Special Issue 3
2021

The GH-Method

Using Distributional Data Analysis Tools to Investigate the Glucose Density Distribution of the Mean Daily Glucose Values (eAG) for Three Type 2 Diabetes Patients Over an 18-Month Period Based on GH-Method: Math-Physical Medicine (No. 510)

Gerald C. Hsu*

eclaireMD Foundation, USA

Abstract

The author read an article recently, "Glucodensities: a new representation of glucose profiles using distributional data analysis", dated August 19, 2020. Incidentally, he has also made two further improvements on his glucose data analysis with his collected big data of sensor glucose via a continuous glucose monitoring (CGM) sensor device. First, in addition to using the HbA1C, which is the mean value of the past 115 days of red blood cells carried glucose, of a patient is the golden standard in evaluating diabetes condition. He investigates the glucose fluctuation (GF) (glucose excursion or glycemic variability) and then transforms the GF values from a wave's time-space into energy's frequency-space via the Fourier Transform operations. Using this approach, he can then guesstimate the degree of damage on internal organs caused by the energies associated with GF. Although the GF research is one step deeper compared to the study of mean value of glucose, such as HbA1C, it is still not deep enough to provide additional details and useful information hidden within the glucose waves. Second, he realized that the average values or mean values of glucose defined by the American Diabetes Association (ADA) such as the HbA1C or time in range (average glucose within a range) can only provide partial overviews of diabetes condition. However, these basic biomarkers are still missing some hidden internal turmoil, i.e., glucose vibrations or severe stimulations, throughout certain selected timeframes due to all types of external and/or internal stimulators. Therefore, he has defined another term known as the glucose density (GD) in order to explore more and different information hidden within the glucose data and their waveforms. GD is defined as the occurrence frequency at a specific glucose value, for example 2.1% occurrence rate at 110 mg/dL glucose value over a selected time period of collected sensor glucose. In this way, he can then calculate and examine each glucose value's occurrence rate within a glucose range that is suitable to a specific patient. If this glucose examination method is accepted by the medical community, it would be an extremely beneficial tool for doctors to be able to quickly study the conditions of their diabetes patients. Furthermore, the author has also programmed his algorithm into an iPhone app software. Through the combination of his published papers and medical books along with a widely distributed app for patient's use in the future, he believes that worldwide type 2 diabetes (T2D) patients can benefit from his research work. As a part of his follow-on research tasks, he plans to further examine his GD% resulted from certain food/diet nutritional types and exercise intensity levels. Hopefully, his research papers would not be limited within the scope of a "descriptive style using 26 alphabets" but instead as a "quantitative style using 10 digits". Numbers do not lie as long as we don't use fake, unorganized, and/or uncleaned data. Statistics is a tricky tool to use because it has the obvious characteristics of garbage in and garbage out (GIGO). It is also important to know that by using statistics with different selected time-windows for certain studies, the results would provide varying conclusions. This part of the introduction assists the author to organize and summarize his thoughts and forces him to express the abstract ideas and theoretical concepts by writing them on paper, which has helped him before. Actually, there is nothing fancy about the above-mentioned analysis approaches; however, he would like to re-iterate what he has learned in the past and apply the mathematical tools to interpret certain interesting biophysical phenomena or solving some biomedical challenges. In summary, the author has chosen to perform research work using the GD concept but with his own developed proprietary research software algorithm and app

Received: 28 October 2021, Accepted: 15 November 2021, Available online: 17 November 2021

