

A predictive tool based on DNA methylation data for personalized weight loss through different dietary strategies [†]

Nereyda Carolina García-Álvarez¹, J. Ignacio Riezu-Boj², J. Alfredo Martínez³, Sonia García-Calzón⁴ and Fermín I. Milagro^{5,*}

¹ Center for Nutrition Research, Department of Nutrition, Food Sciences and Physiology, Faculty of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain. ngarciaalva@alumni.unav.es

² Center for Nutrition Research, Department of Nutrition, Food Sciences and Physiology, Faculty of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain; jiriezu@unav.es

³ Center for Nutrition Research, Department of Nutrition, Food Sciences and Physiology, Faculty of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain; jalfmtz@unav.es

⁴ Center for Nutrition Research, Department of Nutrition, Food Sciences and Physiology, Faculty of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain; sgcalzon@unav.es

⁵ Center for Nutrition Research, Department of Nutrition, Food Sciences and Physiology, Faculty of Pharmacy and Nutrition, University of Navarra, Pamplona, Spain; fmilagro@unav.es

* Correspondence: Tel.: +34 948425600.

Abstract: Background and aims: Obesity is a public health problem. The usual treatment is a reduction in calorie intake and an increase in energy expenditure, but not all individuals respond equally to these treatments. Epigenetics could be a factor that contributes to this heterogeneity. The aim of this research was to determine the association between DNA methylation at baseline and the percentage of BMI loss (%BMIL) after two dietary interventions, in order to design a prediction model to evaluate %BMIL based on methylation data. Methods and Results: Spanish participants with overweight or obesity (n=306) were randomly assigned to two lifestyle interventions with hypocaloric diets: one moderately high in protein (MHP) and the other low in fat (LF) during 4 months (Obekit study). DNA methylation was analyzed in white blood cells using the Infinium MethylationEPIC array. After identifying those methylation sites associated with %BMIL, ($p < 0.05$ and $SD > 0.1$), two weighted methylation sub scores were constructed for each diet: 15 CpGs were used for MHP diet and 11 CpGs for LF diet. Afterwards, a total methylation score was made by subtracting the previous sub scores. These data were used to design a prediction model for %BMIL through a linear mixed effect model in which the interaction between diet and total score. Conclusion: Overall, DNA methylation predicts %BMIL of two hypocaloric diets after 4 months and was able to determine which type of diet is the most appropriate for each individual. These results confirm that epigenetic biomarkers may be further used for precision nutrition and the design of personalized dietary strategies against obesity.

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Biol. Life Sci. Forum* **2022**, *2*, x. <https://doi.org/10.3390/xxxxx>

Academic Editor: Firstname Lastname

Published: date

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Obesity; Epigenetic score; Precision nutrition; Hypocaloric diet; Predictive model

1. Introduction

DNA methylation is considered a key part of the pathogenesis and clinical manifestations of obesity [1]. It has previously been shown that methylation can influence the development of obesity but also the response to dietary weight loss treatments [2],[3]. This is due to the different complex metabolic mechanisms in which methylation may be involved[4]. It has also been associated that methylation in genes related to hunger and satiety mechanisms contribute to the development of obesity [5]. This is why the use of precision nutrition is increasingly being considered[6], Under this perspective, the aim of this research was to determine the association between basal DNA methylation and the

percentage of BMI loss after a dietary intervention, in order to design a model that, based on a methylation score, predicts the percentage loss of BMI with two different types of diets.

2. Materials and Methods

2.1. Study population

The population selected in this research was from the Obekit study in which 314 Spanish individuals with overweight and obesity were initially recruited. The study lasted from October 2015 to February 2017 and was carried out in the metabolic unit of the nutrition research center of the University of Navarra. Major exclusion criteria were type 1 diabetes mellitus, pregnant or breastfeeding women, cardiovascular disease, cancer, eating and cognitive disorders. Of the 314 subjects initially recruited, 8 did not meet the inclusion criteria. The intervention began with 306 participants, who were randomly assigned to two types of hypocaloric diet, 146 participants with a moderately high-protein diet (MHP) and 160 participants with a low-fat diet (LF). The intervention lasted 4 months. All research procedures were carried out following the ethical principles of the Declaration of Helsinki of 2013[7]. The study protocol was approved by the Research Ethics Committee of the University of Navarra (ref.132/2015). All participants gave their written informed consent before their inclusion in the study.

