



OPEN ACCESS

EDITED BY
Li Zhang,
University of London, United Kingdom

REVIEWED BY
Marcos Matabuena,
University of Santiago de Compostela,
Spain
Francisco Gude,
Health Research Institute of Santiago de
Compostela (IDIS), Spain

*CORRESPONDENCE
Oleksandr Sverdlov
✉ oleksandr.sverdlov@novartis.com

RECEIVED 22 June 2023
ACCEPTED 14 August 2023
PUBLISHED 11 September 2023

CITATION

Cui EH, Goldfine AB, Quinlan M, James DA
and Sverdlov O (2023) Investigating the
value of glucodensity analysis of
continuous glucose monitoring data in
type 1 diabetes: an exploratory analysis.
Front. Clin. Diabetes Healthc. 4:1244613.
doi: 10.3389/fcdhc.2023.1244613

COPYRIGHT

© 2023 Cui, Goldfine, Quinlan, James and
Sverdlov. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Investigating the value of glucodensity analysis of continuous glucose monitoring data in type 1 diabetes: an exploratory analysis

Elvis Han Cui¹, Allison B. Goldfine², Michelle Quinlan³,
David A. James⁴ and Oleksandr Sverdlov^{3*}

¹Department of Biostatistics, University of California, Los Angeles, Los Angeles, CA, United States, ²Division of Translational Medicine, Cardiometabolic Disease, Novartis Institutes for Biomedical Research, Cambridge, MA, United States, ³Early Development Analytics, Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States, ⁴Methodology and Data Science, Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States

Introduction: Continuous glucose monitoring (CGM) devices capture longitudinal data on interstitial glucose levels and are increasingly used to show the dynamics of diabetes metabolism. Given the complexity of CGM data, it is crucial to extract important patterns hidden in these data through efficient visualization and statistical analysis techniques.

Methods: In this paper, we adopted the concept of glucodensity, and using a subset of data from an ongoing clinical trial in pediatric individuals and young adults with new-onset type 1 diabetes, we performed a cluster analysis of glucodensities. We assessed the differences among the identified clusters using analysis of variance (ANOVA) with respect to residual pancreatic beta-cell function and some standard CGM-derived parameters such as time in range, time above range, and time below range.

Results: Distinct CGM data patterns were identified using cluster analysis based on glucodensities. Statistically significant differences were shown among the clusters with respect to baseline levels of pancreatic beta-cell function surrogate (C-peptide) and with respect to time in range and time above range.

Discussion: Our findings provide supportive evidence for the value of glucodensity in the analysis of CGM data. Some challenges in the modeling of CGM data include unbalanced data structure, missing observations, and many known and unknown confounders, which speaks to the importance of—and provides opportunities for—taking an approach integrating clinical, statistical, and data science expertise in the analysis of these data.

KEYWORDS

CGM, functional data analysis, glucodensity, pharmacodynamics, visualization

1 Introduction

Type 1 diabetes (T1D) is a life-threatening autoimmune disease with no cure (1). Insulin replacement is lifesaving and treats symptoms of T1D, but does not alter disease progression. Glycated hemoglobin A_{1c} (HbA_{1c}) is an established glycemic physiological biomarker, which is widely used for diagnosing, monitoring, and informing treatment decisions in the management of T1D. However, HbA_{1c} measurements may exhibit high variability and can be affected by various biological and analytic factors, which may complicate an accurate assessment of blood glucose and glycemic control in T1D (2).

Continuous glucose monitoring (CGM) is a technology increasingly used to capture the dynamics of metabolism in T1D (3). CGM provides a more complete assessment of glycemia than what is possible with intermittent evaluations by standard home blood glucose monitoring or HbA_{1c}-based approaches. Modern CGM systems are wearable devices that can transmit glucose readings every 1 min–15 min to connected technologies, such as a smartphone, portable reading device, computer, thereby enabling patients, caregivers, and physicians to monitor glucose levels over time and make informed decisions in diabetes management. CGM is also an integral component of artificial pancreas systems—more advanced technologies that provide the means for an automatic insulin delivery at the right dose and at the appropriate times for the individual patient (4). The use of dashboard systems—software technologies integrating relevant clinical information, including CGM data into a unified display—hold promise for improving population-level management of T1D, especially in pediatric patients (5).

The structure of CGM data is complex. CGM measurements per individual represent high-frequency time series data capturing the dynamics of interstitial glucose concentrations. Although these data (in addition to HbA_{1c} measurements and other relevant biomarkers) can provide important information on diabetes metabolism, they require careful preprocessing and are challenging to analyze statistically (6, 7). As CGM data are frequently acquired under free-living conditions, they are subject to various sources of variability and confounding effects (8). Traditional statistical methods may not adequately capture the dynamic nature of the CGM data and the underlying patterns. Functional data analysis may provide a potentially useful alternative approach (9, 10). By considering the entire individual glucose trajectory as a functional unit and incorporating appropriate statistical models, functional data analysis may help researchers to characterize and compare glucose profiles, identify temporal patterns, and assess the impact of various factors on glucose dynamics (11, 12). Recently, Matabuena et al. (13) introduced a new functional representation of CGM data, termed “glucodensity”, with the corresponding statistical toolkit for analyzing glucodensities. They gave several important examples from both clinical practice and biomedical research, where the concept of glucodensity may be promising and provide a more accurate characterization of glucose metabolism than the more standard approaches. The predictive potential of glucodensity was also validated in the AEGIS study of long-term changes in glucose levels (14). Furthermore, the distributional presentation and

analysis of high-frequency biomedical data can be useful in other areas of clinical research. For example, Matabuena et al. (15) showed the utility of this approach by deriving clinical phenotypes of older adults based on their accelerometry data and demonstrating the value of such phenotyping in predicting 5-year mortality and survival.

CGM data are now frequently collected in T1D clinical trials, commonly as a part of exploratory objectives and potentially as an outcome measure (16–19). In contrast to self-measured glucose using a fingerstick, CGM can provide continuous real-time measurement of glucose levels in an ambulatory setting, with the assessment of glycemic variability, comprehensive measurement of hyperglycemic and hypoglycemic exposure, and safety alerts for glycemic extremes. This extensive glycemic data should be of value for clinical trials of diabetes pharmacotherapies or strategies including nutrition and exercise. However, to date, the utility of CGM data acquired during a clinical trial has been somewhat limited, in part because the value of CGM-derived outcomes has not yet been fully recognized by regulators (i.e., the US Food and Drug Administration) or payers (i.e., insurance companies) as an indicator of safety or effectiveness.

This report provides insights on important statistical analysis issues with CGM data that arise in the context of randomized clinical trials of new investigational drugs for T1D. We highlight some challenges and outline ways to manage them using real data from an ongoing randomized, placebo-controlled phase 2a proof-of-concept study of an investigational medicinal product (IMP) in pediatric individuals and young adults with new-onset T1D. This study aimed to advance valuable approaches for clinical investigators, statisticians, and data scientists involved in the design and analysis of clinical trials collecting CGM data; however, the paper is not meant to provide comprehensive solutions.