*Corresponding author: Gerald C. Hsu, eclaireMD Foundation, USA

software along with the collected CGM sensor glucose data of 3 T2D patients over the same period of 1.5 years from 3/8/2020 through 9/8/2021. Furthermore, he has also selected a consistent glucose range covering 39 mg/dL to 249 mg/dL with a total of 211 glucose points on the x-axis of GD diagram and the GD% amplitude on the y-axis. For most clinical practices, medical doctors use HbA1C as the golden standard to evaluate the conditions of their T2D patients. The HbA1C value represents the average glucose value of all glucose over the past 90 to 120 days or perhaps 115 days based on the red blood cell's lifespan; however, the HbA1C alone cannot tell doctors additional information other than the mean value. In fact, many other useful glucose information is still hidden within. Biomarkers such as glucose variability (GV) or GF can provide more information regarding the damage to a patient's internal organ via glucose excursion which causes many diabetic complications, including both micro-vascular and macro-vascular diseases. Furthermore, the ADA has issued certain guidelines on time in range (TIR), time above range (TAR), and time below range (TBR) that can offer a general idea of how glucose is distributed in three different ranges: TIR for normal conditions, TAR for hyperglycemic conditions, and TBR for hypoglycemic conditions. However, they are still the mean values within each range, e.g., <70, 70-180, >180. Therefore, these three biomarkers, HbA1C, GF, and time in/above/below range (TxR) are still missing the needed ability to provide more detailed glucose variations. Even the GV or GF are still dealing with another kind of mean values. Based on these observed shortcomings, GD can fill in certain gaps of "missing information" from these three biomarkers, HbA1C, GF, and TxR. By using the CGM sensor glucose and the author developed app software program on the iPhone, he can generate three sets of GD data and GD curves for 3 different T2D patients individually and then combine them into one diagram with the same scales of both x-axis and y-axis. Through a closer examination of the combined diagrams of GD% for the 3 T2D patients, the author can describe his 4 key observations: (1) The waveforms of the 3 T2D patients are extremely similar to each other with their GD% peak levels and associated glucose ranges within the same proximity to each other. This observation means that all 3 patients have similar diabetes condition over the past 18 months. Actually, both the male and female patients, who are 70+ years old, have long-term and severe T2D history for over 20 years. However, through a stringent lifestyle management program, they have been able to put their diabetic conditions under control. According to the ADA 2021 consensus report, both male and female patients belong to a "remission stage" at present time (see paper No. 505 and No. 506). The young patient is a 48-year-old male with a 6-year history of T2D whose glucose conditions are not as severe as the 70+ year old diabetes patients. (2) The author has selected a rather strict definition of "normal glucose range" within 70-120 mg/dL. Using this definition, the majority of their glucose data are located within this normal glucose range: male patient has 52%, female patient has 52%, and young patient has 59% of their CGM glucose within this 70-120 mg/dL range. Furthermore, their average GD% value within this "normal glucose range" are also low, 1.7% to 1.9%. This indicates that their average GD% peak values are low as well. The finding signifies that all of them have higher percentages of glucose points which are lower glucose. (3) From examining the possibility of having hyperglycemia (>180 mg/dL) or hypoglycemia (<70 mg/dL), all three clinical cases have shown low possibilities (within the range of 1.2% to 2.3%). This finding reveals that they have lower risks of having either hyperglycemia or hypoglycemia. (4) The 3 GD waveforms are extremely similar in shape with the GD waveforms from the paper in reference 1, except for their peak GD% are within 1% to 2.5% (with only 3 T2D patients over an 18-month timeframe) while the peak GD% in reference 1 are within 2% to 6% (with many more severe T2D patients over relatively shorter timeframes).

Keywords: Glucose density; Type 2 diabetes; Glucose fluctuation; Sensor glucose; Hyperglycemia; Hypoglycemia

Abbreviations: CGM: continuous glucose monitoring; HbA1C: hemoglobin A1C; GF: glucose fluctuation; ADA: American Diabetes Association; GD: glucose density; T2D: type 2 diabetes; GV: glucose variability; TIR: time in range; TAR: time above range; TBR: time below range; TxR: time in/above/below range; MPM: math-physical medicine; FPG: fasting plasma glucose; PPG: postprandial plasma glucose

1. INTRODUCTION

The author read an article recently, "Glucodensities: a new representation of glucose profiles using distributional data analysis", dated August 19, 2020[1].

Incidentally, he has also made two further improvements on his glucose data analysis with his collected big data of sensor glucose via a continuous glucose monitoring (CGM) sensor device.

First, in addition to using the HbA1C, which is the mean value of the past 115 days of red blood cells carried glucose, of a patient is the golden standard in evaluating diabetes condition. He investigates the glucose fluctuation (GF) (glucose excursion or glycemic variability) and then transforms the GF values from a wave's time-space into an energy's frequency-space via Fourier Transform operations. Using this approach, he can then guesstimate the degree of damage on internal organs caused by the energies associated with GF. Although the GF research is one step deeper compared to the study of mean value of glucose, such as HbA1C, it is still not deep enough to provide additional details and useful information hidden within the glucose waves.

Second, he realized that the average values or mean values of glucose defined by the American Diabetes Association (ADA) such as the HbA1C or time in range (average glucose within a range) can only provide partial overviews of diabetes condition. However, these basic biomarkers are still missing some hidden internal turmoil, i.e., glucose vibrations or severe stimulations, throughout certain selected timeframes due to all types of external and/or internal stimulators. Therefore, he has defined another term known as the glucose density (GD) in order to explore more and different information hidden within the glucose data and their waveforms. GD is defined as the occurrence frequency at a specific glucose value, for example 2.1% occurrence rate at 110 mg/dL glucose value over a selected time period of collected sensor glucose. In this way, he can then calculate and examine each glucose value's occurrence rate within a glucose range that is suitable to a specific patient.

