Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 53-2 (2020) 16335–16340

Artificial Intelligence Based Insulin Sensitivity Prediction for Personalized Glycaemic Control in Intensive Care

Balázs Benyó * **Béla Paláncz** * **Ákos Szlávecz** * **Bálint Szabó** *
Yahia Anane * **Katalin Kovács** *** **J. Geoffrey Chase** **

* *Budapest University of Technology and Economics, Budapest, Hungary.*
** *University of Canterbury, Christchurch, New Zealand*
*** *Széchenyi István University, Győr, Hungary*

Abstract:
Stress-induced hyperglycaemia is a frequent complication in the intensive therapy that can be safely and efficiently treated by using the recently developed model-based tight glycaemic control (TGC) protocols. The most widely applied TGC protocol is the STAR (Stochastic-TARgeted) protocol which uses the insulin sensitivity (SI) for the assessment of the patients state. The patient-specific metabolic variability is managed by the so-called stochastic model allowing the prediction of the 90% confidence interval of the future SI value of the patients. In this paper deep neural network (DNN) based methods (classification DNN and Mixture Density Network) are suggested to implement the patient state prediction. The deep neural networks are trained by using three years of STAR treatment data. The methods are validated by comparing the prediction statistics with the reference data set. The prediction accuracy was also compared with the stochastic model currently used in the clinical practice. The presented results proved the applicability of the neural network based methods for the patient state prediction in the model based clinical treatment. Results suggest that the methods' prediction accuracy was the same or better than the currently used stochastic model. These results are the initial successful step in the validation process of the proposed methods which will be followed by in-silico simulation trials.

Keywords: machine learning, artificial intelligence, mixture density network, deep neural neural network, insulin sensitivity, tight glycaemic control, intensive care, STAR protocol

1. INTRODUCTION

Stress-induced hyperglycaemia is a frequent complication in the intensive therapy (McCowen et al. (2001); Ali et al. (2008)). Forcing the blood glucose (BG) level of these hyperglycaemic patients into the normal, so-called normoglycaemic range shows definite clinical benefits (Van Den Berghe et al. (2001); Krinsley (2004); Reed et al. (2007); Chase et al. (2010)). This therapy is called in general as tight glycaemic control (TGC) that includes insulin therapy and occasionally moderation of the nutrition intake of the patient.

The recently developed model-based TGC protocols successfully implement safe and efficient patient treatment (Benyo et al. (2012a); Stewart et al. (2016); Dubois et al. (2017); Le Compte et al. (2012); Schultz et al. (2012)). The STAR (Stochastic-TARgeted) TGC protocol is the most widely applied among them, it is used in four different countries (Stewart et al. (2016)). STAR uses a clinically validated physiological model, called Intensive Control Insulin-Nutrition-Glucose (ICING) to describe the glucose-insulin dynamics, and a population-based stochastic model to manage patient-specific metabolic variability (Evans et al. (2012)).

STAR uses the patient-specific insulin sensitivity (SI) as a key parameter (Chase et al. (2011); Suhaimi et al. (2010)) to define the state of the patient. SI describes the patient metabolic response to insulin. SI is identified from the clinical 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 1.

In the so-called 2D stochastic model of insulin sensitivity, the conditional density function defining the conditional probability distribution of $SI(t+1)$ for a given $SI(t)$ (shown in Figure 2) (Le Compte et al. (2011); Lin et al. (2008)) is used to define the 90% confidence interval of SI in the future, in one, two, or three hours. This is a key step in the STAR protocol to handle the future variability of the patients' state directly. The 2D stochastic model was created based on the treatment data of the SPRINT protocol using kernel fitting (Lonergan et al. (2006); Benyo et al. (2012a); Lin et al. (2008)).

In this paper two versions of Artificial Intelligence (AI), especially Neural Network (NN) based methods are presented to create an alternative stochastic model for the STAR protocol. The primary aim of this research is to

10.1016/j.ifacol.2020.12.659

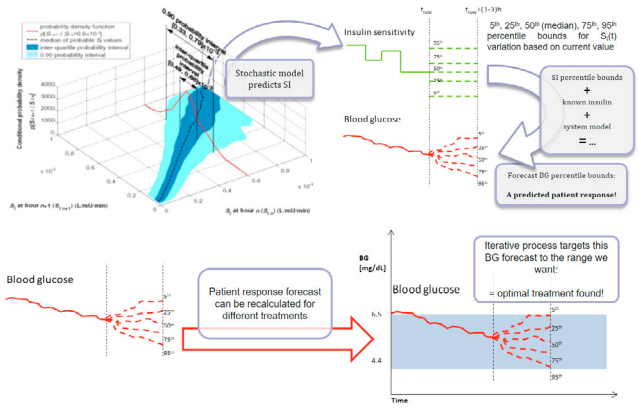

Fig. 1. Illustration of optimal treatment selection method of STAR tight glycaemic control protocol.

