Forecasting postbariatric hypoglycaemia in Roux-en-Y gastric bypass patients using model-based algorithms fed by continuous glucose monitoring data: a proof-of-concept study.

Authors:

Francesco Prendin[1], Giacomo Cappon[1], Afroditi Tripyla[2], David Herzig[2], Lia Bally[2*] and Andrea Facchinetti[1*]

[1]Department of Information Engineering (DEI), University of Padova, Padova, Italy

[2]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

*These authors equally contributed

Corresponding author:

Prof. Andrea Facchinetti

Department of Information Engineering (DEI), University of Padova, Via G. Gradenigo 6/B, 35131 Padova, Italy, Tel: +39-049-8277669

E-Mail: andrea.facchinetti@unipd.it

Short running title:

Forecasting severe hypoglycaemia in postbariatric population using only continuous glucose monitoring data

Accepted Article

1. BACKGROUND

Postbariatric hypoglycaemia (PBH) is an increasingly recognized late complication of bariatric surgery, particularly Roux-en-Y gastric bypass (RYGB)[1,2]. Between 20-80% of RYGB patients may develop the condition, which is characterized by postprandial hypoglycaemic episodes with increased severity after ingestion of carbohydrates with a high glycaemic impact[3]. Prevalence estimates range widely, owing to the current lack of standardised diagnostic criteria. Although incompletely understood, accelerated nutrient absorption alongside with excessive postprandial incretin and insulin exposure are key pathophysiological features[4].

In the absence of an approved pharmacotherapy, dietary management is the first-line treatment of PBH[5]. Dietary measures, however, can be very restrictive, insufficiently effective and challenging to implement in the long-term. Given these limitations, continuous glucose monitoring (CGM) devices, which provide real-time (RT) information on current glucose levels and rate of change, have the potential to support PBH management[6]. CGM can be leveraged to develop RT predictive algorithms allowing for preventive or timely corrective actions (e.g., carbohydrate intake), which may be particularly useful for the frequently encountered PBH patients with hypoglycaemia unawareness and related safety concerns[7,8].

Whilst hypoglycaemia forecasting has been widely studied in type 1 diabetes (T1D)[9], the topic remains understudied in the PBH population. The first and only contribution in this field was the development of a heuristic-based predictive algorithm for a glucose-responsive glucagon delivery system in an experimental inpatient session[10,11].

To address this gap, the purpose of this work was to assess the feasibility of forecasting PBH episodes by exploiting three different predictive algorithms using only CGM data.

2. METHODS

2.1 Dataset and PBH event definition

Data were generated by 39 adults with confirmed PBH after RYGB (defined as symptomatic plasma or sensor glucose (SG) < 54 mg/dL relieved by glucose administration) wearing the Dexcom G6 (Dexcom Inc.,

San Diego, CA, USA) CGM sensor for a median of 10 days (IQR 9-30) in daily life conditions. Data were obtained from usual care and research settings (NCT04330196, NCT04334161, NCT04332289). Overall, the percentage of days using CGM in blinded, unblinded, unknown mode is: 14.5%, 62.7%, and 22.8%, respectively. Based on a PBH event definition of sensor glucose < 54 mg/dL for at least 15 min[7,12], we identified a total of 542 PBH episodes (≈4 every 10 days per subject) with an average duration of 25 min. Participants' details are summarized in Table S1 (see Supplementary Material, Appendix A). Following preprocessing for anomalies and noise (for more details see Supplementary Material, Appendix A), the dataset was split into a training (31 subjects and 489 PBH events) and a test (8 remaining subjects and 53 PBH events) set. In addition, given that the CGM monitoring length may significantly vary between individuals, to create a test set that is as balanced as possible and to avoid any bias on the results, we applied the following criteria to include/exclude a subject in/from the test set: i) the patient has more than 5 and less than 20 consecutive monitoring days and ii) the patient showed at least 1 PBH episodes over 4 monitoring days.

Clinical and demographic information about the training-test partition are detailed in Table S1 (Supplementary Material, Appendix A).

