Behavior Analysis of Sex based Cohorts Using the Toolset of Artificial Intelligence Based Insulin Sensitivity Prediction Methods

Bálint Szabó * **Ákos Szlávecz** * **Béla Paláncz** * **Péter Somogyi** *
Geoff Chase ** **Balázs Benyó** *

* *Department of Control Engineering and Information Technology, Budapest University of Technology and Economics, Budapest, Hungary*

** *University of Canterbury, Christchurch, New Zealand*

Abstract: Tight glycaemic control (TGC) is a treatment in the intensive care in order to avoid stress-induced hyperglycaemia. The insulin sensitivity (SI) prediction is an essential step of the best performing, clinically applied so-called STAR (Stochastic-TARgeted) TGC protocol. Previous results showed performance improvement of the SI prediction using artificial intelligence methods. This study analyses the clinical performance of distinct artificial intelligence based SI prediction methods (2 different neural network based prediction methods: Classification Deep Network and Mixture Density Network with 3 different parametrizations and 2 variants: sex-specific and non sex-specific for each). In-silico validation was used for evaluation simulating the treatment of 171 virtual patients. Based on the results the number of input parameters involved into the prediction can effectively increase the reliability of the SI prediction. Improvements in the performance are also experienced in several cases by using sex-specific models.

Keywords: Insulin sensitivity prediction, Model based Tight Glycaemic Control, Artificial intelligence, In-silico validation, Deep neural network, Mixture Density Network

1. INTRODUCTION

Applying tight glycaemic control (TGC) is an essential treatment in the intensive care therapy in order to avoid stress-induced hyperglycaemia (McCowen et al. (2001); Ali et al. (2008)) resulting in definite clinical benefits (Van den Berghe et al. (2001); Krinsley (2018)).

The STAR (Stochastic-TARgeted) protocol is a model-based TGC protocol successfully implementing safe and efficient patient treatment (Benyó et al. (2012); Stewart et al. (2016); Dubois et al. (2017); Le Compte et al. (2012); Schultz et al. (2012)). STAR uses the clinically validated physiological model, called Intensive Control Insulin-Nutrition-Glucose (ICING) to describe the glucose-insulin dynamics, and – in its original version, that is used in this study – a 2D stochastic model to manage patient-specific metabolic variability (Evans et al. (2012)) displayed in Figure 1.

Most of the parameters in the ICING model are set to a population based constant except insulin sensitivity. The patient-specific insulin sensitivity (SI) is the only dynamic parameter calculated hourly based and aimed to capture the current state of the patient (Chase et al. (2011); Suhaimi et al. (2010)). SI is identified from the clinical

⋆ The research was supported by the Hungarian National Scientific Research Foundation, Grant No. K116574, by the NRDI Fund (TKP2020 IES,Grant No. BME-IE-BIO) and by H2020 MSCA-RISE DCPM (#872488) grant.

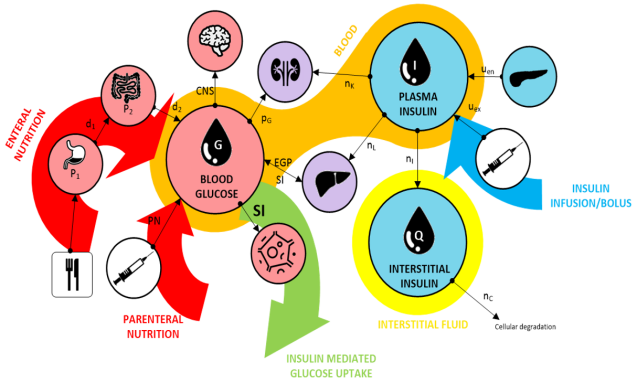

Fig. 1. Schematic representation of the physiological processes described by the ICING model.

treatment data (insulin dosing, nutrition intake and BG measurements) during the treatment of the patients.

The optimal treatment selection method consists of three main steps shown in Figure 2:

(1) Identification of current SI;
(2) Future SI range prediction with a 90% likelihood;
(3) Best treatment option selection.

