ISSN: 2689-1204

Research Article

Journal of Applied Material Science & Engineering Research

Applying Multiple Regression Analyses to Compare the Regression Predicted and Measured Body Weight Using Food Quantity and Sleep Score as Inputs over a 6.5-Year Period for a type 2 Diabetes Patient Based on GH-Method: Math-Physical Medicine (No. 547)

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 to Compare the Regression Predicted and Originally Calculated CVD/Stroke Risk Probabilities Using Medical Condition and Lifestyle Detail Scores as Inputs over a 10-year Period for a type 2 Diabetes Patient Based on GH-Method: Math-Physical Medicine (No. 546). 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 body weight as the output (dependent variable) by using his foods quantity and sleep score as inputs (independent variables).

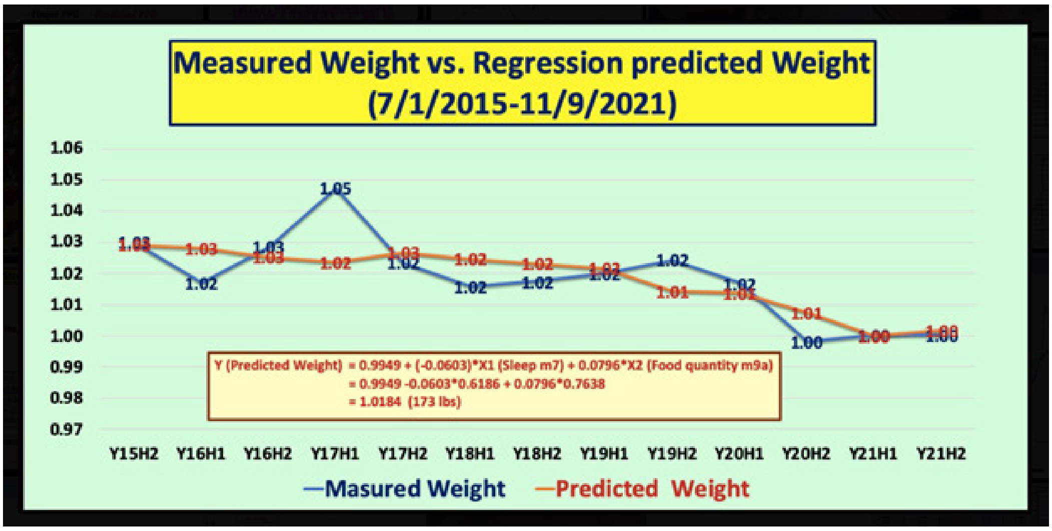
Since 5/1/2015, the author has been collecting various data related to his food nutrition (~0.5 million data) and sleep conditions. The Food Details (FD) category includes both food quantity (m9a) and food quality (nutrition, m9b).

Previously, he has researched and published a few articles regarding the relationship between body weight and quantity of food consumption. During 11/8/2020 - 11/10/2021, he experimented with intermittent fasting experiments to study their inter-connectivity. In this paper, he will combine sleep score (m7) with food quantity (m9a) and connected them with his body weight in early morning (m1) using the multiple regression analysis method.

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 to be considered as statistically significant.

In summary, there are three specific conclusions worth mentioning:

1. *The multiple regression based predicted weight (orange curve) and his measured weight (blue curve) have a strong correlation coefficient (R=73%) and good variance value (R^2=54%). In addition, it has an 100% prediction accuracy rate.* ***These findings have proven the usefulness and accuracy of the predicted dependent variable, body weight, by using multiple regression analysis results from two independent input variables, food quantity and sleep score.***
2. *Furthermore, he has analyzed the connectivity between weight vs. food quantity (R= 73%, R^2=53%) and weight vs. sleep (R= 61%, R^2=37%). It is obvious that for his case and other cases, the contribution of food quantity on body weight is higher than the contribution of sleep on body weight. This conclusion is based on higher correlation and lower p-value associated with food quantity. This is logical to him because* ***food and meals contribute directly to body weight; however, while sleep does influence body weight, it must go through an indirect and more complicated biochemical process.***
3. *Using multiple regression models based on a weight prediction equation, he can reproduce an average weight prediction of 173 lbs. (m1=1.0184) over the same 6.5-years which is the same value as his previous measured average weight of 173 lbs. (m1=1.1084).*



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 body weight as the output (dependent variable) by using his foods quantity and sleep score as inputs (independent variables).

