34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Personalized Blood Glucose Models for Exercise, Meal and Insulin Interventions in Type 1 Diabetic Children *

Naviyn P. Balakrishnan[1], *Member, IEEE,* Gade P. Rangaiah[2], and Lakshminarayanan Samavedham[3]

Abstract— **Modern healthcare is rapidly evolving towards a personalized, predictive, preventive and participatory approach of treatment to achieve better quality of life (QoL) in patients. Identification of personalized blood glucose (BG) prediction models incorporating the lifestyle interventions can help in devising optimal patient specific exercise, food, and insulin prescriptions, which in turn can prevent the risk of frequent hypoglycemic episodes and other diabetes complications. Hence, we propose a modeling methodology based on multi-input single-output time series models, to develop personalized BG models for 12 type 1 diabetic (T1D) children, using the clinical data from Diabetes Research in Children's Network. The multiple inputs needed to develop the proposed models were rate of perceived exertion (RPE) values (which quantify the exercise intensity), carbohydrate absorption dynamics, basal insulin infusion and bolus insulin absorption kinetics. Linear model classes like Box-Jenkins (1 patient), state space (1 patient) and process transfer function models (7 patients) of different orders were found to be the most suitable as the personalized models for 9 patients, whereas nonlinear Hammerstein-Wiener models of different orders were found to be the personalized models for 3 patients. Hence, inter-patient variability was captured by these models as each patient follows a different personalized model.**

I. Introduction

The medical profession is increasingly discarding a "one size fits all" approach to medical treatment and rapidly embracing a personalized approach to medical treatment and care. This change has been brought about not only due to improvements in measurement technology but also because of an improved understanding of diseases that has been catalyzed by the application of mathematical modeling techniques. In the development of blood glucose (BG) models for diabetic subjects, intra- and inter-patient variability in glucose metabolism necessitates the need for personalization. On the other hand, optimal lifestyle interventions in diabetes care itself can serve as a therapy to prevent the occurrence of several short and long term complications. Particularly, optimal exercise interventions in type 1 diabetes (T1D) children are important to prevent hypoglycemic risks during or immediately following the exercise. A personalized model including the effects of lifestyle interventions can aid in optimal prescription of exercise, insulin and meal in diabetes care.

* Research supported by National University of Singapore.

Authors 1, 2 and 3 are with the National University of Singapore, Department of Chemical & Biomolecular Engineering Singapore 117576.

[1](Phone: +65-84395094; e-mail: naviyn@nus.edu.sg);

[2](e-mail: chegpr@nus.edu.sg).

[3](e-mail: chels@nus.edu.sg).

Mechanistic and empirical BG models have been developed since early 1960s [1]. Most of the mechanistic models predict BG variations for meal and insulin inputs. A review on these mechanistic models can be found in [2, 3]. There are also a few models which incorporate exercise effects [4, 5]. Data based modeling techniques like artificial neural networks (ANN) and time series models have also been used to forecast the BG variations using various inputs like meal, insulin, activity, skin impedance, heart rate etc. A brief overview on such data based models can be seen in [2]. Most of the available exercise related models on T1D quantify the exercise intensity or activity related variables using the percentage of oxygen consumption (VO_{2max}) as a means to quantify the exercise intensity [4, 5] or using various other measurements from sensors like sense wear as in [6]. Exercise intensity can also be quantified by using rate of perceived exertion (RPE) values, which can easily be obtained via simple speech or pictorial tests without depending on any special sensors or devices. The validity of revised scaled RPE values (varying between 0 and 10) in children was studied and confirmed in [7].

Hybrid model structure involving compartmental models for meal and insulin prediction and ANN models for glucose-insulin metabolism prediction was investigated in [8]. Recently, a similar hybrid model structure has been used to develop personalized models for the virtual patients of UVa simulator [9]. However, these studies have not incorporated any exercise effects into the personalized models, and these models have been developed using the virtual data only.

Hence, in this paper, we propose a methodology that involves personalized time series BG models with inputs related to exercise (quantified by 0 to 10 revised scale RPE values), meal and insulin (basal and bolus). Clinical data of T1D children from [10] were utilized for developing personalized BG models.

