A minimal C-peptide sampling method to capture peak and total pre-hepatic insulin secretion in model-based experimental insulin sensitivity studies

Thomas Lotz (PhD), Uli Göltenbott, J Geoffrey Chase (PhD),
Paul Docherty and Christopher E Hann (PhD)

May 29, 2009

Department of Mechanical Engineering, Centre for Bio-Engineering
University of Canterbury
Private Bag 4800
Christchurch, New Zealand

Corresponding Author: Prof J. G. Chase
Email: geoff.chase@canterbury.ac.nz

Financial Support: Dr Hann was supported by a NZ Foundation for Research Science and Technology Post-Doctoral Fellowship Grant (Grant: PD20030223) and Dr Lotz by a University of Canterbury Doctoral Scholarship.

Conflicts of Interest: None.

Short Running Title: Minimal C-Peptide Sampling Methods for insulin sensitivity testing

Abstract

Aims and Background: Model-based insulin sensitivity testing via the intravenous glucose tolerance test (IVGTT) or similar is clinically very intensive due to the need for frequent sampling to accurately capture the dynamics of insulin secretion and clearance. The goal of this study is to significantly reduce the number of samples required in intravenous glucose tolerance test protocols to accurately identify C-Peptide and insulin secretion characteristics.

Methods: Frequently sampled, intravenous glucose tolerance test (IVGTT) data from 12 subjects (5 Normal Glucose Tolerant (NGT), 7 Type 2 Diabetes (T2D)) are analyzed to calculate insulin and C-Peptide secretion using a well-accepted C-Peptide model. Samples are reduced in a series of steps based on the critical IVGTT profile points required for accurate estimation of C-Peptide secretion. The full data set of 23 measurements is reduced to sets with 6 or 4 measurements. Peak secretion rate and total secreted C-peptide during 10 and 20 minutes post glucose input, and during the total test time are calculated. Results are compared to those from the full data set Wilcoxon Rank Sum to assess any differences.

Results: In each case, the calculated secretion metrics are largely unchanged, within expected assay variation, and not significantly different from the results obtained using the full 23 measurement data set ($P < 0.05$).

Conclusions: Peak and total C-peptide and insulin secretory characteristics can be accurately estimated in an IVGTT from as few as 4 systematically chosen samples, providing an opportunity to minimize sampling, cost, and burden.

Keywords: Insulin, C-Peptide, secretion, IVGTT

1 Introduction

Assessing pancreatic insulin secretion is important in the diagnosis and monitoring of type-2 diabetes [1–4]. Different tests and markers have been proposed to quantify pre-hepatic insulin secretion. These tests include intravenous tests, such as the hyperglycaemic clamp [5] and the IVGTT [6], oral tests like the OGTT [2, 7], and fasting state assessments [8, 9]. They also vary in resolution and the range of information provided, with the intravenous tests generally providing more details about the bi-phasic secretory characteristics [2].

Good estimation of pre-hepatic insulin secretion can be achieved by estimating C-peptide secretion through modelling of its kinetics [10–14]. This approach is unbiased by first pass hepatic extraction of insulin and is a valid marker due to the equimolar secretion of both peptides [11]. A two compartment model initially proposed by Eaton et al. [11] has been shown to accurately represent C-peptide kinetics. To avoid individual model parameter estimation, Van Cauter et al. [12] proposed a regression model to calculate population parameters from known subject specific characteristics, such as height, weight, age, gender and diagnosis of diabetes. This population methodology has been validated in several studies with peak errors of $10\ \% - 20\ \%$ [12, 15–17].

Accurate estimation of peak secretion rate and total first phase secreted insulin (first 10 minutes) is currently only possible with very frequent sampling during this interval. However, precisely capturing the peak C-peptide concentration and timing is crucial for accurate assessment. Especially given the relatively fast first phase secretion dynamics. Frequent sampling protocols during an IVGTT or similar test, sample C-peptide up to every minute, making these protocols burdensome to the patient, and difficult and costly to perform, as well as requiring significant blood sampling.

For a method to be useful in a clinical diagnostic setting, simplicity, robustness and cost of the protocol are important factors. In this study, a simple

method to estimate C-peptide secretion is proposed, using integrals instead of a typical deconvolution approach. Furthermore, errors introduced by reduced sampling are assessed by comparing different reduced sampling approaches to the full, original frequently sampled data set estimations and values. The analysis is performed on frequently sampled C-peptide data during an IVGTT in five Normal Glucose Tolerant (NGT) and seven subjects with Type 2 Diabetes (T2D).

2 Subjects, Materials and Methods

The C-peptide data from IVGTT studies in this research have been kindly provided by Dr. Andrea Mari (Institute of Biomedical Engineering, Padova, Italy) and Dr. Angelo Avogaro (Department of Clinical and Experimental Medicine, University of Padova, Italy). The data have been previously published [18], with full description of subjects and experimental protocol. The critical aspects are briefly reproduced here for clarity.

2.1 Subjects

The study was performed on 12 subjects, 5 with normal glucose tolerance (NGT, age 24 ± 2, weight 73 ± 6 kg, fasting glucose 5.2 ± 0.1 mmol/l, fasting insulin 50 ± 5 pmol/l) and 7 with type 2 diabtes (T2D, age 49 ± 5, weight 81 ± 3 kg, fasting glucose 8.6 ± 0.8 mmol/l, fasting insulin 125 ± 27 pmol/l). Pharmacological treatment in T2D was stopped 3 days before the study and all subjects received a 2000 kcal/day diet (50% carbohydrate, 35% fat, 15% protein) for at least 30 days prior to the study.

