Impact of sensor and measurement timing errors on model-based insulin sensitivity

Christopher Pretty,* Aaron Le Compte,* Geoffrey Shaw, J. Geoffrey Chase***

** University of Canterbury, Centre for Bioengineering, Christchurch, New Zealand (Tel: +64 3 364 2596; e-mail: Christopher.Pretty@pg.canterbury.ac.nz).*

***University of Otago, Dept of Intensive Care, Christchurch, New Zealand*

Abstract: A model-based insulin sensitivity parameter (*SI*) is often used in glucose-insulin system models to define the glycaemic response to insulin. As a parameter identified from clinical data, insulin sensitivity can be affected by blood glucose (BG) sensor error and measurement timing error, which can subsequently impact analyses or glycaemic variability during control. This study assesses the impact of both measurement timing and BG sensor errors on identified values of *SI* and its hour-to-hour variability within the ICING-type glucose-insulin system models.

Retrospective clinical data was used from 270 patients of the Christchurch Hospital intensive care unit (ICU). An error model was created for the Arkray Super-Glucocard II glucometer used in Christchurch from manufacturer supplied data. Timing error was estimated from recent, computerised clinical data. Monte Carlo analysis was used to quantify the impact of these random errors by identifying *SI* profiles from data incorporating errors and comparing them to the 'true' *SI* profile (without additional errors) at each patient hour. To consolidate comparisons over the n = 100 Monte Carlo simulations, the width of the interquartile range (IQR) was used for percentage difference from the true *SI* level and for percentage hour-to-hour variability.

The results of the study show that timing errors in isolation have little clinically significant impact on identified *SI* level or variability. The clinical impact of changes to *SI* level induced by combined sensor and timing errors is likely to be limited during glycaemic control. Identified values of *SI* were typically within 12% of the true value when influenced by both sources of error. In contrast, for variability, 95% of patient hours had an IQR of 34.9%, indicating that for half the simulations the hour-to-hour variability of *SI* was within ±17.5%.

The results of this study indicate that the impact of sensor or timing errors on *SI* level is unlikely to be clinically significant. The effects are probably overshadowed by physiological factors arising from the critical condition of the patients or other under-modelled or un-modelled dynamics. In contrast, the impact of errors on hour-to-hour *SI* variability is more pronounced and may impact the way the *SI* parameter is utilised for control and analysis

Keywords: **Biomedical system modeling, simulation and visualization; Control of physiological and clinical variables;**

1. INTRODUCTION

Physiological glucose-insulin system models typically rely on some form of insulin sensitivity parameter to characterise the patient-specific glycaemic response to exogenous insulin (Lin et al., 2011, Hovorka et al., 2008). This model-based insulin sensitivity parameter (*SI*) is identified for some period of time using blood glucose (BG) concentration and insulin and nutrition administration data. Errors in blood glucose concentration and measurement timing can thus affect the identified values of *SI*.

In the busy intensive care unit (ICU) environment, BG measurements are rarely taken and recorded at an exact, scheduled time. Sensor errors add uncertainty to the measured BG concentration. Both errors propagate through to *SI* during parameter identification, which in turn may impact subsequent analyses or glycaemic variability during control.

Typical point-of-care glucometers have measurement errors in the range 2-10% (Abbott, 2010, Arkray, 2007, Roche, 2007, Roche, 2008, Solnica et al., 2003). The glucometer used in the Christchurch Hospital ICU (Super-Glucocard II, Arkray Inc., Japan) typically has an error of less than 10% (Arkray, 2007). The uncertainty in BG concentration resulting from sensor error impacts the identified values of *SI* through altering the glucose flux that must be balanced by the insulin-mediated glucose disposal term in the glucose-insulin system model.

The objective of this investigation was to assess the impact of both measurement timing and sensor errors on identified

values of *SI* and its hour-to-hour variability. Specifically, the *SI* parameter from a glucose-insulin system model similar to that described by Lin et al. (2011) was investigated with patient data from the Christchurch Hospital ICU.