II. Nature of Dataset and Modeling Methodology

A. Nature of T1D Children Dataset

In this work, we used clinical data obtained from one of the Diabetes Research in Children's Network (DIRECNET) exercise studies. There are totally 55 T1D children in this cohort, out of which the clinical data of 12 randomly-chosen patients were presented in this work. The dataset includes the clinical data of two outpatient visits lasting about 7 hr with a 75 min exercise in the late afternoon. In one of the visits, the basal insulin supply was stopped during the exercise period whereas it was continued during the other visit. In this DIRECNET study [10], the patients visited the hospital prior

to lunch and left after dinner. Pre-lunch bolus of rapid acting insulin analogue was administered to each patient based on the usual insulin-carbohydrate (CHO) ratio and correction factor calculations performed at home. The late afternoon exercise was performed on a treadmill in 4 phases (P1, P2, P3, and P4) at a target heart rate of 140 bpm. Each phase consisted of 15 min exercise followed by 5 min resting period. The intensity of exercise during each phase was quantified based on 0 to 10 revised scale RPE values. A statistical summary of the clinical data used in this work can be found in Table I. Also, the mean ± standard deviation (SD) plots of CGMS BG readings (sampling interval = 5 min) during the two outpatient visits can be seen in Fig. 1.

B. Modeling Methodology

The system comprises of 4 input variables related to exercise (RPE), meal (CHO), and insulin (basal and bolus doses) and a single output (CGMS BG readings). A hybrid multi-input single-output (MISO) model structure (illustrated in Fig. 2), which involves time series models predicting BG dynamics and mechanistic models characterizing the input dynamics (like meal absorption dynamics and subcutaneous (S.C.) insulin kinetics) is proposed in this work. The type and order of the BG time series models was not the same for all patients due to the inter-patient variability in glucose-insulin metabolism. Hence, they are called as personalized BG models. On the other hand, the type of the mechanistic models used in characterizing the input dynamics was not varied for different patients. Hence, they are termed as generalized models. The clinical data related to the inputs were pre-treated and converted into time series data before being used as inputs for the purpose of personalized model identification.

1) Pretreatment of RPE values:

The clinical dataset contains 0-10 revised scale RPE values measured at 10^{th} min of each exercise phase. The RPE values at the 5^{th} and 15^{th} min of each exercise phase are missing in the current dataset. Hence, in order to predict the missing values, a regression model was developed with the available data. Regression was performed with RPE values as output (*Y*) and gender (*X1*), height (*X2*), weight (*X3*), exercise phase (*X4*), basal insulin status during exercise (*X5*), treadmill speed (*X6*) and inclination (*X7*) as input variables. Interaction between treadmill speed and inclination (*X8* = *X6***X7*) was also considered as one of the input variables. The data of each exercise phase was considered as a sample. Different combinations of testing and training datasets were used for developing the regression model. The final regression model used in the prediction of missing RPE values is summarized in Table II. Variables *X1*, *X3*, *X6*, and *X8* were found to be statistically significant. The negative coefficient value for *X6* indicates that increase in speed alone will not increase the RPE, whereas the positive coefficient for the interaction variable *X8* indicates that simultaneous increase/decrease of both speed and inclination leads to increase/decrease in RPE. The results of regression model are not discussed in detail here due to scope of the article. The missing RPE values were predicted by using the available treadmill speed and inclination data into the model.

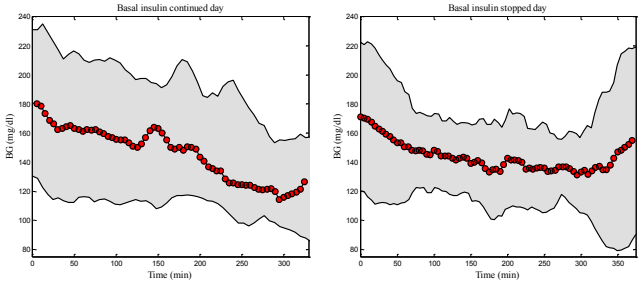

Fig. 1 Mean ± SD values of CGMS BG readings of 12 patients

2) Pretreatment of meal data:

Meal data includes the lunch and snacks given till dinner. The amount of CHO given for every snack is available in the clinical dataset. In case of lunch, the amount of CHO was calculated using the insulin–CHO ratio and the usual pre-lunch bolus dose. Hence, the meal data includes the time and amount of CHO in each intake. In order to convert this data into dynamic (thereby making it suitable for time series analysis), the meal absorption model from Hovorka et al. was used [11]. The mathematical representation of this meal absorption dynamics model (U_G) can be found in equation (4) of [11].

