“GENYAL” Study to Childhood Obesity Prevention: Methodology and Preliminary Results

Helena Marcos-Pasero 1,2, Elena Aguilar-Aguilar 1, Rocío de la Iglesia 3, Isabel Espinosa-Salinas 4, Susana Molina 5, Gonzalo Colmenarejo 6, J. Alfredo Martínez 7,8,9, Ana Ramírez de Molina 10, Guillermo Reglero 11,12 and Viviana Loria-Kohen 1,13*

Objective: This article describes the methodology and summarizes some preliminary results of the GENYAL study aiming to design and validate a predictive model, considering both environmental and genetic factors, that identifies children who would benefit most from actions aimed at reducing the risk of obesity and its complications.

Design: The study is a cluster randomized clinical trial with 5-year follow-up. The initial evaluation was carried out in 2017. The schools were randomly split into intervention (nutritional education) and control schools. Anthropometric measurements, social and health as well as dietary and physical activity data of schoolchildren and their families are annually collected. A total of 26 single nucleotide polymorphisms (SNPs) were assessed. Machine Learning models are being designed to predict obesity phenotypes after the 5-year follow-up.

Settings: Six schools in Madrid.

Participants: A total of 221 schoolchildren (6–8 years old).

Results: Collected results show that the prevalence of excess weight was 19.0, 25.4, and 32.2% (according to World Health Organization, International Obesity Task Force and Ortega Foundation criteria, respectively). Associations between the nutritional state of children with mother BMI \( \beta = 0.21 \) (0.13–0.3), \( p \) (adjusted) < 0.001, geographical location of the school \( \text{OR} = 2.74 \) (1.24–6.22), \( p \) (adjusted) = 0.06, dairy servings per day \( \text{OR} = 0.48 \) (0.29–0.75), \( p \) (adjusted) = 0.05] and 8 SNPs \([\text{rs1260326, rs780094, rs10913469, rs328, rs7647305, rs3101336, rs2568958, rs925946}] \); \( p \) (not adjusted) < 0.05] were found.
CONCLUSIONS: These baseline data support the evidence that environmental and genetic factors play a role in the development of childhood obesity. After 5-year follow-up, the GENYAL study pretends to validate the predictive model as a new strategy to fight against obesity.

CLINICAL TRIAL REGISTRATION: This study has been registered in ClinicalTrials.gov with the identifier NCT03419520, https://clinicaltrials.gov/ct2/show/NCT03419520.

KEYWORDS: pediatric obesity, early intervention, single nucleotide polymorphisms, machine learning, nutrition

INTRODUCTION

Obesity is a complex, chronic and multifactorial disease, originated as an interaction between genetic and environmental factors (1). The prevalence of overweight and obese children is rising every year. Specifically, in compliance with the WHO, the number of overweight and obese children aged 0–5 years increased from 32 million globally in 1990 to 41 million in 2016. And it is expected to increase to 70 million by 2025 if these trends continue (2). The situation in Spain is also alarming, with a prevalence of 23.2% of overweight (22.4% boys and 23.9% girls), and 18.1% of obesity (20.4% boys and 15.8% girls) according to data from the ALADINO study carried out by the Spanish Agency for Consumer Affairs, Food Safety and Nutrition (3).

Childhood obesity usually leads to adulthood obesity, which increases the risk of developing certain diseases, such as hypertension, type 2 diabetes and cardiovascular diseases, prematurely (4–6). This early age has been identified as a key point for the implementation of healthy dietary and lifestyle patterns. Thus, the home and schools provide a useful environment to develop educational and lifestyle interventions for school-age children (7).

There is no doubt about the multifactorial etiology of obesity in which socio-cultural, dietetic, environmental and genetic factors are involved (8–11). However, current knowledge is still insufficient to determine the relative importance of these different factors, having a complex network of associations between them (12). In this regard, machine learning techniques represent a powerful prediction tool through their great ability to big data analysis. Thus, Machine learning represents a tool based on a set of algorithms that can characterize, adapt, learn, predict, and analyze data, increasing the knowledge of obesity and offering possibilities of predicting the disease with unprecedented precision. These techniques have been proposed as a potential tool to predict a future excess of body weight and its comorbidities. There are several predictive machine learning algorithms such as neural networks, decision tree analysis or random forest. Each of them should be used according to the purpose and nature of the study variables (13).

Considering all the above-mentioned aspects, the main objective of the GENYAL study is to design and validate a machine learning-based predictive model that identifies children who would benefit most from actions aimed at reducing the risk of obesity and its complications, considering both environmental and genetic factors, and applicable at the beginning of the school stage. The nutritional education developed in the intervention's schools will be also evaluated as part of the predictive model. This article describes the methods and analyses that will be applied. In addition, it summarizes some preliminary results obtained after the first year of the data collection.

METHODS

TYPE OF STUDY AND DURATION

The present study is a cluster randomized clinical trial with 5-year follow-up intervention based on nutritional education, annual anthropometric measurement evaluations and data collection from questionnaires. Saliva samples were collected for all the schoolchildren in the initial evaluation (2017) in order to obtain genetic information. The final evaluation will be carried out 4 years after the initial intervention, which corresponds to the end of the primary school (Figure 1). The study is therefore expected to last 5 years, from 2016 to 2021 academic year to 2021–2022. Table 1 provides a schedule of activities and interventions throughout the study.

RECRUITMENT, SAMPLE SIZE, AND SAMPLE CHARACTERISTICS

Due to the nature of the study as a clinical trial, the large number of variables necessary for the design of the preventive model with machine learning (each of them of a very different nature), and the duration of 5 years, a statistically robust sample size could not be implemented. Furthermore, the Consejería de Educación e Investigación de la Comunidad de Madrid was responsible for the selection of six representative schools of the Autonomous Community of Madrid (ACM) (Spain) (two in the north, two in the center and two in the city’s south zone), considering the number of students per center and the average socioeconomic level of the districts and neighborhoods. Therefore, the selection was representative of the average income of the ACM households (14). All the School Boards approved the participation in the study and included a total of 569 potential children participants from different districts of Madrid: Chamberi, Hortaleza, Carabanchel, Puente de Vallecas, and Moncloa-Aravaca.
**Inclusion and Exclusion Criteria**

The inclusion criteria to participate in the study were: being in 1st or 2nd grade of primary school and having an informed consent signed by at least one of the parents.

Exclusion criteria were not attending school during the evaluation days or having planned not to stay at the school the following years.

**Randomization**

In order to avoid cross-contamination between intervention and control subjects, randomization was carried out by school center instead of individually. Thus, participating schools were randomly and proportionally stratified into two groups: intervention schools and control schools, considering the number of participants per center, their geographic area and their socioeconomic status.

The randomization procedure was carried out with the statistical software R version 3.4 (www.r-project.org).

**Ethical Aspects and Data Processing**

Protocols and methodology used in the present study comply with the ethical principles for research involving human subjects laid down in the Declaration of Helsinki (1964) and its modifications. The study was approved by the Research Ethics Committee of the IMDEA Food Foundation (PI:IM024; Approval date: March 29th, 2016) and it has been registered in ClinicalTrials.gov with the identifier NCT03419520. School centers and families were informed in detail about the different stages of the project both, orally and in writing. Signed informed consent from at least one of the parents were collected by the researchers prior to the first evaluation. This document included a specific consent to DNA extraction and the evaluation of polymorphisms from the saliva samples. In addition, it included a section on the storage of the remaining samples as a collection registered, according to Spanish legislation (Royal Decree 1716/2011, of November 18th).

