

# **A 3D INSULIN SENSITIVITY PREDICTION MODEL ENABLES MORE PATIENT-SPECIFIC PREDICTION AND MODEL-BASED GLYCAEMIC CONTROL**

Authors:

- Vincent Uyttendaele<sup>a,b</sup> [vincent.uyttendaele@pg.canterbury.ac.nz](mailto:vincent.uyttendaele@pg.canterbury.ac.nz) (corresponding author)
- Jennifer L. Dickson<sup>a</sup> [jennifer.dickson@canterbury.ac.nz](mailto:jennifer.dickson@canterbury.ac.nz)
- Kent W. Stewart<sup>a</sup> [kent.stewart@pg.canterbury.ac.nz](mailto:kent.stewart@pg.canterbury.ac.nz)
- Thomas Desaive<sup>b</sup> [tdesaive@ulg.ac.be](mailto:tdesaive@ulg.ac.be)
- Balazs Benyo<sup>c</sup> [bbenyo@iit.bme.hu](mailto:bbenyo@iit.bme.hu)
- Noemi Szabo-Nemedi<sup>d</sup> [noemi.szabo.nemedi@gmail.com](mailto:noemi.szabo.nemedi@gmail.com)
- Attila Illyes<sup>d</sup> [attila.illyes@diabet.hu](mailto:attila.illyes@diabet.hu)
- Geoffrey M. Shaw<sup>e</sup> [Geoff.Shaw@cdhb.health.nz](mailto:Geoff.Shaw@cdhb.health.nz)
- J. Geoffrey Chase<sup>a</sup> [geoff.chase@canterbury.ac.nz](mailto:geoff.chase@canterbury.ac.nz)

<sup>a</sup>*Department of Mechanical Engineering, University of Canterbury, Private Bag 4800, Christchurch, New Zealand.*

<sup>b</sup>*GIGA – In silico Medicine, University of Liège, Allée du 6 Août 19, Bât. B5a, 4000 Liège, Belgium.*

<sup>c</sup>*Budapest University of Technology and Economics, Department of Control Engineering and Information Technology, Budapest, Hungary*

<sup>d</sup>*Kalman Pandy County Hospital, Dept of Intensive Care, Gyula, Hungary*

<sup>e</sup>*Christchurch Hospital, Dept of Intensive Care, Christchurch, New Zealand and University of Otago, School of Medicine, Christchurch, New Zealand*

## 0.0 Abstract:

**Background:** Insulin therapy for glycaemic control (GC) in critically ill patients may improve outcomes by reducing hyperglycaemia and glycaemic variability, which are both associated with increased morbidity and mortality. However, initial positive results have proven difficult to repeat or achieve safely. STAR (Stochastic TARgeted) is a model-based glycaemic control protocol using a risk-based dosing approach. STAR uses a 2D stochastic model to predict distributions of likely future changes in model-based insulin sensitivity ( $SI$ ) based on its current value, and determines the optimal intervention.

**Objectives:** This study investigates the impact of a new 3D stochastic model on the ability to predict more accurate future  $SI$  distributions, which would allow more aggressive insulin dosing and improved glycaemic control.

**Methods:** The new 3D stochastic model is built using both current  $SI$  and its prior variation to predict future  $SI$  distribution from 68629 hours of clinical data (819 GC episodes). The 5<sup>th</sup>-95<sup>th</sup> percentile range of predicted  $SI$  distribution are calculated and compared with the 2D model.

**Results:** Results show the 2D model is over-conservative compared to the 3D case for more than 77% of the data, predominantly where  $SI$  is stable ( $|\% \Delta SI| \leq 25\%$ ). These formerly conservative prediction ranges are now ~30% narrower with the 3D model, which safely enables more aggressive insulin dosing for these patient hours. In addition, distributions of predicted  $SI$  within the 5<sup>th</sup> – 95<sup>th</sup> percentile range are much closer to the ideal value of 90% for more patients with the 3D model.

**Conclusions:** The new 3D model better characterises patient specific metabolic variability and patient specific response to insulin, allowing more optimal insulin dosing to increase performance and safety.

**Keywords:** *Critical Care, Insulin Sensitivity, Glycaemic Control, Blood Glucose, Hyperglycaemia, Insulin.*

## 1.0 Introduction:

Critically ill patients in intensive care units (ICUs) often experience abnormally elevated blood glucose (BG) concentrations (hyperglycaemia), as a stress response to illness and injury [1-3]. Hyperglycaemia, glycaemic variability, and hypoglycaemia are all independently associated with increased morbidity and mortality [3-10]. Glycaemic control (GC) using insulin therapy has shown beneficial outcomes, reducing organ failure and costs [11-18]. However, other studies failed to reproduce these results [19-24], and all but two studies [25] had increased risk of hypoglycaemia with tight control. GC has been hard to achieve both safely and effectively (e.g. [26]). Fixed or ad hoc protocols are still typically used in hospitals, but fail to capture and fully account for patient variability impacting performance and safety [27]. This issue has led to the emergence of more complex, model-based GC protocols [28-30].

STAR (Stochastic Targeted) is a clinically-validated model-based GC framework, capable of adapting treatment to patient-specific insulin requirements while managing the risk of hypo glycaemia [25, 31-33]. STAR uses a patient-specific time-varying model-based insulin sensitivity ( $SI$ ) to estimate patient metabolic condition. Likely future changes in  $SI$  are assessed using population-based stochastic models [34]. The 5<sup>th</sup>-95<sup>th</sup> percentile interval of BG outcomes is calculated from the 5<sup>th</sup>-95<sup>th</sup> percentile interval in  $SI$  outcomes, allowing forward prediction of likely BG outcomes for any given insulin-nutrition intervention. STAR thus selects an insulin-nutrition treatment to best overlap the clinically specified target BG range, while also managing and mitigating hypoglycaemic risk [32, 35], a unique risk-based dosing approach.

The stochastic model currently used by STAR forecasts future  $SI$  ( $SI_{n+1}$ ) distributions based on the identified current  $SI$  value ( $SI_n$ ). A Markov process is used, where outcome  $SI_{n+1}$  only depends on input  $SI_n$  [34]. This study expands this existing 2D stochastic approach by adding the most recent change in  $SI_n$  as an input parameter for forward prediction of outcome  $SI_{n+1}$ . The new 3D model will now predict future  $SI_{n+1}$  based on current  $SI_n$  and the percentage change in  $SI$  from  $SI_{n-1}$  to  $SI_n$ . The old 2D model and the new 3D model are compared to assess the new model's ability to tighten  $SI$  prediction ranges for tighter forward prediction of future BG. Better forward prediction of  $SI$  allows better characterisation of future metabolic variability, thus improving patient-specific glycaemic control without compromising

safety. Narrower future *SI* prediction ranges enable more targeted insulin dosing for these patients who are more stable. This study also assesses whether more stable patients have less future metabolic variability.

## 2.0 Material and Methods:

### 2.1 Glucose-insulin model and insulin sensitivity

The ICING (Intensive Control Insulin-Nutrition-Glucose) physiological model describing glucose-insulin dynamics is defined [25, 36, 37]:

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

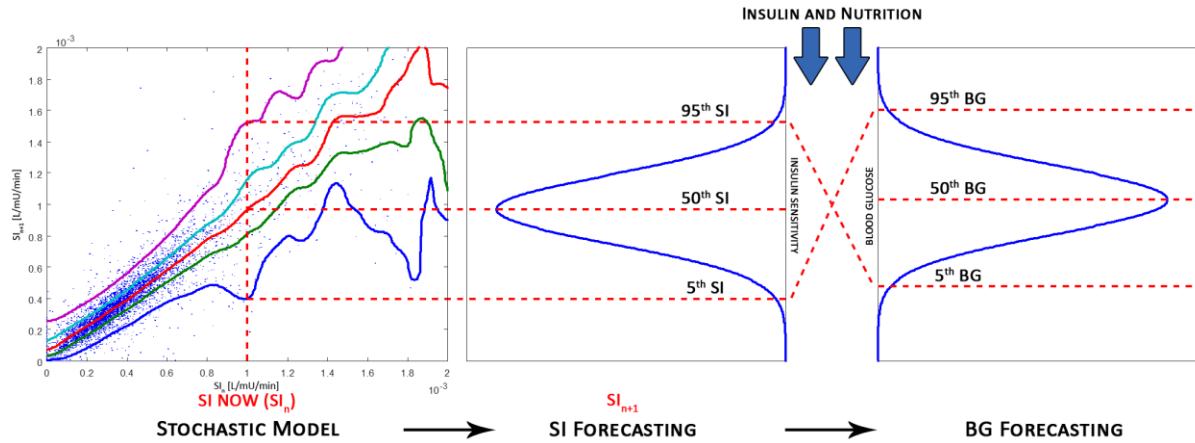
$$\dot{I} = n_K \cdot I(t) - n_L \frac{I(t)}{1 + \alpha_I \cdot I(t)} - n_I (I(t) - Q(t)) + \frac{u_{ex}(t)}{V_I} + (1 - x_L) \frac{u_{en}(G)}{V_I} \quad (2)$$

$$\dot{Q} = n_I (I(t) - Q(t)) - n_C \frac{Q(t)}{1 + \alpha_G Q(t)} \quad (3)$$

Where  $G(t)$  is the blood glucose level (mmol/L),  $I(t)$  is the plasma insulin concentration (mU/L),  $Q(t)$  is the interstitial insulin concentration (mU/L),  $P(t)$  is the glucose appearance in plasma from enteral and parenteral dextrose intake (mmol/min), and  $SI$  is insulin sensitivity (L/mU/min). Other parameters, rates and constants are given in [25, 36, 37] and can be found in the **Additional File 1**.

Model-based insulin sensitivity ( $SI$ ) is patient-specific and time varying, characterising patient-specific glycaemic system response to glucose and insulin administration.  $SI$  is identified hourly from clinical BG, and insulin and nutrition input data, using an integral-based fitting method [38, 39]. This approach is robustly identifiable [40].

STAR currently uses a cohort-based 2D stochastic model to forecast future  $SI$ . As shown in Figure 1, for any current  $SI$  ( $SI_n$ ), the probability of  $SI$  ( $SI_{n+1}$ ) at 1-3 hours in future is determined based on a clinical data model using kernel density methods [34]. Future  $SI$  distributions used in conjunction with Equations 1 – 3 can be used to derive likely future BG distributions for a specific insulin and nutrition intervention. The 5<sup>th</sup> percentile BG prediction is used to ensure safety, limiting the maximum risk of BG < 4.4 mmol/L to 5% and enabling risk-based, rather than target-value-based, dosing [32].



**Figure 1** - Future insulin sensitivity ( $SI_{n+1}$ ) is **forecasted** from current insulin sensitivity ( $SI_n$ ). The distribution of future  $SI$  is used to predict likely BG outcomes for a given insulin-nutrition treatment intervention.

## 2.2 Patients and cohorts

This study uses data from 3 clinical ICU data cohorts totalling 819 GC episodes (606 patients) and 68629 hours of treatment [13, 25]:

1. Patients treated using STAR in Christchurch Hospital ICU, New Zealand, from June 2011 to May 2015.
2. Patients treated using SPRINT in Christchurch Hospital ICU, New Zealand, from July 2005 to May 2007
3. Patients treated using STAR in Kalman Pandy Hospital ICU, Hungary, from December 2011 to May 2015.

Demographics are summarised in **Table 1**.

**Table 1** – Summary of patient demographics for the three cohorts. Results are given as median [IQR] where relevant.

	SPRINT Christchurch	STAR Christchurch	STAR Gyula
# episodes	442	330	47
# patients	292	267	47
# hours	39838	22523	6268
% male	62.7	65.5	61.7
Age (years)	63 [48, 73]	65 [55, 72]	66 [58, 71]
APACHE II	19.0 [15.0:24.5]	21.0 [16.0:25.0]	32.0 [28.0:36.0]
LOS - ICU (days)	6.2 [2.7,13.0]	5.7 [2.5,13.4]	14.0 [8.0,20.5]

### 2.3 Analysis

$SI$  level is fit on an hour-to-hour basis for each patient [38], and the forward  $SI$  variability ( $\% \Delta SI$ ) is defined as the hour-to-hour percentage change in  $SI$ :

$$\% \Delta SI_n = 100 \times \frac{SI_n - SI_{n-1}}{SI_{n-1}}$$

The existing 2D stochastic model uses the input  $SI_n$  to determine the outcome distribution of  $SI_{n+1}$ . This study builds a new 3D model to determine the outcome distribution of  $SI_{n+1}$  based on input of patient-specific current metabolic state,  $SI_n$ , and  $SI$  variability to current time,  $\% \Delta SI_n$ .

A total of 66991 data triplets ( $\% \Delta SI_n, SI_n, SI_{n+1}$ ) are created from the original 68629 hours of treatment, where the number of triplets is lower because no triplets are created for the first and last hour of data. The data triplets were binned with bin sizes of  $\% \Delta SI = 10\%$  and  $SI_n = 0.5e-4$ . These bins are limited to a range of  $\% \Delta SI = [-100\%, 200\%]$  and the 1<sup>st</sup>-99<sup>th</sup> percentile range in identified  $SI$  ( $[1.0e-7, 2.1e-3]$  L/mU/min) values, bringing the total number of triplets considered to 65051, as those few outside these ranges are excluded, corresponding to 97.1% of the original 66991 triplets.

The minimum number of data points required for adequate data density in each bin was arbitrarily defined to be 100 data triplets to ensure any distributions were not influenced by outliers. To improve data density and smooth model extremes, bins not meeting this criterion are summed together along the  $\% \Delta SI$  axis at the same  $SI_n$  level, allowing data triplets to influence neighbouring bins where there is insufficient data density. The summation process is described below, and an example is shown in **Table 2**.

Starting from the bin centred at 0% and going down, and at the bin centred at 10% and going up:

1. Check data density.
  - a. If the number of data triplets is  $\geq 100$ , move to the next bin.

- b. If the number of triplets < 100 add the triplets from the 'outer' adjacent bin(s) until data density is reached. If summation of bins still results in a failure to reach the data density, stop here.

2. Repeat step 1 until the model limits are reached.

**Table 2** – Example of data density before and after the merging bin process, where the joined lines show the bins that are merged to create the final surface.

Bin along $\% \Delta SI$	# triplets in bin for a given $SI_n$	→	# triplets in bin after summing bins together.
[-45, 55] %	0		-
[-35, 45] %	2		-
[-25, 35] %	10		-
[15, 25] %	25		37
[5, 15] %	80		105
[-5, 5] %	150		150
[-15, -5] %	102		102
[-25, -15] %	10		11
[-35, -25] %	1		-

The 5<sup>th</sup>, 50<sup>th</sup>, and 95<sup>th</sup> percentiles of  $SI_{n+1}$  are computed for each bin. These values define a non-parametric 5<sup>th</sup> – 95<sup>th</sup> percentile distribution range and median likely outcome of future variation of  $SI$  values based on a specific current  $SI_n$  level and the previous change  $\% \Delta SI_n$ . They are interpolated linearly between bins to give a percentile surface.

To compare this new 3D model to the previous 2D model, the percentage change in the 5<sup>th</sup>, 50<sup>th</sup>, and 95<sup>th</sup> percentiles are analysed, as well as the percentage change in the 5<sup>th</sup> – 95<sup>th</sup> percentile prediction range width used in STAR GC to select insulin dose in **Figure 1**. This width effectively determines the outcome BG range for a given intervention. Hence, a narrower width allows better prediction performance.

This analysis will identify regions of the model that are either conservative or have higher risk of hyper- and hypo- glycaemia. A narrowing of the 5<sup>th</sup> – 95<sup>th</sup> percentile range in the 3D model compared to the 2D model would allow the STAR controller to dose insulin more aggressively as outcome likelihoods



are tighter. In contrast, wider 5<sup>th</sup> – 95<sup>th</sup> percentile range than in the 2D stochastic model indicates increased risks, when using the 2D model, and insulin is more conservatively dosed based on this new information.

