

# Universal Glucose Models for Predicting Subcutaneous Glucose Concentration in Humans

Adiwinata Gani, Andrei V. Gribok, Yinghui Lu, W. Kenneth Ward, Robert A. Vigersky, and Jaques Reifman

**Abstract**—This paper tests the hypothesis that a “universal,” data-driven model can be developed based on glucose data from one diabetic subject, and subsequently applied to predict subcutaneous glucose concentrations of other subjects, even of those with different types of diabetes. We employed three separate studies, each utilizing a different continuous glucose monitoring (CGM) device, to verify the model’s universality. Two out of the three studies involved subjects with type 1 diabetes and the other one with type 2 diabetes. We first filtered the subcutaneous glucose concentration data by imposing constraints on their rate of change. Then, using the filtered data, we developed data-driven autoregressive models of order 30, and used them to make short-term, 30-min-ahead glucose-concentration predictions. We used same-subject model predictions as a reference for comparisons against cross-subject and cross-study model predictions, which were evaluated using the root-mean-squared error (RMSE) and Clarke error grid analysis (EGA). We found that, for each studied subject, the average cross-subject and cross-study RMSEs of the predictions were small and indistinguishable from those obtained with the same-subject models. These observations were corroborated by EGA, where better than 99.0% of the paired sensor-predicted glucose concentrations lay in the clinically acceptable zone A. In addition, the predictive capability of the models was found not to be affected by diabetes type, subject age, CGM device, and interindividual differences. We conclude that it is feasible to develop universal glucose models that allow for clinical use of predictive algorithms and CGM devices for proactive therapy of diabetic patients.

**Index Terms**—Diabetes, glucose prediction, glucose regulation, inverse problems, mathematical model, regularization, system identification.

## I. INTRODUCTION

**M**INIMALLY invasive continuous glucose monitoring (CGM) devices are becoming the new state-of-the-art

Manuscript received March 13, 2009; revised July 14, 2009. First published October 23, 2009; current version published January 15, 2010. This work was supported by the U.S. Army Medical Department, Advanced Medical Technology Initiative, funded by the Telemedicine and Advanced Technology Research Center of the U.S. Army Medical Research and Materiel Command, Fort Detrick, MD.

A. Gani was with the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702 USA. He is now with Asia Pulp and Paper Sinarmas Group, Jakarta, Indonesia (e-mail: adiwinata.gani@gmail.com).

A. V. Gribok is with the Department of Nuclear Engineering, University of Tennessee, Knoxville, TN 37996 USA (e-mail: agribok@bioanalysis.org).

Y. Lu and J. Reifman are with the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702 USA (e-mail: ylu@bioanalysis.org; jaques.reifman@us.army.mil).

W. K. Ward is with iSense Corporation and Oregon Health and Sciences University, Wilsonville, OR 97070 USA (e-mail: kenward503@msn.com).

R. A. Vigersky is with the Diabetes Institute, Walter Reed Army Medical Center, Washington, DC 20307 USA (e-mail: Robert.Vigersky@amedd.army.mil).

Digital Object Identifier 10.1109/TITB.2009.2034141

standard to measure and record a patient’s glycemic state as frequently as every minute [1]. This information can be utilized by patients to alter or improve their lifestyle, to tighten their glycemic control, or alternatively, by a clinician to adjust therapy. These frequent measurements can also be used by data-driven models to forecast future values of subcutaneous glucose concentration and avoid undesired hypo- and hyperglycemic episodes [1]–[4].

In contrast to intermittent measurements, CGM devices collect information so frequently that consecutive measurements retain a large degree of temporal correlation. This correlation is exploited by data-driven models to infer future values as a function of previous measurements [2]–[4]. However, because of the availability of glucose signals at high sampling rates, developers of data-driven models implicitly assume that the models need to be tuned for a specific individual, thus increasing the burden of model development and reducing their practical applicability.

For example, Sparacino *et al.* [3] use an autoregressive (AR) model of order one, AR(1), which continuously adapts the model coefficient to predict glucose concentrations up to 30 min ahead. Although such a model can produce acceptable predictions, it needs to be continuously adapted for every individual. Additionally, in spite of the adaptive nature of the model, it introduces a significant delay between predicted and measured values. This delay is caused by the low order of the AR model, because a single AR coefficient is not sufficient to capture the temporal variations of the time-series glucose signal. Separately, Dua *et al.* [5] employ a Kalman filter to predict blood glucose by continuously adjusting parameters of a first-principle model. Although the first-principle model is significantly more flexible than the AR(1) model, the continuous adaptation also makes this model individual-specific.

Alternatively, in our recent work [4], we report an AR(30) model with fixed coefficients, which successfully predicts the glucose level of patients with type 1 diabetes. The AR(30) model is able to yield accurate 30-min-ahead glucose-concentration predictions with an average root-mean-squared error (RMSE) of 0.10 mmol/L (SD = 0.02) and a negligible prediction time lag of 0.2 min (SD = 0.4). Moreover, we observe that, due to the robust nature of our algorithm with respect to the modeling parameters, the proposed approach yields similar AR coefficients for different individuals, suggesting the potential existence of “universal” (or “portable”) models, which could be applied without modification to different individuals.

Here, we hypothesize that the short-term (30 min or less) dynamics of glucose regulation is quite similar for different diabetic individuals, where a single, universal AR model can be used to make predictions across different patients. We

TABLE I  
LIST OF THE THREE CGM DEVICES AND CORRESPONDING STUDIES

CGM Device	iSense	Guardian RT	DexCom
Manufacturer	iSense Corp.	Medtronic Inc.	DexCom Inc.
# of Subjects	9	18	7
Diabetes Type	1	1	2
Sampling Interval (minutes)	1	5	5
Collection Time (days)	5	6	56

tested this hypothesis with data collected from three different studies, involving subjects with both type 1 and 2 diabetes, and using three different CGM devices: iSense (iSense Corporation, Wilsonville, OR) [6], Guardian RT (Medtronic, Inc., Northridge, CA) [7], and DexCom (DexCom, Inc., San Diego, CA) [8].<sup>1</sup> We developed data-driven AR models of a fixed order for each subject, as described in Gani *et al.* [4], and tested them on data from other subjects from the same and from different studies. We used the RMSE, prediction time lag, and Clarke error grid analysis (EGA) as metrics to quantify the models' performance, and compared the resulting AR coefficients from the different models developed for each subject.

Our results confirmed that the developed AR models (i.e., the AR model coefficients) were not significantly dependent on a given individual, diabetes type, or CGM device. This suggests the feasibility of universal, individual-independent predictive models, which can significantly reduce the burden of model development as one model can be used to predict any one individual, and opens the door for clinical use of predictive glucose models together with CGM devices for proactive regulatory therapy.

