Journal of Diabetes and Its Complications 31 (2017) 1158–1163

Contents lists available at ScienceDirect

Journal of Diabetes and Its Complications

journal homepage: WWW.JDCJOURNAL.COM

Predicting the 6-month risk of severe hypoglycemia among adults with diabetes: Development and external validation of a prediction model

Emily B. Schroeder [a,b,*], Stan Xu [a], Glenn K. Goodrich [a], Gregory A. Nichols [c], Patrick J. O'Connor [d], John F. Steiner [a,b]

[a] Institute for Health Research, Kaiser Permanente Colorado, Denver, CO
[b] Department of Medicine, University of Colorado Denver, Aurora, CO
[c] Center for Health Research, Kaiser Permanente Northwest, Portland, OR
[d] HealthPartners Research Institute and HealthPartners Center for Chronic Care Innovation, Minneapolis, MN

ARTICLE INFO

Article history:
Received 24 February 2017
Received in revised form 30 March 2017
Accepted 3 April 2017
Available online 11 April 2017

Keywords:
Diabetes mellitus
Hypoglycemia
Prediction model
Risk factors
Electronic health records

ABSTRACT

Aims: To develop and externally validate a prediction model for the 6-month risk of a severe hypoglycemic event among individuals with pharmacologically treated diabetes.
Methods: The development cohort consisted of 31,674 Kaiser Permanente Colorado members with pharmacologically treated diabetes (2007–2015). The validation cohorts consisted of 38,764 Kaiser Permanente Northwest members and 12,035 HealthPartners members. Variables were chosen that would be available in electronic health records. We developed 16-variable and 6-variable models, using a Cox counting model process that allows for the inclusion of multiple 6-month observation periods per person.
Results: Across the three cohorts, there were 850,992 6-month observation periods, and 10,448 periods with at least one severe hypoglycemic event. The six-variable model contained age, diabetes type, HgbA1c, eGFR, history of a hypoglycemic event in the prior year, and insulin use. Both prediction models performed well, with good calibration and c-statistics of 0.84 and 0.81 for the 16-variable and 6-variable models, respectively. In the external validation cohorts, the c-statistics were 0.80–0.84.
Conclusions: We developed and validated two prediction models for predicting the 6-month risk of hypoglycemia. The 16-variable model had slightly better performance than the 6-variable model, but in some practice settings, use of the simpler model may be preferred.

1. Introduction

Hypoglycemia is a potentially life-threatening complication of diabetes treatment, particularly in individuals treated with insulin or sulfonylurea drugs that stimulate insulin secretion.[1–4] A clinically useful prediction tool for hypoglycemia would potentially enable providers and healthcare systems to identify patients at high risk and initiate anticipatory interventions to reduce the risk of severe hypoglycemia through goal modification, medication changes, or focused patient education.

A recent review by the VA QUERI program examined predictors of severe hypoglycemia in adults with type 2 diabetes.[1] Important risk factors for severe hypoglycemia included intensive glycemic control, history of hypoglycemia, renal insufficiency, history of microvascular complications, longer diabetes duration, lower education level, African American race, and history of dementia. Gender, age, and lower body mass index (BMI) were not consistently associated with risk of hypoglycemia, although higher age and lower BMI were associated with increased risk of hypoglycemia in the two largest individual studies. However, this review did not describe a clinically useful and externally validated prediction rule. Using data from DCCT/EDIC on 1441 individuals with type 1 diabetes, Lagani et al. developed a prediction rule for hypoglycemia.[2] They used a separate cohort of 393 individuals with type 2 diabetes as a validation data set. Their model included five variables, several of which are not easily collected from electronic medical records (marital status, strict vs. standard insulin regimen, total insulin daily dose, family history of type 2 diabetes, and past history of severe hypoglycemia).

For individuals using frequent home blood glucose monitoring or continuous glucose monitors, either with or without an insulin pump, computer algorithms have been developed to predict the very short term risk of hypoglycemia (over the next minutes to hours).[5,6] However, models based on clinical risk factors commonly available in electronic health records (EHRs) to predict the longer term risk of severe hypoglycemia (over days to months) are lacking.

The aims of this study were: (1) to develop a multivariable model to predict the 6-month risk of severe hypoglycemia requiring medical intervention among individuals receiving pharmacologic treatment for diabetes, within one integrated health care delivery system (Kaiser

The other authors have no conflicts of interest to disclose.

* Corresponding author at: 10065 E Harvard Ave, Suite 300, Denver, CO, 80231. Tel.: +1 303 614 1396; fax: +1 303 614 1395.
E-mail address: Emily.x.schroeder@kp.org (E.B. Schroeder).

http://dx.doi.org/10.1016/j.jdiacomp.2017.04.004
1056-8727/

Permanente Colorado) using information available in the electronic health record and other available clinical data sources; and (2) to externally validate the prediction model at two other sites (Kaiser Permanente Northwest and HealthPartners).

2. Subjects, materials, and methods

2.1. Study population

This study included three members of the SUPREME-DM (SUrveillance, PREvention, and ManagEment of Diabetes Mellitus) consortium, a group of 11 member organizations of the Health Care Services Research Network (HCSRN).[7,8] The development cohort was based in Kaiser Permanente Colorado (KPCO) which serves the Denver-Boulder metropolitan areas. We used two validation cohorts, based in Kaiser Permanente Northwest (KPNW; serving the Portland, OR and Vancouver, WA metropolitan areas) and HealthPartners (HP; Minneapolis, Minnesota). Research institutes embedded in these health systems have developed a distributed virtual data warehouse that contains information on demographics, outpatient pharmacy dispensing, laboratory tests and laboratory results, and diagnosis and procedure codes from outpatient and inpatient health care encounters from their electronic health record and administrative data systems.[8] The distributed virtual data warehouse allows for common variable definitions to be applied across study sites.