2.2 Predictive algorithms

Based on our previous work on the prediction of hypoglycaemia in T1D[13], we considered the following three algorithms: an Autoregressive model with recursive parameter estimation (AR1)[14], which represents a good example of consolidated adaptive method; an Autoregressive Integrated Moving Average (ARIMA) model[13], which turned out to be the best linear predictor in T1D; and a feed forward Neural Network (NN)[15], as representative of nonlinear methodologies. These methods, besides being considered as state-of-art glucose predictive algorithms for T1D, were also shown to be the best performing for short-term prediction when CGM data is the only available source of information[13].

Regarding the prediction horizon (PH, i.e., how far ahead the method predicts the event), we considered 15, 20, 25 and 30 min.

For each combination of algorithm and PH, model parameters and/or hyperparameters were estimated in the training set. Then, the algorithms were applied to the test set, simulating the acquisition of CGM data in RT (see Supplementary Material, Appendix B). Two examples of RT PBH forecasting using the proposed algorithms with PH=20 min are visualized in Figure 1.
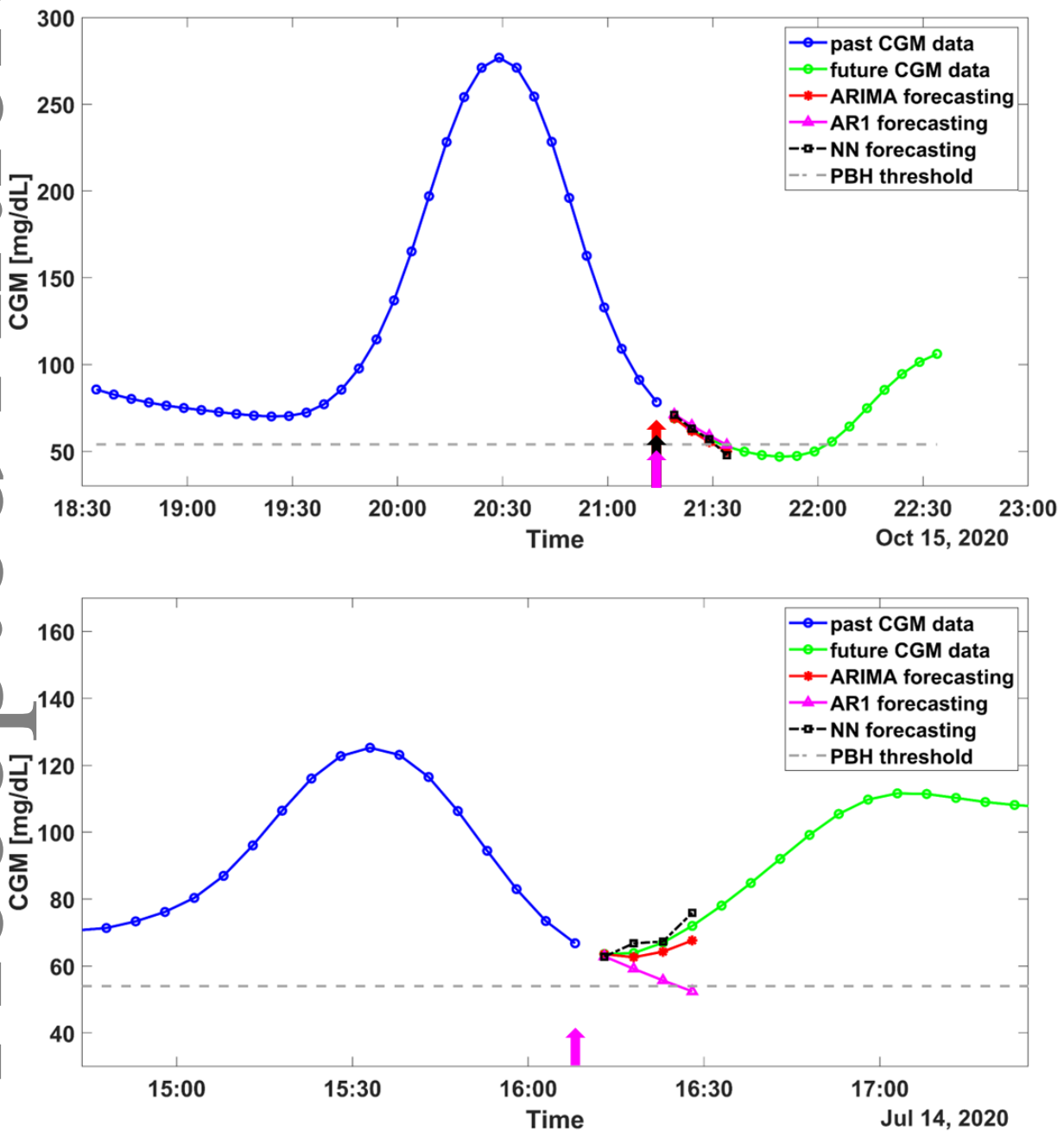

