

Value of Risk Stratification to Increase the Predictive Validity of HbA1c in Screening for Undiagnosed Diabetes in the US Population

Adit A. Ginde, MD, MPH^{1,2}, Enrico Cagliero, MD³, David M. Nathan, MD³, and Carlos A. Camargo, Jr., MD, DrPH⁴

¹Department of Emergency Medicine, University of Colorado Denver School of Medicine, Aurora, CO, USA; ²Department of Emergency Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA; ³Diabetes Center, Massachusetts General Hospital, Boston, MA, USA; ⁴Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, USA.

BACKGROUND: Opportunistic screening using hemoglobin A1c (HbA1c) may improve detection of undiagnosed diabetes but remains controversial.

OBJECTIVE: To evaluate the predictive validity of HbA1c as a screening test for undiagnosed diabetes in a risk-stratified sample of the US population.

DESIGN: Weighted cross-sectional analysis of diabetes risk factors, HbA1c, and fasting plasma glucose (FPG) in National Health and Nutrition Examination Survey (NHANES), 1999–2004.

SUBJECTS: Six thousand seven hundred and twenty-three NHANES participants from morning examination session, aged ≥ 18 years and without prior physician-diagnosed diabetes.

MEASUREMENTS: HbA1c and undiagnosed diabetes defined by FPG ≥ 7.0 mmol/l (126 mg/dl).

RESULTS: The estimated prevalence of undiagnosed diabetes in the US population was 2.8% (5.5 million people). HbA1c had strong correlation with undiagnosed diabetes, with an area under the receiver-operating characteristic curve of 0.93. Independent predictors of undiagnosed diabetes were older age, male sex, black race, hypertension, elevated waist circumference, elevated triglycerides, and low high-density lipoprotein cholesterol. We derived a risk score for undiagnosed diabetes and stratified participants into low (0.44% prevalence), moderate (4.1% prevalence), and high (11.1% prevalence) risk subgroups. In moderate and high risk groups, a threshold HbA1c value $\geq 6.1\%$ identified patients requiring confirmatory FPG; HbA1c $\leq 5.4\%$ identified patients for whom diabetes could be reliably excluded. Intermediate HbA1c (5.5–6.0%) may exclude diabetes in moderate, but not high risk groups).

CONCLUSIONS: Risk stratification improves the predictive validity of HbA1c in screening for undiagnosed diabetes in the US population. Although opportunistic screening with HbA1c would improve detection of undiagnosed diabetes, cost-effectiveness studies are

needed before implementation of specific screening strategies using HbA1c.

KEY WORDS: diabetes mellitus; glycohemoglobin; screening; validity; NHANES.

J Gen Intern Med 23(9):1346–53

DOI: 10.1007/s11606-008-0661-6

© Society of General Internal Medicine 2008

INTRODUCTION

Of the 19 million Americans with type 2 diabetes, 6 million remained undiagnosed.^{1,2} Early diagnosis, lifestyle modification, and tight glycemic control reduce complications; however, these cannot occur if diabetes is undiagnosed. Indeed, diabetes may be present for up to 7 years before diagnosis.³

Although improving diabetes detection remains prominent in the nation's public health agenda,⁴ there is insufficient evidence for or against routine diabetes screening.^{5,6} One reason may be the burden and inconvenience of fasting visits to meet the diagnostic criteria. Diabetes is usually established by fasting plasma glucose (FPG), values which require confirmation on a second visit.⁶

Opportunistic screening for high-risk individuals during unscheduled outpatient, urgent care, or hospital visits may improve rates of diagnosis. Hemoglobin A1c (HbA1c) offers the potential for a more practical alternative, as it is an established measure of long-term glycemia and correlates directly with subsequent development and progression of microvascular complications.^{7,8} In addition, HbA1c measurement does not require fasting and remains unaffected by transient hyperglycemia from acute stress or illness.⁹

For these reasons, investigators have examined HbA1c for diabetes screening in undifferentiated populations,^{10–21} but utility as a diagnostic or screening test remains controversial, primarily due to limited diagnostic accuracy.^{22–26} For instance, Rohlfing et al. found 63% sensitivity and 97% specificity for threshold HbA1c $\geq 6.2\%$, and 83% sensitivity and 84% specificity for HbA1c $\geq 5.7\%$, using the NHANES III dataset.¹¹ These studies do not, however, explore diabetes risk factor data to additionally stratify populations, and many were limited by prior lack of HbA1c reliability, a problem rectified by the National Glycohemoglobin Standardization Program.²⁷ In the present analysis, we evaluated the performance of HbA1c as a screening test for undiagnosed diabetes in a risk-stratified, nationally representative sample of the US population.

Received August 29, 2007

Revised January 29, 2008

Accepted April 25, 2008

Published online June 10, 2008

METHODS

Study Design and Participants

The National Health and Nutrition Examination Survey (NHANES) is a nationally representative probability sample of the noninstitutionalized US civilian population, released data in two-year cycles. We received a waiver from our institutional review board to analyze the three most recent cycles of released NHANES data (i.e., data collected between January 1999 and December 2004).

Details of survey methodology are described elsewhere.²⁸ Briefly, the sample is obtained by using a complex, stratified, multistage design, with oversampling of certain subgroups. NHANES uses a four-stage sampling strategy covering geographic primary sampling units (PSUs).

During 1999 to 2004, NHANES collected household interview data, including demographic characteristics, medical conditions, and physical activity, for 31,126 (87%) of 35,960 invited participants. The majority (29,402 [94%]) subsequently received physical and laboratory examination, including body measurements, plasma glucose, and HbA1c. NHANES randomly assigned these participants to morning, afternoon, or evening examination sessions. Of the 7,768 participants aged ≥ 18 years assigned to the morning session, 7,345 (95%) had valid FPG results; the remaining 5% had fasted < 8 hours or had missing data. For the present analysis, we excluded 622 participants who reported a physician-diagnosis of diabetes. Thus, we analyzed 6,723 adult participants without known diabetes, who represent 195 million Americans.

