



OPEN

## Development of a Genetic Risk Score to predict the risk of overweight and obesity in European adolescents from the HELENA study

Miguel Seral-Cortes<sup>1✉</sup>, Sergio Sabroso-Lasa<sup>2</sup>, Pilar De Miguel-Etayo<sup>1,3,4</sup>, Marcela Gonzalez-Gross<sup>4,5,6</sup>, Eva Gesteiro<sup>6</sup>, Cristina Molina-Hidalgo<sup>7</sup>, Stefaan De Henauw<sup>8</sup>, Frederic Gottrand<sup>9</sup>, Christina Mavrogianni<sup>10</sup>, Yannis Manios<sup>10</sup>, Maria Plada<sup>11</sup>, Kurt Widhalm<sup>12</sup>, Anthony Kafatos<sup>13</sup>, Éva Erhardt<sup>14</sup>, Aline Meirhaeghe<sup>15</sup>, Diego Salazar-Tortosa<sup>16</sup>, Jonatan Ruiz<sup>17</sup>, Luis A. Moreno<sup>1,3,4</sup>, Luis Mariano Esteban<sup>18</sup> & Idoia Labayen<sup>19</sup>

Obesity is the result of interactions between genes and environmental factors. Since monogenic etiology is only known in some obesity-related genes, a genetic risk score (GRS) could be useful to determine the genetic predisposition to obesity. Therefore, the aim of our study was to build a GRS able to predict genetic predisposition to overweight and obesity in European adolescents. A total of 1069 adolescents (51.3% female), aged 11–19 years participating in the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) cross-sectional study were genotyped. The sample was divided in non-overweight (non-OW) and overweight/obesity (OW/OB). From 611 single nucleotide polymorphisms (SNP) available, a first screening of 104 SNPs univariately associated with obesity ( $p < 0.20$ ) was established selecting 21 significant SNPs ( $p < 0.05$ ) in the multivariate model. Unweighted GRS (uGRS) was calculated by summing the number of risk alleles and weighted GRS (wGRS) by multiplying the risk alleles to each estimated coefficient. The area under curve (AUC) was calculated in uGRS (0.723) and wGRS (0.734) using tenfold internal cross-validation. Both uGRS and wGRS were significantly associated with body mass index (BMI) ( $p < .001$ ). Both GRSs could potentially be considered as useful genetic tools to evaluate individual's predisposition to overweight/obesity in European adolescents.

<sup>1</sup>Growth, Exercise, NUTrition and Development (GENUD) Research Group, Instituto Agroalimentario de Aragón (IA2), Instituto de Investigación Sanitaria Aragón (IIS Aragón), Universidad de Zaragoza, Zaragoza, Spain. <sup>2</sup>Spanish National Cancer Research Centre (CNIO), Madrid, Spain. <sup>3</sup>Instituto de Investigación Sanitaria Aragón (IIS Aragón), Zaragoza, Spain. <sup>4</sup>Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y la Nutrición (CIBEROBn), Instituto de Salud Carlos III, Madrid, Spain. <sup>5</sup>Institute of Nutritional and Food Sciences, Nutritional Physiology, University of Bonn, Bonn, Germany. <sup>6</sup>ImFine Research Group, Department of Health and Human Performance, Facultad de Ciencias de la Actividad Física y del Deporte-INEF, Universidad Politécnica de Madrid, Madrid, Spain. <sup>7</sup>EFFECTS 262 Department of Medical Physiology, School of Medicine, University of Granada, 18071 Granada, Spain. <sup>8</sup>Department of Public Health, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium. <sup>9</sup>Faculty of Medicine, University Lille, Lille, France. <sup>10</sup>Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece. <sup>11</sup>University of Crete School of Medicine, Crete, Greece. <sup>12</sup>Division of Gastroenterology and Hepatology, Department of Internal Medicine III, Medical University of Vienna, Austria and Austrian Academic Institute for Clinical Nutrition, Vienna, Austria. <sup>13</sup>Faculty of Medicine, University of Crete, Crete, Greece. <sup>14</sup>Department of Pediatrics, Medical School, University of Pécs, Pécs, Hungary. <sup>15</sup>UMR1167, RID-AGE, Risk Factors and Molecular Determinants of Aging-Related Diseases, Centre Hosp. Univ Lille, Institut Pasteur de Lille, Université de Lille, Lille, France. <sup>16</sup>Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA. <sup>17</sup>Departamento de Actividad Física y Deporte, Faculty of Sport Sciences, Universidad de Granada, Granada, Spain. <sup>18</sup>Escuela Politécnica de La Almunia, Universidad de Zaragoza, Zaragoza, Spain. <sup>19</sup>Department of Health Sciences, Public University of Navarra, Pamplona, Spain. ✉email: mseral@unizar.es

Childhood obesity is a major public health problem<sup>1</sup>. Pediatric obesity increases the risk of physical and psychological health problems already in childhood, and later in adulthood<sup>2</sup>. More so, adiposity related disorders predominantly diagnosed in adults such as type 2 diabetes mellitus (T2DM) and cardiovascular diseases might originate in early life, and potentially reduce life expectancy. Over the last two decades, large-scale studies have been unveiling new common variants in locus of certain genes related with childhood and adult obesity<sup>3–5</sup>. At least 97 loci have been associated with obesity<sup>6</sup>. Currently, the *FTO* gene still remains the locus explaining the largest association with obesity in adults, children and adolescents<sup>7,8</sup>. In this regard, previous studies have shown that each copy of the *FTO* rs9939609 polymorphism A allele is associated with 2.8% higher body fat in European adolescents<sup>9,10</sup>. Some studies have found some associations between single nucleotide polymorphisms (SNPs) with obesity risk factors, being potentially useful as early life risk indicators in children and adults<sup>11</sup>. However, individual SNPs can explain little of disease variance<sup>12</sup>. Several studies have demonstrated the potential value of other genetic approaches that combine a number of SNPs to develop a genetic risk score (GRS) by summing the number of risk alleles: unweighted GRS (uGRS) or by multiplying the number of risk alleles to each estimated coefficient: weighted GRS (wGRS)<sup>13–15</sup>. The creation and validation of obesity-specific GRS sets a landmark in personalised genetic risk prediction for obesity and obesity-related diseases<sup>16</sup>. Different obesity-related GRS have been constructed in adults<sup>15–18</sup> and children<sup>19,20</sup> with significant obesity-gene associations, being implemented on a variety of ethnic population backgrounds. Within European populations, Seyednasrollah et al.<sup>21</sup> computed two weighted GRS (wGRS) of 97 and 19 SNPs previously related to the risk of obesity in two cohorts including 2262 Finnish children and adolescents (3–18 years). Further, Viljakainen et al.<sup>22</sup> developed a wGRS to predict the risk of overweight and obesity in a cohort of 1142 Finnish preadolescents (11.3 ± 0.2 years) considering body mass index (BMI) and 30 BMI-related SNPs from previous genome-wide association studies (GWAS). As only few studies testing obesity risk in European adolescents with GRSs have been conducted, the aim of the present study was to develop a GRS for overweight and obesity in adolescents participating in the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) cross-sectional study.