2.2. Nutritional intervention

The diets used for the study had a 30% calorie restriction. The individual energy requirements of each participant were estimated at the beginning, calculating their energy expenditure at rest and during physical activity, to prescribe hypocaloric diets in a random manner. The macronutrient distribution for the moderately high protein (MHP) diet was 40% carbohydrate, 30% protein, 30% fat and for the low fat (LF) diet it was 60% carbohydrate, 18% protein, and 22% fat. Both the LF and MHP diets were designed on the basis of a food exchange system. Participants received detailed information from trained dietitians on portion sizes, dietary patterns/eating schedules, and food preparation techniques.

2.3. DNA isolation and bisulfite conversion

Blood samples taken at the beginning of the study were centrifuged at 4 °C for 15 minutes to obtain plasma and isolate the buffy coat fraction. DNA extraction was performed with the "MasterPure" DNA purification kit" for blood version II (Epicentre Biotechnologies, WI, USA), and it was quantified with a spectrophotometer (Nanodrop, Thermo Scientific, DE, USA) and stored at -80 °C. In a second step, 500 ng of DNA was treated with sodium bisulfite using the EZ-96 DNA methylation kit (Zymo Research Corporation, CA, USA), to convert unmethylated cytosine residues to uracil, while methylated cytosines remain unchanged.

2.4. Array analysis

The levels of methylated DNA were evaluated using the "Infinium MethylationEPIC BeadChip" kit (Illumina, CA, USA), which includes 850,000 methylation sites. Samples were scanned with an "Illumina HiScanSQ" system and image intensities were extracted with "GenomeStudio v1.9" methylation software (Illumina, CA, USA). The within-array quantile subset normalization (SWAN) method was used to improve the results obtained from the platform, reducing technical variation within and between arrays. The ComBat method was used to adjust for batch effects and remove technical variation. In addition, DNA methylation was corrected for cellular composition (granulocytes, monocytes, B cells, CD8+ cytotoxic cells, CD4+ T helper cells, and natural killer cells) using Houseman's algorithm [8].

2.5. Design of the BMI percentage loss prediction model based on MHP and LF diet methylation data

- Selection of CpG sites for prediction model obtained in the “Illumina” methylation array.
- Weighted sub score and total CpG site methylation score for prediction model
- Design of a linear mixed effect model for the prediction of percentage loss of BMI in which the interaction between diet and total Score is analyzed.
- Information on the CpG sites obtained from the “Illumina” methylation array of each diet to characterize the CpG sites.

2.6. Statistical analysis for the prediction model

Standard deviation of the CpG sites obtained in the Illumina array was performed, then a Spearman measurement analysis was performed with the CpG sites correlating their methylation with the percentage of BMI loss and with these CpG sites the “furnival-Wilson” algorithm was applied. “by leaps and bounds” (vselect in stata)[9], to obtain the best combinations in the multiple linear regression of the percentage of BMI loss with the methylation sites for the MHP diet and the LF diet. With the values of the beta coefficients obtained in the linear regressions, the weighted methylation sub scores for each diet and the total methylation score were constructed. This total methylation score was used for the design of the BMI percent loss prediction model, which was performed using a linear mixed-effects model. The prediction model was drawn applying marginals. A Z test was used to evaluate the distribution of subjects in the prediction model. Statistical analyzes were performed using Stata MP 14 (StataCorp LLC, College Station, TX, USA; <http://www.stata.com>).

3. Results

3.1. Selection of CpG sites for the prediction model

Of the 850,000 CpG sites obtained in the illumina, those that had a standard deviation >0.1 were selected, obtaining 1,233 CpG sites. With these CpG sites, a Spearman correlation was performed between the methylation of the 1,233 selected CpG sites and the percentage of BMI loss for each of the diets after 4 months of intervention, selecting those CpG sites that presented a significant correlation with $P < 0.05$. The result was 34 CpG sites for the MHP diet and 20 CpG sites for the LF diet. To better predict the percentage of BMI loss with each of the diets, the “furnival-Wilson leaps and bounds” algorithm (“vselect” in stata)[9] was used. The algorithm recognized 19 CpG sites for the MHP diet and 14 CpG sites for the LF diet. Then, the CpG sites that presented a $P < 0.19$ in the last regression model were selected and those that presented multicollinearity were eliminated, obtaining 15 CpG sites for the MHP diet and 11 CpG sites for the LF diet, which were included in the next analytical step, which is the calculation of the subscores.