Figure 1. Examples of real-time forecasting of postbariatric hypoglycaemia (PBH) events and preventive alert generation using the 3 model-based algorithms fed by the past continuous glucose monitoring (CGM) values (blue dotted line, sampling time 5-min). Green circles indicate future CGM samples. **Top panel**: at time 21:15 (actual time), the three algorithms fed by the past CGM values,

*predict the next four CGM values (ARIMA: red asterisks; AR1: magenta triangles; NN: black squares). Since the last predicted CGM values is below th_{PBH} and there are no recent alarms, a preventive PBH alarm (red, black and magenta arrow for ARIMA, NN and AR1, respectively) is triggered. **Bottom panel:** at time 16:07 (actual time), AR1 predicts a value below th_{PBH} and raises a false alarm (magenta arrow), whereas ARIMA and NN correctly predict the increase in glucose concentration and do not generate any alert.*

Algorithm performance was evaluated as the ability to predict/detect PBH events[a] (see Supplementary Appendix B for details). For each raised PBH alarm, we counted: a true positive (TP) if a PBH event occurred in the following 45 min; a false positive (FP) if no PBH events occurred in the following 45 min. A false negative (FN) was counted when no alarms were generated despite the occurrence of a PBH event [13]. Based on TP, FP and FN, the following aggregated metrics were calculated: precision (P), recall (R), F1-score (F1).

P can be seen as the percentage of the correct alarms over the total number of raised alarms. Recall, also known as sensitivity or true positive ratio, is the ratio of correctly predicted hypoglycaemic events over the total number of events. F1-score is the harmonic mean of the two previous metrics. Additionally, we evaluated the daily number of false alarms (FP/day) raised by the algorithms, and the time gain (TG) defined as the temporal distance between a TP alarm and the corresponding PBH event onset, thus representing the time window for a preventive intervention.

Due the short CGM recording period of the test set (median 10 days) and consequently low prevalence of PBH events, the value of hypoglycaemic prediction metrics was obtained by considering all hypoglycaemic events of different subjects according to a population-based approach. The results are expressed as a single value for all the considered metrics except for TG, that is expressed as median [25th – 75th], since it can be computed for each TP.

To contrast our results with the previously published work, we reimplemented and trained the PBH Detection System (PBH-DS) algorithm developed by Laguna Sanz et al[10], in particular we referred to the version denoted as PBH-DS v002 in [10]. All the implementations were done in MATLAB (2021a version).

3. RESULTS

---

[a] This work does not consider Dexcom Urgent Low Soon (ULS) alerts algorithm and it does not provide any evaluation of its performance.

Performance metrics of the AR1, ARIMA and NN algorithms for each considered PH as well as of the previously published PBH-DS are shown in Table 1 (for details on parameter identification see Supplementary Material, Appendix C). The ARIMA configuration with PH=20 min performed best, achieving P=79.10%, R=100%, F1=88.33%, FP/day=0.17 and median TG=20 min. In practical terms, provided that CGM reflects blood glucose precisely and accurately, this means that PBH episodes can be effectively predicted 20 min before happening, with no missed events and generating only one false alert every 6 days.