## II. METHODS

### A. Subject Selection and CGM Measurements

Deidentified data for this investigation were obtained from three different independent studies using three different CGM systems: iSense [6], Guardian RT [7], and DexCom [8]. A total of 34 subjects gave their voluntary and written informed consent to participate in these studies, which had received approval by the appropriate institutional review boards. Table I summarizes the three studies.

1) *iSense Dataset*: The purpose of this study was to evaluate the accuracy of the iSense CGM system [6]. Nine subjects were confined to the investigational site for the whole duration of the study and limited to mild physical activity. Subjects were included if they were between 18 and 70 years of age, had been diagnosed with type 1 diabetes and treated with insulin for at least 12 months, had body mass index  $<35.0$  kg/m<sup>2</sup>, and had glycated hemoglobin (HbA1c)  $>6.1\%$ . Subjects were excluded if they had acute and severe illness apart from diabetes, clinically significant abnormal electrocardiogram, hematology or biochemistry screening test, or any disease requiring use of anticoagulants. In addition, subjects were excluded if they were

pregnant or lactating. Subcutaneous glucose measurements were collected on a minute-by-minute basis for each of the nine subjects for approximately five days with the iSense CGM system. This minute-by-minute dataset was also used in our previous work [4], where we describe the AR modeling framework applied here. However, to standardize the sampling rate across studies, we downsampled the data to 5-min sampling intervals. The 5-min sampling interval is half the "optimal" sampling interval (10 min) recommended in the literature [9]. Additional information about this independent study can also be found in the work by Reifman *et al.* [10].

2) *Guardian RT Dataset*: The dataset with the Guardian RT [7] CGM system was retrieved from the Diabetes Research in Children Network (DirecNet) Website [11], which makes continuous glucose data for six different studies involving children with type 1 diabetes publicly available, along with the corresponding protocols. For this paper, we used the DirecNet study entitled "A pilot study to evaluate the navigator continuous glucose sensor in the management of type 1 diabetes in children," which included 30 subjects. Subjects were included if they were between 3 and 7 years old or between 12 and 18 years old, had been diagnosed with type 1 diabetes for more than one year, had been using insulin pump, and had HbA1c  $\leq 10.0\%$ . Subjects were excluded if they had significant medical disorder, had a severe hypoglycemic event resulting in seizure or loss of consciousness in the last month, had used systemic or inhaled corticosteroids in the last month, or had cystic fibrosis. Subjects were provided with the Guardian RT CGM system for home usage, which collected subcutaneous glucose concentration every 5 min for six days. For the analysis in this paper, 12 subjects, out of the total 30, were excluded because they did not possess any consecutive 4000-min segment (i.e., 800 data points) without data gaps.

3) *DexCom Dataset*: The purpose of the study was to investigate the short- and long-term effectiveness and benefits of frequent CGM measurements (versus the infrequent, only before each meal and at bedtime, fingerstick blood glucose measurements). Seven subjects were studied and included on an ongoing investigation from an independent study by R. A. Vigersky. Subjects were included if they were older than 18 years of age, had been diagnosed with type 2 diabetes for at least three months and treated with insulin, and had HbA1c between 7% and 12%. Subjects were excluded if they had been taking glucocorticoids, amphetamines, anabolic, or weight-reducing agents. In addition, subjects were also excluded if they were pregnant, lactating, or planning to become pregnant. For this study, subjects continued to take all medications that had been prescribed for diabetes and other medical conditions, and followed their usual meal plans and activity schedules. Investigators of the study did not make any recommendations to the subjects regarding medications, weight, diet, or exercise at any time during the study. Subjects were instructed to contact their primary care provider for all treatment decisions and consultations. Subcutaneous glucose measurements with the DexCom CGM system [8] were collected every 5 min for each of the seven subjects for approximately eight weeks on 4 two-week cycles. A complete protocol for this study is available online [8].

<sup>1</sup>Protocol for the study is available from <http://clinicaltrials.gov/ct2/show/NCT00529815?term=continuous+glucose+moni>

The reason for selecting three different studies was to test the hypothesis that universal models can be developed and applied across different individuals, with different types of diabetes, employing distinct CGM devices, i.e., we intended to test the invariance of the dynamics of the glucose signal in diabetic patients and the ability of AR models to capture it.

### B. Predictive Algorithm

We applied a previously developed methodology [4] to obtain glucose models for each one of the 34 subjects and predict their glucose concentrations 30 min ahead of time. We used the first 4000 min (i.e., 800 data points with a 5-min sampling interval) of available CGM data of each subject and filtered the data to remove high-frequency noise. The filtering constrained the glucose rate of change to ensure that the first-order time derivative of the glucose signal was consistent with clinically observed values reported by Kovatchev *et al.* [12], i.e.,  $\pm 0.2$  mmol/(L·min) [ $\pm 4$  mg/(dL·min)], while avoiding the introduction of time lags between the filtered and the original CGM signals.

We used the first half of each subject's filtered data to develop an AR model for each subject. An AR model is a type of linear model that infers a future signal  $\hat{y}_n$ , at time  $n$  ( $n = m + 1, \dots, N$ , where  $N$  denotes the total number of data samples available for modeling), based on a linear combination of antecedent samples  $\tilde{y}_{n-i}$  weighted by a fixed set of coefficients  $b_i$

$$\hat{y}_n = \sum_{i=1}^m b_i \tilde{y}_{n-i} \quad (1)$$

where  $m$  denotes the order of the model, i.e., the number of previously observed and filtered glucose concentrations  $\tilde{y}_{n-i}$  used to predict a future glucose concentration  $\hat{y}_n$ . This fixed set of coefficients  $b_i, i = 1, 2, \dots, m$ , which defines a model of order  $m$ , describes the correlations in the signal. We calculated these coefficients by the method of least squares [13] with an added smoothness constraint to insure physiologic plausibility of the obtained coefficients [4].

Once these coefficients were inferred, the model was subsequently used for predicting glucose concentrations [4], where model performance was quantified by computing prediction time lags and RMSEs, as well as by performing Clarke EGA. The RMSE is defined as the square root of the mean of the squared differences between the predicted signal  $\hat{y}_i$  and the filtered observed signal  $\tilde{y}_i, i = 1, 2, \dots, N$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - \tilde{y}_i)^2} \quad (2)$$

and the prediction time lag is calculated based on the cross-correlation between the filtered and predicted signals. The lag, characterized by the peak of the cross-correlation function, provides an accurate estimate of the delay in the predictions [14].