2.2 Experimental Protocol

An insulin-modified IVGTT was performed on all subjects in the morning after an overnight fast. After three fasting samples at -30, -15 and 0 min, a 0.3 g/kg glucose bolus was injected. At 20 minutes, insulin was infused for 5 minutes, totalling 0.03 U/kg (NGT) and 0.05 U/kg (T2D). Blood samples were collected

at 2, 3, 4, 5, 6, 8, 10, 15, 20, 25, 30, 40, 60, 80, 100, 120, 140, 160, 180, 210, and 240 min, and analyzed for C-peptide, glucose and insulin concentrations. Only the C-peptide samples are of interest in this study.

2.3 C-peptide model

A well accepted two compartment model of C-peptide kinetics is employed, as initially described by Eaton et al. [11]. The equations describing the mass transport between compartments are defined:

$$\dot{C}(t) = -(k_1 + k_3)C(t) + k_2 Y(t) + \frac{S(t)}{V_C} \tag{1}$$

$$\dot{Y}(t) = k_1 C(t) - k_2 Y(t) \tag{2}$$

where $C(t)$ is the concentrations in the central (or plasma) compartment (pmol/l), $Y(t)$ is the concentration in the peripheral (or interstitial) compartment (pmol/l), k_1 and k_2 are transport rates between the compartments (1/min), k_3 is the renal loss from the central compartments (1/min), $S(t)$ is the pancreatic secretion rate (pmol/min), and V_C is the central distribution volume (l). A-priori identification of the kinetic parameters is done with known subject information, as described by Van Cauter et al. [12], which is a well utilized, validated and accurate methodology [15–17].

2.4 Integral-based estimation of C-peptide secretion

Estimation of C-peptide secretion rate $S(t)$ is performed with an integral-based method, previously employed in real-time parameter identification in glycemic control trials in the critically ill [19–21] and related biomedical applications. To best compute the integrals in all time steps, the profile of C-peptide is approximated using linear interpolation between data points, which introduces no additional error over model error [21]. The integral functions also have the ad-

vantage of being robust to noise in the measured data, effectively providing a low-pass filter in the summations involved in numerical integrations [21].

C-peptide secretion rate, $S(t)$, is estimated as a step function, with stepsize of 1 min. Thus, during any given 1 min time interval $t \in [t_0, t_1 = t_0 + 1]$, $S(t)$ is assumed constant. Integrating Equation 1 in the interval $[t_0, t_1]$ yields:

$$\int_{t_0}^{t_1} \dot{C}(t)dt = -(k_1 + k_3) \int_{t_0}^{t_1} C(t)dt + k_2 \int_{t_0}^{t_1} Y(t)dt + \frac{1}{V_C} \int_{t_0}^{t_1} S(t)dt \qquad (3)$$

Solving Equation 2 analytically for $Y(t)$ yields:

$$Y(t) = \int_0^t C_{est}(\tau) e^{-k_2(t-\tau)} d\tau \qquad (4)$$

where C_{est} represents the interpolated C-peptide values estimated from the discrete sampled measurements. Combining Equations 3 and 4 and solving for the assumed constant secretion rate $S_{0,1}$ in this time interval yields:

$$\begin{aligned} S_{0,1} \cdot \frac{t_1 - t_0}{V_C} &= C_{est}(t_1) - C_{est}(t_0) + (k_1 + k_3) \int_{t_0}^{t_1} C_{est}(t)dt \\ &\quad -k_2 \int_{t_0}^{t_1} \int_0^t C_{est}(\tau) e^{-k_2(t-\tau)} d\tau dt \end{aligned} \qquad (5)$$

Repeating this process for the intervals $[t_1, t_2]$, $[t_2, t_3]$ etc, results in a 1-min stepwise constant secretion profile $S(t)$. This estimated $S(t)$ profile is (physiologically) constrained to be non-negative. Smoothing the estimated stepwise

constant profile with a zero-phase 3-point moving average is done to avoid overfitting to noisy data and interpolated measurements [21]. This particular filter was picked as a simple choice that does not require further assumptions and does not introduce a phase lag. This last step is not required in frequently sampled data, but results in a more physiological profile between more sparsely sampled data.

2.5 Points of discontinuity

To minimize the number of samples required to describe secretion characteristics, it is crucial to identify key points of physiological discontinuity in the C-peptide concentration profile. These points of discontinuity are caused by sudden changes in C-peptide concentration due to either endogenous or exogenous input. Common changes in C-Peptide secretion that occur during an IVGTT are shown in Figure 1 and defined:

1. **Injection of glucose (D_1)**: A sudden increase in plasma glucose triggers a secretion burst of stored insulin (first phase) lasting 5-10 minutes, which is often reduced or blunted in type-2 diabetes [22, 23]. In the C-peptide concentration profile this dynamic is seen as a very steep rise immediately after administration of glucose. As glucose is administered between $t = 0$ and $t = 1$ minutes, a lag of one minute is chosen here to account for glucose injection and pancreatic response time.

2. **Peak first phase secretion rate (D_2)**: Peak C-peptide secretion rate determines peak C-peptide concentration during the first 10 min post glucose input. In the concentration profile this point is the maximum value CP_{max}, located at t_{CPmax}, assumed between 0-10 minutes.

3. **End of first phase/Start of second phase secretion (D_3)**: First phase secretion ends after approximately 10 minutes. If high glucose concentrations persist, pancreatic insulin secretion continues to rise or remains elevated over basal levels (second phase) [24]. In the concentration profile, this point can be identified as a local minimum around 10 minutes.

4. **Injection of insulin (D_4)**: A sudden increase in plasma insulin inhibits pancreatic insulin secretion [24]. This response can be significantly delayed or not evident in type-2 diabetes [24]. In the concentration profile this point can be seen as a steepening of the negative slope soon after an exogenous insulin input.

5. **Return to basal secretion rate (D_5)**: This step varies widely in individuals, but is usually more gradual than the preceding factors. In the concentration profile this change is evident when the slopes are tending towards zero and the C-peptide concentrations return to fasting values. To pick a clear point in the curve, this study uses the time of the first value to reach the fasting level.

