Original Paper

Predicting Current Glycated Hemoglobin Values in Adults: Development of an Algorithm From the Electronic Health Record

Brian J Wells[1], MD, PhD; Kristin M Lenoir[1], MPH; Jose-Franck Diaz-Garelli[2], PhD; Wendell Futrell[3], BS; Elizabeth Lockerman[4], MD; Kevin M Pantalone[5], DO; Michael W Kattan[6], PhD

[1]Division of Public Health Sciences, Department of Biostatistics and Data Science, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[2]Department of Physiology and Pharmacology, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[3]Clinical and Translational Science Institute, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[4]Department of Internal Medicine, Loyola University Medical Center, Maywood, IL, United States

[5]Endocrinology and Metabolism Institute, Department of Endocrinology, Diabetes and Metabolism, Cleveland Clinic, Cleveland, OH, United States

[6]Lerner Research Institute, Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, United States

Corresponding Author:
Brian J Wells, MD, PhD
Division of Public Health Sciences
Department of Biostatistics and Data Science
Wake Forest University School of Medicine
525 Vine Street, 4th Floor
Medical Center Boulevard
Winston-Salem, NC,
United States
Phone: 1 336 416 5185
Email: bjwells@wakehealth.edu

Abstract

Background: Electronic, personalized clinical decision support tools to optimize glycated hemoglobin (HbA$_{1c}$) screening are lacking. Current screening guidelines are based on simple, categorical rules developed for populations of patients. Although personalized diabetes risk calculators have been created, none are designed to predict current glycemic status using structured data commonly available in electronic health records (EHRs).

Objective: The goal of this project was to create a mathematical equation for predicting the probability of current elevations in HbA$_{1c}$ (≥5.7%) among patients with no history of hyperglycemia using readily available variables that will allow integration with EHR systems.

Methods: The reduced model was compared head-to-head with calculators created by Baan and Griffin. Ten-fold cross-validation was used to calculate the bias-adjusted prediction accuracy of the new model. Statistical analyses were performed in R version 3.2.5 (The R Foundation for Statistical Computing) using the rms (Regression Modeling Strategies) package.

Results: The final model to predict an elevated HbA$_{1c}$ based on 22,635 patient records contained the following variables in order from most to least importance according to their impact on the discriminating accuracy of the model: age, body mass index, random glucose, race, serum non–high-density lipoprotein, serum total cholesterol, estimated glomerular filtration rate, and smoking status. The new model achieved a concordance statistic of 0.77 which was statistically significantly better than prior models. The model appeared to be well calibrated according to a plot of the predicted probabilities versus the prevalence of the outcome at different probabilities.

Conclusions: The calculator created for predicting the probability of having an elevated HbA$_{1c}$ significantly outperformed the existing calculators. The personalized prediction model presented in this paper could improve the efficiency of HbA$_{1c}$ screening initiatives.

KEYWORDS

electronic health records; risk prediction; clinical decision support; hemoglobin A1c; diabetes

Introduction

Many prediction tools have been created to assess the risk of undiagnosed diabetes and related outcomes such as impaired glucose tolerance, prediabetes, risk of future diabetes, and hyperinsulinemia. Most of these tools are not practical in the setting of the electronic health record (EHR) because they include predictor variables not readily available in structured formats [1-20]. Examples of impractical variables include waist circumference, fasting time, physical activity, review of systems, diet, pregnancy-related variables, and detailed ethnicity. Tools typically leverage fasting glucose level as a predictor, which is simple to obtain in practice, but documentation of fasting time in a structured fashion in EHRs is generally absent. The authors identified two tools that accurately predict the presence of diabetes using structured variables routinely present in the EHR [21,22].

Current guidelines from the United States Preventive Services Task Force (USPSTF) recommend screening for abnormal blood glucose in adults aged 40 to 70 years who are overweight or obese. The USPSTF acknowledges that patients with other high-risk characteristics (eg, family history of diabetes, personal history of gestational diabetes) may need to be screened sooner but this is left up to the physician's discretion [23]. The guidelines published by the American Diabetes Association (ADA) recommend glucose screening of adult patients with an elevated body mass index (BMI; ≥ 25 kg/m^2) plus another risk factor (eg, hypertension, physical inactivity, family history of diabetes) at any age and for all patients beginning at 45 years of age at 3-year intervals [24].

