rate, $K_{(x)}$ is the affinity constant for BMI, $BMI_{(x)}$ is the average BMI of nondiabetic individuals, and $n_{(x)}$ is the sigmoidicity parameter or Hill coefficient.

Data and calculations

Diagnosed diabetes and BMI data were obtained from six separate NHANES surveys conducted during 1999 through 2010 [11]. The diabetes prevalence (observed prevalence) and incidence rates (observed $k_{H(x)\_D(x)}$) in each age group were calculated for each survey year (Table 1). Incidence rates were calculated using the 'current age' and 'age when diabetes was first diagnosed' data. The number of years each person had been diagnosed with diabetes was calculated by subtracting the latter from the former. The participants who had a value of zero were identified as having been newly diagnosed within the survey year. The observed $Q_{H(x)}$ and $Q_{D(x)}$ in each age group were then calculated pertaining to a 10,000 population and adjusted for U.S. population age structures in respective survey years [12]. Moreover, three BMI categories: (i) below 25, (ii) 25 to 29.9, and (iii) at least 30 kg/m^2, were formed within each adult age group. Such group formation was required as the model predicts diabetes prevalence of a group or a population using an average BMI. Diabetes prevalence and incidence rates, average BMI of nondiabetic individuals, and $Q_{H(x)}$ and $Q_{D(x)}$ were again calculated for each BMI category in each age group. Annual birth rates ($F_{\text{birth}} = 141 \pm 3.0$ per10,000, during 2000–2010), and the immigration rates of each age group (averages given in Table 1) were calculated from the U.S. Census Bureau statistics [13,14] considering the age structures of immigrants [15] and diabetes prevalence estimates of U.S. immigrants reported in Oza-Fran et al. [16].

Model simulation

The model was separately run to predict age-specific diabetes prevalence in each NHANES survey year using each year's birth (F_{birth}) and migration rate ($F_{H(x)\text{migr}}$ or $F_{D(x)\text{migr}}$) estimates. For these preliminary simulations, the Hill equation parameter values for each age group were assigned in a trial-and-error manner by plotting the NHANES incidence rate estimates of the BMI categories against the average BMI of the nondiabetic individuals. Fractional death rates were taken from Gu et al. [17]. In every simulation, total population size was assumed to be 10,000. All 10,000 individuals were assigned to $Q_{H(1)}$ initially (time = 0 years) and the rest of the pools set to ~0 (Fig. 2). The model was run to achieve steady state, and $Q_{H(x)}$ and $Q_{D(x)}$ were being calculated iteratively using the Runge-Kutta fourth-order method. The $Q_{H(x)}$ and $Q_{D(x)}$ at the steady state correspond respectively with nondiabetic and diabetic population sizes in each age group in the survey year of interest (1999–2000, 2001–2002, etc.). The steady-state $Q_{H(x)}$ and $Q_{D(x)}$ were then used to calculate diabetes prevalence (predicted prevalence) in each age group in each survey year. The $Q_{H(x)}$ and $Q_{D(x)}$

