



Open camera or QR reader and scan code to access this article and other resources online.



ORIGINAL ARTICLE

# The Type 1 Diabetes and EXercise Initiative: Predicting Hypoglycemia Risk During Exercise for Participants with Type 1 Diabetes Using Repeated Measures Random Forest

Simon Bergford, MS,<sup>1</sup> Michael C. Riddell, PhD,<sup>2</sup> Peter G. Jacobs, PhD,<sup>3</sup> Zoey Li, MS,<sup>1</sup> Robin L. Gal, MSPH,<sup>1</sup> Mark A. Clements, MD, PhD,<sup>4</sup> Francis J. Doyle, PhD,<sup>5</sup> Corby K. Martin, PhD,<sup>6</sup> Susana R. Patton, PhD, CDE,<sup>7</sup> Jessica R. Castle, MD,<sup>8</sup> Melanie B. Gillingham, PhD,<sup>9</sup> Roy W. Beck, MD, PhD,<sup>1</sup> Michael R. Rickels, MD, MS,<sup>10</sup> Peter Calhoun, PhD<sup>1</sup>; for the T1DEXI Study Group

## Abstract

**Objective:** Exercise is known to increase the risk for hypoglycemia in type 1 diabetes (T1D) but predicting when it may occur remains a major challenge. The objective of this study was to develop a hypoglycemia prediction model based on a large real-world study of exercise in T1D.

**Research Design and Methods:** Structured study-specified exercise (aerobic, interval, and resistance training videos) and free-living exercise sessions from the T1D Exercise Initiative study were used to build a model for predicting hypoglycemia, a continuous glucose monitoring value <70 mg/dL, during exercise. Repeated measures random forest (RMRF) and repeated measures logistic regression (RMLR) models were constructed to predict hypoglycemia using predictors at the start of exercise and baseline characteristics. Models were evaluated with area under the receiver operating characteristic curve (AUC) and balanced accuracy.

**Results:** RMRF and RMLR had similar AUC (0.833 vs. 0.825, respectively) and both models had a balanced accuracy of 77%. The probability of hypoglycemia was higher for exercise sessions with lower pre-exercise glucose levels, negative pre-exercise glucose rates of change, greater percent time <70 mg/dL in the 24 h before exercise, and greater pre-exercise bolus insulin-on-board (IOB). Free-living aerobic exercises, walking/hiking,

<sup>1</sup>JAEB Center for Health Research, Tampa, Florida, USA.

<sup>2</sup>Muscle Health Research Centre, York University, Toronto, Canada.

<sup>3</sup>Artificial Intelligence for Medical Systems Lab, Department of Biomedical Engineering, Oregon Health and Science University, Portland, Oregon, USA.

<sup>4</sup>Children's Mercy Hospital, Kansas City, Missouri, USA.

<sup>5</sup>Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA.

<sup>6</sup>Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, Louisiana, USA.

<sup>7</sup>Nemours Children's Health, Jacksonville, Florida, USA.

<sup>8</sup>Harold Schnitzer Diabetes Health Center, Oregon Health and Science University, Portland, Oregon, USA.

<sup>9</sup>Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, Oregon, USA.

<sup>10</sup>Institute for Diabetes, Obesity and Metabolism, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.

and physical labor had the highest probability of hypoglycemia, while structured exercises had the lowest probability of hypoglycemia.

**Conclusions:** RMRF and RMLR accurately predict hypoglycemia during exercise and identify factors that increase the risk of hypoglycemia. Lower glucose, decreasing levels of glucose before exercise, and greater pre-exercise IOB largely predict hypoglycemia risk in adults with T1D.

**Keywords:** Type 1 diabetes, Risk, Prediction, Hypoglycemia, Exercise.

## Introduction

REGULAR EXERCISE IS important for maintaining physical fitness, insulin sensitivity, and a healthy cardiovascular risk profile in individuals living with type 1 diabetes (T1D).<sup>1</sup> However, exercise-associated hypoglycemia remains a major barrier for many children<sup>2</sup> and adults<sup>3–5</sup> living with T1D. Based on one recent survey, ~50% of active adults with T1D have frequent episodes of exercise-associated hypoglycemia (during, soon after, or overnight). Rates of “in exercise” hypoglycemia were surprisingly higher in those on continuous subcutaneous insulin infusion (CSII) and in those who report to be knowledgeable about hypoglycemia prevention strategies for exercise (i.e., insulin dose reductions, carbohydrate feeding), compared with those on multiple daily injections (MDI) or those less knowledgeable about the typical hypoglycemia prevention strategies.<sup>6</sup>

The reasons for the higher hypoglycemia reporting rates during exercise in those on CSII are unclear but might be related to lower glucose levels before exercise start time and/or higher levels of circulating insulin levels pre-exercise compared with those on MDI.

Prediction models for when exercise-associated hypoglycemia will occur for individuals with T1D may be useful to incorporate into a hybrid closed-loop (HCL) system or into a mobile Health (mHealth) application that provides decision support. These models can identify clinical and behavioral targets that can be modified to reduce the risk of exercise-associated hypoglycemia. Most forms of physical activity can be detected, and perhaps even classified (i.e., aerobic, anaerobic, mixed), by the continuous monitoring of heart rate and/or accelerometry.<sup>7–10</sup> Once activity is detected, an HCL system can reduce or suspend the insulin delivery even before glucose begins to fall,<sup>11</sup> and the algorithms may even be enhanced by machine learning<sup>12</sup> or adaptive learning of personalized models.<sup>10</sup>

Even if a reduction in insulin delivery is insufficient to eliminate exercise-associated hypoglycemia, because of the slow pharmacokinetic changes of current insulin formulations and delivery strategies, the system could also either recommend carbohydrate intake based on prediction models that include estimated insulin on board (IOB) and glucose rates of change or automatically deliver small doses of glucagon to help prevent hypoglycemia.<sup>13,14</sup> For non-HCL users, an mHealth application could provide decision support related to carbohydrate feeding or modification of insulin dosing pre-exercise.

In one study, Turksoy et al.<sup>15</sup> described a method for predicting hypoglycemia using exercise metrics such as energy expenditure and exercise type. In another study, based on a decision tree and a random forest model, Reddy et al.<sup>16</sup>

identified that at least two factors (baseline glucose less than 180 mg/dL and a heart rate threshold of 120 beats per minute) were important in predicting exercise-associated hypoglycemia. However, this analysis was limited to in-clinic observations of aerobic exercise only, and it is likely that several variables other than pre-exercise glucose level influence the glucose rate of change during exercise such as the type of exercise, the duration of exercise performed, bolus IOB levels, and the exercise time of day. Recently, the Type 1 Diabetes and EXercise Initiative (T1DEXI) study group found that several participant-level and exercise-level factors influenced the glucose rate of change during study video exercises.<sup>17</sup>

However, whether these factors are useful for predicting exercise-associated hypoglycemia during exercise is unclear. In addition, a sophisticated model that can assess many factors simultaneously, handle nonlinear relationships, and consider multiple interactions may improve predictive performance. The purpose of this study was to build an accurate model that can predict hypoglycemia during exercise and identify key physiologic measures that are important predictors of hypoglycemia during exercise. Structured (i.e., video sessions of aerobic, resistance, or interval-type exercises) and free-living exercise sessions from the T1DEXI study cohort were used to train and evaluate machine learning algorithms for predicting hypoglycemia during exercise.