## Methods

**Study design and population.** The data were extracted from the HELENA multicentric and cross-sectional study containing a total sample of 4356 adolescents (51.6% females), aged 11–19 years old, from 10 European cities located in separated geographical points in Europe in 2006–2007. Their size of the cities was large enough to ensure participants diversity<sup>23</sup>. The main objective of the HELENA study was to obtain comparable data of a large sample of European adolescents on nutrition and health-related parameters by a standardised procedure<sup>24</sup>. More so, the study was performed following the ethical guidelines of the Declaration of Helsinki 1964 (revision of 2013), the Good Clinical Practice, and the legislation about clinical research in humans in each of the participating countries and was approved by the Ethics Committee of each city participating in the study<sup>25</sup>. The protocol was approved by the Ethical Committee (Comité de Ética de la Investigación de la Comunidad Autónoma de Aragón: CEICA). Written informed consent and assent to participate in the study were obtained from adolescents and their parents before being enrolled. One third of the subjects (N = 1172) from the total sample were randomly selected for blood sampling<sup>24</sup>. After including specific inclusion criteria from genomic parameters (SNPs) and anthropometry (BMI), a total of 1069 adolescents (51.3% females) were finally considered for the analysis in the present study. The flow chart of the selected sample is displayed in Supplementary Fig. 1.

**Physical examination.** All measurements were performed by trained researchers following standard protocols. Weight and height were measured following standard procedures<sup>26</sup>. BMI was calculated from height and weight (kg/m<sup>2</sup>)<sup>27</sup> and categorised into non-overweight (non-OW) and overweight, including obesity (OW/OB), according to the age- and sex-specific BMI international cut-offs proposed by the World Obesity Federation<sup>28</sup>. FMI was calculated dividing fat mass (FM) by height squared (in meters).

**Blood collection and genotyping.** The blood samples were collected in overnight fasting state. A standardised methodology for blood collection, transport and analysis was performed by a certified laboratory<sup>29</sup>. Blood for DNA extraction was collected in EDTA K3 tubes and stored at the Institute of Nutritional and Food Sciences (IEL) of the University of Bonn, and sent to the Laboratoire d'Analyse Genomique Centre de Ressources Biologiques (LAG-CRB) e BB- 0033-00071 Institut Pasteur de Lille, F-59000 Lille, France. DNA was extracted from white blood cells with the Puregene kit (QIAGEN, Courtaboeuf, France) and stored at – 20 °C. The genotyping was done by an Illumina system (Illumina, Inc, San Diego, California) using the GoldenGate technology (Sampling procedure scheme, GoldenGate; Software, Inc, San Francisco, California). In terms of gene selection within the HELENA study, a candidate gene approach was used. First, a certain number of behaviour and metabolic pathways related to the adolescence's health were identified. These included food intake, eating behaviour, food choices and preferences, energy and adipose tissue metabolism, glucose, insulin, lipid and lipoprotein metabolism among others. Finally, SNPs playing a key role in genes coding for the abovementioned pathways were selected. The HapMap database was used to select tag and independent SNPs. SNPs were selected with a minor allele frequency (MAF) above 0.1 and tag SNPs with  $r^2$  above 0.8. If tag SNPs described for a single gene exceeded in number (more than ~ 20), only SNPs significantly associated with appropriate phenotypes in previous studies were selected, if available. Lastly, SNPs from the NCBI database were used when a limited number of SNPs were available in the HapMap database.

**Statistical analysis.** Descriptive characteristics by sex are shown as median and interquartile range (IQR) for continuous variables and as absolute and relative frequency for categorical ones. The statistical tests used to

compare differences by sex were Pearson's chi-square for categorical variables and Mann–Whitney–Wilcoxon test for continuous variables. Pearson's chi-square statistic test was used to analyse the Hardy–Weinberg equilibrium. Shapiro–Wilk non-parametric test to check normality of variables was performed.

All statistical analyses were performed using RStudio Version 1.2.5001 (RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>) and the significance level was set at  $p < 0.05$ .

**Development of the genetic risk score.** Candidate gene approach was the procedure based to select the genes in the HELENA study. First, relevant adolescent's behaviours and metabolic pathways related to health were identified. Second, key proteins that, according to the literature, play a role in these pathways were also identified. Third, a selection of SNPs coding for these proteins was performed. Fourth, in order to select and tag SNPs independently, the HapMap database (2007 release) was used. SNPs with a minor allele frequency (MAF) above 0.1 and tag SNPs with  $r^2$  above 0.8 were selected. If too many tag SNPs described for a single gene (more than ~20) were identified, only SNPs significantly associated with appropriate phenotypes in previous publications were selected, if available. Finally, SNPs from the NCBI database were included when a too limited number of SNPs were available in the HapMap database. Based on the above, a total of 611 SNPs related to obesity and obesity-related phenotypes available in the HELENA dataset<sup>30</sup> were used to build a GRS considering BMI as obesity-related variable in order to predict a major predisposition of overweight/obesity in European adolescents<sup>24</sup>. Each SNP was recoded as 0, 1, or 2 depending on the number of risk alleles defined in previous literature, respectively. A further selection of SNPs was performed using generalised linear model (GLM) to establish an initial cut off point ( $p < 0.20$ ) to refine the search to 104 SNPs. Then, a step by step algorithm was applied to select the significant SNPs under the  $p < 0.05$  threshold in a multivariate model to shortlist a final number of 21 SNPs significantly associated with BMI. The correspondence between actual and predicted probabilities of this model was analysed by a calibration curve. The unweighted GRS (uGRS) was calculated by summing the number of risk alleles from the 21 SNP variants with a rescaling, considering the SNPs that appear as protector factors. The wGRS was the result of multiplying the number of risk alleles at each locus (0, 1, 2) for each estimated coefficient of the multivariate model. Participants with missing data were dismissed in the GRS analysis ( $N = 3287$ ). Receiver operating characteristics (ROC) curve analysis<sup>31</sup> was applied to test the diagnostic accuracy of the GRS to classify potential participants for obesity associated disturbances<sup>32</sup>. The area under curve (AUC) was calculated in uGRS and wGRS considering weight status as binary variable (i.e., non-OW vs. OW/OB). Selection of uGRS over wGRS to proceed with the design of the final model was performed by the higher value of the AUC compared using the DeLong test. The model was internally validated performing tenfold cross validation analysis. For this analysis, the whole dataset was divided in 10 groups, using 9 of them to build the predictive model and the one to remains to validate this model. This procedure was repeated taking into account all possible ways to select the 9 subgroups, ensuring different forms to validate GRS with data not used in the building model process. Moreover, we evaluated the distribution of uGRS and wGRS values for NON-OW and OW/OB in a boxplot to graphically analyse the performance of the GRS. In order to provide the best cut-off for the use of the GRS as a dichotomic variable, the maximisation of the Youden index<sup>33</sup> was explored (see Table 3). Lastly, to test the GRS reliability with a general adiposity estimate other than BMI, simple linear regression models (LRM) were performed to evaluate the association between fat mass index (FMI) and both wGRS and uGRS.