For the prediction of future SI range STAR applies a conditional density function defining the conditional probability distribution of $SI(t+1)$ for a given $SI(t)$ shown in Figure

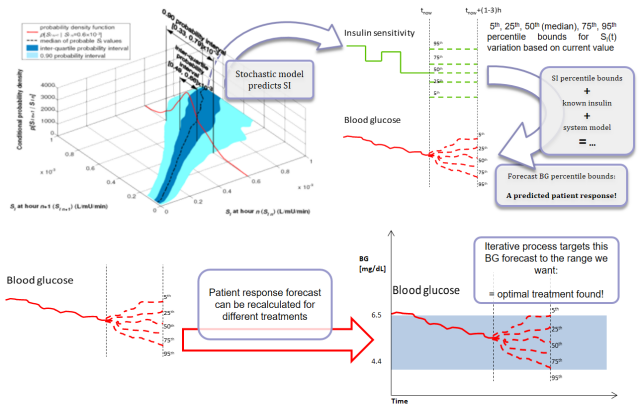

Fig. 2. Illustration of optimal treatment selection method of STAR tight glycaemic control protocol.

3 (Le Compte et al. (2010); Lin et al. (2008a)). Similar density functions are used to define the 90% confidence interval of SI in the future, in one, two, or three hours.

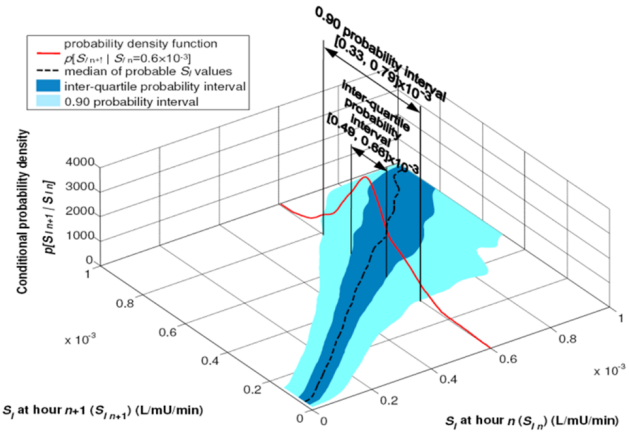

Fig. 3. Conditional density function defining the conditional probability distribution of $SI(t+1)$ for a given $SI(t)$.

In our previous work we implemented and compared two new, artificial intelligence (AI) based prediction methods for the STAR protocol: a Classification Deep Network (CDN) based method and a Mixture Density Network (MDN) based method (Benyó et al. (2020)). In the previous work we have evaluated the methods based on metrics describing the prediction power and accuracy (Szabó (2020)). In this work we investigated the clinical performance of the STAR protocol applying the new, AI based prediction methods using in-silico simulations.

In a recent study (Uyttendaele et al. (2021)) Uyttendaele et al. investigated whether there is any difference between men and woman in insulin resistance. Thus, in this work we trained dedicated models using sex based separation and made performance evaluations with the STAR protocol. This experimental setup allowed us not only to validate the AI based approach but also compare the performance of the sex specific models.

Input parameters	All	Female	Male
$SI(t)$	CDN/MDN	CDN/MDN	CDN/MDN
$SI(t-1), SI(t)$	CDN/MDN	CDN/MDN	CDN/MDN
$SI(t-1), SI(t),$ $BG(t), I(t), N(t)$	CDN MDN	CDN MDN	CDN MDN

Table 1. Overview of neural network models used for prediction of $SI(t+1)$ in this study.

2. METHODS AND DATA

The above mentioned two artificial neural network based methods (MDN, CDN) were used with three different parametrizations. For each of them, there are both sex-specific and non-specific variants. These parametrizations differ in the number of input parameters so they are referred by a dimension number in the name of the model variant (e.g. 1D sex-specific CDN) referring to the number of input parameter it uses. The actual attributes that were used as input parameters in the prediction can be seen in Section 2.3. Table 1 shows a summary of the applied models.