Current approaches do not take advantage of advanced statistical modeling. Hyperglycemia risk prediction with simultaneous consideration of numerous independent variables and nonlinear effects is statistically ideal in the context of a multifactorial pathology. Creating strict cutoffs of individual variables or combinations of a limited number of variables for clinical guidelines does not take advantage of what can now be reasonably achieved. The USPSTF and ADA guidelines encourage physician judgment in the application of glucose screening but do not provide specific guidance. Simplified classification methods used in cancer have been notoriously poor at discriminating between high- and low-risk patients [25]. Moreover, many existing models for predicting hyperglycemia risk are also likely reducing their prediction accuracy by categorizing continuous variables, which reduces granularity and may miss potentially complex associations between a continuous variable and the outcome. This issue was highlighted by Kattan [26] when he showed that traditional regression techniques that incorporated restricted cubic splines to reduce linearity assumptions were found to produce more accurate risk prediction models when compared with classification methods such as classification and regression trees and artificial neural networks.

Predicting the date of onset of hyperglycemia is difficult due to the lack of symptoms early in the course of the disease and inconsistent testing and/or documentation in clinical practice (particularly in a structured fashion). Previous studies indicate that the onset of type 2 diabetes frequently occurs more than 5 years before diagnosis [27,28]. In contrast, blood measurements of glycated hemoglobin (HbA$_{1c}$) provide an easy and accurate method for determining current mean glycemia over the previous 8 to 12 weeks without the need for fasting. HbA$_{1c}$ testing is standardized according to specifications defined by the National Glycohemoglobin Standardization Program (NGSP). HbA$_{1c}$ levels are the primary blood marker used for guiding the management of type 2 diabetes, and the ADA has approved HbA$_{1c}$ testing for diabetes screening [24]. The increasing use of HbA$_{1c}$ as a screening tool in patients without prediabetes or diabetes provides data for prediction modeling from EHR records.

The authors strongly believe that the identification of patients with elevated HbA$_{1c}$ is important clinically despite previous studies that have not shown a mortality benefit from screening for diabetes [29]. The early detection of elevated blood sugar can have other significant benefits:

- Behavioral counseling can lead to reductions in cardiovascular disease risk [30].
- Treatment of prediabetes, which affects approximately 35% of the adult population in the United States [31], has been shown to delay progression to diabetes [32].
- Diabetic-specific retinopathy is present in up to 21% of patients with newly diagnosed type 2 diabetes [33], while peripheral neuropathy and nephropathy are present in 21.5% and 26.5%, respectively, of patients with undiagnosed diabetes [34]. Aggressive blood sugar and blood pressure control among patients with diabetes reduces the risk of microvascular complications [35,36].
- Early detection of diabetes allows for the allocation of proven preventive strategies (eg, fundoscopic screening for retinopathy, pneumococcal vaccination, screening for nephropathy, and aggressive prevention of cardiovascular disease) [37].
- Appropriate documentation of elevated blood sugar and diabetes allows health systems and payers to improve the risk stratification of patients and increases the potential pool of patients available for participation in clinical research.

Therefore, an accurate tool for predicting the current probability that a specific patient has an elevation in HbA$_{1c}$ levels would constitute a major advancement in finding patients with the most probable need of screening interventions. To address this gap, we created a calculator for predicting the probability that a given patient with no history of diabetes or elevated blood sugar currently has an elevated HbA$_{1c}$ value ($\geq 5.7\%$). This cutoff was chosen because it corresponds with the current guidelines published by the ADA that indicate values <5.7%

are considered to be normal. Importantly, the calculator presented in this paper was restricted only to structured variables typically available in EHRs. This focus on common structured variables will enable the tool to be integrated into EHRs for implementation.

Methods

This study was conducted on all adult patients who have undergone HbA_{1c} testing prior to evidence of hyperglycemia (random blood sugar ≥200 mg/dL), any diabetes-related diagnostic code, or prescription for an antihyperglycemic medication. Data were extracted from the Epicare EHR at Wake Forest Baptist Medical Center in Winston-Salem, North Carolina, for the dates between September 2012 and September 2016. The study was approved by the institutional review board and granted a waiver of informed consent. Data were limited to structured data located in these areas of the EHR: encounter diagnoses, problem list, past medical history, procedures, prescriptions, vital signs, demographics, social history, and laboratory values. Candidate predictor variables were chosen based on their theoretical association with hyperglycemia. Textbox 1 shows a list of the candidate predictor variables included in the complete statistical model.