## Methods

### *Study cohort and design*

The T1DEXI study was an observational, at-home study designed to collect a variety of data around structured and free-living exercise for adults with T1D. The study has been described elsewhere<sup>17</sup> and summarized herein. The protocol received institutional review board approval and participants provided informed consent. In brief, participants  $\geq 18$  years of age with a minimum 2-year duration of T1D who were using a commercially approved HCL system, a standard CSII pump, or MDI to administer insulin were randomly assigned to complete one of three types of study-designed structured exercise videos (aerobic, resistance, intervals) ~30 min in length, at home for at least six sessions over 4 weeks. Participants also continued their typical forms of daily physical activity, including structured and free-living exercise sessions, and used a study-developed, cloud-connected smartphone application (T1DEXI app) to enter information about their nutrient intake and exercise.<sup>18</sup>

Diabetes history, glycated hemoglobin A1c (HbA1c), and demographics were self-reported and collected via an online portal. Participants used their personal Dexcom G6 continuous glucose monitoring (CGM), or a study-provided blinded

Dexcom G6 CGM (San Diego, CA) if they did not use a personal Dexcom G6 CGM, and an investigational Verily Study Watch (South San Francisco, CA) to collect continuous glucose and heart rate data throughout the 4-week study period.

### Statistical methods

Exercise sessions included in the analysis met the following requirements: (1) exercise duration between 20 and 90 min, (2)  $\geq 15$  min of CGM data during exercise, (3) a glucose level  $\geq 70$  mg/dL at start of exercise, and (4) non-missing data for the predictors included in the model. Participant-level predictors included age, sex, T1D duration, HbA1c, body mass index, insulin modality (MDI, pump, or HCL), International Physical Activity Questionnaire (IPAQ) metabolic equivalent (MET)-minutes per week, and Pittsburgh Sleep Quality Index measured as participant-reported hours spent asleep. Exercise-level predictors included type of exercise, self-reported exercise intensity, exercise start time entered into the smartphone application, glucose level at start of exercise, glucose rate of change at start of exercise, IOB at the start of exercise, heart rate at start of exercise, and percent time  $< 70$  mg/dL in the 24 h before start of exercise.

Exercise type was reported by the participants through the T1DEXI app by selecting from a prespecified list of 19 activities or by entry of a custom activity. Of the exercise sessions that were included in the analysis, 48 physical activity/exercise activities were initially reported and then grouped into the following eight categories: study exercise video aerobic, interval, or resistance; sports; chores and/or physical labor; walking/hiking; other workout aerobic; and other workout strength. Exercise intensity was based on self-report without a guide to represent low, medium, or high intensity. A description of each predictor can be found in Supplementary Table S1.

The objective was to predict whether an exercise session will have at least one CGM value  $< 70$  mg/dL during the activity, which is deemed low enough according to clinical practice guidelines/consensus to stop exercise and treat with carbohydrates (i.e., level 1 hypoglycemia).<sup>19</sup> Two models were trained to predict hypoglycemia during exercise: a repeated measures random forest (RMRF) model and a repeated measures logistic regression (RMLR) model. Exercise sessions were split into “model training” and “model testing” data sets by participant: 80% of the participants with all their exercise sessions were included in the model training data set, and the remaining 20% of participants with all their exercise sessions were included in the model testing data set.

The best RMRF and RMLR models were chosen based on performance on the training data, and the predictive performance of each was compared using the test data. Area under the receiver operating characteristic curve (AUC; higher values representing better predictions) and Brier score (lower values representing better predictions) were calculated for the training and test data. Balanced accuracy (higher values representing better predictions) was also determined for the test data when classifying exercises as high or low risk using Youden’s index.

A post hoc analysis compared the predictive performance of RMRF and RMLR on two additional hypoglycemia outcomes: (1) at least one CGM value  $< 60$  mg/dL during the activity, and (2) at least one CGM value  $< 54$  mg/dL or at least

one CGM value  $< 100$  mg/dL with a snack during the activity. The best RMRF and RMLR models were retrained for the new outcome, and the predictive performance was evaluated using the test data.

### RMRF model

The RMRF model<sup>20</sup> extends the random forest model by Breiman<sup>21</sup> by analyzing binary outcomes that are correlated within participant. An RMRF model grows several large classification trees by subsampling the participants and re-iteratively splitting the predictors. An unbiased error rate can be derived from the RMRF model. Details of the RMRF model and some of its advantages are provided in the Supplementary Materials and further described in Calhoun et al.<sup>20</sup> The predictors included in the final random forest were selected by starting with all 16 predictors, removing the predictor with the lowest variable importance, and then fitting a new RMRF. These steps were repeated until there was only one predictor remaining. The AUC was calculated for each RMRF model, and the model with the greatest AUC was selected. The final model was optimized based on tuning parameters to grow the tree (Supplementary Materials).

Only one final model was chosen, and we did not reassess the model after predicting the test data. Partial dependence plots (PDPs) were constructed to show how the predictors affected the probability of hypoglycemia. PDPs were created for the variables in the final RMRF model and combined both the training and test data sets. Importantly, PDPs with RMRF illustrate the relationship between predictors and hypoglycemia risk allowing for nonlinear relationships and interactions.

### RMLR model

An RMLR model was fitted using a generalized estimating equation to estimate the parameters using an exchangeable correlation structure. The error rate with the training data was estimated by using fivefold cross-validation. A backward variable selection method was used to determine the optimal number of predictors in the RMLR model. This method started with the full model with 16 predictors, removed the predictor with the greatest robust Wald statistic *P*-value, fitted a new RMLR model, and repeated these steps until there was only one predictor remaining. The AUC was calculated for each RMLR model using cross-validation, and the model with the greatest AUC was chosen as the final RMLR model.

## Results

The analysis included 459 participants with a total of 8827 exercise sessions. Mean  $\pm$  standard deviation age was  $37 \pm 14$  years; 73% were female, 19% used MDI, 36% using standard CSII, and 46% using HCL (Table 1). Mean HbA1c was  $6.6\% \pm 0.8\%$ . Mean glucose level at start of exercise was  $149 \pm 50$  mg/dL and the mean glucose rate of change at the start of exercise was  $0.0 \pm 1.2$  mg/dL/min (Table 2). Walking or hiking was the most common form of exercise (33% of all exercise sessions), and 39% of exercise sessions were reported as low intensity, 48% as medium intensity, and 13% as high intensity. Median (quartiles) CGM use was high, with 99% (99%, 100%) of CGM data available during exercise.