Data compiled along the study are going to be processed using a web application that applies dissociation criteria making the volunteers’ data anonymous, in compliance with the current...
Spanish legislation (Organic Law 15/1999 of December 13th, on the protection of Personal Data) and may be used for scientific purpose as publications and conferences. Only the researchers directly related to the study will be allowed to access data.

Selection of Single Nucleotide Polymorphisms
A total of 26 single nucleotide polymorphisms (SNPs) associated with a higher risk of early-age onset of obesity and its comorbidities were selected. The selection was made considering the biological activity of each SNP, Caucasian allele frequencies and the scientific evidence that supports the association between the presence of the polymorphism and the risk of developing overweight, obesity or its complications. The sum of the risk alleles will further be used to design a genetic risk score.

Different databases such as 1,000 Genomes, HapMap, Pubmed, GWAS Central, GWAS Catalog or Ensembl were used. Table 3 shows the selected 26 SNPs, which will be included in the predictive model.

Questionnaires
Different questionnaires were designed based on other surveys used in similar studies to facilitate the comparison of the results. All of them are annually sent to families by email or in the paper format according to the parents’ preference and are filled by at least one of the parents. The information collected is summarized in Table 3.

Regarding social, health and demographic data, parents annually complete a self-reported questionnaire that includes different personal questions based on the surveys used in the ALADINO and ELOIN studies (3, 112).

Dietary information is gathered using a 48-h food record of 2 non-consecutive days, a weekday and a weekend day, as recommended by the European Food Safety Authority guidelines (113). Afterwards, the data are tabulated and analyzed using the DIAL software (Alce Ingeniería, Madrid, Spain) (114) in order to obtain information about macro and micronutrients.

Moreover, the adherence to the Mediterranean diet pattern is assessed using the “KIDMED Mediterranean Diet Quality Index” in addition to general questions about the dietary habits of the children and their parents. The KIDMED questionnaire consists of a total of 16 dichotomous questions that must be answered affirmatively or negatively to obtain a score (115).

Physical activity and free time data about the children and their parents are gathered using a questionnaire with different sections adapted and modified from the ALADINO (3) and the ELOIN (112) studies. In addition, a 48-h physical activity record is collected, corresponding to 24 h of a weekday and a complete weekend day (116). In the physical activity record, parents had to specify the time that their children spent during 24 h of a week day and 24 h of a weekend day doing different activities, including resting hours and activities with a variable level of intensity (very light, light, moderate and intense). The time spent doing each activity is multiplied by the corresponding activity coefficient defined by the WHO (117), added and divided by 24, obtaining the Individual Physical Activity Coefficient (IPAC). Then, the IPAC corresponding to a weekday is multiplied by 5 and the weekend IPAC by 2, both results are added and divided by 7, thus, obtaining the median physical activity per individual. Afterwards, it is necessary to convert the IPAC into a Physical Activity Coefficient (PAC) according to sex, therefore an equivalence is made between the IPAC and the PAC proposed by the Institute of Medicine (118). Finally, participants are classified into sedentary, lightly active, active and very active in line with their PAC.

All these data and information are collected every year on equal terms.

Anthropometric and Blood Pressure Measurements
These data are collected in the school centers, early in the morning, by previously trained nutritionists, following standardized protocols and WHO international instructions for this age group (117). For the anthropometric measurements, children had to wear a T-shirt and gym shorts. All measures are taken twice, and the average is used for the analyses.

Height is determined using a Leicester height rod with an accuracy of 1 mm (Biological Medical Technology SL, Barcelona, Spain). Body weight and fat mass percentage are assessed using a BFR511 Body Composition Monitor (BF511- OMRON HEALTHCARE UK, LT, Kyoto, Japan). Furthermore, fat mass percentage is classified according to the tables offered by OMRON Healthcare (119). Waist and brachial circumferences measurements are taken using a non-elastic tape (KaWe Kirchner & Wilhelm GmbH, Asperg, Germany; range 0–150 cm, 1 mm of precision). The waist circumference measurements obtained are classified by percentiles in compliance with Fernández et al. (120). Triceps skinfolds are taken following the International Society for the Advancement of Kinanthropometry guidelines (121) using a mechanic caliper (HOLTAIN LTD. CRYMYCH UK 10 g/mm² constant pressure; range 0–39 mm and 0.1 mm of precision) and the results obtained are ranked according to percentiles proposed by Frisancho AR (122).

Using these data, other variables of interest are calculated. In particular, BMI is calculated as the body weight divided by the squared height (kg/m²). There is not a universal technique to classify the BMI values in the pediatric collective, therefore, the results are ranked according to the percentiles of Faustino Orbeigozo Eizaguirre Foundation, reviewed in 2011 (120), International Obesity Task Force reviewed in 2000 (123), and WHO reviewed in 2007 (124). The results of overweight and obesity rates are unified as a single category called excess weight (EW). The arm muscular and fat areas are obtained using the equations proposed by Mataix Verdu and López Jurado (125) and López-Sobaler and Quintas Herrero (126), respectively. The protein and caloric reserves are calculated by Frisancho AR equations (122). Waist/height ratio is calculated as waist circumference (cm)/height (cm) and it was classified according to Panjikkaran et al. and Ashwell investigations (127, 128). Height/age index is rated in percentiles according to Fernández et al. (120).