Since 5/1/2015, the author has been collecting various data related to his food nutrition (~0.5 million data) and sleep conditions. The Food Details (FD) category includes both food quantity (m9a) and food quality (nutrition, m9b).

Previously, he has researched and published a few articles regarding the relationship between body weight and quantity of food consumption. During 11/8/2020 - 11/10/2021, he experimented with intermittent fasting experiments to study their inter-connectivity. *In this paper, he will combine sleep score (m7) with food quantity (m9a) and connected them with his body weight in early morning (m1)* ***using the multiple regression analysis method.***

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 to be considered as statistically significant.

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 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 analysis of his body weight versus bi-annual food consumption quantity and bi-annual sleep scores 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 0.021; therefore, the results are still statistically significant.

The 5 key data are listed below:

Correlation (R) = 73%
Variance (R^2) = 54%
Measured average weight = 173 #
Regression Predicted weight = 173 #
Prediction accuracy = 100%

11/10/21	Y	X1	X2			
m1 vs. m7 & m9a	Weight m1	Sleep m7	Food Quantity m9a	m1 vs. m7 & m9a	Masured Weight	Predicted Weight
Y15H2	1.0299	0.6399	0.9141	Y15H2	1.03	1.03
Y16H1	1.0171	0.6427	0.9025	Y16H1	1.02	1.03
Y16H2	1.0284	0.6385	0.8638	Y16H2	1.03	1.03
Y17H1	1.0474	0.6212	0.8294	Y17H1	1.05	1.02
Y17H2	1.0233	0.6182	0.8672	Y17H2	1.02	1.03
Y18H1	1.0158	0.6237	0.8453	Y18H1	1.02	1.02
Y18H2	1.0175	0.6265	0.8284	Y18H2	1.02	1.02
Y19H1	1.0201	0.6156	0.8024	Y19H1	1.02	1.02
Y19H2	1.0242	0.6338	0.7241	Y19H2	1.02	1.01
Y20H1	1.0169	0.5969	0.6886	Y20H1	1.02	1.01
Y20H2	0.9982	0.6081	0.6194	Y20H2	1.00	1.01
Y21H1	1.0004	0.5811	0.5052	Y21H1	1.00	1.00
Y21H2	1.0003	0.5961	0.5390	Y21H2	1.00	1.00
Average	1.0184	0.6186	0.7638	Average	1.0184	1.0184
Correlation	100%	61%	73%	Correlation	73%	
Variance	100%	37%	53%	Variance	54%	
Slope		-0.0603	0.0796	Pediction Accuracy	100%	
Y-Intercept	0.9949			Average Weight (lbs)	173	173

SUMMARY: Regression Statistics								
Multiple R	0.733							
R Square	54%							
Adjusted R Square	44%							
Standard Error	0.010							
Observations	13.000							
ANOVA	df	SS	MS	F	Significance F			
Regression	2.000	0.001	0.001	5.790	0.021			
Residual	10.000	0.001	0.000					
Total	12.000	0.002						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.99487	0.164	6.066	0.000	0.629	1.360	0.629	1.360
Sleep m7	-0.06026	0.309	-0.195	0.849	-0.749	0.629	-0.749	0.629
Food Quantity m9a	0.07964	0.043	1.870	0.091	-0.015	0.175	-0.015	0.175
RESIDUAL OUTPUT								
Observation	Predicted Weight m1	Residuals						
1	1.029	0.001						
2	1.028	-0.011						
3	1.025	0.003						
4	1.023	0.024						
5	1.027	-0.003						
6	1.025	-0.009						
7	1.023	-0.006						
8	1.022	-0.002						
9	1.014	0.010						
10	1.014	0.003						
11	1.008	-0.009						
12	1.000	0.000						
13	1.002	-0.002						