TABLE 1. FREQUENCY OF HYPOGLYCEMIA BY PARTICIPANT CHARACTERISTICS

	No. of exercise sessions n (%)	% of exercise sessions with CGM <70 mg/dL
Overall	8,827 (100%)	8%
Age		
18–29 years	3266 (37%)	8%
30–39 years	1992 (23%)	8%
40–49 years	1503 (17%)	7%
≥50 years	2066 (23%)	10%
Sex		
Female	6470 (73%)	8%
Male	2357 (27%)	10%
HbA1c		
≤6.0%	1732 (20%)	10%
6.1%–6.4%	2182 (25%)	9%
6.5%–6.9%	2315 (26%)	7%
≥7.0%	2598 (29%)	8%
T1D duration		
<10 years	2416 (27%)	6%
10–<20 years	3408 (39%)	9%
≥20 years	3003 (34%)	9%
BMI		
Underweight (<18.5 kg/m <sup>2</sup> )	18 (<1%)	11%
Normal (18.5–<25.0 kg/m <sup>2</sup> )	5014 (57%)	8%
Overweight (25.0–<30.0 kg/m <sup>2</sup> )	2780 (31%)	9%
Obese (≥30.0 kg/m <sup>2</sup> )	1015 (11%)	8%
Insulin modality		
MDI	1567 (18%)	7%
Pump	3229 (37%)	10%
Hybrid closed-loop	4031 (46%)	8%
IPAQ MET-minutes per week		
<1500 MET	2503 (28%)	6%
1500–<3000 MET	3017 (34%)	8%
≥3000 MET	3307 (37%)	10%
PSQI hours asleep		
<7.5 h	1414 (16%)	10%
7.5–<8.5 h	3517 (40%)	9%
≥8.5 h	3896 (44%)	8%

BMI, body mass index; CGM, continuous glucose monitoring; HbA1c, glycated hemoglobin A1c; IPAQ, International Physical Activity Questionnaire; MDI, multiple daily injections; MET, metabolic equivalent; PSQI, Pittsburgh Sleep Quality Index; T1D, type 1 diabetes.

The AUC was consistently higher for RMRF compared with RMLR for the training data set when using the same number of predictors except for when the only predictor was glucose at the start of exercise (Fig. 1; Supplementary Table S2). The selected RMRF model with the greatest AUC had the following 10 predictors (by order of importance): glucose level at start of exercise, glucose rate of change at start of exercise, percent time <70 mg/dL 24 h before exercise, IOB at the start of exercise, type of exercise, exercise start time, T1D duration, insulin modality, IPAQ MET-minutes per week, and sex. The optimal RMRF parameters were a minimum *P*-value threshold of 0.10 and minimum

TABLE 2. FREQUENCY OF HYPOGLYCEMIA BY EXERCISE CHARACTERISTICS

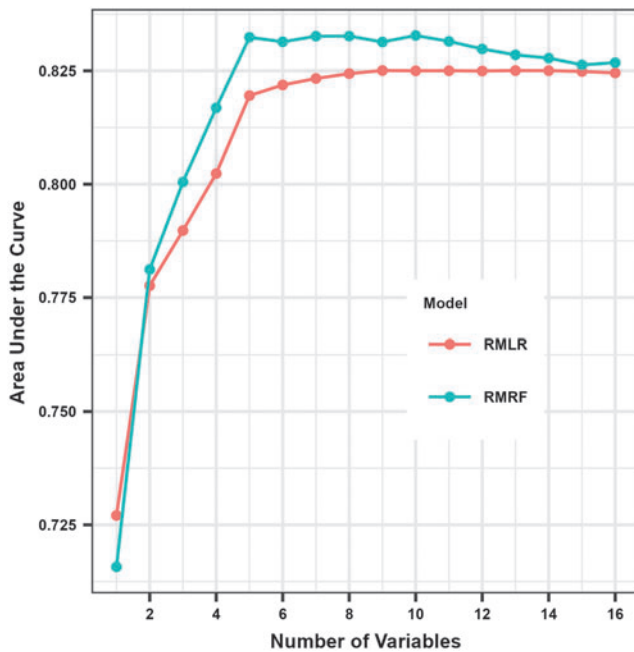
	No. of exercise sessions n (%)	% of exercise sessions with CGM <70 mg/dL
Overall	8,827 (100%)	8%
Type of exercise		
Study video aerobic	657 (7%)	4%
Study video interval	652 (7%)	3%
Study video resistance	668 (8%)	3%
Sports	580 (7%)	9%
Chores and physical labor	605 (7%)	10%
Walking/hiking	2897 (33%)	6%
Workout aerobic	2191 (25%)	4%
Workout strength	577 (7%)	3%
Exercise intensity level		
High	1145 (13%)	9%
Medium	4239 (48%)	9%
Low	3443 (39%)	8%
Exercise time of day		
Morning (3 AM–<12 PM)	3111 (35%)	8%
Afternoon (12 PM–<6 PM)	3621 (41%)	6%
Evening (6 PM–<9 PM)	1624 (18%)	9%
Night (9 PM–<3 AM)	471 (5%)	12%
Glucose at start of exercise		
70–<100 mg/dL	1332 (15%)	23%
100–<140 mg/dL	3158 (36%)	9%
140–<180 mg/dL	2318 (26%)	5%
180–<250 mg/dL	1637 (19%)	2%
≥250 mg/dL	382 (4%)	2%
Glucose rate of change at start of exercise		
<-0.5 mg/dL/min	2490 (28%)	16%
-0.5 to <0.5 mg/dL/min	3909 (44%)	6%
≥0.5 mg/dL/min	2428 (28%)	5%
% time <70 mg/dL 24 hours before exercise		
<1%	3961 (45%)	5%
1%–<4%	2505 (28%)	8%
≥4%	2361 (27%)	14%
IOB at start of exercise		
<1 U	3874 (44%)	6%
1–<2 U	1788 (20%)	9%
2–<3 U	1191 (13%)	10%
≥3 U	1974 (22%)	12%
Heart rate at start of exercise		
<75 beats/min	2066 (23%)	7%
75–<95 beats/min	4256 (48%)	9%
≥95 beats/min	2505 (28%)	9%

IOB, insulin on board.

number of exercise sessions to continue splitting of 10, but the AUC only slightly increased when optimizing the parameters (Supplementary Table S3).