All five of these points are typically very pronounced and consistent in healthy individuals, but can be very gradual, blunted or non-existent in individuals with diabetes, who have an impaired first phase secretion and often have delays in pancreatic response to glucose and insulin concentration changes. Figure 1 shows examples for NGT and T2D subjects with the identified points of discontinuity. Note that points D_2, D_3 and D_5 can be very variable in different individuals and may introduce errors when generic points are chosen.

2.6 Minimal sampling options

The minimal sample optimization analysis is performed in 5 steps. The original complete data set (Step 1) is the reference to which all of the following steps or simplifications are compared to assess any loss in accuracy or utility. This full data set consists of 23 samples.

Steps 2 and 3 are sample-reduced to keep only the optimal median points of discontinuity identified in NGT (Step 2) and T2D (Step 3) subjects. These two reduced sets require only 6 samples. The points of discontinuity chosen are the median time point values observed over all subjects in the data set utilized.

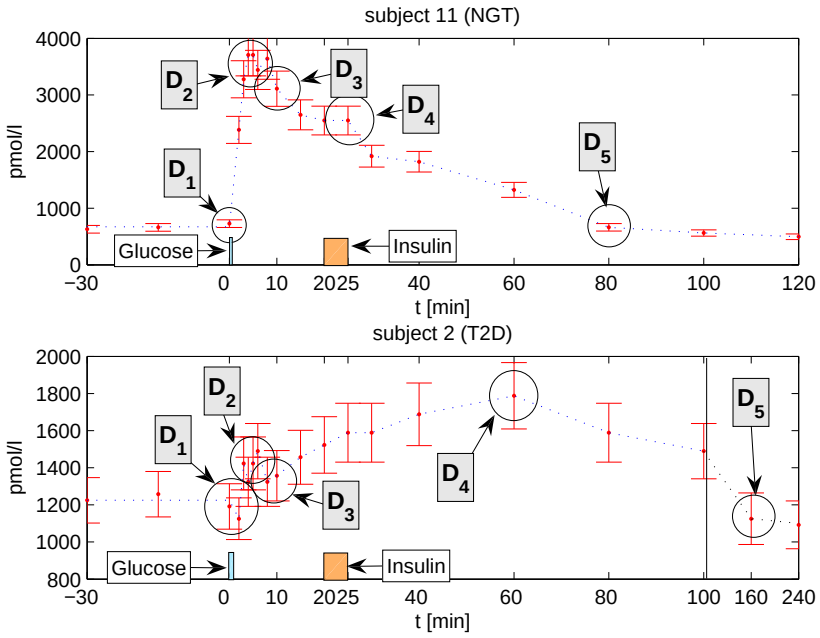

Figure 1: Example of points of discontinuity identified in the C-peptide profile during an IVGTT in NGT (top) and T2D (bottom) subjects. The time axis in T2D is not to scale between 100-240 min.

Using these median values over this diverse data set creates a generic approach that will generalize or extrapolate to any similar data set or study.

Step 4 analyzes a further reduction to 4 samples. This set thus keeps only the most critical points for identifying the dynamics. Specifically, the peak and the return to basal points.

Whereas Steps 1-4 keep the maximal C-peptide sample (D_2)during the first phase response, Step 5 assesses a different approach. More specifically, it is a method that does not rely on exactly capturing the peak concentration. The first sample taken is the sample 2-3 minutes after the median peak time observed over all subjects. To correct for the missing peak sample and timing, an estimated 'correction' sample is introduced at 3 minutes. This estimated point is given a value 10% larger than the actual sample taken 2-3 minutes later. Thus, this estimated value is used to increase the area under the concentration curve to a more physiological value without having to capture it explicitly. Note that while the timing of 3 minutes works well in the data used in this study, this might not be the case for all people and could be a potential source of error. Its validity would have to be assessed in a larger validation study.

These five steps are further clarified in Figure 2 and summarized:

- **Step 1**: Original data set without reduction of samples.
- **Step 2**: Optimized for NGT subjects. 6 Samples at $D_1, t_{CPmax}, D_3, D_4, D_5, t_{end}$.
- **Step 3**: Optimized for T2D subjects. 6 Samples at $D_1, t_{CPmax}, D_3, D_4, D_5, t_{end}$.
- **Step 4**: Further reduction of samples to only include most critical points. 4 Samples at $D_1, t_{CPmax}, D_5, t_{end}$.
- **Step 5**: Sampling missing peak by 2-3 minutes, with 'correction' sample introduced at 2 minutes. 6 Samples at $D_1, (t_{CPmax} - 3), t_{CPmax} + 3, D_3, D_4, D_5, t_{end}$.

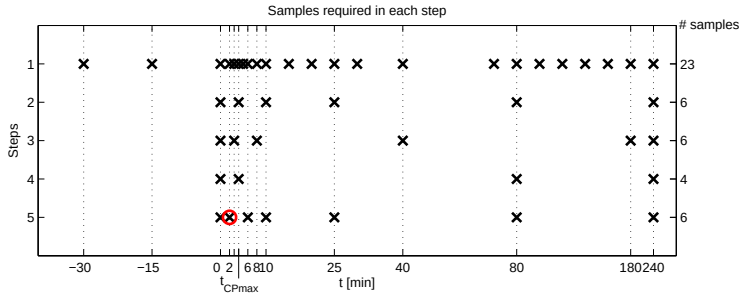

Figure 2: Sample optimization Steps 1-5 and samples used for the calculations in each step. Real samples are marked as 'x', the introduced 'correction' sample in Step 5 is marked with a circle.

The results from each are compared to Step 1 to assess the performance of these reduced sampling schemes in comparison to the original data set.