2.1 CDN network

The Classification Deep Network method – suggested by us in Benyó et al. (2020) – uses multi class classification to predict the confidence interval of $SI(t+1)$. The $SI(t+1)$ domain was separated to disjunct intervals and the data points was labeled with the number of the subinterval which contains the $SI(t+1)$ value of the data point.

The values of the softmax output layer of the network assign a probability for each subinterval that means the probability of containing the $SI(t+1)$ value. This histogram can be interpreted as a discrete distribution of the $SI(t+1)$ value. Thus the confidence interval can be calculated by combining the probable subintervals. For this combination a normal fitting can be used by calculating the mean and standard deviation of the histogram.

More details of the applied CDN network can be found in Benyó et al. (2020).

2.2 MDN network

The Mixture Density Network (Bishop (1994)) is a method to approximate a conditional distribution. The network approximates the distribution with a Gaussian mixture distribution. Therefore its output consists of the parameters of each sub-distribution such as means, deviation and weights. The number of the distributions that builds up the mixture can be specified as hyper parameter of the method.

This method also requires a post processing step to calculate the confidence interval. Endpoints of the interval are calculated by numerical inversion on the cumulative density function of the network output.

More details of the applied MDN network can be found in Benyó et al. (2020).

2.3 Data used

The analysis involved 354 patients from 3 different clinical settings. In the perspective of the neural network training

this means 62433 data point. These data points consist of 15657 data points coming from male patients and 8264 data point coming from female patients. For the remaining 38512 data point there was no sex information.

For each data point the following attributes has been calculated:

- current BG value linearly interpolated from measured BG,
- current insulin sensitivity value of the patient (SI is identified for every hour),
- previous insulin sensitivity value of the patient (one hour before),
- carbohydrate intake (N) during previous hour (in milligram),
- amount of administrated insulin (I) during previous hour (in milliunit).

Based on the data point attributes above 3 model parametrizations have been defined for both the CDN and MDN prediction method. The 1D models use only the current SI as input parameter, the 2D models use the one hour early SI in addition to that and the 5D models use all the earlier mentioned data point attributes as input in the prediction. Even though in the 5D case some of the parameters may correlate with each other, we assume potential benefits of using additional parameters as the human metabolic system shows stochastic behaviour components (Benyó et al. (2016); Lin et al. (2008b); Paláncz et al. ("2016"); Benyó (2019)).

Instead of connecting the sex information as input parameter to the prediction, separate models were created to use in the sex sensitive protocol. These models were trained on the sex selected data points mentioned above. The sex sensitive protocol uses the adequate model based on the patient's sex info.

More details about the effects of involving additional input parameters into the prediction can be found in Uyttendaele et al. (2019); Davidson et al. (2020); Szabó et al. (2021).

2.4 In-Silico validation

In this work we applied in-silico virtual trials (Chase et al. (2010)), to test the clinical performance of the protocol using different methods on different cohorts.

When using virtual trials the clinical environment from historical patient data are simulated by creating virtual patients and then treat them based on the decisions of the protocol. The base idea behind the trial is that the insulin sensitivity of the patient is independent from the nutrition and insulin dosage, it depends only on the state of the patient. Therefore, historical SI values of the patients remain relevant even with modified insulin and nutrition dosage. During the in-silico validation patients are processed one by one. The first BG measurement is taken from historical data then new BG measurements are calculated by the ICING model using the historical SI values.

At the end of simulations clinically important results are extracted, such as hours spent in different clinically relevant BG regions (as it is listed in Table 2 and Table 3) and count of hypo events. For this calculation the BG measurements were resampled hourly using linear interpolation to make them evenly distributed in time.

