Applying Linear Regression Analysis Model to Predict HbA1C Values Using the Daily Estimated Average Glucose (eAG) Values from a 14-Month Period for a type 2 Diabetes Patient based on GH-Method: Math-Physical Medicine (No. 559)

Gerald C Hsu

EclaireMD Foundation, USA

***Corresponding author**
Gerald C Hsu, EclaireMD Foundation, USA

Submitted: 01 Dec 2021; **Accepted:** 08 Dec 2021; **Published:** 28 Dec 2021

Abstract
Since 1/1/2012, the author has been collecting various biomedical and lifestyle data related to his health conditions (~3 million data) which includes 4 categories of medical conditions, 4 chronic diseases consisting of obesity, diabetes (via finger-piercing glucose and HbA1C), hypertension, and hyperlipidemia (m1 through m4), along with 6 categories of lifestyle details, including exercise, water intake, sleep, stress, food, and daily life routines (m6 through m10). Starting on 1/1/2013, he accumulates 4 glucose data per day, one for fasting plasma glucose (FPG) and 3 for postprandial plasma glucose (PPG).

In addition, beginning on 5/8/2018, his glucoses are automatically measured using a continuous glucose monitoring (CGM) sensor device to collect 96 glucose data per day.

Over the past 12 years, he has tested for his HbA1C value each quarter since 2010. Through research work on type 2 diabetes (T2D) from 2015 to 2017, he developed several math-physical models to predict his HbA1C values prior to lab-tests for HbA1C at a clinic or hospital. He compares his predicted HbA1C values against the lab-tested HbA1C values and has achieved a 99% to 100% prediction accuracy.

In this article, he investigates his predicted HbA1C values via the daily estimated average glucose (eAG) values using a linear regression model from the past 14-month period from 10/1/2020 to 11/24/2021. This report is part of a series in his research work based on statistical regression model for the month of November 2021. Several written regression articles were initiated from his body temperature (BT) which he started to measure on 10/1/2020. In order to compare some of his research results from this period, he selected the same 14-month window.

In conclusion, the regression predicted HbA1C versus measured HbA1C, based on the daily glucose value, can be described by the following statistically derived equation:

Regression predicted HbA1C (Y)
= 0.0167*measured eAG (X) + 4.2327

This linear regression equation has the following conclusive key data:

Correlation
= 61% (not very high, but enough)
Variance
= 37% (sufficient)
Significance F & p-value
= 2% (<5%)

Furthermore, when we put the 3 datasets (measure A1C, predicted A1C, and lab-tested A1C) together, we observe that all three average HbA1C values during this 14-month period are equal to 6.1, which achieved 100% prediction accuracy.

Measured A1C vs. Predicted A1C based on Measured eAG (14 months: 10/1/2020-11/24/2021)
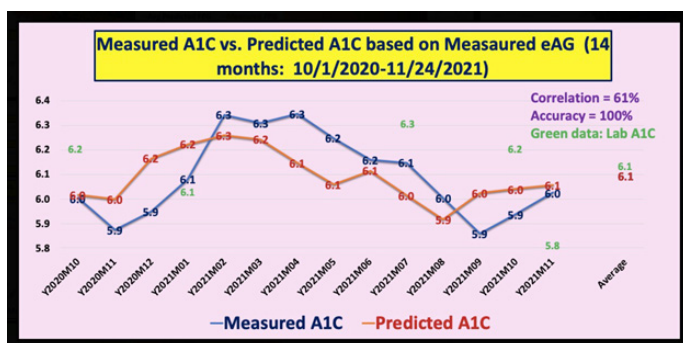
Introduction

Since 1/1/2012, the author has been collecting various biomedical and lifestyle data related to his health conditions (~3 million data) which includes 4 categories of medical conditions, 4 chronic diseases consisting of obesity, diabetes (*via finger-piercing glucose and HbA1C*), hypertension, and hyperlipidemia (m1 through m4), along with 6 categories of lifestyle details, including exercise, water intake, sleep, stress, food, and daily life routines (m6 through m10). Starting on 1/1/2013, he accumulates 4 glucose data per day, one for FPG and 3 for PPG.

In addition, beginning on 5/8/2018, his glucoses are automatically measured using a continuous glucose monitoring (CGM) sensor device to collect 96 glucose data per day.

Over the past 12 years, he has tested for his HbA1C value each quarter since 2010. Through research work on type 2 diabetes (T2D) from 2015 to 2017, he developed several math-physical models to predict his HbA1C values prior to lab-tests for HbA1C at a clinic or hospital. He compares his predicted HbA1C values against the lab-tested HbA1C values and has achieved a 99% to 100% prediction accuracy.

