






Development of a prediction protocol for the screening of metabolic associated fatty liver disease in children with overweight or obesity

Maddi Oses^{1,2,3}  | Cristina Cadenas-Sanchez^{1,2,3}  | María Medrano^{1,2,3}  |
 Arkaitz Galbete⁴  | Emiliano Miranda-Ferrua^{1,2,3} | Jonatan R. Ruiz^{5,6} |
 Felix Sánchez-Valverde⁷ | Francisco B. Ortega⁵ | Rafael Cabeza⁸ |
 Arantxa Villanueva^{2,8}  | Fernando Idoate^{3,9} | Idoia Labayen^{1,2,3}

¹Institute for Sustainability & Food Chain Innovation (ISFOOD), University of Navarra, Pamplona, Spain

²Navarra Institute for Health Research, IIdISA, Pamplona, Spain

³Department of Health Sciences, Public University of Navarra, Campus de Arrosadía, Pamplona, Spain

⁴Navarrabiomed-Hospital Complex of Navarra and Public University of Navarra IIdISA, REDISSEC, Pamplona, Spain

⁵PROMoting FITness and Health through Physical Activity Research Group (PROFITH), Sport and Health University Research Institute (iMUDS), Department of Physical and Sports Education, School of Sports Science, University of Granada, Granada, Spain

⁶Instituto de Investigación Biosanitaria, IBS, Granada, Granada, Spain

⁷Gastroenterology and Paediatric Nutrition Unit, Virgen del Camino Hospital, Pamplona, Spain

⁸Department of Electrical, Electronic and Communications Engineering, Institute of smart cities (ISC), Public University of Navarre, Pamplona, Spain

⁹Department of Radiology, Mutua Navarra, Pamplona, Spain

Correspondence

Maddi Oses, Institute for Sustainability & Food Chain Innovation (ISFOOD), Public University of Navarra, Calle Tajonar 22, Jerónimo de Ayanz Research Building, Pamplona, Spain.
 Email: maddi.oses@unavarra.es

Summary

Background: The early detection and management of children with metabolic associated fatty liver disease (MAFLD) is challenging.

Objective: To develop a non-invasive and accurate prediction protocol for the identification of MAFLD among children with overweight/obesity candidates to confirmatory diagnosis.

Methods: A total of 115 children aged 8–12 years with overweight/obesity, recruited at a primary care, were enrolled in this cross-sectional study. The external validation was performed using a cohort of children with overweight/obesity ($N = 46$) aged 8.5–14.0 years. MAFLD ($\geq 5.5\%$ hepatic fat) was diagnosed by magnetic resonance imaging (MRI). Fasting blood biochemical parameters were measured, and 25 candidates' single nucleotide polymorphisms (SNPs) were determined. Variables potentially associated with the presence of MAFLD were included in a multivariate logistic regression.

Results: Children with MAFLD (36%) showed higher plasma triglycerides (TG), insulin, homeostasis model assessment of insulin resistance (HOMA-IR), alanine aminotransferase (ALT), aspartate transaminase (AST), glutamyl-transferase (GGT) and ferritin ($p < 0.05$). The distribution of the risk-alleles of *PPARGrs13081389*, *PPARGrs1801282*, *HFErs1800562* and *PNLPLA3rs4823173* was significantly different between children with and without MAFLD ($p < 0.05$). Three biochemical- and/or SNPs-based predictive models were developed, showing strong discriminatory capacity (AUC-ROC: 0.708–0.888) but limited diagnostic performance (sensitivity 67%–82% and specificity 63%–69%). A prediction protocol with elevated sensitivity (72%) and specificity (84%) based on two

Abbreviations: ALT, Alanine aminotransferase; AST, Aspartate transaminase; AUC-ROC, Area under the receiver-operating characteristic curve; BMI, Body mass index; GGT, Glutamyl transferase; HOMA-IR, Homeostasis model assessment of insulin resistance; MAFLD, Metabolic associated fatty liver disease; MRI, Magnetic resonance imaging; SNP, Single nucleotide polymorphisms; SPSS, Statistical package for social sciences; SSB, Sugar-sweetened beverage; T2D, Type 2 diabetes mellitus; TG, Triglycerides; WHtR, Waist to height ratio.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Pediatric Obesity* published by John Wiley & Sons Ltd on behalf of World Obesity Federation.