ARIMA predictors with PH=25 and 30 min, despite achieving a larger TG (i.e. window for intervention), resulted in inferior P, R and FP/day. This is illustrated by the F1-score trend, which decreased as the PH horizon increased (F1=83.21%, 70.20%, 58.76%) due to the critical decrease of P (P=72.15%, 54.08%, 41.94%).

Compared to ARIMA, AR1 was inferior for all PHs, particularly in terms of P (P=36.11%, 35.97%, 42.24%, and 44.45% for PH=15, 20, 25, and 30 minutes, respectively). Additionally, AR1 provides the largest FP/day=1.15, in line with its known susceptibility for unstable predictions[14]. The NN configuration performed similarly to ARIMA but yielded lower TG.

Re-implementation of the previously published PBH-DS resulted in R=100% with median TG of 25 min and P=23.87%. As consequence, the FP/day was 2.11, which is 10 times the number of FP raised by ARIMA for PH=20 min (FP/day=0.17).

Table 1 Postbariatric hypoglycaemia (PBH) prediction metrics for the algorithms under investigation (ARIMA, AR1, NN and PBH-DS) according to different prediction horizons (PH) based on a test set containing 53 PBH events. Results of TG are reported as median [25th-75th] percentile.

Algorithm	PH (min)	Metrics				
		P (%)	R (%)	F1 (%)	FP/day	TG (min)
ARIMA		72.15	98.28	83.21	0.27	15 [15-15]
AR1	15	36.11	98.11	52.79	1.15	10 [10-15]
NN		68.29	96.55	80	0.32	15 [10-15]

		P	R	F1	FP/day	TG
ARIMA		79.10	100	88.33	0.17	20 [15-20]
AR1	20	35.97	94.34	52.08	1.11	10 [5-10]
NN		82.26	96.23	88.70	0.14	15 [15-20]
ARIMA		54.08	100	70.20	0.56	25 [20-25]
AR1	25	42.24	92.45	57.99	0.84	10 [5-10]
NN		62.32	81.13	70.49	0.32	20 [15-25]
ARIMA		41.94	98.11	58.76	0.89	25 [20-30]
AR1	30	44.45	90.57	59.63	0.76	10 [5-10]
NN		54.67	77.36	64.06	0.43	25 [20-30]
PBH-DS	-	23.87	100	38.55	2.11	25 [20-30]

Abbreviations: P, precision; R, recall; F1, F1-score; FP/day, false positives per day; TG, time gain; ARIMA, Autoregressive Integrated Moving Average; AR1, Autoregressive model; NN, Neural Network; PBH-DS, Postbariatric Hypoglycaemia Detection System.

In addition, we analysed the performance of the predictive algorithms for individuals wearing the CGM sensor in blinded and unblinded mode (full results are reported in the supplementary Material, Appendix D). In particular, the test set comprises 4 patients with blinded recordings (for a total of 24 hypoglycaemic episodes) and 4 patients with unblinded recordings (for a total of 29 hypoglycaemic episodes). The results are consistent with those reported for the complete dataset (Table 1): i) the best performing algorithm is confirmed to be ARIMA with a PH = 20 min, granting high precision (77.42% and 80.56% for blinded and unblinded subsets, respectively), high recall (100% in both cases) and low FP/day (0.17 vs 0.18 for blinded and unblinded sets, respectively); ii) AR1 is inferior to ARIMA for all the PHs for both blinded and unblinded subsets; and iii) NN performed similarly to ARIMA but it yields to a slightly inferior median TG (15 minutes) both in blinded and unblinded subsets. Of note, the number of FP/day is slightly larger in the unblinded than in the blinded subset.