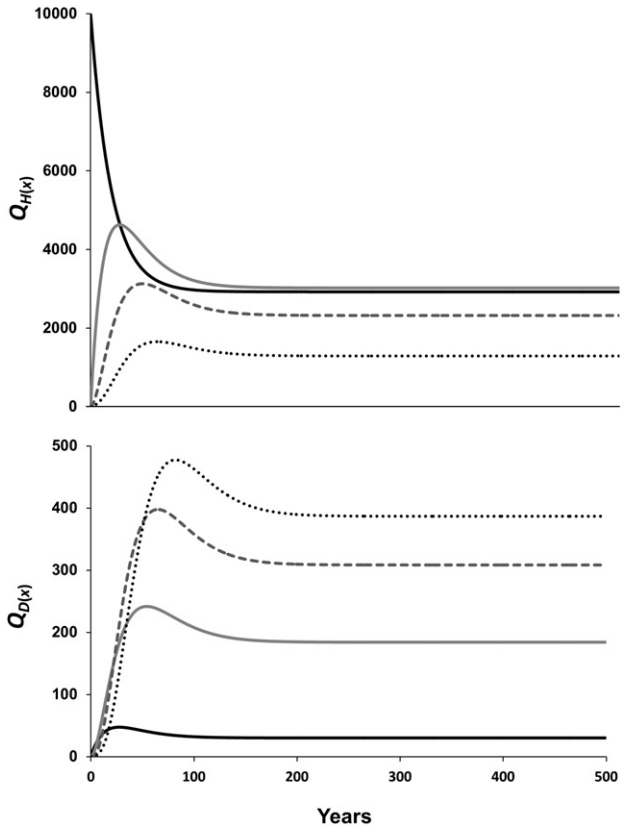

Fig. 2. Simulated time course to steady state of nondiabetic ($Q_{H(x)}$) and diabetic ($Q_{D(x)}$) individuals in younger than 20 (solid black), 20 to 39 (solid gray), 40 to 59 (dashed gray), and 60 years or older (dotted black) age groups for a population of 10,000.

(Fig. 2) and thereby the predicted prevalence achieved a steady state by 300 years (time = 300 years) in all cases.

Model fitting and parameter estimation

Hill equation parameters of adult age groups and fractional death rates of all the age groups were estimated by fitting the model simultaneously to observed $Q_{H(x)}$ and $Q_{D(x)}$ of each BMI category in each age group of each NHANES year. Time for these observed $Q_{H(x)}$ and $Q_{D(x)}$ was set to 300 years. The initial parameter values were calculated from data in Gu et al. [17] and using the plots of NHANES incidence rates versus average BMI. Model fitting was carried out using the Nelder-Mead search algorithm [18] while maximizing the log-likelihood function. The diabetes prevalence in each age group pertaining to each survey year was predicted using final parameter estimates of fractional death rates and $k_{H(x)\_D(x)}$ from the Hill equations. An additional round of prevalence predictions was carried out using the NHANES $k_{H(x)\_D(x)}$ estimates (averages in Table 1). These predicted values were used in internal model evaluations as described below.

Sensitivity analysis

A sensitivity analyses was conducted to quantify changes in diabetes prevalence (response variable) for unit increases in diabetes incidence; birth, death, and migration rate parameters; and BMI. Responsiveness of diabetes incidence (predicted by the Hill equations) to BMI was additionally assessed. The parameter values and BMI were set to age group averages and sensitivities were analyzed simultaneously. The changes in the response variables were calculated as partial derivatives and referred to as sensitivity coefficients. Sensitivity coefficients were normalized for both independent and response variables allowing comparison of sensitivities across age groups and across parameters.

Internal model evaluation

Prediction errors associated with diabetes prevalence in each age group were calculated as:

Prediction error = observed prevalence − predicted prevalence.

As mentioned, the observed prevalences were from the six NHANES surveys (n = 6) conducted from 1999 through 2010. Mean square prediction error statistics were calculated to determine model accuracy [19]. Square root of mean square prediction error (RMSPE) is directly comparable with the observed variable of interest (diabetes prevalence herein) so that RMSPE was calculated and expressed as a percentage of average observed diabetes prevalence to indicate uncertainty of prediction. Model development, simulations, optimizations, sensitivity analysis, and evaluation were carried out using acslXtreme software (AEgis Technologies, Huntsville, AL).

Comparison of model predictions with literature data

The fractional diabetes incidence rates ($k_{H(x)\_D(x)}$, y^{-1}) predicted by the Hill equation are equivalent to the average annual probability that an individual in a particular adult age group develops diabetes. As per NHANES 1999–2010 data, average BMI of middle-aged adults increased by 4% compared with the average BMI of younger adults (27.9 vs. 28.9). When the adults became older (≥60 years) average BMI decreased and was similar to the value they had when young (Table 1). For an adult presently 20 years old, having a BM I of 28.0 kg/m^2 and assumed to live for 80 years, our model predicts an average annual diabetes development probability of 0.0008 (0.08%) and consequently a total probability of 0.016 (0.0008 × 20) in the 20- to 40-year age group. Similarly, the model predicts total probabilities of 0.106 (0.0053 × 20, considering a 4% BMI increase) and 0.206 (0.0103 × 20) in the 40- to 60-year and 60- to 80-year age groups, respectively. Thus, the projected total lifetime diabetes development risk of an individual is 0.328 (32.8%).