3) Pretreatment of insulin data:

Insulin data involves the basal and bolus insulin doses. The basal insulin dose (U_B) was computed by calculating the total daily dose (*TDD*) using the correction ratio (*TDD* = 1800 ÷ correction ratio) and by subtracting the total bolus dose (pre-breakfast, lunch, dinner and bedtime snack doses) from *TDD*. The pre-lunch bolus dose involves the sum of usual bolus dose and correction dose computed on the visit day. The correction dose was calculated based on the available clinical data like pre-lunch BG level correction ratio and pre-meal BG target (correction factor). The subcutaneous insulin absorption kinetics (U_{SCI}) was obtained by using equation (14) in [12].

TABLE I. SUMMARY STATISTICS OF IMPORTANT VARIABLES (N = 12)

Variable	Mean Value	SD Value	Unit
Height	171.23	13.3	cm
Weight	61.96	11.89	kg
Insulin-CHO Ratio (lunch)	10.61	3.1	no unit
Usual Pre-lunch Insulin Dose	6.74	3.2	U
Correction Ratio	49.9	15.8	mg/dl
Pre-meal Target BG (Correction Factor)	128.7	20.32	mg/dl
RPE P1(Visit 1 / Visit 2)	3.1 / 3.9	1.75 / 1.4	no unit
RPE P2 (Visit 1 / Visit 2)	3.9/ 3.6	1.4 / 1.04	no unit
RPE P3 (Visit 1 / Visit 2)	3.8 / 3.7	1.1 / 0.99	no unit
RPE P4 (Visit 1 / Visit 2)	4.1/ 4.2	1.04 / 0.99	no unit
Treadmill Speed (Visit 1 / Visit 2)[a]	3.48 / 3.5	0.33 / 0.28	speed units
Treadmill Inclination (Visit 1 / Visit 2)[b]	6.9 / 6.96	2.7 / 2.6	inclination units

a, b Only overall mean and SD values of the 4 exercise phases in 12 patients are provided for brevity

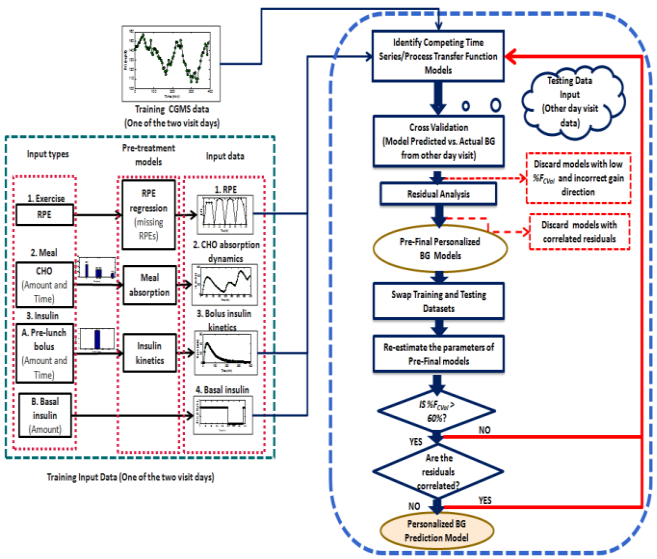

Fig. 2 Proposed hybrid modeling methodology

The pre-treated input data (such as shown in Fig.3) along with CGMS data were used to identify the personalized time series models. The MISO data were preprocessed by removing the means. The system identification tool box in MATLAB was used for model identification purpose. Initially, the data obtained during one of the two visits were used for training and the other day visit data were employed for testing (cross validation). Using the initial training and test datasets, pre-final personalized BG models for each patient were selected from a group of competing models based on cross validation plots (of validation data vs. model prediction values), cross validation percentage fitness values ($\%F_{CVal}$), and residual analysis. Following this, the training and testing datasets of the initial datasets were swapped and the parameters of the pre-final models were re-estimated. If a pre-final model captured important BG trends with $\%F_{CVal}$ > 60% (a reasonable percentage for noisy clinical/biological datasets) after swapping, then that model was selected as the personalized model for a particular patient. $\%F_{CVal}$ values were calculated by:

$$\%F_{CVal} = \left\{ 1 - \frac{\sum_{t=1}^{n} \left((BG_{t,Val})_{scaled} - (BG_{t,Pred})_{scaled} \right)^2}{\sum_{t=1}^{n} \left((BG_{t,Val})_{scaled} \right)^2} \right\} \times 100\% \quad (1)$$

Here, $(BG_{t,Val})_{scaled}$ and $(BG_{t,Pred})_{scaled}$ represent the scaled validation and model predicted BG values at time, t.