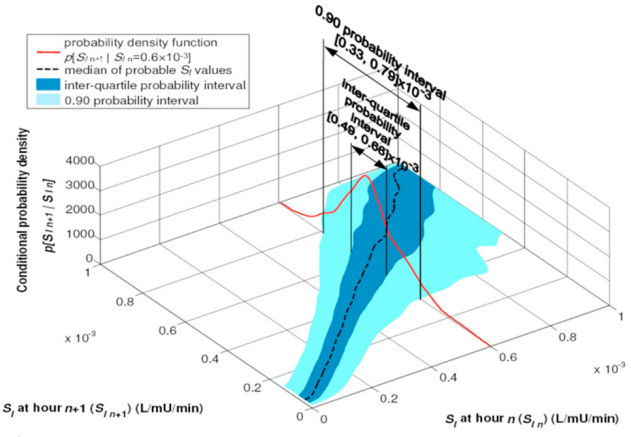

Fig. 2. Conditional density function defining the conditional probability distribution of $SI(t+1)$ for a given $SI(t)$.

develop and validate an alternative methodology for SI prediction and asses its accuracy in the patient treatment.

The potential benefits of using an NN based method are the flexibility of the method to involve further patient parameters into the prediction and the opportunity to incrementally modify the stochastic model based on the results of recent patient treatments. The benefits of involving further patient parameters into the SI prediction have been already shown by the development of the so-called 3D stochastic model (Uyttendaele et al. (2018, 2019)).

The option of regularly modifying the stochastic model by using the recently treated patients is logical and may have the benefit to follow the trends of general behaviour of the patients' state change. The NN based stochastic model creation approach provides this opportunity as well. However, the prerequisite of the development of these opportunities the development of the methodology that provides so accurate SI prediction than the one is used in the original version of the STAR protocol.

In the subsequent sections, after the brief introduction of the ICING model the neural networks used for the prediction of future SI values are defined. The prediction accuracy of the proposed methods are compared with each other and with the reference data in the section Results. The section Discussion will analyze the clinically relevant

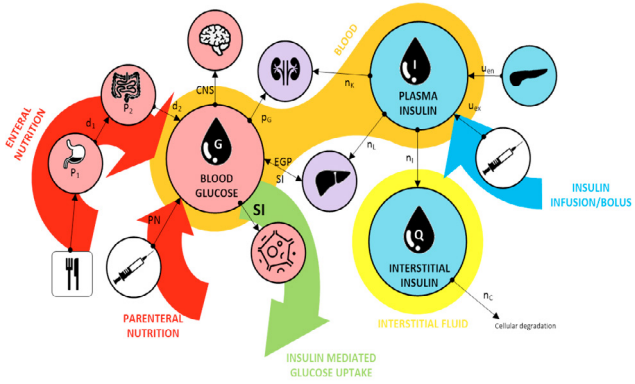

Fig. 3. Schematic representation of the physiological processes described by the ICING model.

aspects of the outcomes. The results are summarized in the section Conclusion.

2. METHODS AND DATA

2.1 ICING model

ICING is a pharmacokinetic-pharmacodynamic model developed by Lin et al. (2011) defining glucose-insulin kinetics and dynamics in the human body. Schematic representation of the physiological processes described by the model is shown in Figure 3. Complete description of the model and the model parameters are given in Lin et al. (2011), the equation describing the blood glucose (G), interstitial insulin (Q), and plasma insulin (I) are defined as follows:

$$\frac{dG(t)}{\mathrm{dt}} = -p_G G(t) - S_I(t) G(t) \frac{Q(t)}{1+\alpha_G Q(t)} + \frac{P(t) + EGP - CNS}{V_G}, \tag{1}$$

$$\frac{dQ(t)}{\mathrm{dt}} = n_I (I(t) - Q(t)) - n_C \frac{Q(t)}{1+\alpha_G Q(t)}, \tag{2}$$

$$\frac{dI(t)}{\mathrm{dt}} = -n_K I(t) - n_L \frac{I(t)}{1+\alpha_I I(t)} - n_I (I(t) - Q(t)) + \frac{u_{\mathrm{ex}}(t)}{V_I} + (1 - x_L) \frac{u_{\mathrm{en}}(t)}{V_I}. \tag{3}$$