2 Background and research questions

Consider a clinical research study with CGM data collection. As the actual IMP itself is not relevant to the CGM data, and because the study is ongoing, details of the IMP are not discussed. In brief, the trial is a non-confirmatory, randomized 2:1, placebo-controlled, investigator- and subject-blinded, parallel-arm, phase 2a trial to assess safety, tolerability, pharmacokinetics (PK), and the early efficacy of the IMP on the preservation of residual pancreatic beta-cell function in new-onset T1D in pediatric and young adult subjects. Eligible participants (newly diagnosed T1D patients, 12–21 years of age weighing between 30 kg and 125 kg) are enrolled into the trial within 8 weeks of the time of diagnosis based on the results of both screening and baseline visits [ClinicalTrials.gov Identifier: NCT04129528].

In this study, CGM data are collected as an exploratory objective for describing the dynamics of T1D metabolism. Study participants are provided with CGM devices (Dexcom G6) and supplies and asked to wear a CGM device for at least 3 consecutive days (10 days are preferred and feasible with the single sensor) at baseline prior to the first dose, at day 1 after receiving the first dose of the study drug, and at months 3, 6, 9, and 12 during the 1 year of treatment. Sufficient supplies are provided for continuous wear,

which is encouraged. The acquired CGM data evaluate glycemic parameters, such as the mean glucose level, variability, time in the glycemic range, hyper and hypoglycemia, and duration, on the assessment days (20).

In this context, several important questions warrant investigation:

1. How can one perform an informative visualization of the longitudinal CGM data that would also be appealing from the clinician's perspective? This is important as current visualizations are generally snapshots of 10 days to 14 days of data.
2. Based on the CGM data, can clinically meaningful subgroups be identified through the clustering of study participants?
3. Using the CGM data before and after randomization (i.e., "pretreatment" and "posttreatment"), can the treatment effect within a subject be quantified and can its significance be tested statistically (e.g., using an analog of a paired t -test)?
4. Is there a way to perform a statistical significance test comparing treatment effects (IMP vs. placebo) with respect to the CGM data, possibly accounting for some important covariates [e.g., an analog of a two-sample t -test or an analysis of covariance (ANCOVA)]?

To address the first two questions, we performed exploratory data analysis using a subset of data from the described phase 2a trial. These results are based on blinded data (i.e., the individual randomized treatment group information is unavailable), and they are presented in Section 4. To address the last two questions, some additional information on the individual treatment assignments and treatment periods is required. At this point, this information is unavailable because the study is still ongoing. Therefore, we present only the relevant data analysis strategy as part of the discussion in Section 5.

3 Materials and methods

Table 1 below shows an example of the CGM data structure. The first column corresponds to the subject ID, the second column

refers to the time at which the glucose levels were recorded, and the third column contains the values of the glucose levels (mg/dL). In our example, the glucose levels are recorded every 5 min. In all conducted analyses, we used all available valid CGM observations (i.e., observations not flagged as erroneous in the dataset) from the study participants, without the knowledge of their treatment assignments and without the knowledge of treatment periods within the subjects. Therefore, our analyses can be viewed as *unsupervised* learning.

3.1 Data visualization

Visualization of CGM data is of great importance to provide accessibility for the comprehension of the information in large datasets. Broll et al. (21) provided a comprehensive list of CGM-based metrics and an R package, *iglu*, for visualization purposes. We used the *iglu* package to obtain lasagna plots of average glucose levels per subject over the 24-h period. Also, we created plots of the available individual raw CGM data over 24 h, along with the mean ambulatory glucose profiles (AGPs) per subject. To visualize dynamic changes of glycemia within subjects, we created plots of individual time-in-range values compared with the study day.

3.2 Clustering using glucodensities

We adopt the concept of glucodensity (13) and demonstrate how one could benefit from a cluster analysis based on estimated glucodensities. The glucodensity is a probability density function describing the distribution of glucose levels over time. Formally, let $Y(t)$ denote the glucose level measured by wearable devices at time $t \in [0, T]$. The glucodensity $f(s)$ is defined as follows (13):

$$f(s) = \frac{\partial}{\partial s} F(s), \text{ where } F(s) = \frac{1}{T} \int_0^T I(Y(t) \leq s) dt. \quad (1)$$

In Equation 1, $I(\cdot)$ stands for the indicator function. This formulation implicitly assumes that the glucose level $Y(t)$ has a common density $f(s)$ for all $t \in [0, T]$ (22). A kernel density estimator (KDE) is applied for estimating $f(s)$. Let $y_i, i = 1, \dots, m$

TABLE 1 Example of a CGM data structure.

Subject ID	Time	Glucose level (mg/dL)
Subject 00000001	2022-01-01 10:55:00 UTC	155
Subject 00000001	2022-01-01 11:00:00 UTC	147
Subject 00000001	2022-01-01 11:05:00 UTC	138
Subject 00000002	2022-05-01 15:55:00 UTC	89
Subject 00000002	2022-05-01 16:00:00 UTC	97
Subject 00000002	2022-05-01 16:05:00 UTC	98
...

The first column corresponds to the subject's ID, the second column refers to the time at which the glucose levels were recorded, and the third column contains the values of the glucose levels (mg/dL).

be the realizations of $Y(t)$ at time points t_1, \dots, t_m and $K_h(\cdot)$ be a kernel with bandwidth h . Then $f(s)$ is estimated by:

$$\hat{f}(s) = \frac{1}{m} \sum_{i=1}^m K_h(s - y_i). \quad (2)$$

Note that the measurements y_i in Equation 2 are not independent. Theoretical justifications of the KDE for dependent data can be found in Hall et al. (23) and Bosq (22). Furthermore, Matabuena et al. (13) showed how the concept of the 2-Wasserstein distance (24) and energy-based methods (25, 26) can be applied to cluster-estimated glucodensities. We used the R package *biosensors.usc*, provided by Matabuena et al. (13), to estimate individual glucodensities and perform the corresponding cluster analysis.

To facilitate a clinical interpretation of the identified clusters, we performed the statistical comparisons of the clusters using analysis of variance (ANOVA) with respect to the baseline level of the C-peptide (an established biomarker of pancreatic beta-cell function) and with respect to conventional CGM-derived parameters such as the time in range (TIR), time above range, and time below range.