The predictions from the model were compared with those reported in Narayan et al. [9], who used a comprehensive Markov scheme to predict lifetime diagnosed diabetes risk in different BMI categories among U.S. adults. Similarly, total diabetes incidence pertaining to middle-aged (average age, 45 years) Caucasian, Asian, Hispanic, and Black women (baseline average BMI of 24.3, 22.7, 24.3, and 26.0, respectively) during the next 20 years were predicted using the model, and compared with observed incidence values from a long-term (1980–2000) female cohort in the United States [20]. Further, the incidence rates predicted for normal weight, overweight, and obese middle-aged men (average age, 50 years) were compared with the observed rates from a long-term (20 years) Swedish cohort study reported by Arnlov et al. [21].

Results

The final estimates of the fractional death rates are given in Table 2. The small standard deviations of the estimates indicate that the observed $Q_{H(x)}$ and $Q_{D(x)}$ data from NHANES 1999–2010 were adequate to determine the parameters. The death rate estimates for younger (20- to 40-year-old) and middle-aged (40- to 60-year-old) U.S. adults, diagnosed with diabetes were 0.0097 (9.7 per 1000) and 0.0156 (15.6 per 1000), respectively. These fractional rate estimates represent average death risk of corresponding groups during the last decade. The death risk in diabetic adults in younger and

Table 2
Death rate estimates (± standard deviation), and normalized sensitivity coefficients of diabetes prevalence and incidence rate with respect to model parameters and body mass index (BMI) for different age groups

	Age group (yrs)			
	<20	20–39	40–59	≥60
Fractional death rate parameter estimates (y^{-1})				
Nondiabetic	0.0045 ± 0.00003	0.0048 ± 0.00002	0.0056 ± 0.00005	0.0791 ± 0.00012
Diabetic	0.0145 ± 0.00035	0.0097 ± 0.00011	0.0156 ± 0.00009	0.0732 ± 0.00068
Sensitivity coefficients of diagnosed diabetes prevalence for model parameters				
Fractional incidence rate	0.057	0.044	0.052	0.034
Birth rate	−0.033	−0.031	−0.006	0.008
Fractional death rates				
Diabetic	−0.014	−0.009	−0.008	−0.067
Nondiabetic	0.003	0.009	0.014	0.096
Net migration rate				
Diabetic	0.035	0.021	0.004	0.000
Nondiabetic	−0.002	−0.012	−0.009	0.000
Sensitivity coefficients of diagnosed diabetes prevalence and incidence rate for BMI				
Diabetes prevalence	NA	3.384	2.568	1.869
Diabetes incidence rate	NA	3.992	3.012	2.232

Note: Nondiabetic individuals also included undiagnosed cases. The sensitivity coefficients were normalized with respect to both independent (rate parameters and body mass index) and response variables (diabetes prevalence and incidence).

middle-aged groups were two- to three-fold greater than the risk among their nondiabetic counterparts. However, the nondiabetic older adults were associated with a slightly greater death risk than those diagnosed with diabetes. Average annual death rate of nondiabetic older adults was 79 per 1000, whereas the estimate for those with diabetes was 73 per 1000.

Sensitivity coefficients given in Table 2 show that diabetes prevalence positively responded to diabetes incidence in every age group. However, it was more sensitive to diabetes incidence in younger and middle-aged than older adults (0.044 and 0.052 vs. 0.034). As expected, diabetes prevalence was negatively associated with death rates of diabetic individuals, whereas increasing death rates among nondiabetics were associated with increasing diabetes prevalence in every adult age group. The sensitivities to death rate became more pronounced as age progressed. For example, compared with younger and middle-aged adults, sensitivity of prevalence to diabetic death rate was about eight-fold greater (−0.009 and −0.008 vs. −0.067) among older adults. The sensitivities of diabetes prevalence to migration rate were negligible in older adults indicating a potential to simplify the model by removing the migration effects from older adult groups. Effects of birth and migration rates were notably stronger in younger than older adults. Diabetes prevalence and incidence were more strongly associated with BMI in younger adults than middle-aged and older adults. The sensitivities to BMI gradually declined as age progressed.