In this article, he investigates his predicted HbA1C values via the daily estimated average glucose (eAG) values using a linear regression model from the past 14-month period from 10/1/2020 to 11/24/2021. This report is part of a series in his research work based on statistical regression model for the month of November 2021. Several written regression articles were initiated from his body temperature (BT) which he started to measure on 10/1/2020. In order to compare some of his research results from this period, he selected the same 14-month window.

Methods
MPM Background

To learn more about the author's developed GH-Method: math-physical medicine (MPM) methodology, readers can select the following three papers 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 biochemical medicine (BCM) approach versus the MPM approach. The third paper, No. 397 depicts a general flow diagram containing ~10 key MPM research methods and different tools.

In particular, paper No. 453 illustrates his GH-Method: math-physical medicine 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".

The Author'S Case of Diabetes and Complications

The author has been a severe type 2 diabetes (T2D) patient since 1996 and weighed 220 lbs. (100 kg, BMI 32.5) at that time. By 2010, he still weighed 198 lbs. (BMI 29.2) with an average daily glucose of 250 mg/dL (HbA1C of 10%). During that year, his triglycerides reached to 1161 (diabetic retinopathy or DR) and albumin-creatinine ratio (ACR) at 116 (chronic kidney disease or CKD). He also suffered five cardiac episodes within a decade. In 2010, three independent physicians warned him regarding his needs of kidney dialysis treatment and future high risk of dying from severe diabetic complications. Other than cerebrovascular disease (stroke), he has suffered most known diabetic complications, including both macro-vascular and micro-vascular complications.

In 2010, he decided to launch his self-study on endocrinology, diabetes, and food nutrition in order to save his own life. During 2015 and 2016, he developed four prediction models related to diabetes conditions: weight, PPG, FPG, and A1C. As a result, from using his developed mathematical metabolism index (MI) model in 2014 and the four prediction tools, by end of 2016, his weight was reduced from 220 lbs. (100 kg, BMI 32.5) to 176 lbs. (89 kg, BMI 26.0), waistline from 44 inches (112 cm, non-alcoholic fatty liver disease or NAFLD) to 33 inches (84 cm), average finger glucose reading from 250 mg/dL to 120 mg/dL, and lab-tested A1C from 10% to ~6.5%. One of his major accomplishments is that he no longer takes any diabetes medications since 12/8/2015.

In 2017, he has achieved excellent results on all fronts, especially his glucose control. However, during the pre-COVID period of 2018 and 2019, he traveled to approximately 50+ international cities to attend 65+ medical conferences and made ~120 oral presentations. This hectic schedule inflicted damage to his diabetes control, through dinning out frequently, post-meal exercise disruption, jet lag, and along with the overall metabolism impact due to his irregular life patterns through a busy travel schedule; therefore, his glucose control and overall metabolism state were somewhat affected during this two-year heavy travel period.

During 2020 and 2021 with a strict COVID-19 quarantine lifestyle, not only has he written and published ~400 medical papers in 100+ journals, but he has also reached his best health conditions for the past 26 years. By the beginning of 2021, his weight was further reduced to 165 lbs. (BMI 24.4) along with a 6.1% A1C value (daily average glucose at 105 mg/dL), without having any medication interventions or insulin injections. These good results are due to his non-traveling, low-stress, and regular daily life routines. Due to the knowledge of chronic diseases, practical lifestyle management experiences, and his developed various high-tech tools, they contributed to his excellent health status since 1/19/2020, which is the start date of being self-quarantine.

On 5/5/2018, he applied a CGM sensor device on his upper arm and checks glucose measurements every 5 minutes for a total of ~288 times each day. He has maintained the same measurement pattern to present day. In his research work, he uses the CGM sensor glucose at time-interval of 15 minutes (96 data per day). Incidentally, the difference of average sensor glucoses between 5-minute intervals and 15-minute intervals is only 0.4% (average glucose of 114.81 mg/dL for 5-minutes and average glucose of 114.35 mg/dL for 15-minutes with a correlation of 93% between these two sensor glucose curves) during the period from 2/19/20 to 8/13/21.

Therefore, over the past 11 years, he could study and analyze the collected ~3 million data regarding his health status, medical conditions, and lifestyle details. He applies his knowledge, models, and tools from mathematics, physics, engineering, and computer science to conduct his medical research work. His medical research work is based on the aims of achieving both "high precision" with "quantitative proof" in the medical findings.