Diabetes Risk Factors

From the household interview data, we analyzed information on self-reported age, sex, race/ethnicity, education, and income. We also examined two healthcare access variables: presence of health insurance and healthcare visit within the past 3 years. Finally, we evaluated six factors that are strongly related to diabetes risk: 1) physical activity based on moderate or vigorous activity for at least 10 minutes during the past 30 days, 2) body mass index (BMI, kilograms/meters²[kg/m²]) 3) hypertension (defined as systolic blood pressure ≥ 130 mmHg, diastolic blood pressure ≥ 85 mmHg, or current use of anti-hypertensive medications, 4) elevated waist circumference (≥ 102 cm for males or ≥ 88 cm for females), 5) elevated triglycerides (≥ 1.7 mmol/l [150 mg/dl]), and 6) low high-density lipoprotein (HDL) cholesterol (< 1.0 mmol/l [40 mg/dl for males or < 1.3 mmol/l [50 mg/dl] for females). The latter four criteria used thresholds defined by the metabolic syndrome.²⁹ Nearly all of the study subjects had data on these variables ($> 97\%$ for all). We excluded missing values as recommended by NHANES; derivation of imputed values was not required. Family history of diabetes was missing in nearly half of the subjects so was not included in the analysis.

Laboratory Measurements

The Diabetes Diagnostic Laboratory at the University of Missouri-Columbia measured plasma glucose using a hexokinase enzymatic method. Based on FPG results, we classified participants as normal (< 5.6 mmol/l [100 mg/dl]), impaired fasting glucose (5.6 to 7.0 mmol/l [100–125 mg/dl])

or diabetes (≥ 7.0 mmol/l [126 mg/dl]), according to standard diagnostic criteria endorsed by the American Diabetes Association (ADA).⁶ Our primary outcome measure was undiagnosed diabetes, as defined by FPG results.

The same laboratory also performed the HbA1c measurement using Primus CLC 330 and Primus CLC 385 (Primus Corporation, Kansas City, MO). The instruments utilize a boronate affinity high performance liquid chromatography (HPLC) system, which has been standardized to the reference method used for the Diabetes Control and Complications Trial.⁷ The affinity chromatographic method has demonstrated excellent, long-term precision with coefficient of variation $< 3\%$, and is not affected by the presence of hemoglobin variants S, C, D or elevated fetal hemoglobin (HbF).

Statistical Analysis

We performed statistical analyses using Stata 9.0 (College Station, TX). We applied the recommended subsample weights to account for unequal probabilities of selection and to represent the US population based on 2000 Census data. We estimated variance based on NHANES-provided masked variance units, using Taylor Series linearization method. All p-values are two-tailed with $p < 0.05$ considered statistically significant.

We calculated proportions with 95% CI for diabetes risk factor data for the entire sample and stratified by FPG results (undiagnosed diabetes (FPG ≥ 7.0 mmol/l [126 mg/dl]) and impaired fasting glucose (FPG ≥ 5.6 – 6.9 mmol/l [100–125 mg/dl])). To improve interpretability of the analysis, we converted age and BMI from continuous to categorical variables. We determined univariate associations between predictors and FPG outcome using Pearson's chi-square test.

Variables with $p < 0.10$ on univariate analysis were eligible for entry into a weighted multivariate logistic regression model for the outcomes of undiagnosed diabetes. We utilized a backward elimination technique and with reference strata defined as age 18–44, non-Hispanic white, high school graduate, and BMI 20–24.9 kg/m²). Only significant predictors ($p < 0.05$) were retained in the final model, unless a variable significantly confounded the estimated associations. No interaction terms were evaluated for this model. We presented the results of two final models, one that evaluated all variables and the second that evaluated only variables that could be obtained during a non-fasting visit (excluded triglycerides and HDL). We reported results as OR with 95% CI for significant predictors of each outcome.

Using the results of the multivariate analysis, we derived a risk score for predicting undiagnosed diabetes based on the ORs in the final model, using previously described methodology.³⁰ We created a point score by dividing multivariate OR by 1.8 and rounding to the nearest 0.5 point. To determine categories for risk scores (low, moderate and high risk), we examined the prevalence of undiagnosed diabetes for each risk score and created categories to maximize homogeneity within each group.

We measured the positive predictive value (PPV) and negative predictive value (NPV) for a range of threshold HbA1c values (5.2 to 6.4%) in the total sample and risk-stratified subgroups (\geq threshold value considered abnormal, $<$ threshold value considered normal). We recommended a positive threshold to approximate a true positive rate (i.e., PPV) of 50%, and a

negative threshold to approximate a false negative rate (i.e., 1 – NPV) of 1%. Our rationale was to recommend further confirmatory FPG for patients above the positive threshold and no further testing for those below the negative threshold. We demonstrated these recommendations by creating Standards for Reporting of Diagnostic Accuracy (STARD) flow diagram for the total sample.³¹

Additionally, we determined sensitivity and specificity with 95% CI, and plotted a receiver operating characteristic (ROC) curve. We calculated area under the curve (AUC) as a measure of performance and determined 95% CI by the Bamber method.³²

RESULTS

The overall weighted prevalence was 6.3% for diagnosed diabetes, 2.8% for undiagnosed diabetes, and 27% for impaired fasting glucose. Subsequent analysis is based only on participants without known diabetes.

