Glycaemic events during exercise can be effectively predicted with machine learning using only start glucose and duration

Catherine L Russon[1, 5], Neil Vaughan[1, 5, 3]*, Richard M Pulsford[4], Robert C Andrews[1, 2, 5], Michael Allen[5]

Background and Aims

When people with Type 1 diabetes (T1D) exercise, we suggest they stay in the range of 7-15 mmol/L and take action if glucose moves outside of this range. In a study of 39 people with T1D who took part in a half marathon, 70% had to stop to deal with a low or high glucose event. If people with T1D were warned pre-exercise of the risk of a glycaemic event, it would allow them to take countermeasures to prevent these events occurring. Using standard demographic data, blood tests and information about the exercise bout, we aimed to examine whether machine learning (ML) could predict low or high glucose events during exercise.

Material and methods

Data came from 2 exercise studies (EXTOD education and EXTOD 101) with each having data on age, sex, length of diabetes, body mass index, HbA1c, C-peptide and time, length, intensity (BORG scale) and type (aerobic, anaerobic and mixed) of exercise sessions. EXTOD education had 2 weeks of Dexcom G6 data from 54 participants and EXTOD 101 8 weeks of Freestyle Libre data from 34 participants. In total there were 976 bouts of exercise.

The thresholds for low glucose and high glucose were set to 7 and 15mmol/L – the point at which intervention is advised. A single reading below or above the threshold was considered as a positive event. 486 bouts contained a glucose reading below 7 and 151 had glucose reading above 15 mmol/L.

80% of this data was used to train an ML algorithm (XGBoost) which then determined how important each measure was and what was the best combination. This was then tested on the remaining 20% of data using area under the receiver operator curve (ROC AUC) score as the validation metric.

Results

For predicting a glucose reading below 7mmol/L, the addition of the best 6 measures identified by ML resulted in a ROC AUC score 0.902. These, in order of importance, were glucose at start of exercise (starting glucose), duration of exercise, type of exercise, intensity, time of day and C-peptide. For predicting glucose above 15 mmol/L, incorporating the best 7 measures resulted in a ROC score of 0.973. These were start glucose, duration, intensity, sex, years since diagnosis, C-peptide, and time of day.

The two most important measures were start glucose and duration of exercise bout in both contexts. Using only these features, a ROC AUC score of 0.890 was achieved for predicting low glucose and 0.968 for high glucose. The model accuracy and sensitivity-specificity intersect was 0.804 and 0.804 for low glucose and 0.949 and 0.900 for high.

Using logistic analysis, we have taken these two measures to develop a heat map that can help patients predict their risk of going below 7 or above 15mmol/L during an exercise (see figure 1).

Conclusions

ML has produced a simple heat map to predict risk of glucose going below 7 or above 15mmol/L during exercise. We will go on to look at whether it can predict lower glucose levels during exercise and at times further away from the start of exercise.