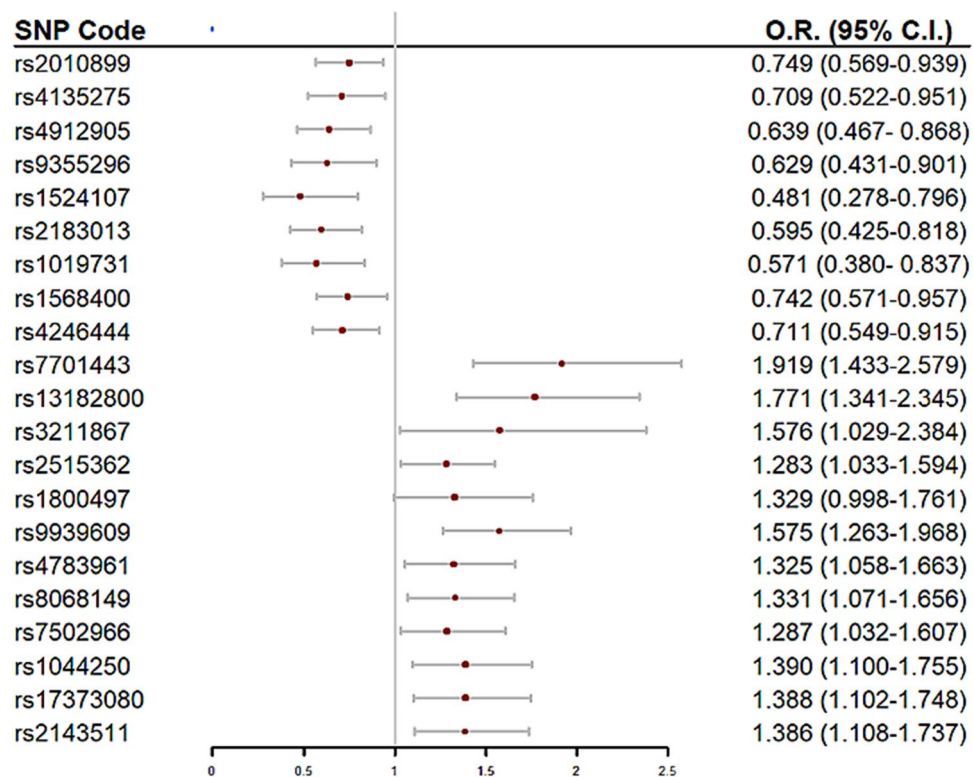
**Informed consent.** Informed consent was signed by parents of all participants. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

## Results

**Description of the study sample: demographics.** The sample was composed of 520 boys and 549 girls. Main characteristics of study participants are shown in Supplementary Table 1. The median age of participants barely differed between males (14.6 years, IQR: 13.6–15.7) and females (14.6 years, IQR: 13.5–15.7) ( $p = 0.722$ ). The prevalence of OW/OB was higher in males (22.2%), than in females (18.8%) ( $p = 0.009$ ).

**Associations between SNPs and overweight/obesity: Building and validation of GRS.** Initially, 104 SNPs potentially associated with BMI were selected (Supplementary Table 2 and were entered in the multivariate model to build the GRS. From them, we found a final number of 21 SNPs significantly associated with OW/OB in the HELENA study Table 1. Table 2 shows the univariate and multivariate model's odds ratio (OR) of each of the selected SNPs for the GRS build up. A forest plot is displayed in Fig. 1 to present the OR's direction (protective/risk) of each SNP. Within our GRS, *AMPD1* rs2010899, *PPARG* rs4135275, *NR3C1* rs4912905, *LPAr*s 9,355,296, *IL-6* rs1524107, *CNTFR* rs2183013, *IGF1* rs1019731, *THRA* rs1568400 and *FASN* rs4246444 had a protective role in the prediction of OW/OB whereas *NR3C1* rs7701443, *NR3C1* rs13182800, *CD36* rs3211867, *CNTF* rs2515362, *DRD2* rs1800497, *FTO* rs9939609, *CETP* rs4783961, *NOS2A* rs8068149, *THRA* rs7502966, *ANGPTL4* rs1044250, *LXRβ* rs17373080 and *PTPN1* rs2143511 increased the risk of OW/OB. Supplementary Fig. 2 shows the calibration curves analysing the correspondence between probabilities of overweight and the real outcome. It can be observed that there is a good agreement between predicted and actual probabilities, thus, the panel of SNPs shows a good adjustment in order to predict OW/OB.

Using the predictABEL R package<sup>34</sup>, from the multivariate logistic regression model built to predict OW/OB, an uGRS and a wGRS were derived. The predictive ability of the GRS by means of the ROC curve, AUCs and Youden index of uGRS and wGRS models are displayed in Fig. 2. The results of the GRSs' AUC (uGRS: 0.7198, wGRS: 0.7336) showed a moderate ability to discriminate OW/OB status. AUC's comparisons displayed statistically significant differences between uGRS and wGRS to predict OW/OB ( $p = 0.043$ ). The discrimination



**Figure 1.** Forest plot of single nucleotide polymorphisms (SNPs) negatively ( $OR < 1$ ) and positively ( $OR > 1$ ) associated with risk of OW/OB. Legend: SNPs ordered by chromosome number. Protective SNPs against risk of overweight/obesity are shown in the upper part of the forest plot; SNPs with predisposition risk to overweight/obesity are shown in the bottom part. Multivariate model Odds Ratio (O.R.) and 95% confidence intervals (C.I.) displayed.

ability of wGRS and uGRS was internally validated by cross-validation using 10 folds. Both GRS provide robust predictions as showed by the AUC results (0.723 and 0.734 for uGRS and wGRS respectively). The distribution of uGRS and wGRS values for the groups of NON-OW and OW/OB by boxplots is displayed in Fig. 3. Both GRS discriminate between groups, but there is no a cut-off point that can clearly separate NON-OW and OW/OB groups. The Youden index was 23.5 for uGRS (specificity 69.4%, sensitivity 63.6%), and  $-0.126$  for wGRS (specificity 61.1%, sensitivity 74.6%). A more general analysis of sensitivity, specificity, positive and negative predictive value, and accuracy is shown in Table 3.

**Associations between GRS and adiposity.** LRM showed a significant association between GRS and FMI ( $p \leq 2e-16$ ) in both weighted ( $\beta = 0.877$ ) and unweighted ( $\beta = 0.312$ ) variants for the total sample.

## Discussion

In the present study, two GRSs, uGRS and wGRS, including 21 SNPs associated with BMI, were successfully developed to assess the risk of overweight and obesity in European adolescents. Hence, to the best of our knowledge, this are the first GRSs with these characteristics in a diversely distributed sample of European adolescents.

There are few previous studies focusing on BMI-specific GRSs with overweight and obesity in European pediatric populations and none exclusively in European adolescents, which reinforces the potential of our GRS analysis. In a cohort of 1142 Finnish preadolescents, Viljakainen et al.<sup>22</sup> constructed a wGRS to predict the risk of overweight (1.39-fold increased odds) and obesity (1.41-fold increased odds) using 30 BMI-related SNPs, stating that their GRS was poor in predicting short-term longitudinal changes in BMI. In two Finnish children and adolescent cohorts, Seyednasrollah et al.<sup>21</sup> developed two wGRS of 97 and 19 SNPs previously related to the risk of obesity, and obtained a lightly better prediction accuracy with the 19-SNP GRS than in our study (AUC = 0.769 vs. 0.734). However, none of the SNPs used in the two GRSs in European adolescents above mentioned concur with the SNPs utilised to develop our GRS.