2.2 SI Prediction Problem

During the patients' STAR treatment the SI value - representing the actual state of the patient - is identified every hour using equation (1). The stepwise time function of SI is used to create the $\{SI(t); SI(t+1)\}$ data pairs. The $SI(t+1)$ prediction from $SI(t)$ was achieved via parameter estimation of a stochastic variation of the ICING model represented by stochastic differential equations, see (Paláncz et al. (2016); Benyó et al. (2016)) and by using a stochastic model. In current research the stochastic model will be defined by an Neural Network based Artificial Intelligence method. Thus, our problem is to define the function giving $SI(t+1)$ for given $SI(t)$ based on this data set.

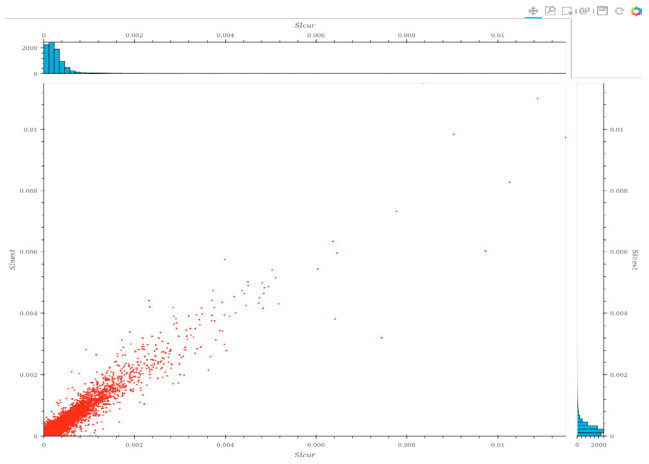

Fig. 4. Input data points defined by the $SI(t)$ and $SI(t+1)$ data pairs. The histogram of the values are shown on the top and right side of the figure.

Deep neural networks as supervised machine learning will be suggested to learn the mapping between the input $SI(t)$ and the output $SI(t+1)$ values. The NN based classification models are used to predict the class of a given input and the NN based regression models are used to predict continuous output values. One of the presented methods will apply classification model and the other regression. The classification model is a classification deep neural network (CDN) with several hidden layers. Since we have to predict the distribution of the $SI(t+1)$ value instead of a single $SI(t+1)$ value a Mixture Density Network (MDN) will be applied as a regression model.

2.3 Patient Selection and SI Data Set Used

All the patents treated by STAR between June 2016 and August 2019 in Christchurch Hospital, New Zealand were included into the study cohort. The following exclusion criteria were applied:

- patients treated less than 10 hours by STAR;
- sections of treatments where the higher border of the BG target band was above 9 mmol/L;
- sections of treatments where lower border of the BG target band was above 6 mmol/L.

Data points used for the creation of the prediction models are created for each real BG measurements. The actual $SI(t)$ and $SI(t+1)$ values are identified using equation 1 based on the treatment data. The input data points are shown in Figure 4. The histogram of the data points are shown on the top and right side of the figure. The total number of data points was $26,033$.

It can be seen that the distribution of the input data points is uneven, 97% of the SI values are below 0.001 mmol/l. Thus in the evaluation phase the lowest 10% of the SI codomain will be used.

2.4 SI Prediction Based on Deep Neural Network

To apply the classification deep neural network (CDN) for the prediction of the $SI(t+1)$ distribution based on $SI(t)$ the codomain of SI was divided into 100 equal size intervals. Each interval was associated with one output class. The input of the CDN is the $SI(t)$ value. The output layer of the CDN consists of 100 nodes, associated with the classes defined above. The CDN nodes will define for each output class the probability that the SI domain associated with the given class includes the predicted $SI(t+1)$ value.

Thus, the size of the output layer of the network is defined, it is equal with the number of classes which is 100 in our case. The input of the network is the $SI(t)$ value so there is only a single node in the input layer. The training data set contains pairs of a real number, as an input and a vector of 100 numbers between 0 and 1, interpreted as a probability, as an output. The real number is the $SI(t)$ value which will be the input of the CDN, the 100-element vector is the (discrete) probability distribution of $SI(t+1)$ which will be the output of the CDN.