2. SUBJECTS AND METHODS

2.1 Patients

This study was conducted as retrospective analyses of data from 270 patients admitted to the Christchurch Hospital ICU between 2005 and 2007. All patients were on the SPRINT protocol for at least 24 hours (Chase et al., 2008). Table 1 shows a summary of the cohort details. The Upper South Regional Ethics Committee, New Zealand granted approval for the audit, analysis and publication of this data.

Table 1. Cohort details summary. Data are shown as median [interquartile range] where appropriate.

N	270
Age (years)	65 [49-73]
Gender (M/F)	165/105
Operative/Non-Operative	104/166
Hospital mortality (%)	27%
APACHE II score	19 [16-25]
APACHE II ROD (%)	30 [17-53]
Diabetic status (T1DM/T2DM)	10/34
ICU length of stay (hrs)	160 [77-346]

2.2 Model-based insulin sensitivity

The glucose-insulin system model used in this study was an enhanced version of the ICING model described by Lin et al. (2011), with a new endogenous insulin secretion sub-model (7) derived from data not yet published. The model is defined below in (1)-(7). Model parameters, rates and constants were generally as described in Lin et al. (2011), except for n_I, n_C and V_I which have been adjusted to 0.06 min^{-1}, 0.06 min^{-1} and 4.0 L, respectively. These changes were made based on an analysis of results from several microdialysis studies and the population parameters from Van Cauter et al. (1992).

Endogenous insulin secretion was modelled as a function of BG. Sub-model parameters, u_{min} and u_{max} are 16.7 mU.min^{-1} and 266.7 mU.min^{-1}, respectively. For non-diabetic patients, k_1 and k_2 take the values 14.9 mU.L.mmol^{-1}.min^{-1} and -49.9 mU.min^{-1}.

The model was implemented in MATLAB (2011b, Natick, MA), and a value of *SI* identified each hour for every patient using clinical BG, insulin and nutrition records. The parameter identification was performed using an integral method that ensured the global optimum value was located (Hann et al., 2005). Hour-to-hour *SI* variability is defined in (8).

$$\dot{G}(t) = -p_G G(t) - \mathrm{SI}.\, G(t)\frac{Q(t)}{1+\alpha_G Q(t)} + \frac{P(t) + EGP - CNS}{V_G} \quad (1)$$

$$\dot{Q}(t) = n_I\big(I(t) - Q(t)\big) - n_C\frac{Q(t)}{1+\alpha_G Q(t)} \quad (2)$$

$$\dot{I}(t) = n_K I(t) - n_L\frac{I(t)}{1+\alpha_I I(t)} - n_I\big(I(t) - Q(t)\big) + \cdots \frac{u_{ex}(t)}{V_I} + (1-x_L)\frac{u_{en}(G)}{V_I} \quad (3)$$

$$P(t) = \min(d_2 P2,\ P_{\max}) + PN(t) \quad (4)$$

$$\dot{P1}(t) = -d_1 P1 + D(t) \quad (5)$$

$$\dot{P2}(t) = -\min(d_2 P2,\ P_{\max}) + d_1 P1 \quad (6)$$

$$u_{en}(G) = \min\left(\max\left(u_{min}, k_1 G(t) + k_2\right), u_{max}\right) \quad (7)$$

The model-based insulin sensitivity parameter (*SI*) has been shown to correlate well with the insulin sensitivity index (ISI) determined by the gold-standard euglycaemic-hyperinsulinaemic clamp (r >0.90) (Lotz et al., 2008).

$$\Delta\%SI_k = 100 \times \frac{(SI_{k+1} - SI_k)}{SI_k} \quad (8)$$

2.3 BG sensor error

The point-of-care glucometer used in the Christchurch Hospital ICU is the Arkray Super-Glucocard II (Arkray Inc., Japan). Sensor bias was determined from the correlation data reported for the glucometer test strips, (Arkray, 2007) and was therefore known at all BG concentrations. Precision was reported as a coefficient of variation (CV) at three BG concentrations, 4.3, 6.9 and 21.0 mmol/L. For this analysis, the CV was linearly interpolated for BG concentrations within the reported range and held constant outside this range. These data were used to construct an additive error model for the glucometer used in this investigation. Table 2 presents the bias and precision components for a range of glycaemia.