Funding information

Economic Development of Government of Navarra, Grant/Award Numbers: 0011-1365-2019-000000, 0011-1365-2020-00140; Fondos Estructurales de la Unión Europea (FEDER); Instituto de Salud Carlos III, Grant/Award Number: PI13/01335; Ministerio de Economía, Industria y Competitividad, Gobierno de España, Grant/Award Number: DEP2016-78377-R; Secretaría de Estado de Investigación, Desarrollo e Innovación, Grant/Award Number: BES-2017-080770

consecutive steps was developed. The external validation showed similar results: sensitivity of 70% and specificity of 85%.

Conclusions: The HEPAKID prediction protocol is an accurate, easy to implant, minimally invasive and low economic cost tool useful for the early identification and management of paediatric MAFLD in primary care.

KEYWORDS

fatty liver, metabolic diseases, Paediatric obesity, primary health care

1 | INTRODUCTION

Metabolic associated fatty liver disease (MAFLD) has become a global health burden with an increasingly prevalence in paediatric population.¹ The term MAFLD has been recently agreed between expert groups in order to reflect more accurately the current knowledge of fatty liver diseases associated with metabolic dysfunction.^{2,3} The definition of paediatric MAFLD is based on evidence of intrahepatic fat accumulation in addition to one of the three criteria: excess adiposity, presence of pre-diabetes or type 2 diabetes or evidence of metabolic dysregulation (presence of at least two metabolic risks according to sex and age percentiles increased waist circumference, hypertension, hypertriglyceridemia, low serum HDL or impaired fasting glucose).³

MAFLD is considered a major risk factor for T2D and cardiovascular diseases, already in childhood.⁴ It is estimated that MAFLD is present in 3%–10% of general paediatric population, and this can increase to 80% in children with overweight/obesity.⁵ MAFLD can progress to steatohepatitis, fibrosis and cirrhosis over time, and it is one of the most common chronic liver diseases in the world that increase liver- and non-liver-related mortality.⁶

Lifestyle-based treatments are effective in reducing hepatic fat in children with overweight.⁷ Therefore, early detection and management of children with MAFLD is the most important step to prevent the progression of the disease.^{7,8} Unfortunately, MAFLD in childhood is often asymptomatic which, together with its high prevalence, long-term health risks and costly and/or invasive diagnostic methods, set up a challenge to clinicians and scientists for developing early diagnosis methods.⁹ Nowadays, the most widely used screening test in paediatric units is based on elevated alanine aminotransferase (ALT) levels,^{10–12} but, in children, this blood test shows very low sensitivity. Of note is that MAFLD may be present even with normal blood ALT results, leaving many children without further screening and clinical supervision.¹³ Thus, at present, available screening methods for MAFLD in children have two major disadvantages: (1) those algorithms or ALT cut-points with elevated specificity^{10,11,14,15} have very low sensitivity (<50%) and (2) the need of biochemical analysis in every child with overweight or obesity, which is certainly a large amount of blood testing in children that very often are apparently healthy. Therefore, the development of non-invasive, sensitive and accurate screening methods is of clinical interest.

Our group recently developed a simple, non-invasive, inexpensive and easy-to-perform pre-screening tool (the HEPAKID index) to

identify MAFLD among preadolescent children with overweight/obesity (i.e., children at risk of MAFLD).¹⁶ The HEPAKID index does not require blood sampling; it is based on the recording of sociodemographic factors (ethnic minority status and gestational age at birth), anthropometric data (waist circumference and height) and lifestyle variables (sugar-sweetened beverage [SSB] consumption and screen time) and shows high sensitivity (82%) to identify children who should be referred for additional diagnostic tests, being appropriate for a pre-screening method.¹⁶ However, a second-step screening tool should improve its limited specificity (63%) before conducting very invasive (biopsy) or costly (magnetic resonance imaging [MRI]) confirmatory diagnosis.