2.7 Performance metrics

The performance metrics used in this study try to capture all possible secretory characteristics of interest. The goal is to assess possible errors introduced by a reduced sampling protocol. These metrics are defined:

1. **First phase peak secretion rate (S_{max}) and timing of peak (t_{peak})**: Missing samples in the first minutes after glucose input can lead to large errors in the estimated peak secretion rate due to a slower observed increase in C-peptide concentration than actually occurs.

2. **Total C-peptide secreted in first phase, 0-10 min (AUC_{10})**: The area under curve (AUC), or also described as acute insulin response (AIR), is a common metric to describe total insulin secreted during first phase response [2, 9]. It is calculated by integrating estimated secretion rate between 0-10 minutes.

3. **Total C-peptide secreted between glucose and insulin inputs, 0-20 min (AUC_{20})**: As the exogenous insulin inhibits pancreatic insulin secretion, it could be of interest to assess the endogenously secreted insulin until it is inhibited by exogenous insulin.

4. **Total C-peptide secreted during the IVGTT (AUC_{total})**: Calculated by integrating over the complete test, this metric assesses total pancreatic effort.

C-Peptide assays also introduce errors for any data set that will affect the outcome values assessed. These expected error ranges are assessed by Monte Carlo analysis of the estimated secretion rate (10^4 runs), employing normally distributed, zero-mean noise with a coefficient of variation (CV) of 3%. This CV matches currently reported state of the art assays ([25]), and is thus a conservative choice, as older radio immunoassays have CV's of up to twice this value [26], which would result in larger allowable errors from the reduced sampling protocol.

Therefore, the Monte Carlo analysis provides an expected variation for the full set of Step 1 due to assay error. Reduced sampling schemes with results within this assay error range of the Step 1 results would be considered not different. The use of a small, state of the art CV thus restricts this allowable variation to a minimum value.

2.8 Statistical analysis

Non-parametric hypothesis testing with the Wilcoxon Rank Sum test is used to assess if Steps 2-5 are significantly different to Step 1. Normality of results is assessed by the single sample Kolmogorov-Smirnov (KS) test. Where results were log-normally distributed, the log-normal geometric mean and multiplicative standard deviation [27] are used, and specifically noted in the respective results presented.

Table 1: Integral method performance compared to metrics obtained by Mari [18] using a deconvolution approach on the same data. ISR_b is basal secretion rate, ISR_1 mean secretion rate over basal in the 6 minutes post glucose injection, ISR_2 mean secretion rate over basal from 7 minutes until glucose reaches basal levels. Values from this study are converted to match units used by Mari [18].

	ISR_b	ISR_1	ISR_2
	Mean (SEM) in pmol $\cdot$ min$^{-1}\cdot$ m^{-2}		
NGT			
Deconvolution	71 (7)	900 (233)	127 (37)
Integral method	71 (10)	851 (216)	132 (25)
Correlation (P<0.001)	0.93	1.00	0.95
T2D			
Deconvolution	141 (29)	218 (120)	121 (31)
Integral method	136 (30)	277 (136)	130 (34)
Correlation (P<0.001)	0.98	1.00	0.99

3 Results

Pre-hepatic insulin secretion rate was estimated well with the full data set using the integral-based method, resulting in the stepwise constant profiles in Figure 3. The qualitative shape of the secretory curves compare well to the clinical data in the original publication [18]. Mean peak secretion rate is slightly higher in this study for both subgroups, likely due to the smaller stepsize (1 minute vs. 2 minutes) for the estimated secretion rate fitting in this study. Performance of the presented integral method is equivalent to the deconvolution method used by Mari [18], as can be seen by the matching metrics shown in Table 1.

Points of discontinuity are partly given by the protocol, as the timing of glucose and insulin inputs ($D_1 = 1$ min), and are otherwise identified from the sampled C-peptide profile in each subject (D_2, D_3, D_4, D_5). All identified points

are given in Table 2. More variability in all points and especially a distinct lag in D_4 (response to insulin input) and D_5 (return to basal) are evident in T2D subjects, as expected.

Resulting deviations in performance metrics for Steps 2-5, compared to the original sample sets, are shown in Table 3. None of the metrics in Steps 2-4 were significantly different to the corresponding reference metrics in Step 1 in both subgroups (all $P < 0.05$). The distribution of the resulting performance metrics is log-normal and results are thus given using log-normal statistics. Relative differences are normally distributed and are described using normal statistics. Correlations of Steps 2-5 compared to Step 1 are also shown in Table 3.

Errors in performance metrics due to assay errors were assessed by Monte Carlo analysis (10^4 runs) and are given as a CV for each metric, with the median and 100% range over all 12 subjects:

- S_{max}: CV=5.47 %, range $2.97 - 11.01$ %
- AUC_{10}: CV=4.10 %, range $1.92 - 9.39$ %
- AUC_{20}: CV=3.13 %, range $1.90 - 4.26$ %
- AUC_{total}: CV=1.11 %, range $0.97 - 1.25$ %

Reconstruction of C-peptide concentrations from the identified secretion profiles during the analyzed steps resulted in the residuals shown in Figure 4. Residuals are given as relative values (decimal percentages), compared to the complete sampling protocol of Step 1. Deviations from the original sample set are caused by smoothing of the estimated secretion profile, by errors introduced through linear interpolation and, obviously, by the reduced number of sampling steps being examined. The ideal goal is to have all variation within the dashed lines due to assay error.