Table 2. Error components of the Arkray Super-Glucocard II glucometer (Arkray, 2007).

Blood glucose (mmol/L)	2.0	4.3	6.9	10.0	21.0	30.0
Bias (%)	+7.9	+2.1	+0.2	-0.8	-2.0	-2.3
Precision, CV (%)	3.5	3.5	2.8	2.8	2.7	2.7

The data used in the study was collected by trained clinical staff, minimising the potential for additional error through device misuse (Bergenstal, 2008). Blood samples tested were typically arterial, although, when an arterial line was not present, capillary blood was used.

2.4 Timing error

Measurements and interventions during the SPRINT protocol were 1 or 2-hourly and intended to be taken on the hour. These measurements were recorded by hand and attributed to the nearest hour on the standard paper 24-hour charts used in the Christchurch Hospital ICU. Hence, any discrepancies between the actual measurement time and the 'nearest hour' were lost.

Recent pilot trials of the STAR (Stochastic TARgeted) protocol at Christchurch Hospital ICU (Evans et al., 2011) provide data to generate a timing error model (1651 measurements on 20 patients). The STAR protocol is implemented on a tablet computer, thus the exact time when BG measurements are entered is recorded and can be compared to the written records. Using the discrepancies between scheduled and actual BG measurements, a model of timing error can be generated and applied to data from the SPRINT protocol. Although the STAR protocol differs from SPRINT, particularly with its computerised interface, it is used by the same clinical staff, in the same unit. Hence, it may be assumed that timing errors in making measurements will be similar.

Timing errors were limited to a maximum of 20 minutes. The empirical error distributions are shown in Figure 1. Errors from these distributions were applied additively to the SPRINT data by randomly sampling from the error vectors. The errors were applied to both the measurement and intervention timing. Thus, the measurements and interventions remained synchronised, as they would in the hospital.

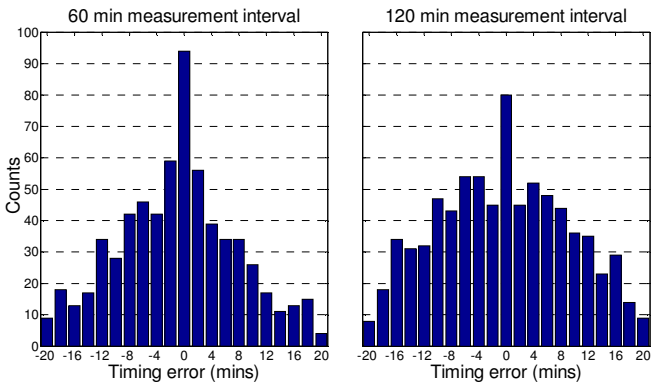

Figure 1. Timing error models based on data from the STAR pilot trials (Evans et al., 2011). Errors from 1- and 2 hour measurements are shown on the left and right, respectively.

2.5 Analyses

To assess the effects of random timing and sensor errors on *SI*, Monte Carlo simulations were performed. The *SI* profile of each patient in the cohort was refitted n = 100 times with randomly sampled errors applied to the observed timing and BG concentrations. The *SI* profiles identified without additional random errors were considered the 'true' profiles, and the Monte Carlo profiles were compared to these to assess the impact each of the sources of error, both individually and combined. Comparisons of both *SI* level and hour-to-hour variability were made.

To facilitate comparisons when timing errors were applied, *SI* was identified in 60-minute intervals, rather than between BG measurements. This use of fixed, 60-minute fitting intervals is consistent with the methodology used for glycaemic control by the STAR protocol.

To analyse the impact of errors on the identified *SI* level, the variation induced by the simulated errors at each hour was assessed across the Monte Carlo simulations. To characterise the distribution of percentage differences in *SI* level at each hour, between the true and simulated profiles, the width of the interquartile range (IQR) was used. Figure 2 illustrates the methodology for *SI* level comparisons between the n = 100 Monte Carlo simulations and the true data.