For blood pressure monitoring, an automatic digital monitor is used (OMRON M3-Intellisense) using a cuff suitable for
| Gene  | Name         | Function | Location       | Alleles | WT | MAF<sup>b</sup> | Frequency<sup>b</sup> | Association                                      | References |
|-------|--------------|----------|----------------|---------|----|----------------|-----------------------|--------------------------------------------------|------------|
| APOA5 | rs662799     | Upstream gene variant | Chromosome 11:116792991 | G       | A  | 0.16 (G)     | G: 0.076/A: 0.924     | Overweight and obesity, lipid profile, CVD, MS     | (15–19)   |
| BDNF  | rs925946     | Intron variant | Chromosome 11:27645655  | T       | G  | 0.25 (T)     | T: 0.338/G: 0.662     | Overweight and obesity, calcium consumption, MS, waist/hip index, abdominal obesity | (20–25)   |
| ETV5  | rs7647305    | Intron variant | Chromosome 3:186116501  | T       | C  | 0.24 (T)     | T: 0.227/ C: 0.773    | Overweight and obesity, early menarche, BP         | (20, 24, 26, 27) |
| FTO   | rs3751812    | Intron variant | Chromosome 16:53784548  | G       | T  | 0.22 (T)     | G: 0.556/T: 0.444     | Overweight and obesity, lipid profile              | (28–31)   |
| FTO   | rs7190492    | Intron variant | Chromosome 16:53794840  | A       | G  | 0.30 (A)     | A: 0.343/G: 0.657     | Overweight and obesity                             | (28, 29, 32) |
| FTO   | rs8050136    | Intron variant | Chromosome 16:53782363  | C       | A  | 0.32 (A)     | C: 0.556/A: 0.444     | Overweight and obesity, T2DM                      | (28, 29, 33–36) |
| FTO   | rs9939609    | Intron variant | Chromosome 16:53786615  | T       | A  | 0.34 (A)     | T: 0.556/A: 0.444     | Overweight and obesity, extreme childhood obesity, T2DM, MS, early menarche, fat mass | (20, 27, 37–41) |
| GCKR  | rs1260326    | Missense variant | Chromosome 2:27508073   | T       | C  | 0.29 (T)     | T: 0.429/C: 0.571     | Overweight and obesity, MS, lipid profile, glycemic profile | (42–45)   |
| GCKR  | rs780094a    | Intron variant | Chromosome 2:27518370   | T       | C  | 0.30 (T)     | T: 0.409/C: 0.591     | Overweight and obesity, lipid profile, MS, glycemic profile, fatty liver, T2DM | (42, 43, 46–50) |
| GNPDA2| rs10938397   | Intergenic variant | Chromosome 4:45180510   | A       | G  | 0.33 (G)     | A: 0.576/G: 0.424     | Overweight and obesity, extreme childhood obesity, early menarche, T2DM, glycemic profile, fat mass, waist circumference, waist-height ratio, central obesity | (20, 27, 40, 41, 51) |
| KCTD15| rs368794     | Intergenic variant | Chromosome 19:38298647   | T       | A  | 0.42 (T)     | T: 0.296/A: 0.702     | Overweight and obesity, fat mass                   | (20, 41, 52, 53) |
| LEPR  | rs1137101    | Missense variant | Chromosome 1:65592830   | A       | G  | 0.42 (A)     | A: 0.515/G: 0.485     | Obesity, lipid profile, glycemic profile, caloric intake, T2DM | (54–58)   |

(Continued)
| Gene | Name | Function | Location\(^b\) | Alleles | WT | MAF\(^b\) | Frequency\(^b\) | Association | References |
|------|------|----------|----------------|--------|-----|--------|----------------|-------------|------------|
| LPL  | rs328| Stop gained | Chromosome 8:19962213 | C G C | 0.09 (G) | C: 0.874/G: 0.126 | Lipid profile, CVD | (59–63) |
| MC4R | rs17782313 | Intergenic variant | Chromosome 18:60183864 | T C T | 0.24 (C) | T: 0.742/C: 0.258 | Overweight and obesity, extreme childhood obesity, MS, T2DM, abdominal obesity, fat mass, waist circumference, waist-height ratio | (22, 41, 64–69) |
| NEGR1| rs2568958 | Upstream gene variant | Chromosome 1:72294933 | G A G | 0.32 (G) | G: 0.364/A: 0.636 | Overweight and obesity, extreme childhood obesity, fat mass. | (20, 41, 70) |
| NEGR1| rs3101336 | Upstream gene variant | Chromosome 1:72285502 | T C T | 0.32 (T) | T: 0.364/C: 0.636 | Overweight and obesity, severe early-onset obesity. | (70–72) |
| NPY  | rs16147 | Upstream gene variant | Chromosome 7:24283791 | T C T | 0.48 (T) | T: 0.485/C: 0.515 | Overweight and obesity | (73–75) |
| PPAR\(\gamma\) | rs1801282 | Missense variant | Chromosome 3:12351626 | C G C | 0.07 (G) | C: 0.904/G: 0.096 | Overweight and obesity, glycemic profile, CVD, T2DM, Skin folds | (76–82) |
| SEC16B| rs10913469 | Intronic variant | Chromosome 1:177944384 | T C C | 0.23 (C) | T: 0.758/C: 0.242 | Overweight and obesity, waist circumference, fat mass, early menarche | (27, 28, 30, 55, 53, 84) |
| TCF7L2 | rs7930146 | Intronic variant | Chromosome 10:112996590 | C T T | 0.23 (T) | C: 0.687/T: 0.313 | IMC, T2DM, MS, glycemic profile, height, lipid profile in DM patients | (85–93) |
| TMEM18 | rs2867125 | Intergenic variant | Chromosome 2:622827 | T C C | 0.14 (T) | T: 0.162/C: 0.838 | Overweight and obesity | (30, 94, 95) |
| TMEM18 | rs4854344 | Intragenic variant | Chromosome 2:638144 | G T G | 0.15 (G) | G: 0.167/T: 0.833 | Overweight and obesity, waist circumference, visceral fat and SBP | (83, 96–98) |
| TMEM18 | rs6548238 | Regulatory variant | Chromosome 2:634906 | T C C | 0.12 (T) | T: 0.167/C: 0.833 | Overweight and obesity, extreme childhood obesity, fat mass, early menarche, waist circumference | (20, 27, 41, 55, 65, 97, 99, 103) |
| TMEM18 | rs7561317 | Intergenic variant | Chromosome 2:644953 | A G A | 0.16 (A) | A: 0.162/G: 0.838 | Overweight and obesity | (21, 22, 24, 33, 40, 101, 102) |

(Continued)
TABLE 2 | Continued

| Gene | Alleles | WT | Location | Frequency | Association |
|------|---------|----|----------|-----------|-------------|
| UCP2 |          |     | Chromosome11:73983709 | C T C | 0.41 (T) C: 0.636/T: 0.364 |
| UCP3 |          |     | Chromosome11:74009120 | G A G | 0.19 (A) G: 0.768/A: 0.232 |

Approved name (111). Ensembl (https://www.ensembl.org/index.html). Location, for forward strand. WT, wild type; MAF, Minor allele frequency; Frequency, Allele and genotype; APOA5, apolipoprotein A5; BDNF, brain derived neurotrophic factor; ETV5, ETS variant 5; FTO, FTO, alpha-ketoglutarate dependent dioxygenase; GCKR, glucokinase regulator; GNPDA2, glucosamine-6-phosphate deaminase 2; KCTD15, potassium channel tetramerization domain containing 15; LEPR, leptin receptor; LPL, lipoprotein lipase; MC4R, melanocortin 4 receptor; NEGR1, neuronal growth regulator 1; NPY, neuropeptide Y; PPARγ, peroxisome proliferator activated receptor gamma; SEC16B, SEC16 homolog B; TCF7L2, transcription factor 7 like 2; TMEM18, transmembrane protein 18; UCP2, uncoupling protein 2; UCP3, uncoupling protein 3; CVD, cardiovascular disease; MS, metabolic syndrome; BP, blood pressure; T2DM, type 2 diabetes mellitus; SBP, systolic blood pressure.