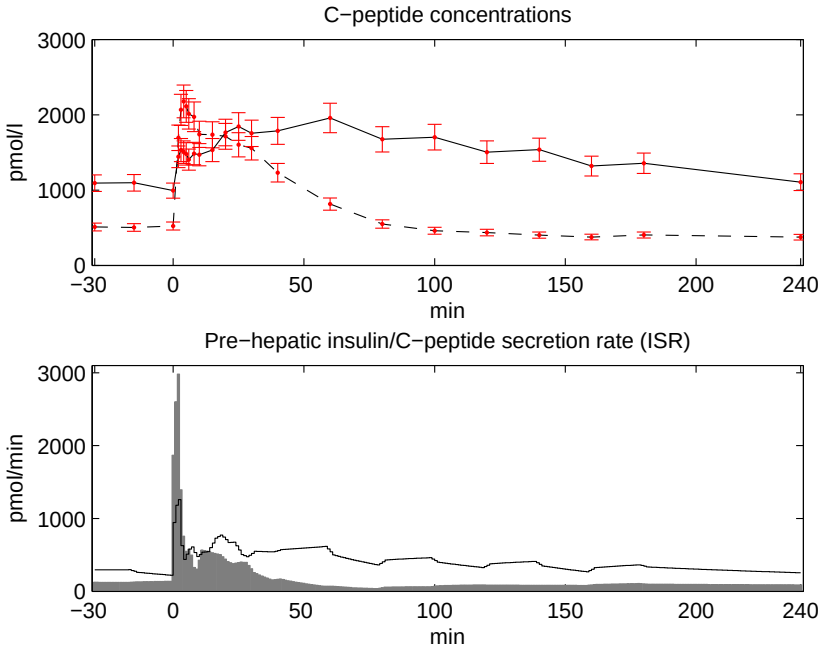

Figure 3: Top: Mean C-peptide concentration in NGT (dashed) and T2D (solid) subjects. Samples from [18] are shown with error bars of $\pm 2\sigma$. Bottom: Mean estimated C-peptide secretion rate (ISR) in NGT (grey area) and T2D (solid line) subjects.

Table 2: Points of discontinuity identified in all subjects. Note that points $D_2 - D_5$ have significant differences between subgroups. The median values are used in the generic selection of points for reduced sample analysis to enable a generalizable approach.

Subject	$\mathbf{D_1}$	$\mathbf{D_2}$	$\mathbf{D_3}$	$\mathbf{D_4}$	$\mathbf{D_5}$
NGT					
7	1	4	10	20	100
8	1	4	10	25	60
9	1	3	10	25	80
10	1	3	8	25	60
11	1	4	10	25	80
median	1	4	10	25	80
SD	0.00	0.55	0.89	2.24	16.73
T2D					
1	1	3	10	60	240
2	1	6	8	60	160
3	1	2	8	30	100
4	1	2	15	25	240
5	1	4	8	40	180
6	1	2	8	40	120
12	1	3	10	20	240
median	1	3	8	40	180
SD	0.00	1.46	2.57	15.92	59.36
Overall					
median	1	3	10	25	100
SD	0.00	1.15	1.98	14.22	71.07

Table 3: Outcomes of sample reduction steps. Shown are the relative percentile changes of Steps 2-5 compared to the reference Step 1 (t_{peak} is given as absolute difference in min). Results from Step 1 have a log-normal distribution and are described by the log-normal geometric mean (geom) and the multiplicative standard deviation (multipl). Relative changes in Steps 2-5 are normally distributed and are described by the mean and standard deviation (SD). Correlations shown are Steps 2-5 compared to Step 1 ($P < 0.001$).

Steps (# samples)		**1** (23)		**2** (6)	**3** (6)	**4** (4)	**5** (6)
		Reference		Percentile change [%]			
NGT							
S_{max}	mean (geom)	2578.8	mean	-5.52	-7.58	-5.52	-1.80
[pmol/min]	SD (multipl)	1.8	SD	1.61	5.68	1.59	5.27
	correlation			1.00	0.99	1.00	1.00
t_{peak}	median	3	median	0	0	0	0
[min]	SD	0.0	SD	0.0	0.0	0.0	0
AUC_{10}	mean (geom)	10301.8	mean	-5.21	2.23	10.10	-0.71
[pmol]	SD (multipl)	1.8	SD	1.56	3.10	9.28	2.243
	correlation			1.00	1.00	0.99	1.00
AUC_{20}	mean (geom)	15110.8	mean	-3.24	-0.54	8.40	-0.62
[pmol]	SD (multipl)	1.7	SD	9.97	9.84	9.34	9.56
	correlation			0.97	0.97	0.97	0.97
AUC_{total}	mean (geom)	42648.9	mean	4.95	19.44	9.30	5.70
[pmol]	SD (multipl)	1.4	SD	4.42	6.53	3.98	3.76
	correlation			0.99	0.99	1.00	1.00
T2D							
S_{max}	mean (geom)	826.3	mean	-15.64	-2.36	-15.30	-0.61
[pmol/min]	SD (multipl)	2.5	SD	14.32	12.07	13.91	21.47
	correlation			1.00	1.00	1.00	0.96
t_{peak}	median	3	median	0	0	0	0
[min]	SD	0.8	SD	0.8	0.8	3.0	1.11
AUC_{10}	mean (geom)	4600.7	mean	-6.40	-0.73	5.42	-2.39
[pmol]	SD (multipl)	2.5	SD	5.83	9.27	27.14	10.88
	correlation			1.00	1.00	0.99	1.00
AUC_{20}	mean (geom)	9439.8	mean	-4.23	-4.47	-6.76	-2.23
[pmol]	SD (multipl)	2.4	SD	9.38	12.74	20.27	8.66
	correlation			0.99	0.98	0.95	0.99
AUC_{total}	mean (geom)	82523.7	mean	0.43	-0.52	-1.32	0.63
[pmol]	SD (multipl)	2.2	SD	5.18	4.36	6.21	5.39
	correlation			1.00	1.00	1.00	1.00

