Validation of a Model-based Virtual Trials Method for Tight Glycaemic Control in Intensive Care

Fatanah Suhaimi *, J. Geoffrey Chase *, Aaron J. Le Compte *, Jean-Charles Preiser**
Jessica Lin *, Geoffrey M. Shaw***

* Dept of Mechanical Eng, University of Canterbury, Christchurch, New Zealand.
(Tel: +64-3-364-2987; e-mail:fatanah.suhaimi@pg.canterbury.ac.nz, geoff.chase@canterbury.ac.nz.)
** Dept of Intensive Care, Centre Hospitalier Universitaire de Liege (CHU), Liege, Belgium.
(e-mail:jean-charles.preiser@chu.ulg.ac.be)
*** Department of Intensive Care, Christchurch Hospital, Christchurch, New Zealand.   -

Abstract: In-silico virtual trials offer significant advantages in cost, time and safety. However, no such method has been truly validated with clinical data. This study tests 2 matched cohorts from an independent ICU treated with 2 different glycaemic control protocols. The goal is to validate the in-silico virtual trials model and methods, including the underlying assumptions. A retrospective analysis used records from a 211 patient subset of the Glucontrol trial in Liege, Belgium. Glucontrol-A (N = 142) targeted a BG range of 4.4-6.1 mmol/L and Glucontrol-B (N = 69) targeted 7.8-10.0 mmol/L. Cohorts were matched by APACHE II score, but the Glucontrol A cohort was slightly older (p = 0.0352). Results showed high correlation between self- and cross-validation virtual trials and clinical results. The virtual trials models and methods are thus validated on independent data.

Keywords: Intensive Care, in-silico trials, virtual trials, hyperglycaemia, model, validation.

1. INTRODUCTION

Stress-induced hyperglycaemia and high insulin resistance are prevalent in critical care. Hyperglycaemia worsens outcomes, increasing the risk of severe infection, myocardial infarction (Capes et al. 2000), and critical illness polyneuropathy and multiple organ failure (Van den Berghe et al. 2001). Hence, hyperglycaemia is strongly associated with increased mortality.

Van den Berghe et al (2001) showed that tight glucose control (TGC) reduced intensive care unit (ICU) patient mortality up to 45% using a target of 6.1mmol/L. Other studies with similar or higher targets have successfully reduced mortality (Krinsley 2004; Chase et al. 2008). Hence, the data indicate that a control algorithm that safely provides TGC to reduce hyperglycaemia and glycaemic variability has the potential to reduce mortality and cost (Krinsley and Jones 2006).

In this study, "virtual trials" are performed using a clinically validated model (Chase et al. 2007) of the glucose-insulin system. Insulin sensitivity, S_I, is used as the critical marker of a patient's metabolic state and is assumed independent of the model inputs. Virtual trials can be used to simulate a TGC protocol using a S_I profile identified from clinical data and different insulin and nutrition inputs. Virtual trials enable the rapid testing of new TGC intervention protocols and analysing control protocol performance. They are thus a means of safely optimising protocols prior to implementation.

Virtual patient trials have been used in design of TGC protocols (Chase et al. 2007). The clinical results of SPRINT (Chase et al. 2008) showed very close agreement to expected results from simulation (Chase et al. 2007). However, SPRINT was implemented in Christchurch Hospital ICU and was designed using clinical data and virtual patients from the same unit.

Thus, the performance of virtual trials on separate cohorts, independent of the ICU used to generate the virtual patients, has not yet been evaluated. In addition, the virtual trials assumption of the model independence of a virtual patient's insulin sensitivity profile from intervention inputs, identified from patient data and the system model, has never been validated.

In this study, the Glucontrol study (Preiser et al. 2009) provides a source of patient data independent of the Christchurch ICU. Only data from one Glucontrol centre (University Hospital of Liege, Belgium) were used. Self-validation and cross-validation virtual trial simulations are used to assess model error and validate the overall virtual trials approach.

2. METHODS

2.1 Glucontrol Protocol and Patient Cohorts

The Glucontrol trial (Preiser et al. 2009) randomised patients into two groups: A and B. Group A received intensive insulin therapy and Group B received

Table 1. Glucontrol A and B comparison (Median [IQR]). P-values computed using chi-squared and Mann-Whitney tests.