This study was approved by the Kaiser Permanente Colorado Institutional Review Board (IRB), and each participating site ceded oversight to the Kaiser Permanente Colorado IRB.

2.2. Cohort identification

We first identified a population of adults with diabetes using previously described methods.[7,9] Specifically, we defined the diabetes recognition date as the earlier of one inpatient diagnosis (ICD-9-CM 250.x, 357.2, 366.41, 362.01–362.07, either primary or secondary) or any combination of two of the following events, using the date of the first event in the pair as the identification date: 1) Hemoglobin A1c ≥ 6.5% (48 mmol/mol); 2) fasting plasma glucose ≥126 mg/dl; 3) random plasma glucose ≥200 mg/dl; 4) outpatient diagnosis code (same codes as for inpatient); 5) any single anti-hyperglycemic medication dispensing. When the two events were from the same source (e.g., two outpatient diagnoses or two elevated laboratory values), we required them to occur on separate days no more than 2 years apart. Dispensings of metformin or thiazolidinediones with no other indication of diabetes were not included because these agents may be used for diabetes prevention or to treat polycystic ovarian syndrome. Periods of pregnancy were excluded. Information on diabetes status was collected starting in 2000.

For each individual, the index date was the earliest date on or after 1/1/2007 when the following criteria were met: (1) diabetes case definition was satisfied, (2) enrolled continuously in the health plan for at least 6 months after the index date, (3) at least 12 months of continuous enrollment prior to the index date, (4) received any glucose lowering drug class medication on or within 100 days preceding the index date, and (5) at least 20 years of age on the index date. Individuals were censored at the earliest of: disenrollment for greater than 90 days, death, pregnancy, or 9/30/2015. For HealthPartners, which provides health insurance for patients receiving care outside HealthPartners Medical Group (HPMG) clinics, we additionally required that individuals have an available body mass index at the initial index date, indicating those who were receiving medical care within HPMG (and therefore would have complete EHR data).

2.3. Outcome assessment

We defined severe hypoglycemic events using a modified version of the algorithm initially developed by Ginde.[10] Our definition is based on ICD-9 codes (primary and secondary) collected in emergency departments and inpatient encounters. Events were included that met at least one of the following criteria: 1) a code for 251.0, 251.1, 251.2, or 962.3, or 2) a code for 250.8× without a code for 259.8, 272.7, 681.xx, 682.xx, 686.9x, 707.1–707.9, 709.3, 730.0–730.2, or 731.8. We also excluded events that were the result of intentional overdoses (ICD-9 codes: E950.x-E958.x, E980.x-E988.x, or V62.84 accompanied by 960.xx-989.xx or 870.xx-897.xx), as the predictors for those events likely differ from the predictors for unintentional hypoglycemic events. Finally, we grouped all events that occurred within 7 days and considered them to be a single event.

2.4. Candidate predictor variables

We started with the list of potential predictors identified in the systematic review by Bloomfield et al.[1] Because our population included individuals with both type 1 and type 2 diabetes, we also included diabetes type as a potential predictor. Using subject matter expertise, we narrowed the potential predictor variables to the 16 variables we felt were most likely to have good predictive abilities (Table 1) and would be readily available in most EHR systems. These included demographic variables (age, race/ethnicity), type of diabetes, body mass index, hemoglobin A1c, estimated glomerular filtration rate (eGFR), utilization (recent hospitalization, emergency department, and severe hypoglycemic event), important comorbidities (retinopathy, atherosclerotic cardiovascular disease, depression, and heart failure), and medication use (insulin, metformin, and number of classes of glucose lowering medications). Diabetes type was defined using a modification of an algorithm developed by Klompas et al.[11] For BMI, hemoglobin A1c, and serum creatinine, we used the most recent value in the 2 years preceding the initial index date. We estimated GFR using the CKD-EPI estimating equation, assuming non-black race if race was unknown.[12] We assessed whether there had been any hospitalization or utilization of the emergency department in the previous 365 days. Comorbidities were defined using relevant ICD-9 codes, as defined in Table 1. Medication use was defined as prescription dispensing for that medication type in the previous 100 days. In addition to the 16-variable model, we considered a simplified model with 6 variables: age, diabetes type (type 1 or 2), hemoglobin A1c, eGFR, history of a hypoglycemic event in the prior 365 days, and insulin use. These variables were selected a priori using expert knowledge, and were considered to be the minimum set of variables that clinicians would accept in a prediction model.

2.5. Observation periods

For both the 16- and 6-variable sets, we developed a model to predict the risk of a severe hypoglycemic event in the next 6-months, allowing individuals to have multiple 6-month observation periods. For the first observation period, we used the index date as the baseline date (beginning of the observation period), and defined the predictor variables at the time of the baseline date. For the second observation period, we used the index date + 182 days as the baseline date, and again defined the predictor variables at the time of the new baseline date. This process was continued until the individual was censored. For all variables except for diabetes type, we redefined the variables using the most recent information available at the time of the new baseline date. For BMI, hemoglobin A1c, and serum eGFR, we carried forward the most recent non-missing value, in order to minimize missing data. Within a given observation period, we only counted the first hypoglycemic event. However, individuals were not censored if