Independent variables were defined on the date of the HbA_{1c} of interest. For missing continuous variables (eg, systolic blood pressure), the most recent prior value was used instead. Patients completely lacking values for independent variables were excluded. The investigators did not impute missing data because it was felt that imputation would not be appropriate at the point of implementation. Comorbidities were considered to be present if the patient had any structured instances of the diagnostic code on or before the date of the first HbA_{1c}. Medications with start dates on or before the date of the HbA_{1c} and end dates on or after the date of HbA_{1c} were considered to be active. Medication order dates were used when the start dates were missing. Medications missing both start and order dates were excluded. Medication categories (eg, antihyperglycemics) were provided by First Databank Inc. Multiple logistic regression was used to model the association between the independent variables and the outcome of HbA_{1c} ≥5.7%. Continuous variables were fit using restricted cubic splines with 3-knots. Due to collinearity, the model could not be fit with the simultaneous inclusion of serum non–high-density lipoprotein and high-density lipoprotein. Therefore, high-density lipoprotein was removed from the complete model. The model was reduced using Harrell's model approximation method [39]. For parsimony, the diagnosis of obesity variable was removed after variable selection. The diagnosis had little impact on the prediction accuracy and was redundant since BMI is also in the model. The reduced model was compared head-to-head with the calculators created by Baan and Griffin [21,22]. The head-to-head comparisons were performed using 10-fold cross-validation in order to calculate the bias-adjusted prediction accuracy of the new model. Prediction model metrics included measures of discrimination (concordance statistic), calibration (calibration curves), and decision curves [40]. Statistical analyses were performed in R version 3.2.5 (R Foundation for Statistical Computing) using the rms (Regression Modeling Strategies) package.

Textbox 1. Candidate variables in the complete model prior to variable selection.

Laboratory measurements:

- Serum triglycerides

- Random blood glucose

- Serum non–high-density lipoprotein

- Serum high-density lipoprotein (dropped due to an inability to fit model)

- Serum total cholesterol

- Estimated glomerular filtration rate (estimated from serum creatinine using the modified Chronic Kidney Disease Epidemiology Collaboration formula [38])

Active prescription medication categories:

- Antihypertensive

- First generation antipsychotic

- Second generation antipsychotic

- 3-hydroxy-3-methyl-glutaryl-*coenzyme A* reductase inhibitor (statin)

- Fibrate

- Valproic acid

- Beta-blocker

- Thiazide diuretic

- Niacin

- Oral glucocorticoid

- Protease inhibitor

- Nucleoside reverse transcriptase inhibitor

- Oral contraceptive

- Injectable medroxyprogesterone acetate

- Cyclosporine

- Sirolimus

- Tacrolimus

Diagnosis codes (see Multimedia Appendix 1):

- Hypertension

- Ischemic heart disease

- Peripheral vascular disease

- Neuropathy

- Obesity

- Hyperlipidemia

Vital signs:

- Systolic blood pressure

- Diastolic blood pressure

- Body mass index

Demographics:

- Race

- Age

- Gender

Family history:

- Number of first degree relatives with diabetes

Social history:

- Smoking status

Results

The record search identified 22,635 patients for model building and validation of which 26% were found to have an elevated HbA_{1c} (≥5.7%). Figure 1 shows the number of patients included and excluded from the model building.

The final model included the following 8 variables ordered from most important to least important: age, BMI, random glucose, race, serum non–high-density lipoprotein, serum total cholesterol, estimated glomerular filtration rate (eGFR), and smoking status. Table 1 shows descriptive statistics for the variables included in the final model by HbA_{1c} results. As expected, patients found to have elevated HbA_{1c} levels were older, had higher BMI, lower eGFR, and higher random glucose values.

The coefficients along with instructions for calculating the probability of an elevated HbA_{1c} (≥5.7%) and sample calculations for 2 patient scenarios are shown in Multimedia Appendix 2.

The 3 models were compared in their ability to accurately rank patients according to risk as measured by the concordance statistic (c-stat) and bias-adjusted using 10-fold cross-validation. The current model (c-stat 0.765, 95% CI 0.762 to 0.769) demonstrated statistically significant improvements in discrimination when compared to the models created by Baan (c-stat 0.637, 95% CI 0.633 to 0.641) and Griffin (c-stat 0.668, 95% CI 0.665 to 0.672).

The calibration curve shown in Figure 2 reveals that the current model is well calibrated. The predicted probabilities tend to overestimate risk at the right tail of the distribution, but the wide confidence intervals allude to the scarcity of the data at these extreme high levels of risk. Error bars represent the 95% confidence interval around the point estimate.

Decision curves are displayed in Figure 3 and also demonstrate the superiority of this model. Our model shows a net benefit up to a probability of 0.73 for an elevated HbA_{1c} (≥5.7%) without significant net harms above this threshold. This model confers a net benefit that is equal to or greater than the net benefit offered by the other models at all probability thresholds.

