ISSN: 2689-1204

Research Article

Journal of Applied Material Science & Engineering Research

Applying Multiple Regression Analyses Model to Predict Fasting Plasma Glucose in Early Morning based on 3 Independent Variables, Sleep Score, Weight in the Early Morning, and Daily HbA1C over a 6.5-Year Period for a type 2 Diabetes Patient Based on GH-Method: Math-Physical Medicine (No. 548)

Gerald C Hsu

EclaireMD Foundation, USA

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

Submitted: 01 Dec 2021; **Accepted:** 08 Dec 2021; **Published:** 17 Dec 2021

Citation: *Gerald C Hsu (2021) Applying Multiple Regression Analyses Model to Predict Fasting Plasma Glucose in Early Morning based on 3 Independent Variables, Sleep Score, Weight in the Early Morning, and Daily HbA1C over a 6.5-Year Period for a type 2 Diabetes Patient Based on GH-Method: Math-Physical Medicine (No. 548). J App Mat Sci & Engg Res, 5(4), 01-05.*

Abstract

In the author's previous medical research reports, he mainly applied physics theories, engineering models, mathematical equations, computer big data analytics and artificial intelligence (AI) techniques, as well as some statistical approaches to explore and interpret various biophysical phenomena. However, the majority of medical research papers he has read thus far are primarily based on statistics. As a result, in this article, he selects some basic statistical tools, such as correlation, variance, p-values, and multiple regression analyses, to study the predicted fasting plasma glucose (FPG) in early morning as the output (dependent variable) by using three independent variables, sleep score, weight in the early morning, and daily HbA1C as inputs.

Since 5/1/2015, the author has been collecting various data related to his medical conditions including body weight, blood pressure, blood lipids, and glucose along with sleep conditions and many other lifestyle details.

The original intent was to use insulin resistance as one of the three inputs for his predicted FPG level in the early morning. It is not easy for a patient to obtain specific tests for measuring insulin resistance conditions. He next considered utilizing the TyG biomarker which requires triglycerides (TG) data; however, he only has permission to conduct a quarterly blood lipid test, not on a daily basis. Therefore, the quarterly lipid test results create a data compatibility and integrity issue when they are mixed with other daily collected data. Finally, he decided to use the predicted daily HbA1C data to indicate his status of insulin resistance. His A1C value reflects not only the insulin resistance condition but also the glycemic fluctuation due to many influential factors. At least, the predicted daily A1C (not lab-tested A1C) can serve as a replacement tool for indicating his daily index of insulin resistance.

In this study, he will not repeat the detailed introduction of the regression analysis in the Method section because it is available in many statistics textbook. It should be noted that in regression analysis, the correlation coefficient R should be > 0.5 or 50% to indicate a strong inter-connectivity and the p-value should be <0.05 or 5% to be considered as statistically significant.

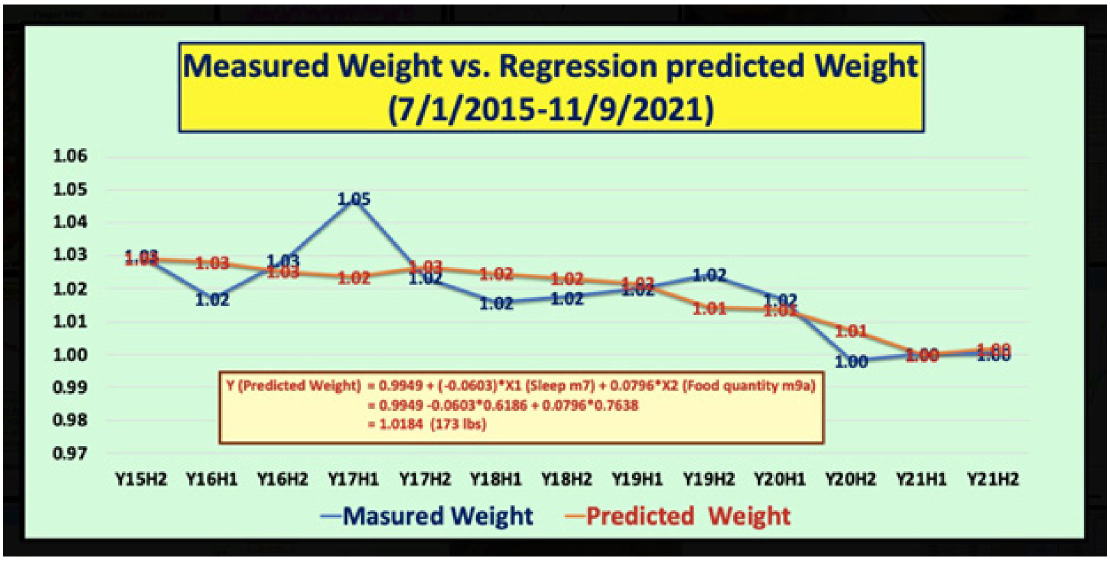
In this paper, he will combine the sleep score, body weight, and A1C (three inputs or independent variables) and then connect them with his predicted FPG in the early morning (output or dependent variable) by using the multiple regression analyses method.