Risk Stratification

Table 1 shows the prevalence and univariate associations for diabetes risk factors, stratified by FPG results. Table 2 presents the results of the weighted multivariate logistic regression models and risk score development, based on the ORs in

Model 3 (all factors evaluated for inclusion). We calculated risk scores for participants with all required data (306 [4%] had at least one missing element) to predict the outcome of undiagnosed diabetes.

We assigned points based on the OR of significant predictors—age 45–64 (2.5 points), age ≥65 (4 points), male sex (1.5 points), black race (1 point), hypertension (1 point), elevated waist circumference (1.5 points), elevated triglycerides (1 point), and low HDL (1 point). We stratified resulting scores into low (0 to 4.5 points), moderate (5 to 6.5 points), and high (≥7.0 points) risk groups. The prevalence of undiagnosed diabetes increased across the risk groups: 0.44% for low, 4.1% for moderate, and 11.1% for high. Thus, the high risk group was 25 times more likely to have undiagnosed diabetes than the low risk group. The distribution of risk scores, stratified by risk factor data are presented in Table 3.

Hemoglobin A1c

The weighted mean HbA1c was 5.32% for the overall sample, 5.19% for normal FPG, 5.46% for impaired fasting glucose, and 7.07% for undiagnosed diabetes. Table 4 shows the PPV and NPV for a range of HbA1c threshold values (5.2–6.4%) in predicting undiagnosed diabetes for the entire sample and stratified by risk groups, using FPG as the gold standard. Additionally, we present the sensitivity and specificity values in Table 4 and plot the ROC curve for threshold values of HbA1c

Table 1. Prevalence of Impaired Fasting Glucose and Undiagnosed Diabetes Stratified by Risk Factors

Variable	n	US Population Estimate (95% CI) (in millions)*	IFG [†] % (95% CI)	Diabetes [†] % (95% CI)	p-value [‡]
Age					<0.001
18–44	3,565	107 (102–112)	18% (16–20)	0.8% (0.4–1.5)	
45–64	1,728	61 (55–67)	35% (32–39)	4.4% (3.4–5.8)	
≥65	1,430	27 (24–30)	45% (41–49)	7.1% (5.5–9.3)	
Male Sex	3,208	93 (91–96)	34% (31–37)	3.7% (3.0–4.5)	<0.001
Race/ethnicity					0.024
Non-Hispanic white	3,290	141 (133–148)	28% (25–31)	2.9% (2.4–3.5)	
Non-Hispanic black	1,321	21 (17–24)	19% (16–22)	3.0% (2.1–4.2)	
Mexican-American	1,596	14 (11–18)	28% (25–31)	2.1% (1.2–3.4)	
Other Hispanic	283	10 (5.5–15)	28% (21–35)	3.4% (1.6–7.2)	
Other/Multiracial	233	9.1 (6.8–11)	28% (21–36)	2.2% (0.8–5.8)	
Education level					<0.001
< High school	2,209	40 (36–44)	32% (29–36)	4.6% (3.6–5.7)	
High school grad	1,657	51 (46–57)	27% (24–31)	2.7% (2.0–3.8)	
> High school	2,843	103 (99–108)	25% (22–28)	2.1% (1.5–3.0)	
Income <\$20,000/year	1,491	32 (27–37)	30% (27–34)	3.9% (2.8–5.5)	0.02
No Health Insurance	1,475	36 (31–40)	25% (20–30)	2.5% (1.8–3.5)	0.21
Last healthcare visit ≥3 years ago/never	403	11 (9.1–13)	32% (24–41)	3.4% (1.7–6.9)	0.19
No exercise in past month	2,841	70 (62–77)	30% (28–32)	3.9% (3.1–4.9)	<0.001
BMI, kilograms/meters ²					<0.001
<20	399	11 (9.3–13)	12% (8–17)	0.4% (0.1–1.3)	
20–24.9	2,022	61 (56–66)	17% (14–20)	1.1% (0.7–1.7)	
25–29.9	2,281	65 (59–71)	32% (29–35)	2.8% (2.2–3.6)	
≥30	1,877	54 (49–59)	35% (33–38)	5.0% (4.0–6.2)	
Hypertension	2,629	72 (65–79)	38% (34–41)	5.5% (4.6–6.7)	<0.001
Waist circumference, ≥102 cm (M) or ≥88 cm (F)	3,161	89 (81–97)	34% (31–37)	4.6% (3.9–5.4)	<0.001
Triglycerides, ≥150 mg/dl	2,128	60 (54–66)	37% (34–40)	5.1% (4.1–6.4)	<0.001
HDL, <40 mg/dl (men) <50 mg/dl (women)	2,104	63 (57–69)	32% (29–36)	4.4% (3.5–5.7)	<0.001
TOTAL	6,723	195	27% (25–30)	2.8% (2.4–3.3)	

Abbreviations: IFG, impaired fasting glucose; BMI, body mass index; cm, centimeters; M, male; F, female; HDL, high-density lipoprotein

* Estimates based on weighted data from National Health and Nutrition Examination Survey (NHANES) 1999–2004

[†] Fasting plasma glucose classified according to American Diabetes Association threshold (5.6–7.0 mmol/l [100–125 mg/dl] for IFG; ≥7.0 mmol/l [126 mg/dl] for diabetes)

[‡] p-value based on Pearson's chi-square with <0.05 considered statistically significant

Table 2. Significant Predictors and Risk Score Development for Undiagnosed Diabetes by Multivariate Logistic Regression