### C. Validation Scenarios

To test the hypothesis that we can develop one model, based on one subject's signals, and universally apply this model to other diabetic subjects, we examined three scenarios using 34

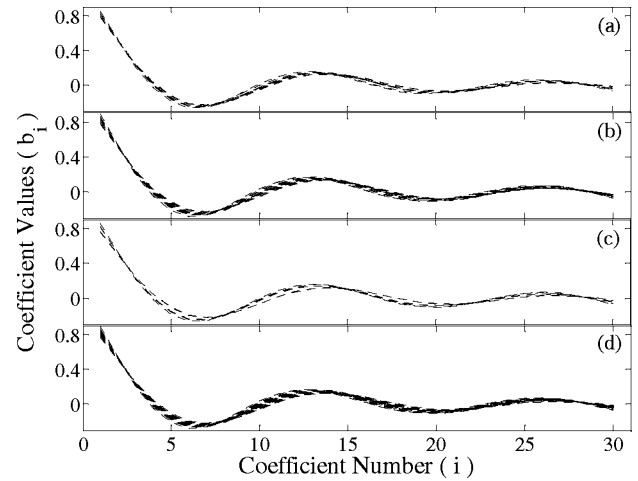


Fig. 1. Coefficients of the AR models for (a) the nine iSense subjects, (b) the 18 Guardian RT subjects, (c) the seven DexCom subjects, and (d) the combined 34 subjects for all three studies.

subjects from the three studies. We used the first 2000 min of the filtered signals of each subject to train the AR models (training dataset) and the next 2000 min to test the predictions (testing dataset). The three scenarios were created to allow for the comparison of model performance on the same testing datasets by applying distinct models derived from different training datasets as follows.

- 1) *Scenario I [Same-subject model (same subject, same CGM device)]*: For each of the 34 subjects, a model was trained on each subject's training dataset, resulting in 34 different models. Each model was tested on that subject's testing dataset. For example, the Guardian RT training data for subject 5 were used to derive a model for that subject, which was subsequently used to predict that subject's testing data.
- 2) *Scenario II [Cross-subject model (different subjects, same CGM device)]*: For each subject within a given study, the models developed in scenario I for the remaining subjects of that same study were applied to the testing dataset of the subject. For example, the models developed for the remaining 17 subjects in the Guardian RT study were applied to the testing dataset of subject 5 of the *same* Guardian RT study.
- 3) *Scenario III [Cross-study model (different subjects, different CGM devices)]*: Similarly, for each subject within a given study, the models developed in scenario I for the subjects in the other two studies were applied to the testing dataset of the subject. For example, the models developed for the nine subjects in the iSense study and the seven subjects in the DexCom study were applied to the testing dataset of subject 5 of the Guardian RT study.

## III. RESULTS

Fig. 1 shows the AR model coefficients  $b_i, i = 1, 2, \dots, 30$ , for nine iSense subjects (a), 18 Guardian RT subjects (b), seven DexCom subjects (c), and the combined 34 subjects for all three studies (d). Fig. 1(d) shows that the coefficients  $b_i$ , and hence

TABLE II  
MEAN AND STANDARD DEVIATION (SD) OF THE 30 AR MODEL COEFFICIENTS  
FOR EACH OF THE THREE STUDIES

AR Coeff.	iSense		Guardian		DexCom	
$b_i$	Mean	(SD)	Mean	(SD)	Mean	(SD)
1	0.8123	(0.0246)	0.8271	(0.0314)	0.8039	(0.0342)
2	0.5135	(0.0069)	0.5176	(0.0086)	0.5103	(0.0100)
3	0.2375	(0.0101)	0.2324	(0.0118)	0.2387	(0.0109)
4	0.0108	(0.0194)	-0.0003	(0.0235)	0.0148	(0.0245)
5	-0.1470	(0.0216)	-0.1602	(0.0263)	-0.1421	(0.0291)
6	-0.2289	(0.0173)	-0.2402	(0.0209)	-0.2247	(0.0250)
7	-0.2401	(0.0093)	-0.2465	(0.0106)	-0.2377	(0.0144)
8	-0.1960	(0.0075)	-0.1960	(0.0072)	-0.1956	(0.0024)
9	-0.1178	(0.0153)	-0.1115	(0.0168)	-0.1190	(0.0132)
10	-0.0283	(0.0212)	-0.0171	(0.0236)	-0.0302	(0.0231)
11	0.0526	(0.0226)	0.0663	(0.0248)	0.0510	(0.0277)
12	0.1104	(0.0193)	0.1237	(0.0207)	0.1099	(0.0263)
13	0.1375	(0.0129)	0.1479	(0.0131)	0.1383	(0.0198)
14	0.1335	(0.0074)	0.1389	(0.0072)	0.1355	(0.0099)
15	0.1039	(0.0106)	0.1035	(0.0115)	0.1063	(0.0037)
16	0.0582	(0.0165)	0.0522	(0.0174)	0.0601	(0.0127)
17	0.0077	(0.0200)	-0.0026	(0.0202)	0.0083	(0.0198)
18	-0.0369	(0.0200)	-0.0494	(0.0190)	-0.0381	(0.0230)
19	-0.0674	(0.0168)	-0.0798	(0.0148)	-0.0705	(0.0216)
20	-0.0795	(0.0116)	-0.0894	(0.0098)	-0.0839	(0.0165)
21	-0.0729	(0.0073)	-0.0786	(0.0077)	-0.0776	(0.0088)
22	-0.0515	(0.0085)	-0.0520	(0.0100)	-0.0552	(0.0029)
23	-0.0216	(0.0123)	-0.0171	(0.0123)	-0.0233	(0.0091)
24	0.0089	(0.0148)	0.0172	(0.0130)	0.0098	(0.0146)
25	0.0325	(0.0150)	0.0427	(0.0121)	0.0360	(0.0170)
26	0.0439	(0.0128)	0.0534	(0.0102)	0.0488	(0.0155)
27	0.0404	(0.0085)	0.0471	(0.0075)	0.0452	(0.0105)
28	0.0231	(0.0037)	0.0250	(0.0047)	0.0260	(0.0027)
29	-0.0039	(0.0077)	-0.0082	(0.0066)	-0.0047	(0.0074)
30	-0.0354	(0.0161)	-0.0463	(0.0130)	-0.0406	(0.0179)

the glucose models, did not vary significantly from subject to subject and from study to study, i.e., they were independent of a subject's age, diabetes type, and CGM device used to measure the glucose concentration. The figure also validated our hypothesis that our methodology yields portable glucose models. Table II provides additional information to assess the similarity of the AR coefficients  $b_i$  across the three studies. The columns in the table show the mean values (and standard deviations) of each of the 30 AR coefficients for each study, which were obtained by averaging each coefficient  $b_i$  across the models in each