In summary, there are three specific conclusions worth mentioning:

1. *The multiple regression based predicted FPG (orange curve) and his measured FPG (blue curve) have an extremely strong correlation coefficient (R=98%) and variance value (R^2=95%). In addition, it has a 100% prediction accuracy rate.* ***These findings have proven the usefulness and accuracy of the predicted dependent variable, FPG, by using multiple regression analyses model from three inputs, which are the independent variables of sleep score, weight, and HbA1C.***
2. *Furthermore, he has analyzed the connectivity between FPG vs. sleep (R= 73%, R^2=54%); FPG vs. weight (R=89%,*

R^2=80%); and FPG vs. A1C (R= 93%, R^2=87%). It is obvious that for his case and other cases, insulin resistance via A1C on FPG is the highest contributor, weight is in the middle, and sleep is the lowest. This conclusion is based on the correlation and p-value of each of the three inputs associated with the output, FPG. This seems logical since the ***insulin resistance condition affects both FPG and postprandial plasma glucose (PPG). Incidentally, he has already identified the strong connection between FPG and weight (see his previously published papers). The contribution of sleep on FPG is the most questionable input due to the lower correlation and higher p-value.***

3. *By using multiple regression models based predicted FPG equation, he can reproduce an average FPG prediction of 110.7 mg/dL which is 100% identical to his measured finger-piercing FPG value.*



Introduction

In the author's previous medical research reports, he mainly applied physics theories, engineering models, mathematical equations, computer big data analytics and artificial intelligence (AI) techniques, as well as some statistical approaches to explore and interpret various biophysical phenomena. However, the majority of medical research papers he has read thus far are primarily based on statistics. As a result, in this article, he selects some basic statistical tools, such as correlation, variance, p-values, and multiple regression analyses, to study the predicted FPG in early morning as the output (dependent variable) by using three independent variables, sleep score, weight in the early morning, and daily HbA1C as inputs.

Since 5/1/2015, the author has been collecting various data related to his medical conditions including body weight, blood pressure, blood lipids, and glucose along with sleep conditions and many other lifestyle details.

The original intent was to use insulin resistance as one of the three inputs for his predicted fasting plasma glucose (FPG) level in the early morning. It is not easy for a patient to obtain specific tests for measuring insulin resistance conditions. He next considered utilizing the TyG biomarker which requires triglycerides (TG) data; however, he only has permission to conduct a quarterly blood lipid test, not on a daily basis. Therefore, the quarterly lipid test results create a data compatibility and integrity issue when they are mixed with other daily collected data. Finally, he decided to use the predicted daily HbA1C data to indicate his status of insulin resistance. His A1C value reflects not only the insulin resistance condition but also the glycemic fluctuation due to many influential factors. At least, the predicted daily A1C (not lab-tested A1C) can serve as a replacement tool for indicating his daily index of insulin resistance.