## 2.4 Validation

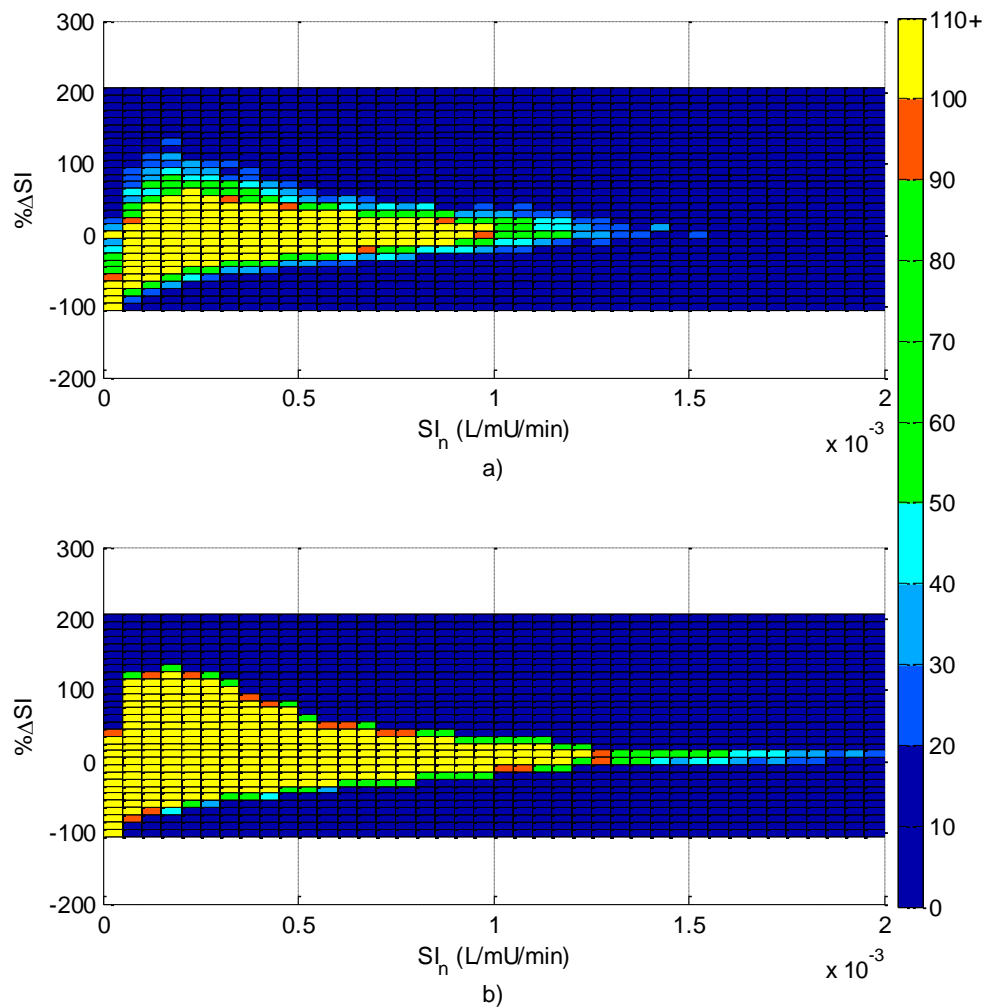
A preliminary validation of the model is carried out by assessing its ability to predict *SI* in clinical data episodes longer than 24 hours. The two models are compared by evaluating the per-patient percentage of *SI* outcomes falling into the model-predicted 5<sup>th</sup> – 95<sup>th</sup> the 25<sup>th</sup> – 75<sup>th</sup> percentile ranges for each model. Ideally, all episodes have exactly 90% and 50% within these ranges indicating a cohort derived model that is also perfect for each patient episode. Overall, this metric gives a measure of the model's ability to capture clinically observed patient-specific changes in insulin sensitivity [41], as well as quantifying the cohort-derived model's level of patient-specificity. The overall goal of this comparison is a more patient-specific stochastic model.

To validate model consistency in a cohort different from which it was developed, cross simulations were used. A 3D model was built from *SI* traces from a randomly selected group of episodes comprising of 70% of all episodes, and tested on the remaining 30%. Per-patient percent time in the 25<sup>th</sup> – 50<sup>th</sup> and 5<sup>th</sup> – 95<sup>th</sup> percentile ranges are computed, as well as the ratio between the widths of the 5<sup>th</sup> – 95<sup>th</sup> percentile prediction ranges for both models. This process is repeated 50 times using episodes with at least 24 hours of clinical glycaemic control data. Significant variability in results indicates a model based on too little data and/or dominated by selected patients or episodes. Consistent results indicate the model is built of enough data and/or is not skewed by outlying data. This analysis thus assesses model robustness to its underlying data.

### 3.0 Results:

#### 3.1 3D model for forward prediction of $SI$ :

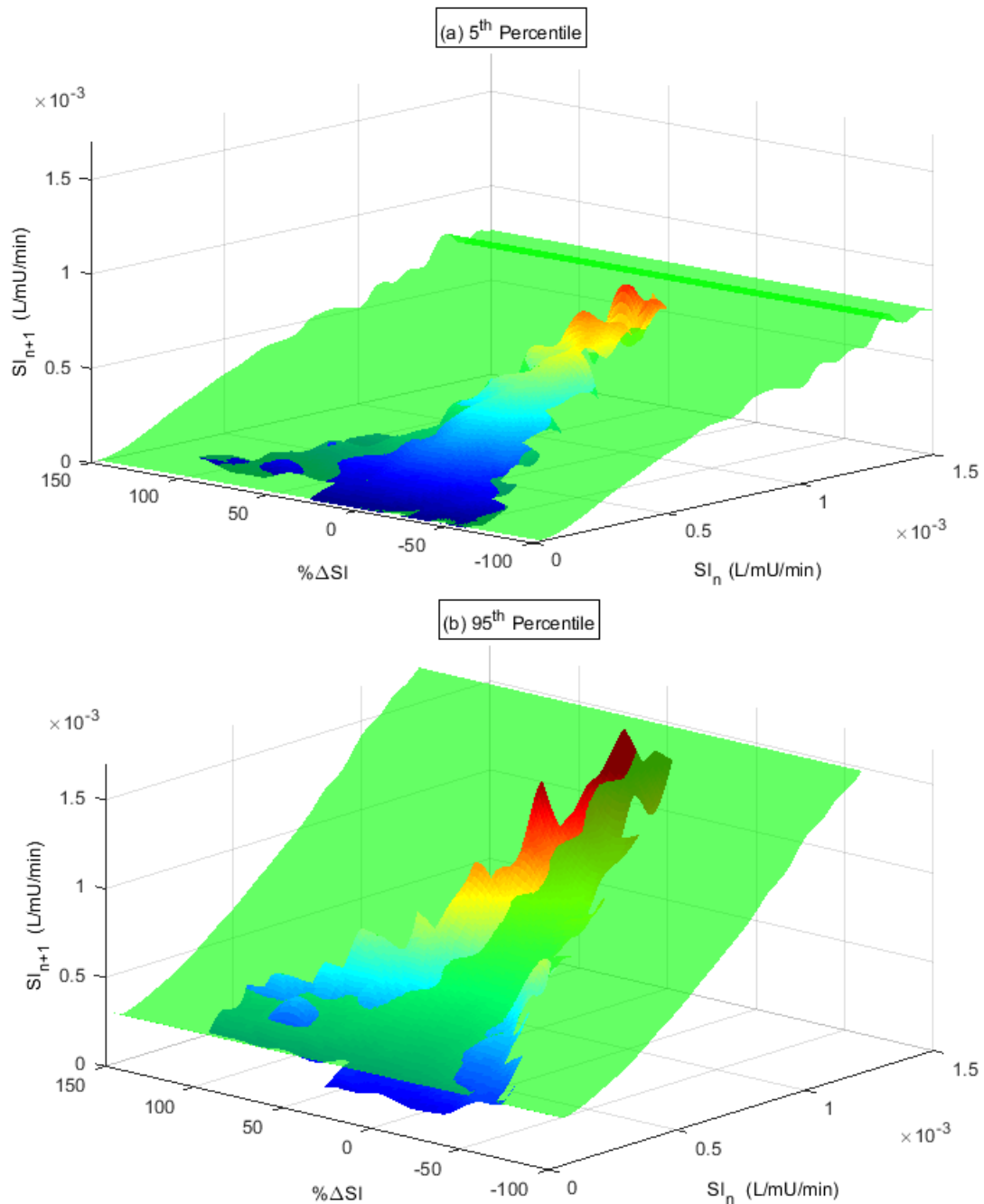
The number of data triplets ( $\% \Delta SI_n, SI_n, SI_{n+1}$ ) per bin is shown in **Figure 2**, before and after bin merging to improve data density. The triangular shape of binned data suggests greater variability at lower  $SI$  where changes constitute a larger percentage change relative to absolute  $SI$  value. Yellow areas represent bins with enough data density (at least 100 data triplets) and thus the bins which will be used to build the model. These bins represent 62316 (95.8%) of the original 65051 triplets.