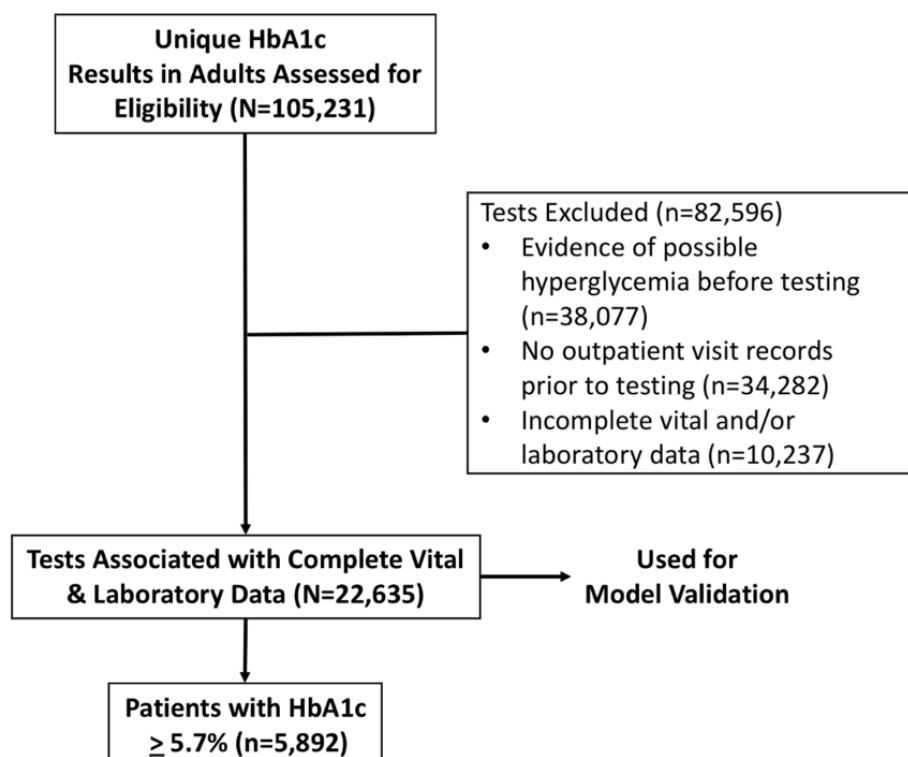**Figure 1.** Data flowsheet. HbA_{1c}: glycated hemoglobin.

Table 1. Descriptive statistics by glycated hemoglobin outcome.

Characteristics	HbA$_{1c}$[a] <5.7% (n=16,743)	HbA$_{1c}$ ≥5.7% (n=5892)	*P* value
Age (years) mean (SD)	48.1 (15.4)	54.8 (14.0)	<.001
Race, n (%)			<.001
Black	3692 (22.05)	2183 (37.05)	
Other	1178 (7.00)	487 (8.30)	
White	11873 (70.91)	3222 (54.68)	
BMI[b] (kg/m^2), mean (SD)	30.1 (7.44)	33.0 (8.41)	<.001
Smoking status, n (%)			<.001
Current smoker	2747 (16.41)	1393 (23.64)	
Former smoker	3867 (23.10)	1480 (25.11)	
Never smoker	10129 (60.50)	3019 (51.23)	
eGFR[c] (mL/min/1.73 m^2), mean (SD)	92.0 (33.0)	87.9 (30.8)	<.001
Random blood glucose (mg/dL), mean (SD)	88.4 (12.7)	96.1 (16.0)	<.001
Non-HDL[d] cholesterol (mg/dL), mean (SD)	135 (37.4)	144 (41.7)	<.001
Total cholesterol (mg/dL), mean (SD)	186 (39.4)	192 (43.1)	<.001

[a]HbA$_{1c}$: glycated hemoglobin.

[b]BMI: body mass index.

[c]eGFR: estimated glomerular filtration rate, calculated using the Chronic Kidney Disease Epidemiology Collaboration formula (CKD-EPI) [38].

[d]HDL: high-density lipoprotein.