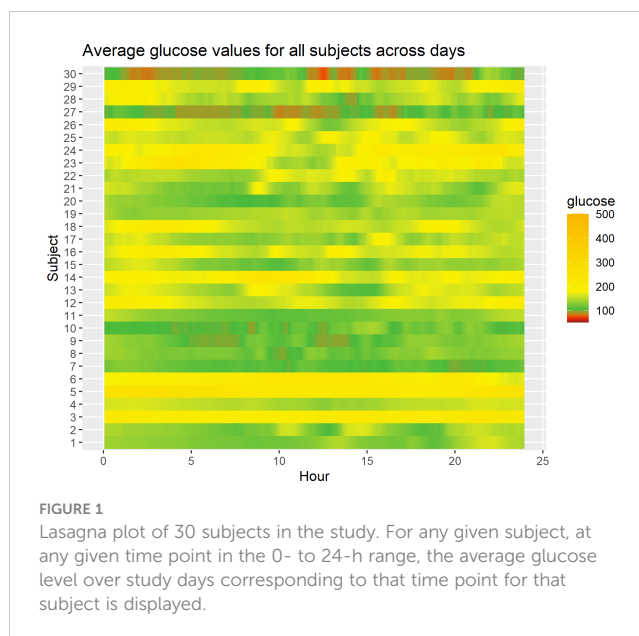
4 Results

4.1 Analysis dataset

Our analysis dataset included 30 subjects and 760,510 valid records in total. During the data preprocessing step, device error values were identified and flagged in the database (they represented < 5% of the observations) and discarded from the analysis. An assessment of potential bias due to temporary loss of data capture is outside the scope of the present paper. The overall time range from the first to the last observation for the 30 subjects was 4 days to 518 days, with a mean of 158.9 days and a median of 115.5 days. We defined the metric *coverage* as the number of days with valid observations covering at least 70% of each day, in line with the International Consensus on the Use of Continuous Glucose Monitoring (19, 20), which supports a minimum of approximately 70% of possible CGM readings over 14 days to enable an optimal glycemic assessment for real-time decision-making. Larger gaps in a 24-h period are generally considered insufficient for describing a daily profile. We note that our analysis differed in that we were assessing prespecified glycemic observation intervals to estimate glycemia over a broader interval of time, but we employed the same standard for a minimal definition of a complete day capture. In our data, the CGM coverage achieved by the 30 subjects ranged from 3 days to 468 days, with a mean of 84.7 days, a median of 45.5 days, a first quartile of 16.3 days, and a third quartile of 120 days.

4.2 Data visualization

Figure 1 displays a lasagna plot (27) of the average glucose level per subject over the 24-h period (i.e., for any given hour in the 0- to 24-h range, what is displayed is the average over the study days



corresponding to that hour for that subject). One can observe some differences, both between subjects and within subjects, as highlighted by different colors in the heatmap. For example, some subjects have predominantly green profiles corresponding to glucose levels between ≈ 100 mg/dL and 180 mg/dL, whereas others have yellow profiles corresponding to glucose levels > 200 mg/dL. Also, one can see that some subjects have alternating patterns, for example periods of green and yellow and sometimes red, which correspond to low blood glucose levels, that is, roughly < 70 mg/dL. One limitation of the lasagna plot is that it reflects the average values only and not the variability of the measurements. Furthermore, it does not account for the fact that some participants may have more data than others, for example, because of the different length of time in the study or the different patterns of wearing the CGM devices. Finally, the lasagna plots do not demonstrate improving or worsening glycemic trends over time for the individual, rather the plots demonstrate glycemic averages for a given time of day.

Figure 2 displays all available CGM data for 30 subjects in the study. In each panel, the x-axis represents hours within the 0- to 24-h interval, and the y-axis refers to the glucose levels (mg/dL). Within a panel, the blue curves correspond to the raw CGM data acquired from a subject on different study days, and a red curve represents the mean AGP, obtained using smoothing techniques (28). The black dashed lines represent the target range of 70 mg/dL to 180 mg/dL (3.9 mmol/L–10 mmol/L). One can see from Figure 2 that there is different amount of data per subject, reflecting the fact that some participants have been longer in the study or been wearing the CGM device for more days, and contributed more CGM data than others. For some subjects, there is evidence that the AGP is within the desired range, whereas for others it is crossing the upper limit of 180 mg/dL over the 24-h period. Notably, we observed some periodicity of AGPs over the day for some individuals. One should be mindful that the results depend on the amount of data, and the findings for an individual subject may change as more data for this subject are

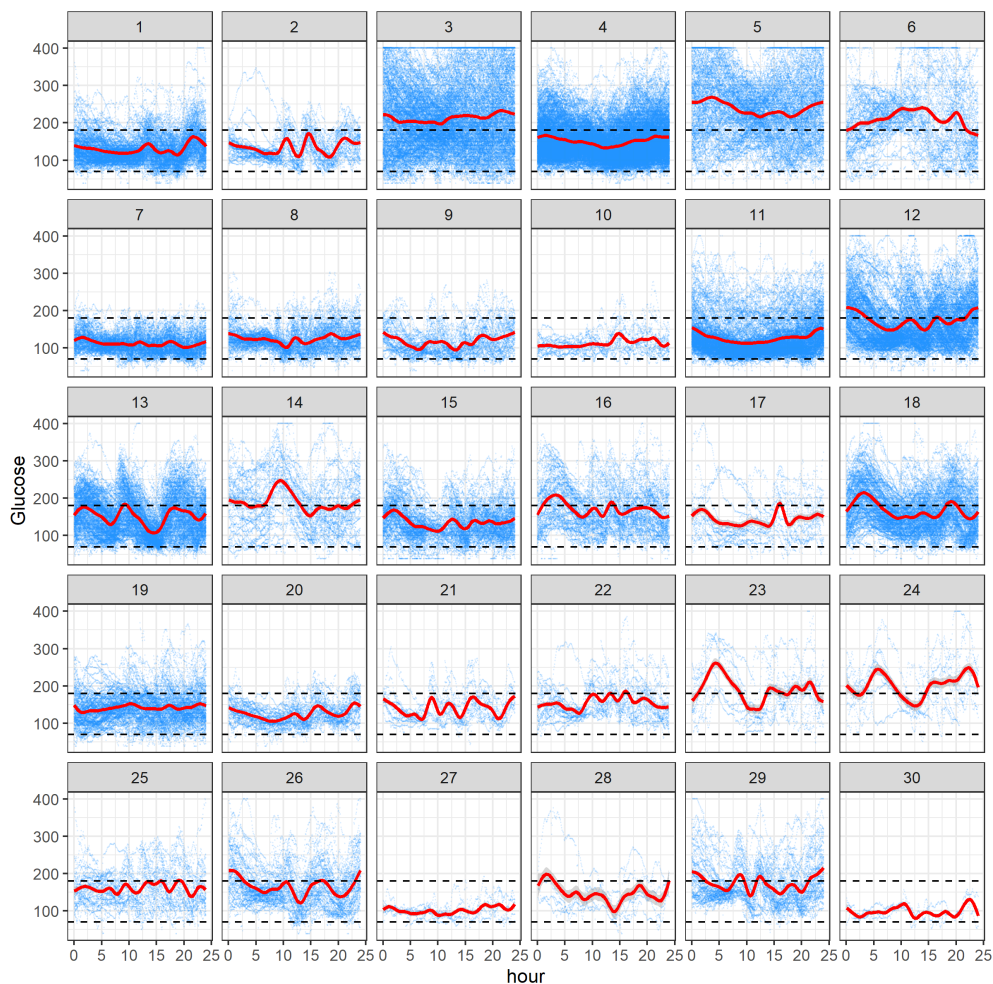


FIGURE 2

Ambulatory glucose profile (AGP) plots of 30 subjects in the study. Within each subject panel, the blue curves correspond to the raw CGM data acquired from the subject on different study days, and a red curve represents the mean AGP. The black dashed lines represent the range of 70 mg/dL to 180 mg/dL.

acquired during the study. Similarly to lasagna plots, these plots do not demonstrate improvement or worsening glycemic trends over time for the individual, rather they show glycemic averages for a given time of day.