Cohort	A	B	*P* value
Baseline Variables			
Number	142	69	
Percent male (%)	64.8	56.5	0.25
Age	71 [61-80]	69 [53-77]	0.035
Weight (kg)	72 [62-85]	75 [68-81]	0.38
BMI	25.4 [22.6-29.3]	26.0 [23.2-29.3]	0.46
APACHE II	17 [14-22]	17 [14-21]	0.76
Initial BG	6.6 [5.56-8.56]	6.6 [5.65-9.36]	0.58
Glucose Control			
Total hours	16, 831	12, 946	
BG measures	4, 571	2, 820	
BG (mmol/L)	6.3 [5.3-7.6]	8.2 [6.9-9.4]	
Insulin (U/hr)	1.5 [0.5-3.0]	0.7 [0.0-1.7]	
All glucose admin (mmol/min)	0.30 [0.00-0.90]	0.60 [0.10-1.00]	

$$\dot{G} = -p_G G - S_I G \frac{Q}{1 + \alpha_G Q} + \frac{P(t) + EGP_{max} - CNS}{V_G(t)} \tag{1}$$

$$\dot{Q} = -kQ + kI \tag{2}$$

$$\dot{I} = -\frac{nI}{1 + \alpha_I I} + \frac{u_{ex}(t)}{V_I} + e^{-k_I U_{ex}(t)} I_B \tag{3}$$

Where: $G(t)$ is the total plasma glucose, $I(t)$ is the plasma insulin, and $Q(t)$ is the effect of previously infused insulin being utilized over time. EGP_{max} is the theoretical maximum endogenous glucose production (EGP) for a patient, which is suppressed with increasing glucose concentrations. This suppression, independent of non-insulin mediated glucose uptake by the central nervous system (CNS) is captured by the term p_G. In contrast, patient-specific insulin mediated glucose removal is captured with insulin sensitivity, S_I, which is identified from clinical data as a time-dependent variable that reflects evolving patient condition. Exogenous inputs are glucose appearance $P(t)$ from the carbohydrate content of nutrition infusions via a two compartment model (Wong et al. 2006) and intravenous insulin administration $u_{ex}(t)$. The remaining parameters are physiologically defined population constants for transport rates (n, k), saturation

conventional insulin therapy, with target BG ranges of 4.4-6.1 mmol/L and 7.8-10.0 mmol/L respectively. The protocols are defined in (Preiser et al. 2009). Insulin was administered as a continuous IV infusion. Hourly BG measurements were taken when the glycaemic level was not within the target range. Otherwise, 2-hour measurements were taken in the case of limited variation of glycaemia, defined as less than a 50% change from the previous glycaemia in 2-hour range. Finally, 4-hour measurements were taken when the glycaemic level was less than 50% of the highest glycaemia of the four last hours. The protocol specified insulin infusion rates whereas nutritional input was left to local and/or clinician standards, and was not explicitly considered in the Glucontrol study.

In this study, data was used from 350 patients treated using the Glucontrol protocol at CHU de Liege, Belgium, between March 2004 and April 2005. Patients were eliminated from the analysis if they received no insulin for their entire stay (per protocol), had less than 5 BG measurements or received little or no (recorded) carbohydrate administration (in any form) for more than 48 hours of their stay. Clinical details of the resulting Glucontrol cohorts are in Table 1. Patients in Group A were slightly older than Group B. However, there were no significant differences in weight, BMI or severity of illness as measured by APACHE II score.