Results from model prediction accuracy analyses are presented in Table 3. When NHANES fractional incidence rate estimates were used, the model notably overpredicted diabetes prevalence in younger and middle-aged adults (RMSPE = 173% and 54.2% of the average observed value; Table 3). Fractional incidence rates were then estimated with the Hill equation (parameter estimates ± standard deviations are given in Fig. 3) using BMI data. Consistent with the overpredicted prevalence, the NHANES incidence rate estimates were considerably larger than the estimates obtained from the Hill equation in middle-aged and older groups (Table 3). This confirms the greater sensitivity of diabetes prevalence to incidence rates in these groups. Fractional incidence rate estimates from the Hill equation significantly improved the accuracy of diabetes prevalence in both younger and middle-aged adults as the RMSPE declined substantially (from 173% to 20.0% and from 54.2% to 9.64%, respectively). The model reasonably predicted diabetes prevalence in older adults (RMSPE, 18.5%), even when using the diabetes incidence rate estimates from NHANES. However, the estimates from the Hill equation further reduced prediction errors (RMSPE, 7.58%) of this group as well.

The model predicted lifetime diagnosed diabetes risk of adults fairly close to the estimates from the Markov scheme used by Narayan et al. [9]. The model predicted 14.8%, 29.5%, and 55.5% lifetime risks for 18-year-old, normal weight, overweight, and obese individuals, respectively. The corresponding projections by Narayan et al. [9] were 14.2%, 32.6%, and 64.1% respectively. Our model projections for the 45- and 65-year-old individuals in respective BMI categories were 13.0, 26.0%, and 50.7%, and 7.8%, 16.5%, and 27.8%, respectively. The corresponding projections from Narayan et al. [9] were 12.5%, 28.3%, and 54.8%, and 6.5%, 16.3%, and 31.9%, respectively. The model overpredicted the 20-year diabetes risk of Caucasian women (86 vs. 58 per 1000) and underpredicted the 20-year diabetes risk of Asian (74 vs. 84 per 1000), Hispanic (90 vs. 100 per 1000), and Black (109 vs. 127 per 1000) middle-aged

Table 3
Fractional diabetes incidence rates predicted by the Hill equation, and accuracy of diabetes prevalence predicted using NHANES incidence rate estimates versus prevalence predicted using the rate estimates obtained from the Hill equation

Age group (yrs)	Fractional diabetes incidence rate		Diabetes prevalence				
	From NHANES	From Hill equations	Mean	With NHNES incidence rates		With the Hill equations	
			Observed	Mean predicted	RMSPE %	Mean predicted	RMSPE %
20–39	0.0027	0.0008	0.019	0.049	173	0.019	20.0
40–59	0.0086	0.0052	0.102	0.156	54.2	0.099	9.64
≥60	0.0108	0.0103	0.208	0.240	18.5	0.202	7.58

RMSPE % = square root of mean square prediction error estimate, expressed as a percentage of the average observed value; the errors were calculated pertaining to incidence and prevalence from six separate National Health and Nutrition Examination Survey (NHANES; 1999–2010) datasets.