Figure 3 visualizes the TIR (i.e., percentage of CGM glucose readings in the target range of 70 mg/dL to 180 mg/dL) for 30 study participants. In each panel, the x -axis represents the study day, and the y -axis represents the TIR (proportion, measured on a scale of 0 to 1). Note that the x -axis range is different across the subjects, as it reflects the different number of days a given participant contributed CGM data in the study. The y -axis range is also displayed in a subject-specific manner, as there was a substantial variation in the TIR values across the subjects. Within a panel, the blue curve represents individual TIR values, the red curve is a modeled mean profile obtained using smoothing techniques, and the gray area quantifies uncertainty around the mean (28). The value of the TIR $\geq 70\%$ is consistent with the clinically desirable target, as suggested by the American Diabetes Association (29, 30). From Figure 3, one can get some useful insights into the individual glycemic control over time. Some subjects have relatively stable patterns (e.g., subject 7), whereas

for others there is evidence of fluctuation—increasing or decreasing the TIR values (e.g., subject 11)—indicating improved or worsened glycemic control over time. Clearly, the results depend on the amount of CGM data per subject, with more data enabling a more robust exploratory assessment. Importantly, these plots demonstrate the variability of the TIR across the duration of CGM use.

4.3 Clustering using glucodensities

Figure 4 shows the estimated and clustered glucodensities (left three panels) and the corresponding cumulative distribution functions (right three panels) for our dataset. Each glucodensity curve in a cluster corresponds to an individual subject, and they were constructed based on pooled CGM data from the individual. There are three identified clusters: red (six subjects), with the highest average and most variable levels of blood glucose; blue (11 subjects), with the somewhat better glycemic control; and, green (13 subjects), with the lowest average and least variable levels of blood glucose.

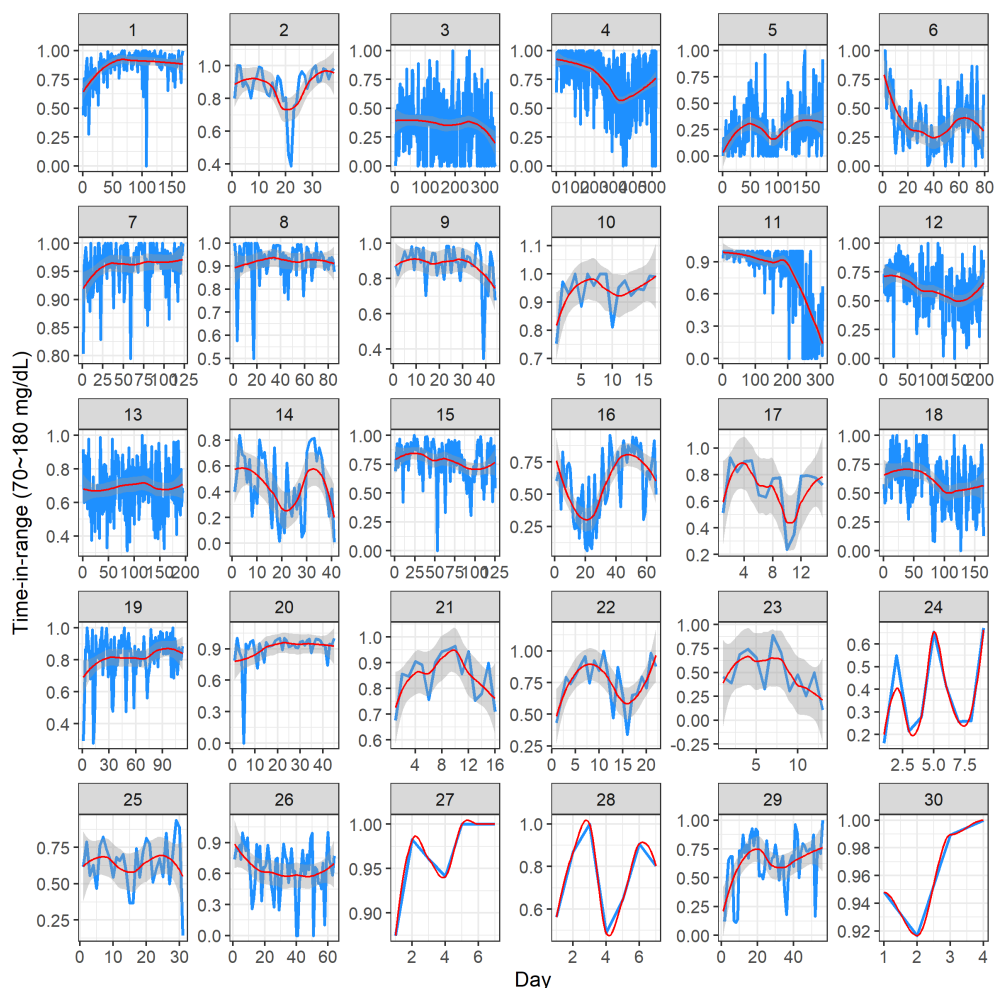


FIGURE 3

Dynamic time-in-range plots of 30 subjects in the study. Within each subject panel, the blue curve represents the individual time-in-range value on a given study day, the red curve is a modeled mean profile, and the gray area represents the uncertainty around the mean. The x-axis range is different across the subjects, as it reflects the different number of days of CGM data per subject. The y-axis range is displayed in a subject-specific manner because there was a substantial variation in the time-in-range values across the subjects.

Figure 5 shows boxplots of individual values and the results of non-parametric ANOVA comparing three clusters with respect to the baseline C-peptide levels (top left plot), TIR (the proportion of CGM measurements in the range of 70 mg/dL to 180 mg/dL; top right plot), time above range (the proportion of CGM measurements > 180 mg/dL; bottom left plot), and time below range (the proportion of CGM measurements < 70 mg/dL; bottom right plot).

From Figure 5 (top left plot), there is evidence of an overall difference among the three clusters with respect to the baseline C-peptide levels (Kruskal–Wallis $p = 0.042$), with the difference being pronounced between clusters 1 and 3 (the red and green clusters, $p = 0.024$) and clusters 2 and 3 (the blue and green clusters, $p = 0.065$), but not between clusters 1 and 2 (the red and blue clusters, $p = 0.52$).

