Blood Glucose Prediction using Physiological Models and Support Vector Regression

Razvan Bunescu    Nigel Struble    Kevin Plis

Cindy Marling    Frank Schwartz    Jay Shubrook

Ohio University, Athens, OH
Management of Type I Diabetes

• Approximately 20 million people have Type 1 Diabetes:
  – In type 1 diabetes, the pancreas produces no insulin.
  – Patients depend upon external supplies of insulin, via injections or insulin pumps.

• Diabetes can not be cured, but it can be treated and managed:
  – To delay or prevent long-term complications, patients try to keep Blood Glucose Levels (BGL) as close to normal as possible.
  – Patients monitor blood glucose using:
    • Glucometers (fingerstick measurements).
    • Continuous Glucose Measurement Systems (CGMS).
Chronic Complications vs. Blood Glucose Control

- Foot Ulcers
- Angina
- Heart Attack
- Coronary Bypass Surgery
- Stroke
- Kidney Transplant
- Dialysis
- Blindness
- Amputation

- Albuminuria
- Macular Edema
- Proliferative Retinopathy
- Periodontal Disease
- Impotence
- Gastroparesis
- Depression

- Microalbuminuria
- Mild Retinopathy
- Mild Neuropathy
Monitoring BGL: Glucometer

A blood sample is taken and put on test strip.

Strip is put into blood glucose meter.

A log book is a helpful aid in keeping track of blood glucose levels.
Data Overload
Continuous Glucose Monitoring (CGM) in Insulin Pump Therapy Systems

CGM Sensor:
- interstitial BGL.
- every 5 minutes.

Insulin Pump delivers insulin through boluses and basal rate:

<table>
<thead>
<tr>
<th>Three Types of Bolus Insulin</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NORMAL BOLUS</strong></td>
</tr>
<tr>
<td><img src="image_url" alt="Image of three types of bolus insulin" /></td>
</tr>
</tbody>
</table>
More Data Overload
Achieving Good Blood Glucose Control

• Patients must continually monitor their blood glucose levels and adjust insulin doses, striving to keep blood glucose levels as close to normal as possible:
  – Requires significant effort from patients and doctors.

• Try to avoid especially:
  – Hypoglycemia
  – Hyperglycemia
  – Excessive Glycemic Variability
Achieving Good Blood Glucose Control

- Patients must continually monitor their blood glucose levels and adjust insulin doses, striving to keep blood glucose levels as close to normal as possible:
  - Requires significant effort from patients and doctors.

- Try to avoid especially:
  - Hypoglycemia
  - Hyperglycemia
  - Excessive Glycemic Variability
Automatic BGL Prediction

• Design a time series forecasting model that predicts BGL 30 or 60 minutes into the future:
  – Accurate predictions up to 60m in advance would allow plenty of time to take preventive action, to avoid hypo- or hyper-glycemia.

• Inputs for the prediction model:
  – Previous blood glucose measurements taken at 5-minute intervals through a CGM system.
  – Daily event data:
    • Insulin dosages, recorded in the CGM device.
    • Life events, collected through a smartphone interface.
Input: Blood Glucose Levels and Insulin Dosages
Input: Life Events

• Developed a smartphone interface to collect relevant **life events**:
  – Meals (carb amounts, glycemic index).
  – Sleep (start, end).
  – Work (start, end).
  – Exercise (intensity, start, duration).
  – Hypoglycemic event.
  – Health events (stress, depression, ...).
  – Other events.

• Designed to encourage entering events immediately before/after they happen:
  – to minimize incorrect/incomplete data.
Evaluation Dataset

- Total of 1,400 days worth of clinical patient data:
  - CGMS + insulin + life events.

- Human performance on the task of BGL prediction:
  - Asked 3 diabetes experts to manually label an evaluation dataset with their 30/60 min predictions:
    - 200 timestamps, coming from 5 patients with T1D.
      - 40 points per patient.
      - Manually selected to reflect a diverse set of situations.
  - Built a GUI to facilitate navigating the data and labeling.
Physician Performance

• Compared the 3 physicians against 2 baselines:
  – $t_0$ predicts that future BGL is the same as current BGL.
  – Auto-Regressive Integrated Moving Averages (ARIMA), trained on past BGL data.