The majority of the GRSs developed in pediatric and adolescent populations have been performed in non-European subjects, based on SNPs associated with obesity risk from previous GWAS. Comparatively to the European GRSs, the non-European studies provided a fewer number of selected SNPs related to obesity risk to develop their GRS. In a cross-sectional study of Brazilian children and adolescents (mean age  $11.9 \pm 2.8$  years)<sup>19</sup>, the BMI-specific wGRS composed of 3 SNPs was associated with a 2.65-fold increased risk of overweight and obesity. In a Chinese cohort of children aged 7.3 to 11.1 years, Fang J et al.<sup>35</sup> developed an uGRS of 11 BMI-related SNPs which explained  $0.11 \text{ kg/m}^2$  BMI increase in those children carrying BMI susceptibility alleles. Lv D et al.<sup>36</sup>

rs code	Nearest gene	Alleles (major/minor)	MAF	<i>p</i>	Genotyping success rate	HWE
rs2010899	<i>AMPD1</i>	A/C	0.44	0.012	99.8	0.349
rs4135275	<i>PPARG</i>	A/G	0.19	0.024	99.9	0.655
rs4912905	<i>NR3C1</i>	C/G	0.23	0.004	100.0	0.665
rs7701443	<i>NR3C1</i>	A/G	0.40	$1.35 \times 10^{-4}$	100.0	0.127
rs13182800	<i>NR3C1</i>	C/A	0.24	$5.96 \times 10^{-4}$	99.9	0.183
rs9355296	<i>LPA</i>	G/A	0.13	0.013	100.0	0.216
rs1524107	<i>IL-6</i>	G/A	0.07	0.006	100.0	0.879
rs3211867	<i>CD36</i>	C/A	0.07	0.033	100.0	0.728
rs2183013	<i>CNTFR</i>	C/G	0.17	0.001	100.0	0.934
rs2515362	<i>CNTF</i>	A/G	0.44	0.024	99.9	0.184
rs1800497	<i>DRD2</i>	G/A	0.18	0.049	99.8	0.763
rs1019731	<i>IGF1</i>	C/A	0.11	0.005	100.0	0.744
rs9939609	<i>FTO</i>	T/A	0.40	$5.81 \times 10^{-4}$	100.0	0.322
rs4783961	<i>CETP</i>	A/G	0.50	0.014	100.0	0.558
rs8068149	<i>NOS2A</i>	G/A	0.46	0.010	100.0	0.136
rs7502966	<i>THRA</i>	A/G	0.44	0.025	99.7	0.264
rs1568400	<i>THRA</i>	A/G	0.26	0.023	100.0	0.349
rs4246444	<i>FASN</i>	C/A	0.27	0.008	94.7	0.461
rs1044250	<i>ANGPTL4</i>	G/A	0.29	0.005	99.6	0.051
rs17373080	<i>LXRβ</i>	G/C	0.32	0.005	99.7	0.523
rs2143511	<i>PTPNI</i>	A/G	0.43	0.004	99.9	0.337

**Table 1.** Main characteristics of the 21 single nucleotide polymorphisms (SNPs) included in the genetic risk score. Association of SNPs in relation to body mass index (BMI) displayed in *p* values (*p*). *AMPD1* (Adenosine Monophosphate Deaminase 1); *ANGPTL4* (Angiopoietin like protein 4); BMI (Body Mass Index); *CD36* (cluster of differentiation 36); *CETP* (Cholesteryl ester transfer protein); *CNTF* (Ciliary Neurotrophic Factor); *CNTFR* (Ciliary Neurotrophic Factor Receptor); *DRD2* (Dopamine Receptor D2); *FASN* (Fatty acid synthase); *FTO* (Fat mass and obesity-associated gene); HWE (Hardy–Weinberg Equilibrium); *IGF1* (Insulin Like Growth Factor 1); *IL-6* (Interleukin-6); *LPA* (Lipoprotein A); *LXRβ* (Liver X receptor beta); MAF (Minor allele frequency); *NOS2A* (Nitric oxide synthase 2); *NR3C1* (Nuclear receptor subfamily 3, group C, member 1); *PPARG* (Peroxisome Proliferator Activated Receptor Gamma); *PTPNI* (Protein Tyrosine Phosphatase Non-Receptor Type 1); *THRA* (Thyroid Hormone Receptor Alpha).

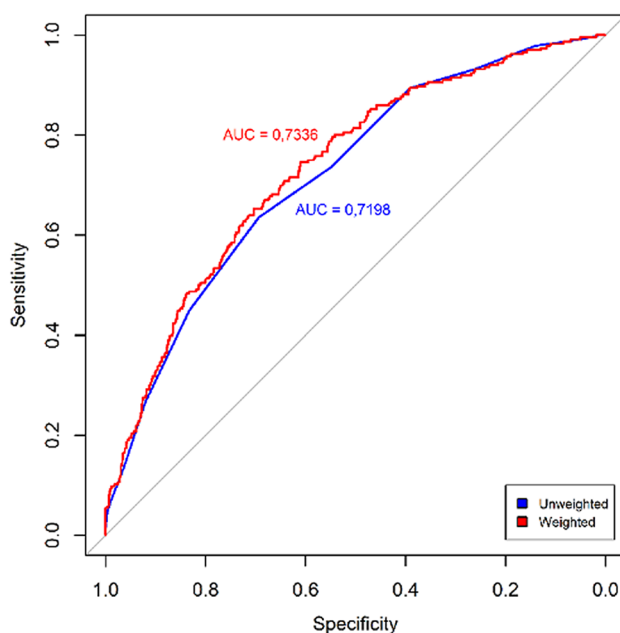
positively associated the cumulative effect of 5 BMI-related SNPs GRS to obesity risk by more than sevenfold increased odds in individuals carrying 5–7 risk alleles (age  $11.6 \pm 2.5$ ,  $N = 2977$ ). Finally, SNPs previously related to other cardiometabolic risk factors (hypertension) in a Chinese adolescent population (aged  $12.2 \pm 3.0$  years) were used to develop a 3 SNP GRS positively associated to obesity risk<sup>37</sup>. The increased risk of obesity associated to the mentioned GRSs in non-European subjects seem to have a consistency to some SNPs showed in the present study, despite acknowledging that the subjects origin does not allow to make comparisons in this regard.

The elaboration of a BMI-GRS comprised protector and risk in the same model. Some of those SNPs have been significantly associated with the risk of obesity in previous studies. The A allele of *FTO* rs9939609 polymorphism has been consistently associated with higher BMI and waist circumference in several studies in adults<sup>11</sup>, adolescents<sup>9</sup> and children<sup>38</sup>. In cohorts of children and adults with European ancestry, Frayling et al.<sup>11</sup> and Willer et al.<sup>39</sup> found the strongest associations of the A risk allele of *FTO* rs9939609 with BMI. These findings confirm the role of *FTO* rs9939609 in our GRS as risk factor to OW/OB. In a study by Bokor et al.<sup>40</sup>, *CD36* rs3211867 increased the risk of obesity by almost two folds in a cohort of Hungarian obese ( $N = 307$ ) and normal weight ( $N = 339$ ) adolescents. Although the study had two independent samples with limited sample size, the findings are consistent with the results of our study. Moreover, in a pooled-study by Solaas et al.<sup>41</sup>, authors observed significant association between *LXRβ* rs17373080 and the risk of T2DM and OW/OB by 1.59-fold increased odds. Equally, in our study, *LXRβ* rs17373080 SNP was associated with a 1.38 fold higher risk of OW/OB. Of note, the last two studies included participants from the HELENA cohort. On the other hand, the present study showed that *THRA* rs1568400 SNP was negatively associated with the risk of OW/OB whereas the same SNP was associated with higher BMI in a cohort of Spanish adults<sup>42</sup>.