In this study, he will not repeat the detailed introduction of the regression analysis in the Method section because it is available in many statistics textbook. It should be noted that in regression analysis, the correlation coefficient R should be > 0.5 or 50% to indicate a strong inter-connectivity and the p-value should be <0.05 or 5% to be considered as statistically significant.

In this paper, he will combine the sleep score, body weight, and A1C (three inputs or independent variables) and then connect them with his predicted FPG in the early morning (output or dependent variable) by using the multiple regression analyses method.

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, fasting plasma glucose (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, nonalcoholic fatty liver disease /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 continuous glucose monitoring (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 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 536 medical papers 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 a summarized data table of the multiple regression analyses of his semi-annual FPG versus semi-annual sleep, weight, and HbA1C over the past 6.5-year period from 7/1/2015 to 11/9/2021. There are 13 observations (half-years) with the significance F value of 0.000003; therefore, the results are statistically significant.

The 5 key data are listed below:

Correlation (R) = 98%
Variance (R^2) = 95%
Measured FPG = 110.7 mg/dL
Predicted FPG = 110.7 mg/dL
Prediction accuracy = 100%

11/11/21	Y	X1	X2	X3	FPG vs. Sleep/Weight/A1C		
Period	FPG	Sleep	Weight	Daily A1c	Period	Measured FPG	Predicted FPG
Y15H2	119	0.64	172	7.2	Y15H2	119	120
Y16H1	115	0.64	172	7.1	Y16H1	115	118
Y16H2	119	0.64	174	6.9	Y16H2	119	119
Y17H1	124	0.62	176	7.0	Y17H1	124	126
Y17H2	116	0.62	172	6.8	Y17H2	116	113
Y18H1	117	0.62	171	6.9	Y18H1	117	113
Y18H2	110	0.63	171	6.8	Y18H2	110	111
Y19H1	113	0.62	172	6.8	Y19H1	113	114
Y19H2	116	0.63	173	6.7	Y19H2	116	112
Y20H1	110	0.60	172	6.6	Y20H1	110	109
Y20H2	92	0.61	168	6.1	Y20H2	92	93
Y21H1	97	0.58	169	6.2	Y21H1	97	96
Y21H2	91	0.60	168	6.1	Y21H2	91	94
Average	111	0.6186	171.6	6.71	*Averaged FPG*	*110.7*	*110.7*
Correlation	100%	73%	89%	93%	Correlation	98%	
Variance	100%	54%	80%	87%	Variance	95%	
Slope		-39.7433	1.9744	18.9828	Pediction Accuracy	100%	
Y-Intercept	-330.9539						

SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.976227497					
R Square	0.953020126					
Adjusted R Square	0.937360168					
Standard Error	2.666714627					
Observations	13					
ANOVA	*df*	*SS*	*MS*	*F*	*Significance F*	
Regression	3	1298.332175	432.7773916	60.85713164	0.000003	
Residual	9	64.00230211	7.111366901			
Total	12	1362.334477				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-330.9538634	80.93725698	-4.08901754	0.003	-514.046659	-147.8610678
Sleep	-39.74334844	73.38688491	-0.541559278	0.601	-205.7560158	126.2693189
Weight	1.97441768	0.519778795	3.798572967	0.004	0.798596356	3.150239005
Daily A1c	18.98284378	4.574526371	4.149685069	0.002	8.634546185	29.33114138
Rgression Y: (Predicted FPG)	*110.7*					

Figure 1: Data table of multiple regression analysis results of FPG resulting from 3 inputs, sleep, weight, and HbA1C (7/1/2015-11/9/2021)

Figure 2 illustrates the comparison of his multiple regression analyses predicted FPG (orange curve) versus measured FPG (blue curve). It is very clear that his regression predicted FPG values are highly correlated with the measured FPG over the 13 semi-annual periods.

As a reference, here is the multiple regression generated semi-annual weight prediction equation:

Y (regression predicted FPG)
= -330.95 + (-39.74)* X1 (sleep) + (1.97)*X2 (weight) + (18.98)*X3 (A1C)

Using averaged input values of
sleep = 0.6186
weight = 171.6
A1C = 6.71

We will obtain a predicted average FPG of 110.7 mg/dL which is identical with the measured average FPG value.