**Figure 2** – Number of data triplets ( $\% \Delta SI_n, SI_n, SI_{n+1}$ ) per bin, before (a) and after (b) merging side bins (along y-axis). Bins in yellow reach minimum data density and will be used to build the 3D model.

The new interpolated 3D model is shown in **Figure 3** and compared to the original stochastic model (green) for the 5<sup>th</sup> (a) and 95<sup>th</sup> (b) percentiles. The previous 2D model forms a plane in the new 3D

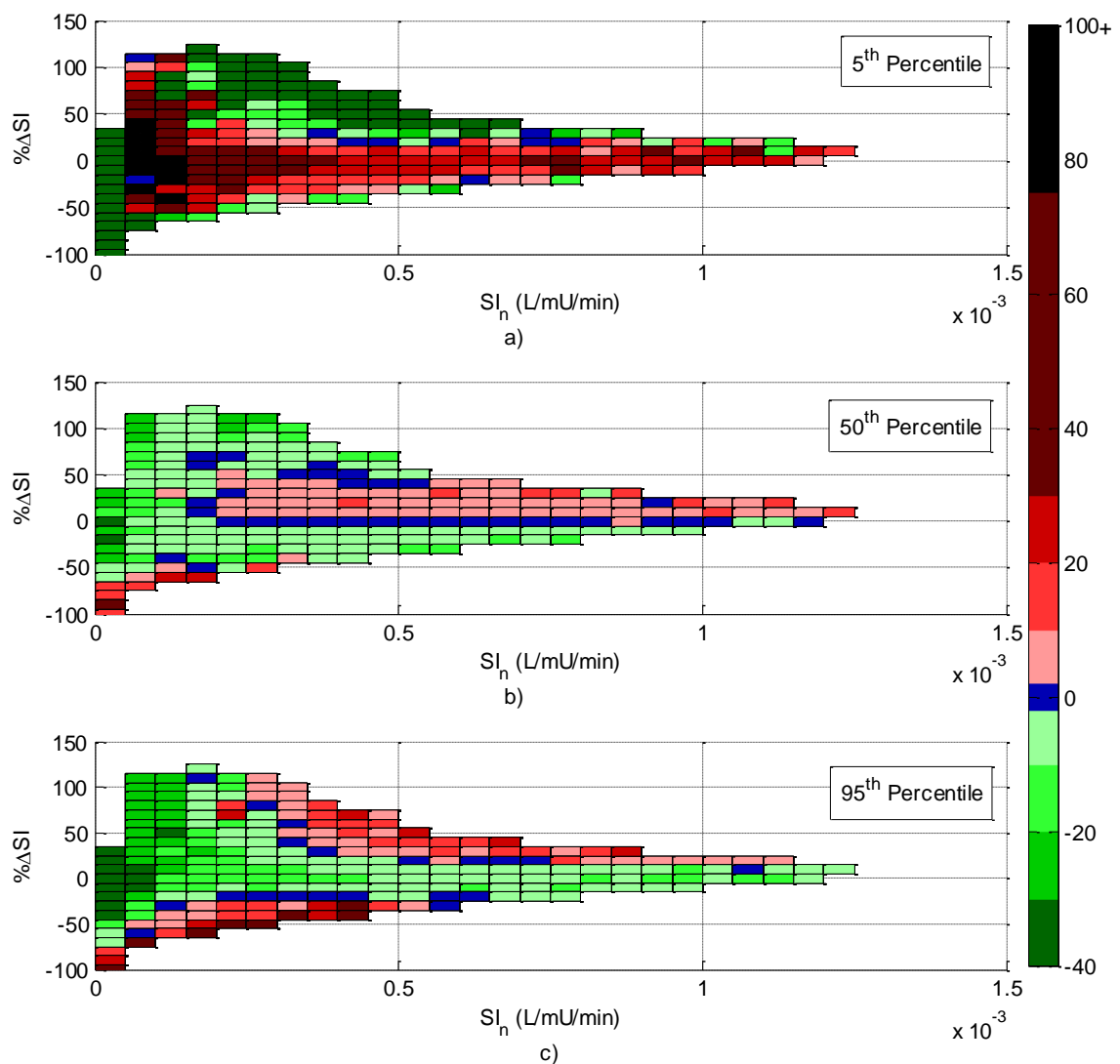
model space as it is constant across all  $\% \Delta SI$ . Where the new 3D model sits above the 95<sup>th</sup> percentile or below the 5<sup>th</sup> percentile planes indicates where the 2D model was too narrow. The reverse case of above the 5<sup>th</sup> and/or below the 95<sup>th</sup> percentile indicates the 2D model was over conservative and the 3D model 5<sup>th</sup> – 95<sup>th</sup> percentile prediction range is narrower.



**Figure 3** – Comparison between new 3D model (colour) and the original 2D model (green) for the 5<sup>th</sup> (a) and 95<sup>th</sup> (b) percentiles.

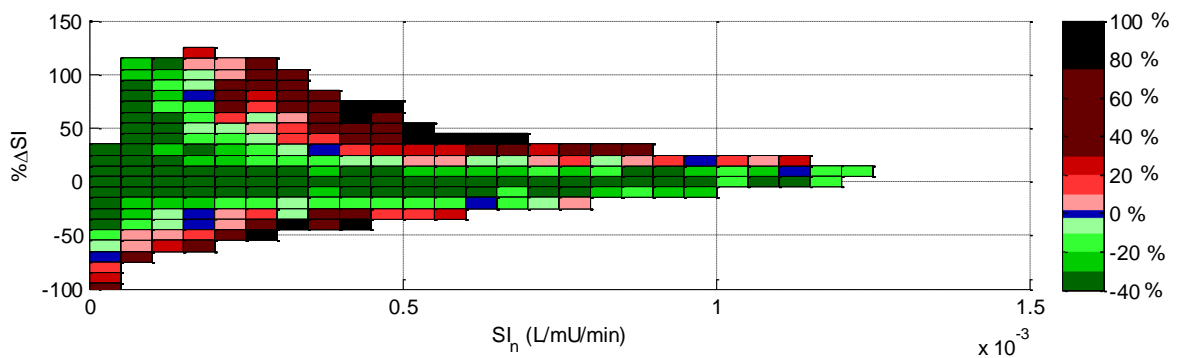
The percentage change in the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentiles prediction of  $S_{I_{n+1}}$  is shown in **Figure 4**. Two main regions can be identified:

1. A conservative region, mainly between  $\% \Delta S_{I_n} = \pm 25\%$ . The 5<sup>th</sup> percentile is higher than the previous 2D model, while the 95<sup>th</sup> is lower, regardless of  $S_{I_n}$ , describing thus a narrower **5<sup>th</sup> – 95<sup>th</sup> percentile range** in the forward prediction of  $S_{I_{n+1}}$ . This region represents 77.2% of the data triplets.
2. A non-conservative region outside  $\% \Delta S_{I_n} = \pm 25\%$ . The 5<sup>th</sup> percentile is lower and the 95<sup>th</sup> is higher, indicating higher resulting risks of hyper- and hypo- glycaemia than predicted by the 2D model.