Variable	Model 1* OR (95% CI)	Model 2* OR (95% CI)	Model 3* OR (95% CI)	Risk Score Points†
Age				
18–44	1.0 (referent)	1.0 (referent)	1.0 (referent)	
45–64	4.3 (2.7–9.4)	4.0 (1.9–8.1)	4.5 (2.2–9.5)	+2.5
≥65	6.7 (2.7–17)	5.7 (2.5–13)	7.2 (3.1–17)	+4
Male sex	2.4 (1.7–3.4)	2.4 (1.7–3.3)	2.4 (1.7–3.3)	+1.5
Race/ethnicity				
Non-Hispanic white	1.0 (referent)	–	1.0 (referent)	
Non-Hispanic black	1.4 (0.84–2.4)	–	1.8 (1.2–2.9)	+1
Mexican-American	0.83 (0.43–1.6)	–	1.3 (0.74–2.4)	
Other Hispanic	1.1 (0.37–3.5)	–	1.4 (0.58–3.6)	
Other/ Multiracial	1.7 (0.57–4.9)	–	1.5 (0.51–4.4)	
Education level				
< High school	1.3 (0.69–2.4)	–	–	
High school grad	1.0 (referent)	–	–	
> High school	0.84 (0.46–1.5)	–	–	
Income	1.3 (0.80–2.1)	–	–	
No exercise in past month	1.3 (0.78–2.2)	1.6 (1.1–2.4)	–	
BMI, kilograms/ meters ²				
<20	0.53 (0.14–2.0)	–	–	
20–24.9	1.0 (referent)	–	–	
25–29.9	0.98 (0.54–1.8)	–	–	
≥30	1.2 (0.61–2.5)	–	–	
Hypertension	2.0 (1.2–3.1)	2.0 (1.4–3.0)	2.1 (1.4–3.1)	+1
Waist circumference, ≥102 cm (M) or ≥88 cm (F)	2.5 (1.6–3.9)	3.5 (2.4–5.3)	3.0 (1.9–4.6)	+1.5
Triglycerides, ≥150 mg/dl	1.7 (1.1–2.7)	–	1.7 (1.1–2.5)	+1
HDL, <40 mg/dl (M) or <50 mg/dl (F)	1.8 (1.2–2.8)	–	2.0 (1.3–3.0)	+1

Abbreviations: OR, odds ratio; BMI, body mass index (kilograms/meters²); cm, centimeters; M, male; F, female; HDL, high-density lipoprotein

Variables with *p*<0.10 on univariate analysis eligible for entry into models. Backward stepwise elimination used, *p*<0.05 considered significant. Categorical variables dummy coded using indicated reference groups. Only significant predictors listed were included in the final models

* Model 1, all eligible risk factor variables included; Model 2, significant predictors only for non-fasting variables (excluded triglycerides and high-density lipoprotein); Model 3, significant predictors only with all variables evaluated

† Risk points based on OR/1.8 (rounded to nearest 0.5 point). Risk stratification: 0–4.5 points=low risk (0.44% prevalence); 5–6.5 points=moderate risk (4.1% prevalence); ≥7.0 points=high risk (11.1% prevalence)

in predicting undiagnosed diabetes in Figure 1. The AUC of 0.93 (95% CI, 0.91–0.95) revealed a high diagnostic accuracy for HbA1c.

We chose a positive threshold of HbA1c≥6.1%, which retained a PPV of approximately 50% for predicting undiagnosed diabetes in the entire sample. Additionally, the PPV for detecting impaired fasting glucose or undiagnosed diabetes (i.e., FPG≥5.6 mmol/l [100 mg/dl]) was 91% at this threshold. This threshold also had an overall 99% NPV, 98% specificity and 68% sensitivity. We chose the negative threshold of HbA1c ≤ 5.4%, which maintained a false negative rate of less than 1% for the entire and had an overall 73% specificity and 91% sensitivity. The intermediate values of HbA1c (5.5 to 6.0%) had variable performance depending on pretest probability, and risk stratification was useful in interpreting results in this range.

Figure 2 is a STARD flow diagram for this proposed algorithm of risk stratification and HbA1c in screening for undiagnosed diabetes. This figure displays data from individual NHANES participants to demonstrate the algorithm's potential utility in clinical practice.

Table 3. Distribution of Diabetes Risk Score by Demographic Variables and Significant Predictors of Undiagnosed Diabetes

Variable	Low Risk* % (95% CI)	Moderate Risk* % (95% CI)	High Risk* % (95% CI)
Age			
18–44	93% (92–94)	7.0% (5.9–8.3)	0.1% (0–0.3)
45–64	32% (29–34)	50% (47–53)	19% (16–21)
≥65	3.2% (2.0–5.2)	46% (42–49)	51% (48–55)
Male sex	54% (51–56)	30% (27–32)	17% (15–18)
Race/ethnicity			
Non-Hispanic white	59% (57–61)	27% (25–29)	14% (13–15)
Non-Hispanic black	61% (59–64)	23% (21–26)	15% (14–17)
Mexican-American	77% (73–81)	17% (14–20)	6% (4.5–7.7)
Other Hispanic	69% (59–78)	21% (14–29)	10% (6.8–14)
Other/Multiracial	69% (63–74)	25% (19–32)	7% (4.3–9.8)
Hypertension	27% (24–30)	43% (41–46)	30% (28–32)
Waist circumference, ≥102 cm (M) or ≥88 cm (F)	40% (38–43)	36% (34–38)	24% (22–25)
Triglycerides, ≥150 mg/dl	39% (35–42)	31% (28–34)	30% (28–33)
HDL, <40 mg/dl (M) <50 mg/dl (F)	53% (51–56)	23% (21–26)	23% (21–26)
TOTAL	62% (60–64%)	26% (24–27)	13% (12–14)

Abbreviations: cm, centimeters; M, male; F, female; HDL, high-density lipoprotein

Variables with *p*<0.10 eligible on univariate analysis for entry into model. Backward stepwise elimination used, *p*<0.05 considered significant. Categorical variables dummy coded using indicated reference groups. Only significant predictors listed were included in the final model