3.2. Design of weighted sub-scores that contain the CpG sites of each diet and the calculation of the total methylation score for the prediction model

Weighted sub-scores were made for each diet, using the sum of the previously selected CpG sites and multiplying them by the beta coefficients obtained in each of the multiple linear regressions of the MHP diet and the LF diet (**Table 1**). To obtain a total score for each individual that would allow to be included as a term for the interaction with the diet variable, the MHP diet subscore was subtracted from the LF diet subscore as shown in **Table 1**.

Table 1. Design of weighted sub-scores for diet MHP, LF and Total Score.

Score	Calculating formula
Sub score MHP Diet	$(cg22355889*9.82)+(cg24433124*-5.81)+(cg14317533*-7.58)+(cg07104639*6.25)+(cg19148731*-7.88)+(cg01097406*5.98)+(cg15263617*8.97)+(cg23377942*-4.19)+(cg25340688*-5.79)+(cg19053046*5.08)+(cg11460778*4.22)+(cg03447554*-4.08)+(cg09768983*3.98)+(cg20684491*4.80)+(cg20315590*-6.71)$
Sub score LF Diet	$(cg15572235*7.57)+(cg19424457*-8.04)+(cg18493115*-6.26)+(cg00481382*5.78)+(cg16600909*5.38)+(cg15837280*5.83)+(cg03188948*-4.30)+(cg16078649*-4.16)+(cg15011943*4.64)+(cg07167872*3.98)+(cg14222729*-3.40)$
Total Methylation Score	Sub score MHP – Sub Score LF

A linear mixed effect model was used to predict, based on the total methylation score of each individual, which diet will be the best for them based on the greatest percentage loss of BMI. Therefore, a linear mixed effect model was designed with the percentage of BMI loss as the dependent variable and total score, diet (MHP/LF) and the interaction term between the total score and the diet as a fixed effect. Moreover the id of the participants were included as a random effect. The model was adjusted for sex and age. **Table 2** shows the independent variables, the beta coefficient with the standard error and the P value. As it can be seen in the table, the model is not affected when adjusting for sex and age, since none of these variables showed statistical significance.

Table 2. Linear mixed effect model for the prediction of the percentage loss of BMI.

Prob > chi2 = < 0.001	Percentage of BMI loss	
Independent variable	Beta Coefficient ± SEM	P Values Z test
Age	-0.1 ± <0.001	0.943
Sex	-0.1±0.1	0.954
Diet (MHP or LF)	-1.21±0.1	<0.001
Total methylation score	0.2±.01	<0.001
Diet## total methylation score	-0.2±.02	<0.001
Cons	-7.9±0.6	<0.001

Dependent variable percentage of BMI loss, fixed effect variable diet, total score, interaction of diet and total score. Adjusted for age, sex. Diet##total score indicates the interaction between both variables. SEM: standard error.

3.3 Representation of the prediction model

Based on the linear mixed effect model, a "diet x total score" interaction graph was made showing the marginal percentage loss of BMI for each diet. As shown in **Figure 1**, where the "X" axis represents the total methylation score and the "Y" axis shows the percentage of BMI loss, the percentage of BMI loss with each diet is estimated (MHP in blue and LF in red) according to the methylation score presented by the subjects before starting the intervention. Thus, for methylation score values between -10 and -1, the error bars that predict the percentage of BMI loss for each diet overlap, indicating that both diets are effective in losing similar BMI values. It should be noted that between -13 and -10 the overlap is less. Conversely, when an individual has a Score that is between 2 and 11, the error bars separate, suggesting that the methylation score is effective in predicting the type of diet that is the best for weight loss in that individual.

