



Using real-time machine learning to prevent in-hospital hypoglycemia: a prospective study

Michael Fralick^{1,2} · Meggie Debnath² · Chloe Pou-Prom² · Patrick O'Brien² · Bruce A. Perkins^{3,4} · Esmeralda Carson⁵ · Fatima Khemani⁵ · Muhammad Mamdani^{2,4}

Received: 9 September 2022 / Accepted: 27 October 2022
© The Author(s), under exclusive licence to Società Italiana di Medicina Interna (SIMI) 2022

Dear Editor,

There are many examples of machine learning-based algorithms with impressive diagnostic characteristics [1], but few published studies have evaluated how well they perform when deployed into clinical care [2]. We recently published the results of a machine-learned model developed to predict inpatient hypoglycemia [3]. The objective of our current study was to evaluate its performance following implementation into clinical care on cardiovascular and vascular surgery ward. Patients on these wards are at particularly high risk of hypoglycemia because guidelines recommend tight glycemic control post-operatively [4].

Methods

We conducted a prospective analysis of a machine learning algorithm to predict hypoglycemia. The algorithm was trained, validated, and tested using data from 2013 to 2019. The details of the machine learning methods have been published, but in brief we employed multiple supervised machine learning techniques (e.g., extreme gradient

boosting) to predict inpatient hypoglycemia and severe hypoglycemia using a wide-range of patient-level data (i.e., features) including medications, labs, nursing notes, comorbid conditions, among others. Our deployed model was an extreme gradient boosting model. [3] Our study was part of a quality improvement initiative to reduce inpatient hypoglycemia and did not require informed consent.

The pre-implementation period for the model was Jan 1, 2018, to May 31, 2020, and the model was implemented on the cardiovascular surgery and vascular surgery ward at St. Michael's Hospital of Unity Health January 1, 2021 and evaluated until April 30, 2022. This ward was selected because diabetes is common (~30% of patients) and because the ward is managed by a small number of nurse practitioners which improves the feasibility of model deployment. Prior to implementation we met with the nurse practitioners to understand how best to provide them with the results of the algorithm. The nurse practitioners are responsible for the day-to-day clinical care of patients during their hospitalization from Monday to Friday. They requested that a daily email that included the names of the patients at highest risk of hypoglycemia would be the ideal approach. The email itself was generated by the algorithm and thus was entirely automated and we were able to track how often the email was opened. The email included a list of patients identified as highest risk by the model in the next 24 h and an additional list of patients that experienced a blood glucose level below 6.0 mmol/L [108 mg/dL] in the previous 24 h. This was to support the nurse practitioners in reviewing both patients anticipated for a hypoglycemia event and those who were actively trending low independent of the algorithm. The intent was to support our clinicians with daily actionable information for patients that were identified as high-risk at that point in time. No other information was provided in the email such as approaches to reduce the risk of hypoglycemia. This was purposeful because our end-users were

✉ Michael Fralick
mike.fralick@mail.utoronto.ca

¹ Division of General Internal Medicine, Sinai Health System, ON, Toronto, Canada

² Data Science and Advanced Analytics, Unity Health Toronto, Toronto, ON, Canada

³ Division of Endocrinology, Sinai Health System, Toronto, ON, Canada

⁴ Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

⁵ Division of Vascular and Cardiovascular Surgery, St Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada

already experts in preventing hypoglycemia based on their years of clinical experience.

Our primary outcome was the proportion of weeks before and after model implementation where more than 5% of patients experienced hypoglycemia (glucose < 3.9 mmol/L [70 mg/dL]) per week. Given the small number of patients on the study units we anticipated significant variability in rates of hypoglycemia and aggregated to weekly estimates. A clinically relevant metric of proportion of patients on the units experiencing hypoglycemia was chosen. The rates were calculated as the sum of (# encounters experiencing hypoglycemia, each day of the week) divided by sum of (daily patient census, for each day of the week) * 100. We used daily event and patient census estimates since the intervention was delivered on a daily basis. We assessed changes in the primary outcome graphically and using a Chi-square test. Secondary outcomes included changes in the variance of hypoglycemia rates, weekly rates of hypoglycemia annual rates of severe hypoglycemia (glucose < 2.2 mmol/L [40 mg/dL]), and weekly rates of hyperglycemia. Because severe hypoglycemia is rare we assessed the rate on a yearly basis. Levene's test was used to assess changes in variance between the pre-implementation and post-implementation periods and segmented regression was used to examine changes in weekly rates of hypoglycemia and hyperglycemia [5]. All statistical analyses were performed using R version 3.6.3.