The selected RMLR model had the following nine predictors (by order of importance): glucose level at start of exercise, glucose rate of change at start of exercise, IOB at the start of exercise, percent time <70 mg/dL 24 h before exercise, type of exercise, exercise start time, insulin modality, IPAQ MET-minutes per week, and heart rate at start of exercise.

By selecting the models with the greatest AUC when evaluated on the training data set, the estimated predictive



**FIG. 1.** AUC of RMRF and RMLR in training data set based on number of predictors included. Individual data points (i.e., dots) indicate training AUC for RMRF and RMLR based on number of predictors in the model, with higher numbers indicating better prediction performance. The training AUC is estimated using the out-of-bag sample for RMRF and using fivefold cross-validation for RMLR. AUC, area under the receiver operating characteristic curve; RMLR, repeated measures logistic regression; RMRF, repeated measures random forest.

performance may be overly optimistic. Thus, the true error rate was estimated with a test data set: the AUC was 0.833 with RMRF and 0.825 with RMLR (Table 3, Supplementary Fig. S1). Balanced accuracy was 77% for both models, and Brier score was also similar for the two models (0.055 for

**TABLE 3.** PREDICTIVE PERFORMANCE FOR REPEATED MEASURES RANDOM FOREST AND REPEATED MEASURES LOGISTIC REGRESSION FOR TEST DATA SET

Model	AUC	Brier score	Optimal cut point		
			Sensitivity	Specificity	Balanced accuracy
RMLR <sup>a</sup>	0.825	0.056	75%	78%	77%
RMRF <sup>b</sup>	0.833	0.055	70%	84%	77%

<sup>a</sup>RMLR model with an exchangeable covariance structure and glucose level at start of exercise, glucose rate of change at start of exercise, insulin on board at the start of exercise, % time <70 mg/dL 24 h before exercise, type of exercise, exercise start time, insulin modality, IPAQ MET-minutes per week, and heart rate at start of exercise as covariates.

<sup>b</sup>RMRF model with an exchangeable covariance structure, maximum depth of 10, 500 trees, and glucose level at start of exercise, glucose rate of change at start of exercise, % time <70 mg/dL 24 h before exercise, insulin on board at the start of exercise, type of exercise, exercise start time, T1D duration, insulin modality, IPAQ MET-minutes per week, and sex as covariates.

AUC, area under the receiver operating characteristic curve; RMLR, repeated measures logistic regression; RMRF, Repeated measures random forest.

RMRF vs. 0.056 for RMLR). The predictive performance of RMRF and RMLR for two other hypoglycemic outcomes during exercise found similar results with a slightly higher AUC for RMRF compared with RMLR, but a similar balanced accuracy and Brier score (Supplementary Table S4).

The variable importance of the predictors for the final RMRF is shown in Figure 2. Lower glucose values at the start of exercise (i.e., values <125 mg/dL) and a greater negative glucose rate of change at the start of exercise (i.e., <-0.5 mg/dL/min) were the most important predictors for hypoglycemia (Figs. 3 and 4). The probability of hypoglycemia was also increased for greater percent time <70 mg/dL in the 24 h before exercise and for higher IOB at the start of exercise. The mean probability of hypoglycemia was highest for free-living aerobic workouts (9.2%), walking/hiking (8.9%), and chores and physical labor (8.8%), while the mean probability of hypoglycemia was lowest for the study videos (6.6% interval and resistance, 6.8% aerobic; Fig. 5).

Higher IPAQ MET-minutes per week, T1D duration, exercise start time, insulin modality, and sex also had a small effect on the probability of hypoglycemia and were included in the final model (Supplementary Figs. S2–S5). Other factors including exercise intensity and heart rate at start of exercise did not affect the probability of hypoglycemia during exercise.

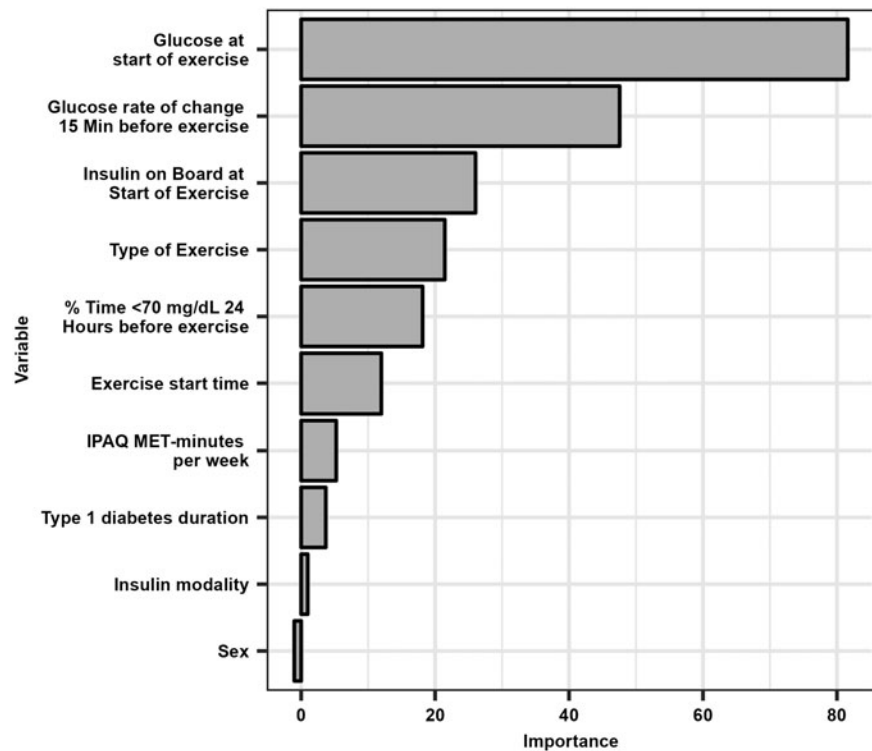
## Discussion

We present results on two glucose forecasting algorithms that can be used to predict hypoglycemia risk with some degree of precision/accuracy during real-world exercise in adults living with T1D. Importantly, these algorithms were trained and tested on the largest known data set of real-world exercise data in an adult T1D cohort currently available (<https://doi.org/10.25934/PR00008428>).<sup>17</sup> The AUC of RMRF and RMLR was similar (0.833 vs. 0.825, respectively) and both models had a balanced accuracy of 77%.

Hypoglycemia forecasting algorithms such as the RMRF and the RMLR models may be used by people with T1D, or future mHealth applications, to estimate the likelihood of hypoglycemia before performing exercise. If the algorithms are accurate, they can help the person avoid hypoglycemia and maintain better glycemic management during exercise while helping to alleviate fears about hypoglycemia. Overall, the risk for hypoglycemia during all forms of physical activity/exercise was ~8% in the T1DEXI data set. We identified that lower glucose levels just before exercise and a greater negative glucose rate of change at the start of exercise as highly associated with hypoglycemia during exercise. Specifically, the results from RMRF, shown in Figure 3, demonstrate that CGM-based glucose values at the start of exercise <125 mg/dL, and a glucose rate of change below -0.5 mg/dL/min had higher-than-average risk of hypoglycemia after adjusting for the other predictors in the model.