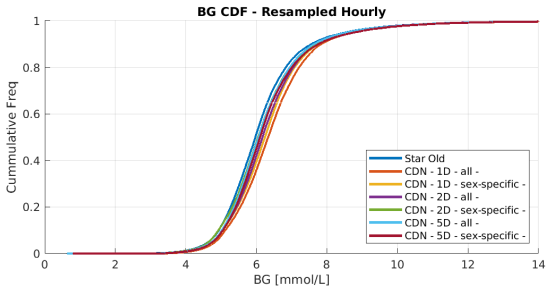

Fig. 4. CDF curve based comparison of CDN networks by dimension and sex specificity

To compare the models based on the statistics it has to be considered that both hypoglycaemic and hyperglycaemic events have to be avoided. Thus, the lower number of episodes and relative times in these regions belong to higher safety. Between these two event types the hypoglycaemic is more dangerous as it can cause severe pathophysiological symptoms in relatively short time.

From the evaluation two types of statistics have been created. The cumulative distribution function (CDF) of the BG values and the cohort statistics.

3. RESULTS

The CDF curve based comparison of the model variants can be seen in Figure 4. The curves are similar in shape but tends to be translated on the x axis compared to the original STAR protocol results meaning that the BG values resulted by the AI prediction are shifted to higher BG regions.

The hypoglycaemic events that should be avoided from safety aspects are the BG measurements below $4.4 mmol/L$. The corresponding part of the CDF curves can be seen in Figure 5 and 7. Decreased number of hypoglycaemic events are resulted by all the AI prediction methods in different measures. In general the largest improvement is resulted by the 1D version predictions. These hypoglycaemic events and the most serious hypoglycaemic events – the BG measurement below $2.5 mmol/L$ – can be also seen in Table 2 and 3. In these tables we can see that the differences between the methods are relatively small: The number of episodes below $2.2 mmol/L$ are between 4 and 9 representing 0.02 and 0.04 percent of the treatment hours and the number of episodes below $4.0 mmol/L$ are between 83 and 125 representing 0.76 and 1.33 percent of the total treatment hours.

In Table 2 and Table 3 there is also information about the relative time spent in the hyperglycaemic (above 10 mmol/L) range. The application of the AI based prediction moderately increases the BG measurements in higher regions compared to the original STAR protocol: The number of hours spent in the $6.0 - 9.0 mmol/L$ increased up to 15%. However, the number of serious hyperglycaemic events (above $10 mmol/L$) has not increased significantly.

	STAR Current	1D all	1D sex-spec.	2D all	2D sex-spec.	5D all	5D sex-spec.
Num BG measurements	13186	13210	13367	13241	13551	13365	13801
Num episodes < 4.0 mmol/L	105	83	92	88	105	101	99
Num episodes < 2.22 mmol/L	9	4	8	5	4	4	4
BG median (mmol/L)	6.0	6.3	6.2	6.2	6.0	6.1	6.1
[IQR]	[5.39-6.62]	[5.69-7.00]	[5.61-6.88]	[5.58-6.84]	[5.46-6.73]	[5.50-6.74]	[5.54-6.78]
BG mean (mmol/L)	6.1	6.4	6.3	6.2	6.1	6.1	6.2
BG StDev (mmol/L)	1.21	1.21	1.21	1.21	1.22	1.21	1.21
% BG < 2.22 (mmol/L)	0.04	0.02	0.03	0.02	0.02	0.02	0.02
% BG < 4.0 (mmol/L)	1.30	0.76	0.98	0.95	1.22	1.07	1.01
% BG < 4.4 (mmol/L)	3.03	1.87	2.25	2.23	2.79	2.47	2.25
% BG within 4.4 - 6.1 (mmol/L)	68.65	55.62	59.54	61.97	65.40	65.56	64.76
% BG within 4.4 - 7.0 (mmol/L)	80.31	72.97	76.10	76.94	78.31	78.82	77.79
% BG within 4.4 - 8.0 (mmol/L)	89.96	89.45	89.74	90.01	89.53	90.07	89.53
% BG within 4.4 - 9.0 (mmol/L)	93.35	94.11	93.91	94.03	93.27	93.77	93.51
% BG within 6.0 - 9.0 (mmol/L)	44.49	59.74	56.32	53.74	47.81	48.56	50.23
% BG within 8.0 - 10 (mmol/L)	5.05	6.65	6.00	5.77	5.54	5.41	6.01
% BG > 10 (mmol/L)	1.97	2.04	2.03	1.99	2.15	2.06	2.23

Table 2. Statistics for Classification Deep Network (CDN) has been generated showing relative time spent in the different BG ranges. BG levels have been resampled except in the first three rows. "1D", "2D", "5D" refers to the number of prediction inputs, while in case of "sex-spec.", and "all" the sex specific models were applied or not applied during in-silico simulation. The column "STAR" contains the results of the original protocol.