Several clinical biomarkers such as elevated levels of cholesterol, triglycerides (TG), aspartate aminotransferase (AST) and ALT, as well as hypertension or insulin resistance, are associated with MAFLD.^{17–19} In addition, there is evidence that MAFLD is strongly associated with excess adiposity.¹⁸ Yet, MAFLD and obesity are not concomitant, and not every child with overweight/obesity develops the disease. Different ethnic groups display differences in MAFLD prevalence, indicating that genetics plays a role.^{20,21} In this line, several genetic polymorphisms have shown to confer higher susceptibility to MAFLD in children and adults.²²

The present study aims to develop a second-step screening tool gathering together elevated sensitivity (>80%) and specificity (>80%) with high predictive potential for identifying children at a high risk of MAFLD. To accomplish with this objective, the current work extends the search from sociodemographic, lifestyle and anthropometric data used in the HEPAKID index¹⁶ to biochemical and genetic variables potentially associated with MAFLD. Further, the present study seeks to develop a decision tree for the identification of children with overweight/obesity candidates to confirmatory diagnosis that can be useful in clinical practice.

2 | METHODS**2.1 | Study design and participants**

This cross-sectional study uses baseline data from the EFIGRO project ([ClinicalTrials.gov](https://clinicaltrials.gov) ID: NCT02258126) whose overall aim was to examine the effect of combined family-based lifestyle plus exercise program on hepatic fat in children with overweight or obesity. Details of

sample calculation, randomization, characteristics of the study participants, methodological procedures and the measurements taken are available elsewhere.²³ The study protocol was approved by the Ethic Committee of Clinical Investigation of Euskadi (PI2014045) and complies with the ethical guidelines of the Declaration of Helsinki (2013 revision). Before being enrolled in the study, all parents/legal guardians signed an informed written consent, and all children gave their assent.

For the current purpose, the data of 115 preadolescent children with overweight or obesity, aged between 8.5 and 12.0 years, were analysed. Overweight and obesity status was defined according to the body mass index (BMI) international age- and sex-specific cut-off values provided by the World Obesity Federation.²⁴ Having other hepatic disease or/and any other disease accompanied with elevated blood transaminase levels, such as viral hepatitis, toxic hepatitis or autoimmune diseases, was considered as exclusion criteria. The present study followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines.²⁵

2.2 | Measurements

2.2.1 | Hepatic fat

Hepatic fat percentage was assessed by MRI using a Magnetom Avanto system (Siemens Healthcare, Erlangen, Germany). The details of the hepatic fat measurement protocol have been published elsewhere.²³ Thereafter, children were categorized as having or not having MAFLD ($\geq 5.5\%$ or $< 5.5\%$ percentage hepatic fat, respectively).²⁶

2.2.2 | Sociodemographic, lifestyle and anthropometric characteristics

Body mass (SECA 760), height (SECA 220) and waist circumference (SECA 201) were measured in duplicate following standard protocols. Then, the BMI (kg/m^2) and the waist to height ratio (WHtR) were then calculated.²⁷

The sociodemographic information was obtained via self-reported questionnaire. Belonging to an ethnic minority was defined as having a foreign-born mother from a low- or middle-income country or belonging to a recognized ethnic minority for Spain (i.e., Roma) according to the categories provided by the European Commission for Spain.²⁸ Perinatal variables, such as gestational age at birth (weeks), were collected from clinical records. SSB consumption was determined as the ingestion of soft drinks, sweetened juices and energetic drinks²⁹ in g/day. Then, children were categorized as consumers or non-consumers of SSB. Dietary intake was assessed by two non-consecutive 24-h recalls within a period of 7 days. A self-reported sedentary behaviour questionnaire³⁰ was completed in order to determine the frequency of specified sedentary behaviours such as watching TV, playing on-screen games and

surfing the Internet; the children were then categorized as meeting (< 2 h/day) or not-meeting (≥ 2 h/day) the World Health Organization recommendations regarding screen time for children.³¹