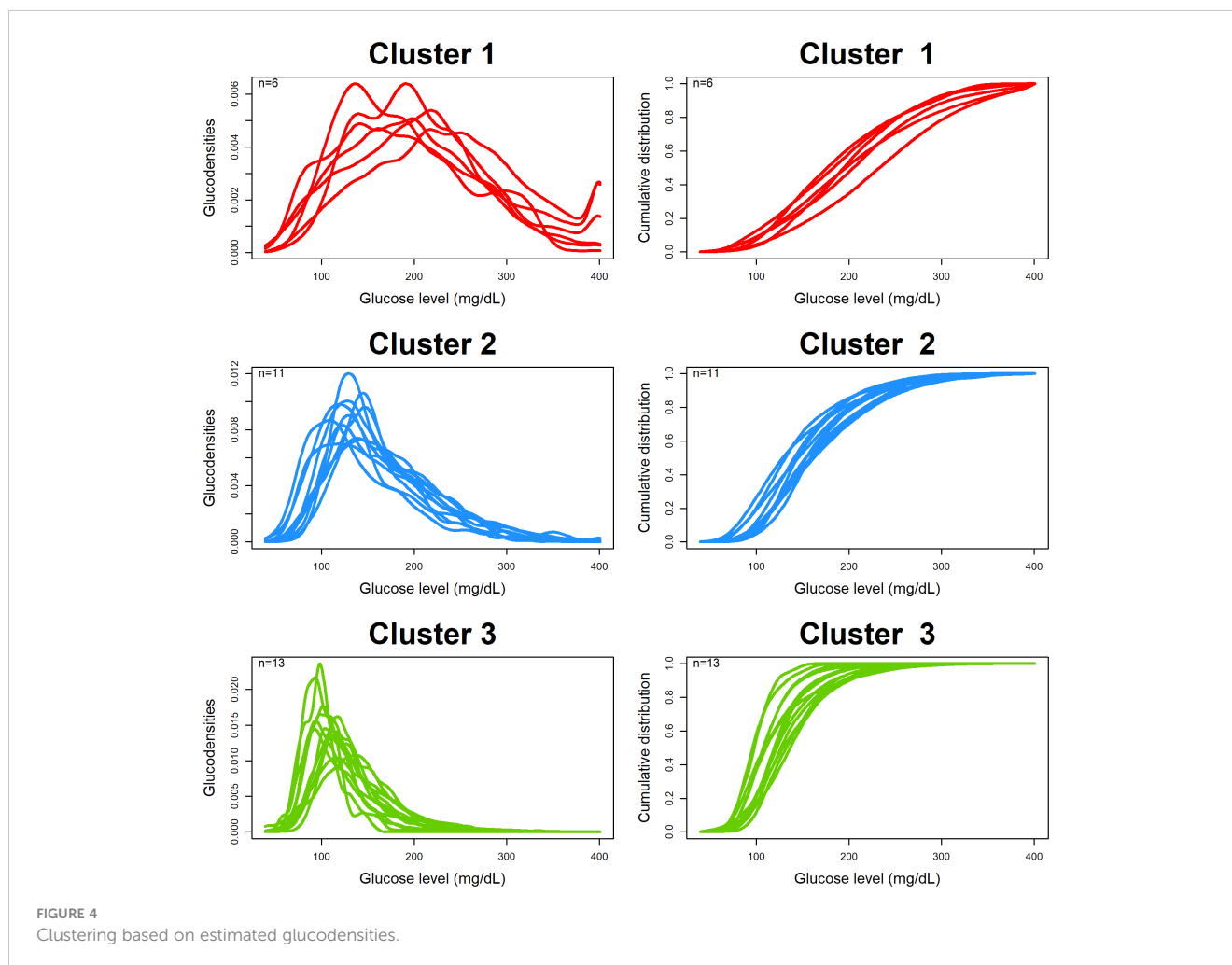
Also, one can see statically significant evidence of the difference between the three clusters with respect to the TIR (Figure 5, top right plot) and time above range (Figure 5, bottom left plot). In regard to the time below the range (Figure 5, bottom right plot), the overall difference among the three clusters is not statistically significant (Kruskal–Wallis $p = 0.11$); however, there is some

evidence of difference between clusters 2 and 3 (blue and green clusters, $p = 0.047$).

5 Discussion

5.1 CGM data provides important information on T1D metabolism

CGM data collection is increasingly common in diabetes research (6, 21, 31–33). Recently, many statistical software packages have been developed to facilitate CGM data visualization and analysis. Table 2 provides a summary of some of these packages. The *iglu* package (21) was useful for obtaining visual displays (e.g., lasagna plots) of individual CGM-derived metrics over a 24-h period. However, such plots reflect an overall measurement, but not the dynamics of individual glycemic parameters over time. We propose displaying relevant CGM metrics, such as the TIR versus the study day, to capture the



dynamics of glycemic control and variability of the selected CGM parameters over time and simultaneously display the fact that some subjects contributed more CGM data in the study than others.

In the present work, we evaluated glucodensity (13) and performed a cluster analysis of estimated glucodensities based on the blinded CGM data of 30 participants from an ongoing study. Our analysis indicates that a clinically meaningful clustering using glucodensities is possible. We identified three clusters of patients with different shapes of glucodensities indicating different levels of blood glucose control. We found evidence of a statistically significant difference among the clusters with respect to some established biomarkers of T1D, such as residual pancreatic beta-cell function (C-peptide levels) at baseline, thereby indicating a potential grouping of patients with different severities of T1D. We also found evidence of cluster differences with respect to conventional CGM-derived parameters such as the TIR and time above range. Overall, the added value of statistical comparisons of the identified clusters using ANOVA, with respect to established biomarkers of pancreatic beta-cell function, is that it aids the clinical interpretation of the clusters. Although our findings are limited by the small sample size and the exploratory nature of the analysis, they are promising to characterize pancreatic function in stage 3 recent-onset T1D. Further analyses to validate these results are warranted once the study achieves the final database lock.

In our analysis, we focused only on the estimation and clustering of individual glucodensities. We did not attempt to estimate an average glucodensity curve for each cluster and the corresponding uncertainty. The estimation of “average glucodensity” from a sample of individual glucodensities and quantifying the associated uncertainty requires applying more advanced functional data analysis techniques such as the 2-Wasserstein distance and the Fréchet mean-of-the-density functions (13, 39). In addition, as the estimated glucodensities can be used as a basis for deriving CGM parameters (e.g., mean, standard deviation, coefficient of variation, TIR), another important problem is establishing a statistical agreement between empirical (i.e., obtained directly from the CGM data) and glucodensity-based CGM parameters. These additional questions require further careful investigation, and we defer them to future work.

5.2 Next step: quantifying treatment effects and treatment contrasts

An important objective of a randomized controlled trial in T1D is to compare treatment effects (IMP vs. placebo) by applying a statistical test on some clinically meaningful outcome measure. For