These results support current consensus recommendations to target both a stable glucose concentration at the start of exercise and a pre-exercise glucose of ~150 mg/dL when using CGM.<sup>19</sup> However, the results from RMRF, shown in Figure 4, also demonstrate that a lower glucose value at the start of exercise could be safe if the glucose rate of change is positive (i.e., glucose is increasing), perhaps because of pre-exercise insulin dose adjustments and/or carbohydrate

**FIG. 2.** Variable importance for the selected RMRF model. The gray bars represent the scaled permutation variable importance.



feeding, which are also recommended for planned exercise, particularly when CGM systems are not used.<sup>22</sup> These findings are important as they identify ways in which the risk of exercise-associated hypoglycemia can be minimized. Such risk mitigation measures may ultimately be integrated into future HCL and mHealth decision-support systems to help people with T1D better manage glucose during exercise. In addition, the PDPs produced from RMRF allow for an easy determination of the risk of hypoglycemia before exercise.

However, it should also be noted that other important factors that could be assessed during the exercise itself, such as relative exercise intensity based on heart rate levels or ratings of perceived exertion, competition stress, and estimated circulating insulin levels, may also be of value for integration into a hypoglycemia risk algorithm for automated insulin delivery systems designed for exercise.<sup>23</sup>

Besides glucose value and glucose rate of change at start of exercise, HCL systems could also incorporate the patient's prandial and/or correction IOB and prior time <70 mg/dL in the previous 24 h before exercise. The simplest RMRF model that included these four predictors plus exercise type had an out-of-bag AUC of 0.8324, similar to the highest out-of-bag AUC of 0.8328. Understanding how these factors interact with each other and affect the probability of hypoglycemia during exercise can be difficult for a person, but use of HCL systems could simplify the process by computing the probability of hypoglycemia and possibly suggesting carbohydrate intake or by automatically suspending insulin delivery when hypoglycemia risk increases.

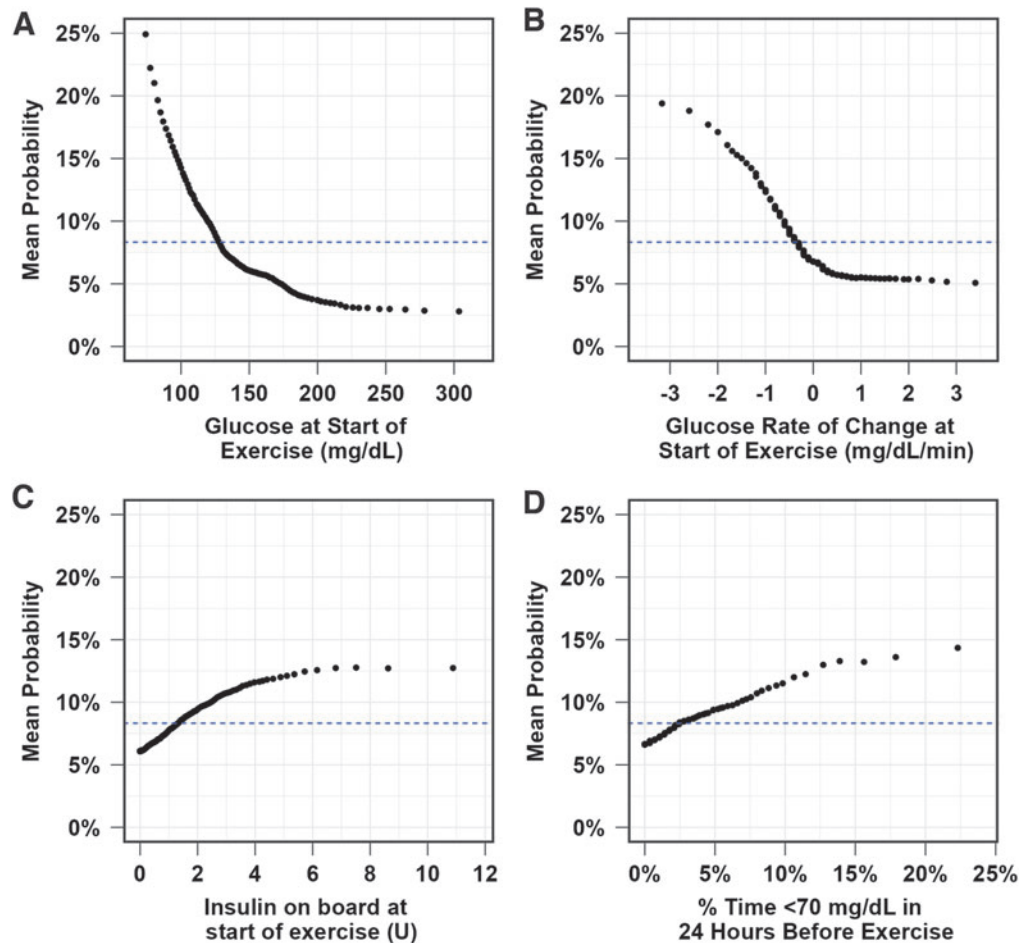
Experimental and commercially available HCL systems have been found to improve glycemic control in laboratory settings, such as minimizing hypoglycemia while maximizing time in glycemic target range,<sup>24</sup> and the results presented in this analysis can help generalize this ability as the exercise

sessions were free-living. However, in this analysis, the insulin modality had little effect on the probability of hypoglycemia during exercise, although the probability of developing hypoglycemia was indeed numerically less in those on MDI than for those on both types of CSII systems (Supplementary Fig. S5), which is in line with the recent survey data by Paiement et al.<sup>6</sup>

The reasons for the apparent differences in hypoglycemia risk rate during exercise between MDI (7.8%), HCL (8.2%), and pump (8.7%) may be related to behavioral differences on how individuals manage planned and unplanned exercise<sup>6</sup> and/or differences in IOB levels (median IOB at exercise start is 0.90 U for MDI, 1.51 U for HCL, and 1.12 U for pump). In any event, we found little evidence here that HCL systems were superior to MDI or CSII pump for the prevention of exercise-associated hypoglycemia when starting with the same glucose concentration, glucose rate of change, IOB, and prior time with glucose <70 mg/dL.