TABLE II. FINAL RPE REGRESSION MODEL USED FOR PREDICTION OF MISSING RPE VALUES IN THE DIRECNET'S CLINICAL DATABASE

Variable	Coefficient value	t-stat value	p-value
Gender (X1)	-2.11	-6.742	<0.0001
Height (X2)	0.095	2.02	0.0294
Weight (X3)	-0.38	-5.69	<0.0001
Treadmill speed (X6)	-3.31	-3.78	0.016
*[(X6) * inclination] (X7)*	0.332	5.92	<0.0001
Intercept	16.14		
R-squared/Adjusted R-squared	0.69 / 0.64		

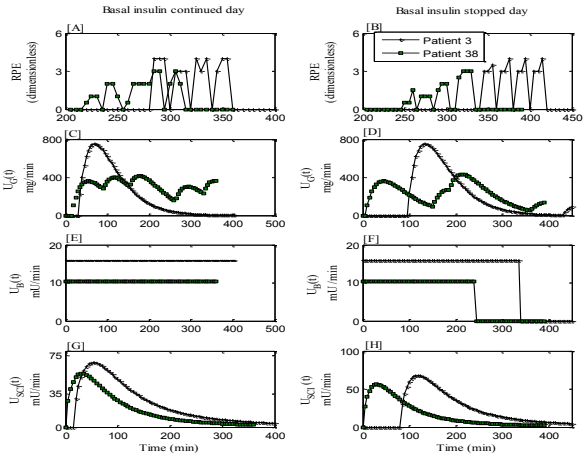

Fig. 3 Pretreated dynamic input profiles obtained for two outpatient visits.

III. RESULTS AND DISCUSSION

The identified personalized model for each patient along with the $\%F_{CVal}$ obtained using initial and swapped datasets can be found in Table III. It can be inferred that the personalized models identified for 9 (out of 12) patients showed a high $\%F_{CVal}$ (>70%) in both initial and swapped datasets. Also, each patient followed a different personalized model (different model class or order), indicating that the inter-patient variability was captured well. After swapping and re-estimating the parameters of pre-final models, there was about 0.03-12% decrease in $\%F_{CVal}$ of 9 patients and about 2-7% increase in $\%F_{CVal}$ of 3 patients. In case of Patient 47, the personalized model structure along with the parameter values estimated before and after swapping is summarized in Table IV. The directions of the gains were correct for both the initial and swapped dataset cases. Also, swapping and parameter re-estimation resulted in almost similar magnitudes of gains (k_i) and exactly the same time delay (T_{di}) as that of the initial model (which can be observed in bold text numbers of Table IV).

TABLE III. IDENTIFIED PERSONALIZED MODELS WITH $\%F_{CVAL}$ VALUES

Pt id.	Input delay (samples)				Personaliz--ed model	% F_{CVal} [a]	
	1	2	3	4		Before swapping	After swapping
3	1	2	15	10	P3DZ[b]	88.35	80.76
47	1	10	14	7	P3DZ[b]	84.2	86.6
14	29	1	1	1	P1D[b]	72	75
19	15	3	7	2	State space	86.2	86.17
22	1	11	19	6	BJ	78.3	71.34
26	1	8	2	1	HW[c]	73.4	65.5
29	1	21	3	21	P2DZ[b]	77.3	75.7
32	1	2	19	2	P2DZ[b]	84	73
38	1	3	14	1	HW[c]	72.5	79.2
44	1	27	7	14	P2DZ[b]	71.5	60.5
6	1	19	18	8	P2DZ[b]	71	62
54	1	7	3	5	HW[c]	79.8	71.31

a. % fitness is based on the validation trends; b. Process transfer function model with P-Poles, D-Delay, and Z-Zero;
c. Hammerstein-Wiener (HW);
d. Box Jenkins (BJ);
Inputs 1, 2, 3, 4 denote U_B, U_G, U_{SCI}, *RPE*, respectively