Figure 1: Data table of multiple regression analysis results of body weight (m1) resulting from food quantity (m9a) and sleep score (m7) during 7/1/2015-11/9/2021

Figure 2 shows two diagrams of multiple regressions based predicted weight versus food quantity and sleep score as its two inputs (7/1/2015 - 11/9/2021). Weight is more dependent on food quantity (higher R & R^2 with low p-value) than sleep (lower correlation & variance with high p-value).

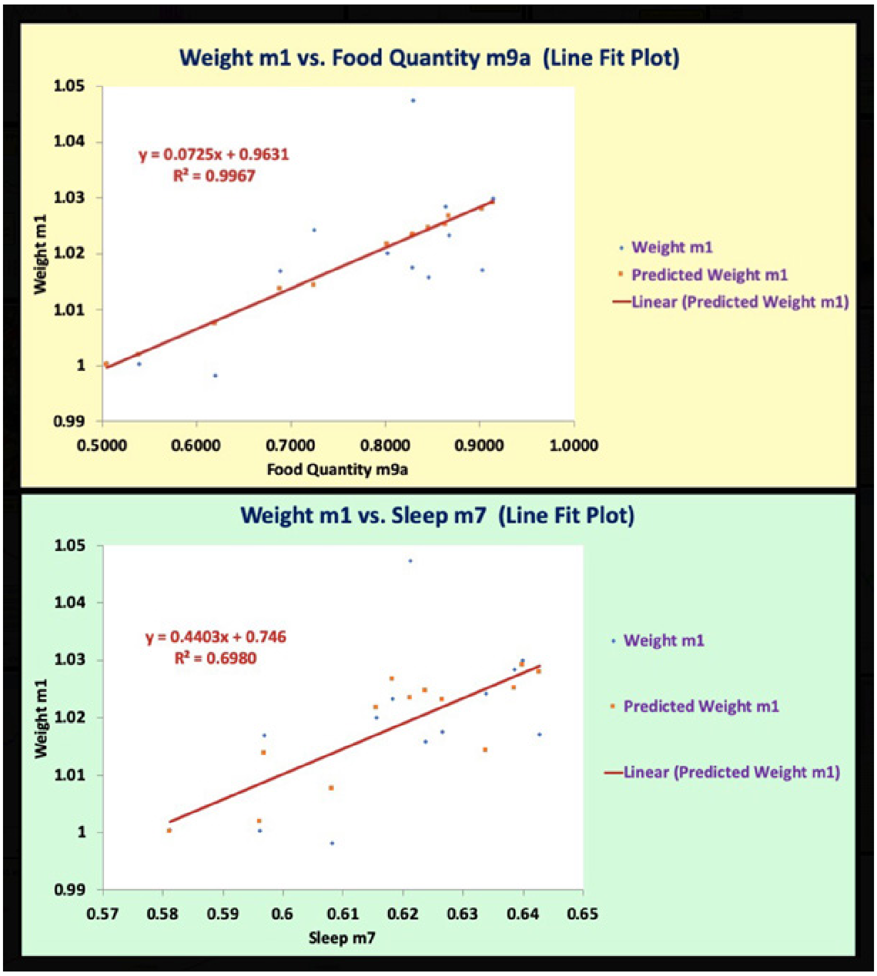


Figure 2: Multiple Regression based Predicted weight using food quantity and sleep score as its two inputs during 7/1/2015 - 11/9/2021

Figure 3 illustrates the comparison of his multiple regression analysis predicted weight (orange curve) versus measured weight (blue curve). It is very clear that his regression predicted weight values are highly correlated but with a narrower band of variation than the measured weight during the 13 half-annual periods.