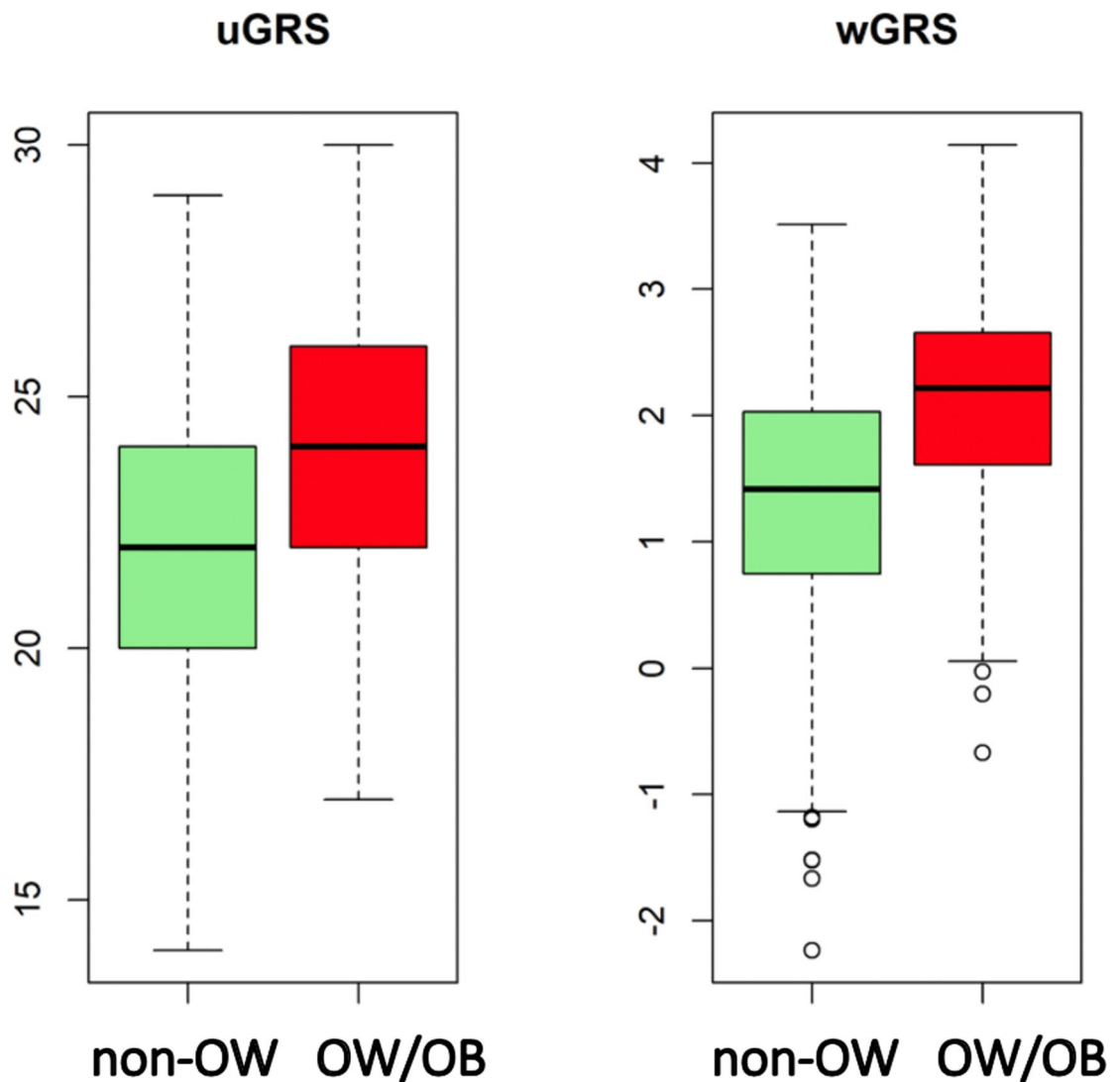
Other SNPs included in the GRS developed in the present study have also been associated with obesity-related cardiometabolic risk factors in ethnically diverse adults, but not with overweight or obesity risk. Thus, in a European population, *DRD2* rs1800497<sup>43</sup> modified the relationship between birth weight and adulthood educational attainment in Finnish subjects and *FASN* rs4246444<sup>44</sup> attenuated the effect on low density lipoproteins (LDL) peak particle diameter when consuming a high amount of fat in a Canadian cohort. Within non-European background, in Chinese population, *NR3C1* rs7701443<sup>45</sup> was significantly associated with a higher risk of metabolic syndrome and CC alleles of *IL-6* rs1524107<sup>46</sup> had a higher risk of developing nephropathy in T2DM subjects. In addition, *PPARG* rs4135275<sup>47</sup> was positively associated with glycated hemoglobin and fasting plasma glucose in Taiwanese mental health patients. In pregnant Turkish women, *LPA* rs9355296<sup>48</sup> was positively related to vascular

rs code	Chromosome	Univariate*		Multivariate*	
		O.R. (95% C.I.)	<i>p</i>	O.R. (95% C.I.)	<i>p</i>
rs2010899	1	0.822 (0.665–1.014)	0.068	0.749 (0.596–0.939)	0.012
rs4135275	3	0.738 (0.556–0.967)	0.031	0.709 (0.522–0.951)	0.012
rs4912905	5	0.810 (0.630–1.034)	0.096	0.639 (0.467–0.867)	0.024
rs9355296	6	0.636 (0.448–0.886)	0.009	0.629 (0.430–0.901)	<0.001
rs1524107	7	0.602 (0.363–0.950)	0.038	0.481 (0.278–0.795)	0.014
rs2183013	9	0.674 (0.495–0.903)	0.010	0.595 (0.425–0.818)	0.033
rs1019731	12	0.681 (0.468–0.967)	0.038	0.571 (0.379–0.837)	0.049
rs1568400	17	0.794 (0.623–1.006)	0.059	0.742 (0.571–0.957)	0.023
rs4246444	17	0.756 (0.592–0.957)	0.022	0.711 (0.548–0.914)	0.009
rs7701443	5	1.170 (0.946–1.447)	0.146	1.919 (1.432–2.579)	0.005
rs13182800	5	1.374 (1.094–1.721)	0.006	1.771 (1.340–2.344)	<0.001
rs3211867	7	1.487 (1.009–2.159)	0.040	1.576 (1.029–2.384)	0.006
rs2515362	11	1.189 (0.973–1.452)	0.090	1.283 (1.032–1.594)	0.002
rs1800497	11	1.204 (0.924–1.559)	0.162	1.329 (0.998–1.761)	0.024
rs9939609	16	1.496 (1.219–1.838)	<0.001	1.575 (1.263–1.968)	0.005
rs4783961	16	1.272 (1.034–1.568)	0.023	1.325 (1.058–1.663)	0.014
rs8068149	17	1.190 (0.974–1.456)	0.088	1.331 (1.071–1.656)	0.010
rs7502966	17	1.165 (0.952–1.427)	0.138	1.287 (1.031–1.607)	0.025
rs1044250	19	1.211 (0.976–1.500)	0.079	1.390 (1.100–1.754)	0.006
rs17373080	19	1.313 (1.061–1.624)	0.012	1.388 (1.102–1.748)	0.005
rs2143511	20	1.294 (1.056–1.588)	0.013	1.386 (1.108–1.736)	0.004

**Table 2.** Association of the 21 single nucleotide polymorphisms (SNPs) of the genetic risk score with OW/OB. SNPs ordered by chromosome number and protector/risk nature of the SNP. \*Univariate (individual model of SNP-Overweight/Obesity association) and multivariate model (multiple model SNP-Overweight/Obesity association) displayed with Odds Ratio (O.R.) and 95% confidence interval (C.I.).



**Figure 2.** Receiver operating characteristics (ROC) curves of the two genetic risk scores, unweighted (uGRS) and weighted (wGRS). Areas under curves (AUC) are indicated. The straight line represents the ROC expected by chance only.