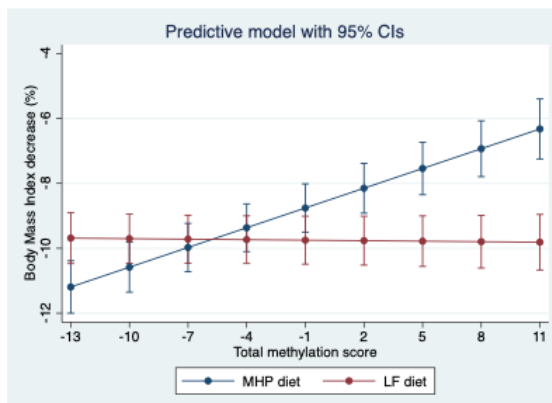


Figure 1. Prediction model of percentage of BMI loss according to the total methylation score of each individual. "Y" axis Percentage loss of BMI, "X" axis Methylation score. According to the total score, it can be predicted with which diet the greatest percentage of BMI would be achieved for each individual.

The statistical differences between the predictions of the percentage loss of BMI with both diets for each individual were analyzed using the "Z" test that involves the standard errors of each estimation of the Score. A significant P (<0.05) allows prescribing the most appropriate diet for each individual. However, if the P is not significant (P>0.05), the two diets have a similar effect, so they can be prescribed interchangeably.

With this test it was observed that, of the 201 participants, it cannot be predicted which diet is better in 126 participants, but 75 participants can be advised one diet as better than the other based on BMI loss. As shown in Figure 2, in which the "X" axis represents the total score and the "Y" axis shows the P value obtained in the Z test, the non-significant P values (>0.05) are between the Total methylation score values of -10 and -2; this is the population for which it is not possible to predict which diet is better (n=126). While the individuals (n=75) who have a score between -13 and -11 or between -1 and 11, a type of diet for losing weight can be recommended to them according to their baseline methylation score value (p<0.05).

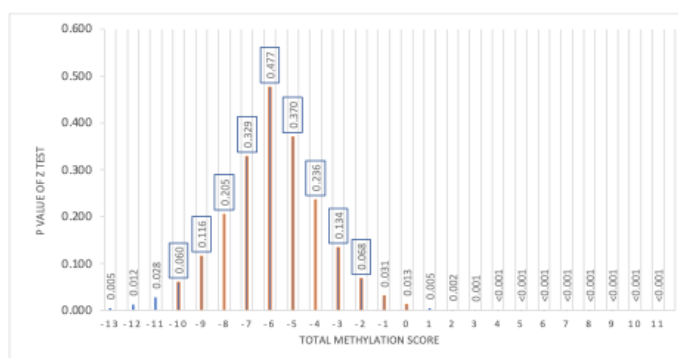


Figure 2. Representation of the distribution of the P value of the Z test with respect to the total methylation score. "X" axis shows the total score. "Y" axis shows P value of Z test. A "P" value <0.05 was considered significant. From -10 to -1 of the Methylation Score a non-significant P is presented.

4. Discussion

In this research, CpG sites whose baseline methylation was associated with BMI reduction after 4 months of dietary intervention in a population were identified, then used

to construct a total methylation score that was finally included in a model along with the two types of diet. This model adequately predicts the percentage of BMI loss.

Currently, the usual management for overweight and obesity is done by calculating energy expenditure, establishing daily nutritional requirements and performing caloric restriction, in addition to giving recommendations for lifestyle changes specific to each individual [10], [11]. The main limitation of this routine management is that it does not take into account the variability of each individual's response to these interventions. Therefore, a prediction model was designed based on methylation data from the studied population that predicts the percentage of BMI loss with the MHP diet and the LF diet. With the aim of obtaining a methylation score that predicts the most appropriate diet for percentage of BMI loss to an intervention for the dietary management of obesity or overweight. The model that predicts the percentage of BMI loss was made with a total methylation score that was constructed considering the MHP diet methylation score in which 15 CpG sites were selected that were better associated with the reduction of BMI percentage and 11 CpG sites for LF diet. With this model it is possible to predict with which diet the highest percentage of BMI would be lost for 75 participants corresponding to 37.3% of the total population of this study. This total methylation score could be used to predict which diet is more appropriate for each individual or if it can be chosen by dietary preference, since there are subjects that with both diets would lose a similar percentage of BMI. The potential of DNA methylation to predict BMI loss has been described before, as demonstrated in a study in which a prediction score was performed based on 83 CpG sites that found association with BMI. The result they obtained was a prediction that represented 29% of the variation in BMI in the studied population [12]. The prediction model of the percentage of BMI loss has no similarity in CpG sites and genes related to these CpG sites for both diets. This may be due to the differences in the investigations, such as different methods and selection parameters of the CpG sites, sample size, population characteristics, and experimental protocols.