**Figure 4** – Percentage change in the 5<sup>th</sup> (a), 50<sup>th</sup> (b), and 95<sup>th</sup> (c) percentiles between the original 2D stochastic model and the new 3D model.

Summarising **Figure 4**, the percentage change in the 5<sup>th</sup> – 95<sup>th</sup> percentile prediction range is shown in **Figure 5**. In the 2D model's conservative region, a significant decrease of ~25-40% in this range is observed for the 3D model, suggesting the new model allows more aggressive insulin treatment than the previous model and thus provides improved information for dose selection. For the non-conservative region, increases in  $SI$  of up to 80% or more are observed, allowing the model to more safely predict and cope with large changes in  $SI$  in regions of high metabolic variability (high  $\% \Delta SI$ ). The reduced width green region contains 51871 data triplets (79.7%) of the total.



**Figure 5** – Percentage change in the width of the 5<sup>th</sup>-95<sup>th</sup> percentile range when the new 3D model is compared to the previous 2D model. Green and red areas suggest overly conservative and under conservative behaviour respectively within the 2D model.

### 3.2 Self Validation

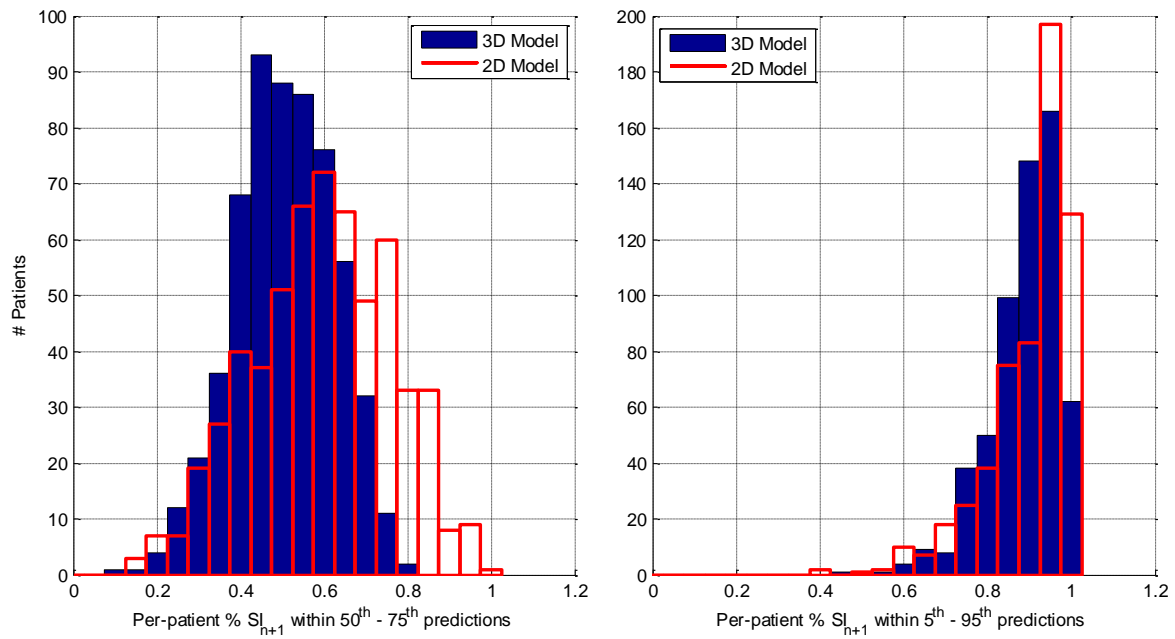
The predictive power of the new model is tested on 587 episodes of minimum 24 hours, representing a total of 63508 triplets. In total, 2.7% of these data triplets were in the initial exclusion definition and thus 97.3% of  $SI_{n+1}$  predictions fell within the model. Of these included values, 90.4% of  $SI_{n+1}$  predictions were within the 5<sup>th</sup> – 95<sup>th</sup> percentile prediction range, and 52.6% fell within the 25<sup>th</sup> – 75<sup>th</sup> percentile range, which is very close to the expected values of 90% and 50%, respectively. The 5<sup>th</sup> – 95<sup>th</sup> percentile interval is more critical due to its use in dosing (**Figure 1**). The larger error in the 25<sup>th</sup> - 75<sup>th</sup> percentile interval versus the 5<sup>th</sup> – 95<sup>th</sup> percentile interval indicates a small mismatch in the distribution shapes across  $SI_n$ .

**Figure 6** presents a histogram of per-patient percentage forward prediction in the 25<sup>th</sup>-75<sup>th</sup> and 5<sup>th</sup>-95<sup>th</sup> percentile bands, showing the accuracy at a per-patient level, rather than an overall cohort level. **Table 3** shows the corresponding per-patient median [IQR] percentage time in these bands. The per-patient median [IQR] percentage time in the 25<sup>th</sup> – 75<sup>th</sup> percentile prediction range is higher in the previous 2D model (60.3% [47.8%, 71.5%] vs. 51.2% [42.9%, 59.2%]), while the percentage time in the 5<sup>th</sup> – 95<sup>th</sup> percentile prediction range is similar (93.6% [85.7%, 97.3%] vs. 90.7% [84.4%, 94.6%]) between the 2D and 3D models. However, as seen in **Figure 6**, the per-patient distributions are tighter to the ideal values (50% and 90%) for the 3D model, reducing over-conservatism.

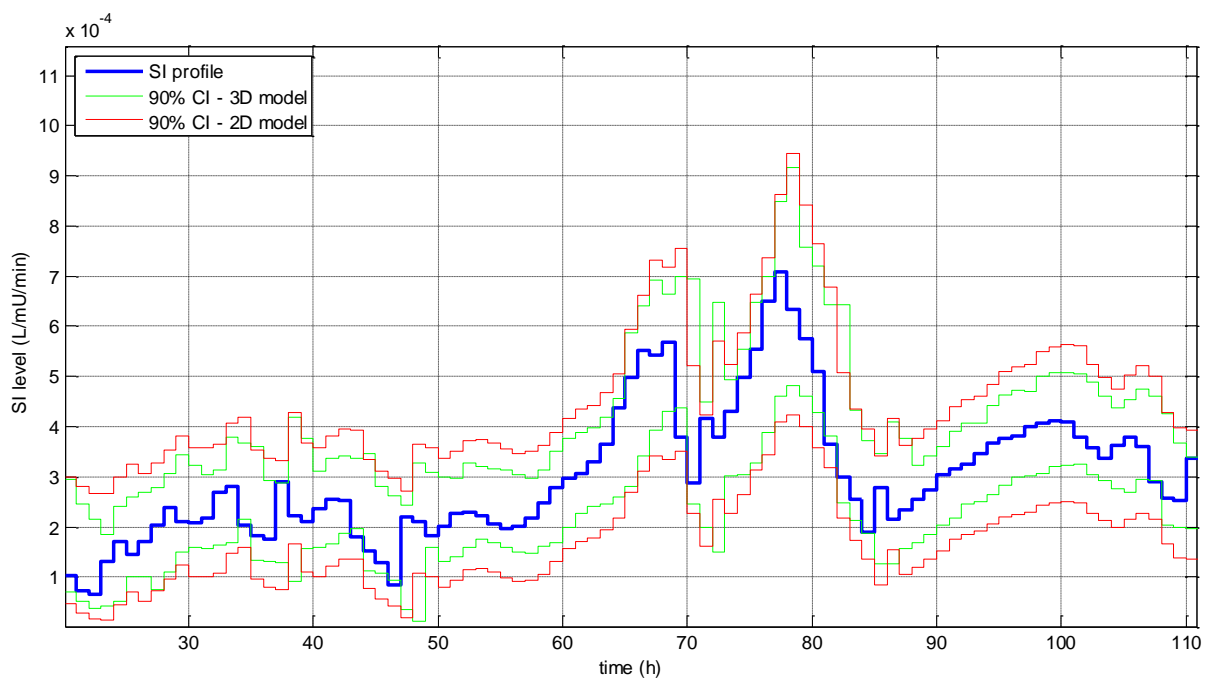
More importantly, there was a significant reduction in the width of the 5<sup>th</sup>-95<sup>th</sup> percentile ranges for each patient, with the 3D model reducing this width by median 28.9% [21.6%, 33.0%] per bin. These results suggest the new 3D model is able to account for changes in *S/I* equally well in comparison to the 2D model, but with significantly narrowed prediction range for many hours of care. This outcome should allow safe application of more aggressive insulin treatments for more stable patients. An example comparison of the predictive power of the two models is shown in **Figure 7**.