Figure 2. Calibration curve of the new model for predicting glycated hemoglobin ≥5.7%.
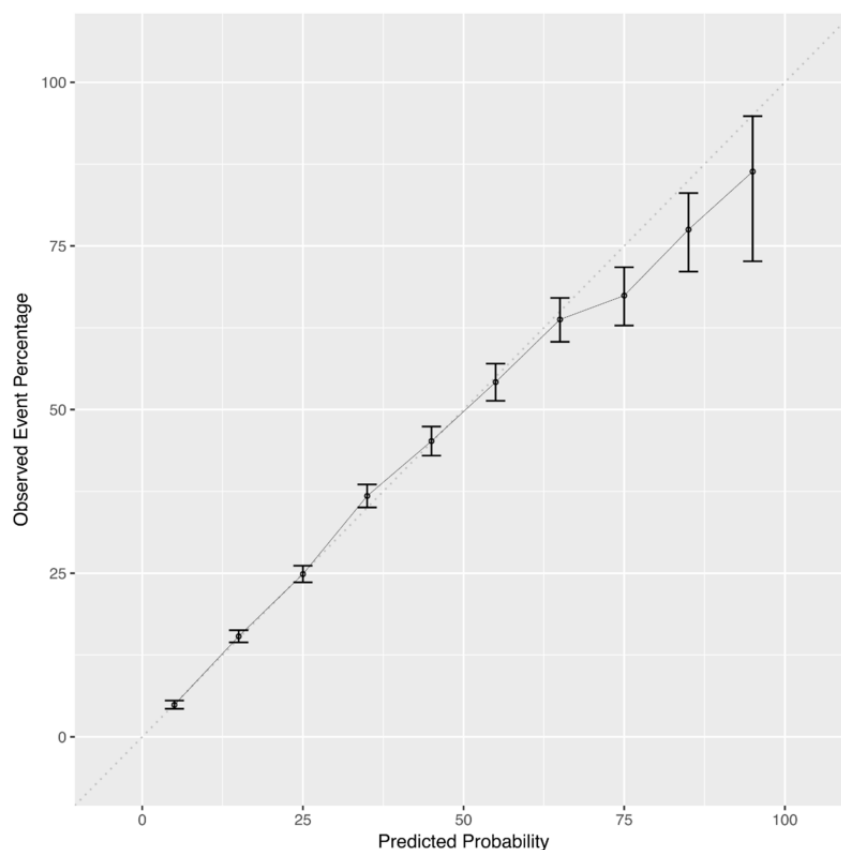
Figure 3. Decision curve analysis.
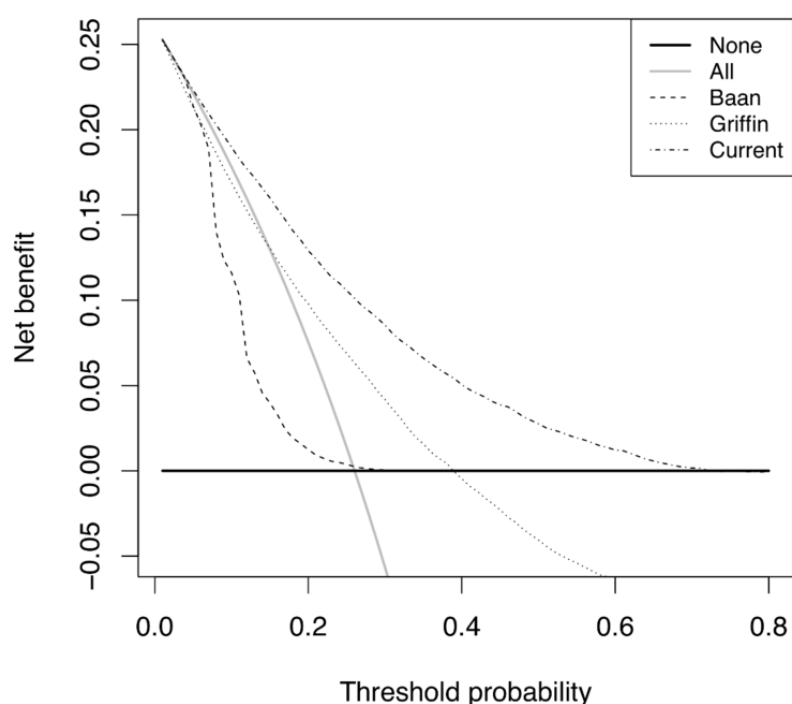
Discussion

Principal Findings

The calculator created for predicting the probability of having an elevated HbA_{1c} significantly outperformed the existing calculators. It should be noted that the calculators created by Baan and Griffin were designed for predicting current glucose tolerance test results and were not specifically calibrated to predict HbA_{1c} values. However, any potential issues with calibration should not impact the ability to discriminate patients according to risk. The authors chose not to develop a simple risk score, which would be easier to calculate without a computer but would be less accurate and would not provide an absolute probability. One of the benefits of using multivariable regression over many machine learning methods is that the mathematical output of this model can be integrated into an EHR using common mathematical operations. In contrast, classification-based methods like random forest, artificial neural networks, and classification and regression trees would increase the complexity of implementation by requiring separate software outside the EHR to calculate probabilities. The movement of data into and out of the EHR also raises concerns about security and privacy.

Limitations and Strengths

Limitations of the study include the lack of external validation. The model was validated internally using resampling and may not reflect the prediction accuracy that would be achieved in a prospective fashion at the current institution or when validated in a different health system. However, the authors used 10-fold cross-validation in which patients in the test data for each fold were not used to build the model. Another limitation pertains to the lack of data from outside health systems. Patients may

have additional medication or laboratory results outside of the health system that could alter the predicted risk or change the patient's status in terms of hyperglycemia. In order to ensure that patients have a minimal amount of data to guide the calculator's creation, the investigators required that patients had at least one value for each of the independent variables. Future research and quality improvement projects may need to query patients about health history prior to implementation.