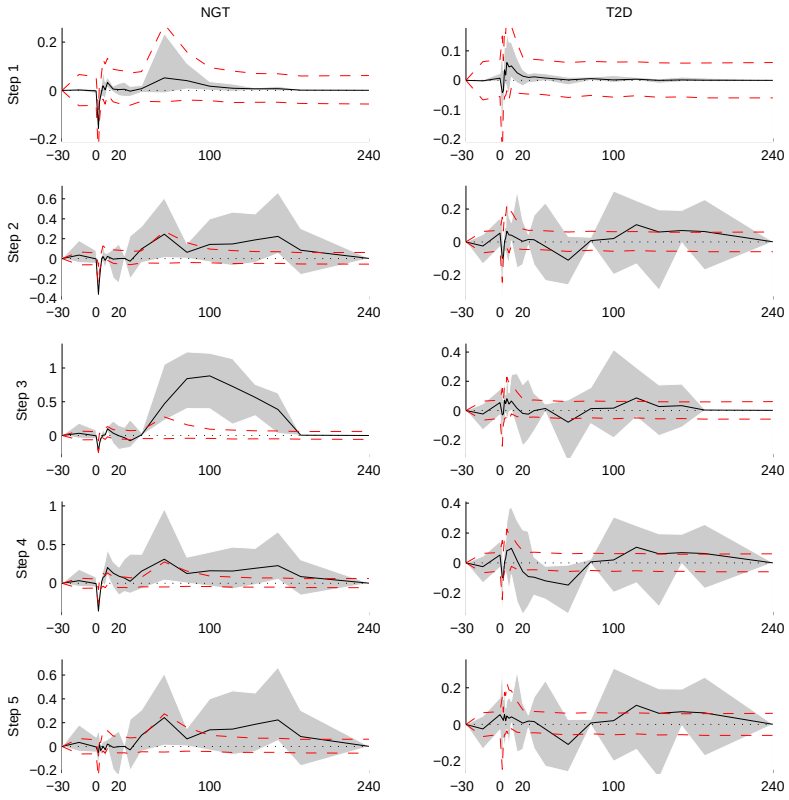

Figure 4: Residuals introduced when reproducing the full C-peptide sample profile with the sample reduction steps. The solid line shows the mean residuals and the grey area shows the full 100 % range of residuals in that step. The dashed lines show the 95 % range of residuals for Step 1 (full sample set) that is introduced by assay error, as estimated by Monte Carlo analysis.

4 Discussion

Estimating pre-hepatic insulin secretion through modelling of C-peptide kinetics has been a common methodology and is relatively less-invasive to perform in research settings [12–14, 28]. In particular, the population method proposed by Van Cauter et al. [12] enables the estimation of secretion rate with a single experiment. By employing this method, model parameters are consistent across studies, enabling a better comparison, as tradeoffs between estimated parameters and secretion rates are reduced. Nonetheless, the estimation of peak secretion rate and insulin secreted during first phase is still highly dependent on assay errors and sampling frequency during the initial minutes. Ideally, sampling should be performed every minute to assess an accurate profile, which introduces significant labor, cost, and burden, as well as reducing the robustness of the method.

It is important to keep in mind that significant errors are also introduced due to assay inaccuracy. Thus, for example, peak estimated secretion rate, S_{max}, has a median CV= 5.47 % and can therefore vary between ±11 % (±2 σ), even with a 1-minute step sampling protocol. Most of the performance metrics are within, or slightly outside of ±2 σ of assay error. This result implies that they are, in fact, just within the natural variability that can be identified [26].

Using the same model and Van Cauter et al.'s parameter estimation method, estimation of secretion rate has previously been proposed by deconvolution [11] and a more elaborate constrained regularization method [14]. The main drawbacks of these methods are the individualized method adjustments required for each subject, including knot placements for cubic spline interpolations [11], or a separate step to find the optimal proportionality constant in each subject [14]. All of these extra steps introduce time, computation and human variability into the results. While criteria are available to determine optimal regularization parameters, such as Maximum Likelihood, these methods require underlying a-priori assumptions about the parameter solutions and thus add complexity to

the approach.

In contrast, the integral-based method described in this study is a single step, computationally convex and fast method that only requires linearly interpolated data. By constraining the resulting linear least squares estimation to non-zero values and smoothing the estimated secretion rate, the resulting profile is physiologically accurate and the effects caused by noisy data are reduced [21]. First and second phase secretion characteristics were clearly identified, with slight quantitative, but not qualitative, deviations from the profiles originally reported with this data by Mari [18]. In addition, these smaller deviations can be readily explained by the longer stepsize used in that study. In a quantitative comparison, the secretion metrics obtained with the integral-based method compare very well to the secretion metrics calculated by Mari [18] using a deconvolution method, as seen in Table 1. Correlations of the subgroup results are all very high, showing equality in performance of both methods when the full sample data set is used.

The reduction of samples was approached by identifying key points of discontinuity and reducing the sample set to those points. Points of discontinuity varied only slightly during first phase secretion in both subgroups, but were significantly delayed in T2D during the second secretion phase. This delay can be attributed to a delay in pancreatic response to insulin input in the case of D_4 [24] and an increased total demand and production rate during this stage, in D_5 [1].

Comparing Steps 2 and 3, which are optimized for the NGT and T2D subgroups respectively, it can be seen that maximum secretion rate S_{max} is more accurate in the subgroup for which it was optimized. This result is especially valid for the T2D subjects in Step 3. This result also holds is for total secreted C-peptide AUC_{total} in NGT and T2D, but is not the case in the other metrics.

The standard deviations of the metrics are mostly very broad in T2D, especially in the metrics during the first minutes. This result indicates a very broad variability in the estimated metric. This variability may be due in part to the strongly blunted first phase response in T2D, resulting in a weak signal to noise ratio and thus exaggerating the effects of assay errors. Nonetheless, none of the sample reduced steps were statistically significantly different than the reference Step 1 ($P < 0.05$).