Results

Our study included 3989 hospitalizations during the pre-implementation period and 1916 post-implementation. Baseline characteristics of patients, including comorbid conditions, were similar prior to and following implementation (Table 1). Approximately one-third of patients were women, the median age was 66 years, 23% received metformin in hospital, 7% received a sulfonylurea, and the median length of stay was 6 days. In Fig. 1, we provided a visual representation of the changes in the rate of the outcomes overtime. Following implementation of our model, we observed reductions in the rate of hypoglycemia. During the pre-implementation period 12 of the 127 weeks (i.e., 9.4% of weeks) there were more than 5% of patients who developed hypoglycemia. During the post-implementation period, 0 of the 79 weeks (i.e., 0% of weeks) had more than 5% of patients who developed hypoglycemia. The weekly variability in the rates of hypoglycemia decreased by approximately 50% from the pre-implementation (standard deviation 1.8, variance 3.4) to implementation phase (standard deviation 1.3, variance 1.6; $p=0.03$). There was a week-to-week decrease in hypoglycemia rates by 0.03 events per week [95% CI: - 0.04, - 0.01] ($p=0.004$) but no significant change in

Table 1 Baseline characteristics of included patients

	Pre-implementation	Implementation
Number of encounters	3989	1916
Median age (IQR)	67 (59–75)	66 (59–74)
Female sex	1052 (26%)	564 (29%)
Charlson comorbidity index ^a		
0	1288 (32%)	583 (30%)
1	922 (23%)	414 (22%)
2	871 (22%)	393 (21%)
3	538 (13%)	267 (14%)
4	185 (5%)	101 (5%)
Number of encounters with diabetes	1229 (30.8%)	572 (29.9%)
Insulin administration in-hospital	1942 (48.7%)	769 (40.1%)
Long-acting	540 (13.5%)	235 (12.3%)
Intermediate-acting	24 (0.6%)	8 (0.4%)
Short-acting	1899 (47.6%)	753 (39.3%)
Mix	69 (1.7%)	20 (1.0%)
Metformin use in-hospital	956 (24%)	445 (23%)
Sulfonylurea use in-hospital	384 (10%)	132 (7%)
Median length of stay in days (IQR)	6.3 (4.5–9.0)	6.6 (4.7–10.1)

^aA commonly used score for classifying comorbid conditions which higher scores indicating a higher number of comorbid conditions. Scores 5 and higher not included and constitute the remaining patients (~5%)

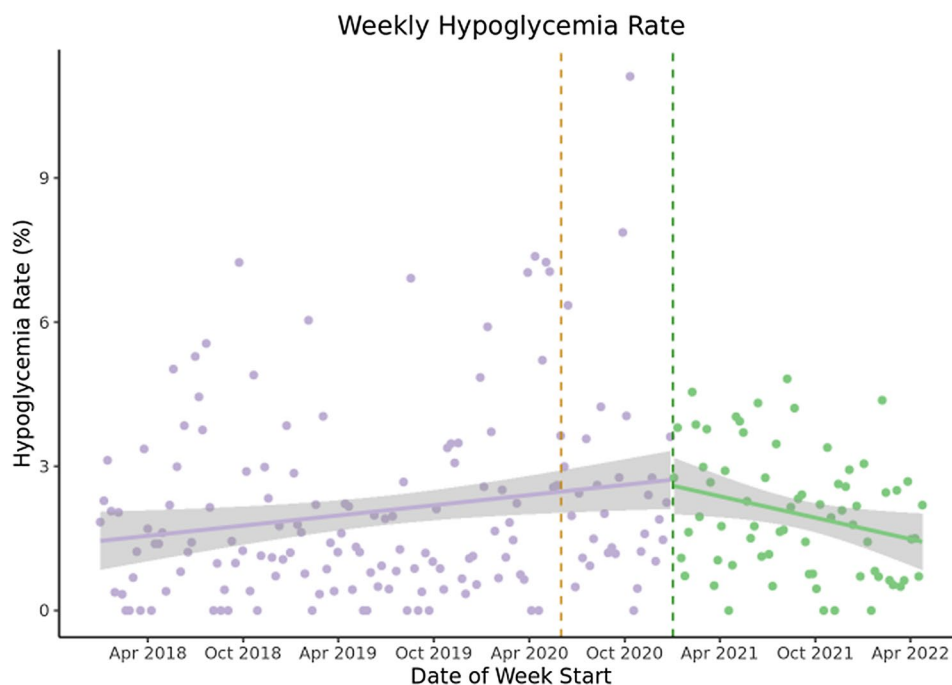
weekly rates of hyperglycemia (- 0.04 [95% CI: - 0.10, 0.01]; $p=0.102$). The severe hypoglycemia events per 100 patients per year was 1.3 pre-implementation and 1.1 following implementation.