**Table 3** – Per-patient predictive power comparison between old and new stochastic model. Results are given as median [IQR].

	2D Model	3D model
Median per-patient % prediction within 25 <sup>th</sup> -75 <sup>th</sup> percentile range	60.3% [47.8%, 71.5%]	51.2% [42.9%, 59.2%]
Median per-patient % prediction within 5 <sup>th</sup> -95 <sup>th</sup> percentile range	93.6% [85.7%, 97.3%]	90.7% [84.4%, 94.6%]
Median per-patient % reduction in 5 <sup>th</sup> -95 <sup>th</sup> percentile range width	28.9% [21.6%, 33.0%]	



**Figure 6** – Per-patient predictive power within the 25<sup>th</sup>-75<sup>th</sup> percentile prediction range (left) and within the 5<sup>th</sup>-95<sup>th</sup> percentile prediction range (right) of new (blue) and old (red) models.



**Figure 7** – Excerpt from a patient showing fitted  $SI$  (blue) as well as 5<sup>th</sup> and 95<sup>th</sup> percentile prediction for the new 3D model (green) and the old 2D model (red). The new model predictive range is generally narrower than the old model.

**Table 4** compares prediction outcomes for patients who had increased and decreased time in prediction ranges, respectively. A total of 101 episodes increased percentage time in the 5<sup>th</sup> – 95<sup>th</sup> percentile prediction range by ~5% (87.3% [80.0%, 92.7%] vs. 82.3% [71.8%, 88.9%] for the 2D and 3D models respectively). Conversely, for the remaining 486 episodes, the new model shows slightly lower performance (91.1% [85.7%, 95.0%] vs. 94.6% [88.9%, 97.6%]). However, the percentage time in range for these patients has been brought closer to the intended 90%, ensuring the 3D model treats patients more consistently across the cohort.

**Table 4** – Per-patient predictive power comparison between old and new stochastic model for two groups: 101 patients whose % prediction increased with the new model and 486 patients whose % prediction decreased. Results are given as median [IQR].

	2D Model	3D model
Median % prediction within 5 <sup>th</sup> -95 <sup>th</sup> percentile range for 101 patients who increased time in range.	82.3% [71.8%, 88.9%]	87.3% [80.0%, 92.7%]
Median % prediction within 5 <sup>th</sup> -95 <sup>th</sup> percentile range for 486 patients who decreased time in range.	94.6% [88.9%, 97.6%]	91.13% [85.7%, 95.0%]

### 3.3 Cross validation:

Cross-validation was carried out using patient episodes 24 hours or longer, and results are shown in **Table 5**. Compared to the original 2D stochastic model, the new 3D model has consistent, 12% absolute lower median [IQR] percentage forward prediction in the 25<sup>th</sup>-75<sup>th</sup> and 5<sup>th</sup>-95<sup>th</sup> percentiles ranges (3D: 51.8% [51.5%, 52.1%] vs. 2D: 63.1% [62.8%, 63.4%] and 89.8% [89.6%, 90.0%] vs. 92.5% [92.4%, 92.6%]). Additionally, the 5<sup>th</sup>-95<sup>th</sup> percentile range width from 2D to 3D model is reduced by median 30.8% [30.5%, 31.1%]. These results suggest both models generalise well to other ICU patients when developed from an independent, but similar cohort of patients, matching similar tests across cohorts [25].



**Table 5** – Cross-validation per-patient results for old 2D stochastic model and the new 3D model, on all *S/* values from episodes of minimum 24 hours. Results are given as median [IQR].

	2D Model	3D model
Median % prediction within 25 <sup>th</sup> -75 <sup>th</sup> percentile range	63.1% [62.8%, 63.4%]	51.8% [51.5%, 52.1%]
Median % prediction within 5 <sup>th</sup> -95 <sup>th</sup> percentile range	92.6% [92.5%, 92.7%]	89.7% [89.6%, 90.0%]
Median % reduction of the 5 <sup>th</sup> -95 <sup>th</sup> percentile range width from 2D to 3D model	30.8% [30.5%, 31.1%]	

## 4.0 Discussion:

### 4.1 Discussion of main results

Forecasting changes in insulin sensitivity underpins the ability of STAR to respond in a patient-specific manner to potential future changes in patient glycaemic control requirements, resulting in safe and effective control strategies. If the distributions of forecast likely  $SI$  changes are narrower, then control can be further improved, with tighter control in more stable patients, and better avoidance of hypoglycaemia in patients that exhibit high glycaemic variability. In this study, current probabilistic forecasting methods have been extended to include the change in  $SI$  ( $\% \Delta SI$ ) as an input predictor for future  $SI$  alongside current  $SI$ .

The previous 2D stochastic model is shown to be conservative for  $\sim 77\%$  of the data, where the  $\% \Delta SI_n$  is within  $\pm 25\%$  change. While conservatism results in wider prediction ranges in likely BG outcomes, thus reducing the risk of hypoglycaemia, it also inhibits the controllers' ability to reduce BG to the normal range using more aggressive control dosing. This issue particularly affects patients who tend to remain stable, but are mildly hyperglycaemic as a result. Hence, such conservatism, while safe, has a potentially negative clinical impact, as well.

Compared to the 2D stochastic model, a region that is over-conservative equally implies a region that is under-conservative. This under-conservative region means there are less stable patients, outside  $|\% \Delta SI| \leq 25\%$ , who have increased hypoglycaemic risk from relatively over-aggressive dosing, resulting from prediction bands that are too narrow. This trade off of conservatism is seen in **Figures 4** and **Figure 5**, and offers unintended increased risk for these patients in using the 2D model.

The new model utilises change,  $\% \Delta SI$ , as an additional model input to better predict a more patient-specific future  $SI_{n+1}$ , with narrowed prediction ranges for 77% of hours, and overall similar ability to meet the expected 90% of  $SI$  outcomes within the 5<sup>th</sup>-95<sup>th</sup> percentile prediction range. The new model is thus more patient-specific, and better predicts likely BG outcomes. These results should translate into more aggressive insulin dosing where patients are more stable and  $SI$  outcomes are more certain, and less

aggressive, lower insulin doses in patients who are more variable. Greater patient-specificity also reduces risk for more variable patients. This model could thus lead to tighter and less variable control with greater safety from hypoglycaemia, and thus improved outcomes [18, 42].

Both the self-validation and the cross-validation tests have shown the predictive power of this new model to be more consistent, with closer to 90% of forward prediction of *SI* within the 5<sup>th</sup>-95<sup>th</sup> percentile range. These results align well with the expected 90% of *SI* outcomes falling within this prediction range. However, in comparison to the previous 2D model, the new 3D model achieves this performance with overall narrower and tighter prediction distributions for 77% of hours. A median reduction in the 5<sup>th</sup>-95<sup>th</sup> percentile prediction range width of approximately 30% was achieved with the 3D model, indicating the 3D model is better able to predict future *SI* outcomes, and thus safely allowing significantly increased insulin dosing.

The new 3D model thus treats patients more consistently across the patient cohort. Previously, with the 2D model, some patients within the cohort had much more than 90% of their *SI* outcomes within the 5<sup>th</sup> – 95<sup>th</sup> percentile prediction range, and others much less (**Figure 6**). While greater conservatism is advantageous for avoiding extreme BG outcomes, it also implies an over conservatism and inability of the model to meet its design specifications (90% of *SI* within the 5<sup>th</sup> -95<sup>th</sup> percentile range). Equally, it prevents aggressive dosing and better control where it could be warranted for specific patients, and is thus less patient-specific than the new 3D model. Given the improved 5<sup>th</sup> - 95<sup>th</sup> percentile range performance of the new 3D model, it is clear it is better able to consider patients more consistently across the cohort with the added  $\% \Delta SI$  model input.