The type of exercise had a small effect on the probability of hypoglycemia with the study video interval exercises having the lowest mean probability of hypoglycemia (6.6%), while common physical activity including walking, hiking, chores, and physical labor had the highest mean probability of hypoglycemia at ~9.2%. It is possible that participants were more cognizant of hypoglycemia risk mitigation when performing the study video exercises, and they prepared for exercise differently and responded accordingly when hypoglycemia was imminent, while they may have overlooked the risk of hypoglycemia related to routine or spontaneous physical exertion. Indeed, the study videos recommended that pre-exercise glucose level be monitored before exercise start time.

One limitation with this analysis is that we did not collect information on if exercise was stopped prematurely because



**FIG. 3.** Probability of “in exercise” hypoglycemia (defined as one or more CGM values <70 mg/dL during exercise) based on CGM-based glucose level at the start of exercise (**A**), glucose rate of change at start of exercise (**B**), IOB at the start of exercise (**C**), and % time <70 mg/dL in 24 h before exercise (**D**). Black dots indicate the mean probability of hypoglycemia during exercise at each level of the variable. The dashed blue line represents the mean incidence of hypoglycemia during exercise. CGM, continuous glucose monitoring; IOB, insulin on board.

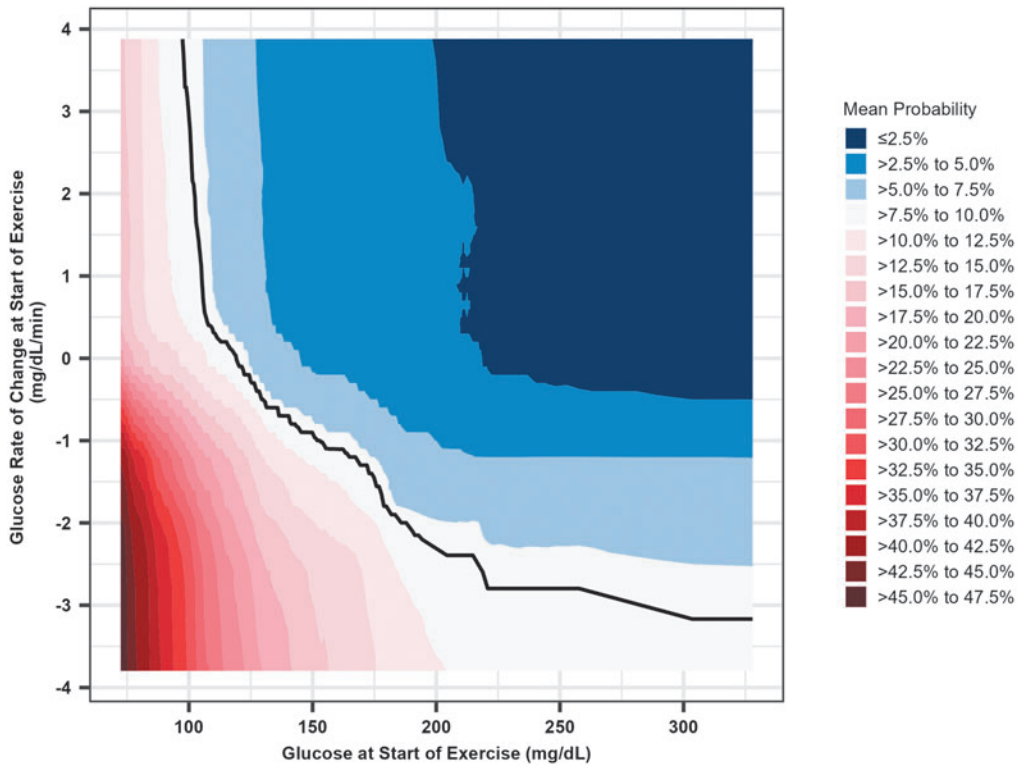
of an individual’s concerns about glucose levels during exercise. In other words, we do not know if participant volunteers stopped exercising because they achieved an a priori fitness or time goal or if they possibly stopped to treat an imminent hypoglycemic event. Thus, we purposefully did not use exercise duration as a predictor in our model because duration could be a cause of a hypoglycemic event or be impacted by the hypoglycemic event if adults stopped exercising to treat a low glucose.

We also did not include factors that could have been measured during exercise (e.g., mean heart rate during exercise) as the objective was to predict hypoglycemia based on information available at the start of exercise. Insulin sensitivity was unavailable for many participants, but incorporating both the IOB and insulin sensitivity factor may better predict hypoglycemia during exercise. And finally, each participant’s fitness level may also have been an important metric to consider a predictive feature. Prior work by Tyler et al.<sup>12</sup> has shown that people with higher levels of fitness tended to experience faster glucose drops and more hypoglycemia during aerobic exercise. In the current study, the probability of hypoglycemia increased from 7.5% to 9.0%

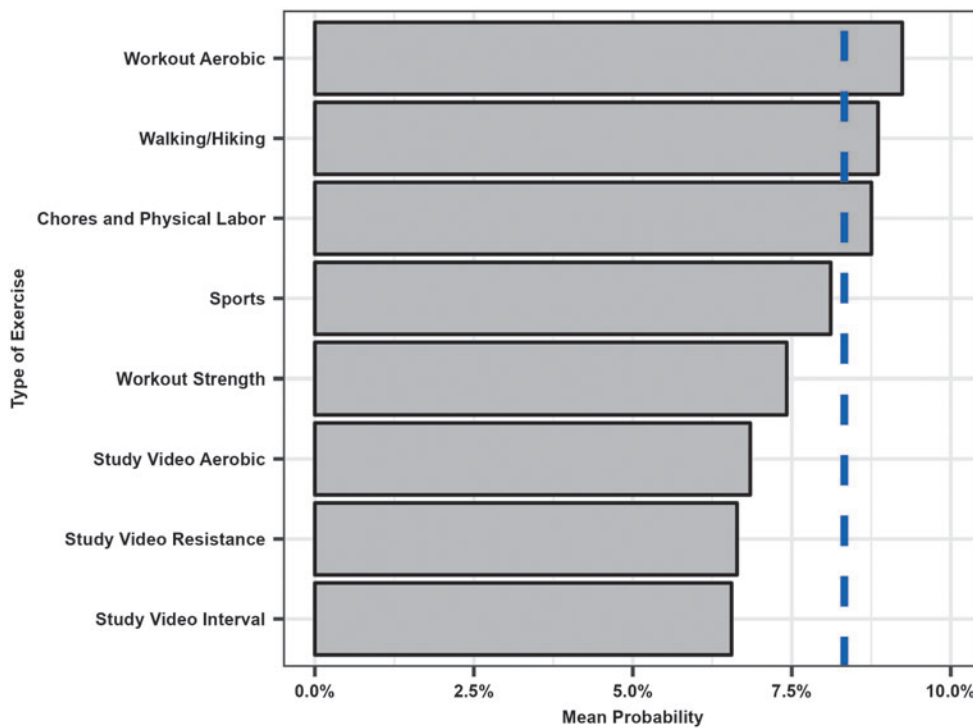
when the IPAQ score increased from 1500 to 3000 MET-minutes per week.