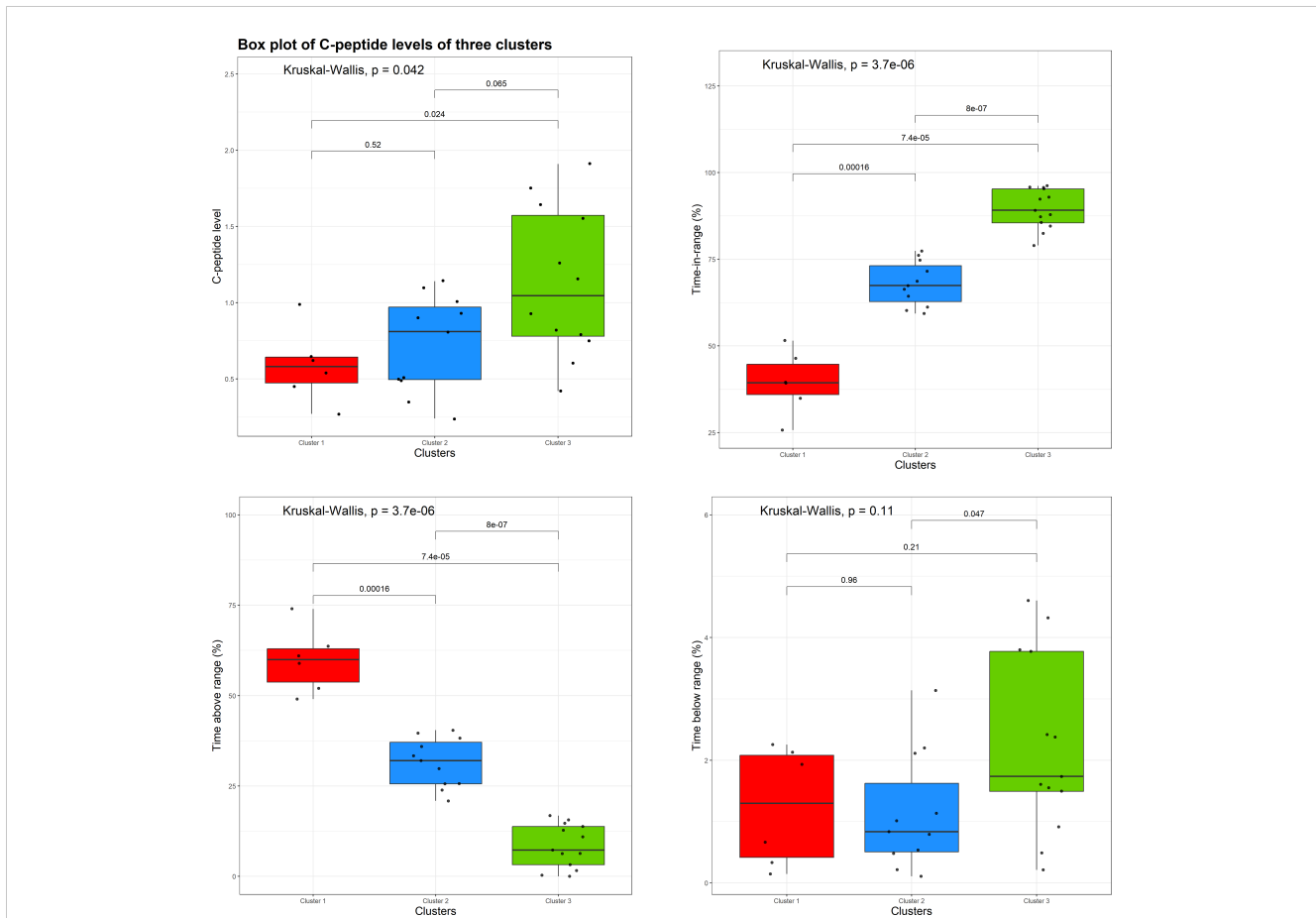


FIGURE 5 Boxplots of individual values and the results of a non-parametric analysis of variance (ANOVA) comparing three clusters with respect to the baseline C-peptide level (top left plot), time in range (top right plot), time above range (bottom left plot), and time below range (bottom right plot).

this purpose, one can consider different CGM-derived metrics, such as the TIR, which is defined as either the percentage of CGM glucose values or the number of hours spent in the range of 70 mg/dL to 180 mg/dL during the measurement period (20, 40). One recent study found that 14 days of CGM data correlated well with 3 months of CGM data, and, within these, 14 days of at least 70% of CGM wear should provide robust data for deriving the TIR (41).

In the final analysis of our T1D clinical trial, for each trial participant, one can derive the baseline (pretreatment) TIR value and the TIR after 1 year of treatment, by taking their difference and then testing the significance of this difference (separately for the subjects in the IMP group and subjects in the placebo group) using a paired *t*-test. Furthermore, to assess treatment contrast (IMP vs. placebo), one can apply a two-sample *t*-test on the change from

TABLE 2 R packages for analyzing CGM data.

Package	Reference	Usage	Remarks
<i>Iglu</i>	Broll et al. (21)	Exploratory data analysis	Comprehensive and user-friendly for visualization; limited input size
<i>CGManalyzer</i>	Zhang et al. (34)	Exploratory data analysis	Designed to read and organize the CGM data; no R Shiny available
<i>cgmanalysis</i>	Vigers et al. (35)	Exploratory data analysis	Designed to read and organize the CGM data; no R Shiny available
<i>biosensors.usc</i> (glucodensity)	Matabuena et al. (13)	Density estimation and cluster analysis	The first R shiny for modeling the CGM data; regression and clustering are available
<i>fda</i>	Ramsay et al. (36)	Functional data analysis	A standard package for analyzing functional data; preprocessing is required for the CGM data
<i>Refund</i>	Kokoszka and Reimherr (37)	Functional data analysis	A standard package for analyzing functional data; preprocessing is required for the CGM data
<i>fdANOVA</i>	Górecki and Smaga (38)	Functional data analysis	Designed to perform functional ANOVAs for intensive longitudinal data; useful for hypothesis testing with the CGM data

baseline values of the TIR. Alternatively, one can consider fitting an ANCOVA model as follows:

$$TIR_{1,i} = \mu + \alpha \cdot \delta_i + \beta \cdot TIR_{0,i} + \varepsilon_i, i = 1, \dots, n, \quad (3)$$

where $TIR_{1,i}$ is the i th participant's outcome after 1 year of treatment and $TIR_{0,i}$ is the similar value at baseline, $\delta_i = 1$ (or 0), if the i th participant is assigned to the IMP (or the placebo), μ is the overall mean, α is the effect due to treatment (IMP), β is the effect due to TIR at baseline, and ε_i 's are independent, normally distributed measurement errors. By testing the hypothesis $H_0: \alpha = 0$ we can address the following question: does the active treatment (IMP) have a significant effect on the TIR compared with placebo after 1 year of treatment?

Furthermore, extensions of the model (3) could be considered. In our T1D study example, the CGM assessments are made every 3 months during the 1 year of treatment. Although the primary interest is in the outcome at 12 months, the outcomes at earlier time points may be of interest as well. In this case, one could consider fitting a mixed-effects model with repeated measurements (MMRM) (42).

The TIR is one CGM-based outcome measure that can be analyzed using the approach described above. Other CGM-based outcome measures, such as the average glucose level, glucose measurement index, and glucose coefficient of variation (CV), can be similarly analyzed. To obtain a more complete assessment of the treatment effect, one can perform statistical analyses on multiple clinically important CGM-based measures and present the results (estimated treatment effects with corresponding confidence intervals and p -values) in one table. However, a limitation of such an approach is at least twofold. First, different CGM metrics are likely to be correlated, but the analyses of individual metrics do not account for such correlations. In fact, proper statistical adjustments for multiplicity are required to mitigate the risk of false-positive findings. Second, by deriving these different metrics, some important information may be lost. For example, the TIR does not account for the dynamic nature of the CGM data, nor the possibly unequal number of CGM data from individuals. This prompts the consideration of more complex CGM-based objects, such as glucodensities, for analyses.

Matabuena et al. (13) noted that glucodensity can be used "to establish if there are statistically significant differences between patients subjected to different interventions, for example, in a clinical trial." Furthermore, Matabuena et al. (14, 15, 43) proposed methods and estimators for handling missing data and exploring their potential values in analyzing CGM data. In our considered setting, the main objects of interest are probability density functions, and functional data analysis techniques may be useful for statistical inference. For the i th individual in the sample, we assume there exists a "true" glucodensity f_{0i} that describes the distribution of the individual's glucose levels at baseline. It is plausible to assume that the individual glucodensity may change over time (e.g., due to treatment intervention received in the study, environmental factors, or disease progression). Let f_{1i} denote the i th subject's glucodensity after 1 year of treatment. An important question is:

how could we measure "similarity" of f_{0i} and f_{1i} ? One could consider the 2-Wasserstein distance (13):

$$d_{W^2}(f_{0i}, f_{1i}) = \sqrt{\int_0^1 (F_{0i}^{-1}(s) - F_{1i}^{-1}(s))^2 ds}, \quad (4)$$

where F_{0i} and F_{1i} stand for the cumulative distribution functions of the glucodensities f_{0i} and f_{1i} . Metric (4) has some computational and modeling advantages, and it has a physical interpretation in the theory of optimal transport.