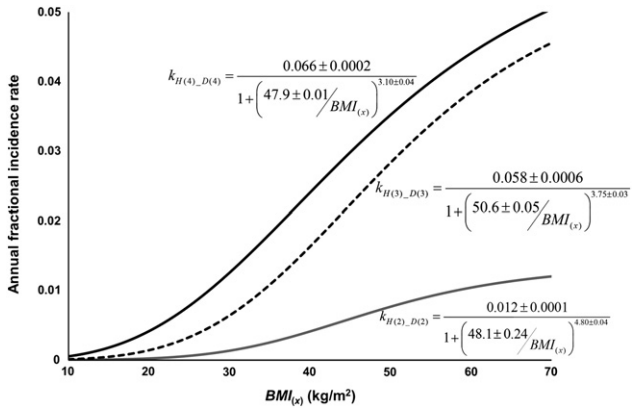

Fig. 3. Hill equation parameter estimates and relationship of predicted fractional diabetes incidence rate ($k_{H(x)\_D(x)}$) to average body mass index ($BMI_{(x)}$) for 20 to 39 (gray solid, $x = 2$), 40 to 59 (black dashed, $x = 3$), and 60 years or older (black solid, $x = 4$) adult age groups. The parameter estimates (± standard deviation) of each Hill equation were obtained by fitting the model to diagnosed diabetes prevalence from six separate datasets of NHANES 1999–2010.

women, compared with the observed rates in Shai et al. [20]. The model predicted the 20-year diabetes risk of middle-aged, overweight (144 vs. 140 per 1000) and obese (270 vs. 306 per 1000) middle-aged Swedish men with reasonable accuracy [21].

Conclusions

As expected, diabetes incidence rates were strongly associated with diabetes prevalence in every age group. However, death rate even dominated incidence rate in determining prevalence in older adults. This is consistent with previous observations by Monesi et al. [5]. Diabetes has been found to be the seventh leading cause of death among U.S. adults [22]. Diabetes can cause death indirectly by leading to strokes, heart attacks, and kidney failure, and as such it may not be recorded as a principal cause on death certificates. For this reason, the number of deaths attributable to diabetes could, potentially, be underestimated [6]. The annual death rates of younger and middle-aged diabetic adults considerably decreased (by 27% and 60%, respectively) from estimates from 1971 through 1993 [17]. Nonetheless, diabetic death rates in these age groups still remained at least two-fold greater than those of the nondiabetic adults. Interestingly, older diabetic adults were found to be under a slightly lesser (by 8%) death risk than older nondiabetic adults. Gulliford and Charlton [23] found similar death rate comparisons between elderly diabetic individuals and the general population in the United Kingdom. These authors implied that this could be a result of improved standards of chronic illness care leading to increased uptake of medical interventions controlling mortality risk factors. Gu et al. [17], on the other hand, suggested a greater prevalence of more life-threatening diseases than diabetes among older adults.

Because obesity is strongly associated with diabetes development risk in the United States [24,25], we chose BMI of the nondiabetic individuals for estimating diabetes development risk in terms of fractional diabetes incidence rates. Colditz et al. [26] and Ford et al. [27] clearly showed a curvilinear relationship between BMI and risk of developing diabetes mellitus among U.S. adults. Moreover, Chiu et al. [28] demonstrated a sigmoidal relationship between diabetes development risk and BMI in Canadian adults. We therefore chose the Hill equation to represent the relationship between BMI and diabetes incidence rate in our model. As expected, diabetes incidence rates estimated with the Hill equation led to more accurate prevalence predictions. This further strengthens the idea that BMI is promising predictor of diabetes risk [29]. Furthermore, this model can be extended to include other diabetes risk factors, such as fasting blood glucose and blood pressure. Our results showed that sensitivity of diabetes risk to BMI in younger adults was nearly twice that in older adults. These results agree with previous observations that weight gain in early adulthood is related to a higher risk of developing diabetes than weight gain in older age groups [30]. Furthermore, this model could provide a framework to develop a generic model for predicting risk and prevalence of the other noncommunicable diseases, such as cardiovascular diseases and cancer.

The absence of representation of ethnic and gender effects could be a significant limitation of the model [17,28]. Presently, it predicts average diabetes incidence and diabetes prevalence across ethnic groups and genders, so that the predictions are positioned, for instance, between the values of the Caucasian population in the United States having less prevalence (overpredictions), and the Hispanic and Black populations having greater prevalence (underpredictions). The model seemed to significantly overpredict (by 48%) diabetes risk in Caucasian adults compared with the extents of underprediction in Hispanic (by 10%) and Black (by 14%) adults. This is potentially an artifact of oversampled minority groups in NHANES data (i.e., 40% White, 32% Hispanic, and 23% Non-Hispanic Black), for which the model parameter were estimated. Hence, ethnic group-specific model parameter estimates can be expected to improve corresponding model predictions. Diabetes incidence estimates calculated from NHANES data seemed to be doubtful specifically in younger and middle-aged adults. This is not surprising; the NHANES surveys were not specifically designed to estimate incidence rates. Surveys targeting reliable diabetes incidence rates need to increase sample size for younger populations compared with older populations.