	STAR Current	1D all	1D sex-spec.	2D all	2D sex-spec.	5D all	5D sex-spec.
Num BG measurements	13186	13228	13283	13157	13301	13421	13693
Num episodes < 4.0 mmol/L	105	101	106	116	125	119	110
Num episodes < 2.22 mmol/L	9	5	7	8	7	5	5
BG median (mmol/L)	6.0	6.1	6.2	6.1	6.0	5.9	6.0
[IQR]	[5.39-6.62]	[5.51-6.74]	[5.57-6.80]	[5.47-6.69]	[5.40-6.64]	[5.39-6.56]	[5.40-6.60]
BG mean (mmol/L)	6.0	6.2	6.2	6.1	6.0	6.0	6.0
BG StDev (mmol/L)	1.21	1.21	1.21	1.22	1.22	1.21	1.21
% BG < 2.22 mmol/L	0.04	0.02	0.03	0.04	0.04	0.02	0.02
% BG < 4.0 mmol/L	1.30	1.12	1.10	1.29	1.40	1.33	1.26
% BG < 4.4 mmol/L	3.03	2.57	2.61	3.08	3.35	3.20	3.09
% BG within 4.4 - 6.1 mmol/L	68.65	64.60	60.98	65.97	67.63	69.83	68.94
% BG within 4.4 - 7.0 mmol/L	80.31	78.66	77.67	79.19	79.83	81.37	81.18
% BG within 4.4 - 8.0 mmol/L	89.96	90.13	89.86	89.77	89.82	90.19	90.23
% BG within 4.4 - 9.0 mmol/L	93.35	93.79	93.73	93.33	93.10	93.32	93.36
% BG within 6.0 - 9.0 mmol/L	44.49	50.88	55.19	49.13	45.95	42.83	44.60
% BG within 8.0 - 10 mmol/L	5.05	5.31	5.54	5.22	4.91	4.69	4.75
% BG > 10 mmol/L	1.97	1.99	2.00	1.94	1.94	1.93	1.93

Table 3. Statistics for Mixture Density Network (MDN) has been generated showing relative time spent in the different BG ranges. BG levels have been resampled except in the first three rows. "1D", "2D", "5D" refers to the number of prediction inputs, while in case of "sex-spec.", and "all" the sex specific models were applied or not applied during in-silico simulation. The column "STAR" contains the results of the original protocol.

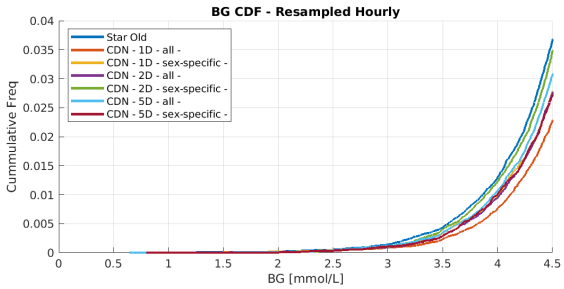

Fig. 5. CDF curve based comparison of CDN networks by dimension and sex specificity in the low BG range

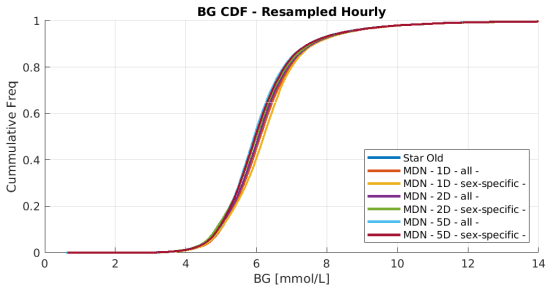

Fig. 6. CDF curve based comparison of MDN networks by dimension and sex specificity