If this glucose examination method is accepted by the medical community, it would be an extremely beneficial tool for doctors to be able to quickly study the conditions of their diabetes patients. Furthermore, the author has also programmed his algorithm into an iPhone app software. Through the combination of his published papers and medical books along with a widely distributed app for patient's use in the future, he believes that worldwide type 2 diabetes (T2D) patients can benefit from his research work.

As a part of his follow-on research tasks, he plans to further examine his GD% resulted from certain food/diet nutritional types and exercise intensity levels. Hopefully, his research papers would not be limited within the scope of a "descriptive style using 26 alphabets" but instead as a "quantitative style using 10 digits". Numbers do not lie as long as we don't use fake, unorganized, and/or uncleaned data. Statistics is a tricky tool to use because it has the obvious characteristics of garbage in and garbage out (GIGO). It is also important to know that by using statistics with different selected time-windows for certain studies, the results would provide varying conclusions.

This part of the introduction assists the author to organize and summarize his thoughts and forces him to express the abstract ideas and theoretical concepts by writing them on paper, which has helped him before. Actually, there is nothing fancy about the above-mentioned analysis approaches; however, he would like to re-iterate what he has learned in the past and apply the mathematical tools to interpret certain interesting biophysical phenomena or solving some biomedical challenges.

2. METHODS

2.1 MPM background

To learn more about his developed GH-Method: Math-Physical Medicine (MPM) methodology, readers can read the following three papers selected from his ~500 published medical papers.

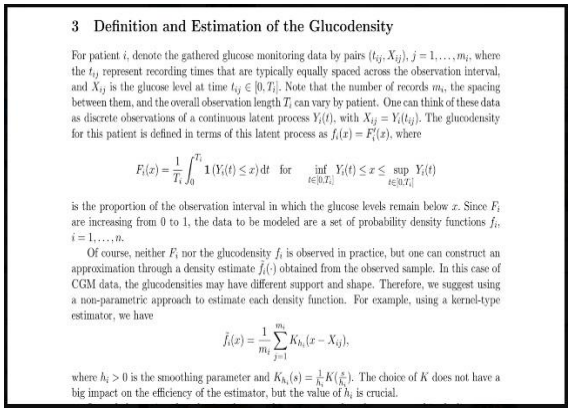
The first paper, No. 386 describes his MPM methodology in a general conceptual format.

The second paper, No. 387 outlines the history of his personalized diabetes research, various application tools, and the differences between the biochemical medicine (BCM) approach vs. the MPM approach. The third paper, No. 397 depicts a general flow diagram containing ~10 key MPM research methods and different tools.

In particular, his paper No. 453 illustrates his GH-Method: MPM in great details, "Using Topology Concept of Mathematics and Finite Element Method of Engineering to Develop a Mathematical Model of Metabolism in Medicine in Order to Control Various Chronic Diseases and their Complications via Overall Health Conditions Improvement".

2.2 Glucose density (GD%)

The author took the following photo directly from the beginning part of section 3 in the "Glucodensities" paper[1] because he is not quite familiar with how to write English articles with LATAX math symbols using Page application on iPad.



3 Definition and Estimation of the Glucodensity

For patient i, denote the gathered glucose monitoring data by pairs (t_{ij}, X_{ij}), $j = 1, \ldots, m_i$, where the t_{ij} represent recording times that are typically equally spaced across the observation interval, and X_{ij} is the glucose level at time $t_{ij} \in [0, T_i]$. Note that the number of records m_i, the spacing between them, and the overall observation length T_i can vary by patient. One can think of these data as discrete observations of a continuous latent process $Y_i(t)$, with $X_{ij} = Y_i(t_{ij})$. The glucodensity for this patient is defined in terms of this latent process as $f_i(x) = F_i'(x)$, where

$$F_i(x) = \frac{1}{T_i}\int_0^{T_i} \mathbf{1}(Y_i(t) \le x)\,\mathrm{d}t \quad \text{for} \quad \inf_{t\in[0,T_i]} Y_i(t) \le x \le \sup_{t\in[0,T_i]} Y_i(t)$$

is the proportion of the observation interval in which the glucose levels remain below x. Since F_i are increasing from 0 to 1, the data to be modeled are a set of probability density functions f_i, $i = 1, \ldots, n$.