**Figure 3.** Boxplot of the distribution of unweighted genetic risk score (uGRS) and weighted genetic risk score (wGRS). Legend: Values for the groups of Non-overweight (NON-OW) and Overweight/Obesity (OW/OB).

inflammation as future cardiovascular event indicator. Finally, in a large adult cohort of 5 different ancestry groups, *CETP* rs4783961<sup>49</sup> was involved in sleep-associated adverse high density lipoproteins (HDL) profiles.

In contrast, as far as we know, several SNPs included our GRSs (i.e., *AMPD1* rs2010899, *NR3C1* rs4912905, *CNTFR* rs2183013, *IGF1* rs1019731 as protector factors and *NR3C1* rs13182800, *CNTF* rs2515362, *NOS2A* rs8068149, *THRA* rs7502966, *ANGPTL4* rs1044250 and *PTPN1* rs2143511 as risk factors) are new predictive factors, as they had not previously been associated with obesity or obesity related diseases nor had been significantly relevant in previous studies.

Additionally, the present GRS was positively tested to evaluate its ability to predict the risk of overweight and obesity in other adiposity estimates (FMI). Previous studies have also identified potential interactions between an obesity-GRS and diet on FMI in English children (9yrs)<sup>50</sup>. Monnereau et al.<sup>51</sup> constructed 15 SNPs-wGRS related to child BMI in children (6yrs) from Netherlands significantly associated to total fat mass. More so, another study<sup>52</sup> showed a BMI-based GRS significantly associated to higher body fat mass in Finnish children and adolescents.

Although using the external weight from meta-analyses is the gold standard to build a GRS, when the external weights are not available, the uGRS is commonly used<sup>35,53</sup>. In the present approach, internal weights from the genetic effects of the same study were used. The wGRS outperformed the uGRS in terms of statistical power (0.734 vs. 0.723). Conventionally, it is accepted that the AUC in a ROC analysis should be > 0.8 to be of clinical value for screening<sup>54</sup>. When constructing the GRS models, AUC fell short of this threshold combining genetic factors alone. As SNPs themselves have little predictor capacity, we should consider the results obtained to construct the uGRS and wGRS as statistically acceptable, so our wGRS could be replicated in other cohorts with similar characteristics. Thus, our findings add a significant contribution to obesity-specific GRS that may improve the predictive values of obesity biomarkers in adolescents.

Cut-off	Spec	Sens	NPV	PPV	Acc	Cut-off	Spec	Sens	NPV	PPV	Acc
uGRS						wGRS					
30	100	0.42	77.99	100	78.01	4.14	100	0.42	77.99	100	78.01
29	99.88	3.81	78.56	90.00	78.67	3.51	99.88	5.50	78.86	92.85	79.04
28	99.15	6.35	78.89	68.18	78.67	3.30	99.15	8.89	79.34	75.00	79.23
27	96.39	13.55	79.74	51.61	78.11	2.98	96.63	15.25	80.09	56.25	78.67
26	92.07	26.69	81.59	48.83	77.64	2.61	92.07	27.54	81.76	49.61	77.82
25	83.31	44.91	84.22	43.26	74.83	2.26	83.31	48.72	85.15	45.27	75.67
24	69.38	63.55	87.04	37.03	68.10	1.88	69.38	65.25	87.57	37.65	68.47
23	54.74	73.72	88.03	31.57	58.93	1.50	54.74	79.23	90.29	33.15	60.14
22	39.13	89.40	92.87	29.38	50.23	1.10	39.13	89.40	92.87	29.38	50.23
21	26.17	93.22	93.16	26.34	40.97	0.77	26.17	93.22	93.16	26.34	40.97
20	14.04	97.88	95.90	24.39	32.55	0.35	14.04	97.03	94.35	24.23	32.36
19	7.92	98.72	95.65	23.30	27.97	0.03	7.92	98.72	95.65	23.30	27.97
18	2.88	99.57	96.00	22.50	24.22	-0.54	2.88	99.57	96.00	22.50	24.22
17	1.32	100	100	22.30	23.10	-0.82	1.32	100	100	22.30	23.10
16	0.48	100	100	22.15	22.45	-1.20	0.48	100	100	22.15	22.45
15	0.00	100	100	22.09	22.17	-1.67	0.12	100	100	22.09	22.17

**Table 3.** Specificity, sensitivity, negative and positive predictive value and accuracy analysis presented in percentages (%). Abbreviations: uGRS (Unweighted Genetic Risk Score); wGRS (Weighted Genetic Risk Score); Acc (accuracy); NPV (negative predictive value); PPV (positive predictive value); Sens (sensitivity); Spec (specificity).

Other authors<sup>55</sup> suggest that traditional predictors, such as family history and childhood obesity have stronger predictive power than models based on the established genetic variants. Nonetheless, the limited predictive ability of genetic variants does not undervalue the role of gene discovery for obesity as, based on the literature, genetic analyses have already provided with promising insights involving BMI regulation<sup>6</sup>. As such, the present GRSs, or future GRS comprising additional SNPs from other genes which do not have (a priori) previous reported associations with obesity, could yield promising results to minimise the risk of cardiovascular events related to obesity.

However, the present study has some limitations. The results should be validated in larger pediatric study populations, also using obesity incidence, in order to test the reliability of this obesity-specific GRS in other populations with similar ethnicity. Despite that some studies on children reporting the prediction of adulthood obesity efficiently 20 to 30 years later<sup>21</sup>, different genetic factors might affect the short-term changes in BMI, especially during periods of rapid growth<sup>56</sup>. Due to the cross-sectional nature of the study, no cause-effect relationship can be determined. Additionally, although the model to develop the GRSs was internally validated performing tenfold cross validation analysis, we understand that the optimal situation would have been an external validation in an independent cohort. Moreover, only selected risk loci are available in the HELENA study. Since the established common variants from GWAS explain a small proportion of the BMI variation<sup>6</sup>, it is likely that other loci from rarer variants, still to be discovered, will emerge when larger sample sizes are included in GWAS. Furthermore, there is no data available regarding the relatedness or the ethnic origins among the studied participants; the allele frequencies and their effect size might be different from non-European populations and the outcome should not be reproduced to other ethnicities. Since the HELENA study selected genes based on candidate genes instead of GWAS, the overlapping effect observed between SNPs of European and non-European adolescents could be possible. More common SNPs in non-European GRSs were found than in European GRSs. This finding could be due to the higher number of GRSs developed in other ethnicities in comparison to the number of GRS performed in European population. Also, we used the same data in the SNPs selection process and in the building model, thus little bias can be produced. Therefore, the results showed in the present study should be considered carefully. Further studies with larger sample size could provide key information of this potential genetic predisposition to obesity. On the other hand, the present study has also some strengths. The multicentric design of HELENA study involved the participation of adolescents from 10 European cities. This allowed the researchers to use a large database with relevant and diverse information from different populations across Europe. Additionally, only few GRSs to predict the overweight and obesity risk have been developed particularly in European adolescents, an understudied population from the early treatment and prevention perspective<sup>21,22</sup>. Similarly to Viljakainen et al.<sup>22</sup>, the proposed genetic score of predisposition to obesity defined in this study might efficiently contribute to discern population at risk for overweight and obesity and not just obesity alone.

In conclusion, our findings suggest that the GRSs developed in the present study (uGRS and wGRS) could be considered as a useful genetic tool to evaluate individual's predisposition to OW/OB, allowing to advance in the prevention and management of the disease from early stages in life. Future GRS development with larger samples will be able to detect new variants previously not related to obesity that could influence the genetic risk of obesity, other than the common ones.