• Evaluation measures:
  – Root Mean Square Error (RMSE).
  – Total cost of ternary classification:
    • Future BGL is Same (S), Lower (L), Higher (H) as current BGL.
      – Same means within 5 (10) mg/dl for 30 (60) min prediction.
    • cost(L, S) = cost (H, S) = 1; cost(L, H) = 2.
Physician Performance

Physicians, who use daily event data, outperform ARIMA. Physicians regularly refer to daily events:

- Timing of meal events and boluses, carb amounts, bolus types.
Physician Performance

- Physicians, who use daily event data, outperform ARIMA.
- Physicians regularly refer to daily events:
  - Timing of meal events and boluses, carb amounts, bolus types.

Use daily events to extract features for automatic BGL prediction.
Physiological Modeling of BG Dynamics

- Use equations from literature [6, 7, 8, 9] to model dynamics of variables that are relevant to BG behavior:
  - Almost identical equations (based on the same data).
  - Characterize the overall dynamics into 3 compartments:
    - **Meal absorption** dynamics.
    - **Insulin** dynamics.
    - **Glucose** dynamics.

- Update some equations and their parameters to better match published data and feedback from our doctors.
A Physiological Model of BG Dynamics

- A continuous dynamical model that is described by:
  1) The input variables $U$.
  2) The state variables $X$.
  3) The state transition function $f$ that computes the next state given the current state and input i.e. $X_{t+1} = f(X_t, U_t)$.

1) The vector of input variables $U$ contains:
   - $U_C(t)$, the carbohydrate intake measured in grams (g).
   - $U_I(t)$, the amount of insulin measured in insulin units (U):
     - Computed from bolus events and basal rate data.
A Physiological Model of BG Dynamics

1) **The state variables** $X$ are organized according to the 3 compartments:

1) **Meal Absorption Dynamics:**
   - $C_{g1}(t) =$ carbohydrate consumption (g).
   - $C_{g2}(t) =$ carbohydrate digestion (g).

2) **Insulin Dynamics:**
   - $I_S(t) =$ subcutaneous insulin ($\mu$U).
   - $I_m(t) =$ insulin mass ($\mu$U).
   - $I(t) =$ level of active plasma insulin ($\mu$U/ml).

3) **Glucose Dynamics:**
   - $G_m(t) =$ blood glucose mass (mg).
   - $G(t) =$ blood glucose concentration (mg).
A Physiological Model of BG Dynamics

2) The state transition function $f$ captures dependencies among variables in $X$ and $U$ at consecutive time steps:
2) The state transition function $f$ captures dependencies among variables in $X$ and $U$ at consecutive time steps:
A Physiological Model of BG Dynamics

2) The state transition function $f$ captures dependencies among variables in $X$ and $U$ at consecutive time steps:
Glucose Dynamics: Insulin Dependent Utilization
Glucose Dynamics: Insulin Dependent Utilization

\[ G_m(t+1) = G_m(t) - \Delta_{dep} - \Delta_{ind} - \Delta_{clr} - \Delta_{dep} + \Delta_{abs} + \Delta_{egp} \]

\[ \Delta_{dep} = \alpha_1 \times I(t) \times [G(t) + \alpha_2] \]
A Physiological Model of BG Dynamics

- The state transition equations were used in an Extended Kalman Filter (EKF) model:
  - Run a state prediction step every 1 minute.
  - Run a correction step every 5 minutes.

- The EKF model itself can be used to make 30 or 60 minute predictions:
  - Performance is lower even than the $t_0$ baseline.
  - Could improve by tuning the $\alpha$ parameters for each patient:
    - Time consuming, unfeasible due to large number of params.
    - Difficult to incorporate other types of life events in the model.
A Support Vector Regression (SVR) Model with Physiological Features

- The state vector $X(t)$ computed by the physiological model is $X(t) = [C_{g1}(t), C_{g2}(t), I_s(t), I_m(t), I(t), G_m(t), G(t)]$:
  - Run the EKF model up to time $t_0$, with a correction step every 5 minutes $\Rightarrow X(t_0)$.
  - Run the EKF model in prediction mode for 60 more minutes $\Rightarrow X(t_0 + 30)$ and $X(t_0 + 60)$.