* Risk stratification: low risk=0.44% prevalence; moderate risk=4.1% prevalence; high risk=11.1% prevalence for undiagnosed diabetes

Table 4. Predictive Validity of Hemoglobin A1c in Screening for Undiagnosed Diabetes, Based on Threshold Values and Risk Stratification

HbA1c Threshold (n)	Total Sample				Risk Subgroups*		
	Sensitivity % (95% CI)	Specificity % (95% CI)	LR %	PPV, NPV %	Low Risk† PPV, NPV %	Moderate Risk† PPV, NPV %	High Risk† PPV, NPV %
		n=6,717			n=3,689	n=1,641	n=1,087
5.2% (776)	98 (94-99)	38 (35-41)	1.6	4, 100	1,100	5, 100	12, 98
5.3% (754)	98 (94-99)	50 (47-53)	2.0	5, 100	1, 100	6, 100	14, 99
5.4% (729)	94 (87-97)	62 (59-65)	2.5	7, 100	2, 100	7, 100	15, 97
5.5% (631)	91 (84-95)	73 (70-75)	3.3	9, 100	2, 100	9, 100	17, 97
5.6% (432)	88 (81-92)	82 (80-84)	4.9	12, 100	4, 100	12, 100	21, 97
5.7% (317)	88 (81-92)	88 (87-90)	7.5	18, 100	7, 100	17, 99	25, 97
5.8% (215)	86 (80-91)	93 (92-94)	12	25, 100	13, 100	23, 99	31, 98
5.9% (166)	79 (70-86)	95 (94-96)	16	32, 99	21, 100	27, 99	39, 97
6.0% (102)	70 (62-78)	97 (97-98)	24	41, 99	27, 100	38, 99	46, 96
6.1% (71)	68 (60-76)	98 (98-99)	36	51, 99	42, 100	49, 98	54, 96
6.2% (49)	62 (54-70)	99 (98-99)	47	58, 99	45, 100	52, 98	62, 97
6.3% (34)	59 (49-67)	99 (99-99)	73	68, 99	52, 100	61, 98	75, 95
6.4% (32)	55 (45-64)	99 (99-100)	105	75, 99	54, 100	76, 98	81, 95

Abbreviations: HbA1c, hemoglobin A1c; LR, likelihood ratio; PPV, positive predictive value; NPV, negative predictive value
Accuracy based on the gold standard of fasting plasma glucose ≥ 7.0 mmol/l (126 mg/dl)

* Unable to calculate risk score for 306 participants (missing values)

† low risk=0.44% prevalence; moderate risk=4.1% prevalence; high risk=11.1% prevalence for undiagnosed diabetes

DISCUSSION

We report the predictive validity of HbA1c as a screening test for undiagnosed diabetes in a large, nationally representative and risk-stratified population. Given that the overall prevalence of undiagnosed diabetes is relatively low (2% to 3%), risk stratification, although not performed in prior HbA1c screening studies,¹⁰⁻²¹ is an important tool to stratify pre-test probabilities in order to better estimate the true predictive value of HbA1c for subgroups of patients.

We found similar overall performance of HbA1c as in earlier reports. For example, in a meta-analysis of 18 smaller studies prior to 1996, Peters et al. found a comparable 66% sensitivity and 98% specificity for HbA1c $\geq 6.1\%$, with PPV 63%, using oral glucose tolerance test (OGTT) as the gold standard¹⁰. Rohlffing et al. analyzed the NHANES III dataset (collected 1988 to 1994) using FPG as the gold standard and found an area under the ROC curve of 0.90, which was similar but slightly lower than the current NHANES analysis.¹¹ Additionally, they reported comparable 63% sensitivity and 97% specificity for threshold HbA1c $\geq 6.2\%$, and 83% sensitivity and 84% specificity for HbA1c $\geq 5.7\%$. Davidson et al. also evaluated NHANES III but with OGTT as the gold standard, and found similar results.¹² These reports generated significant controversy on both sides of the debate.²²⁻²⁶ Although other investigators have found similar parameters for HbA1c in a variety of clinical settings,¹³⁻²¹ these data have been insufficient to recommend its implementation in clinical practice.⁶

To help address these concerns, we have presented a strategy of stratification by risk profile, before HbA1c testing, and interpretation of HbA1c based on the pre-test probability of diabetes. First, we recommend risk stratification by the proposed or other mechanisms to create risk groups of low (<0.5%), moderate (4% to 6%), and high (>10%) prevalence rates. Patients in the low risk group may generally be excluded, without further testing, based on the very low pre-test

probability of diabetes. The moderate and high risk groups may receive HbA1c screening at a variety of healthcare visit since fasting is not required.

We propose that patients in the moderate risk group receive FPG testing for HbA1c values in the positive range ($\geq 6.1\%$), and the diagnosis may be reliably excluded for intermediate and negative values ($\leq 6.0\%$). Patients in the high risk group may receive FPG testing if HbA1c $\geq 6.1\%$; the diagnosis can be reliably excluded if HbA1c $\leq 5.4\%$. For high-risk patients with intermediate HbA1c values (5.5% to 6.0%), clinical judgment and cost-effectiveness analysis are needed to guide decision-making, since the 6% prevalence of undiagnosed diabetes is

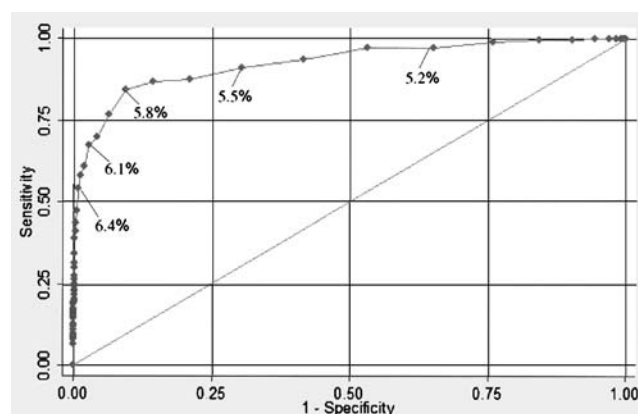


Figure 1. Receiver operating characteristic (ROC) curve for accuracy of hemoglobin A1c in predicting undiagnosed diabetes. Based on gold standard of fasting plasma glucose ≥ 7.0 mmol/l (126 mg/dl) to define undiagnosed diabetes. Area under the curve=0.93 (95% CI, 0.91-0.95).