Of course, neither F_i nor the glucodensity f_i is observed in practice, but one can construct an approximation through a density estimate $\hat{f}_i(\cdot)$ obtained from the observed sample. In this case of CGM data, the glucodensities may have different support and shape. Therefore, we suggest using a non-parametric approach to estimate each density function. For example, using a kernel-type estimator, we have

$$\hat{f}_i(x) = \frac{1}{m_i}\sum_{j=1}^{m_i} K_{h_i}(x - X_{ij}),$$

where $h_i > 0$ is the smoothing parameter and $K_{h_i}(s) = \frac{1}{h_i}K(\frac{s}{h_i})$. The choice of K does not have a big impact on the efficiency of the estimator, but the value of h_i is crucial.

For the case of any individual patient (i = 1), the author can then ignore the index i and only use j = 1,.....T, where T is the overall observation length of glucose, or in his case, the total T is 211 (from 39 mg/dL to 249 mg/dL).

His gathered CGM glucose data by pairs (tj, Xj), j = 1,…,T, where the Xj = Y(tj) = CGM glucose and the tj represents recording times (every 15 minutes for 96 times each day). Therefore, he can simplify the above equation in the photo further into a simplified equation for one patient only. The GD for himself can be defined in terms of a continuous format as follows:

$$GD(x) = \left(\int_1^T Y(t)\,dt\right) / T$$

with $x1 < Y(t) < x2$

where x1 and x2 are boundaries of his selected glucose range.

The GD for himself can also be defined in terms of a discrete format as follows:

$$GD(x) = \left(\sum_{j=1}^{T} Y(tj)\right) / T$$

with $x1 < Y(t) < x2$

where x1 and x2 are boundaries of his selected glucose range.

He uses above equations to develop his app software program on iPhone device to calculate three GD% values of these 3 T2D patients and then draw the three associated GD% curves.

3. Results

Figure 1 shows the 90-days moving average CGM sensor glucose (eAG) curves for the 3 T2D patients during a period of 1.5 years from 3/8/2020 to 9/8/2021. The related information of these T2D patients is listed below in a format of (age, years of having T2D, average eAG):

Male: 74, 26, 115 mg/dL
Female: 73, 21, 108 mg/dL
Young: 48, 6, 107 mg/dL

Top diagram of Figure 2 shows the combined GD% curves of these 3 patients. They have similar GD% "peak-lumps" located within the glucose range of 90 mg/dL to 120 mg/dL with GD% amplitudes within the range of 1.5% to 2.5%. However, most of the GD% values near the glucose data point x1 (39-65 mg/dL) and the glucose data point x2 (200-249 mg/dL) having 0% of GD% values. Therefore, the author has truncated most of the 0% GD values near point x1 and point x2, and then combined the 3 GD% curves into one diagram with identical glucose ranges from 39 mg/dL (x1) to 249 mg/dL (x2) which has a total glucose range value of 211, i.e., T = 211.

The bottom diagram in Figure 2 illustrates the 3 GD% waveforms in comparison to the GD% waveforms from reference 1 showing similar waveforms. Except, these 3 patients' peak-lumps of GD% are located within 1.5% to 2.5% over 18 months, while the peak-lumps of GD% in reference 1 are within 2% to

6% (with many more severe T2D patients over relatively shorter timeframes, e.g., 1-2 weeks).

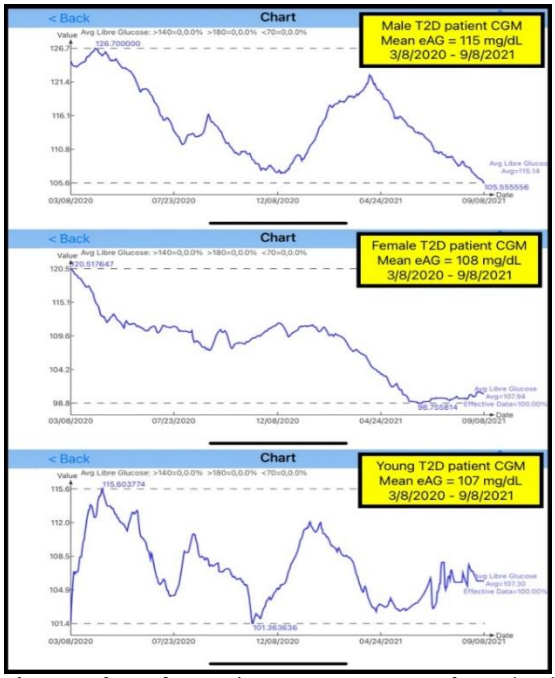


Figure 1: The 90-days moving average CGM sensor glucose (eAG) curves for the 3 T2D patients during a period of 1.5 years from 3/8/2020 to 9/8/2021.