TABLE IV. PERSONALIZED MODEL STRUCTURE AND PARAMETER VALUES OF PATIENT 47 (BEFORE AND AFTER DATASET SWAPPING)

Personalized model: P3DZ

$$BG(t) = G_1(s)u_1 + G_2(s)u_2 + G_3(s)u_3 + G_4(s)u_4$$

$$G_i(s) = \frac{k_i(1-T_{zi}s)}{(1+T_{p1i}s)(1+T_{p2i}s)(1+T_{p3i}s)} * exp(-T_{di}s)$$

where $i = 1,2,3,4$

Case	Parameter values before swapping				Parameter values after swapping			
i	*1*	2	*3*	4	1	2	3	4
k_i (mg/dl)	**-0.01**	**1.5**	**-1.7**	**-1.8**	**-0.01**	**1.2**	**-1.9**	**-1.8**
T_{zi} (min)	-0.01	9.9	-8.6	-5.4	-0.02	7.6	-8.9	-6
T_{di} (min)	**5**	**50**	**70**	**35**	**5**	**50**	**70**	**35**
T_{p1i} (min)	48.1	297	41	37.2	42.5	221	58	26
T_{p2i} (min)	380	19.8	1.9	115	400	19	1.2	132
T_{p3i} (min)	3.25	19	251	188	2.6	36	298	131

Note: k_i=Gain; T_{di}=delay; T_{zi}=zeros; T_{p1i}, T_{p2i}, T_{p3i}-Poles;

In case of patient 3, before swapping the training and test datasets, Box-Jenkins or BJ model and process transfer function model (P3DZ) were found to closely mimic the patient's BG dynamics with $\%F_{CVal}$ of 78.4% and 88.4%, respectively (see Fig. 4(A)). Also, Hammerstein-Wiener (HW) model showed $\%F_{CVal}$ of about 17% with the initial training and testing datasets. However, after swapping the training and test datasets, HW model showed a much higher $\%F_{CVal}$ of about 90%, followed by P3DZ and BJ models with $\%F_{CVal}$ of 80.8% and 63.4%, respectively (see Fig. 4(B)). In this case, HW cannot be chosen as a personalized model owing to its inconsistency in BG prediction for initial and swapped datasets. Although both P3DZ and BJ models showed about 7% and 15% decrease in $\%F_{CVal}$ for the swapped data, P3DZ captured most of the crests and troughs in the BG trend correctly (with $\%F_{CVal}$ of 81%), except for BG values between 220 and 230 minutes as seen in Fig. 4(B)). Hence, P3DZ model was selected as the personalized model for patient 3 by discarding its counter parts.

In case of patient 38, before swapping the datasets, autoregressive moving average exogenous input (ARMAX), ARX and HW models were identified as pre-final personalized BG models with $\%F_{CVal}$ of 55%, 52% and 72.5%, respectively. Unlike patient 3, $\%F_{CVal}$ of these three competing models increased by 13% (ARMAX), 11% (ARX) and 9% (HW) after swapping and re-estimation. From $\%F_{CVal}$ values and cross validation plots (Figs. 4(c) and (d)), the nonlinear HW model was selected as the personalized model. Input nonlinearities for meal and bolus insulin absorption dynamics and the output nonlinearity for BG dynamics were estimated using piecewise linear functions. The poles and zeros+1 values for the linear block in HW model were [3 4 4 4] and [1 1 1 1], respectively.

In a similar way, the personalized models were identified for other patients. Linear models like BJ and state-space models were identified as personalized models for patients 22 and 19, respectively. For brevity, results of some patients are shown and discussed here. Detailed discussion on the personalized models for all 55 patients will be presented in future.

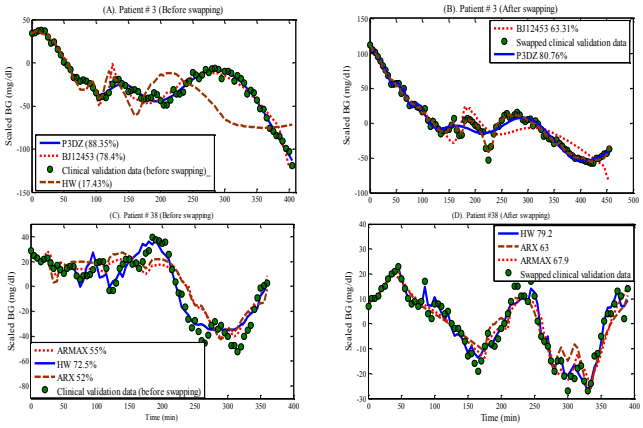

Fig. 4 Measured CGMS BG vs. model predicted BG profiles before and after swapping the training and test data for patients 3 and 38

IV. CONCLUSION

This work is a first step towards a giant leap of achieving personalized, predictive, preventive and participatory medicine (P4) in diabetes care. The predictive nature of the personalized models developed in this work can be utilized to develop personalized prescriptions for exercise, meal, and insulin injection.