The training data set is created in a way that all of these vectors contain 99 zero values and a single one value. The one value belongs to the class which has an associated SI domain that includes the $SI(t+1)$ value predicted from the $SI(t)$.

Several network topologies have been tested, in Table 1 the best performing CDN topology is defined.

Table 1. Classification deep neural network topology definition

layer type	size	activation function
input	1	-
hidden	10	tanh
hidden	20	tanh
hidden	30	tanh
hidden	40	tanh
hidden	60	tanh
output	100	softmax

The deep neural network used for SI prediction was implemented in Python using TensorFlow and Keras Gulli and Pal (2017). 80% of the input data set was used for training. The training consisted of 20 epochs.

The final output of the SI prediction is calculated by fitting a Gaussian distribution to the output of the deep neural network which is considered in this case as a histogram. This calculation is illustrated in Figure 5. In this figure the output of the deep neural network prediction is shown as a blue histogram. The fitted Gaussian distribution is the green line. The mean value of the Gaussian will be the predicted $SI(t+1)$. The 5% and 95% percentile values are also shown by blue and red lines.

2.5 SI Prediction based on Mixture Density Network

The Mixture Density Network (Bishop (1994)) is a deep network with input $SI(t)$ and with Mean values, Standard Deviations, and the Weights of the Normal Distributions providing the statistical mixture model for $SI(t+1)$. For finding these parameters in the teaching phase this network was embedded in a larger network having additional input $SI(t+1)$, see Figure 6. Teaching this network the parameters of the embedded network can be computed, and then this network can be employed for the $SI(t+1)$ prediction.

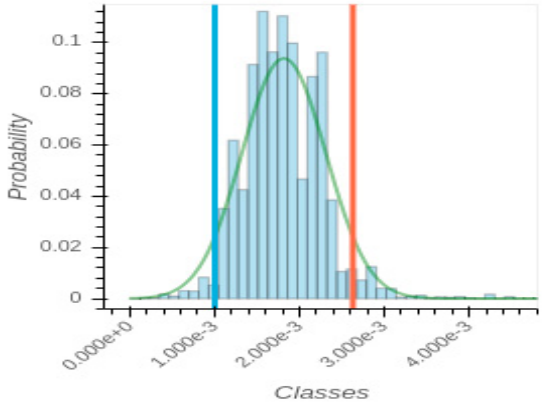

Fig. 5. Output of the CDN network prediction for a given $SI(t)$ value. The fitted Gaussian distribution (green) and the 5% and 95% percentile values (blue and red bars) are also shown.

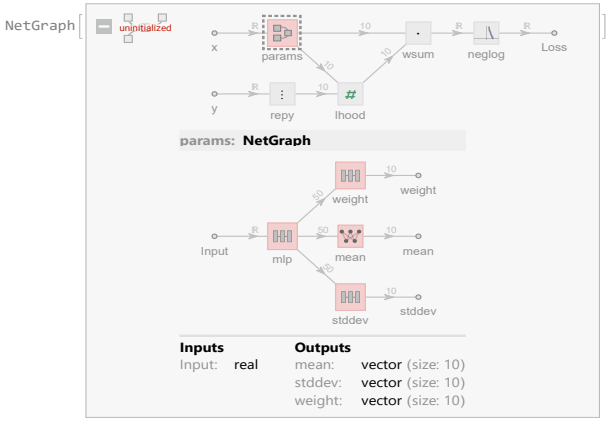

Fig. 6. Topology of Mixture Density Networks proposed for the SI prediction for learning the statistical parameters of the model of the mixture distribution of the ten normal distributions.

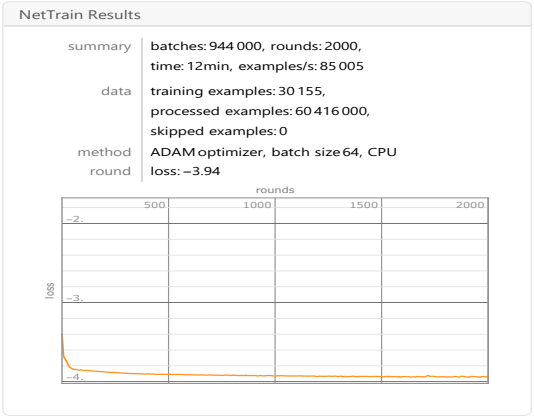

Fig. 7. Training parameters and training results of the MDN.