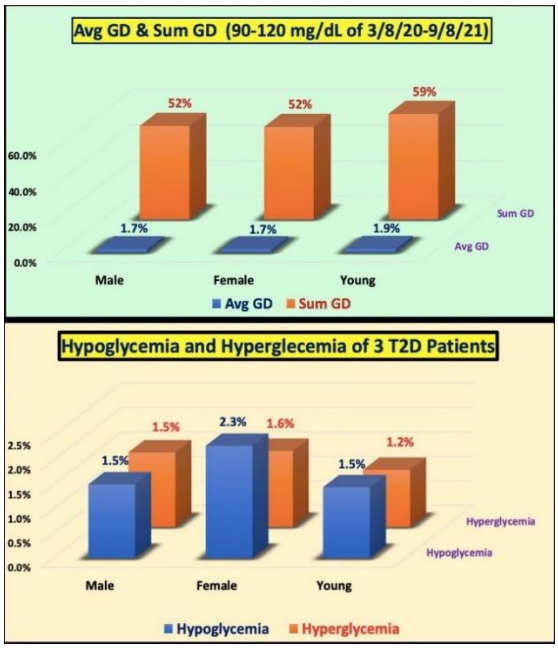


Figure 2: Glucose density (GD%) curves of FPG, PPG, and eAG (5/8/2018-9/8/2019).

Figure 3 depicts two summarized observations of GD%. The top diagram demonstrates the summation of GD% and the average GD% of 3 patients within the "normal glucose range" of 70-120 mg/dL. The following table lists in the format of summation GD%, averaged GD%:

Male: 52%, 1.7%
Female: 52%, 1.7%
Young: 59%, 1.9%

Both male and female patients have identical summation of GD% value (52%) and average of GD% value (1.7%). The young patient has higher values for summation of GD% value (59%) and average of GD% value (1.9%) which means he has a better glucose condition than the elderly male and female patients.

The bottom diagram demonstrates the possibilities of having either hyperglycemia (>180 mg/dL) or hypoglycemia (<70 mg/dL). The following table lists in the format of hyperglycemia %, hypoglycemia %:

Male: 1.5%, 1.5%
Female: 1.6%, 2.3%
Young: 1.2%, 1.5%

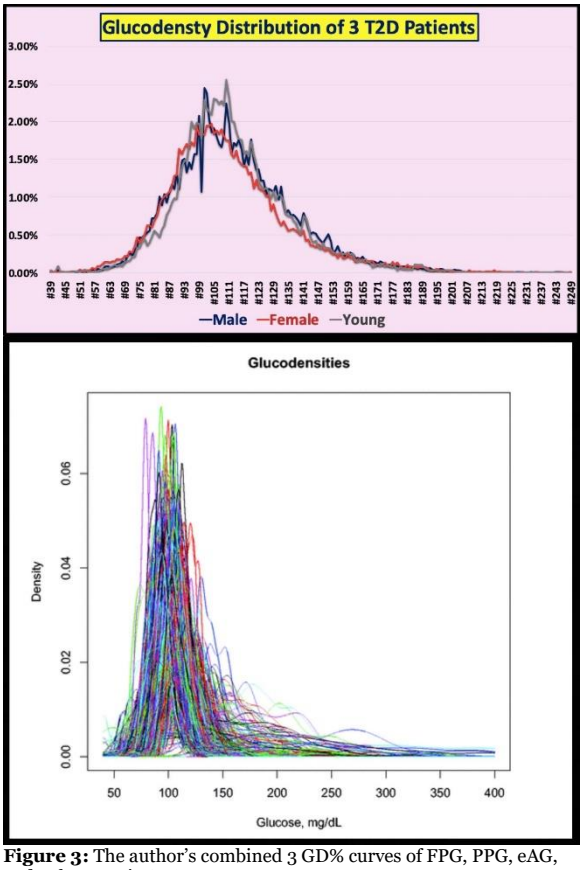


Figure 3: The author's combined 3 GD% curves of FPG, PPG, eAG, and reference 1's GD% curve.