Received: 21 July 2020; Accepted: 22 January 2021

Published online: 04 February 2021

## References

1. WHO | Childhood overweight and obesity. WHO. <https://www.org/entity/dietphysicalactivity/childhood/en/index.html> (2017).
2. Gurnani, M., Birken, C. & Hamilton, J. Childhood obesity: causes, consequences, and management. *Pediatr Clin North Am* **62**, 821–840. <https://doi.org/10.1016/j.pcl.2015.04.001> (2015).
3. Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* **42**, 937–948. <https://doi.org/10.1038/ng.686> (2010).
4. Bradfield, J. P. *et al.* A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet* **44**, 526–531. <https://doi.org/10.1038/ng.2247> (2012).
5. Meyre, D. *et al.* Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat Genet* **41**, 157–159. <https://doi.org/10.1038/ng.301> (2009).
6. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–206. <https://doi.org/10.1038/nature14177> (2015).
7. Loos, R. J. & Yeo, G. S. The bigger picture of FTO: the first GWAS-identified obesity gene. *Nat. Rev. Endocrinol.* **10**, 51–61. <https://doi.org/10.1038/nrendo.2013.227> (2014).
8. Kilpeläinen, T. O. *et al.* Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med.* <https://doi.org/10.1371/journal.pmed.1001116> (2011).
9. Labayen, I. *et al.* Dietary fat intake modifies the influence of the FTO rs9939609 polymorphism on adiposity in adolescents: the HELENA cross-sectional study. *Nutr. Metab. Cardiovasc. Dis.* **26**, 937–943. <https://doi.org/10.1016/j.numecd.2016.07.010> (2016).
10. Ruiz, J. R. *et al.* Attenuation of the effect of the FTO rs9939609 Polymorphism on total and central body fat by physical activity in adolescents: the HELENA study. *Arch. Pediatrics Adolesc. Med.* <https://doi.org/10.1001/archpediatrics.2010.29> (2010).
11. Frayling, T. M. *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894. <https://doi.org/10.1126/science.1141634> (2007).
12. Wang, S., He, S., Yuan, F. & Zhu, X. Tagging SNP-set selection with maximum information based on linkage disequilibrium structure in genome-wide association studies. *Bioinformatics* **33**(14), 2078–2081. <https://doi.org/10.1093/bioinformatics/btx151> (2017).
13. Kathiresan, S. *et al.* Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med* **358**, 1240–1249. <https://doi.org/10.1056/NEJMoa0706728> (2008).
14. Janssens, A. C. *et al.* Predictive testing for complex diseases using multiple genes: fact or fiction?. *Genet Med* **8**, 395–400. <https://doi.org/10.1097/01.gim.0000229689.18263.f4> (2006).
15. Morrison, A. C. *et al.* Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. *Am J Epidemiol* **166**, 28–35. <https://doi.org/10.1093/aje/kwm060> (2007).
16. Corella, D. *et al.* Polymorphism of the Transcription Factor 7-Like 2 Gene (TCF7L2) interacts with obesity on Type-2 diabetes in the PREDIMED study emphasizing the heterogeneity of genetic variants in type-2 diabetes risk prediction: time for obesity-specific genetic risk scores. *Nutrients* <https://doi.org/10.3390/nu8120793> (2016).
17. Belsky, D. W. *et al.* Development and evaluation of a genetic risk score for obesity. *Biodemogr. Soc. Biol.* **59**, 85–100. <https://doi.org/10.1080/19485565.2013.774628> (2013).
18. Goumidi, L. *et al.* Effects of established BMI-associated loci on obesity-related traits in a French representative population sample. *BMC Genet.* <https://doi.org/10.1186/1471-2156-15-62> (2014).
19. Todendi, P. F. *et al.* Genetic risk score based on fat mass and obesity-associated, transmembrane protein 18 and fibronectin type III domain containing 5 polymorphisms is associated with anthropometric characteristics in South Brazilian children and adolescents. *Br. J. Nutr.* **121**, 93–99. <https://doi.org/10.1017/s0007114518002738> (2019).
20. Zhao, H., Wilkinson, A., Shen, J., Wu, X. & Chow, W. H. Genetic polymorphisms in genes related to risk-taking behaviours predicting body mass index trajectory among Mexican American adolescents. *Pediatr. Obes.* **12**, 356–362. <https://doi.org/10.1111/ijpo.12151> (2017).
21. Seyednasrollah, F. *et al.* Prediction of adulthood obesity using genetic and childhood clinical risk factors in the cardiovascular risk in young Finns study. *Circ. Cardiovasc. Genet.* <https://doi.org/10.1161/circgenetics.116.001554> (2017).
22. Viljakainen, H. *et al.* Genetic risk score predicts risk for overweight and obesity in Finnish preadolescents. *Clin. Obes.* **9**, e12342. <https://doi.org/10.1111/cob.12342> (2019).
23. Moreno, L. A. *et al.* Nutrition and lifestyle in European adolescents: the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) study. *Adv. Nutr.* **5**, 615s–623s. <https://doi.org/10.3945/an.113.005678> (2014).
24. Moreno, L. A. *et al.* Design and implementation of the Healthy Lifestyle in Europe by Nutrition in Adolescence Cross-Sectional Study. *Int. J. Obes. (Lond.)* **32**(Suppl 5), S4–11. <https://doi.org/10.1038/ijo.2008.177> (2008).
25. Beghin, L. *et al.* Quality assurance of ethical issues and regulatory aspects relating to good clinical practices in the HELENA Cross-Sectional Study. *Int. J. Obes. (Lond.)* **32**(Suppl 5), S12–18. <https://doi.org/10.1038/ijo.2008.179> (2008).
26. Nagy, E. *et al.* Harmonization process and reliability assessment of anthropometric measurements in a multicenter study in adolescents. *Int. J. Obes. (Lond.)* **32**(Suppl 5), S58–65. <https://doi.org/10.1038/ijo.2008.184> (2008).
27. Cole, T. J., Bellizzi, M. C., Flegal, K. M. & Dietz, W. H. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* **320**, 1240–1243. <https://doi.org/10.1136/bmj.320.7244.1240> (2000).
28. Cole, T. J. & Lobstein, T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatr. Obes.* **7**, 284–294. <https://doi.org/10.1111/j.2047-6310.2012.00064.x> (2012).
29. Gonzalez-Gross, M. *et al.* Sampling and processing of fresh blood samples within a European multicenter nutritional study: evaluation of biomarker stability during transport and storage. *Int. J. Obes. (Lond.)* **32**(Suppl 5), S66–S75. <https://doi.org/10.1038/ijo.2008.185> (2008).
30. Goumidi, L. *et al.* Healthy lifestyle by nutrition in adolescence (HELENA). A new EU funded project. *Therapie* <https://doi.org/10.2515/therapie:2007050> (2007).
31. Hanley, J. A. & McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29–36. <https://doi.org/10.1148/radiology.143.1.7063747> (1982).
32. Carayol, J., Tores, F., König, I. R., Hager, J. & Ziegler, A. Evaluating diagnostic accuracy of genetic profiles in affected offspring families. *Stat. Med.* **29**, 2359–2368. <https://doi.org/10.1002/sim.4006> (2010).
33. Liu, X. Classification accuracy and cut point selection. *Stat. Med.* **31**, 2676–2686. <https://doi.org/10.1002/sim.4509> (2012).
34. Kundu, S., Aulchenko, Y. S., van Duijn, C. M. & Janssens, A. C. PredictABEL: an R package for the assessment of risk prediction models. *Eur. J. Epidemiol.* **26**, 261–264. <https://doi.org/10.1007/s10654-011-9567-4> (2011).
35. Fang, J. *et al.* Polygenic risk, adherence to a healthy lifestyle, and childhood obesity. *Pediatr. Obes.* **14**, e12489. <https://doi.org/10.1111/ijpo.12489> (2019).
36. Lv, D. *et al.* Genetic variations in SEC16B, MC4R, MAP2K5 and KCTD15 were associated with childhood obesity and interacted with dietary behaviors in Chinese school-age population. *Gene* **560**, 149–155. <https://doi.org/10.1016/j.gene.2015.01.054> (2015).