Cross simulation tested the ability of the model to predict *SI* in patients that were not used to build the model. Cross validation results were very consistent and close to expected values from the whole-cohort analysis, suggesting that the model would generalise well to ICU patients from different protocols and/or units. It also indicates the model is not dominated by smaller outlying subsets of patients.

This result also reflects previous results where similar and consistent  $SI$  variability was seen across 3 different ICU cohorts in 3 different countries [25, 43]. This previous analysis uses slightly different patient cohorts from those presented here, and includes an additional cohort not used here. Thus, these results suggest the model generalises well to other ICU patients and cohorts, when the model is built from an independent, but overall similar cohort of mixed medical ICU patients.

## 4.2 Limitations

One limitation of this analysis is the binning process, which creates a discretized 3D space, where percentile surfaces were linearly interpolated between bins. Kernel-density methods, used to build the previous 2D stochastic model [34], could also be applied to generate smoother prediction percentile surfaces with respect to data density. However, this approach introduces assumptions around the shape of data density distributions and the effect of surrounding data on percentile surfaces. In this analysis, bin sizes were chosen to balance increased resolution against sufficient data density, as well as to first prove the concept.

The bin width of 10% in  $\% \Delta SI$  was chosen based on a previous analysis, which assessed the impact of BG measurement error on  $SI$  [44]. Given over 68,000 hours of patient data are used to build the new 3D model, the percentile surfaces of the full model are likely to be sufficiently reflective of  $SI$  dynamics, as data density was typically sufficient across typical areas of interest. Thus, limitations due to data size and density are likely to be minimal.

A further possible limitation of the model is that ~7% of data points fell outside the model range, having been discarded as outliers in regions of insufficient data density. These unusually large changes in  $SI$  or extremely high  $SI$  levels are thus not included in the model, and may reflect inaccuracies in data recording or patient-specific deviations from model-dynamics. It is also possible these points highlight times of extreme glycaemic change or measurement error, where the best clinical practice could be to discontinue insulin for an hour and come back and re-measure. In essence for these, outlying, potentially unexplained events, discontinuing insulin for a short period is a safe course of action. Within the STAR framework, the clinical usage would be to utilise the original proven 2D model, using the

safest and most conservative intervention [31, 43, 45]. Hence, points that fall in this range could be used to warn of outlying events that might not otherwise be noted, and be used to take no or more conservative action.

Overall, the new 3D model shows similar predictive performance and much tighter predictive bound when compared to the previous model. Cross-validation shows the new predictive model, constructed from bin sizes likely reflective of  $S/I$  dynamics, to accurately predict  $S/I$  for data not used to develop the model. These improvements in prediction should translate to tighter glycaemic control, without compromising safety from hypoglycaemia. Future work will investigate kernel-density methods to generate smoothed model surfaces, as well as exploring the management of outliers and their clinical significance.

## 5.0 Conclusions

Insulin sensitivity plays a major role in any model-based glycaemic control protocol and insulin sensitivity forecasting is particularly important for managing dynamic ICU patients. It thus plays a leading role in the model-based STAR protocol. In particular, it enables a patient-specific approach to achieve better control and the use of forward stochastic prediction models enables safety and performance to be explicitly balanced in determining optimal insulin dosing.

This study has shown the positive impact of identifying prior change in a proven model-based insulin sensitivity metric on the prediction of likely future insulin sensitivity distribution ranges. A new 3D model was developed, achieving similar predictive power as the previous model, while significantly reducing the width of the 5<sup>th</sup>-95<sup>th</sup> percentile prediction range for more than 77% of the hours of data. This outcome ensures that over three-quarters of patient hours will be treated less over-conservatively. Equally, it also ensures that the remaining quarter of patient hours are not treated aggressively (under-conservatively), and thus improves safety. Both outcomes will improve the performance, safety and patient specificity of glycaemic control, and thus patient outcomes.

## 6.0 Abbreviations

%ΔSI	Hour-to-hour percentage change in SI
APACHE	Acute Physiology and Chronic Health Evaluation
BG	Blood Glucose
GC	Glycaemic Control
ICING	Intensive Control Insulin-Nutrition-Glucose
ICU	Intensive Care Unit
LOS	Length of stay
SI	Insulin Sensitivity
SPRINT	Specialised Relative Insulin Nutrition Tables
STAR	STochastic-TARgeted

## 7.0 Acknowledgements

The authors declare no competing interests.

VU drafted the manuscript and carried out the main analysis and results. JLD, KS, TD, and JGC expertise input in the field was necessary for results interpretation and discussion. Clinical data and clinical interpretation of the results were provided by GMS, BB, NSN, and AI. All authors had input into the redaction process.

## 8.0 Funding sources

The authors acknowledge the support of the EUFP7 and RSNZ Marie Curie IRSES program, the MedTech CoRE and TEC, the NZ National Science Challenge 7, Science for Technology and Innovation, the European Erasmus+ Student Mobility program, and the FRIA – Fund for Research Training in Industry and Agriculture grant.

## 9.0 References

1. Capes SE, Hunt D, Malmberg K, Gerstein HC: **Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview.** *Lancet* 2000, **355**:773-778.
2. McCowen KC, Malhotra A, Bistrian BR: **Stress-induced hyperglycemia.** *Crit Care Clin* 2001, **17**:107-124.
3. Krinsley JS: **Association between hyperglycemia and increased hospital mortality in a heterogeneous population of critically ill patients.** *Mayo Clin Proc* 2003, **78**:1471-1478.
4. Ali NA, O'Brien JM, Jr., Dungan K, Phillips G, Marsh CB, Lemeshow S, Connors AF, Jr., Preiser JC: **Glucose variability and mortality in patients with sepsis.** *Crit Care Med* 2008, **36**:2316-2321.
5. Bagshaw SM, Bellomo R, Jacka MJ, Egi M, Hart GK, George C, Committee ACM: **The impact of early hypoglycemia and blood glucose variability on outcome in critical illness.** *Crit Care* 2009, **13**:R91.
6. Egi M, Bellomo R, Stachowski E, French CJ, Hart G: **Variability of blood glucose concentration and short-term mortality in critically ill patients.** *Anesthesiology* 2006, **105**:244-252.
7. Egi M, Bellomo R, Stachowski E, French CJ, Hart GK, Taori G, Hegarty C, Bailey M: **Hypoglycemia and outcome in critically ill patients.** *Mayo Clin Proc* 2010, **85**:217-224.
8. Krinsley JS: **Glycemic variability and mortality in critically ill patients: the impact of diabetes.** *J Diabetes Sci Technol* 2009, **3**:1292-1301.
9. Krinsley JS, Grover A: **Severe hypoglycemia in critically ill patients: risk factors and outcomes.** *Crit Care Med* 2007, **35**:2262-2267.
10. Reed CC, Stewart RM, Sherman M, Myers JG, Corneille MG, Larson N, Gerhardt S, Beadle R, Gamboa C, Dent D, et al: **Intensive insulin protocol improves glucose control and is associated with a reduction in intensive care unit mortality.** *J Am Coll Surg* 2007, **204**:1048-1054; discussion 1054-1045.
11. Van den Berghe G, Wouters P, Weekers F, Verwaest C, Bruyninckx F, Schetz M, Vlasselaers D, Ferdinande P, Lauwers P, Bouillon R: **Intensive insulin therapy in critically ill patients.** *N Engl J Med* 2001, **345**:1359-1367.
12. Krinsley JS: **Effect of an intensive glucose management protocol on the mortality of critically ill adult patients.** *Mayo Clin Proc* 2004, **79**:992-1000.
13. Chase JG, Shaw G, Le Compte A, Lonergan T, Willacy M, Wong XW, Lin J, Lotz T, Lee D, Hann C: **Implementation and evaluation of the SPRINT protocol for tight glycaemic control in critically ill patients: a clinical practice change.** *Crit Care* 2008, **12**:R49.
14. Van den Berghe G, Wouters PJ, Kesteloot K, Hilleman DE: **Analysis of healthcare resource utilization with intensive insulin therapy in critically ill patients.** *Crit Care Med* 2006, **34**:612-616.
15. Krinsley J: **Glucose control reduces ICU stay and mortality.** *Perform Improv Advis* 2005, **9**:4-6, 1.
16. Krinsley JS, Jones RL: **Cost analysis of intensive glycemic control in critically ill adult patients.** *Chest* 2006, **129**:644-650.
17. Chase JG, Pretty CG, Pfeifer L, Shaw GM, Preiser JC, Le Compte AJ, Lin J, Hewett D, Moorhead KT, Desai T: **Organ failure and tight glycemic control in the SPRINT study.** *Crit Care* 2010, **14**:R154.
18. Penning S, Pretty C, Preiser JC, Shaw GM, Desai T, Chase JG: **Glucose control positively influences patient outcome: A retrospective study.** *J Crit Care* 2015, **30**:455-459.
19. Finfer S, Chittock DR, Su SY, Blair D, Foster D, Dhingra V, Bellomo R, Cook D, Dodek P, Henderson WR, et al: **Intensive versus conventional glucose control in critically ill patients.** *N Engl J Med* 2009, **360**:1283-1297.