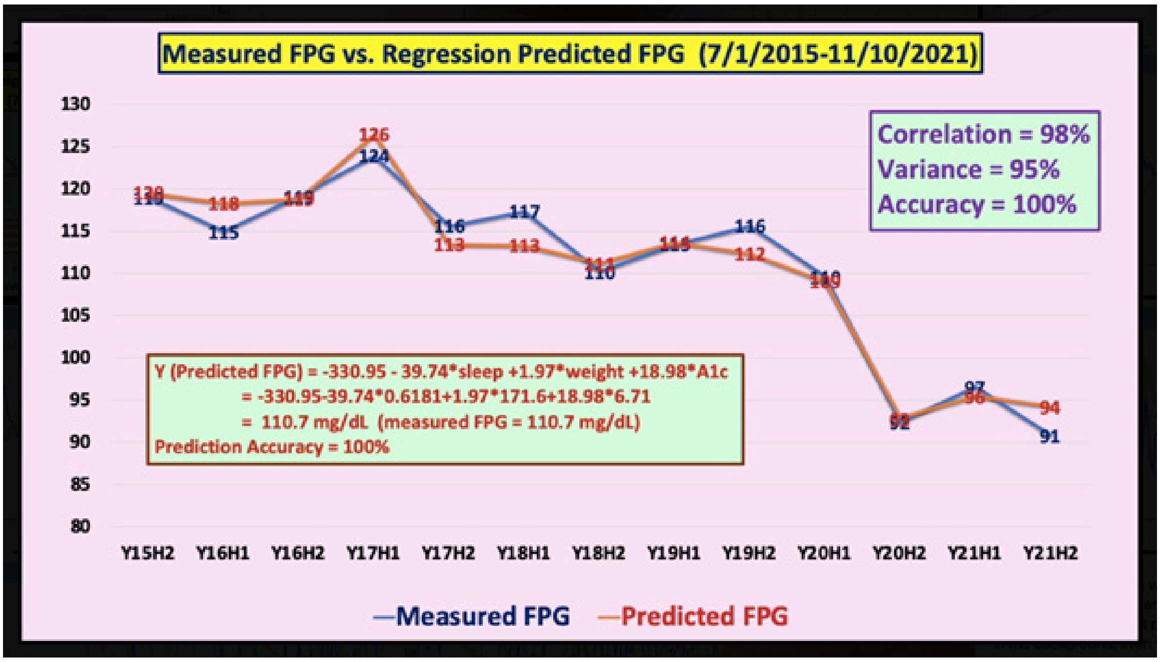


Figure 2: Multiple Regression based Predicted FPG versus measured FPG (7/1/2015 - 11/9/2021)

Conclusions

In summary, there are three specific conclusions worth mentioning:

1. The multiple regression based predicted FPG (orange curve) and his measured FPG (blue curve) have an extremely strong correlation coefficient (R=98%) and variance value (R^2=95%). In addition, it has a 100% prediction accuracy rate. ***These findings have proven the usefulness and accuracy of the predicted dependent variable, FPG, by using multiple regression analyses model from three inputs, which are the independent variables of sleep score, weight, and HbA1C.***
2. Furthermore, he has analyzed the connectivity between FPG vs. sleep (R= 73%, R^2=54%); FPG vs. weight (R=89%, R^2=80%); and FPG vs. A1C (R= 93%, R^2=87%). It is obvious that for his case and other cases, insulin resistance via A1C on FPG is the highest contributor, weight is in the middle, and sleep is the lowest. This conclusion is based on the correlation and p-value of each of the three inputs associated with the output, FPG. This seems logical since the ***insulin resistance condition affects both FPG and PPG. Incidentally, he has already identified the strong connection between FPG and weight (see his previously published papers). The contribution of sleep on FPG is the most questionable input due to the lower correlation and higher p-value.***
3. By using multiple regression models based predicted FPG equation, he can reproduce an average FPG prediction of 110.7 mg/dL which is 100% identical to his measured finger-piercing FPG value.

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.