Midroni et al.<sup>25</sup> and Ben Brahim et al.<sup>26</sup> found that prior glucose values were the most important factor when predicting future glucose values. Ben Brahim et al.<sup>26</sup> and Rodríguez-Rodríguez et al.<sup>27</sup> also found that IOB was an important predictor, and Ben Brahim et al.<sup>26</sup> found that greater IOB values predicted decreases in future glucose values. Reddy et al.<sup>16</sup> found that a glucose <182 mg/dL at the start of exercise was highly predictive of hypoglycemia risk. They also found that higher heart rates within the first 10 min after the start of exercise increased the likelihood of hypoglycemia; however, RMRF did not consider heart rate readings during exercise as the point was to predict hypoglycemia before starting exercise and the heart rate at the start of exercise was not selected in the final model.

The methods used in these previous studies did not account for correlated data, and the studies specifically examining glucose dynamics during exercise were limited to only one exercise classification (i.e., treadmill walking or jogging) performed in a highly controlled clinic setting. The RMRF model handles nonlinear relationships and interactions and handles



**FIG. 4.** Heat map probability of hypoglycemia risk during exercise based on glucose levels and glucose rate of change at the start of exercise. Shaded region represents mean probability of hypoglycemia based on glucose level and glucose rate of change at start of exercise. The blue regions in the upper right indicate below average risk of hypoglycemia, and the red regions in the bottom left denote above average risk of hypoglycemia. The white region between the blue and red regions indicates average hypoglycemia risk. The solid black line indicates the average risk of hypoglycemia.



**FIG. 5.** Probability of hypoglycemia during exercise based on type of exercise. The gray bars represent the mean probability of hypoglycemia during exercise for each type of exercise. The dashed blue line represents the mean incidence of hypoglycemia during exercise.



correlated outcomes within a participant, which can better identify predictors' effect on hypoglycemia during exercise.

A major strength of the study is the large database of real-world data with 8827 exercise sessions analyzed from 459 participants. However, a limitation of the analysis is that exercise sessions with a duration between 20 and 90 min were analyzed, so predicting hypoglycemia for exercises longer than 90 min may not be reliable. The analyses only included exercises with at least 15 min of CGM data, but CGM use was high during exercise and similar results would be seen with a stricter CGM data requirement (e.g., as a certain percentage of exercise duration). In addition, this cohort may not represent the general T1D population as many of the participants exercised regularly, maintained a low HbA1c level, and used CGM technology, so the results may not generalize to individuals who do not exercise regularly or exhibit higher glycemic levels.

However, the exercise-level predictors were assessed within participant, and we believe their effect on hypoglycemia would likely be similar for adults with T1D who do not regularly exercise or have higher glycemic levels. Another limitation of our analysis approach is that multicollinearity can affect the relationship between factors and risk of hypoglycemia during exercise, but having correlated variables is less of an issue with random forests and often has little effect on predictive performance. Partial least-squares regression could be a good alternative over RMLR if the goal was to interpret some of the correlated factors in the regression context. Finally, the computational burden of building the RMRF model limited the parameter optimization and hindered formally testing if the predictors were significantly associated with increased hypoglycemia. However, once the RMRF model was built, predictions could be returned immediately for a new exercise session.

In summary, accurate predictions of hypoglycemia during exercise are possible with either an RMRF or an RMLR model. Moreover, the study identified factors that can be modified to reduce the risk of exercise-associated hypoglycemia. In our model, lower glucose values, a steeper negative glucose rate of change, greater IOB, and greater prior time <70 mg/dL at start of exercise were associated with increased risk of hypoglycemia during exercise. Notably, these variables would all be available to HCL systems, suggesting that algorithms incorporating these variables can make it possible to help people with T1D exercise safely.

### Authors' Contributions

S.B. and P.C. wrote/edited the article. S.B. performed statistical analysis. M.C.R., P.G.J., Z.L., R.L.G., M.A.C., F.J.D., C.K.M., S.R.P., J.R.C., M.B.G., R.W.B., and M.R.R. reviewed, edited, and contributed to discussion.

### Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of The Leona M. and Harry B. Helmsley Charitable Trust.

### Author Disclosure Statement

S.B. reports no conflicts of interest: M.C.R. reports receiving consulting fees from the JAEB Center for Health Research, Eli Lilly, Zealand Pharma, and Zucara Therapeutics; speaker

fees from Sanofi Diabetes, Eli Lilly, Dexcom Canada, and Novo Nordisk; and stock options from Supersapiens and Zucara Therapeutics. P.G.J. reports receiving grants from the National Institutes of Health, The Leona M. and Harry B. Charitable Trust, JDRF, Dexcom, and the Oregon Health and Science University Foundation; consultancy fees from CDISC; U.S. patents 62/352,939, 63/269,094, 62/944,287, 8810388, 9,480,418, 8,317,700, 61/570382, 8,810,388, 7,976,466, and 6,558,321; and reports stock options from Pacific Diabetes Technologies, outside submitted work. Z.L. reports no conflict of interests. R.L.G. reports no conflict of interest. M.A.C. is Chief Medical Officer of Glooko, Inc., and has received grants or contracts from Dexcom, Abbott Diabetes Care, National Institutes of Health, JDRF, the Emily Rosebud Foundation, Eli Lilly, Tolerion, and Garmin.

F.J.D. reports no conflict of interests. C.K.M. reports no conflict of interests. S.R.P. reports receiving grants from The Leona M. and Harry B. Helmsley Charitable Trust, the National Institutes of Health, and the JAEB Center for Health Research; and honorarium from the American Diabetes Association, outside the submitted work. J.R.C. reports receiving grants from JDRF, the National Institutes of Health, Dexcom, and Medtronic; consultancy fees from Novo Nordisk, Insulet, and Zealand, outside the submitted work. M.B.G. reports no conflict of interest. R.W.B. reports receiving consulting fees, paid to his institution, from Insulet, Bigfoot Biomedical, vTv Therapeutics, and Eli Lilly, grant support and supplies, provided to his institution, from Tandem and Dexcom, and supplies from Ascensia and Roche. M.R.R. reports consultancy fees from Zealand Pharma. P.C. reports no conflict of interests.

### Funding Information

Research reported in this publication was supported by The Leona M. and Harry B. Helmsley Charitable Trust. One of the author's institutions (C.K.M., Pennington Biomedical Research Center) is supported by the NORC Center grant P30 DK072476 entitled "Nutrition and Metabolic Health Through the Lifespan" sponsored by NIDDK and by grant U54 GM104940 from the National Institute of General Medical Sciences, which funds the Louisiana Clinical and Translational Science Center. Verily (South San Francisco, CA) provided the Study Watch at no cost. Dexcom provided continuous glucose monitors at a discounted rate.