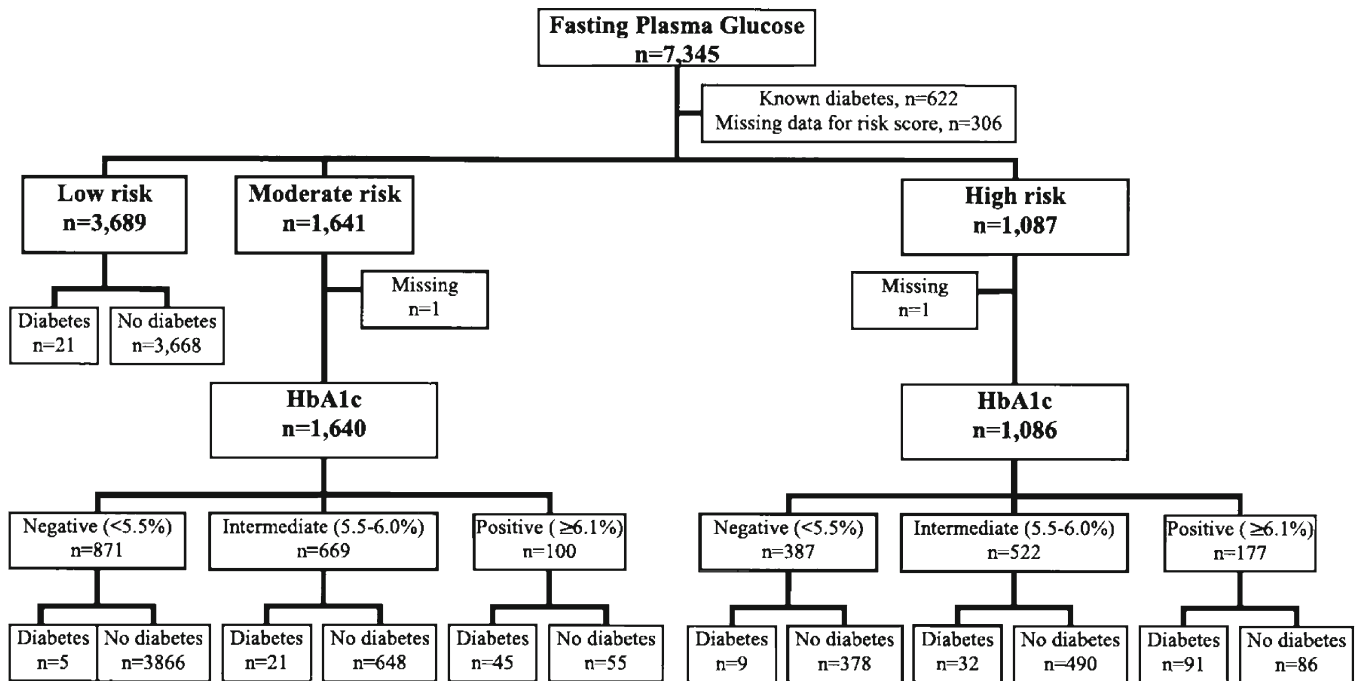


Figure 2. STARD diagram for detection of undiagnosed diabetes by risk stratification and hemoglobin A1c. Abbreviations: STARD, Standards for Reporting of Diagnostic Accuracy; HbA1c, hemoglobin A1c. Risk groups based on significant predictors of undiagnosed diabetes in multivariate analysis.

high enough that diabetes may not be excluded, but low enough that further testing may be relatively low yield. One strategy could be to recommend diet and lifestyle modification and aggressive management of hypertension and dyslipidemia for this group and then repeat screening (either HbA1c or FPG) after 1 to 3 years.

The risk factors of age, black race, and elements of the metabolic syndrome (hypertension, waist circumference, triglycerides, and HDL cholesterol) are commonly recognized as significant risk factors for diabetes. Interestingly, male sex was an independent risk factor for undiagnosed diabetes. Although not explored further, this finding has been reported in prior NHANES analysis.^{33,34} While the prevalence of undiagnosed diabetes varied by sex, the overall prevalence of diabetes (diagnosed plus undiagnosed) was similar. This disparity may merely reflect that males have 50% fewer outpatient preventive care than females,³⁵ but warrants further exploration. Although the overall prevalence of diabetes is higher among Hispanic individuals, we did not find an association between Hispanic ethnicity and the rate of undiagnosed diabetes, which is consistent with prior analyses.³⁴

While providers may choose to use different tools for risk stratification, the principle of deriving a low (<0.5%), moderate (4% to 5%) and high pre-test probability (>10%) could remain similar. Edelman et al. suggested a similar concept of the utility of HbA1c in a risk-stratified cohort from one outpatient clinic.^{14,17} Since fasting and well visits are not required for reliable HbA1c testing, the proposed screening algorithm would be more convenient for patients and providers and have potential for generalizability to a variety of unscheduled healthcare settings. Because of limited availability of some diabetes risk factor data in NHANES (i.e., family history,

gestational diabetes, and history of IFG/IGT), we chose to risk stratify based on best available data, which included fasting lipid profile. However, the concept of risk stratification prior to ordering and interpreting HbA1c still holds, and providers should risk stratify based on data available to them.