- Create the following features for the SVR model:
  - All predicted state variables in $X(t_0 + 30)$ and $X(t_0 + 60)$.
  - The difference vectors $X(t_0) - X(t_0 + 30)$ and $X(t_0) - X(t_0 + 60)$.
  - 12 features $\text{delta}_i = BG(t_0) - BG(t_0 - 5i)$.
  - Optionally, train ARIMA on 4 days before $t_0$, and use the 12 predictions in the one hour after $t_0$ as features.
SVR Evaluation

• Train SVR on the week of data preceding each test point $t_0$:
  – Use a Gaussian kernel:
    • Tune parameters $\gamma$, $\varepsilon$, and $C$ using grid search on the week preceding the training week.
    • If not enough tuning examples, use generic parameters tuned on another patient.

• Compare the best doctor performance with:
  – ARIMA and the $t_0$ baselines.
  – SVR model using physiological features, with ($SVR\phi+A$) and without ($SVR\phi$) ARIMA features.
  – A previous SVR system ($SVR\pi+A$) that uses CGM data, ARIMA, and an ad-hoc implementation of daily event features.
Experimental Results on BGL Prediction

<table>
<thead>
<tr>
<th>Horizon</th>
<th>$t_0$</th>
<th>ARIMA</th>
<th>Phys$_1$</th>
<th>SVR$_{\pi}$$+$A</th>
<th>SVR$_{\phi}$</th>
<th>SVR$_{\phi}$$+$A</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 min</td>
<td>27.5</td>
<td>22.9</td>
<td>19.8</td>
<td>22.2</td>
<td>19.6</td>
<td><strong>19.5</strong></td>
</tr>
<tr>
<td>60 min</td>
<td>43.8</td>
<td>42.2</td>
<td>38.4</td>
<td>41.3</td>
<td>36.1</td>
<td><strong>35.7</strong></td>
</tr>
</tbody>
</table>
Experimental Results on BGL Prediction

<table>
<thead>
<tr>
<th>Horizon</th>
<th>$t_0$</th>
<th>ARIMA</th>
<th>$\text{Phys}_1$</th>
<th>SVR$_{\pi+A}$</th>
<th>SVR$_\phi$</th>
<th>SVR$_{\phi+A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 min</td>
<td>27.5</td>
<td>22.9</td>
<td>19.8</td>
<td>22.2</td>
<td>19.6</td>
<td><strong>19.5</strong></td>
</tr>
<tr>
<td>60 min</td>
<td>43.8</td>
<td>42.2</td>
<td>38.4</td>
<td>41.3</td>
<td>36.1</td>
<td><strong>35.7</strong></td>
</tr>
</tbody>
</table>

Both SVR$_\phi$ systems outperform the 3 diabetes experts!
Current Work

Dawn Phenomena
Current Work

Hypoglycemia
Hypoglycemia Prediction Dataset

- Evaluation dataset of 5,816 hypo and non-hypo events:
  - Collected 152 hypo events:
    - Consecutive CGM readings under 70 mg/dl were combined into one event.
      - Events that lasted less then 20 minutes were ignored.
    - A prediction time $t_0$ is selected such that $t_0 + 30$ is the first point where blood glucose drops below 70.
  - Sampled 5,664 non-hypo events to reflect empirical distribution.
Hypoglycemia Prediction: Evaluation

• Each of the 5,816 test points serves as the prediction time $t_0$:
  – The SVR system is trained on data before $t_0$.
  – Predictions are made for future timestamps in 5 minute increments, from $t_0 + 5$ to $t_0 + 60$.
  – If any one of the 12 predictions is less than 70 mg/dl, we consider that the system predicted a hypo event.
Hypoglycemia Prediction: Evaluation

- Each of the 5,816 test points serves as the prediction time $t_0$:
  - The SVR system is trained on data before $t_0$.
  - Predictions are made for future timestamps in 5 minute increments, from $t_0 + 5$ to $t_0 + 60$.
  - If any one of the 12 predictions is less than 70 mg/dl, we consider that the system predicted a hypo event.

