FIRST ASSESSMENT OF THE PERFORMANCE OF A PERSONALIZED MACHINE LEARNING APPROACH TO PREDICTING BLOOD GLUCOSE LEVELS IN PATIENTS WITH TYPE 1 DIABETES: THE CDDIAB STUDY

Stéphane BIDET[1], Nicolas CALECA[1], Eric RENARD[2], Thibault CAMALON[1], Lucas DE LA BROSSE[1], Mickael REHN[1], Omar DIOURI[2] and Jérôme PLACE[2]

[1] hillo ai SAS, Department of data science, Palaiseau, France,
[2] Montpellier University Hospital, Department of Endocrinology, Diabetes, Nutrition and Institute of Functional Genomics, INSERM, CNRS, University of Montpellier, Montpellier, France

BACKGROUND AND AIMS

To maintain normal blood glucose level, and therefore prevent complications, it is critical for patients with T1DM to self-administer insulin according to their fluctuating needs. But assessing the most appropriate dosing is highly complex since there are many parameters governing blood glucose levels, and their impact varies widely. Patients must make those decisions on-the-spot, using data at their disposal, personal knowledge and experience, and what they lack is a way to anticipate.

Therefore, in order to improve glycaemic control, there is a need for a system capable of producing reliable and personalised prediction, as the gathering and accurate interpretation of the data is the most relevant aspect of the patient's decision-making process.

We conducted an observatory study at the Montpellier University Hospital in France. The CDDIAB study's main objective was to evaluate a new machine learning approach to predicting BG levels of each individual from a collection of personal BG measurements with contextual data.

MATERIALS AND METHODS

The sample included 14 patients with T1DM, who volunteered to track BG measurements, meal intakes and insulin doses in real life conditions. The study ran over 30 days, and no specific intervention on the usual diabetes treatment was undertaken during this period.

Sample Characteristics	
Group Size	n = 14
M / F	6 (43%) / 8 (57%)
Age (years)	51 ± 15
T1DM Duration (years)	26 ± 17
Body Mass Index (kg/m²)	24.5 ± 3.8
HbA1c (%)	7.09 ± 0.82
Treatment Characteristics	
Multiple Daily Injections / Insulin Pump	3 (21%) / 11 (79%)
CGM	
• Abbott FreeStyle Libre	12 (86%)
• Medtronic Enlite	1 (7%)
• Dexcom G4	1 (7%)

Table 1: Characteristics [M ± SD, n (%)]

Based on those data, predictive algorithms were trained to estimate future blood glucose fluctuations, with 30- to 90-minutes prediction horizons. The method combines pharmacokinetic modelling with machine learning algorithms.

Each prediction model is patient-specific. It is initially fitted on a training dataset corresponding to an average of 9 days, using a 5-fold cross-validation method, and the remaining days of available data are used to provide an unbiased evaluation of final models.

RESULTS

Our algorithms have been evaluated using Parkes error grid analysis for 30-, 60- and 90-minutes prediction horizons. Parkes EGA is a standardised metric to measure the performance of CGM signals in relation to reference measurements, as described in **[A]**. It breaks down a scatterplot of reference glucose readings and evaluated glucose reading into five regions:

- Zone A: Clinically accurate – No effect on clinical action;
- Zone B: Little to no effect on clinical outcome;
- Zone C: Likely to affect clinical outcome;
- Zone D: Could have significant medical risk;
- Zone E: Could have dangerous consequences.