TABLE 3 | Content of the questionnaires used in the study.

| Questionnaire | Variables |
|---------------|-----------|
| Characteristics of parents and family | People living with the schoolchildren, number of siblings, family structure, level of education, employment situation, profession, country of birth, years living in Spain, income level, diagnosis of disease, smoking habit, body weight, height and date of birth. |
| Diet | Daily intakes, frequency of breakfast per week, breakfast foods, reasons for stop breastfeeding, desire to extend the breastfeeding period, perception of their child’s diet. |
| Physical activity | Distance between school and home, weekly hours of physical activities, frequency of physical activity shared with their child, town conditions for doing outdoor activities. |
| Characteristics of children | Country of birth, years living in Spain, adopted or foster child, diagnosis of disease, body weight and height reported by their parents and date of birth. |
| Diet | Diet changes over the last year and reasons, number of daily intakes, frequency of breakfast per week, place where they have breakfast, breakfast foods, place where they have lunch, frequency of eating in fast food restaurants, breastfeeding, beikost’s start date. 48-h dietary record |
| Physical activity | Sleeping hours, type of transport used to go to school and back home, outdoor activities, time usually spend watching TV, playing video games or using the computer, presence of TV or computer games in their bedroom, weekly hours of physical activities and children’s physical activity level according to their parents 48-h physical activity questionnaire |

Compiling Saliva Samples, DNA Extraction and Genotyping

Buccal smears were collected for DNA extraction following standardized protocols. For this purpose, a sterile swab free of human RNAse, DNase and DNA (300263DNA-Hisopos Deltalab polystyrene and polyester) was used. Children had to have their mouth clean and avoid eating or drinking 30 min prior to collection. Three samples were taken per children, each one identified with the number corresponding to the order of extraction, to ensure traceability. As the samples were collected, they were directly stored in refrigeration until all the children were evaluated. Immediately after, they were frozen at −80°C until their processing.

Genomic DNA was extracted from the buccal swabs using the INVISORB® SPIN TISSUE MINI KIT (Stratec), according to the manufacturer’s instructions. Samples were lysed in the presence of proteinase K and a specific lysis buffer. The lysate was then purified and finally, it was eluted in a free EDTA solution.

children. The results are classified according to the percentiles established by the Spanish Association of Pediatrics (129).

All these measurements are repeated every year on equal terms.
TABLE 4 | Educational components of GENYAL study.

| Format       | Module                     | Objective                                                                 |
|--------------|----------------------------|---------------------------------------------------------------------------|
| Children     | Notebook activities         | Essential nutrients: definition, function and sources                      |
|              | The food pyramid            | To identify the different food categories and explore which foods are in each group |
|              | The food pyramid            | To identify the different food categories and which foods are in each group |
|              | Introduction                | To know the bases and objectives of the study                             |
| Parents      | Nutritional guide           | Essential nutrients: definition, function and sources                      |
|              | Food and nutrients          | To identify the different food categories and which foods are in each group |
|              | The food pyramid            | To know about the preliminary results of the study and the recommendations derived from them |
|              | Preliminary results         | To know about the preliminary results of the study and the recommendations derived from them |
| Teachers     | Nutritional guide           | Essential nutrients: definition, function and sources                      |
|              | Food and nutrients          | To know the bases and objectives of the study                             |
|              | The food pyramid            | To identify the different food categories and which foods are in each group |
|              | Preliminary results         | To know about the preliminary results of the study and the recommendations derived from them |

For genotyping, the DNA samples were loaded in TaqMan® OpenArray® Real-Time PCR plates (Life Technologies Inc., Carlsbad, CA) already configured with the specific selected SNPs with specific waves for each allele, marked with a different fluorophore to determine the genotype. This process was made using the OpenArray® AccuFill™ System (Life Technologies Inc., Carlsbad, CA). Once it was charged, a PCR performed and the chips were read in the QuantStudio® 12K Flex Real-Time PCR Instrument (Life Technologies Inc., Carlsbad, CA). Results were analyzed using the TaqMan® Genotyper software (Life Technologies Inc., Carlsbad, CA), which automatically assigns the genotype to each sample according to the amount of detected signal for each fluorophore.

The duplicate analysis was used to validate the genotyping result.

**Statistical Analysis**

Descriptive analyses of the baseline data were performed by computing for the categorical variables the class’s absolute and relative frequencies, and for the quantitative variables the mean, median, standard deviation, interquartile range, maximum and minimum. To check the homogeneity of the two groups in the case of quantitative variables, t-tests were used for normally-distributed variables, or Mann-Whitney U-test as non-parametric alternative. In the case of categorical variables, Chi-Square or Fischer exact tests were used. The association between anthropometric and dietary, social, health and SNP variables were performed by linear or logistic regressions. The Bonferroni correction was applied for multiple tests. In addition, for the SNPs variables, the Hardy-Weinberg equilibrium condition was tested by means of Chi-Square tests. All analyses were conducted with R Statistical Software version 3.4.1. Statistical tests used a 0.05 significance level, in two-tailed tests.

Regarding Machine Learning models, they will be derived to predict the BMI from all the analyzed variables after the 5-year follow-up. Both classification (after dichotomization of the BMI) and regression models will be considered, and Random Forest will be applied. It has been observed that the use of Random Forest improves the predictive model’s performance, creating a more effective predictive model than the one that could be obtained using decision tree or logistic regression techniques (13). The predictive power of the models will be evaluated and internal cross-validation and external validations with external datasets will be implemented. Variable importance analyses will be performed in order to quantify the relative weights of the different variables in the prediction of BMI. The model will be iteratively improved by refitting with new data along with the successive yearly evaluations during the study.

**RESULTS**

Parents of 224 children (116 girls and 105 boys) accepted to participate in the study and signed the informed consent. It shows a collaboration rate of 39%. Finally, 221 children were evaluated, since three did not attend the initial evaluation (Figure 2).
Among the total number of students enrolled in the study, 115 belonged to intervention schools while 106 were in control schools. Tables 5, 6 show the basal main characteristics of the study sample according to the corresponding control or intervention group.

According to the preliminary results obtained after the first year of the evaluation, 32.2% of the students presented EW, taking into account the WHO criteria. These figures were higher when IOFT standard (25.4%) or the national criteria of the Orbegozo Foundation (19.0%) were applied.

A significative association between the nutritional state of children and mother BMI \( \beta = 0.21 (0.13–0.3), p (\text{adj}) < 0.001 \), geographical location of the school \( \text{OR} = 2.74 (1.24–6.22), p (\text{adj}) = 0.06 \), dairy servings per day \( \text{OR} = 0.48 (0.29–0.75), p (\text{adj}) = 0.05 \) and eight of the total SNPs studied \( \text{rs1260326, rs780094, rs10913469, rs328, rs7647305, rs3101336, rs2568958, rs925946; p (\text{not adj}) < 0.05 } \) were found (data adjusted for sex and age).

**DISCUSSION**

The GENYAL study to prevent child obesity is, to our knowledge, the first interventional trial in Spanish schoolchildren aiming to provide preventive and therapeutic approach based on a high degree of evidence for early obesity through machine learning.

The baseline data and the associations observed after the first analysis support the evidence that environmental and genetic factors play a role in the development of childhood obesity.

According with the results, for each point in the BMI of the mother, the BMI of the child was increased 0.21 kg/m\(^2\). It shows that among the multiple risk factors for the development of obesity in children, parental obesity is one of the most impactful as a result of both genetic and environmental interactions. Children imitate their parents, therefore, the parents’ dietary habits and PA are more likely to be reproduced by their descendants (130).