2.2 Glucose-Insulin System Model

The analysis of patient-specific insulin sensitivity uses a glucose insulin system model that has been clinically validated in several studies (Wong et al. 2006; Chase et al. 2007):
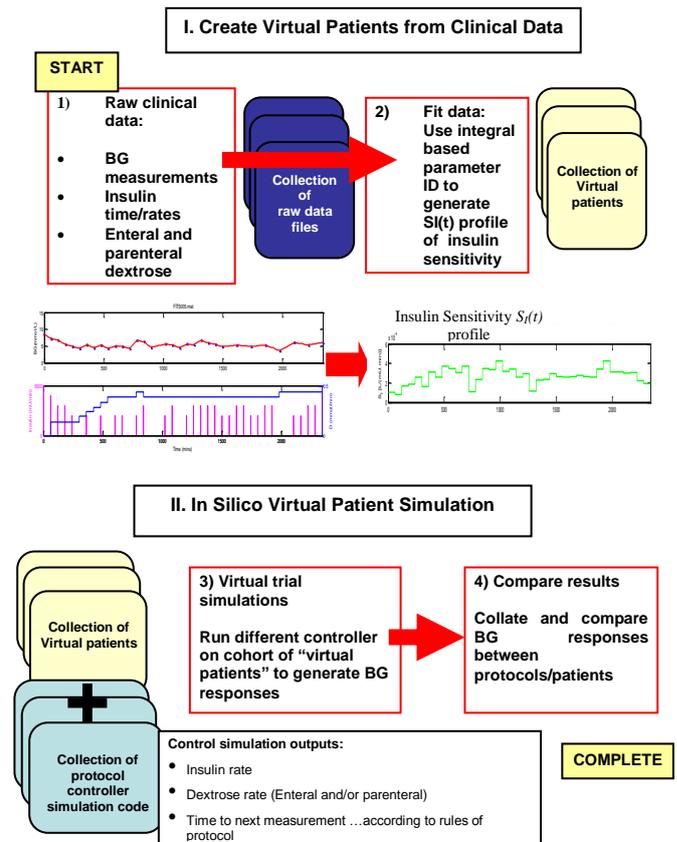

Fig. 1. Virtual patient development and in-silico simulation method. (I) Clinical data is used for fitting insulin sensitivity profiles to create 'virtual patients'. (II) These virtual patients can be used for simulating different protocols.
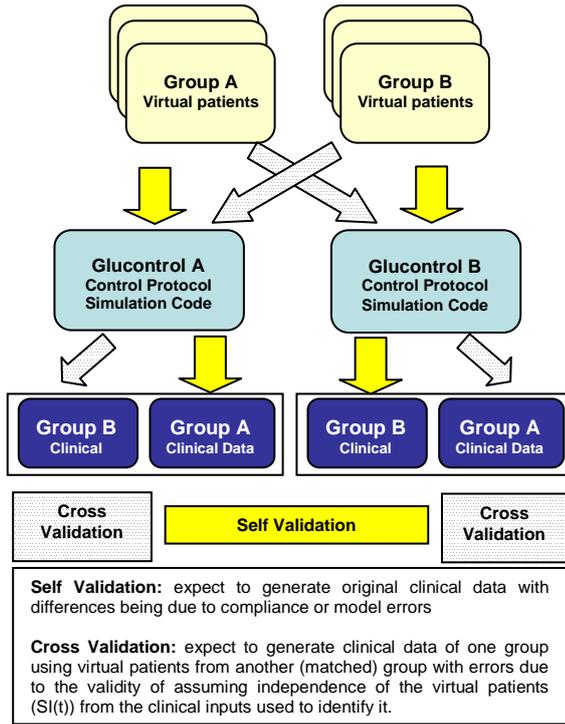
Fig. 2. Virtual trial validation method. Glucontrol A and Glucontrol B virtual trials are compared to the appropriate clinical results.

parameters α_G, α_I), endogenous insulin secretion (I_B, k_I) or volumes (V_G, V_I).

The essential parameter that drives the observed patient-specific glycaemic response to insulin and nutrition inputs is insulin sensitivity, S_I. The S_I parameter is identified by fitting the model to retrospective blood glucose measurements, and insulin and carbohydrate administration inputs from clinical data for each protocol. The resulting insulin sensitivity profile, $S_I(t)$, varies hourly.

2.3 Virtual Trials

The "virtual trial" method is used to simulate a clinical trial using patient specific data such as insulin sensitivity, S_I. The insulin sensitivity profile $S_I(t)$ captures a patient's time varying glycaemic response to insulin and nutrition inputs and is assumed to be an intrinsic property of the patient, independent of their glycaemic control interventions. This profile can then be used to simulate the blood glucose response to other combinations of insulin and dextrose inputs specified by a modified TGC protocol to obtain new glycaemic responses. Hence, an expected blood glucose profile can be generated for each patient to simulate glycaemic response to a specific protocol. Thus, virtual trials can be used to analyse, in-silico, the effect of different protocols on patient performance.

In this analysis, 2 groups of virtual patients are created from clinical data: Groups A and B, defined by their treatment with the Glucontrol A (intensive) or Glucontrol B (conventional) protocol.

This study performs two major forms of validation using virtual trials, as described schematically in Figs. 1-2:

1. **Self validation:** This test assesses the ability of the in-silico virtual trials to reproduce the clinical data. Simulation results are compared to clinical results. Differences can be ascribed to model errors and/or lack of perfect compliance in the clinical study versus the perfect compliance and timing in-silico.