NHANES data do not distinguish between types 1 and 2 diabetes, so predicted prevalence gave general diabetes prevalence in U.S. adults. Nonetheless, about 95% of diabetes cases in adults are type 2. Therefore, the predictions can reasonably be ascribed to type 2. As mentioned, the nondiabetic pools ($Q_{H(x)}$) in this model also include undiagnosed cases, so the predictions are underestimates of absolute diabetes prevalence. Given the significantly declining undiagnosed diabetes prevalence in the United States [31,32], effects of this discrepancy could be minimal in the future. Additionally, the Hill equation in this model describes a likelihood of overweight or obese adults having their diabetes diagnosed rather than the pathophysiologic link between BMI and diabetes development risk. Projected diabetes risk changes in response to changes in BMI given by the model need to be interpreted carefully. Nonetheless, Gregg et al. [32] and Wee et al. [33] demonstrated the ratio of undiagnosed diabetes prevalence to diagnosed diabetes prevalence remains unchanged across different BMI categories. Therefore, the model provides opportunities to assess the pathophysiological link between BMI and diabetes development risk at least in relative terms.

We have proposed a generic mathematical framework for predicting age-specific diabetes development risk and prevalence across ethnic groups and genders. The Hill equation appropriately represents the link between BMI and diabetes risk and we have parameterized three Hill equations to project diabetes development risk in younger, middle-aged, and older adult groups. This generic age-specific layout should predict lifetime diabetes risk successfully. The sensitivity estimates obtained from this type of mechanistic model would help select critical parameters in an epidemiology model. This model also allows critical parameters (i.e., death rates) to be estimated by fitting the model to observed prevalence, which can be obtained more conveniently. Firm

conclusions regarding the prediction accuracy and model applicability should be made after a comprehensive model evaluation.

Acknowledgments

The authors thank the Advanced Foods and Materials Network of Canada, the Canada Research Chairs Program and the Sesnon Endowed Chairs Program (UC Davis) for funding support.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.annepidem.2013.03.011.