	30-minutes		60-minutes		90-minutes	
ID	MARD*	Parkes A+B	MARD*	Parkes A+B	MARD*	Parkes A+B
#01	7.78%	99.69%	17.75%	97.50%	25.27%	92.60%
#02	5.95%	99.91%	13.93%	98.40%	21.26%	96.40%
#03	6.76%	99.99%	14.79%	98.70%	20.96%	96.90%
#04	9.43%	99.96%	21.55%	97.40%	30.09%	94.80%
#05	7.29%	100.00%	9.63%	100.00%	13.21%	99.90%
#06	4.99%	100.00%	15.58%	98.50%	20.87%	98.90%
#07	7.93%	99.89%	14.21%	99.10%	19.74%	96.60%
#08	6.09%	100.00%	14.22%	97.20%	19.80%	92.50%
#09	6.55%	100.00%	19.99%	99.50%	24.63%	95.50%
#10	9.36%	99.58%	13.40%	99.50%	17.28%	98.30%
#11	5.56%	100.00%	10.75%	97.50%	20.41%	95.70%
#12	7.20%	99.98%	14.49%	97.50%	17.77%	96.20%
#13	6.30%	100.00%	14.83%	99.10%	22.22%	95.50%
#14	6.57%	100.00%	11.82%	100.00%	17.40%	98.30%
	6.98% ± 1.30%	**99.93% ± 0.13%**	**14.78% ± 3.25%**	**98.56% ± 1.00%**	**20.78% ± 4.08%**	**96.29% ± 2.15%**

Table 2: Results for 30-, 60-, 90-min prediction horizons [M ± SD]
** MARD = Mean Absolute Relative Deviation*

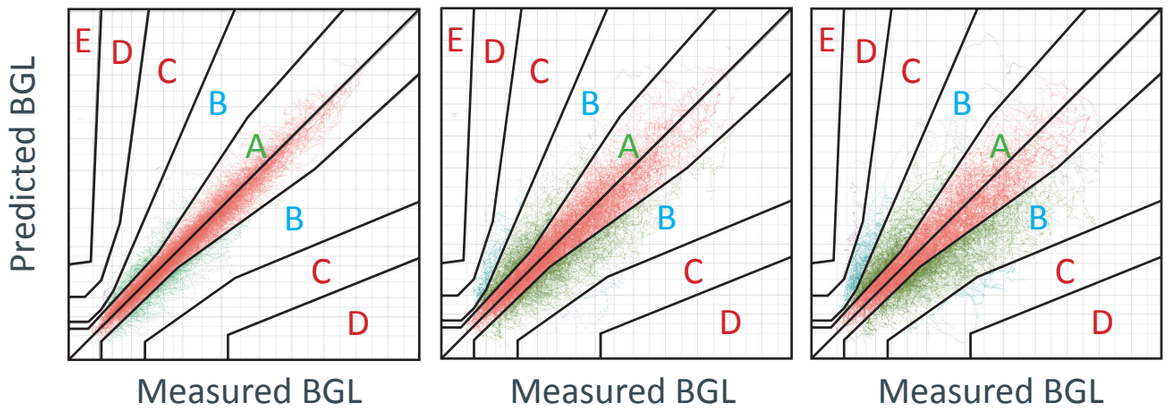

Figure 1: Results using Parkes EGA

DISCUSSIONS

Our prediction algorithms showed promising results since 99.9, 98.6 and 96.3% of computed BG values were in Parkes EGA A+B zones at 30-, 60- and 90-min horizons, respectively. The 60- and 90-minutes models cannot consider events occurring within the prediction horizon and which could impact BG levels, which partly explains the accuracy decreasing. But they are sufficient to give relevant insights on BG trends.

CONCLUSION

Integrated inside a mobile application to support decision-making process, this technology could help patients anticipate and avoid upcoming occurrence of hypoglycaemia and hyperglycaemia, in particular during night time. It could also be used on top of an Artificial Pancreas MPC model, allowing for more personalization and better regulation of the system, particularly during unstable phases with rapid glucose changes.

For future direction, as an activity tracker was used during the study to monitor heart rate of each patient during day and night, we will integrate this data into the training process of the prediction algorithms which could help improving results.

REFERENCES

[A] J.L. Parkes, S.L. Slatin, S. Pardo, B.H. Ginsberg. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care*, 23(8):1143–1148, 2000.