2. **Cross validation:** This test assesses the assumption that the $S_I(t)$ insulin sensitivity profiles accurately capture patient dynamics, independent of the insulin and nutrition inputs used to create them. Hence, it tests the A protocol on B virtual patients, and the B protocol on A virtual patients. Since both cohorts are matched clinically, differences can be ascribed to failures in this critical, underlying independence assumption behind this virtual trials method.

These two tests provide both per-patient and cohort-wide validation of this in-silico approach.

1) Self-Validation: Self validation tests the ability to recreate the clinical data given the cohort of virtual patients and protocol definitions. For the self validation on Glucontrol A, the Glucontrol A protocol is simulated on Group A virtual patients, and these virtual trial results are compared to the clinical data of Group A. This step was repeated for self validation on Glucontrol B. Two virtual trials were simulated on each group considering a) the actual measurement timing used in retrospective trials (actual measurement) and b) actual measurement as suggested from the protocol (per protocol). Use of actual vs. per-protocol measurement timing allows one measure of compliance error.

2) Cross-Validation: Cross validation uses the matched A and B cohorts to determine the ability of the modelling method to reproduce the clinical data on a matched, but independent, cohort. For example, protocol A is simulated on virtual patients derived from Group B clinical data. The results of this test are then compared to the clinical data of Glucontrol Group A. Similarly, protocol B is test on virtual patients of Group A and the results are compared to Group B clinical data. If patients were perfectly matched the in-silico results between groups would also match.

Differences using large matched cohorts can thus be ascribed to how well the assumption holds that these virtual patient $S_I(t)$ profiles are independent of the clinical inputs used to derive them. If cross validation results match the clinical results well for clinically matched cohorts, then this assumption can be considered valid.

3. RESULTS

Fig. 3 shows the CDF of measured blood glucose on a cohort basis, comparing:

1. The distribution of simulated BG from the virtual trials of the Glucontrol A protocol on virtual

patients of Group A (self validation) and Group B (cross validation). Both sets of results are compared to the clinical results of Group A.

2. Simulated BG distribution from the virtual trials of Glucontrol B protocol on virtual patients of Group B (self validation) and Group A (cross validation), compared to Group B clinical results.

Hence, Fig. 3 presents both the self and cross validation in-silico results for both the A and B virtual cohorts.

The breakdown of distributions shows a clear separation between the protocols for Glucontrol A and Glucontrol B for all combinations of simulations, as expected from the clinical results and trial design. The four distributions for the Glucontrol A protocol show particularly close agreement. The Glucontrol A clinical median cohort BG value of 6.2 mmol/L agrees well with the 6.0 mmol/L and 6.2 mmol/L medians for the self validation trials using retrospective and per-protocol BG timing respectively. The cross-validation median BG of 6.5 mmol/L is also in close agreement with the clinical result. Similar agreement between clinical and simulation data are also present in the per-patient results in Table 2.

The four BG distributions in Fig. 3 for the Glucontrol B protocol shows a slightly greater spread in results, particularly below the Group B target of 8 mmol/L. However, the median cohort clinical BG value of 8.1 still agrees well with the medians of 8.5 and 8.7 mmol/L for Glucontrol B self validation with actual measurement and per-protocol measurements respectively, and also agrees well with the cross-validation median result of 8.5 mmol/L.

Finally, the wider error below 8 mmol/L may be due to the fact that the Glucontrol B protocol requires zero exogenous insulin below its target. Hence, model error grows due to the population-constant fixed endogenous insulin rate assumed in this situation (I_B in Equation (3)).

Table 2. Comparison of per-patient clinical results and virtual trial simulations (self-validation and cross validation) on Glucontrol A. Median [IQR] is used where appropriate.

Virtual Trials Simulation	Clinical A	GlucoA on A actual measurement	GlucoA on A per protocol	GlucoA on B per protocol
No. of patients	142	142	142	69
Per Patient				
Average insulin rate (U/hr)	1.4 [0.9 - 2.1]	1.9 [1.2 - 3.0]	2.6 [1.7 - 4.2]	4.6 [2.5 - 6.5]
Average glucose rate (g/hr)	1.1 [0.5 - 7.6]	1.1 [0.5 - 7.6]	1.1 [0.5 - 7.6]	2.9 [0.7 - 7.4]
BG (mmol/L)	6.4 [5.9 - 6.9]	6.1 [5.7 - 6.7]	6.3 [5.8 - 6.8]	6.6 [6.0 - 7.0]
BG measures	4564	4564	9635	7414
Measurement frequency	6.52	6.52	13.87	13.79

Table 3. Comparison of per-patient clinical results and virtual trial simulations (self-validation and cross validation) on Glucontrol B. Median [IQR] is used where appropriate.