ACKNOWLEDGMENT

The authors thank the DIRECNET for publicly sharing the clinical dataset and their timely email correspondence to queries on the dataset. Also, N.P.B thanks NUS for providing financial support for pursuing doctoral studies.

REFERENCES

[1] V.W. Bolie, "Coefficients of normal blood glucose regulation," *J Appl Physiol*, vol. 16, (no. 5), pp. 783-788, September 1, 1961 1961.

[2] N.P. Balakrishnan, G.P. Rangaiah, and L. Samavedham, "Review and Analysis of Blood Glucose (BG) Models for Type 1 Diabetic Patients," *Industrial & Engineering Chemistry Research*, vol. 50, pp. 12041–12066, 2011.

[3] C. Cobelli, C. Dalla Man, G. Sparacino, L. Magni, G. De Nicolao, and B.P. Kovatchev, "Diabetes: Models, Signals, and Control," *Biomedical Engineering, IEEE Reviews in*, vol. 2, pp. 54-96, 2009.

[4] J. Kim, G.M. Saidel, and M.E. Cabrera, "Multi-scale computational model of fuel homeostasis during exercise: Effect of hormonal control," *Annals of Biomedical Engineering*, vol. 35, (no. 1), pp. 69-90, Jan 2007.

[5] A. Roy and R.S. Parker, "Dynamic modeling of exercise effects on plasma glucose and insulin levels," *Journal of diabetes science and technology*, vol. 1, (no. 3), pp. 338-347, 2007.

[6] D.K. Rollins, N. Bhandari, J. Kleinedler, K. Kotz, A. Strohbehn, L. Boland, M. Murphy, D. Andre, N. Vyas, G. Welk, and W.E. Franke, "Free-living inferential modeling of blood glucose level using only noninvasive inputs," *Journal of Process Control*, vol. 20, (no. 1), pp. 95-107, 2010.

[7] A.C. Utter, R.J. Robertson, D.C. Nieman, and J. Kang, "Children's OMNI Scale of Perceived Exertion: walking/running evaluation," *Medicine & Science in Sports & Exercise*, vol. 34, (no. 1), pp. 139-144, 2002.

[8] S.G. Mougiakakou, A. Prountzou, D. Iliopoulou, K.S. Nikita, A. Vazeou, and C.S. Bartsocas, "Neural Network based Glucose - Insulin Metabolism Models for Children with Type 1 Diabetes," in Book Neural Network based Glucose - Insulin Metabolism Models for Children with Type 1 Diabetes, *Series Neural Network based Glucose - Insulin Metabolism Models for Children with Type 1 Diabetes*, Editor ed.^eds., City, 2006, pp. 3545-3548.

[9] K. Zarkogianni, A. Vazeou, S.G. Mougiakakou, A. Prountzou, and K.S. Nikita, "An Insulin Infusion Advisory System Based on Autotuning Nonlinear Model-Predictive Control," *Biomedical Engineering, IEEE Transactions on*, vol. 58, (no. 9), pp. 2467-2477, 2011.

[10] "Diabetes Research in Children Network (DirecNet) public datasets," in:http://public.direc.net/. [cited March 2012] Accession no. Accession Number| Available rom Database Provider|.

[11] R. Hovorka, V. Canonico, L.J. Chassin, U. Haueter, M. Massi-Benedetti, M. Federici, T.R. Pieber, H.C. Schaller, L. Schaupp, T. Vering, and M.E. Wilinska, "Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes," *Physiological Measurement*, vol. 25, (no. 4), pp. 905, 2004.

[12] G. Nucci and C. Cobelli, "Models of subcutaneous insulin kinetics. A critical review," *Computer Methods and Programs in Biomedicine*, vol. 62, (no. 3), pp. 249-257, 2000.