20. Finfer S, Liu B, Chittock DR, Norton R, Myburgh JA, McArthur C, Mitchell I, Foster D, Dhingra V, Henderson WR, et al: **Hypoglycemia and risk of death in critically ill patients.** *N Engl J Med* 2012, **367**:1108-1118.
21. Brunkhorst FM, Engel C, Bloos F, Meier-Hellmann A, Ragaller M, Weiler N, Moerer O, Gruendling M, Oppert M, Grond S, et al: **Intensive insulin therapy and pentastarch resuscitation in severe sepsis.** *N Engl J Med* 2008, **358**:125-139.
22. Preiser JC, Devos P, Ruiz-Santana S, Melot C, Annane D, Groeneveld J, Iapichino G, Leverve X, Nitenberg G, Singer P, et al: **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 Med* 2009, **35**:1738-1748.
23. Wiener RS, Wiener DC, Larson RJ: **Benefits and risks of tight glucose control in critically ill adults: a meta-analysis.** *JAMA* 2008, **300**:933-944.
24. Griesdale DE, de Souza RJ, van Dam RM, Heyland DK, Cook DJ, Malhotra A, Dhaliwal R, Henderson WR, Chittock DR, Finfer S, Talmor D: **Intensive insulin therapy and mortality among critically ill patients: a meta-analysis including NICE-SUGAR study data.** *CMAJ* 2009, **180**:821-827.
25. Stewart KW, Pretty CG, Tomlinson H, Thomas FL, Homlok J, Noemi SN, Illyes A, Shaw GM, Benyo B, Chase JG: **Safety, efficacy and clinical generalization of the STAR protocol: a retrospective analysis.** *Ann Intensive Care* 2016, **6**:24.
26. Finfer S, Chittock DR, Su SY, Blair D, Foster D, Dhingra V: **Intensive versus conventional glucose control in critically ill patients.** *N Engl J Med* 2009, **360**.
27. Chase JG, Le Compte AJ, Suhaimi F, Shaw GM, Lynn A, Lin J, Pretty CG, Razak N, Parente JD, Hann CE, et al: **Tight glycemic control in critical care--the leading role of insulin sensitivity and patient variability: a review and model-based analysis.** *Comput Methods Programs Biomed* 2011, **102**:156-171.
28. Chase JG, Shaw GM, Wong XW, Lotz T, Lin J, Hann CE: **Model-based glycaemic control in critical care - a review of the state of the possible.** *Biomedical Signal Processing and Control* 2006, **1**:3-21.
29. Van Herpe T, De Moor B, Van den Berghe G: **Towards closed-loop glycaemic control.** *Best Pract Res Clin Anaesthesiol* 2009, **23**:69-80.
30. Amrein K, Ellmerer M, Hovorka R, Kachel N, Parcz D, Korsatko S, Smolle K, Perl S, Bock G, Doll W, et al: **Hospital glucose control: safe and reliable glycemic control using enhanced model predictive control algorithm in medical intensive care unit patients.** *Diabetes Technol Ther* 2010, **12**:405-412.
31. Evans A, Le Compte A, Tan CS, Ward L, Steel J, Pretty CG, Penning S, Suhaimi F, Shaw GM, Desaive T, Chase JG: **Stochastic targeted (STAR) glycemic control: design, safety, and performance.** *J Diabetes Sci Technol* 2012, **6**:102-115.
32. Fisk LM, Le Compte AJ, Shaw GM, Penning S, Desaive T, Chase JG: **STAR development and protocol comparison.** *IEEE Trans Biomed Eng* 2012, **59**:3357-3364.
33. Evans A, Shaw GM, Le Compte A, Tan CS, Ward L, Steel J, Pretty CG, Pfeifer L, Penning S, Suhaimi F, et al: **Pilot proof of concept clinical trials of Stochastic Targeted (STAR) glycemic control.** *Ann Intensive Care* 2011, **1**:38.
34. Lin J, Lee D, Chase JG, Shaw GM, Le Compte A, Lotz T, Wong J, Lonergan T, Hann CE: **Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care.** *Comput Methods Programs Biomed* 2008, **89**:141-152.
35. Evans A, Shaw GM, Compte A, Tan CS, Ward L, Steel J: **Pilot proof of concept clinical trials of Stochastic Targeted (STAR) glycemic control.** *Ann Intensive Care* 2011, **1**.
36. Lin J, Razak NN, Pretty CG, Le Compte A, Docherty P, Parente JD, Shaw GM, Hann CE, Geoffrey Chase J: **A physiological Intensive Control Insulin-Nutrition-Glucose (ICING) model validated in critically ill patients.** *Comput Methods Programs Biomed* 2011, **102**:192-205.

37. Pretty CG, Signal M, Fisk L, Penning S, Le Compte A, Shaw GM, Desai T, Chase JG: **Impact of sensor and measurement timing errors on model-based insulin sensitivity.** *Comput Methods Programs Biomed* 2014, **114**:e79-86.
38. Hann CE, Chase JG, Lin J, Lotz T, Doran CV, Shaw GM: **Integral-based parameter identification for long-term dynamic verification of a glucose-insulin system model.** *Comput Methods Programs Biomed* 2005, **77**:259-270.
39. Docherty PD, Chase JG, David T: **Characterisation of the iterative integral parameter identification method.** *Med Biol Eng Comput* 2012, **50**:127-134.
40. Docherty PD, Chase JG, Lotz TF, Desai T: **A graphical method for practical and informative identifiability analyses of physiological models: a case study of insulin kinetics and sensitivity.** *Biomed Eng Online* 2011, **10**:39.
41. Dickson J, Floyd RP, Le Compte A, Fisk L, Chase JG: **External validation and sub-cohort analysis of stochastic forecasting models in NICU cohorts.** *Biomedical Signal Processing and Control* 2013, **8**:409-419.
42. Signal M, Le Compte A, Shaw GM, Chase JG: **Glycemic levels in critically ill patients: are normoglycemia and low variability associated with improved outcomes?** *J Diabetes Sci Technol* 2012, **6**:1030-1037.
43. Dickson JL, Stewart KW, Pretty CG, Flechet M, Desai T, Penning S, Lambermont BC, Benyo B, Shaw GM, Chase G: **Generalisability of a Virtual Trials Method for Glycaemic Control in Intensive Care.** *IEEE Transactions on Biomedical Engineering* 2017, **PP**:1-1.
44. Uyttendaele V, Dickson JL, Shaw GM, Desai T, Chase JG: **Untangling glycaemia and mortality in critical care.** *Crit Care* 2017, **21**:152.
45. Stewart KW, Pretty CG, Tomlinson H, Thomas FL, Homlok J, Noemi SN: **Safety, efficacy and clinical generalization of the STAR protocol: a retrospective analysis.** *Ann Intensive Care* 2016, **6**.