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

Y (regression predicted weight)
= 0.9949 + (-0.0603)* X1 (food quantity m9a) + 0.0796*X2 (sleep score)

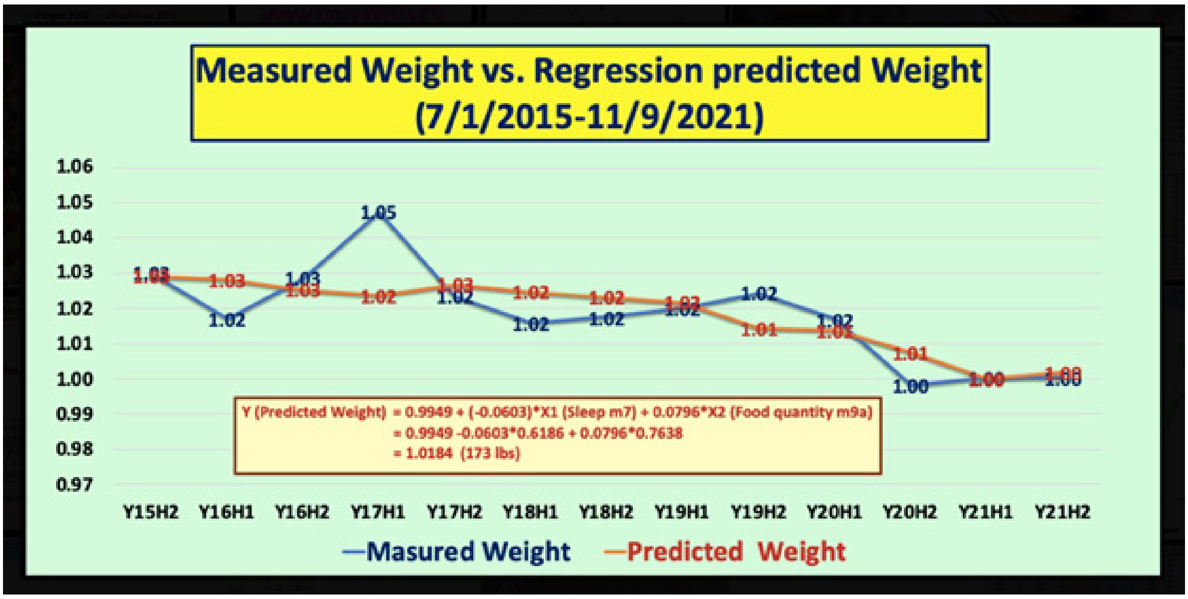


Figure 3: Multiple Regression based Predicted weight versus measured weight during 7/1/2015 - 11/9/2021

Conclusions

In summary, there are three specific conclusions worth mentioning:

1. The multiple regression based predicted weight (orange curve) and his measured weight (blue curve) have a strong correlation coefficient (R=73%) and good variance value (R^2=54%). In addition, it has an 100% prediction accuracy rate. ***These findings have proven the usefulness and accuracy of the predicted dependent variable, body weight, by using multiple regression analysis results from two independent input variables, food quantity and sleep score.***
2. Furthermore, he has analyzed the connectivity between weight vs. food quantity (R= 73%, R^2=53%) and weight vs. sleep (R= 61%, R^2=37%). It is obvious that for his case and other cases, the contribution of food quantity on body weight is higher than the contribution of sleep on body weight. This conclusion is based on higher correlation and lower p-value associated with food quantity. This is logical to him because ***food and meals contribute directly to body weight; however, while sleep does influence body weight, it must go through an indirect and more complicated biochemical process.***
3. Using multiple regression models based on a weight prediction equation, he can reproduce an average weight prediction of 173 lbs. (m1=1.0184) over the same 6.5-years which is the same value as his previous measured average weight of 173 lbs. (m1=1.1084).

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.