With reference to the location of the school, that represent the socioeconomic level, presented a close relationship with the presence of EW, with a decreasing distribution of risk from south to north area. These results are consistent with other studies that have shown how the socioeconomic status of the school correlates with the prevalence of overweight and obesity as it increases the likelihood that schoolchildren will follow a diet rich in energy-dense, low-cost foods, as well as fewer opportunities to practice sport (131).
Regarding dietary aspects, dairy servings per day showed a protective effect against EW. It could be related to several factors such as if this food is a source of calcium, peptides, bioactive compounds, etc. They have been studied due to their relationship in the appetite control and other mechanisms involved in controlling weight (132, 133). These results highlight the important role that this group of foods could have for the prevention of weight overload.

The study of these factors in the child population and the social context of Madrid is minimal, having found a single study with similar sociodemographic characteristics (134). Nevertheless, no intervention was performed in this study, nor was genetic data collection, thus the GENYAL project is shown as a novel study in this regard.

Table 5: Main socio-demographic and anthropometric characteristics of the schoolchildren by groups (intervention schools vs. control schools).

| Socio-Demographic | Intervention group (n = 115) | Control group (n = 106) | p    |
|-------------------|-----------------------------|-------------------------|------|
| Spanish (%)       | 98.10                       | 97.00                   | 0.678|
| Spanish mother (%)| 35.00                       | 23.50                   | 0.074|
| Anthropometry     | 6.73 ± 0.76                 | 6.77 ± 0.69             | 0.506|
| Sex (%)           | 52.20                       | 42.50                   | 0.148|
| Fat mass (%)      | 21.23 ± 7.56                | 19.84 ± 6.71            | 0.159|
| Muscle mass (%)   | 27.97 ± 3.00                | 28.01 ± 2.97            | 0.082|
| BMI (kg/m²)       | 17.21 ± 2.99                | 16.68 ± 2.24            | 0.319|

Table 6: Main dietary and physical activity characteristics of the schoolchildren by groups (intervention schools vs. control schools).

| Physical activity | Intervention group (n = 115) | Control group (n = 106) | p    |
|-------------------|-----------------------------|-------------------------|------|
| Physical activity rate | 1.61 ± 0.11               | 1.55 ± 0.09             | <0.001|
| Number of hours dedicated to moderate and vigorous activities | 3.53 ± 1.72 | 3.96 ± 1.89 | 0.068|
| Diet              |                            |                         |      |
| Total caloric value (kilojoules) | 7,695.97 ± 1,962.52       | 7,678.59 ± 1,463.50    | 0.378|
| Carbohydrates %   | 45.00 ± 5.15               | 43.94 ± 5.43            | 0.154|
| Lipids %          | 38.40 ± 4.91               | 39.52 ± 5.09            | 0.116|
| Proteins %        | 16.57 ± 2.25               | 16.52 ± 2.09            | 0.971|
| Total fiber (g)   | 18.54 ± 6.37               | 17.78 ± 5.19            | 0.481|
| KIDMED score      | 6.33 ± 1.98                | 6.67 ± 1.82             | 0.251|

In this study, a significant association (data adjusted for sex and age) between the nutritional status of schoolchildren and 8 SNPs was found (rs1260326, rs780094, rs10913469, rs328, rs7647305, rs3101336, rs2568958, rs925946). In previous studies, these polymorphisms have been associated with adiposity traits and their related comorbidities (Table 2). Genetic factors play an essential role in the development of obesity (135). Thus, knowing which ones are associated with excess weight early in life could contribute to obesity early detection.

Conversely, it is important to note that current knowledge is insufficient to determine the relative importance of these different factors. Therefore, new techniques are needed to be used as predictive instruments (12). Currently, machine learning is considered an extremely valuable tool in the medical field, since it is capable of providing diagnostic and early detection strategies for diseases through the analysis of large datasets (136). Prevention plays a crucial role in controlling the high obesity prevalence, so machine learning techniques have already been used for the prediction of the BMI in children (137). However, the current predictive model would be the first to include obesity-related SNPs as genetic information, as well as anthropometric, social and lifestyle variables.

Regarding the last results from the Commission on Ending Childhood Obesity, the implementation of integral programs that promote healthy environment in schools is recommended with the objective of ensuring that children grow well and develop healthy habits (138). Nevertheless, although Spain is one of the countries where more intervention studies to prevent obesity have been developed (139), the politic strategies to prevent chronic illnesses such as overweight and obesity are not defined, and even show very low evidence of efficiency, according to the last data revised by the Cochrane Database (140).

According to the latest scientific research, the intervention studies in schools which include family and community spheres, implementing actions to promote healthy food and physical activity, are the most effective (5, 46). This study has been designed to elaborate strategies and to work as a multidisciplinary team reinforcing the educational sphere of the participating children and their environment, school, and family.

One of the strong points of our study consists in the implementation, and subsequent validation, of educational tools for students, their parents and teachers, applying a nutritional education method that promotes healthy dietetic habits and physical exercise, both in schools and outside. The importance of the validation of these educational strategies lay on a large number of studies with contradictory results, which might be partly explained by the fact that many researches may have lacked statistical power to detect changes in the results of interest related to adiposity or children’s nutritional status (5). In the present study, the educational approaches to the intervention schools will be held for 5 years, in line with the annual anthropometric assessments. Their utility will be evaluated taking into account the evolution of the anthropometric and dietetic results annually collected. Moreover, the presence or absence of the educational program (control and intervention schools) will be included as an input dichotomous variable in the predictive model, evaluating its influence on the predicted BMI. This will
allow us to detect differences in the body composition between intervention and control schools, enabling us to assess the impact of educational support.

Another strength of the present study is the selection of 26 SNPs for early prevention of obesity, by making an extensive bibliographic research, as shown in Table 2. The nutritional genomic tools would be very useful in the research and prevention of obesity, and they would be an important support in public health applications. Obesity is a multifactorial illness, where the genetic variants involved are dispersed along the whole genome. Although SNPs have been cataloged as the best indicators to predict obesity risk (141), several studies suggest that much remains to be discovered. There is a lot of interest in predicting the appearance of chronic diseases at an early age (142). According to the last review about precision nutrition (143), the creation of a genetic risk score may let us determine the risk of developing obesity or other chronic pathologies related to the individual genetic component, and even be able to predict the expected weight gain as a consequence of exposure to different variables, such as specific diets.

On the other hand, we consider that the sample size used is one of the weak aspects in our study, pointing to the necessity to include new schools to increase the number of children. Nevertheless, from this first phase of the study, we expect to calculate the sample size needed to increase the statistical power and, consequently, to find solid associations between the studied variables. Furthermore, the classification of schools according to the area and the socioeconomic level widens the scope of the research, making it more representative of the city of Madrid. Similarly, the use of dietary and physical activity questionnaires may lead to reporting bias, but in the absence of better tools with low cost and high throughput, these records can offer valuable information, although it should be interpreted with caution.

After 5 years of follow-up, the GENYAL study aims to validate the machine learning predictive model that considers environmental and genetic factors in the obesity development, as well as the educational tools, to obtain new and potentially valuable data to increase our knowledge of the precipitants of childhood obesity and their relative importance to design preventive protocols at an early age based on machine learning models. With a view to the future perspective of continuity of this study, in addition to increasing the sample size to validate the results obtained, the possibility of implementing personalized nutritional education interventions is proposed to improve adherence and efficacy by applying the novel concepts provided by studies on precision nutrition.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author/s.