References

[1] One adult in ten will have diabetes by 2030, International Diabetes Federation, www.idf.org/media-events/press-releases/2011/diabetes-atlas-5th-edition; 2011 [accessed 16.01.2013].
[2] Zhang P, Zhang X, Brown J, Vistisen D, Sicree R, Shaw J, Nichols G. Global healthcare expenditure on diabetes for 2010 and 2030. Diabetes Res Clin Pract 2010;87:293–301.
[3] Boutayeb A, Chetouani A. A critical review of mathematical models and data used in diabetology. Biomed Eng Online 2006;5:43.
[4] Boutayeb A, Derouich M. Age structured models for diabetes in East Morocco. Mathematics and Computers in Simulation 2002;58:215–29.
[5] Monesi L, Baviera M, Marzona I, Avanzini F, Monesi G, Nobili A, et al. Prevalence, incidence and mortality of diagnosed diabetes: evidence from an Italian population-based study. Diabet Med 2012;29:385–92.
[6] Panzram G. Mortality and survival in type 2 (non-insulin-dependent) diabetes mellitus. Diabetologia 1987;30:123–31.
[7] Geiss LS, Pan L, Cadwell B, Gregg EW, Benjamin SM, Engelgau MM. Changes in incidence of diabetes in US adults, 1997-2003. Am J Prev Med 2006;30:371–7.
[8] Huang ES, Basu A, O'Grady M, Capretta JC. Projecting the future diabetes population size and related costs for the US. Diabetes Care 2009;32:2225–9.
[9] Narayan KM, Boyle JP, Thompson TJ, Gregg EW, Williamson DF. Effect of BMI on lifetime risk for diabetes in the US. Diabetes Care 2007;30:1562–6.
[10] Hill AV. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J Physiol 1910;40:4–6.
[11] National Center for Health Statistics. National Health and Nutrition Examination Survey. http://www.cdc.gov/nchs/nhanes.htm; [accessed 29.08.2011].
[12] U.S. Census Bureau. Resident population by sex and age 1980 to 2010. http://www.census.gov/compendia/statab/; [accessed 12.04.2012].
[13] U.S. Census Bureau. Births, deaths, marriages, and divorces. http://www.census.gov/; [accessed 29.08.2011].
[14] Bhaskar R, Scopilliti M, Hollmann F, Amastrong D. Plans for producing estimates of net international migration for the 2010 demographic analysis estimates. http://www. census.gov/; [accessed 12.04.2012].
[15] Department of Homeland Security. Yearbook of immigration statistics 2001-2009. http://www.dhs.gov/yearbook-immigration-statistics; [accessed 12.04.2012].
[16] Oza-Frank R, Stephenson R, Narayan KM. Diabetes prevalence by length of residence among US immigrants. J Immigr Minor Health 2011;13:1–8.
[17] Gu K, Cowie CC, Harris MI. Mortality in adults with and without diabetes in a national cohort of the US Population, 1971-1993. Diabetes Care 1998;21:1138–45.
[18] Nelder JA, Mead R. A simplex method for function minimization. Computer Journal 1965;7:308–13.
[19] Tedeschi L. Assessment of the adequacy of mathematical models. Agric Syst 2006;89:225–47.
[20] Shai I, Jiang R, Manson JE, Stampfer MJ, Willett WC, Colditz GA, et al. Ethnicity, obesity, and risk of type 2 diabetes in women: a 20-year follow-up study. Diabetes Care 2006;29:1585–90.
[21] Arnlov J, Sundstrom J, Ingelsson E, Lind L. Impact of BMI and the metabolic syndrome on the risk of diabetes in middle-aged men. Diabetes Care 2011;34:61–5.
[22] Heron M, Hoyert DL, Murphy SL, Xu J, Kochanek KD, Tejada-Vera B. Deaths: final data for 2006. Natl Vital Stat Rep 2009;57:1–134.
[23] Gulliford MC, Charlton J. Is relative mortality of type 2 diabetes mellitus decreasing? Am J Epidemiol 2009;169:455–61.
[24] Kelly T, Yang W, Chen CS, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. Int J Obes (Lond) 2008;32:1431–7.
[25] Mokdad AH, Bowman BA, Ford ES, Vinicor F, Marks JS, Koplan JP. The continuing epidemics of obesity and diabetes in the United States. JAMA 2001;286:1195–200.
[26] Colditz GA, Willett WC, Rotnitzky A, Manson JE. Weight gain as a risk factor for clinical diabetes mellitus in women. Ann Intern Med 1995;122:481–6.
[27] Ford ES, Williamson DF, Liu S. Weight change and diabetes incidence: findings from a national cohort of US adults. Am J Epidemiol 1997;146:214–22.
[28] Chiu M, Austin PC, Manuel DG, Shah BR, Tu JV. Deriving ethnic-specific BMI cutoff points for assessing diabetes risk. Diabetes Care 2011;34:1741–8.
[29] Chan JM, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. Diabetes Care 1994;17:961–9.
[30] Schienkiewitz A, Schulze MB, Hoffmann K, Kroke A, Boeing H. Body mass index history and risk of type 2 diabetes: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Am J Clin Nutr 2006;84:427–33.
[31] Smith JP. Nature and causes of trends in male diabetes prevalence, undiagnosed diabetes, and the socioeconomic status health gradient. Proc Natl Acad Sci USA 2007;104:13225–31.
[32] Gregg EW, Cadwell BL, Cheng YJ, Cowie CC, Williams DE, Geiss L, et al. Trends in the prevalence and ratio of diagnosed to undiagnosed diabetes according to obesity levels in the US. Diabetes Care 2004;27:2806–12.
[33] Wee CC, Hamel MB, Huang A, Davis RB, Mittleman MA, McCarthy EP. Obesity and undiagnosed diabetes in the US. Diabetes Care 2008;31:1813–5.