2.3 | Biochemical and genetic variables

Blood extraction and collection details have been published elsewhere.²³ Fasting serum concentrations of biochemical parameters such as total cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (LDL), TG, glucose, insulin, ALT, AST, gamma-glutamyl transferase (GGT) and ferritin concentrations were measured as reported elsewhere.²³ Thereafter, the homeostasis model assessment of insulin resistance ($\text{HOMA-IR} = \text{insulin} [\text{mU}/\text{L}] \times \text{glucose} [\text{mmol}/\text{L}] / 22.5$) was calculated.³² Genomic DNA was extracted from white blood cells using Maxwell[®] RSC Blood DNA Kit and Maxwell[®] RSC Instrument (Promega) equipment. The genotyping was done by an Illumina system (Illumina, Inc, San Diego, CA, USA) using the Golden-Gate technology (sampling procedure scheme, Golden-Gate; Software, Inc, San Francisco, California). Candidate gene approach was the procedure used to select 25 single nucleotide polymorphisms (SNPs) potentially associated with MAFLD^{33,34} in the current study.

2.4 | Statistical analysis

Differences in characteristics between children with or without MRI-diagnosed MAFLD were analysed using the independent t-test (continuous variables) or χ^2 test (categorical variables).

Variables potentially associated with the presence of MAFLD were included as candidates in a multivariate logistic regression of each model. Those independent variables that showed collinearity, and those whose effect was negligible were removed from the final model. Two different models were developed: (1) biochemical model (model I), (2) genetic variants model (model II) and (3) biochemical plus genetic variants model (model III). The probability of having MAFLD was determined from the models, multiplying by 100 to obtain the index of each model, which therefore has a score range from 0 to 100.

The discriminatory capacity of each model was analysed by calculating the area under the receiver-operating characteristic curve (AUC-ROC, with 95% confidence intervals [CI]). The calibration of each model was examined using a calibration plot (plotting the expected probabilities against observed event proportions and smoothing via the Loess method) and the Hosmer-Lemeshow test. Bootstrap resampling with 150 samples was performed as an internal validation and to provide an optimism-corrected AUC-ROC.

The Youden index³⁵ was used to identify the optimal cut-off point for binary classification for the two models, prioritizing high sensitivity ($\geq 80\%$). The performance of the proposed models was

	Non-MAFLD		MAFLD		<i>p</i>
	<i>N</i>	Mean (SD)	<i>N</i>	Mean (SD)	
Anthropometric and body composition characteristics					
Age (years)	74	10.6 (1.1)	41	10.5 (1.1)	0.749
Girls (<i>N</i> , %)	74	42.57	41	20.49	0.077
Body mass index (kg/m ²)	74	25.0 (3.2)	41	26.2 (3.3)	0.059
Hepatic fat (%)	74	3.67 (0.97)	41	9.18 (4.8)	<0.001
Biochemical parameters					
Cholesterol (mg/dl)	73	170.5 (28.0)	41	174.4 (29.0)	0.489
High-density lipoprotein (mg/dl)	73	51.8 (11.8)	41	49.2 (10.3)	0.232
Low-density lipoprotein (mg/dl)	73	104.1 (23.6)	41	105.5 (24.6)	0.768
Triglycerides (mg/dl)	73	73.5 (30.0)	41	98.2 (47.4)	0.004
Glucose (mg/dl)	73	84.7 (4.9)	40	86.8 (6.1)	0.086
Insulin (μU/ml)	73	11.1 (4.3)	41	13.9 (5.5)	0.006
HOMA-IR	73	2.34 (0.95)	40	3.01 (1.28)	0.006
Aspartate aminotransferase (U/L)	73	23.0 (4.3)	41	24.6 (3.9)	0.045
Alanine aminotransferase (U/L)	73	18.0 (5.3)	41	24.8 (11.2)	0.001
Gamma-glutamyl-transferase (U/L)	72	14.9 (3.6)	40	18.8 (5.4)	<0.001
Ferritin (ng/ml)	71	45.0 (22.2)	41	67.8 (64.8)	0.035
Genetic variants (% of risk allele carriers)					
PPARG rs13081389	51	4	24	21	0.019
PPARG rs1801282	51	12	24	33	0.025
HFE rs1800562	51	6	24	21	0.050
PNPLA3 rs4823173	51	22	24	46	0.031

Note: Bold values indicate *p* value < 0.05.