Virtual Trials Simulation	Clinical B	GlucoB on B actual measurement	GlucoB on B per protocol	GlucoB on A per protocol
No. of patients	69	69	69	142
Per Patient				
Average insulin rate (U/hr)	0.6 [0.3 - 1.2]	0.7 [0.4 - 1.3]	1.0 [0.4 - 1.9]	0.5 [0.1 - 1.2]
Average glucose rate (g/hr)	2.9 [0.7 - 7.4]	2.9 [0.7 - 7.4]	3.0 [0.7 - 7.7]	1.1 [0.5 - 7.5]
BG (mmol/L)	8.3 [7.6 - 8.8]	8.5 [8.1 - 8.8]	8.6 [8.3 - 9.2]	8.4 [8.0 - 8.9]
BG measures	2820	2820	4083	5436
Measurement frequency	5.23	5.23	7.62	7.86

Above 8 mmol/L, despite relatively low median insulin doses at 1 U/hr (Table 1), the agreement between BG distributions is much closer. Finally, as with the Group A results, the self- and cross-validation agreement is within measurement error and clinically insignificant at all likelihood (y-axis) levels of the CDF.

Fig. 4 shows the same results for the CDF of the median patient blood glucose levels across all patients in each group. This "per-patient" comparison has the same whole-cohort trend in Fig.3. Interestingly, and as with the cohort results, the largest gap is between the self validation and clinical data for Glucontrol B using retrospective
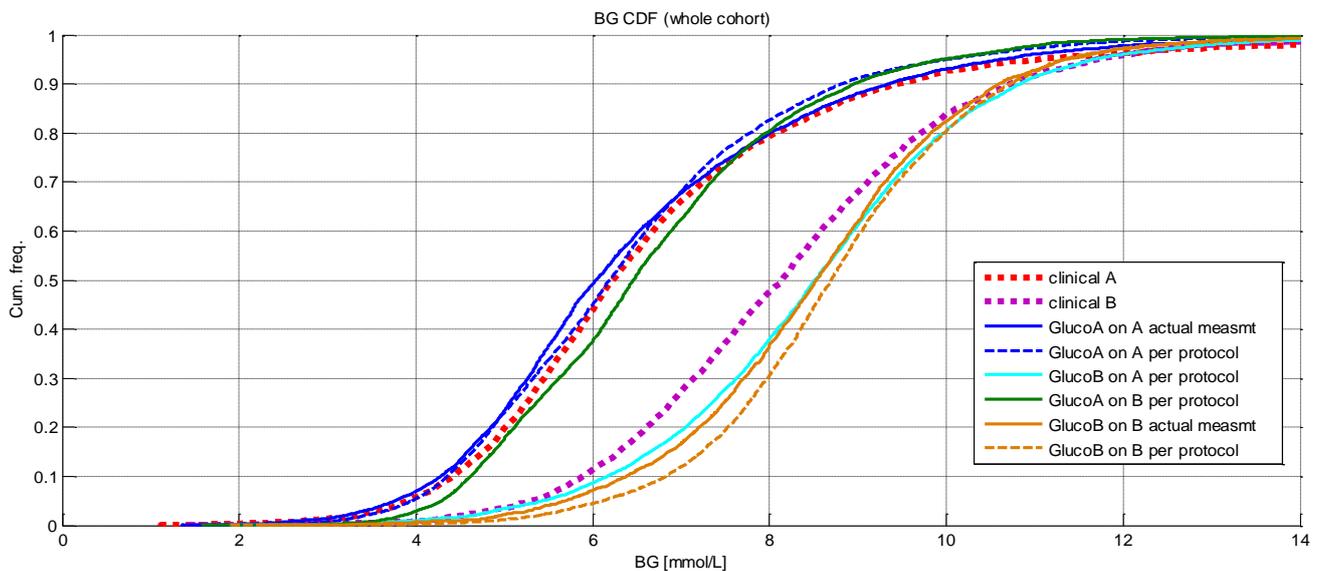