**ETHICS STATEMENT**

The studies involving human participants were reviewed and approved by Ethics Committee of Fundación IMDEA-Food. Written informed consent to participate in this study was provided by the participants’ legal guardian/next of kin.

**AUTHOR CONTRIBUTIONS**

VL-K was the principal investigator and responsible for the study and protocol design. HM-P helped designing the protocol and drafting the manuscript. HM-P, EA-A, RI, and IE-S were responsible for data collection. GC conducted statistical analysis of the data. SM contributed to genetic samples management. JM, GR, and AR supervised the final compilation of the manuscript and provided scientific advice and consultation. All authors read and approved the final manuscript.

**FUNDING**

This study was supported by Conserjería de Educación, Universidades y Ciencia de la Comunidad de Madrid, Dirección General de Educación Infantil, Primaria y Secundaria.

**ACKNOWLEDGMENTS**

The authors would like to acknowledge the help of all the teachers and head teachers of Juan Zaragüeta, Fernando el Católico, Fernández Moratin, La Rioja, Concepción Arenal y Rosa Luxemburgo schools; and, especially, the families and children who participated in the study. We would like to thank all the IMDEA-Food students for their contribution in the nutritional education programme development.

**REFERENCES**

1. Güngör NK. Overweight and obesity in children and adolescents. J Clin Res Pediatr Endocrinol. (2014) 6:129–43. doi: 10.4274/jcrpe.1471
2. WHO. Facts and Figures on Childhood Obesity. Available online at: http://www.who.int/.end-childhood-obesity/facts/en/ (accessed December 22, 2019).
3. Agencia Española de Consumo, Seguridad Alimentaria y Nutrición. Estudio ALADINO 2015: Estudio de Vigilancia del Crecimiento, Alimentación, Actividad Física, Desarrollo Infantil y Obesidad en España 2015. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad (2016)
4. Kolotkin RL, Andersen JR. A systematic review of reviews: exploring the relationship between obesity, weight loss and health-related quality of life. Clin Obes. (2017) 7:273–89. doi: 10.1111/cob.12203
5. Bleich SN, Vercammen KA, Zatz LY, Frelier JM, Ebbeling CB, Peeters A. Interventions to prevent global childhood overweight and obesity: a systematic review. Lancet Diabetes Endocrinol. (2017) 4:332–46. doi:10.1016/S2213-8587(17)30358-3
6. Genoni G, Menegon V, Secco GG, Sonzini M, Martelli M, Castagno M, et al. Insulin resistance, serum uric acid and metabolic syndrome are linked to cardiovascular dysfunction in pediatric obesity. Int J Cardiol. (2017) 249:366–71. doi: 10.1016/j.ijcard.2017.09.031
7. Matwiejczyk L, Mehta K, Scott J, Tonkin E, Coveney J. Characteristics of effective interventions promoting healthy eating for pre-schoolers
in childcare settings: an umbrella review. *Nutrients*. (2018) 10:293. 
doi: 10.3390/nu10030293.

8. Igacel I, Fernández-Alvira JM, Labayen I, Moreno LA, Samper MP, 
Rodríguez G. CALINA study. Social vulnerabilities as determinants of 
overweight in 2-, 4- and 6-year-old Spanish children. *Eur J Public 
Health*. (2017) 27:289–95. doi: 10.1093/eurpubck905.

9. Paparo L, di Costanzo M, di Scala C, Cosenza L, Leone L, Nocerino R, et al. 
The influence of early life nutrition on epigenetic regulatory mechanisms 
of the immune system. *Nutrients*. (2014) 6:4706–19. doi: 10.3390/nu6114706.

10. Kumar S, Kelly AS. Review of childhood obesity: from epidemiology, 
etiology, and comorbidities to clinical assessment and treatment. *Mayo Clin 
Proc*. (2017) 92:251–65. doi: 10.1016/j.mayocp.2016.09.017.

11. Kostovski M, Tasic V, Laban N, Polenakovicki M, Daniilovski D, Guec Z. 
Obesity in childhood and adolescence, genetic factors. *PRILOZI*. (2017) 
38:121–33. doi: 10.2478/prilozi-2018-0013.

12. Qasim A, Turcotte M, Souza RJ de, Samaan MC, Champredon D, Dus hoff J, et al. On the origin of obesity: identifying the biological, environmental and 
cultural drivers of genetic risk among human populations. *Obes Rev*. (2018) 
19:121–49. doi: 10.1111/obr.12625.

13. DeGregory KW, Kuiper P, DeSilvio T, Pleuss JD, Miller R, Roginski JW, 
Qasim A, Turcotte M, Souza RJ de, Samaan MC, Champredon D, Dush off J, et al. 
Association analysis of the FTO gene with obesity in children of Caucasian 
and African ancestry reveals a common tagging SNP. *PLoS ONE*. (2008) 
3:e1746. doi: 10.1371/journal.pone.0001746.

14. Mitchell JA, Hakonarson H, Rebbeck TR, Grant SFA. Obesity-susceptibility 
loci and the tails of the pediatric BMI distribution. *Obesity*. (2013) 21:1256– 
60. doi: 10.1002/oby.20319.

15. Qureshi SA, Mumtaz A, Shahnaz SU, Shabana NA. rs3751812, a common 
variant in fat mass and obesity-associated (FTO) gene, is associated with 
serum high- and low-density lipoprotein cholesterol in Pakistani individuals. 
*Nutrition*. (2017) 39–40:92–5. doi: 10.1016/j.nut.2016.04.008.

16. Wang K, Li W-D, Zhang CK, Wang Z, Glessner JT, Grant SFA, et al. Genome- 
wide association studies on obesity and obesity-related traits. *PLoS ONE*. 
(2011) 6:e18939. doi: 10.1371/journal.pone.0018939.

17. Namjou B, Keddache M, Marsolo K, Wagner M, Lingren T, Cobb B, et al. 
EMR-linked GWAS study: investigation of variation landscape of loci for 
body mass index in children. *Front Genet*. (2013) 4:268. 
doi: 10.3389/fgene.2013.00268.

18. Dwivedi OP, Tabassum R, Chauhan G, Ghosh S, Marwaha RK, Tandon N, et al. Functional glucokinase regulator gene variants have inverse effects on 
fasting plasma lipids and risk of metabolic syndrome: evidence from a 
case-control study and meta-analysis. *PLoS ONE*. (2013) 8:e56216. 
doi: 10.1371/journal.pone.0056216.

19. Willer CJ, Speliotes EK, Chasman DI, Perfetti C, Sulem P, Sulem P, et al. 
Common variants near BDNF is associated with early infancy weight gain and growth. *PLoS Med*. (2010) 7:e1000284. 
doi: 10.1371/journal.pmed.1000284.

20. Dušátková L, Zamrazilová H, Al-Haj H, Hainerová I, et al. Association of obesity susceptibility gene variants with metabolic syndrome and related traits in 1,443 Czech adolescents. *Folia Biol*. (2013) 59:123–33.

21. Kvaløy K, Kulle B, Romundstad P, Solheim O, et al. Social determinants of 
weight-affecting gene variants in a life course perspective–The HUNT Study, 
Norway. *Int J Obes*. (2013) 37:1221–9. doi: 10.1038/ijo.2012.220.