As can be seen in the residuals reported in Figure 4, Step 2 has a clear advantage over Step 3 in NGT subjects. In T2D subjects a slight advantage for Step 3 is evident between 0-20 minutes, but the remaining time is equivalent to Step 2. This behavior could be due to the fact that the points of discontinuity in the time after insulin input are not as distinct in T2D, and thus not as critical if inaccurately chosen. Larger residuals appear after $t = 20$ minutes in all cases, where sampling is less frequent. During the first section up to $t = 20$ minutes, residuals are mostly within the assay variation bounds shown for the full sample set (Step 1), giving accurate estimations of the most dynamic secretory characteristics. Overall, Step 2 seems to be the better choice if one generic setting were chosen for both types of subjects examined.

In Step 4, where samples are further reduced to a total of 4, residuals are more variable, but still within similarly tight ranges, as in the previous steps. In particular, the first phase section is well represented and captured. In NGT subjects, residuals are even tighter than in Step 3, which has two additional samples that are not optimally placed for this group.

Finally, Step 5 analyzes a different approach by introducing a calculated 'correction' sample to make up for the missing concentration peak sample. This step appears to give the tightest residuals during the first phase, even tighter than the full sampling set. This unexpected result is due to a more accurate fast rise in concentration, as the sample is introduced at $t = 2$ minutes, resulting

in a higher secretory peak. Without this correction sample, linear interpolation from 1-6 minutes would result in a far slower secretion rate increase and a more constant and non-physiological estimated secretion rate during these initial 5 minutes. Hence, the resulting area under the concentration curve is more physiological, which results in a more accurate integrated secretion curve and thus better residuals. During the later phase of these tests, the residuals are identical to Step 2 because the same sample timings are used.

Overall, it can be seen that reduced sampling does not necessarily compromise the information that can be gathered from such a test. This is clearly visible by the very high correlations shown in Table 3 between the full and the sample reduced steps. However, smart sample placement is critical and needs to be chosen correctly according to the secretory information of interest to the researcher. Steps 2 and 3 propose optimized sampling protocols for NGT and T2D subgroups respectively, enabling the investigator to decide on an optimal sampling schedule when designing a test protocol. Even a heavily reduced and generic protocol using only 4 (17%) of the original 23 samples (Step 4), results in acceptable accuracy in the stated performance metrics, most of which are still within reported assay errors.

While the methods developed in this study performed well on the presented dataset, it could be argued that the number of subjects is insufficient to validate the approach. We acknowledge that the number of subjects used to derive the presented method is limited. The goal of the study was not to clinically validate the method, but to derive and present a new method to estimate insulin secretion, that is more robust and automated compared to previously presented methods. In that sense, it should be regarded as a pilot study to derive a new method. This new method would have to be validated in a separate study on a different dataset to prove its validity. In addition, the use of physiologically relevant points of discontinuity that are readily recognizable and well-accepted adds weight to the underlying assumption that the results of this limited pilot

analysis would carry through in a larger study.

In a similar approach, the analysis presented in this study could also be applied to a C-peptide dataset without insulin administration. New points of discontinuity would have to be defined as they could differ slightly, particularly around the time when insulin is administered. We believe that the approach would work just as well on such a dataset, but it is out of the scope of this study to analyze different trial protocols. This could be analyzed in a separate study with a corresponding dataset.

We believe that our approach is novel compared to other methods to estimate insulin secretion presented in the past. Strong emphasis was placed on developing a robust and convex method that would allow automated analysis of C-peptide data without requiring manual intervention or a-priori assumptions about the solutions. While methods presented in the past have primarily focused on accuracy on full data sets, our approach has been primarily on a method that could be applied to reduced data sets and thus be more useful in routine clinical testing environments, where time and cost contribute greatly to the success of a test.

5 Conclusions

Estimation of pre-hepatic insulin or C-peptide secretion can be achieved using an easy to apply population model in combination with a simple and consistent integral-based deconvolution method. Reduction of samples to reduce test complexity, clinical burden, and cost can be done without significantly reducing the accuracy of the test. If smart sample placements are chosen by identifying key points of discontinuity these reductions are readily enabled, saving significant cost and burden.

The approaches presented in this study include sampling optimized for NGT

or T2D subjects (6 samples), a further reduction to 4 samples, and a final option that does not require samples during the first 5 minutes after glucose administration by introducing an additional calculated 'correction' sample. Each step further reduces the sampling stress, cost and blood taken. Overall, the results show that reduced sampling has no clinical or research "cost" in the outcome metrics derived, as shown by Monte Carlo and statistical results, but can enable significantly simpler test protocols.

6 Acknowledgements

The authors would like to thank Prof. Andrea Mari and Prof. Angelo Avogaro for kindly providing the IVGTT data used in this study. Uli Göltenbott was supported with a Baden-Württemberg-Scholarship from the Landesstiftung Baden-Württemberg, Germany.

References

[1] E. Ferrannini, A. Gastaldelli, Y. Miyazaki, M. Matsuda, A. Mari, and R. A. DeFronzo. beta-Cell function in subjects spanning the range from normal glucose tolerance to overt diabetes: a new analysis. *J Clin Endocrinol Metab*, 90(1):493–500, 2005.

[2] E. Ferrannini and A. Mari. Beta cell function and its relation to insulin action in humans: a critical appraisal. *Diabetologia*, 47(5):943–56, 2004.

[3] M. Stumvoll, A. Fritsche, and H. U. Haring. Clinical characterization of insulin secretion as the basis for genetic analyses. *Diabetes*, 51 Suppl 1:S122–9, 2002.

[4] R. N. Bergman, M. Ader, K. Huecking, and G. Van Citters. Accurate assessment of beta-cell function: the hyperbolic correction. *Diabetes*, 51 Suppl 1:S212–220, 2002.

[5] R. A. DeFronzo, J. D. Tobin, and R. Andres. Glucose clamp technique:

a method for quantifying insulin secretion and resistance. *Am J Physiol*, 237(3):E214–23, 1979.