37. Fu, L. *et al.* Gene–gene interactions and associations of six hypertension related single nucleotide polymorphisms with obesity risk in a Chinese children population. *Gene* **679**, 320–327. <https://doi.org/10.1016/j.gene.2018.09.019> (2018).
38. Lauria, F. *et al.* Prospective analysis of the association of a common variant of FTO (rs9939609) with adiposity in children: results of the IDEFICS study. *PLoS ONE* **7**, e48876. <https://doi.org/10.1371/journal.pone.0048876> (2012).
39. Willer, C. J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* **41**, 25–34. <https://doi.org/10.1038/ng.287> (2009).
40. Bokor, S. *et al.* Single-nucleotide polymorphism of CD36 locus and obesity in European adolescents. *Obesity (Silver Spring)* **18**, 1398–1403. <https://doi.org/10.1038/oby.2009.412> (2010).
41. Solaas, K. *et al.* Suggestive evidence of associations between liver X receptor beta polymorphisms with type 2 diabetes mellitus and obesity in three cohort studies: HUNT2 (Norway), MONICA (France) and HELENA (Europe). *BMC Med. Genet.* **11**, 144. <https://doi.org/10.1186/1471-2350-11-144> (2010).
42. Fernandez-Real, J. M. *et al.* Thyroid hormone receptor alpha gene variants increase the risk of developing obesity and show gene–diet interactions. *Int. J. Obes. (Lond.)* **37**, 1499–1505. <https://doi.org/10.1038/ijo.2013.11> (2013).
43. Keltikangas-Jarvinen, L. *et al.* Dopamine receptor D2 gene Taq1A (C32806T) polymorphism modifies the relationship between birth weight and educational attainment in adulthood: 21-year follow-up of the Cardiovascular Risk in Young Finns study. *Pediatrics* **120**, 756–761. <https://doi.org/10.1542/peds.2007-0073> (2007).
44. Dolley, G. *et al.* Interactions between dietary fat intake and FASN genetic variation influence LDL peak particle diameter. *J. Nutrigenet. Nutrigenomics* **4**, 137–145. <https://doi.org/10.1159/000327778> (2011).
45. Yan, Y. X. *et al.* Polymorphisms in NR3C1 gene associated with risk of metabolic syndrome in a Chinese population. *Endocrine* **47**, 740–748. <https://doi.org/10.1007/s12020-014-0324-9> (2014).
46. Chang, W. T. *et al.* Interleukin-6 gene polymorphisms correlate with the progression of nephropathy in Chinese patients with type 2 diabetes: a prospective cohort study. *Diabetes Res. Clin. Pract.* **120**, 15–23. <https://doi.org/10.1016/j.diabres.2016.07.013> (2016).
47. Liu, Y. R. *et al.* Association of the PPAR- $\gamma$  gene with altered glucose levels and psychosis profile in schizophrenia patients exposed to antipsychotics. *Psychiatry Investig.* **11**, 179–185. <https://doi.org/10.4306/pi.2014.11.2.179> (2014).
48. Tuten, A. *et al.* Relationship between LPA SNPs and inflammatory burden in patients with preeclampsia to address future cardiovascular risk. *J. Matern. Fetal Neonatal. Med.* <https://doi.org/10.1080/14767058.2019.1622667> (2019).
49. Noordam, R. *et al.* Multi-ancestry sleep-by-SNP interaction analysis in 126,926 individuals reveals lipid loci stratified by sleep duration. *Nat. Commun.* **10**, 5121. <https://doi.org/10.1038/s41467-019-12958-0> (2019).
50. Riedel, C. *et al.* Interactions of genetic and environmental risk factors with respect to body fat mass in children: Results from the ALSPAC study. *Obesity* <https://doi.org/10.1002/oby.20196> (2013).
51. Monnereau, C. *et al.* Associations of genetic risk scores based on adult adiposity pathways with childhood growth and adiposity measures. *BMC Genet.* <https://doi.org/10.1186/s12863-016-0425-y> (2016).
52. Viitasalo, A. *et al.* Genetic predisposition to higher body fat yet lower cardiometabolic risk in children and adolescents. *Int. J. Obes.* <https://doi.org/10.1038/s41366-019-0414-0> (2019).
53. Che, R. & Motsinger-Reif, A. A. Evaluation of genetic risk score models in the presence of interaction and linkage disequilibrium. *Front. Genet.* **4**, 138. <https://doi.org/10.3389/fgene.2013.00138> (2013).
54. English, P. A. *et al.* A case for the use of receiver operating characteristic analysis of potential clinical efficacy biomarkers in advanced renal cell carcinoma. *Future Oncol.* **12**, 175–182. <https://doi.org/10.2217/fon.15.290> (2016).
55. Loos, R. J. F. & Janssens, A. Predicting polygenic obesity using genetic information. *Cell Metab.* **25**, 535–543. <https://doi.org/10.1016/j.cmet.2017.02.013> (2017).
56. Warrington, N. M. *et al.* A genome-wide association study of body mass index across early life and childhood. *Int. J. Epidemiol.* **44**, 700–712. <https://doi.org/10.1093/ije/dyv077> (2015).

## Acknowledgements

This work was part of the HELENA Study (<http://www.helenastudy.com/>). We gratefully acknowledge financial support of the European Community Sixth RTD Framework Programme (contract FOOD-CT-2005-007034). The data for this study were gathered under the auspices of the HELENA project, and further analysis was additionally supported by the Spanish Ministry of Economy and Competitiveness (Grants RYC-2010-05957 and RYC-2011-09011), the Instituto de Salud Carlos III, Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición (CIBEROBn). Miguel Seral-Cortes, the corresponding author, has received funding from the Iberus Talent Pre-doctoral fellowships 2018, under the European Union's H2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No 801586. Pilar De Miguel-Etayo was supported by ISCIII – CB15/00043 (Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición, CIBEROBn). The study was performed following the ethical guidelines of the Declaration of Helsinki 1964 (revision of 2013), the Good Clinical Practice, and the legislation about clinical research in humans in each of the participating countries<sup>25</sup>. We are grateful for the support provided by school boards, headmasters, teachers, school staff and communities, and for the effort of all study nurses and our data managers.

## Author contributions

L.M. was coordinator of the HELENA project, and together with F.G., M.G.G. and S.H. formed the Core Management Group of the project. M.G.G. was responsible for the blood sampling and analysis procedure. P.D.M.-E. and I.L. are co-supervisor of MS. M.S., I.L., L.M.E. and S.S. contributed to the GRS analysis and development. I.L. supervised the genomic data. L.M., I.L., P.D.M.-E., L.M.E. and S.S. did editing of the first draft. E.G., A.M., M.G.G., C.M.H., S.H., F.G., C.M., Y.M., M.P., K.W., A.K., E.E., D.S.T. and J.R. provided critical comments on the paper, draft, and analysis. All authors read the draft and agreed on the final version.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-82712-4>.

**Correspondence** and requests for materials should be addressed to M.S.-C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021