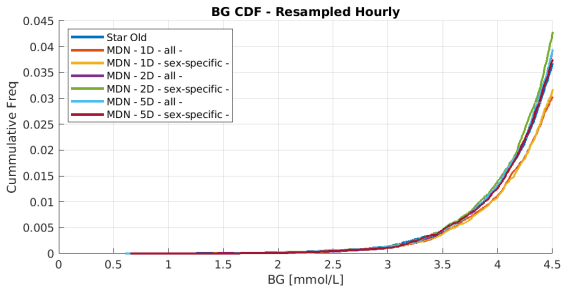

Fig. 7. CDF curve based comparison of MDN networks by dimension and sex specificity in the low blood glucose range $4.4 mmol/L$

4. DISCUSSION

The CDF comparison shows that there are differences in the treatment results depending on the number of inputs and the sex-specificity of the model used for SI prediction. The easily recognisable shift between the CDF curves in the central $(4.0 - 8.0$ mmol/L) BG region – see Figure 4 – is basically caused by the different proportions of the hypoglycaemic events in BG region below 4.0 mmol/L that can be seen in Figure 5.

From the safety aspect all of the new CDN models and most of the MDN models perform better than the currently used method, the 1D non-sex-specific models produce the best result based on the Figures 4 and 6. Measured by the number of episodes (second row of Table 2 and Table 3) and by the proportion of the time spent in hypoglycaemic BG range (row labelled by "% BG 4.4 mmol/L" of Table 2 and Table 3) the performance is two times better in the case of the best prediction method. The absolute numbers are relatively low but the about half percent improvement corresponds to about 120 treatment hours in the case of 350 patient. It is important to mention that the safety improvement is not due to the more frequent BG measurements. The number of measurements has not increased significantly by using AI based prediction as it can be seen in the first row of the tables.

Comparing the MDN methods with each other the involvement of additional parameters did not improve the safety of the protocol. The largest difference between the models can be experienced in the time spent in the BG region within 6.0 to $9.0 mmol/L$.

This analysis showed that there are differences in SI and BG behaviour of the patients based on their sex which confirms the results of previous similar studies (Uyttendaele et al. (2021)). Extracting this information to improve the performance of the SI prediction is not straightforward. In this analysis dedicated models are defined to learn the sex-specific conditional distribution using dataset filtered by sex. This way, unfortunately, the sex-specific models are trained with less data sample than the non-specific variants. Moreover, there were significant amount of patient data lacking sex information that can be used in the training of the non-specific models but not in the sex-specific ones. If we consider that the neural networks will perform better by the bigger volume of training data even if the quality of the data is not so good, than this difference in the volume can lead to significant performance loss.

It was an interesting result that the 1D prediction of both CDN and MDN networks the sex-specific models are less performant than the non sex-specific variants. There are also cases when the sex-specific and non sex-specific model perform very similarly. These results could be also caused by the smaller dataset used by the training of sex-specific models.

Therefore, the subsequent step of this research will be to create neural models using the sex information as an input parameter instead of using dedicated models separately. The sex information can be encoded on a range of -1 to 1 where the two sexes are the border values. A patient with unspecified sex can be coded with zero as sex parameter. By this solution the network can be trained on the whole dataset.

5. CONCLUSION

In this paper two neural network based SI prediction methods with three different parametrizations were evaluated and compared with their sex-specific and non-sex-specific counterparts, and also compared with the currently used SI prediction method of the STAR protocol. The evaluation was done using in-silico validation simulating the treatment of 171 virtual patients. Results show that building dedicated sex-specific models can improve the SI prediction performance but in some cases a non sex-sensitive model can lead to similar results, as well. However, all of the presented AI based prediction models were more efficient from one or more clinical aspects than the currently used 1D prediction. Thus, the flexibility of the AI based methods to easily develop multi dimensional predictions is highly appreciated in the SI prediction and may result in clinical benefits.