[6] R. N. Bergman, L. S. Phillips, and C. Cobelli. Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose. *J Clin Invest*, 68(6):1456–1467, 1981.

[7] M. Stumvoll, A. Mitrakou, W. Pimenta, T. Jenssen, H. Yki-Jarvinen, T. Van Haeften, W. Renn, and J. Gerich. Use of the oral glucose tolerance test to assess insulin release and insulin sensitivity. *Diabetes Care*, 23(3):295–301, 2000.

[8] D. R. Matthews, J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher, and R. C. Turner. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*, 28(7):412–9, 1985.

[9] G. Pacini and A. Mari. Methods for clinical assessment of insulin sensitivity and beta-cell function. *Best Pract Res Clin Endocrinol Metab*, 17(3):305–22, 2003.

[10] O. K. Faber, C. Hagen, C. Binder, J. Markussen, V. K. Naithani, P. M. Blix, H. Kuzuya, D. L. Horwitz, A. H. Rubenstein, and N. Rossing. Kinetics of human connecting peptide in normal and diabetic subjects. *J Clin Invest*, 62(1):197–203, 1978.

[11] R. P. Eaton, R. C. Allen, D. S. Schade, K. M. Erickson, and J. Standefer. Prehepatic insulin production in man: kinetic analysis using peripheral connecting peptide behavior. *J Clin Endocrinol Metab*, 51(3):520–8, 1980.

[12] E. Van Cauter, F. Mestrez, J. Sturis, and K. S. Polonsky. Estimation of insulin secretion rates from C-peptide levels. Comparison of individual and standard kinetic parameters for C-peptide clearance. *Diabetes*, 41(3):368–77, 1992.

[13] K. S. Polonsky, J. Licinio-Paixao, B. D. Given, W. Pugh, P. Rue, J. Galloway, T. Karrison, and B. Frank. Use of biosynthetic human C-peptide in the measurement of insulin secretion rates in normal volunteers and type I diabetic patients. *J Clin Invest*, 77(1):98–105, 1986.

[14] R. Hovorka, P. A. Soons, and M. A. Young. ISEC: a program to calculate insulin secretion. *Comput Methods Programs Biomed*, 50(3):253–64, 1996.

[15] R. Hovorka, E. Koukkou, D. Southerden, J. K. Powrie, and M. A. Young. Measuring pre-hepatic insulin secretion using a population model of C-peptide kinetics: accuracy and required sampling schedule. *Diabetologia*, 41(5):548–54, 1998.

[16] G. Toffolo, F. De Grandi, and C. Cobelli. Estimation of beta-cell sensitivity from intravenous glucose tolerance test C-peptide data. Knowledge of the kinetics avoids errors in modeling the secretion. *Diabetes*, 44(7):845–54, 1995.

[17] C. N. Jones, D. Pei, P. Staris, K. S. Polonsky, Y. D. Chen, and G. M. Reaven. Alterations in the glucose-stimulated insulin secretory dose-response curve and in insulin clearance in nondiabetic insulin-resistant individuals. *J Clin Endocrinol Metab*, 82(6):1834–8, 1997.

[18] A. Mari. Assessment of insulin sensitivity and secretion with the labelled intravenous glucose tolerance test: improved modelling analysis. *Diabetologia*, 41(9):1029–39, 1998.

[19] X. W. Wong, J. G. Chase, G. M. Shaw, C. E. Hann, T. Lotz, J. Lin, I. Singh-Levett, L. J. Hollingsworth, O. S. Wong, and S. Andreassen. Model predictive glycaemic regulation in critical illness using insulin and nutrition input: a pilot study. *Med Eng Phys*, 28(7):665–81, 2006.

[20] J. G. Chase, G. M. Shaw, J. Lin, C. V. Doran, C. Hann, T. Lotz, G. C. Wake, and B. Broughton. Targeted glycemic reduction in critical care using closed-loop control. *Diabetes Technol Ther*, 7(2):274–82, 2005.

[21] C. E. Hann, J. G. Chase, J. Lin, T. Lotz, C. V. Doran, and G. M. Shaw. Integral-based parameter identification for long-term dynamic verification of a glucose-insulin system model. *Comput Methods Programs Biomed*, 77(3):259–270, 2005.

[22] M. J. Davies, G. Rayman, A. Grenfell, I. P. Gray, J. L. Day, and C. N. Hales. Loss of the first phase insulin response to intravenous glucose in subjects with persistent impaired glucose tolerance. *Diabet Med*, 11(5):432–6, 1994.

[23] S. Del Prato, P. Marchetti, and R. C. Bonadonna. Phasic insulin release and metabolic regulation in type 2 diabetes. *Diabetes*, 51 Suppl 1:S109–116, 2002.

[24] Leonard S. Jefferson and Alan Cherrington. *The endocrine pancreas and regulation of metabolism*, volume 2 of *Handbook of physiology - The endocrine system*. Oxford University Press, Oxford, 2001.

[25] Roche. Data Sheet - C-Peptide Immunoassay, Elecsys 1010/2010/Modular Analytics E170. Technical Report 03184897 190, Roche Diagnostics, Mannheim, Germany, 2005.

[26] P. M. Clark. Assays for insulin, proinsulin(s) and C-peptide. *Ann Clin Biochem*, 36 (Pt 5):541–64, 1999.

[27] E. Limpert, W. A. Stahel, and M. Abbt. Log-normal distributions across the sciences: Keys and clues. *Bioscience*, 51(5):341–352, 2001.

[28] R. M. Watanabe, A. Volund, S. Roy, and R. N. Bergman. Prehepatic beta-cell secretion during the intravenous glucose tolerance test in humans: application of a combined model of insulin and C-peptide kinetics. *J Clin Endocrinol Metab*, 69(4):790–7, 1989.