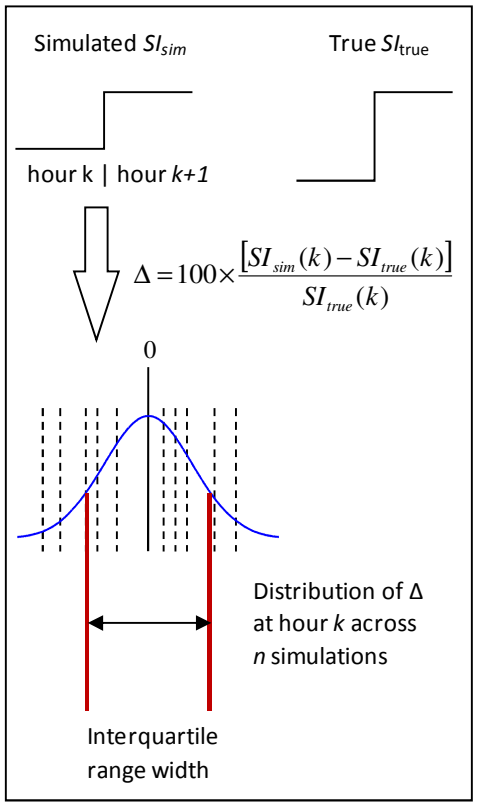

Figure 2. *SI* level comparison method for the Monte Carlo simulations with added sensor and timing error. The width of the interquartile range of differences was used to characterise the variability in level induced by the errors.

This analysis method resulted in one 'IQR width' per patient hour. For the 270 patient SPRINT cohort, there were approximately 43,000 hours of data. These 43,000 IQR widths were presented as cumulative distributions to show the overall effect of the errors on the cohort.

The hour-to-hour variability of simulated data could not be compared to the true variability using ratios, as the centre of the variability distribution for any given patient was approximately zero. Therefore, many values were close to zero, causing comparison ratios to approach zero or infinity, providing an effectively meaningless metric. To quantify the variability of hour-to-hour variability caused by the errors, the width of the IQR across the simulations for each hour was used. This method relies on the assumption that the median value across the simulations for each hour is close to zero.

A link to the actual variability is provided by the distribution of 'ΔMedian *SI*.' For any given hour, this metric is defined as the difference between the median hour-to-hour variability (%), across the n simulations, and the actual value (%). Typically, the distribution was tightly centred about zero, justifying the use of simulation IQR width without explicit reference to the actual variability.

3. RESULTS & DISCUSSION

3.1 Timing error

Figure 3 shows the impact of timing errors on identified *SI* level (left panel) and variability (right panels). For 95% of hours, the IQR width of *SI* level was less than 12.4%. Thus, for those 40,000 hours, half the simulations resulted in *SI* values within approximately ±6.2% of the true value, assuming a symmetrical distribution. Similarly, for variability, 95% of hours had an IQR width of hour-to-hour changes of less than 17.8%, or ±8.9% about the simulation median. The top right panel of Figure 3 shows the simulation median was typically very close to the true value for variability, justifying the assumptions made.

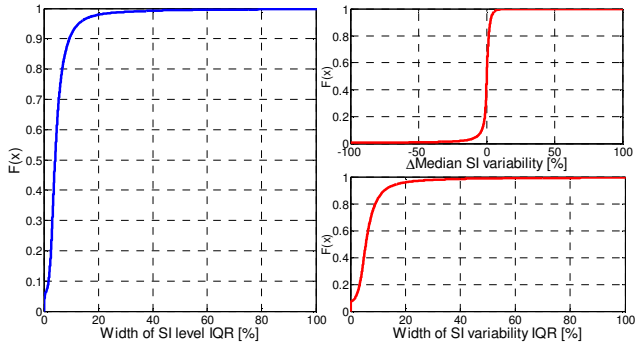

Figure 3. The impact of timing error on *SI* level (left panel) and hour-to-hour variability (right panels), determined by Monte Carlo simulation. The panels on the right show the location of the median simulated hour-to-hour variability, compared to the actual (top) and the variability about that median (bottom).