All 3 patients have low risk in terms of having hyperglycemia or hypoglycemia. With a closer examination, we would discover that the female case is higher than the male case, while the male case is higher than young case.

4. CONCLUSION

In summary, the author has chosen to perform research work using the GD concept but with his own developed proprietary research software algorithm and app software along with the collected CGM sensor glucose data of 3 T2D patients over the same period of 1.5 years from 3/8/2020 through 9/8/2021. Furthermore, he has also selected a consistent glucose range covering 39 mg/dL to 249 mg/dL with a total of 211 glucose points on the x-axis of GD diagram and the GD% amplitude on the y-axis.

For most clinical practices, medical doctors use HbA1C as the golden standard to evaluate the conditions of their T2D patients. The HbA1C value represents the average glucose value of all glucose over the past 90 to 120 days or perhaps 115 days based on the red blood cell's lifespan; however, the HbA1C alone cannot tell doctors additional information other than the mean value. In fact, many other useful glucose information is still hidden within. Biomarkers such as glucose variability (GV) or GF can provide more information regarding the damage to a patient's internal organ via glucose excursion which causes many diabetic complications, including both micro-vascular and macro-vascular diseases. Furthermore, the ADA has issued certain guidelines on time in range (TIR), time above range (TAR), and time below range (TBR) that can offer a general idea of how glucose is distributed in three different ranges: TIR for normal conditions, TAR for hyperglycemic conditions, and TBR for hypoglycemic conditions. However, they are still the mean values within each range, e.g., <70, 70-180, >180.

Therefore, these three biomarkers, HbA1C, GF, and time in/above/below range (TxR) are still missing the needed ability to provide more detailed glucose variations. Even the GV or GF are still dealing with another kind of mean values. Based on these observed shortcomings, GD can fill in certain gaps of "missing information" from these three biomarkers, HbA1C, GF, and TxR.

By using the CGM sensor glucose and the author developed app software program on the iPhone, he can generate three sets of GD data and GD curves for 3 different T2D patients individually and then combine them into one diagram with the same scales of both x-axis and y-axis. Through a closer examination of the combined diagrams of GD% for the 3 T2D patients, the author can describe his 4 key observations:

(1) The waveforms of the 3 T2D patients are extremely similar to each other with their GD% peak levels and associated glucose ranges within the same proximity to each other. This observation means that all 3 patients have similar diabetes condition over the past 18 months. Actually, both the male and female patients, who are 70+ years old, have long-term and severe T2D history for over 20 years. However, through a stringent lifestyle management program, they have been able to put their diabetic conditions under control. According to the ADA 2021 consensus report, both male and female patients belong to a "remission stage" at present time (see paper No. 505 and No. 506). The young patient is a 48-year-old male with a 6-year history of T2D whose glucose conditions are not as severe as the 70+ year old diabetes patients.

(2) The author has selected a rather strict definition of "normal glucose range" within 70-120 mg/dL. Using this definition, the majority of their glucose data are located within this normal glucose range: male patient has 52%, female patient has 52%, and young patient has 59% of their CGM glucose within this 70-120 mg/dL range. Furthermore, their average GD% value within this "normal glucose range" are also low, 1.7% to 1.9%. This indicates that their average GD% peak values are low as well. The finding signifies that all of them have higher percentages of glucose points which are lower glucose.

(3) From examining the possibility of having hyperglycemia (>180 mg/dL) or hypoglycemia (<70 mg/dL), all three clinical cases have shown low possibilities (within the range of 1.2% to 2.3%). This finding reveals that they have lower risks of having either hyperglycemia or hypoglycemia.

(4) The 3 GD waveforms are extremely similar in shape with the GD waveforms from the paper in reference 1, except for their peak GD% are within 1% to 2.5% (with only 3 T2D patients over an 18-month timeframe) while the peak GD% in reference 1 are within 2% to

6% (with many more severe T2D patients over relatively shorter timeframes).

5. REFERENCES

For editing purposes, majority of the references in this paper, which are self-references, have been removed for this article. Only references from other authors' published sources remain. The bibliography of the author's original self-references can be viewed at www.eclairemd.com.

1) Matabuena M, Petersen A, Vidal JC, et al. Glucodensities: A new representation of glucose profiles using distributional data analysis. Stat Methods Med Res. 30(6);1445–1464:2021.

Special Issue | Issue No. 3

The GH-Method

Endocrinology and Diabetes Insights: A New Representation Using Distributional Biomarker Data Density Analysis and TBR/TIR/TAR

Gerald C. Hsu

https://www.theghmethod.com

2021