22. Sandholt CH, Vestmar MA, Bille DS, Broglykke A, Almind K, Hansen L, et al. Studies of metabolic phenotype correlates of 15 obesity associated 
gene variants. *PLoS ONE*. (2011) 6:e23531. doi: 10.1371/journal.pone.00 
23531.
triglyceride and glucose levels, and decrease the risk of obesity in children. *Diabetes Metab.* (2011) 37:432–9. doi: 10.1016/j.diabet.2011.02.003

44. Santoro N, Zhang CK, Zhao H, Pakstis AJ, Kim G, Kursaw R, et al. Variant in the glucokinase regulatory protein (GCKR) gene is associated with fatty liver in obese children and adolescents. *Hepatology.* (2012) 55:781–9. doi: 10.1002/hep.24806

45. Alfred T, Ben-Shlomo Y, Cooper R, Hardy D, Deary IJ, Elliott J, et al. Associations between a polymorphism in the pleiotropic GCKR and Age-related phenotypes: the HALCyon programme. *PLoS ONE.* (2013) 8:e70045. doi: 10.1371/journal.pone.0070045

46. Wong Y, Cai L, Wu Y, Wilson RF, Weston C, Favole O, et al. What childhood obesity prevention programmes work? A systematic review and meta-analysis. *Obes Rev.* (2015) 16:547–65. doi: 10.1111/obr.12277

47. Chang H-W, Lin F-H, Li P-F, Huang C-L, Chu N-F, Su S-C, et al. Association between a glucokinase regulator genetic variant and metabolic syndrome in Taiwanese adolescents. *Genet Test Mol Biomark.* (2016) 20:137–42. doi: 10.1089/gtmb.2015.0241

48. Zain SM, Mohamed Z, Mohamed R. Common variant in the glucokinase regulatory gene rs780094 and risk of nonalcoholic fatty liver disease: a meta-analysis. *J Gastroenterol Hepatol.* (2015) 30:21–7. doi: 10.1111/jgh.12714

49. Fesinmeyer MD, Meigs JB, North KE, Schumacher FR, Bužková P, Franceschini N, et al. Genetic variants associated with fasting glucose and insulin concentrations in an ethnically diverse population: results from the Population Architecture using Genomics and Epidemiology (PAGE) study. *BMC Med Genet.* (2010) 59:2980–9. doi: 10.2337/db10-0370

50. Li H, Xu R, Peng X, Wang Y, Wang T. Association of glucokinase regulatory protein polymorphism with type 2 diabetes and fasting plasma glucose: a meta-analysis. *Mol Biol Rep.* (2013) 40:3935–42. doi: 10.1007/s11033-012-2470-6

51. Xi B, Cheng H, Shen Y, Chandak GR, Zhao X, Hou D, et al. Study of 11 BMI-associated loci identified in GWAS for associations with central obesity in the Chinese children. *PLoS ONE.* (2013) 8:e56472. doi: 10.1371/journal.pone.0056472

52. Beyerlein A, von Kries R, Ness AR, Ong KK. Genetic markers of obesity risk: stronger associations with body composition in overweight compared to normal-weight children. *PLoS ONE.* (2011) 6:e19057. doi: 10.1371/journal.pone.0019057

53. Li S, Zhao JH, Luan J, Ekelund U, Luben RN, Khaw K-T, et al. Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. *PLoS Med.* (2010) 7:e1000332. doi: 10.1371/journal.pmed.1000332

54. Queiroz EM, Cândido APC, Castro IM, Bastos AQA, Bastos-Goytia L, Muñoz-Valle M, et al. Polymorphisms in LEPR, PPARG and APM1 genes: associations with energy intake and metabolic traits in young children. *OMICS.* (2016) 20:137–46. doi: 10.1089/omi.2015.0045731

55. Sun Z, Shi H, Zhao JH, Jiang Y, Luan J, Ekelund U, et al. Common SNP rs5882 polymorphisms and their relationship with lipid profile in human serum of obese individuals. *Gene.* (2015) 558:195–9. doi: 10.1016/j.gene.2014.12.070

56. White MJ, Eren F, Aigbirshi D, Chen J, Hu T, Moore JH, et al. Systems genetics approach to dyslipidemia in children and adolescents. *OMICS.* (2015) 19:248–59. doi: 10.1089/omi.2014.0140

57. Biermann A, Wirth G, Breslow NE, Breslow JA, Goldin J, et al. Interaction of lipoprotein lipase polymorphisms with body mass index and birth weight to modulate lipid profiles in children and adolescents: the CASPIAN-III Study. *São Paulo Med J.* (2016) 134:121–9. doi: 10.1590/1516-3180.2015.00792608
Guan Y, Yan LH, Liu XY, Zhu XY, Wang SZ, Chen LM. Correlation of Wang S, Song K, Srivastava R, Fathzadeh M, Li N, Mani A. The protective effect of transcription factor 7-like 2 risk allele rs7903146 against elevated fasting plasma triglyceride in type 2 diabetes: a meta-analysis. J Diabetes Invest. (2015) 6:24472. doi: 10.1038/rep.2014.24472

Körner A, Berndt SI, Justice AE, Young KL, Feitosa MF, Barata L, et al. Correction: Genome-wide physical activity interactions in adiposity - A meta-analysis of 200,452 adults. PLoS Genet. (2017) 13:e1006972. doi: 10.1371/journal.pgen.1006972

Prats-Puig A, Grau-Cabrera P, Riera-Pérez E, Cortés-Marina R, Fortea E, Soriano-Rodriguez P, et al. Variations in the obesity genes FTO, TMEM18 and NRXN3 influence the vulnerability of children to weight gain induced by short sleep duration. Int J Obses. (2013) 37:182–7. doi: 10.1038/ioj.2012.27

Graff M, Scott RA, Justice AE, Young KL, Feitosa MF, Barata L, et al. Correction: Genome-wide physical activity interactions in adiposity - A meta-analysis of 200,452 adults. PLoS Genet. (2017) 13:e1006972. doi: 10.1371/journal.pgen.1006972

Marcos-Pasero et al. “GENYAL” Study to Childhood Obesity Prevention
abdominal obesity and serum lipids: the Finnish Diabetes Prevention Study. 
*BMC Med Genet.* (2009) 10:94. doi: 10.1186/1471-2350-10-94

111. HGNC Database of Human Gene Names. HUGO Gene Nomenclature Committee. Available online at: https://www.genenames.org/ (accessed January 26, 2019).

112. Ortiz H, Cuadrado J, León K, Esteban M, Galán I, Bravo J, et al. Diseño del estudio ELOID y prevalencia de sobrepeso y obesidad en la población infantil de 4 años de la Comunidad de Madrid. Madrid: Informe de la Dirección General de Atención Primaria. (2014). Available online at: https://www.comunidad.madrid/sites/default/files/doc/sanidad/epid/ejoseno_eol.pdf

113. European Food Safety Authority. General principles for the collection of national food consumption data in the view of a pan-European dietary survey. EFSA J. (2009) 7:51. doi: 10.2903/J.efsa.2009.1435

114. Ortega R, López-Sobaler A, Andrés P, Requejo A, Aparicio A, Molinero L. DHA Software for Assessing Diets and Food Calculations (for Windows, version 3.3.6.0). Madrid: Departamento de Nutrición y Ciencia de los Alimentos (UCM) and Alceingeniería, S.A. (2013).