### Supplementary Material

Supplementary Data S1  
Supplementary Figure S1  
Supplementary Figure S2  
Supplementary Figure S3  
Supplementary Figure S4  
Supplementary Figure S5  
Supplementary Table S1  
Supplementary Table S2  
Supplementary Table S3  
Supplementary Table S4

### References

1. Riddell MC, Peters AL. Exercise in adults with type 1 diabetes mellitus. *Nat Rev Endocrinol* 2023;19(2):98–111.

2. Jabbour G, Henderson M, Mathieu ME. Barriers to active lifestyles in children with type 1 diabetes. *Can J Diabetes* 2016;40(2):170–172.
3. Brazeau AS, Rabasa-Lhoret R, Strychar I, et al. Barriers to physical activity among patients with type 1 diabetes. *Diabetes Care* 2008;31(11):2108–2109.
4. Keshawariz A, Pirovato AR, Brown TL, et al. Lower objectively measured physical activity is linked with perceived risk of hypoglycemia in type 1 diabetes. *J Diabetes Complications* 2018;32(11):975–981.
5. Lascar N, Kennedy A, Hancock B, et al. Attitudes and barriers to exercise in adults with type 1 diabetes (t1dm) and how best to address them: A qualitative study. *PLoS One* 2014;9(9):e108019.
6. Paiement K, Frenette V, Wu Z, et al. Is better understanding of management strategies for adults with type 1 diabetes associated with a lower risk of developing hypoglycemia during and after physical activity? *Can J Diabetes* 2022;46(5):526–534.
7. Breton MD, Brown SA, Karvetski CH, et al. Adding heart rate signal to a control-to-range artificial pancreas system improves the protection against hypoglycemia during exercise in type 1 diabetes. *Diabetes Technol Ther* 2014;16(8):506–511.
8. Dasanayake IS, Bevier WC, Castorino K, et al. Early detection of physical activity for people with type 1 diabetes mellitus. *J Diabetes Sci Technol* 2015;9(6):1236–1245.
9. Jacobs PG, Resalat N, El Youssef J, et al. Incorporating an exercise detection, grading, and hormone dosing algorithm into the artificial pancreas using accelerometry and heart rate. *J Diabetes Sci Technol* 2015;9(6):1175–1184.
10. Turksoy K, Paulino TM, Zaharieva DP, et al. Classification of physical activity: Information to artificial pancreas control systems in real time. *J Diabetes Sci Technol* 2015;9(6):1200–1207.
11. Turksoy K, Quinn LT, Littlejohn E, Cinar A. An integrated multivariable artificial pancreas control system. *J Diabetes Sci Technol* 2014;8(3):498–507.
12. Tyler NS, Mosquera-Lopez C, Young GM, et al. Quantifying the impact of physical activity on future glucose trends using machine learning. *iScience* 2022;25(3):1–19.
13. Castle JR, El Youssef J, Wilson LM, et al. Randomized outpatient trial of single- and dual-hormone closed-loop systems that adapt to exercise using wearable sensors. *Diabetes Care* 2018;41(7):1471–1477.
14. Wilson LM, Jacobs PG, Ramsey KL, et al. Dual-hormone closed-loop system using a liquid stable glucagon formulation versus insulin-only closed-loop system compared with a predictive low glucose suspend system: An open-label, outpatient, single-center, crossover, randomized controlled trial. *Diabetes Care* 2020;43(11):2721–2729.
15. Turksoy K, Bayrak ES, Quinn L, et al. Hypoglycemia early alarm systems based on multivariable models. *Ind Eng Chem Res* 2013;52(35):12329–12336.
16. Reddy R, Resalat N, Wilson LM, et al. Prediction of hypoglycemia during aerobic exercise in adults with type 1 diabetes. *J Diabetes Sci Technol* 2019;13(5):919–927.
17. Riddell MC, Li Z, Gal RL, et al. Examining the acute glycemic effects of different types of structured exercise sessions in type 1 diabetes in a real-world setting: The type 1 diabetes and exercise initiative (t1dexi). *Diabetes Care* 2023;46(1):704–713.
18. Gillingham MB, Li Z, Beck RW, et al. Assessing mealtime macronutrient content: Patient perceptions versus expert analyses via a novel phone app. *Diabetes Technol Ther* 2021;23(2):85–94.
19. Moser O, Riddell MC, Eckstein ML, et al. Glucose management for exercise using continuous glucose monitoring (CGM) and intermittently scanned CGM (ISCGM) systems in type 1 diabetes: Position statement of the European Association for the Study of Diabetes (EASD) and of the International Society for Pediatric and Adolescent Diabetes (ISPAD) endorsed by JDRF and supported by the American Diabetes Association (ADA). *Diabetologia* 2020;63(12):2501–2520.
20. Calhoun P, Levine RA, Fan J. Repeated measures random forests (rmrf): Identifying factors associated with nocturnal hypoglycemia. *Biometrics* 2021;77(1):343–351.
21. Breiman L. Random forest. *Mach Learn* 2001;45:5–32.
22. Riddell MC, Gallen IW, Smart CE, et al. Exercise management in type 1 diabetes: A consensus statement. *Lancet Diabetes Endocrinol* 2017;5(5):377–390.
23. Hobbs N, Samadi S, Rashid M, et al. A physical activity-intensity driven glycemic model for type 1 diabetes. *Comput Methods Programs Biomed* 2022;226:107153.
24. Eckstein ML, Weigluni B, Tauschmann M, et al. Time in range for closed-loop systems versus standard of care during physical exercise in people with type 1 diabetes: A systematic review and meta-analysis. *J Clin Med* 2021;10(11):2445.
25. Midroni C, Leimbigger P, Baruah G. Predicting glycemia in type 1 diabetes patients: Experiments with xgboost. Proceedings of the 3rd International Workshop on Knowledge Discovery in Healthcare Data. 2018. <https://ceur-ws.org/Vol-2148/paper13.pdf> [Last accessed June 14, 2023].
26. Ben Brahim N, Place J, Renard E, et al. Identification of main factors explaining glucose dynamics during and immediately after moderate exercise in patients with type 1 diabetes. *J Diabetes Sci Technol* 2015;9(6):1185–1191.
27. Rodríguez-Rodríguez I, Chatzigiannakis I, Rodríguez JV, et al. Utility of big data in predicting short-term blood glucose levels in type 1 diabetes mellitus through machine learning techniques. *Sensors (Basel)* 2019;19(20):4482.

Address correspondence to:

*Peter Calhoun, PhD*

*JAEB Center for Health Research*

*15310 Amberly Drive, Suite 350*

*Tampa, FL 33647*

*USA*

*E-mail: pcalhoun@jaeb.org*