These results show that typical timing errors in isolation have relatively small impact on the level and variability of *SI*. With a median absolute difference between the simulated and actual measurement intervals of 9 minutes and using bolus insulin delivery, this result is not too surprising. Unlike infused insulin, bolus delivery ensures that the entire prescribed dose is always administered, regardless of the time between measurements. In addition, timing discrepancies only affect the later parts of the interstitial insulin profile, where concentrations are lowest and thus contribute least to the area under the curve used in fitting the *SI* parameter (Hann et al., 2005).

3.2 BG sensor error

Figure 4 shows the impact of BG sensor errors on *SI* level (left panel) and variability (right panels). The variability induced in both *SI* level and variability is greater than that due to timing error. The increases at the median values for level and variability compared to timing error of Figure 3 were 4.9% and 2.2%, respectively ($p = 0$, Wilcoxon rank-sum test). However, with so many data points, a statistically significant difference is almost guaranteed.

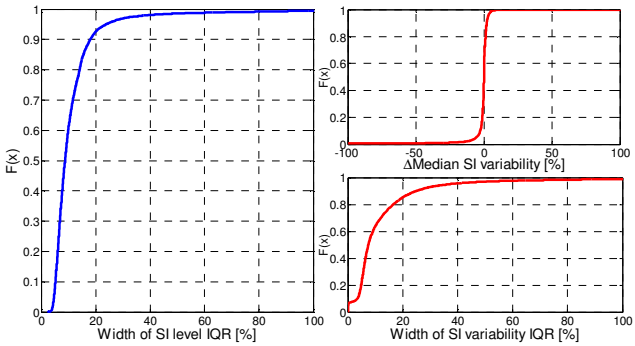

Figure 4. The impact of BG sensor error on *SI* level (left panel) and hour-to-hour variability (right panels), determined by Monte Carlo simulation. The panels on the right show the location of the median simulated hour-to-hour variability, compared to the actual (top) and the variability about that median (bottom).

The non-zero minimum width of *SI* level IQR evident in the left panel of Figure 4 is the result of the sensor bias. This minimum width of 2.7% was not present with the zero-mean timing error simulations of Figure 3, or if the bias is set to zero (results not shown).

3.3 Combined measurement error

Figure 5 shows the impact of the combined timing and BG sensor errors on *SI* level (left panel) and variability (right panels). The previous two sections have characterised the individual contributions of timing and sensor error. This analysis combines them, simulating errors seen in the real, clinical situation.

For 95% of hours, the IQR width of *SI* level was less than 23.9%. Thus, assuming a symmetrical distribution, half the simulations resulted in *SI* values within approximately ±12% of the true value. Similarly for variability, the 95th percentile was 34.9%, indicating that for half the simulations the hour-to-hour variability of *SI* was within ±17.5%.

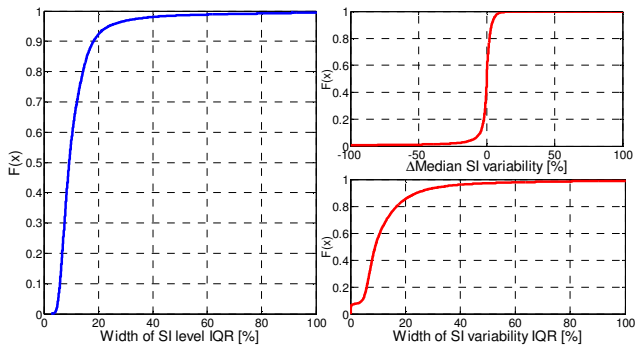

Figure 5. The impact of combined timing and BG sensor error on *SI* level (left panel) and hour-to-hour variability (right panels), determined by Monte Carlo simulation. The panels on the right show the location of the median simulated hour-to-hour variability, compared to the actual (top) and the variability about that median (bottom).

3.4 Implications of results

The clinical impact during glycaemic control of changes to *SI* level induced by sensor and timing errors is likely to be limited. Identified values of *SI* were typically within 12% of the true value when influenced by both sources of error. In contrast, changes in *SI* greater than 20% were seen with glucocorticoid treatment (Pretty et al., 2010) and improving patient condition over the first 18 hours of ICU stay (Pretty et al., 2012).