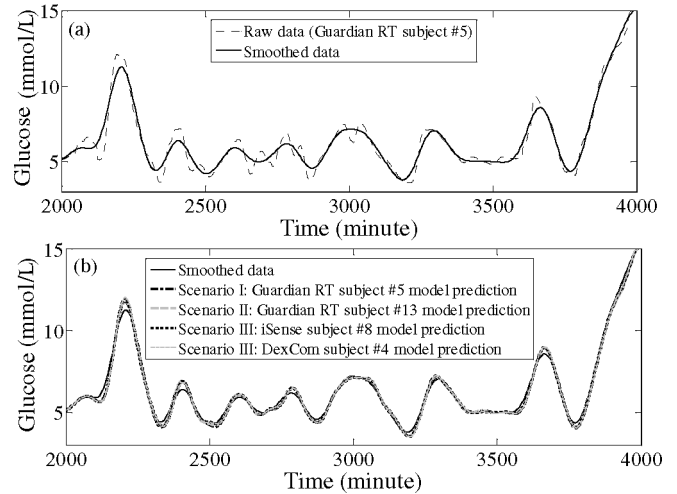


Fig. 2. (a) Raw and filtered data for Guardian RT subject 5, and (b) corresponding 30-min-ahead glucose predictions for this subject based on scenario I: same-subject predictions (RMSE = 0.20 mmol/L, no time lag), scenario II: cross-subject predictions based on a model trained on Guardian RT subject 13's data (RMSE = 0.21 mmol/L, no time lag), and scenario III: cross-study model trained on iSense subject 8's data (RMSE = 0.22 mmol/L, no time lag), as well as on a model trained on DexCom subject 4's data (RMSE = 0.24 mmol/L, no time lag).

study. The table indicates that the coefficients, in particular, the ones with relatively large values ( $>0.05$ ), were indeed similar across the studies and that their differences were, in general, within one standard deviation. This suggests that a universal model could be obtained from one subject's data and used to predict another subject's glucose levels across a short prediction horizon, completely bypassing the need to develop and fine tune the model for that subject. However, because the coefficients  $b_i$  describe the correlations in the time-series signal, their absolute values are a function of the sampling frequency of the data used to develop the model. Thus, models developed on glucose data sampled at different frequencies are expected to yield slightly different coefficients.

For demonstration purposes, we randomly chose a subject, Guardian RT subject 5, and used models developed from this and other subjects' data to predict this subject's glucose concentrations (in the testing dataset). Fig. 2(a) shows the raw and filtered glucose signals, and Fig. 2(b) shows the 30-min-ahead predictions for four different models, which exemplify the models' portability in the three scenarios. Fig. 2(a) shows the difference between the raw and the filtered data, indicating how the filtering algorithm smoothed the sharp excursions in the raw signal. On average, the filtering process removed about 7% of the signal's energy, which constitutes an acceptable loss [15]. In practice, the optimal amount of filtering poses a tradeoff between missed and false alarms for hypo- and hyperglycemic episodes. More filtering produces smoother signals and increases the frequency of missed alarms. Conversely, less filtering retains the sharp excursions of the raw signals, increasing the frequency of false alarms. The results in Fig. 2(b) indicate that the predictions of the Guardian RT subject 5 based on four different models were nearly indistinguishable from one another.

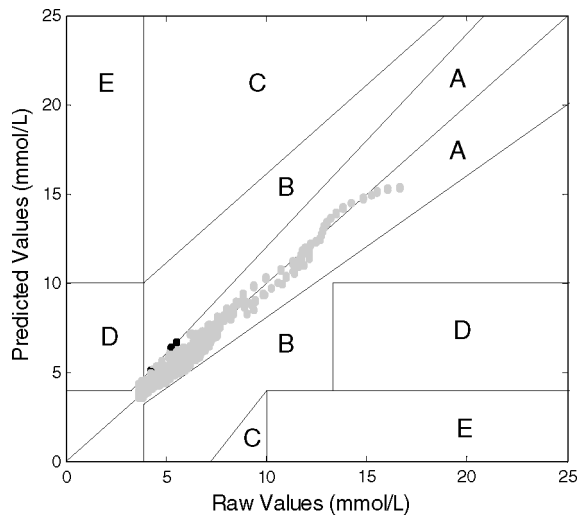


Fig. 3. Clarke EGA plot of paired predicted and raw glucose signals for the Guardian RT subject 5 corresponding to the four models in Fig. 2(b). For the 1600 data points of the four 30-min-ahead model predictions, 99.25% and 0.75% of the data lay in the clinically acceptable zones A and B, respectively.

The predictions based on Guardian RT subject 13's model illustrated model portability across different subjects within the same study (scenario II). Similarly, the predictions based on iSense subject 8's model and DexCom subject 4's model demonstrated portability across different studies and across different types of diabetes (scenario III). The same-subject predictions (Guardian RT subject 5) in scenario I served as a reference for comparison among the different models.

To assess the utility of the predictions using clinically acceptable metrics, we performed Clarke EGA [16], which maps pairs of sensor-predicted glucose concentrations into five zones, A–E, of varying degrees of accuracy and inaccuracy of glucose estimation. Values in zones A and B are clinically acceptable, values in zone C may result in unnecessary corrections, values in zone D could lead to incorrect treatments and detections, and values in zone E represent erroneous treatment. Fig. 3 shows the composite EGA scatter plot for the Guardian RT subject 5 corresponding to the four model predictions in Fig. 2(b). Each of the 1600 predictions, i.e., 400 predictions for each model, was paired with the corresponding raw glucose concentration in Fig. 2(a). Of the 1600 data points, 1588 (or 99.25%) lay in zone A and 12 (or 0.75%) in zone B. For the 12 points in zone B, each of the four models contributed three points, and these points corresponded to predictions at two time instances, 2150 and 2660 min, where the deviations between the raw and the filtered signals were the largest [see Fig. 2(a)]. These results further demonstrated the equivalent predictive power obtained with the same-subject model, the cross-subject model, and the cross-study model.

We also performed Clarke EGA for each of the three studies using the same-subject model predictions (scenario I), where the composite results of each analysis were plotted on separate graphs (not shown). Of the 3600 entries (400 data points  $\times$  9 subjects) for the iSense study, 3564 (99.0%) lay in zone A, 35 in zone B, and 1 in zone D. Of the 7200 entries (400  $\times$  18) for

the Guardian study, 7150 (99.3%), 32, and 18 lay in zones A, B, and D, respectively. Similarly, of the 2800 entries (400  $\times$  7) for the DexCom study, 2787 (99.5%), 12, and 1 lay in zones A, B, and D, respectively. These results demonstrated the clinical utility of our predictive models.