5. Conclusions

This research demonstrates that DNA methylation is an individual characteristic that can be used to have greater precision in the nutritional treatment of obesity. The model designed based on the methylation information through the linear mixed effect model allows predicting the loss of BMI percentage. This model can be useful to determine which diet is more adequate for weight loss for each individual.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used “Conceptualization, F.I.M. and J.A.M.; methodology, F.I.M. and J.I.R.; formal analysis, J.I.R. and S.G-C.; investigation, N.G-A. and S.G-C.; resources, F.I.M. and J.A.M.; data curation, J.I.R.; writing—original draft preparation, N.G-A.; writing—review and editing, F.I.M. and S.G-C.; supervision, F.I.M. and S.G-C.; project administration, F.I.M. and S.G-C.; funding acquisition, F.I.M. and J.A.M. All authors have read and agreed to the published version of the manuscript.” Please turn to the [CRediT taxonomy](#) for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Author Contributions:

Funding: This research was funded by CIBEROBN (Grant number: CB12/03/30002), Government of Navarra (Obekit-PT024 and Microbiota-PI035 projects), the Spanish Ministerio de Ciencia, Innovación y Universidades (reference RTI2018-102205-B-I00).

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement:

Acknowledgments: The authors acknowledge the Obekit team (Miren Iosune Zubieta, Laura Olazarán, Salomé Pérez and Ana Lorente) for collaborating in the recruitment and follow-up of the volunteers, sampling and biochemical analyses.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Wu FY, Yin RX. Recent progress in epigenetics of obesity. *Diabetol Metab Syndr.* **2022**, *14*, 171.
2. Sun D, Heianza Y, Li X, Shang X, Smith SR, Bray GA, et al. Genetic, epigenetic and transcriptional variations at NFATC2IP locus with weight loss in response to diet interventions: The POUNDS Lost Trial. *Diabetes Obes Metab.* **2018**, *20*, 2298–303.
3. Samblas M, Milagro FI, Martínez A. DNA methylation markers in obesity, metabolic syndrome, and weight loss. *Epigenetics.* **2019**, *14*, 421–444.
4. Fuso A, Raia T, Orticello M, Lucarelli M. The complex interplay between DNA methylation and miRNAs in gene expression regulation. *Biochimie.* **2020**, *173*, 12–16.
5. Czogała W, Strojny W, Schab M, Grabowska A, Miklusiak K, Kowalczyk W, et al. Fto and plag1 genes expression and fto methylation predict changes in circulating levels of adipokines and gastrointestinal peptides in children. *Nutrients.* **2021**, *13*, 35–85.
6. Livingstone KM, Ramos-Lopez O, Pérusse L, Kato H, Ordovas JM, Martínez JA. Precision nutrition: A review of current approaches and future endeavors. *Trends in Food Science and Technology.* **2022**, *128*, 253–264.
7. World Medical Association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA.* **2013**, *310*, 2191–4.
8. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. **2012**, *13*, 86.
9. Lindsey C. Stat Softw Components. Boston College Department of Economics. VSELECT: Stata module to perform linear regression variable selection. Available from: <https://ideas.repec.org/c/boc/bocode/s457808.html>. [accessed on 23 august 2023]
10. Thibault R, Abbasoglu O, Ioannou E, Meija L, Ottens-Oussoren K, Pichard C, et al. ESPEN guideline on hospital nutrition. *Clin Nutr.* **2021**, *40*, 5684–5709.
11. Ballesteros Pomar MD, Vilarrasa García N, Ángel M, Herrera R, Barahona MJ, Bueno M, et al. Abordaje clínico integral SEEN de la obesidad en la edad adulta. *Endocrinol Diabetes Nutr (Engl Ed).* **2021**, *68*, 130–136.
12. Do WL, Sun D, Meeks K, Dugué PA, Demerath E, Guan W, et al. Epigenome-wide meta-analysis of BMI in nine cohorts: Examining the utility of epigenetically predicted BMI. *Am J Hum Genet.* **2023**, *110*, 273–283.