Fig. 3. CDF of blood glucose levels of clinical Glucontrol data versus virtual trials on a cohort basis.
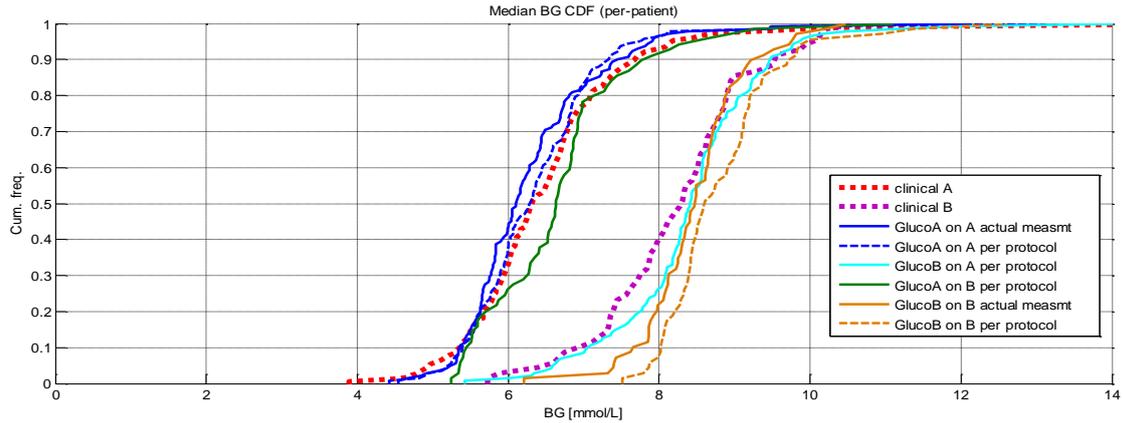
Fig. 4. CDF of median blood glucose measured of Glucontrol clinical data versus virtual trial on a per patient basis.

measurement timing. This result indicates the possibility of compliance error, whereas the difference between self-validation using per-protocol timing to the clinical data provides a combination of model error and compliance error.

Table 2 shows the comparison of clinical trials to the self validation and cross validation on Glucontrol A. Per patient results show a very close agreement between self validation per protocol to the clinical data given that the rate of insulin are almost doubled. For the cross validation, Glucontrol A protocol required higher rates of insulin for Group B, almost 3x compared to the clinical data. Comparison of clinical trials with self validation and cross validation on Glucontrol B is summarised in Table 3. Self validation results show close agreement to the clinical result and for cross validation, Group A insulin requirements are reflected by lower nutrition and higher target BG following the Glucontrol B protocol. Similarly, virtual trials of Glucontrol B per protocol have higher measurement frequency compared to the actual measurement. The differences in BG measurement frequency for clinical and self validation per protocol suggest that the protocol was not being followed clinically.

4. DISCUSSION

This paper focuses on the Glucontrol protocol from one centre (Liege, Belgium; pilot centre). Glucontrol was a multi-centre study stopped early to a high rate of unintended protocol violations (Preiser et al. 2009). This clinical data was independent from the Christchurch Hospital ICU data used in prior development and validation of the model used here. More importantly, there are 2 cohorts matched by severity of illness, weight and sex, which had significantly different glycaemic targets and glycaemic control therapies.

The virtual trials results are close to clinical results in all cases for both self validation and cross validation (Figs. 4 and 5). Referring to the same figures, the obvious separation between two protocols indicates the inter-protocol differences are, as expected, much larger than any inter-group differences. More importantly, the close correlation of self and cross validation results to clinical data validates the idea that these in-silico virtual trial simulations can accurately predict the expected clinical results of a protocol prior to clinical implementation using such virtual patients.

The results in Figs. 3 and 4 illustrate some variation between clinical data and virtual trials. In particular, Glucontrol A simulation results are closer to clinical data compared to Glucontrol B data. The major difference is that protocol B uses much less insulin given its higher glycaemic target. Therefore, the impact of intrinsic and potentially variable patient-specific dynamics, such as endogenous insulin production (I_B and k_I), in Equation (3) are more pronounced with respect to the far lesser exogenous insulin given to Group B, especially at low blood glucose levels. As these metrics are, by necessity, assumed population constants, some of the Group B simulation errors may reflect errors in these population values.