To summarize the performance of the models, we tabulated the results of the 30-min-ahead predictions for the three studies involving the three scenarios. Table III shows the RMSEs and prediction time lags for the nine iSense subjects tested using different models from the three scenarios. In scenarios II and III, the reported RMSEs and time lags were averaged values. For example, for iSense subject 1, the scenario II RMSE of 0.13 mmol/L and time lag of 1.3 min represent averaged values of the predictions over the other eight iSense subjects (not including iSense subject 1). Similarly, for iSense subject 1, the scenario III RMSE of 0.12 mmol/L and time lag of 1.2 min represent averaged values of the predictions over the 18 Guardian RT subjects and the seven DexCom subjects. Equivalent tabulations are shown in Tables IV and V for the 18 Guardian RT subjects and the seven DexCom subjects, respectively. The results in Tables III–V show not only that the glucose models derived using our methodology did not vary significantly [as shown in Fig. 1(d)], but that they also yielded very accurate forecasts (negligible average RMSEs and prediction time lags).

Finally, to verify that the employed datasets did not only represent well-treated diabetic patients with glucose levels within the euglycemic range and that the filtering procedure did not oversmooth the raw data, we calculated the total number of hypo- and hyperglycemic episodes in the raw, filtered, and predicted data. Because there is no consensus on what constitutes a hypo- or hyperglycemic episode [17], we adopted a lower threshold for hypoglycemia of 3.9 mmol/L (70 mg/dL) and an upper threshold for hyperglycemia of 10 mmol/L (180 mg/dL), required an interepisode separation of at least 30 min, and a minimum of 30 min (seven consecutive data points) outside of the euglycemic range to count the excursion as an episode. Table VI lists the cumulative number of hypo- and hyperglycemic episodes, and related statistics (averaged over the corresponding subjects) for the raw, filtered, and predicted data for each of the three studies. The results confirmed that the subjects did exhibit glucose excursions and that the filtering did not significantly smoothed them out. Overall, the model correctly predicted 89 out of 93 hyperglycemic episodes and 20 out of 23 hypoglycemic episodes.

We also performed sensitivity analysis by considering 20 min (five consecutive data points), instead of 30 min, outside of the euglycemic range for counting the excursion as an episode. This more conservative definition yielded a slight decrease in prediction accuracy, with the model correctly predicting 90 out of 97 hyperglycemic episodes and 20 out of 27 hypoglycemic episodes (entries inside the parenthesis in the last two columns in Table VI).

#### IV. DISCUSSION

The portability properties demonstrated by our models can be attributed to two factors: the conserved nature of the

TABLE III  
MODEL PERFORMANCE FOR THE NINE iSense SUBJECTS

iSense Testing Subject #	30-min-ahead Predictions Using Models Trained on					
	iSense				Guardian RT & DexCom	
	Same-Subject (Scenario I)		Cross-Subject (Scenario II)		Cross-Study (Scenario III)	
	RMSE (mmol/L)	Lag (minutes)	RMSE (SD) (mmol/L)	Lag (SD) (minutes)	RMSE (SD) (mmol/L)	Lag (SD) (minutes)
1	0.14	5.0	0.13 (0.01)	1.3 (2.3)	0.12 (0.01)	1.2 (2.2)
2	0.15	0.0	0.19 (0.03)	0.0 (0.0)	0.18 (0.02)	0.0 (0.0)
3	0.20	0.0	0.21 (0.02)	0.0 (0.0)	0.20 (0.02)	0.0 (0.0)
4	0.17	0.0	0.18 (0.02)	1.3 (2.3)	0.17 (0.02)	2.6 (2.5)
5	0.20	0.0	0.23 (0.01)	0.0 (0.0)	0.22 (0.01)	0.0 (0.0)
6	0.19	0.0	0.21 (0.02)	0.0 (0.0)	0.20 (0.02)	0.0 (0.0)
7	0.16	0.0	0.17 (0.01)	0.0 (0.0)	0.16 (0.01)	0.0 (0.0)
8	0.19	0.0	0.18 (0.01)	0.0 (0.0)	0.18 (0.01)	0.0 (0.0)
9	0.17	0.0	0.14 (0.01)	0.0 (0.0)	0.14 (0.01)	0.0 (0.0)
Average (SD)	0.17 (0.02)	0.6 (1.7)	0.18 (0.03)	0.3 (1.2)	0.17 (0.03)	0.4 (1.4)

TABLE IV  
MODEL PERFORMANCE FOR THE 18 GUARDIAN RT SUBJECTS

Guardian RT Testing Subject #	30-min-ahead Predictions Using Models Trained on					
	Guardian RT				iSense & DexCom	
	Same-Subject (Scenario I)		Cross-Subject (Scenario II)		Cross-Study (Scenario III)	
	RMSE (mmol/L)	Lag (minutes)	RMSE (SD) (mmol/L)	Lag (SD) (minutes)	RMSE (SD) (mmol/L)	Lag (SD) (minutes)
1	0.14	0.0	0.14 (0.01)	0.0 (0.0)	0.14 (0.01)	0.0 (0.0)
2	0.30	0.0	0.35 (0.02)	0.0 (0.0)	0.36 (0.02)	0.0 (0.0)
3	0.31	0.0	0.32 (0.02)	0.0 (0.0)	0.33 (0.03)	0.0 (0.0)
4	0.21	0.0	0.17 (0.02)	0.0 (0.0)	0.18 (0.02)	0.0 (0.0)
5	0.20	0.0	0.21 (0.01)	0.0 (0.0)	0.22 (0.01)	0.0 (0.0)
6	0.24	0.0	0.26 (0.02)	0.0 (0.0)	0.26 (0.02)	0.0 (0.0)
7	0.25	0.0	0.24 (0.02)	0.0 (0.0)	0.25 (0.02)	0.0 (0.0)
8	0.25	0.0	0.25 (0.02)	0.0 (0.0)	0.26 (0.02)	0.0 (0.0)
9	0.09	0.0	0.09 (0.01)	0.0 (0.0)	0.09 (0.01)	0.0 (0.0)
10	0.17	0.0	0.19 (0.03)	0.0 (0.0)	0.20 (0.03)	0.0 (0.0)
11	0.11	0.0	0.11 (0.01)	0.6 (1.7)	0.12 (0.01)	0.6 (1.7)
12	0.11	0.0	0.09 (0.01)	0.0 (0.0)	0.09 (0.01)	0.0 (0.0)
13	0.21	0.0	0.21 (0.02)	0.6 (1.7)	0.22 (0.02)	0.0 (0.0)
14	0.27	0.0	0.25 (0.02)	1.5 (2.3)	0.26 (0.02)	3.1 (2.5)
15	0.18	0.0	0.17 (0.01)	0.0 (0.0)	0.18 (0.02)	0.0 (0.0)
16	0.22	0.0	0.24 (0.02)	0.0 (0.0)	0.24 (0.02)	0.0 (0.0)
17	0.25	0.0	0.27 (0.04)	0.0 (0.0)	0.29 (0.04)	0.0 (0.0)
18	0.27	0.0	0.24 (0.02)	0.0 (0.0)	0.25 (0.02)	0.0 (0.0)
Average (SD)	0.21 (0.06)	0.0 (0.0)	0.21 (0.07)	0.1 (0.8)	0.22 (0.08)	0.2 (1.0)