The training results of the network implemented in *Wolfram Mathematica* (Wolfram Language and Documentation Center (2019)) is shown in Figure 7.

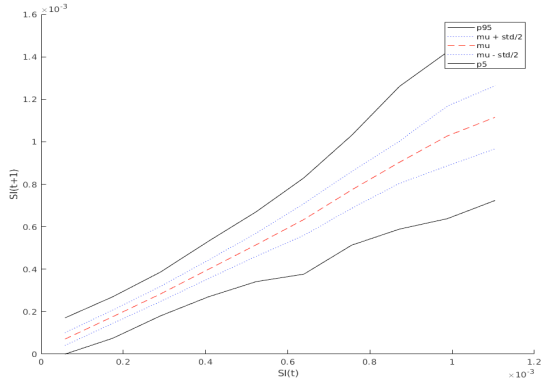

Fig. 8. Mean (red dashed line), standard deviation (two dotted blue lines), and the 5% and 95% percentile values (enveloping black solid lines) of the $SI(t+1)$ in function of $SI(t)$ in the training data set.

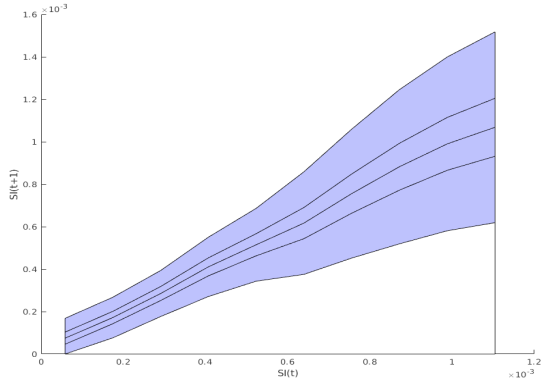

Fig. 9. Output of the CDN network prediction in function of input $SI(t)$ values. Mean value (middle solid line), standard deviation (two solid lines around the mean), and the 5% and 95% percentile values (enveloping solid lines) of the $SI(t+1)$ are shown.

3. RESULTS

The $SI(t+1)$ prediction accuracy of the proposed NN based prediction methods were evaluated by comparing the mean value, the standard deviation and the 5% and 95% percentile values with the reference data in the intervall of $SI(t)$ values including the majority (97%) of the data points. The reference data is shown in Figure 8.

The output of the CDN network prediction in function of input $SI(t)$ values is shown in Figure 9. It can be seen that the dependency of the mean value (middle solid line) is closer to linear from the $SI(t)$ value. The standard deviation (two solid lines around the mean) also grows proportionally with the $SI(t)$. The 5% and 95% percentile values (enveloping solid lines) show similar behavior to the standard deviation.

In Figure 10 the output of the CDN network prediction is compared with the reference data. The prediction accuracy, based on the two momentums are quite impressive.

In Figure 11 the output of MDN network prediction can be seen. The curves of the mean value (middle red dashed line), the standard deviation (two dotted blue lines around

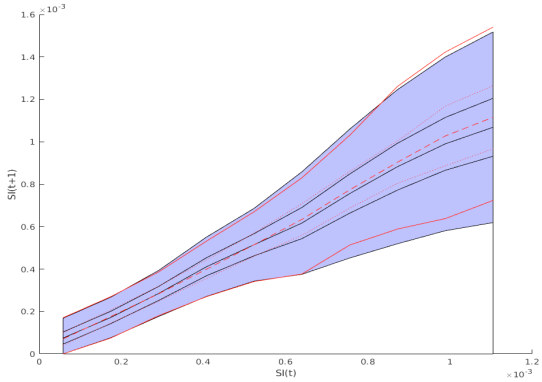

Fig. 10. Comparison of the output of the CDN network prediction with the reference data in function of input $SI(t)$ values (notations are given in Figure 9 and 8.)