Although HbA1c is more expensive than serum glucose, opportunistic screening with HbA1c may prevent unnecessary fasting office visits in over two-thirds of moderate and high risk patients. Moreover, these strategies might advance two important *Healthy People 2010* objectives: to increase the rate that asymptomatic diabetes is diagnosed and to increase the overall proportion of adults whose condition has been diagnosed.⁴

Prior reports of diabetes screening in community and clinical venues have yielded mixed results, often limited by low prevalence rates and poor follow-up.³⁶⁻³⁸ Similar to any disease screening, patient adherence with confirmatory testing and subsequent therapy is vital to the successful implementation. Additionally, the cost-effectiveness of opportunistic diabetes screening is unclear and will require further investigation.³⁹ However, the US Preventive Services Task Force and ADA recommend diabetes screening for patients with defined diabetes risk factors, with ADA adopting broader criteria.^{5,6} Implementation of HbA1c as a screening test for risk-stratified populations may enhance screening efforts in appropriate high-risk patients.

The study has several potential limitations. NHANES 1999-2004 used FPG as the gold standard for diabetes, and OGTT were not performed. Although endorsed by the ADA in the diagnostic algorithm for diabetes,⁴⁰ FPG is not as accurate as OGTT for detecting glucose intolerance, especially in older adults.⁴¹ However, OGTT is more cumbersome and rarely used in US clinical practice. The algorithm developed was based on

persons who had apparently escaped diabetes diagnosis and may not perform similarly if used as a primary diagnostic strategy. We derived a new risk score based on the data available in NHANES in order to risk stratify the population, which would need to be validated in an independent population before implementation.

Additionally, several important known risk factors for diabetes, including family history, history of gestational diabetes, and history of IFG/IGT, were not available in this dataset. Nevertheless, we utilized this risk-stratification to create subgroups of different levels of risk—i.e., the results of our HbA1c analysis should apply regardless of the exact method of risk stratification. The proposed algorithm of risk stratification and HbA1c relied on pragmatic reasoning and interpretation of the data; others may suggest thresholds corresponding to different predictive values, and cost effectiveness analysis would further clarify optimal thresholds for clinical practice. Finally, this analysis provides a proposed algorithm, which, if validated, can serve as a guideline for providers, but should not substitute for sound clinical judgment for individual patients.

In summary, risk stratification improves the predictive validity of HbA1c in screening for undiagnosed diabetes in the US population. Opportunistic screening with HbA1c may be utilized in a variety of healthcare encounters and probably would improve detection of undiagnosed diabetes. Validation of our clinical algorithm and decision analysis should be performed to further understand the accuracy and optimal threshold values of HbA1c and to determine the cost-effectiveness of early detection of asymptomatic diabetes on a population level.

Acknowledgements: Dr. Ginde was supported by the Emergency Medicine Foundation Research Fellowship grant (Dallas, TX).

Conflict of Interest: Dr. Ginde has received investigator-initiated research funding from Bayer and Novo Nordisk in the past 3 years. Dr. Camargo has received financial support from a variety of groups for participation in conferences, consulting, and medical research. Over the past 3 years, industry sponsors with an interest in diabetes were AstraZeneca, Bayer, GlaxoSmithKline, Merck, Novartis, and Novo Nordisk. Dr. Cagliero and Dr. Nathan have no financial disclosures to report.

Corresponding Author: Adit A. Ginde, MD, MPH; Department of Emergency Medicine, University of Colorado Denver School of Medicine, 12401 E. 17th Avenue, B-215, Aurora, CO 80045, USA (e-mail: adit.ginde@uchsc.edu).