frequency content in the glucose signal of diabetic patients and the properties of our modeling approach.

A recent study by Rahaghi *et al.* [18] shows that the dynamics in the blood glucose time-series signal of diabetic patients can be characterized by four distinct frequency ranges. These different frequency ranges characterize different physiologic mechanisms and are best described by the periodicity of their oscillations. The highest frequency range, with periods between 5 and 15 min, is

generated by pulsatile secretion of insulin. The second highest, ultradian glucose oscillations, corresponds to periods between 60 and 120 min. Exogenous inputs, such as meals and insulin, generate oscillations with periods between 150 and 500 min, and finally, circadian oscillations are responsible for the low-frequency range, with periods longer than 700 min.

Analysis of the time-series glucose signals of all subjects in the three studies confirmed these findings and, more importantly,

TABLE V  
MODEL PERFORMANCE FOR THE SEVEN DEXCOM SUBJECTS

DexCom Testing Subject #	30-min-ahead Predictions Using Models Trained on					
	DexCom				iSense & Guardian RT	
	Same-Subject (Scenario I)		Cross-Subject (Scenario II)		Cross-Study (Scenario III)	
	RMSE (mmol/L)	Lag (minutes)	RMSE (SD) (mmol/L)	Lag (SD) (minutes)	RMSE (SD) (mmol/L)	Lag (SD) (minutes)
1	0.18	0.0	0.19 (0.02)	0.0 (0.0)	0.19 (0.02)	0.0 (0.0)
2	0.13	0.0	0.13 (0.01)	0.0 (0.0)	0.14 (0.01)	0.0 (0.0)
3	0.14	0.0	0.17 (0.02)	0.0 (0.0)	0.17 (0.01)	0.0 (0.0)
4	0.19	0.0	0.17 (0.02)	0.0 (0.0)	0.17 (0.01)	0.0 (0.0)
5	0.19	0.0	0.19 (0.02)	0.0 (0.0)	0.19 (0.02)	0.0 (0.0)
6	0.15	0.0	0.17 (0.02)	0.0 (0.0)	0.18 (0.02)	0.0 (0.0)
7	0.13	0.0	0.13 (0.01)	0.0 (0.0)	0.12 (0.01)	0.0 (0.0)
Average (SD)	0.16 (0.03)	0.0 (0.0)	0.16 (0.03)	0.0 (0.0)	0.17 (0.03)	0.0 (0.0)

TABLE VI  
CUMULATIVE NUMBER OF HYPO- AND HYPERGLYCEMIC EPISODES AND RELATED STATISTICS FOR THE RAW, FILTERED, AND PREDICTED DATA FOR EACH OF THE THREE STUDIES

	Min (avg.) (mmol/L)	Max (avg.) (mmol/L)	Mean (avg.) (mmol/L)	SD (avg.) (mmol/L)	Total # Hyperglycemic Episodes	Total # Hypoglycemic Episodes
iSense (9 subjects)						
Raw	3.95	15.81	8.72	2.61	25 (26)	4 (4)
Filtered	4.38	14.70	8.72	2.52	24 (24)	3 (3)
Predicted	4.28	14.87	8.69	2.55	24 (25)	3 (3)
Guardian RT (18 Subjects)						
Raw	3.41	16.75	8.99	3.45	49 (51)	17 (21)
Filtered	3.92	16.30	8.99	3.38	48 (48)	15 (15)
Predicted	3.77	16.40	8.96	3.41	49 (49)	15 (15)
DexCom (7 subjects)						
Raw	4.66	14.00	8.48	1.97	19 (20)	2 (2)
Filtered	5.21	12.61	8.48	1.87	16 (16)	2 (2)
Predicted	5.09	12.59	8.46	1.90	16 (16)	2 (2)

The entries inside the parenthesis indicate the cases when we used a minimum of 20 consecutive minutes, instead of 30 min, outside of the euglycemic range for counting the excursion as an episode.

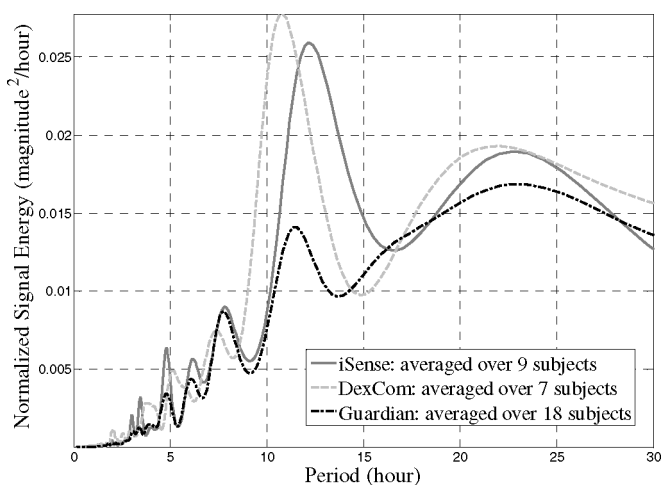


Fig. 4. Power spectrum density profiles for each of the three studies, averaged over the subjects in each study. The abscissa axis is plotted as periods (i.e., the reciprocal of frequencies) of oscillations.

showed that the frequency content in the signals is conserved across subjects. Fig. 4 shows the power spectrum density profiles for each of the three studies, averaged over the subjects in each

study. Note that while the amplitude of the profiles is different for each of the studies, the periodicity (i.e., the location of the peaks on the  $x$ -axis) is conserved across the studies. Perhaps this finding should not be too surprising as the conservation of biological rhythms, such as the circadian rhythm, across species, or even kingdoms, is a well-known phenomenon [19].