A relatively small proportion of historical HbA_{1c} tests were appropriate for use in model building. Some of the tests were obtained before the installation of a comprehensive EHR and, therefore, accompanying information like vital signs were not available. Many of the tests were obtained in patients who already had evidence of possible hyperglycemia, some of whom were already being treated with antidiabetic medication. These patients would be inappropriate to use for the creation of a model aimed at patients with unknown glucose status. Limiting the model building and validation dataset to patients with complete data further reduced the sample size from 32,872 to 22,635. The authors chose not to impute the missing values given the adequate number of patients with complete data. In addition, the authors are not convinced that imputation would be acceptable to patients and providers when the model is implemented into practice. Patients lacking common variables used in the model such as BMI and blood pressure values are probably very new to the health system or are seeking their usual care elsewhere. Serum creatinine and lipid measurements are routinely obtained in clinical practice, especially among older adults. Patients without any creatinine or lipid measurements are likely to be younger and less likely to be at risk for diabetes. The authors felt it was important to identify a population for model building that matches the future population where the model will be implemented. Despite the restrictions

on data inclusion, the dataset contained >5000 patients with the outcome of interest. The size of this dataset is large compared to most similar studies conducted prior to the adoption of EHRs and is more than adequate for regression modeling. Harrel [39] has proposed that 7 to 10 outcomes for each degree of freedom are necessary to prevent overfitting when building a regression equation. The model created in this study contains 28 degrees of freedom, which could have safely been built from a dataset containing only 196 to 280 outcomes according to the aforementioned heuristic. A feasibility analysis was conducted among patients in the Department of Family and Community Medicine, and it was determined that approximately 20% of the adult patients seen in the past 3 years would be appropriate for application of the tool.

Imprecision in the measurement of HbA_{1c} levels could have negatively impacted the model building and could decrease the prediction accuracy of the model upon implementation. Wake Forest Baptist Medical Center is not a certified member of the NGSP but maintains accreditation by the Clinical Laboratory Improvement Amendments Program (identification number 34D0664386). The Wake Forest Baptist Medical Center's core laboratory performs HbA_{1c} testing using ion exchange high performance liquid chromatography, which is highly precise and constituted the vast majority of HbA_{1c} measurements used to create the data for this study. However, the investigators did not exclude HbA_{1c} measurements obtained using different methods at other locations in the health system (eg, point-of-care testing), which likely introduced variability in the HbA_{1c} measurements. Comorbid conditions such as iron deficiency anemia can lead to HbA_{1c} measurements that do not accurately reflect average blood glucose levels [41]. Despite the potential negative impact of imprecise or inaccurate HbA_{1c} measurements, the prediction model performed very well.

Conclusions

Improving the efficiency of diabetes screening should be of great interest in the United States given the increased use of value-based care contracts. Health systems could use our model for diabetes screening initiatives in a variety of ways. The decision curves suggest that using the new algorithm to guide HbA_{1c} testing would provide a net benefit between probabilities of 0.01 to 0.71. The authors will conduct a targeted screening study in which patients with a predicted risk of an elevated HbA_{1c} ≥50% will be notified directly regarding their elevated risk. Coupled with standing laboratory orders, this direct-to-patient design would enable patients to undergo HbA_{1c} testing prior to a physician visit. This is particularly important for patients with infrequent in-person visits. The hope is that patients with subsequent elevations in HbA_{1c} would be more likely to re-engage with the health system.

In summary, the risk equation created in this study is optimized for integration within an EHR and outperforms other similar models. Future research will attempt to integrate the risk calculator into clinical workflows, examine the ability of the calculator to predict risk in other health systems, and evaluate the potential economic savings of using this model for diabetes screening.

Authors' Contributions

BJW and KML were responsible for design of the study. BJW, KML, EL, and MWK were responsible for statistical analyses. BJW, KML, JFDG, EL, and KMP collaborated on the writing of the manuscript. JFDG performed the literature review. WF performed data extraction and editing. KMP contributed his clinical expertise.

Conflicts of Interest

MWK conducts research sponsored by Novo Nordisk and Merck that is not directly related to this project. In the past 12 months, KMP has received research support from Merck and Novo Nordisk that is not directly related to this project. In addition, in the past 12 months, KMP has received speaker honoraria from Merck, Novo Nordisk, and Astra Zeneca and consulting honoraria from Merck, Novo Nordisk, Sanofi, and Eli Lilly. The other authors report no potential conflicts of interest.

Multimedia Appendix 1

Diagnostic codes.

[PDF File (Adobe PDF File), 35KB-Multimedia Appendix 1]

XSL·FO
RenderX

Multimedia Appendix 2

Instructions for calculating the probability of an elevated glycated hemoglobin and sample calculations.