REFERENCES

- Centers for Disease Control and Prevention. National diabetes fact sheet: general information and national estimates on diabetes in the United States, 2005. Rev ed. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, 2005.
- Cowie CC, Rust KF, Byrd-Holt DD, et al. Prevalence of diabetes and impaired fasting glucose in adults in the US population. *Diabetes Care*. 2006; 29:1263–8.
- Harris MI, Klein R, Welborn TA, Knudman MW. Onset of NIDDM occurs at least 4–7 years before clinical diagnosis. *Diabetes Care*. 1992;15:815–9.
- US Department of Health and Human Services. Healthy people 2010. 2nd ed. With understanding and improving health and objectives for improving health. Section 5: Diabetes. Washington, DC: US Government Printing Office, November 2000.
- US Preventive Services Task Force. Screening for type 2 diabetes mellitus in adults: Recommendations and rationale. *Ann Intern Med*. 2003;1:212–4.
- American Diabetes Association. Standards of medical care in diabetes – 2007. *Diabetes Care*. 2007;30:S4–S41.
- The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med*. 1993;329:977–86.
- UK Prospective Diabetes Study Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet*. 1998;352:837–53.
- Woerle HF, Pimenta WP, Meyer C, et al. Diagnostic and therapeutic implications of relationships between fasting, 2-hour postchallenge plasma glucose and hemoglobin A1c values. *Arch Intern Med*. 2004;164:1627–32.
- Peters AL, Davidson MB, Schriger DL, Hasselblad V. A clinical approach for the diagnosis of diabetes mellitus: An analysis using glycosylated hemoglobin levels. *JAMA*. 1996;276:1246–52.
- Rohlfing CL, Little RR, Wiedmeyer HM, et al. Use of GHb (HbA1c) in screening for undiagnosed diabetes in the US population. *Diabetes Care*. 2000;23:187–91.
- Davidson MB, Schriger DL, Peters AL, Lorber B. Relationship between fasting plasma glucose and glycosylated hemoglobin: Potential for false-positive diagnoses of type 2 diabetes using new diagnostic criteria. *JAMA*. 1999;281:1203–10.
- Perry RC, Shankar RR, Fineberg N, McGill J, Baron AD. Early Diabetes Intervention Program (EDIP). HbA1c Measurement improves the detection of type 2 diabetes in high-risk individuals with nondiagnostic levels of fasting plasma glucose: The Early Diabetes Intervention Program (EDIP). *Diabetes Care*. 2001;24:465–71.
- Edelman D, Edwards LJ, Olsen MK, Dudley TK, Harris AC, Blackwell DK, Oddone EZ. Screening for diabetes in an outpatient clinic population. *J Gen Intern Med*. 2002;17:23–8.
- Greci LS, Kailasam M, Malkani S, Katz DL, Hulinsky I, Ahmadi R, Nawaz H. Utility of HbA1c levels for diabetes case finding in hospitalized patients with hyperglycemia. *Diabetes Care*. 2003;26:1064–8.
- Miyazaki M, Kubo M, Kiyohara Y, et al. Comparison of diagnostic methods for diabetes mellitus based on prevalence of retinopathy in a Japanese population: the Hisayama study. *Diabetologia*. 2004;47:1411–5.
- Edelman D, Olsen MK, Dudley TK, Harris AC, Oddone EZ. Utility of Hemoglobin A1c in Predicting Diabetes Risk. *J Gen Intern Med*. 2004;19:1175–80.
- Ellison RL, Elliott R, Moyes SA. HbA1c screening for undiagnosed diabetes in New Zealand. *Diabetes Metab Res Rev*. 2005;21:65–70.
- Peter R, Luzio SD, Dunseath G, Pauwaday V, Mustafa N, Owens DR. Relationship between HbA1c and indices of glucose tolerance derived from a standardized meal test in newly diagnosed treatment naïve subjects with type 2 diabetes. *Diabet Med*. 2006;23:990–5.
- Nakagami T, Tominaga M, Nishimura R, Yoskiike N, Daimon M, Oizumi T, Tajima N. Is the measurement of glycated hemoglobin A1c alone an efficient screening test for undiagnosed diabetes? Japan national diabetes survey [online]. *Diabetes Res Clin Pract*. 2007;76:251–6.
- Buell C, Kermah D, Davidson MB. Utility of hemoglobin A1c for diabetes screening in the 1999–2004 NHANES population. *Diabetes Care*. 2007; [In press].
- Vinacor F. When is diabetes diabetes [editorial]. *JAMA*. 1999;281:1222–4.
- Davidson MB, Schriger DL, Peters AL, Lorber B. Glycosylated hemoglobin as a diagnostic test for type 2 diabetes mellitus [editorial]. *JAMA*. 2000;283:605–6.
- Papoz L, Favier F, Clabé A. GHb (HbA1c) is more sensitive than fasting blood glucose as a screening test for diabetes [editorial]. *Diabetes Care*. 2000;23:1206–7.
- Herman WH, Engelgau MM, Zhang Y, Brown MB. Use of GHb (HbA1c) to screen for undiagnosed diabetes in the US population [editorial]. *Diabetes Care*. 2000;23:1207.
- Davidson MB. Catch-22 [editorial]. *Diabetes Care*. 2001;24:414.
- Little RR. Glycated hemoglobin standardization—National glycohemoglobin standardization program (NGSP) perspective. *Clin Chem Lab Med*. 2003;41:1191–8.
- Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination

- Survey Data. Hyattsville, MD: US Department of Health and Human Services, Centers for Disease Control and Prevention, 1999–2004. <http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm>. Accessed April 25, 2008.
29. Executive summary of the third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA*. 2001;285:2486–97.
 30. **Wells PS, Anderson DR, Rodger M, et al.** Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: Increasing the models utility with the SimpliRED D-Dimer. *Thromb Haemost*. 2000;83:416–20.
 31. **Bossuyt PM, Reitsma JB, Bruns DE, et al.** Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Ann Intern Med*. 2003;138:40–4.
 32. **Bamber D.** The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *J Math Psychol*. 1975;12:387–415.
 33. **Harris MI, Flegal KM, Cowie CC, et al.** Prevalence of diabetes, impaired fasting glucose and impaired glucose tolerance in US adults: The third national health and nutrition examination survey, 1988–1994. *Diabetes Care*. 1998;21:518–24.
 34. **Cowie CC, Rust KF, Byrd-Holt DD, et al.** Prevalence of diabetes and impaired fasting glucose in adults in the US population: National health and nutrition examination survey 1999–2002. *Diabetes Care*. 2006;29:1263–8.
 35. **Hing E, Cherry DK, Woodwell DA.** National ambulatory medical care survey: 2004 summary. *Adv Data*. 2006;374:1–33.
 36. **Ealovega MW, Tabaei BP, Brande M, Burke R, Herman WH.** Opportunistic screening for diabetes in routine medical practice. *Diabetes Care*. 2004;27:9–12.
 37. **Tabaei BP, Burke R, Constance A, et al.** Community-based screening for diabetes in Michigan. *Diabetes Care*. 2003;26:668–70.
 38. **Rolka DB, Narayan KM, Thompson TJ, et al.** Performance of recommended screening tests for undiagnosed diabetes and dysglycemia. *Diabetes Care*. 2001;24:1899–903.
 39. CDC Cost-effectiveness study group. The cost-effectiveness of screening for type 2 diabetes. *JAMA*. 1998;280:1757–63.
 40. **Stern MP, Williams K, Haffner SM.** Identification of persons at high-risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med*. 2002;136:575–81.
 41. **Motta M, Bennati E, Ferlito L, Passamonte M, Malaguarnera M.** Value and significance of new diagnostic criteria of diabetes mellitus in older people. *Arch Gerontol Geriatr*. 2007;45:103–8.