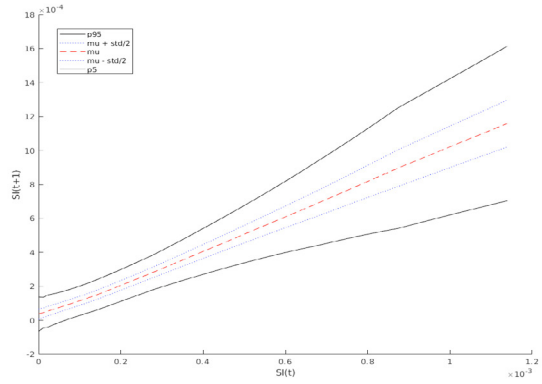

Fig. 11. Output of the MDN network prediction in function of input $SI(t)$ values. Mean value (middle red dashed line), standard deviation (two dotted blue lines around the mean), and the 5% and 95% percentile values (enveloping solid lines) of the $SI(t+1)$ are shown.

the mean) and the 5% and 95% percentile values are very much the same as the prediction output of the CDN network. This similarity can be confirmed in Figure 12. The curves representing the statistical parameters of the predicted distributions of the $SI(t+1)$ values defined by the two networks are side by side, almost covering each other.

The prediction accuracy of one of the neural networks - the CDN network - was compared with the prediction accuracy of the currently used stochastic model of the STAR protocol. In Table 2 the 90% confidence interval of the $SI(t+1)$ prediction was compared with the real $SI(t+1)$ value extracted from the treatment records. The prediction *true rate* in the first row of the table shows the proportion of the cases when the neural network 90% confidence interval includes the real $SI(t+1)$ value. Similarly, the first column of the table shows the proportion of the cases when the currently used STAR stochastic model 90% confidence interval includes the real $SI(t+1)$ value. Table also shows the proportion of the cases when the currently used STAR stochastic model and the CDN 90% confidence intervals do not include the real $SI(t+1)$ value (columns and rows with False heading).

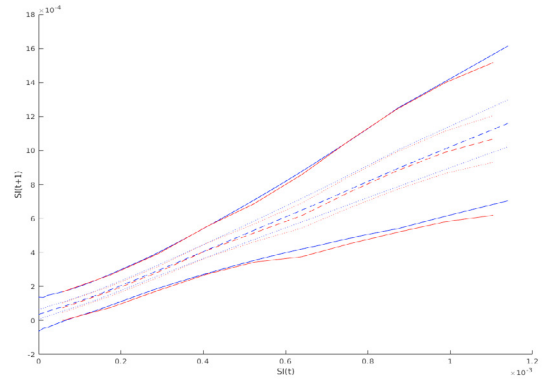

Fig. 12. Comparison of the output of the CDN network prediction with the MDN prediction in function of input $SI(t)$ values. Mean values (middle dashed line), standard deviations (two dotted lines around the mean), and the 5% and 95% percentile values (enveloping solid lines); red: CDN network prediction; blue: MDN prediction.

Table 2. True and False prediction rate table. The total true and false prediction rates of STAR and CDN predictions (bold numbers) and the size of the intersections of the given sets.

		STAR		CDN
		True	False	Total
CDN	True	0.885279	0.053680	**0.938959**
	False	0.015954	0.045087	**0.061041**
STAR	Total	**0.901233**	**0.098767**	

4. DISCUSSION

Considering the comparison of the statistical parameters of the proposed prediction methods and the reference data set (Figure 10 and 12) it can be clearly seen that there is almost no difference in the accuracy, the 90% confidence intervals of the predicted $SI(t+1)$ distribution cover each-other. Thus, both of the suggested neural network based model were able to accurately predict the $SI(t+1)$ distribution from the $SI(t)$ value. These results confirm the applicability of the presented AI methods in the STAR treatment. The study was limited to the statistical evaluation of the prediction results on the given data set. Prior to clinical application of the results extensive validation will be necessary.

The neural network based methods were also compared with the STAR stochastic model currently used in the clinical treatment, see Table 2. It can be seen in this table that the total true rate – i.e. the proportion of cases when the 90% confidence internal defined by the given method – of the neural network is almost 4% higher (93.90% vs. 90.12%) than the true rate of the current STAR prediction. In 5.3% of the cases the classification deep neural network suggests an appropriate confidence interval so that the STAR prediction was not accurate in the given case and only 1.59% of the cases happens it on a revers way. These numbers suggest that the classification deep neural network somewhat better in the prediction accuracy from the aspect of the STAR clinical application. However, the quantitative differences are relatively small

so the accuracy differences should be further investigated by in-silico trials.