This similarity in the frequency content of the glucose signals was exploited by the AR models. Periodic signals like glucose concentration can be fully characterized by three parameters: amplitude, frequency, and phase of the underlying oscillations. However, a unique property of AR models is their invariance with respect to a signal's amplitude and phase, and sole dependency on its frequency [14]. In fact, a sequence of AR coefficients captures and represents the frequency content of a time-series signal. Therefore, the development of AR models from signals with similar frequency content inevitability produced similar (or portable) models, regardless that different time-series signals recorded from different subjects had different amplitudes and initial phases. It is this invariance of the AR coefficients to the glucose signal's amplitude and phase that affords model portability across subjects with type 1 and type 2 diabetes. Type 1 diabetes patients usually have larger

glucose-level variations than type 2 patients. However, as long as these variations contain the same frequency information, AR models shall remain portable across them. Moreover, because of the frequency-dependent nature of the AR coefficients, information concerning exogenous inputs, such as meals and exercise, is automatically incorporated into the models as long as this information is present in the training data.

We expect, however, if some of the subjects were nondiabetic and fasting, that the models' portability could be jeopardized because the glucose dynamics would be different in this case. This would be particularly relevant for the highest frequency component of the glucose time-series signal, i.e., the shortest periods spanning between 5 and 15 min, because while these periods are prominent in nondiabetic, fasting individuals [18], they are drastically attenuated in diabetic patients, in particular, in type 1 patients. In diabetic patients, insulin-generating cells responsible for pulsatile secretion of insulin, and the corresponding 5- to 15-min periods, are severely handicapped, essentially eliminating these periods in the glucose signals [9], [20]. Moreover, the blood-to-interstitial transport acts as a low-pass filter, reducing the high-frequency dynamics in the CGM signals [21], which are further attenuated by our filtering procedure.

Filtering of the high-frequency signal, required to yield consistent AR coefficients and robust models [4], could have the unintended consequence of smoothing out hypo- and hyperglycemic episodes in the glucose data. Our filtering procedure, however, did not significantly impact the ability to capture these episodes and, hence, the clinical usefulness of our approach. Table VI shows that our models correctly predicted 96% of the hyperglycemic episodes and 87% of the hypoglycemic episodes present in the three studies. Sensitivity analysis using a more conservative definition of hypo- and hyperglycemic episodes (i.e., 20 instead of 30 consecutive minutes outside of the euglycemic range) only slightly decreased the prediction accuracy of these cases.

Another contributing property for the AR model portability relates to the limits imposed on the model coefficients by the constrained least-squares method. Besides fitting the AR model to the data, the employed constrained least-squares method also limits the curvature (i.e., the norm of the second derivative) of the AR coefficients [4]. This is illustrated in Fig. 1, where the shape of the coefficients can be loosely described as a dampened sine wave that reflects the periodic nature of the glucose signal, and where the values of the coefficients gradually decrease as the time lag (or coefficient number) increases. This behavior of the AR coefficients is intuitively correct, as we expect the glucose data to gradually lose intersample correlations as the time lag between samples increases. However, had we used the standard least-squares method (instead of the constrained one) to obtain the AR-model coefficients, the resulting coefficients would have exhibited unphysiologic behavior, with their values not gradually decreasing, which would incorrectly indicate that intersample correlations do not decrease as a function of time [4].

Analysis of Fig. 2 is revealing in terms of the models' predictive capabilities, as it shows that predictions produced by different models are virtually indistinguishable with practically

identical prediction errors. For example, the same-subject predictions (Guardian RT subject 5 model, scenario I) resulted in a RMSE of 0.20 mmol/L, while the cross-subject predictions (Guardian RT subject 13 model, scenario II) resulted in a RMSE of 0.21 mmol/L. The two cross-study predictions in scenario III, iSense subject 8 model, and DexCom subject 4 model resulted in RMSEs of 0.22 and 0.24 mmol/L, respectively. These observations were corroborated by Clarke EGA, where more than 99.0% of the paired sensor-predicted glucose concentrations lay in the clinically relevant zone A (see Fig. 3).

Table IV shows that although the models were portable, their performance, in terms of RMSE, may vary from subject to subject. For example, the RMSE for subject 9 in scenario I was 0.09 mmol/L, whereas for subject 2, the RMSE was 0.30 mmol/L. This difference in prediction error for specific subjects was due to the different amounts of noise present in different subjects' data. However, as can be seen from Tables III–V, for a given subject, the models' performance is practically identical.

Analysis of Tables III and IV also reveals that sometimes a small time lag was introduced in the cross-subject and the cross-study scenarios. This small time lag was, most likely, due to small differences in glucose dynamics across different individuals. AR models would exhibit prediction lags if they failed to account for some frequency component present in the test signal. Such small differences in frequency components do exist in our datasets and could be the reason for the small prediction time lags. The introduction of a 5-min lag for iSense subject 1 in scenario I (see Table III) was, most likely, due to small frequency differences between this subject's training and testing data.

We note that our results on model portability are only valid for AR-type models, and we do not claim the nonexistence of interindividual variability in glucose dynamics. As was discussed previously, AR models capture the signal's frequency information, and are invariant to the signal's phase and amplitude. The latter property is not shared by other modeling approaches, such as those based on ordinary differential equations or harmonic regression, which prevents their portability.

## V. CONCLUSION

In conclusion, we have shown that it is possible to develop stable, universal glucose models that capture the correlations in glucose time-series signals of diabetic patients. Given continuous glucose signals from a patient, such universal models could be readily used to make near-future glucose concentration predictions for other patients without any need for model customization. This finding has significant practical and clinical implications, as it opens the door for integration of CGM devices and predictive algorithms for model-based, proactive therapy of diabetic patients. With this proof-of-concept established, future work shall concentrate on prospective, real-time applications.

## DISCLAIMER

The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as



reflecting the views of the U.S. Army or of the U.S. Department of Defense. This paper has been approved for public release with unlimited distribution.