Discussion

Our prospective analysis of a recently validated machine learned model [3] to prevent hypoglycemia demonstrated a reduction in the rates of inpatient hypoglycemia. And while there are other studies that have sought to predict inpatient hypoglycemia [6], most have not been prospectively evaluated to assess their performance in routine care.

There are other strategies to prevent hypoglycemia in hospital such as having a virtual glucose management service [7]. In a study including 3 hospitals in California, the implementation of this service, which consisted of a physician, nurse educator, and pharmacist, reduced rates of hypoglycemia by approximately 40% and the absolute number of severe hypoglycemic events (< 2.2 mmol/L [40 mg/dL]) was reduced from 40 to 15 per year following implementation. The reduction in the rate of both hypoglycemia and severe hypoglycemia was impressive, but it is unclear how cost-effective, sustainable, or generalizable

Fig. 1 Rates of hypoglycemia before and after model implementation. Legend—the green dashed line indicates the time the model was fully implemented whereas the dashed orange line indicates the start of the transition period



this model is. In contrast, our model does not require additional clinical team members and is constantly reviewing all of the available data for each patient, 24 h per day and 7 days per week. One potential unintended impact of our approach is a cognitive bias to only focus on the patients that appeared on the daily email. However, if this were leading to missed cases of hypoglycemia among lower risk patients we would not have expected to observe a reduction in the number of events.

An important limitation of our study is that it occurred at a single hospital in Toronto, Ontario with only 1 year of data to evaluate its implementation. However, this is a necessary step before wider adoption to ensure the tool is achieving adequate performance. With only 1 year of implementation data we are likely under-powered to identify its impact on the rate of severe hypoglycemia because it is a rare event. Another limitation of all non-randomized studies is an inability to rule out unmeasured confounding or temporal changes that may have affected the primary outcome. For example, most of our implementation phase took place during the COVID-19 pandemic and prior data have shown that there was a marked reduction in the number of hospitalizations during this period for non-COVID related illness and increased severity of illness among those who did present with non-COVID related illness. Finally, we lacked qualitative feedback on our tool from the end-user (i.e., nurse practitioners). Collecting these data are an important area of future research to have a broader understanding of both the quantitative and qualitative impact of machine-learned models. Despite these limitations, the results of our study suggest that machine

learning methods can be leveraged to prevent inpatient hypoglycemia.

Author contributions Study concept and design: all authors; acquisition of data: all authors; analysis/interpretation of data: all authors; drafting of the manuscript: MF; critical revision of the manuscript: all authors; statistical analysis: all authors.

Funding This study was funded by Banting and Best Diabetes Centre.

Declarations

Conflict of interest B.A.P. has received speaker honoraria from Abbott, Medtronic, Insulet, and Novo Nordisk; has served as an advisor to Insulet, Sanofi, and Abbott; and has received research support to his research institute from Novo Nordisk, and the Bank of Montreal (BMO). MF is a consultant for ProofDx, a start-up company that has created a point of care diagnostic tests for COVID-19 using CRISPR.

Human and animal rights statement and Informed consent The consent was not required as part of this QI study.

References

1. Dayan I, Roth HR, Zhong A et al (2021) Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat Med* 27(10):1735–1743
2. Escobar GJ, Liu VX, Schuler A, Lawson B, Greene JD, Kipnis P (2020) Automated identification of adults at risk for in-hospital clinical deterioration. *N Engl J Med* 383(20):1951–1960
3. Fralick M, Dai D, Pou-Prom C, Verma AA, Mamdani M (2021) Using machine learning to predict severe hypoglycaemia in hospital. *Diabetes Obes Metab* 23(10):2311–2319

4. Diabetes Canada Clinical Practice Guidelines Expert Committee (2018) Diabetes Canada 2018 clinical practice guidelines for the prevention and management of diabetes in Canada. *Can J Diabetes* 42(Suppl 1):S1–S5
5. Levene H (1960) Contributions to probability and statistics: Essays in Honor of Harold Hotelling. In: I Olkin ed, Stanford University Press 278–292
6. Mathioudakis NN, Abusamaan MS, Shakarchi AF et al (2021) Development and validation of a machine learning model to predict near-term risk of iatrogenic hypoglycemia in hospitalized patients. *JAMA Netw Open* 4(1):e2030913
7. Rushakoff RJ, Sullivan MM, MacMaster HW et al (2017) Association between a virtual glucose management service and glycemic control in hospitalized adult patients: an observational study. *Ann Intern Med* 166(9):621–627

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.