[PDF File (Adobe PDF File), 70KB-Multimedia Appendix 2]

References

1. Al-Lawati J, Tuomilehto J. Diabetes risk score in Oman: a tool to identify prevalent type 2 diabetes among Arabs of the Middle East. Diabetes Res Clin Pract 2007 Sep;77(3):438-444. [doi: 10.1016/j.diabres.2007.01.013] [Medline: 17306410]
2. Colagiuri S, Hussain Z, Zimmet P, Cameron A, Shaw J. Screening for type 2 diabetes and impaired glucose metabolism: the Australian experience. Diabetes Care 2004 Feb;27(2):367-371. [Medline: 14747215]
3. Glümer C, Carstensen B, Sandbaek A, Lauritzen T, Jørgensen T, Borch-Johnsen K, inter99 study. A Danish diabetes risk score for targeted screening: the Inter99 study. Diabetes Care 2004 Mar;27(3):727-733. [Medline: 14988293]
4. Herman W, Smith P, Thompson T, Engelgau M, Aubert R. A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes. Diabetes Care 1995 Mar;18(3):382-387. [Medline: 7555482]
5. Hidvegi T, Hetyesi K, Biro L, Jermendy G. Screening for metabolic syndrome in hypertensive and/or obese subjects registered in primary health care in Hungary. Med Sci Monit 2003;9(7):CR328-CR334.
6. Hippisley-Cox J, Coupland C. Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. BMJ 2017:359.
7. Kanaya A, Wassel F, de Rekeneire N, Shorr R, Schwartz A, Goodpaster B, et al. Predicting the development of diabetes in older adults: the derivation and validation of a prediction rule. Diabetes Care 2005 Feb;28(2):404-408. [Medline: 15677800]
8. Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care 2003 Mar;26(3):725-731. [Medline: 12610029]
9. Meigs J, Williams K, Sullivan L, Hunt K, Haffner S, Stern M, et al. Using metabolic syndrome traits for efficient detection of impaired glucose tolerance. Diabetes Care 2004;27(6):1417-1426.
10. Mohan V, Deepa R, Deepa M, Somannavar S, Datta M. A simplified Indian Diabetes Risk Score for screening for undiagnosed diabetic subjects. J Assoc Physicians India 2005 Sep;53:759-763. [Medline: 16334618]
11. Pearson T, Pronk N, Tan A, Halstenson C. Identifying individuals at risk for the development of type 2 diabetes mellitus. Am J Manag Care 2003 Jan;9(1):57-66 [FREE Full text] [Medline: 12549815]
12. Rauh SP, Heymans MW, Koopman ADM, Nijpels G, Stehouwer CD, Thorand B, et al. Predicting glycated hemoglobin levels in the non-diabetic general population: development and validation of the DIRECT-DETECT prediction model—a DIRECT study. PLoS One 2017 Feb;12(2):e0171816 [FREE Full text] [doi: 10.1371/journal.pone.0171816] [Medline: 28187151]
13. Rolka D, Narayan K, Thompson T, Goldman D, Lindenmayer J, Alich K, et al. Performance of recommended screening tests for undiagnosed diabetes and dysglycemia. Diabetes Care 2001 Nov;24(11):1899-1903. [Medline: 11679454]
14. Ruige J, de Neeling J, Kostense P, Bouter L, Heine R. Performance of an NIDDM screening questionnaire based on symptoms and risk factors. Diabetes Care 1997 Apr;20(4):491-496. [Medline: 9096967]
15. Saydah S, Byrd-Holt D, Harris M. Projected impact of implementing the results of the diabetes prevention program in the US population. Diabetes Care 2002;25(11):1940-1945.
16. Schmidt M, Duncan B, Vigo A, Pankow J, Ballantyne C, Couper D, ARIC Investigators. Detection of undiagnosed diabetes and other hyperglycemia states: the Atherosclerosis Risk in Communities Study. Diabetes Care 2003 May;26(5):1338-1343. [Medline: 12716785]
17. Schulze M, Hoffmann K, Boeing H, Linseisen J, Rohrmann S, Möhlig M, et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. Diabetes Care 2007 Mar;30(3):510-515. [doi: 10.2337/dc06-2089] [Medline: 17327313]
18. Stern M, Williams K, Haffner S. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? Ann Intern Med 2002 Apr 16;136(8):575-581. [Medline: 11955025]
19. Tabaei B, Engelgau M, Herman W. A multivariate logistic regression equation to screen for dysglycaemia: development and validation. Diabet Med 2005 May;22(5):599-605. [doi: 10.1111/j.1464-5491.2005.01467.x] [Medline: 15842515]
20. Wilson PWF, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. Arch Intern Med 2007 May 28;167(10):1068-1074. [doi: 10.1001/archinte.167.10.1068] [Medline: 17533210]
21. Griffin S, Little P, Hales C, Kinmonth A, Wareham N. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. Diabetes Metab Res Rev 2000;16(3):164-171. [Medline: 10867715]
22. Baan CA, Ruige JB, Stolk RP, Witteman JC, Dekker JM, Heine RJ, et al. Performance of a predictive model to identify undiagnosed diabetes in a health care setting. Diabetes Care 1999 Feb;22(2):213-219 [FREE Full text] [Medline: 10333936]
23. Siu AL, US Preventive Services Task Force. Screening for abnormal blood glucose and type 2 diabetes mellitus: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2015 Dec 01;163(11):861-868. [doi: 10.7326/M15-2345] [Medline: 26501513]