5. CONCLUSION

Two neural network based insulin sensitivity prediction methods were presented that can be used in the STAR tight glycaemic control protocol. The suggested methods prediction accuracy was the same or better than the currently used stochastic model accuracy. These results are the initial successful step in the validation process of the proposed methods which will be followed by in-silico simulation trials.

The presented results proved the applicability of the neural network based methods for the patients state prediction in the model based clinical treatment. These methods allow more flexible inclusion of additional patient parameters into the patient state prediction process which is a promising opportunity for better clinical treatment.

ACKNOWLEDGEMENTS

The research was supported by the Hungarian National Scientific Research Foundation, Grant No. K116574, by the BME-Biotechnology FIKP grant of EMMI (BME FIKP-BIO), by the EFOP-3.6.1-16-2016-00017 project, and by H2020 MSCA-RISE DCPM (#872488) grant.

REFERENCES

Ali, N.A., O'Brien Jr, J.M., et al. (2008). Glucose variability and mortality in patients with sepsis. *Critical care medicine*, 36(8), 2316.

Benyo, B., Illyés, A., Némedi, et al. (2012a). 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., Stewart, K., Homlok, J., et al. (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.

Bishop, C.M. (1994). Mixture density networks.

Chase, J.G., Le Compte, A.J., Suhaimi, F., Shaw, et al. (2011). Tight glycemic control in critical care–the leading role of insulin sensitivity and patient variability: a review and model-based analysis. *Computer methods and programs in biomedicine*, 102(2), 156–171.

Chase, J.G., Pretty, C.G., Pfeifer, L., Shaw, G.M., et al. (2010). Organ failure and tight glycemic control in the SPRINT study. *Crit Care*, 14(4), R154.

Dubois, J., Van Herpe, T., et al. (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., et al. (2012). Stochastic targeted (STAR) glycemic control: design, safety, and performance. *Journal of diabetes science and technology*, 6(1), 102–115.

Gulli, A. and Pal, S. (2017). *Deep Learning with Keras*. Packt Publishing Ltd.

Krinsley, J.S. (2004). Effect of an intensive glucose management protocol on the mortality of critically ill adult patients. In *Mayo Clinic Proceedings*, volume 79, 992–1000. Elsevier.

Le Compte, A., Chase, J.G., et al. (2011). Modeling the glucose regulatory system in extreme preterm infants. *Computer methods and programs in biomedicine*, 102(3), 253–266.

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. (2008). Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care. *Computer methods and programs in biomedicine*, 89(2), 141–152.

Lin, J., Razak, N.N., Pretty, C.G., et al. (2011). A physiological Intensive Control Insulin-Nutrition-Glucose (ICING) model validated in critically ill patients. *Computer methods and programs in biomedicine*, 102(2), 192–205.

Lonergan, T., Compte, A.L., et al. (2006). A simple insulin-nutrition protocol for tight glycemic control in critical illness: development and protocol comparison. *Diabetes technology & therapeutics*, 8(2), 191–206.

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

Paláncz, B., Stewart, et al. (2016). Stochastic simulation and parameter estimation of the icing model. *IFAC-PapersOnLine*, 49(5), 218 – 223. 4th IFAC Conference on Intelligent Control and Automation Sciences 2016.

Reed, C.C., Stewart, R.M., Sherman, M., Myers, et al. (2007). Intensive insulin protocol improves glucose control and is associated with a reduction in intensive care unit mortality. *Journal of the American College of Surgeons*, 204(5), 1048–1054.

Schultz, M., Harmsen, R., et al. (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., et al. (2016). Safety, efficacy and clinical generalization of the STAR protocol: a retrospective analysis. *Annals of intensive care*, 6(1),24.

Suhaimi, F., Compte, A.L., et al. (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.

Uyttendaele, V., Knopp, J.L., et al. (2019). 3D kernel-density stochastic model for more personalized glycaemic control: development and in-silico validation. *BioMedical Engineering OnLine*, 18(1), 102.

Uyttendaele, V., Knopp, J.L., et al. (2018). A 3D insulin sensitivity prediction model enables more patient-specific prediction and model-based glycaemic control. *Biomedical Signal Processing and Control*, 46, 192–200.

Van Den Berghe, G., Wouters, P., et al. (2001). Intensive insulin therapy in critically ill patients. *New England journal of medicine*, 345(19), 1359–1367.

Wolfram (2019). Regression with uncertainty - mixture density networks. URL `https://reference.wolfram.com/language/tutorial/NeuralNetworksRegressionWithUncertainty.html`.