4. CONCLUSIONS

In this proof-of-concept study, we assessed the feasibility to forecast PBH events in real-time using various linear and non-linear black-box predictive algorithms fed by CGM data only. The highest performance was

achieved with ARIMA approach using a prediction horizon of 20 min, which was able to effectively predict PBH events with a median lead time of 20 min, with no missed events and only one false alert every 6 days. The ARIMA approach outperformed the previously published hypoglycaemia prediction algorithm, which yielded two false alarms per day when applied on our data[10,11]. Apart from usability aspects, avoidance of false alarms is particularly important for the PBH population as unnecessary corrective ingestion carbohydrates can cause rebound hypoglycaemia and predispose to weight regain[1]. Although comparability is limited, the herein achieved performance metrics for hypoglycaemia prediction can even compete with those reached in T1D and T2D populations using similar methods[13]. Thus, our findings are encouraging and support the feasibility to forecast PBH episodes by leveraging CGM data in combination with an ARIMA-based. Of note, compared to models used in [13], the proposed ARIMA model shows an inferior number of parameters to describe the glucose dynamics (i.e., the autoregressive model order). Thus, suggesting that glucose dynamics are faster in PBH than T1D population.

4.1 Limitations of the study

Despite the promising and encouraging results obtained in this paper we acknowledge various limitations. First, model assessment on original non-processed CGM data would have been interesting, but not solid because of noise, which could have negatively impacted the identification and training procedure of the algorithms. However, it is worth noting that, despite the offline data preprocessing aimed at removing noise/anomalies that could have introduced a bias in the evaluation, all predictive algorithms have been applied simulating a real-time application.

Another potential bias in the analysis may be the presence of CGM recordings acquired in unblinded modality. In fact, we found that the number of FPs is higher in the unblinded subset. This may be explained as follows: low glucose alarms generated by the unblinded CGM sensor and/or the possibility of reading in real-time CGM values may have triggered a preventive carbohydrates intake, thereby mitigating against the impending PBH episode. Unfortunately, lack of information on either alert settings or preventive carbohydrates intakes in the dataset precludes a definitive confirmation. Still, it is important to note that

the inclusion of unblinded CGM recordings could have generated an underestimation of the performance of our algorithms, as a results of false positives.

Finally, we acknowledge that the previously a time-lag of about 10-15 minutes[16] between interstitial and blood glucose concentrations, could reduce the actual time-anticipation of PBH events via CGM sensor data to 5-10 minutes. As intravascular sensors are currently not a viable option, further studies are required to provide an estimation of the blood-to-interstitial fluid time lag during the rapid dynamics of a PBH episode and thus assess the true effectiveness of PBH CGM-based predictive algorithms.

4.2 Future developments

Future work will focus on the development of subject-specific algorithms, which allows considering the large heterogeneity that characterizes the PBH population. This will only be possible once large CGM longitudinal datasets are available. A further and natural extension of this work will be the assessment of the improvement in forecasting PBH events by increasing the input data by additional information such as meal and physical activity. A more thorough understanding of the correlation/causation between the PBH episodes and adverse clinical events will further help to determine practical clinical impact of PBH real-time prediction.

In conclusion, CGM data can be leveraged to forecast PBH and future research and clinical validation trials will unravel whether the technology translates into patient benefits.

5. ACKNOWLEDGEMENTS

F.P, G.C., D.H., A.F. and L.B. designed the analysis, A.T. recruited participants and collected the CGM data, A.T. and F.P. reviewed and prepared the data for analysis, F.P., G.C. and A.F. performed the analysis, F.P., G.C., A.T., D.H., A.F. and L.B. interpreted the data. F.P. and G.C. wrote the first draft of the manuscript. A.T., D.H., A.F. and L.B. critically reviewed the manuscript. L.B. and A.F. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final draft of the manuscript for submission.

Conflict of Interest Statement

Authors declare no conflict of interest.

References

1. Salehi M, Vella A, McLaughlin T, Patti ME. Hypoglycemia after gastric bypass surgery: Current concepts and controversies. *Journal of Clinical Endocrinology and Metabolism*. 2018;103(8):2815-2826. doi:10.1210/jc.2018-00528

2. Kefurt R, Langer FB, Schindler K, Shakeri-Leidenmühler S, Ludvik B, Prager G. Hypoglycemia after Roux-En-Y gastric bypass: detection rates of continuous glucose monitoring (CGM) versus mixed meal test. *Surgery for Obesity and Related Diseases*. 2015;11(3):564-569. doi:https://doi.org/10.1016/j.soard.2014.11.003