A second, potentially clinically significant, impact is on analytical use of *SI* as a marker of injury or change in state. Much of the true hour-to-hour change of this SPRINT cohort was smaller than the variability induced by sensor and timing errors. The IQR of true hour-to-hour variability across the entire cohort was -9.7% to 13.9% and 63% of all values lay in the range ±17.5%. Thus, short-term changes in SI may be a result of measurement timing or sensor errors rather than a true physiological phenomenon. Hence, using changes in *SI* level as a diagnostic must be done with caution, potentially looking at longer-term trends, where the effects of random errors may be cancelled by averaging over time.

In the context of the STAR protocol, the increase in hour-to-hour variability may be clinically significant. STAR uses a stochastic model of expected *SI* hour-to-hour variability to forecast the results of potential interventions and avoid hypoglycaemia. The relatively large, ±17.5%, range of hour-to-hour variability about the median caused by errors, suggests that a significant proportion of this expected variability may be dominated by these errors, rather than physiological variability. Hence, to reduce the impact of these errors, the 5-95 range of the stochastic model should be used. *SI* values at these more extreme percentiles are less likely to be influenced by random sensor or timing errors.

3.5 Potential for reducing error

There is no effective way to reduce the impact of these errors as they are random and apply equally to all patients. The only available option is to reduce the magnitude of the errors. The timing error distribution shown in Figure 1 shows that more than 85% of measurements are within 10 minutes of the scheduled time, which is a very good result in a busy ICU environment. In contrast, BG sensor errors can be reduced with better, more accurate, but likely more expensive equipment.

To compare the impact of sensor errors from glucometers with a state of the art blood gas analyser (BGA), an error model was developed for the Bayer RapidLab 860 (Bayer Diagnostics, Tarrytown, NY) (2004, Peet et al., 2002). Errors for this BGA consist of very little bias (≤ 0.06 mmol/L) and CV in the range 1.7%-4.9%, depending on BG concentration. The RapidLab 860 is used in the Christchurch Hospital Neonatal ICU (Le Compte, 2009), and thus represents a realistic option for the adult ICU in Christchurch. An n = 50 run Monte Carlo simulation was performed with both timing and BG sensor errors from the BGA for a basic comparison.

The results show only a minor improvement by using the BGA. The 95th percentile of IQR widths reduced from 23.9% to 21.0% and from 34.9% to 30.2% for level and variability, respectively. These results appear surprising, but are probably due to the relatively low reported errors of the glucometer. Data for the error model was taken from the manufacturer's data-sheet (Arkray, 2007). However, a published report (Solnica and Naskalski, 2005) failed to recreate these results, possibly due to user error, sample type, abnormal hematocrit or interfering substances (Bergenstal, 2008).

4. CONCLUSION

The objective of this study was to assess and quantify the impact of typical timing and BG sensor errors on the level and variability of model-based *SI*. Specifically, the variability of level and the variability of *SI* hour-to-hour variability were investigated under the influence of these sources of error, both separately and combined. Measurement timing errors were shown to have a relatively small impact on the *SI* parameter. The BG concentration errors of the Arkray glucometer had a larger effect on *SI* and tended to dominate the combined analysis.

The results of this study indicate that the impact of errors on *SI* level is unlikely to be clinically significant and probably overshadowed by physiological or therapy factors arising from the critical condition of the patients. In contrast, the impact of errors on hour-to-hour *SI* variability is more pronounced and may impact the way the *SI* parameter is utilised for control and analysis.

This analysis indicated that for a given *SI* level, variability induced by errors might dominate the IQR of the probability density function describing *SI* for the subsequent hour. Hence, to avoid inadvertently basing control decisions on this artificial variability, the 5-95 percentile range of the stochastic model should be used. In addition, using changes in *SI* level as a diagnostic must be done with caution, potentially looking at longer-term trends, where the effects of random errors may be cancelled by averaging over time.

Given the random nature of these errors, the only feasible method of mitigation is to use BG sensors with smaller errors. However, a comparison with results from a state of the art, clinical blood gas analyser showed that the limited improvement in performance probably doesn't justify the significantly increased cost of the device. However, understanding the effect of sensor and timing errors on *SI* allows their impact to be taken into account when using the parameter for control and analysis.