24.   Professional Practice Committee. Standards of medical care in diabetes—2018. Diabetes Care 2018;41(Supplement 1):S3. [doi: 10.2337/dc18-Sppc01]

25.   Kattan M. Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: preoperative application in prostate cancer. Curr Opin Urol 2003 Mar;13(2):111-116. [doi: 10.1097/01.mou.0000058631.64616.54] [Medline: 12584470]

26.   Kattan MW. Comparison of Cox regression with other methods for determining prediction models and nomograms. J Urol 2003 Dec;170(6 Pt 2):S6-S9. [doi: 10.1097/01.ju.0000094764.56269.2d] [Medline: 14610404]

27.   Harris M, Klein R, Welborn T, Knuiman M. Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis. Diabetes Care 1992 Jul;15(7):815-819. [Medline: 1516497]

28.   Porta M, Curletto G, Cipullo D, Rigault de la Longrais R, Trento M, Passera P, et al. Estimating the delay between onset and diagnosis of type 2 diabetes from the time course of retinopathy prevalence. Diabetes Care 2014 Jun;37(6):1668-1674. [doi: 10.2337/dc13-2101] [Medline: 24705614]

29.   Selph S, Dana T, Blazina I, Bougatsos C, Patel H, Chou R. Screening for type 2 diabetes mellitus: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2015 Jun 02;162(11):765-776. [doi: 10.7326/M14-2221] [Medline: 25867111]

30.   Lin J, O'Connor E, Whitlock E, Beil T. Behavioral counseling to promote physical activity and a healthful diet to prevent cardiovascular disease in adults: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2010 Dec 07;153(11):736-750. [doi: 10.7326/0003-4819-153-11-201012070-00007] [Medline: 21135297]

31.   Karve A, Hayward RA. Prevalence, diagnosis, and treatment of impaired fasting glucose and impaired glucose tolerance in nondiabetic U.S. adults. Diabetes Care 2010 Nov;33(11):2355-2359 [FREE Full text] [doi: 10.2337/dc09-1957] [Medline: 20724649]

32.   Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 7;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]

33.   Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, et al. Retinopathy in diabetes. Diabetes Care 2004 Jan 01:s84. [doi: 10.2337/diacare.27.2007.S84]

34.   Koopman RJ, Mainous AG, Liszka HA, Colwell JA, Slate EH, Carnemolla MA, et al. Evidence of nephropathy and peripheral neuropathy in US adults with undiagnosed diabetes. Ann Fam Med 2006 Sep;4(5):427-432 [FREE Full text] [doi: 10.1370/afm.577] [Medline: 17003143]

35.   UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes. BMJ 1998 Sep 12;317(7160):703-713 [FREE Full text] [Medline: 9732337]

36.   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 Sep 12;352(9131):837-853. [Medline: 9742976]

37.   American Diabetes Association. Standards of medical care in diabetes—2014. Diabetes Care 2014 Jan;37 Suppl 1:S14-S80. [doi: 10.2337/dc14-S014] [Medline: 24357209]

38.   Levey AS, Stevens LA, Schmid CH, Zhang Y, Castro AF, Feldman HI, CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009 May 05;150(9):604-612 [FREE Full text] [Medline: 19414839]

39.   Harrell F, Lee K, Mark D. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-387.

40.   Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006 Sep;26(6):565-574 [FREE Full text] [doi: 10.1177/0272989X06295361] [Medline: 17099194]

41.   Sacks DB, John WG. Interpretation of hemoglobin A1c values. JAMA 2014 Jun 11;311(22):2271-2272. [doi: 10.1001/jama.2014.6342] [Medline: 24915255]

Abbreviations

ADA: American Diabetes Association
CKD-EPI: Chronic Kidney Disease Epidemiology Collaboration formula
c-stat: concordance statistic
eGFR: estimated glomerular filtration rate
EHR: electronic health record
HbA$_{1c}$: glycated hemoglobin
NGSP: National Glycohemoglobin Standardization Program
USPSTF: United States Preventive Services Task Force

XSL•FO
RenderX

XSL·FO

RenderX