Based on the presented results the specific effect of the additional patient parameters on the SI prediction performance can not be clearly identified. The most probable reason for this is the limited dataset. The unbalanced training dataset may also affects the outcome. The sex-specific models were trained on filtered datasets based on sex information causing significantly reduced sized training data which negatively affects the prediction performance. To overcome on these difficulties we encourage the creation of models handling the sex information of the patients as an input parameter. A model like that can be trained on the whole cohort dataset and may reach better predictive performance.

REFERENCES

Ali, N.A., O'Brien Jr, J.M., Dungan, K., Phillips, G., Marsh, C.B., Lemeshow, S., Connors Jr, A.F., and Preiser, J.C. (2008). Glucose variability and mortality in patients with sepsis. *Critical care medicine*, 36(8), 2316.

Benyó, B. (2019). In-silico analysis of stochastic modelling of human blood glucose regulatory system. In *2019 IEEE 23rd International Conference on Intelligent Eng. Systems (INES)*, 000229–000234.

Benyó, B., Stewart, K., Homlok, J., Pretty, C.G., Chase, J.G., and Paláncz, B. (2016). Specific validation analysis

of stochastic icing model based estimation of insulin sensitivity profile using clinical data. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 004317–004324.

Benyó, B., Illyés, A., Szabó Némedi, N., Le Compte, A.J., Havas, A., Kovács, L., Fisk, L., Shaw, G.M., and Chase, J.G. (2012). Pilot study of the sprint glycemic control protocol in a hungarian medical intensive care unit. *Journal of Diabetes Science and Technology*, 6(6), 1464–1477.

Benyó, B., Paláncz, B., Ákos Szlávecz, Szabó, B., Anane, Y., Kovács, K., and Chase, J.G. (2020). Artificial intelligence based insulin sensitivity prediction for personalized glycaemic control in intensive care. *IFAC-PapersOnLine*, 53(2), 16335–16340. 21th IFAC World Congress.

Bishop, C. (1994). Mixture density networks. Workingpaper, Aston University.

Chase, J.G., Le Compte, A.J., Preiser, J.C., Shaw, G.M., Penning, S., and Desaive, T. (2011). Physiological modeling, tight glycemic control, and the ICU clinician: what are models and how can they affect practice? *Annals of Intensive Care*, 1(1), 1–8.

Chase, J.G., Suhaimi, F., Penning, S., Preiser, J.C., Le Compte, A.J., Lin, J., Pretty, C.G., Shaw, G.M., Moorhead, K.T., and Desaive, T. (2010). Validation of a model-based virtual trials method for tight glycemic control in intensive care. *BioMedical Eng. OnLine*, 9(1), 84.

Davidson, S.M., Uyttendaele, V., Pretty, C.G., Knopp, J.L., Desaive, T., and Chase, J.G. (2020). Virtual patient trials of a multi-input stochastic model for tight glycaemic control using insulin sensitivity and blood glucose data. *Biomedical Signal Processing and Control*, 59, 101896.

Dubois, J., Van Herpe, T., van Hooijdonk, R.T., Wouters, R., Coart, D., Wouters, P., Van Assche, A., Veraghtert, G., De Moor, B., Wauters, J., Wilmer, A., Schultz, M.J., Van den Berghe, G., and Mesotten, D. (2017). Software-guided versus nurse-directed blood glucose control in critically ill patients: the LOGIC-2 multicenter randomized controlled clinical trial. *Critical Care*, 21(1), 212.

Evans, A., Le Compte, A., Tan, C.S., Ward, L., Steel, J., Pretty, C.G., Penning, S., Suhaimi, F., Shaw, G.M., Desaive, T., et al. (2012). Stochastic targeted (STAR) glycemic control: design, safety, and performance. *Journal of diabetes science and technology*, 6(1), 102–115.

Krinsley, J.S. (2018). Is It Time to Rethink Blood Glucose Targets in Critically Ill Patients? *Chest*, 154(5), 1004–1005.