In practice, the functions f_{0i} and f_{1i} are unknown, but they can be consistently estimated from the CGM data of the i th individual by \hat{f}_{0i} and \hat{f}_{1i} ; see Equation (2). Then, the distance $d_{W^2}(\hat{f}_{0i}, \hat{f}_{1i})$ can be computed to quantify the difference between the i th subject's empirical glucodensities during pretreatment and posttreatment periods. Note that the pretreatment and posttreatment CGM data of an individual may be highly correlated. A statistical significance test may be considered for testing:

$$H_{0i}: f_{0i} = f_{1i} \text{ for } i = 1, \dots, n. \quad (5)$$

Furthermore, let $\mu_0(s)$ denote the "average" glucodensity in the population of study subjects in the pretreatment period, and $\mu_1(s)$ denote a similar object for the period after 1 year of treatment. A research question is: are $\mu_0(s)$ and $\mu_1(s)$ different? In other words, one may be interested in testing $H_0: \mu_1(s) = \mu_0(s)$ based on the CGM data from the sample of study participants (separately for the IMP group and the placebo group). This problem can be viewed as a generalization of a paired t -test.

Taking a step further, for comparing treatment effects, one can consider the Wasserstein–Fréchet regression model using 1-year posttreatment glucodensity as the outcome, and treatment (IMP or placebo) and other clinically relevant covariates as predictors (39). This approach can be viewed as a generalization of an ANCOVA model (3) to functional outcomes. Relevant technical details, including estimation, statistical testing procedures, and asymptotic results, are beyond the scope of the current work. Petersen et al. (39) provided an example of applying functional regression analysis to post-intracerebral hemorrhage hematoma densities and found it promising in that context. We plan to address similar research problems for the CGM data and glucodensities in future work.

5.3 Conclusion and future work

This article outlined several research questions arising in the context of randomized clinical trials in T1D with CGM data and provided some preliminary ideas on how to tackle them in practice. Efficient visualization of the CGM data and cluster analysis using glucodensities are useful tools for exploratory analysis of the CGM data. Future work is needed to evaluate additional important issues. For instance, the estimand framework is increasingly useful in biopharmaceutical product development (44). Defining estimands based on glucodensities is a challenging but important and necessary next step for rigorously addressing clinical research questions concerning CGM data. Another important area of

research is related to experimental design issues—sample size determination, choice of data collection time windows and time points, quantifying the amount of information from these data (e.g., Fisher's information matrix)—and formulation of various optimal design problems. Finally, we think that the CGM data analysis provides an opportunity for a collaborative effort, integrating clinical, statistical, data science, and pharmacometrics expertise. Such a combination is increasingly viewed as being synergistic in solving complex problems in biomedical research (45, 46).

Data availability statement

The datasets presented in this article are not readily available because of informed consent and confidentiality restrictions. The study is still ongoing. Requests to access the datasets should be directed to alex.sverldov@novartis.com.

Ethics statement

This study has been registered at <https://classic.clinicaltrials.gov/ct2/show/NCT04129528>. The studies were conducted in accordance with local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

AG and MQ were involved in the design of the original study. OS and MQ conceptualized the statistical analysis plan. EC, MQ,

DJ, and OS performed the statistical analysis. OS wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors are grateful to the Editorial Board and the two independent reviewers for their thorough reading of the manuscript and constructive feedback. This work was completed as part of EC's 2022 biostatistics summer internship in Early Development Analytics at Novartis.

Conflict of interest

The analyzed data is from a study sponsored by Novartis Institutes for Biomedical Research. Authors AG, MQ, DJ, and OS are employed by the company Novartis.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Atkinson MA, McGill DE, Dassau E, Laffel L. Type 1 diabetes mellitus. In: Melmed S, Auchus RJ, Goldfine AB, Koenig RJ, Rosen CJ, editors. *Williams textbook of endocrinology, 14th edition*. Philadelphia, PA: Elsevier (2020). p. 1403–37.
- Service FJ. Glucose variability. *Diabetes* (2013) 62(5):1398–404. doi: 10.2337/di12-1396
- Riddle MC, Gerstein HC, Cefalu WT. Maturation of CGM and glycemic measurements beyond HbA1c—A turning point in research and clinical decisions. *Diabetes Care* (2017) 40(12):1611–3. doi: 10.2337/dci17-0049
- Trevitt S, Simpson S, Wood A. Artificial pancreas device systems for the closed-loop control of type 1 diabetes. *J Diabetes Sci Technol* (2016) 10(3):714–23. doi: 10.1177/1932296815617968
- Ferstad JO, Vallon JJ, Jun D, Gu A, Vitko A, Morales DP, et al. Population-level management of type 1 diabetes via continuous glucose monitoring and algorithm-enabled patient prioritization: Precision health meets population health. *Pediatr Diabetes* (2021) 22(7):982–91. doi: 10.1111/pedi.13256
- Clarke JW, Kovatchev B. Statistical tools to analyze continuous glucose monitor data. *Diabetes Technol Ther* (2009) Suppl 1:S45–54. doi: 10.1089/dia.2008.0138
- Gaynanova I. Digital biomarkers of glucose control—reproducibility challenges and opportunities. *Biopharmaceutical Rep Spring* (2022) 2022:21–6.
- Thabit H, Leelarathna L, Wilinska ME, Elleri D, Allen JM, Lubina-Solomon A, et al. Accuracy of continuous glucose monitoring during three closed-loop home studies under free-living conditions. *Diabetes Technol Ther* (2015) 17(11):801–7. doi: 10.1089/dia.2015.0062
- Guo W. Functional mixed effects models. *Biometrics* (2002) 58:121–8. doi: 10.1111/j.0006-341X.2002.00121.x
- Ferraty F, Vieu P. *Nonparametric functional data analysis*. Springer New York, NY: Springer Series in Statistics (2006). doi: 10.1007/0-387-36620-2
- Gecili E, Huang R, Khoury JC, King E, Altaye M, Bowers K, et al. Functional data analysis and prediction tools for continuous glucose-monitoring studies. *J Clin Trans Sci* (2020) 5(1):e51. doi: 10.1017/cts.2020.545
- Mahmoudi Z, Del Favero S, Jacob P, Choudhary P, on behalf of the Hypo-RESOLVE Consortium. Toward an optimal definition of hypoglycemia with continuous glucose monitoring. *Comput. Methods Programs Biomedicine* (2021) 209:106303. doi: 10.1016/j.cmpb.2021.106303
- Matabuena M, Petersen A, Vidal JC, Gude F. Glucodensities: A new representation of glucose profiles using distributional data analysis. *Stat Methods Med Res* (2021) 30(6):1445–64. doi: 10.1177/0962280221998064
- Matabuena M, Felix P, Garcia-Meixide C, Gude F. Kernel machine learning methods to handle missing responses with complex predictors. Application in modelling five-year glucose changes using distributional representations. *Comput Methods Programs Biomedicine* (2022) 221:106905. doi: 10.1016/j.cmpb.2022.106905
- Matabuena M, Félix P, Hammouri ZAA, Mota J, del Pozo Cruz B. Physical activity phenotypes and mortality in older adults: A novel distributional data analysis of accelerometry in the NHANES. *Aging Clin Exp Res* (2022) 34(12):31073114. doi: 10.1007/s40520-022-02260-3
- Beck RW, Calhoun P, Kollman C. Use of continuous glucose monitoring as an outcome measure in clinical trials. *Diabetes Technol Ther* (2012) 14(10):877–82. doi: 10.1089/dia.2012.0079
- Schnell O, Barnard K, Bergenstal R, Bosi E, Garg S, Guerci B, et al. Role of continuous glucose monitoring in clinical trials: Recommendations on reporting. *Diabetes Technol Ther* (2017) 19(7):391–9. doi: 10.1089/dia.2017.0054