## REFERENCES

- [1] D. C. Klonoff, "Continuous glucose monitoring—Roadmap for 21st century diabetes therapy," *Diabetes Care*, vol. 28, pp. 1231–1239, 2005.
- [2] T. Bremer and D. A. Gough, "Is blood glucose predictable from previous values? A solicitation for data," *Diabetes*, vol. 48, pp. 445–451, 1999.
- [3] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, and C. Cobelli, "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 931–937, May 2007.
- [4] A. Gani, A. V. Gribok, S. Rajaraman, W. K. Ward, and J. Reifman, "Predicting subcutaneous glucose concentration in humans: Data-driven glucose modeling," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 2, pp. 246–254, Feb. 2009.
- [5] P. Dua, F. J. Doyle, and E. N. Pistikopoulos, "Model-based blood glucose control for type 1 diabetes via parametric programming," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1478–1491, Aug. 2006.
- [6] (2009, Mar. 11). iSense Corporation, Wilsonville, OR [Online]. Available: <http://www.isensecorp.com>
- [7] (2009, Mar. 11). Medtronic, Inc., Northridge, CA [Online]. Available: <http://www.medtronic.com>
- [8] (2009, Feb. 16). DexCom, Inc., San Diego, CA [Online]. Available: <http://www.dexcom.com>
- [9] D. A. Gough, K. Kreutz-Delgado, and T. M. Bremer, "Frequency characterization of blood glucose dynamics," *Ann. Biomed. Eng.*, vol. 31, pp. 91–97, 2003.
- [10] J. Reifman, S. Rajaraman, A. Gribok, and W. K. Ward, "Predictive monitoring for improved management of glucose levels," *J. Diabetes Sci. Technol.*, vol. 1, pp. 478–486, 2007.
- [11] (2009, Mar. 11). Diabetes Research in Children Network (DirecNet). [Online]. Available: <http://public.direc.net>
- [12] B. P. Kovatchev, W. L. Clarke, M. Breton, K. Brayman, and A. McCall, "Quantifying temporal glucose variability in diabetes via continuous glucose monitoring: Mathematical methods and clinical application," *Diabetes Technol. Ther.*, vol. 7, pp. 849–862, 2005.
- [13] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ: Prentice-Hall PTR, 1999.
- [14] C. Chatfield, *The Analysis of Time Series: An Introduction*. London, U.K./Boca Raton, FL: Chapman & Hall/CRC Press, 2004.
- [15] T. Jolliffe, *Principal Component Analysis*. New York/Secaucus, NJ: Springer-Verlag, 2002.
- [16] W. L. Clarke, "The original Clarke error grid analysis (EGA)," *Diabetes Technol. Ther.*, vol. 7, pp. 776–779, 2005.
- [17] D. Klonoff, private communication, Mills-Peninsula Health Services, San Mateo, CA, Jun. 2009.
- [18] F. N. Rahaghi and D. A. Gough, "Blood glucose dynamics," *Diabetes Technol. Ther.*, vol. 10, pp. 81–94, 2008.
- [19] K. Wager-Smith and S. Kay, "Circadian rhythm genetics: From flies to mice to humans," *Nature Genet.*, vol. 26, pp. 23–27, 2000.
- [20] N. Porksen, M. Hollingdal, C. Juhl, P. Butler, J. D. Veldhuis, and O. Schmitz, "Pulsatile insulin secretion: Detection, regulation, and role in diabetes," *Diabetes*, vol. 51, pp. S245–S254, 2002.
- [21] M. D. Breton, D. P. Shields, and B. P. Kovatchev, "Optimum subcutaneous glucose sampling and Fourier analysis of continuous glucose monitors," *J. Diabetes Sci. Technol.*, vol. 2, pp. 495–500, 2008.



**Adiwinata Gani** received the B.S., M.S., and Ph.D. degrees in chemical engineering from the University of California, Los Angeles, in 2002, 2003, and 2007, respectively.

He was a Postdoctoral Fellow in the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, Fort Detrick, MD. His current research interests include inverse and ill-posed problems in engineering, and fault-detection and fault-tolerant control systems.



**Andrei V. Gribok** received the B.S. degree in systems science and the M.S. degree in nuclear engineering from Moscow Institute of Physics and Engineering, Moscow, Russia, both in 1987, and the Ph.D. degree in biological physics from Moscow Institute of Biological Physics, Moscow, in 1996.

He is currently a Research Assistant Professor in the Department of Nuclear Engineering, University of Tennessee, Knoxville. He was on an interpersonal agreement assignment with the Telemedicine and Advanced Technology Research Center, Fort Detrick, MD.

His current research interests include inverse and ill-posed problems in engineering, statistical learning, and model misspecification in statistics.



**Yinghui Lu** received the B.S. degree from Zhejiang University, Hangzhou, China, in 1996, the M.S. degree from the University of Tennessee, Knoxville, in 2002, and the Ph.D. degree from Georgia Institute of Technology, Atlanta, in 2007, all in electrical engineering.

He is a Postdoctoral Fellow in the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, Fort Detrick, MD. His research interests include digital signal processing, intelligent control systems, ultrasonics, and medical

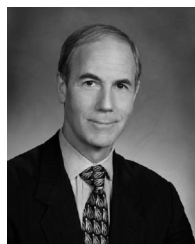
informatics.



**W. Kenneth Ward** received the Medical degree from the University of Colorado, Boulder, in 1978.

He received a Research Fellowship in diabetes and endocrinology from the University of Washington. He is currently a Diabetes Specialist and a Biosensor Scientist with iSense Corporation, Wilsonville, OR, where he is engaged in collecting data used in the study while collaborating with the staff at the Profil Institute, Neuss, Germany. He is also with Oregon Health and Sciences University, Wilsonville. His research interests include development of minimally

invasive and fully implantable biosensors, artificial pancreas algorithms, biomaterials, and the foreign-body response to implanted devices.



**Colonel Robert A. Vigersky** received a combined B.A. and M.D. degree from Boston University, Boston, MA, in 1970.

He was an Internal Medicine Residency at the Johns Hopkins Hospital. In 1975, he was an Endocrinology Fellow with the National Institutes of Health. He is on an Active Duty in the United States Army, and is the Founder and the Director of the Diabetes Institute, Walter Reed Health Care System, Washington, DC. His current research interests include investigating the use of technology to improve

the outcomes of patients with diabetes mellitus.

Dr. Vigersky is the President of the Endocrine Society and its Representative to the American Medical Association.



**Jaques Reifman** received the B.S. degree in civil engineering from Rio de Janeiro State University, Rio de Janeiro, Brazil, in 1980, the B.B.A. degree in business administration from Rio de Janeiro Federal University, Rio de Janeiro, in 1985, and the M.S.E. and Ph.D. degrees in nuclear engineering from the University of Michigan, Ann Arbor, in 1985 and 1989, respectively.

He is currently a Senior Research Scientist in the Department of the Army, U.S. Army Medical Research and Materiel Command (USAMRMC), Fort

Detrick, MD, where he is also the Director of two organizations that he created: the Department of Defense Biotechnology High Performance Computing Software Applications Institute for Force Health Protection and the USAMRMC Bioinformatics Cell, Telemedicine and Advanced Technology Research Center. His research interests include physiologic signal processing, statistical pattern recognition, artificial intelligence, data mining, biomathematical modeling, systems biology, bioinformatics, genomics, and proteomics.