Le Compte, A.J., Lee, D.S., Chase, J.G., Lin, J., Lynn, A., and Shaw, G.M. (2010). Blood glucose prediction using stochastic modeling in neonatal intensive care. *Biomedical Eng., IEEE Trans. on*, 57(3), 509–518.

Le Compte, A.J., Lynn, A.M., Lin, J., Pretty, C.G., Shaw, G.M., and Chase, J.G. (2012). Pilot study of a model-based approach to blood glucose control in very-low-birthweight neonates. *BMC pediatrics*, 12(1), 117.

Lin, J., Lee, D., Chase, J.G., Shaw, G.M., Le Compte, A., Lotz, T., Wong, J., Lonergan, T., and Hann, C.E. (2008a). Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care. *Computer methods and programs in biomedicine*, 89(2), 141–152.

Lin, J., Lee, D., Chase, J.G., Shaw, G.M., Le Compte, A., Lotz, T., Wong, J., Lonergan, T., and Hann, C.E. (2008b). Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care. *Computer methods and programs in biomedicine*, 89(2), 141–152.

McCowen, K.C., Malhotra, A., and Bistrian, B.R. (2001). Stress-induced hyperglycemia. *Critical care clinics*, 17(1), 107–124.

Paláncz, B., Stewart, K., Homlok, J., Pretty, C.G., Chase, J.G., and Benyó", B. ("2016"). "stochastic simulation and parameter estimation of the icing model". *"IFAC-PapersOnLine"*, "49"("5"), "218 – 223". "4th IFAC Conference on Intelligent Control and Automation SciencesICONS 2016".

Schultz, M., Harmsen, R., Korevaar, J., Abu-Hanna, A., Van Braam Houckgeest, F., Van Der Sluijs, J., and Spronk, P. (2012). Adoption and implementation of the original strict glycemic control guideline is feasible and safe in adult critically ill patients. *Minerva anestesiologica*, 78(9), 982—995.

Stewart, K.W., Pretty, C.G., Tomlinson, H., Thomas, F.L., Homlok, J., Szabó Némedi, N., Shaw, G.M., Benyó, B., and Chase, J.G. (2016). Safety, efficacy and clinical generalization of the STAR protocol: a retrospective analysis. *Annals of intensive care*, 6(1), 24.

Suhaimi, F., Le Compte, A., Preiser, J.C., Shaw, G.M., Massion, P., Radermecker, R., Pretty, C.G., Lin, J., Desaive, T., and Chase, J.G. (2010). What makes tight glycemic control tight? The impact of variability and nutrition in two clinical studies. *Journal of diabetes science and technology*, 4(2), 284–298.

Szabó, B. (2020). *Artificial intelligence based prediction of insulin sensitivity of intensive care patients*. Budapest University of Technology and Economics.

Szabó, B., Szlávecz, k.J., Paláncz, B., Geoffrey, C., and Benyó, B.I. (2021). In-silico simulation based evaluation of insulin prediction method for personalized medical treatment. In *Proc. of the Workshop on the Advances of Information Technology (WAIT) 2021*, 145–155.

Uyttendaele, V., Chase, J.G., Knopp, J.L., Gottlieb, R., Shaw, G.M., and Desaive, T. (2021). Insulin sensitivity in critically ill patients: are women more insulin resistant? *Annals of Intensive Care*, 11(1), 2110–5820.

Uyttendaele, V., Knopp, J.L., Davidson, S., Desaive, T., Benyo, B., Shaw, G.M., and Chase, J.G. (2019). 3D kernel-density stochastic model for more personalized glycaemic control: development and in-silico validation. *BioMedical Eng. OnLine*, 18(1), 102.

Van den Berghe, G., Wouters, P., Weekers, F., Verwaest, C., Bruyninckx, F., Schetz, M., Vlasselaers, D., Ferdinande, P., Lauwers, P., and Bouillon, R. (2001). Intensive insulin therapy in critically ill patients. *New England Journal of Medicine*, 345(19), 1359–1367.