The following timetable provides a rough sketch of the emphasis of his medical research during each stage:

- 2000-2013: Self-study diabetes and food nutrition, developing a data collection and analysis software.
- 2014: Develop a mathematical model of metabolism, using engineering modeling and advanced mathematics.
- 2015: Weight & FPG prediction models, using neuroscience.
- 2016: PPG & HbA1C prediction models, utilizing optical physics, AI, and neuroscience.
- 2017: Complications due to macro-vascular research such as cardiovascular disease (CVD), coronary heart disease (CHD) and stroke, using pattern analysis and segmentation analysis.
- 2018: Complications due to micro-vascular research such as CKD, bladder, foot, and eye issues such as DR.
- 2019: CGM big data analysis, using wave theory, energy theory, frequency domain analysis, quantum mechanics, and AI.
- 2020: Cancer, dementia, longevity, geriatrics, DR, hypothyroidism, diabetic foot, diabetic fungal infection, linkage between metabolism and immunity, and learning about certain infectious diseases such as COVID-19.
- 2021: Applications of linear elastic glucose theory (LEGT) and perturbation theory from quantum mechanics on medical research subjects, such as chronic diseases and their complications, cancer, and dementia. Using metabolism and immunity.it's as the base, he expands his research into cancers, dementia, and COVID-19. In addition, he has also developed a few useful analysis methods and tools for his medical research work.

To date, he has collected nearly 3 million data regarding his medical conditions and lifestyle details. In addition, he has written and published 500+ articles in 100+ various medical journals, including 7 special editions with selected 20-25 papers for each edition. Moreover, he has given ~120 presentations at ~65 international medical conferences. He has continuously dedicated time and effort on medical research work to share his findings and knowledge with patients worldwide.

Results

Figure 1 displays the input data table and results of the linear regression analyses of the predicted HbA1C based on his daily eAG versus the measured HbA1C over a 14-month period from 10/1/2012 to 11/24/2021. There are 14 observations (months) with the significance F and p-value of 0.02; therefore, *the results are statistically significant.*

Figure 1: Input data table and results of regression analysis of predicted A1C using eAG as input (10/1/2020 - 11/24/2021)

The 2 key data are listed as follows:

Correlation (R) = 61%
Variance (R^2) = 37%

Figure 2 illustrates the comparison of the regression predicted A1C (orange curve) versus the measured A1C (blue curve), which contains 5 concrete data of lab-tested values. It should be emphasized that all three A1C values, measured A1C, predicted A1C, and lab-tested A1C, are identical with 6.1%. As a result, this indicates that for a 28-year T2D patient without any medication intervention is extremely healthy.
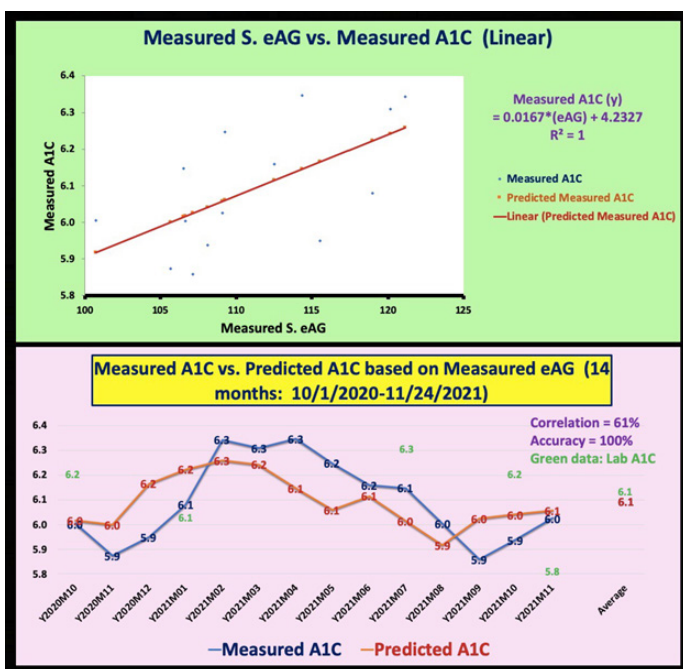
Figure 2: Results using eAG as input and comparison among regression predicted A1C, measured daily A1C, and 5 lab-tested A1C (10/1/2020 - 11/24/2021)

Conclusions

In conclusion, the regression predicted HbA1C versus measured HbA1C, based on the daily glucose value, can be described by the following statistically derived equation:

Regression predicted HbA1C (Y)
= 0.0167*measured eAG (X) + 4.2327

This linear regression equation has the following conclusive key data:

Correlation
= 61% (not very high, but enough)
Variance
= 37% (sufficient)
Significance F & p-value
= 2% (<5%)

Furthermore, when we put the 3 datasets (measure A1C, predicted A1C, and lab-tested A1C) together, we observe that all three average HbA1C values during this 14-month period are equal to 6.1, which achieved 100% prediction accuracy.

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.

Readers may use this article as long as the work is properly cited, and their use is educational and not for profit, and the author's original work is not altered.