5. REFERENCES

Abbott 2010. Optium Blood Glucose Test Strip datasheet. UK: Abbott Diabetes Care Ltd.

Anon 2004. MHRA Blood Gas Analyser Survey 2004:

RapidLab 800 Series. Guildford Medical Device Evaluation Centre.

Arkray 2007. Glucocard™ Test Strip 2 Data Sheet. Japan: Arkray Inc.

Bergenstal, R. M. 2008. Evaluating the accuracy of modern glucose meters. *Insulin,* 3**,** 5-14.

Chase, J. G., Shaw, G., Le Compte, A., Lonergan, T., Willacy, M., et al. 2008. Implementation and evaluation of the SPRINT protocol for tight glycaemic control in critically ill patients: a clinical practice change. *Critical Care,* 12**,** R49.

Evans, A., Shaw, G. M., Le Compte, A., Tan, C. S., Ward, L., et al. 2011. Pilot proof of concept clinical trials of Stochastic Targeted (STAR) glycemic control. *Ann Intensive Care,* 1**,** 38.

Hann, C. E., Chase, J. G., Lin, J., Lotz, T., Doran, C. V., et al. 2005. Integral-based parameter identification for long-term dynamic verification of a glucose-insulin system model. *Comput Methods Programs Biomed,* 77**,** 259-270.

Hovorka, R., Chassin, L. J., Ellmerer, M., Plank, J. & Wilinska, M. E. 2008. A simulation model of glucose regulation in the critically ill. *Physiol Meas,* 29**,** 959-78.

Le Compte, A. 2009. *Modelling the Glucose-Insulin Regulatory System for Glycaemic Control in Neonatal Intensive Care.* PhD thesis, University of Canterbury.

Lin, J., Razak, N. N., Pretty, C. G., Le Compte, A., Docherty, P., et al. 2011. A physiological Intensive Control Insulin-Nutrition-Glucose (ICING) model validated in critically ill patients. *Comput Methods Programs Biomed,* 102**,** 192-205.

Lotz, T. F., Chase, J. G., Mcauley, K. A., Shaw, G. M., Wong, X. W., et al. 2008. Monte Carlo analysis of a new model-based method for insulin sensitivity testing. *Comput Methods Programs Biomed,* 89**,** 215-25.

Peet, A. C., Kennedy, D. M., Hocking, M. D. & Ewer, A. K. 2002. Near-patient testing of blood glucose using the Bayer Rapidlab 860 analyser in a regional neonatal unit. *Ann Clin Biochem,* 39**,** 502-8.

Pretty, C., Chase, J. G., Lin, J., Shaw, G. M., Le Compte, A., et al. 2010. Impact of glucocorticoids on insulin resistance in the critically ill. *Computer Methods and Programs in Biomedicine***,** 2-10.

Pretty, C., Le Compte, A., Chase, J. G., Shaw, G. M., Preiser, J. C., et al. 2012. Variability of insulin sensitivity during the first 4 days of critical illness: Implications for tight glycaemic control. *Annals of Intensive Care,* In Press.

Roche 2007. ACCU-CHEK Comfort Curve Test-strip packet insert. USA: Roche Diagnostics Ltd.

Roche 2008. Evaluation Report of the Accu-Chek Comfort Curve Test Strip as a Plasma-like Test Strip. Roche Diagnostics, Mannheim, Germany.

Solnica, B. & Naskalski, J. W. 2005. Quality control of SMBG in clinical practice. *Scand J Clin Lab Invest Suppl,* 240**,** 80-5.

Solnica, B., Naskalski, J. W. & Sieradzki, J. 2003. Analytical performance of glucometers used for routine glucose self-monitoring of diabetic patients. *Clin Chim Acta,* 331**,** 29-35.

Van Cauter, E., Mestrez, F., Sturis, J. & Polonsky, K. S. 1992. Estimation of insulin secretion rates from C-peptide levels. Comparison of individual and standard kinetic parameters for C-peptide clearance. *Diabetes,* 41**,** 368-77.