115. Serra-Majem L, Ribas L, Ngo J, Ortega RM, García A, Pérez-Rodrigo C, et al. Food, youth and the Mediterranean diet in Spain. Development of KDMDM, Mediterranean Diet Quality Index in children and adolescents. *Public Health Nutr.* (2004) 7:931–5. doi: 10.1079/PHN2004556

116. Ortega R, Requejo A, López-Sobaler A. "Modelos de cuestionario de actividad,". Nurtirgúa. Manual de nutrición clínica en atención primaria. Madrid: Complotense (2006). 468 p.

117. WHO. Physical Status: The Use and Interpretation of Anthropometry. Available online at: http://www.who.int/chldgrowth/publications/physical_status/en/ (accessed December 22, 2019).

118. Lupton JR, Brooks J, Butte NF, Caballero B, Flatt JP, Fried SK. Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids. Washington, DC: National Academy Press (2002).

119. OMRON Healthcare. Available online at: https://www.omron-healthcare.es/en/demandware.static/-/Sites-master-catalog/default/dw50f02349/pdfs/ES/IM-HBF-511B-E-ES-10-08-2017.pdf (accessed December 22, 2019).

120. WHO. Growth Reference Data for 5-19 Years. Available online at: http://www.who.int/growthref/en/ (accessed December 22, 2019).

121. Stewart A, Marfell-Jones M. International Society for Advancement of Kinanthropometry. International Standards for Anthropometric Assessment. LowerHutt: International Society for the Advancement of Kinanthropometry (2011).

122. Frisancho AR. New norms of upper limb fat and muscle areas for OMROM Healthcare. Available online at: https://www.omron-healthcare.es/sites/default/files/documentos/12_hta.pdf

123. Bahreynian M, Qorbani M, Khaniabadi BM, Motlagh ME, Safari O, Asayesh H, et al. Association between obesity and parental weight status in children and adolescents. *J Clin Res Pediatr Endocrinol.* (2017) 9:111–7. doi: 10.4274/jcrpe.3790

124. Bel-Serrat S, Heinen MM, Mehegan J, O’Brien S, Eldin N, Murrin CM, et al. School sociodemographic characteristics and obesity in schoolchildren: does the obesity definition matter? *BMC Public Health.* (2018) 18:337. doi: 10.1186/s12889-018-5246-7

125. Wadowska L, Ulewicz N, Sobas K, Wuestew JW, Słownińska MA, Niedzwiedzka E, et al. Dairy-related dietary patterns, dietary calcium, body weight and composition: a study of obesity in polish mothers and daughters, the MODAF project. *Nutrients.* (2018). 10:980. doi: 10.3390/nu10010090

126. Koca T, Akcam M, Serdaroglu F, Dereci S. Breakfast habits, dairy product consumption, physical activity, and their associations with body mass index in children aged 6–18. *Eur J Pediatr.* (2017) 176:1251–7. doi: 10.1007/s00431-017-2976-y

127. Panjikkaran ST, Kumari K. Augmenting BMI and waist-height ratio on/demandware.static/-/Sites-master-catalog/default/dw50f02349/pdfs/ES/IM-HBF-511B-E-ES-10-08-2017.pdf (accessed January 26, 2021).

128. Ashwell M. Obesity risk: importance of the waist-to-height ratio. *Nutr Rev.* (2016) 74:506–20. doi: 10.1111/nure.12224

129. Wadolowska L, Ulewicz N, Sobas K, Wuestew JW, Słownińska MA, Niedzwiedzka E, et al. Dairy-related dietary patterns, dietary calcium, body weight and composition: a study of obesity in polish mothers and daughters, the MODAF project. *Nutrients.* (2018). 10:980. doi: 10.3390/nu10010090

130. Marginean CO, Marginean C, Melit LE. New insights regarding genetic aspects of childhood obesity: a minireview. *Front Pediatr.* (2018) 6:267. doi: 10.3389/fped.2018.00271

131. Gilvary C, Madhukar N, Elkahder J, Elemento O. The missing pieces of artificial intelligence in medicine. *Trends Pharmacol Sci.* (2019) 40:555–64. doi: 10.1016/j.tips.2019.06.001

132. Wellens Frederik J, Nathan NK, Sutherland R, Wyse RJ, et al. Strategies for enhancing the implementation of school-based policies or practices targeting risk factors for chronic disease. *Cochrane Database Syst Rev.* (2017) 11:CD011677. doi: 10.1002/14651858.CD011677.pub2

133. Dugan TM, Mukhopadhyay S, Carroll A, Downes S. Machine learning techniques for prediction of early childhood obesity. *Appl Clin Inform.* (2015) 6:506–20. doi: 10.4338/ACI-2015-03-RA-0036

134. Nishtar S, Gluckman P, Armstrong T. Ending childhood obesity: a time for action. *Lancet.* (2016) 387:825–7. doi: 10.1016/S0140-6736(16)00140-9

135. Visiedo A, de Baranda PS, Rosselló M, Perez-Llamas F. Programs to prevent obesity in school children 5 to 10 years old: a review. *Nutr Hosp.* (2016) 33:814–24. doi: 10.20960/nh.375

136. Wolfenden L, Nathan NK, Sutherland R, Yoong SL, Hodder RK, Wyse RJ, et al. Strategies for enhancing the implementation of school-based policies or practices targeting risk factors for chronic disease. *Chowr* (2017) 1:51. doi: 10.2903/j.efsa.2009.1435

137. Dugan TM, Mukhopadhyay S, Carroll A, Downes S. Machine learning techniques for prediction of early childhood obesity. *Appl Clin Inform.* (2015) 6:506–20. doi: 10.4338/ACI-2015-03-RA-0036

138. Nishtar S, Gluckman P, Armstrong T. Ending childhood obesity: a time for action. *Lancet.* (2016) 387:825–7. doi: 10.1016/S0140-6736(16)00140-9

139. Visiedo A, de Baranda PS, Rosselló M, Perez-Llamas F. Programs to prevent obesity in school children 5 to 10 years old: a review. *Nutr Hosp.* (2016) 33:814–24. doi: 10.20960/nh.375

140. Wolfenden L, Nathan NK, Sutherland R, Yoong SL, Hodder RK, Wyse RJ, et al. Strategies for enhancing the implementation of school-based policies or practices targeting risk factors for chronic disease. *Chowr Database Syst Rev.* (2017) 1:51. doi: 10.2903/j.efsa.2009.1435

141. de Toro-Martín J, Arsenault BJ, Després J-P, Vohl M-C. Precision nutrition: a review of personalized nutritional approaches for the prevention and management of metabolic syndrome. *Nutrients.* (2017) 9:913. doi: 10.3390/nu9080913

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Marcos-Pasero, Aguilar-Aguilar, de la Iglesia, Espinosa-Salinas, Molina, Colmenaroyo, Martínez, Ramírez de Molina, Reglero and Loria-Kohen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.