18. Fox BQ, Benjamin PF, Aqeel A, Fitts E, Flynn S, Levine B, et al. Continuous glucose monitoring use in clinical trials for on-market diabetes drugs. *Clin Diabetes* (2021) 39(2):160–6. doi: 10.2337/cd20-0049
19. Battelino T, Alexander CM, Amiel SA, Arreaza-Rubin G, Beck RW, Bergenstal RM, et al. Continuous glucose monitoring and metrics for clinical trials: An international consensus statement. *Lancet Diabetes Endocrinol* (2023) 11(1):42–57. doi: 10.1016/S2213-8587(22)00319-9
20. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International consensus on use of continuous glucose monitoring. *Diabetes Care* (2017) 40(12):1631–40. doi: 10.2337/dc17-1600
21. Broll S, Urbanek J, Buchanan D, Chun E, Muschelli J, Punjabi NM, et al. Interpreting blood glucose data with R package iglu. *PLoS One* (2021) 16:e0248560. doi: 10.1371/journal.pone.0248560
22. Bosq D. *Nonparametric statistics for stochastic processes: Estimation and prediction*. Springer New York, NY: Springer Science & Business Media (2012). doi: 10.1007/978-1-4612-1718-3
23. Hall P, Lahiri SN, Truong YK. On bandwidth choice for density estimation with dependent data. *Ann Stat* (1995) 23:2241–63.
24. Villani C. *Optimal transport: Old and new*. Springer Berlin, Heidelberg: Springer (2009). doi: 10.1007/978-3-540-71050-9
25. Székely GJ, Rizzo ML. The energy of data. *Annu Rev Stat Its Appl* (2017) 4:447–79. doi: 10.1146/annurev-statistics-060116-054026
26. França G, Rizzo ML, Vogelstein JT. Kernel k-groups via Hartigan's method. *IEEE Trans Pattern Anal Mach Intell* (2020) 43:4411–25. doi: 10.1109/TPAMI.2020.2998120
27. Swihart BJ, Caffo B, James BD, Strand M, Schwartz BS, Punjabi NM. Lasagna plots: A saucy alternative to spaghetti plots. *Epidemiology* (2010) 21(5):621. doi: 10.1097/EDE.0b013e3181e5b06a
28. Wood SN. *Generalized additive models: An introduction with R*. Boca Raton, FL: Chapman and Hall/CRC (2006). doi: 10.1201/9781315370279
29. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical targets for continuous glucose monitoring data interpretation: Recommendations from the International Consensus on Time in Range. *Diabetes Care* (2019) 42(8):1593–603. doi: 10.2337/dci19-0028
30. Gabbay MAL, Rodacki M, Calliari LE, Vianna AGD, Krakauer M, Pinto MS, et al. Time in range: A new parameter to evaluate blood glucose control in patients with diabetes. *Diabetol Metab Syndrome* (2020) 12:1–8. doi: 10.1186/s13098-020-00529-z
31. Bergenstal RM. Understanding continuous glucose monitoring data. In: *Role of continuous glucose monitoring in diabetes treatment*. Arlington (VA: American Diabetes Association (2018). doi: 10.2337/db20181-20
32. Nguyen M, Han J, Spanakis EK, Kovatchev BP, Klonoff DC. A review of continuous glucose monitoring-based composite metrics for glycemic control. *Diabetes Technol Ther* (2020) 22:613–22. doi: 10.1089/dia.2019.0434
33. Lin R, Brown F, Ekinici EL. The ambulatory glucose profile and its interpretation. *Med J Aust* (2022) 217(6):295–8. doi: 10.5694/mja2.51666
34. Zhang XD, Zhang Z, Wang D. Cgmanalyzer: an R package for analyzing continuous glucose monitoring studies. *Bioinformatics* (2018) 34:1609–11. doi: 10.1093/bioinformatics/btx826
35. Vigers T, Chan CL, Snell-Bergeon J, Bjornstad P, Zeitler PS, Forlenza G, et al. cgmanalysis: An R package for descriptive analysis of continuous glucose monitor data. *PLoS One* (2019) 14:e0216851. doi: 10.1371/journal.pone.0216851
36. Ramsay JO, et al. *Package 'fda'* (2023). Available at: <https://cran.r-project.org/web/packages/fda/fda.pdf>.
37. Kokoszka P, Reimherr M. *Introduction to functional data analysis*. Boca Raton, FL: Chapman and Hall/CRC (2017). doi: 10.1201/9781315117416
38. Górecki T, Smaga L. fdanova: An R software package for analysis of variance for univariate and multivariate functional data. *Comput Stat* (2019) 34:571–97. doi: 10.1007/s00180-018-0842-7
39. Petersen A, Liu X, Divani AA, Wasserstein F-tests and confidence bands for the Fréchet regression on density response curves. *Ann Stat* (2021) 49(1):590–611.
40. Timmons JG, Boyle JG, Petrie JR. Time in range as a research outcome measure. *Diabetes Spectr* (2021) 34(2):133–8. doi: 10.2337/ds20-0097
41. Riddlesworth TD, Beck RW, Gal RL, Connor CG, Bergenstal RM, Lee S, et al. Optimal sampling duration for continuous glucose monitoring to determine long-term glycemic control. *Diabetes Technol Ther* (2018) 20(4):314–6. doi: 10.1089/dia.2017.0455
42. Mallinckrodt CH, Lane PW, Schnell D, Peng Y, Mancuso JP. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. *Drug Inf J* (2008) 42(4):303–19. doi: 10.1177/009286150804200402
43. Matabuena M, Félix P, Ditzhaus M, Vidal J, Gude F. Hypothesis testing for matched pairs with missing data by maximum mean discrepancy: An application to continuous glucose monitoring. *Am Statistician* (2023), 1–13. doi: 10.1080/00031305.2023.2200512
44. Mallinckrodt C, Molenberghs G, Lipkovich I, Ratitch B. *Estimands, estimators and sensitivity analysis in clinical trials*. Boca Raton, FL: CRC Press (2020).
45. Koch G, Pfister M, Daunhawer I, Wilbaux M, Wellmann S, JE JEV. Pharmacometrics and machine learning partner to advance clinical data analysis. *Clin Pharmacol Ther* (2020) 107(4):926–33. doi: 10.1002/cpt.1774
46. Ryzeknik Y, Sverdlov O, Svensson EM, Montepiedra G, Hooker AC, Wong WK. Pharmacometrics meets statistics—a synergy for modern drug development. *CPT Pharmacometrics Syst Pharmacol* (2021) 10(10):1134–49. doi: 10.1002/psp4.12696