<table>
<thead>
<tr>
<th>System</th>
<th>Recall</th>
<th>FP Rate</th>
<th>RMSE</th>
<th>RMSE$_{30}$</th>
<th>RMSE$_{60}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVR</td>
<td>23.0%</td>
<td>0.8%</td>
<td>24.4</td>
<td>22.6</td>
<td>35.8</td>
</tr>
<tr>
<td>ARIMA</td>
<td>9.9%</td>
<td>0.3%</td>
<td>27.1</td>
<td>24.9</td>
<td>39.6</td>
</tr>
<tr>
<td>$t_0$</td>
<td>--</td>
<td>--</td>
<td>28.7</td>
<td>26.6</td>
<td>41.7</td>
</tr>
</tbody>
</table>
Hypoglycemia Prediction: Evaluation

- Each of the 5,816 test points serves as the prediction time $t_0$:
  - The SVR system is trained on data before $t_0$.
  - Predictions are made for future timestamps in 5 minute increments, from $t_0 + 5$ to $t_0 + 60$.
  - If any one of the 12 predictions is less than 70 mg/dl, we consider that the system predicted a hypo event.

<table>
<thead>
<tr>
<th>System</th>
<th>Recall</th>
<th>FP Rate</th>
<th>RMSE</th>
<th>RMSE$_{30}$</th>
<th>RMSE$_{60}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVR</td>
<td>23.0%</td>
<td>0.8%</td>
<td>24.4</td>
<td>22.6</td>
<td>35.8</td>
</tr>
<tr>
<td>ARIMA</td>
<td>9.9%</td>
<td>0.3%</td>
<td>27.1</td>
<td>24.9</td>
<td>39.6</td>
</tr>
<tr>
<td>$t_0$</td>
<td>--</td>
<td>--</td>
<td>28.7</td>
<td>26.6</td>
<td>41.7</td>
</tr>
</tbody>
</table>
Hypoglycemia Prediction: Evaluation

- Each of the 5,816 test points serves as the prediction time $t_0$:
  - The SVR system is trained on data before $t_0$.
  - Predictions are made for future timestamps in 5 minute increments, from $t_0 + 5$ to $t_0 + 60$.
  - If any one of the 12 predictions is less than 70 mg/dl, we consider that the system predicted a hypo event.

<table>
<thead>
<tr>
<th>System</th>
<th>Recall</th>
<th>FP Rate</th>
<th>RMSE</th>
<th>RMSE$_{30}$</th>
<th>RMSE$_{60}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVR</td>
<td>23.0%</td>
<td>0.8%</td>
<td>24.4</td>
<td>22.6</td>
<td>35.8</td>
</tr>
<tr>
<td>ARIMA</td>
<td>9.9%</td>
<td>0.3%</td>
<td>27.1</td>
<td>24.9</td>
<td>39.6</td>
</tr>
<tr>
<td>$t_0$</td>
<td>--</td>
<td>--</td>
<td>28.7</td>
<td>26.6</td>
<td>41.7</td>
</tr>
</tbody>
</table>

All 47 false alarms are under 140 mg/dl, of which 32 are 80 mg/dl or lower.
Conclusions and Future Work

• Built an adaptive model for BGL prediction that outperforms human experts:
  – Physiological modeling was essential to good performance.

• In future work, extend to use richer set of daily events, such as exercise and stress.

• Investigate unobtrusive sensing devices in order to reduce the amount of input required from the patient:
  – Heart rate, galvanic skin response, acceleration of body movements.

• Change to online learning model.
Acknowledgments

• Our dedicated research nurses.
• Current and former graduate students:
  – Nattada Nimisuwan (OU), Melih Altun (OU), and Matthew Wiley (UC-Riverside).
• Over 50 anonymous patients with Type 1 Diabetes on insulin pump therapy.
• Our generous sponsors:
References for Physiological Modeling of Blood Glucose Behavior