3. Capristo E, Panunzi S, de Gaetano A, et al. Incidence of Hypoglycemia after Gastric Bypass vs Sleeve Gastrectomy: A Randomized Trial. *Journal of Clinical Endocrinology and Metabolism*. 2018;103(6):2136-2146. doi:10.1210/jc.2017-01695

4. Goldfine AB, Mun EC, Devine E, et al. Patients with neuroglycopenia after gastric bypass surgery have exaggerated incretin and insulin secretory responses to a mixed meal. *Journal of Clinical Endocrinology and Metabolism*. 2007;92(12):4678-4685. doi:10.1210/jc.2007-0918

5. Lee D, Dreyfuss JM, Sheehan A, Puleio A, Mulla CM, Patti ME. Glycemic Patterns Are Distinct in Post-Bariatric Hypoglycemia after Gastric Bypass (PBH-RYGB). *Journal of Clinical Endocrinology and Metabolism*. 2021;106(8):2291-2303. doi:10.1210/clinem/dgab323

6. Kubota T, Shoda K, Ushigome E, et al. Utility of continuous glucose monitoring following gastrectomy. *Gastric Cancer*. 2020;23(4):699-706. doi:10.1007/s10120-019-01036-5

7. Craig CM, McLaughlin TL. Defining clinically important hypoglycemia in patients with postbariatric hypoglycemia. *Surgery for Obesity and Related Diseases*. 2021;17(11):1865-1872. doi:10.1016/j.soard.2021.06.013

8. Lehmann V, Tripyla A, Herzig D, et al. The impact of postbariatric hypoglycaemia on driving performance: A randomized, single-blind, two-period, crossover study in a driving simulator. *Diabetes, Obesity and Metabolism*. 2021;23(9):2189-2193. doi:10.1111/dom.14456

9. Oviedo S, Vehí J, Calm R, Armengol J. A review of personalized blood glucose prediction strategies for T1DM patients. *International Journal for Numerical Methods in Biomedical Engineering*. 2017;33(6):1-21. doi:10.1002/cnm.2833

10. Laguna Sanz AJ, Mulla CM, Fowler KM, et al. Design and Clinical Evaluation of a Novel Low-Glucose Prediction Algorithm with Mini-Dose Stable Glucagon Delivery in Post-Bariatric Hypoglycemia. *Diabetes Technology and Therapeutics*. 2018;20(2):127-139. doi:10.1089/dia.2017.0298

11. Mulla CM, Zavitsanou S, Sanz AJL, et al. A randomized, placebo-controlled double-blind trial of a closed-loop glucagon system for postbariatric hypoglycemia. *Journal of Clinical Endocrinology and Metabolism*. 2020;105(4):E1260-E1271. doi:10.1210/clinem/dgz197

12. Heller SR. Glucose concentrations of less than 3.0 mmol/L (54 mg/dL) should be reported in clinical trials: A joint position statement of the American diabetes association and the European association for the study of diabetes. *Diabetes Care*. 2017;40(1):155-157. doi:10.2337/dc16-2215

13. Prendin F, del Favero S, Vettoretti M, Sparacino G, Facchinetti A. Forecasting of Glucose Levels and Hypoglycemic Events: Head-to-Head Comparison of Linear and Nonlinear Data-Driven Algorithms Based on Continuous Glucose Monitoring Data Only. *Sensors*. 2021;21(5). doi:10.3390/s21051647

14. Sparacino G, Zanderigo F, Corazza S, Maran A, Facchinetti A, Cobelli C. Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series. *IEEE Transactions on biomedical engineering*. 2007;54(5):931-937.

15. Pérez-Gandìa C, Facchinetti A, Sparacino G, et al. Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring. *Diabetes Technol Ther*. 2010;12(1):81-88.

16. Guillot FH, Jacobs PG, Wilson LM, et al. Accuracy of the dexcom G6 glucose sensor during aerobic, resistance, and interval exercise in adults with type 1 diabetes. *Biosensors (Basel)*. 2020;10(10). doi:10.3390/BIOS10100138