Another likely cause is evident in Fig. 4, where the most mismatched line of the three results is for Glucontrol B using measurements as per-protocol. While no specifics were reported for the compliance issues in Glucontrol, the larger gap seen in the Glucontrol B self and cross validation results is primarily at lower glycaemic levels below the target range of 8.0 mmol/L, where nursing compliance on an otherwise well controlled patient could have been less. The fact that the clinical data is lower than the simulations could indicate non-compliance in timing or dosing, or simple overriding of the protocol recommendations by clinical staff. Protocol compliance and model error together form discrepancies between self validation and clinical results. There are many factors that lead to the protocol compliance, such as timing errors and faulty or non-compliant insulin administration. In contrast, computer simulations will always follow protocols exactly as instructed. Hence, the self validation error captures both model and compliance errors.

For the cross validation, protocol A on Group B is a very good match and close to the self validation results for Group A. In addition, protocol B on Group A virtual

patients is within the tolerance defined in the Group B self validation and closer to the slope and trends of the clinical data. These results indicate that the cross validation is within self validation model and compliance errors. Thus, the insulin sensitivity independence assumption appears to hold for these virtual trials, independently validating this concept and the virtual trial method based on this model.

Differences between self and cross validation results indicate remaining differences between patient groups despite clinical matching. More importantly, the relatively small differences show the strength of model-fitted insulin sensitivity as a description of patient metabolic state, rather than as a therapy-specific parameter value. The tight agreement in the cross validation results for cohorts matched in terms of clinical characteristics and metabolic dynamics (Fig. 3) adds weight to the use of these model-based insulin sensitivity profiles, and thus this model, to accurately generate expected clinical outcomes in testing new protocols.

Finally, this paper shows the potential for TGC to be readily optimised and implemented using model predictive control. Even though some TGC clinical trials have not achieved any benefit from TGC (Finfer et al. 2009), only 2 protocols have been first optimized with virtual trials (Chase et al. 2007; LeCompte et al. 2009). Both delivered safe, effective TGC with reduced or zero hypoglycaemia.

5. CONCLUSIONS

This paper presented the analysis and validation of an in-silico virtual patient and model-based virtual trials methodology. It takes advantage of a set of independent clinical data comprised of two clinically matched cohorts treated with two different TGC protocols with two different glycaemic targets. It has 2 main conclusions that can be drawn:

- The self validations indicated a clinically insignificant error in these virtual patient methods due to model and/or clinical compliance.
- The cross validations clearly show the virtual patients enabled by the identified patient-specific $S_I(t)$ profiles are independent of the clinical inputs used to generate these profiles.

Thus, the virtual patients and in-silico virtual trial methods presented are validated in their ability to accurately simulate, in advance, the clinical results of an independent TGC protocol, directly enabling rapid design and optimisation of safe and effective TGC protocols with confidence.

REFERENCES

Capes, S. E., D. Hunt, et al. (2000). "Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview." *Lancet* **355**(9206): 773-778.

Chase, J. G., G. Shaw, 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**(2): R49.

Chase, J. G., G. M. Shaw, et al. (2007). "Model-based insulin and nutrition administration for tight glycaemic control in critical care." *Curr Drug Deliv* **4**(4): 283-296.

Finfer, S., D. R. Chittock, et al. (2009). "Intensive versus conventional glucose control in critically ill patients." *N Engl J Med* **360**(13): 1283-1297.

Krinsley, J. S. (2004). "Effect of an intensive glucose management protocol on the mortality of critically ill adult patients." *Mayo Clin Proc* **79**(8): 992-1000.

Krinsley, J. S. and R. L. Jones (2006). "Cost analysis of intensive glycemic control in critically ill adult patients." *Chest* **129**(3): 644-650.

LeCompte, A., J. Chase, et al. (2009). "Blood Glucose Controller for Neonatal Intensive Care: Virtual trials development and 1st clinical trials." *Journal of Diabetes Science and Technology (JoDST)* **3**(5): 1066-1081.

Preiser, J., P. Devos, et al. (2009). "A prospective randomised multi-centre controlled trial on tight glucose control by intensive insulin therapy in adult intensive care units: the Glucontrol study." *Intensive Care Medicine* **35**(10): 1738-1748.

Van den Berghe, G., P. Wouters, et al. (2001). "Intensive insulin therapy in the critically ill patients." *N Engl J Med* **345**(19): 1359-1367.

Wong, X. W., I. Singh-Levett, et al. (2006). "A novel, model-based insulin and nutrition delivery controller for glycemic regulation in critically ill patients." *Diabetes Technol Ther* **8**(2): 174-190.