Abbreviations: HOMA-IR, homeostasis model assessment of insulin resistance; MAFLD, children with metabolic associated fatty liver disease; non-MAFLD, children without metabolic associated fatty liver disease; SD, standard deviation; SSB, sugar-sweetened beverages.

expressed as sensitivity, specificity, positive and negative predictive values (with their corresponding 95% CIs) for the proposed cut-off points. All the analyses were performed for the sample as a whole and separately for boys and girls.

All calculations were performed using SPSS software v.23.0 (IBM, Armonk, NY, USA) and R statistical software v.3.6.3. Significance was set at $\alpha = 0.05$.

2.5 | Development of the prediction protocol and external validation

Two steps algorithm was developed for the detection of MAFLD. The first step is based on a short questionnaire punctuation (HEPAKID index),¹⁶ which includes anthropometric data (WHtR), sociodemographic factors (ethnic minority status and gestational age at birth) and lifestyle variables (SSB consumption and screen time). The second step is based on biochemical screening: HOMA-IR, TG, ALT, AST, GGT and ferritin (model I equation).

TABLE 1 Anthropometric and clinical characteristics of overweight/obese children with and without metabolic associated fatty liver disease in the exploratory sample of children

Once the model was developed, the external validation was performed using the baseline data from the MICROKID project ([ClinicalTrials.gov](https://clinicaltrials.gov) ID: NCT04575506) whose overall aim is the study of the influence of the diversity and composition of the microbiota in the development of MRI diagnosed MAFLD. A total of 46 preadolescent children with overweight/obesity (*N* = 20 girls), aged between 8.5 and 14.0 years, were analysed. The inclusion and exclusion criteria were the same than the original sample.

3 | RESULTS

Table 1 shows the characteristics of participants with (36%) and without MAFLD (64%). Children with MAFLD showed higher plasma TG, insulin, HOMA-IR, AST, ALT, GGT and ferritin compared to those peers without MAFLD (*p* < 0.05). Table S1 shows the characteristic of the external validation sample with (28%) and without MAFLD (72%). This sample also showed higher plasma TG, insulin, HOMA-IR, AST, ALT and ferritin in children with MAFLD compared to those peers without MAFLD (*p* < 0.05).

From the 25 SNPs potentially associated with MAFLD (Table S2), four genetic variants (Table S3) were significantly associated with the presence of the disease. The distribution of carriers/non-carriers of the risk-alleles of the *PPARG* rs13081389, *PPARG* rs1801282, *HFE* rs1800562 and *PNPLA3* rs4823173 was significantly different between children with and without MAFLD ($p < 0.05$, Table 1).

3.1 | Development of the models

Table 2 shows the multivariate logistic regression analysis of the three proposed models. The model I was based on six biochemical parameters potentially associated with having MAFLD. The model II was based on four SNPs potentially associated with having MAFLD. The model III was based on the six biochemical parameters plus the four SNPs. All models were defined using the standardized regression coefficients (β) obtained in the multivariate logistic regression analyses (Table 2). The calculators of each model are available on <https://acortar.link/1yeEyY>.

The Hosmer–Lemeshow test (Model I, $p = 0.355$, Model II, $p = 0.830$ and Model III, $p = 0.299$) and the calibration plots (Figure S1) showed the calibration of each model. The AUC-ROC values for the three indexes were 0.824 (Model I), 0.708 (Model II) and 0.888 (Model III). Model I and model III showed strong discriminatory capacity for identifying MAFLD in the study population, while model II showed limited capacity (Figure S1). The optimism corrected AUC-ROC were 0.792 (panel A, Model I), (panel B, Model II) 0.665 and 0.812 (panel C, Model III).

Table S4 shows the diagnostic performance of the three developed models at different cut-off points. For models I, II and III, the optimum cut points were 25.0, 22.0 and 24.0, respectively. Table 3 shows the diagnostic performance of the selected cut points of the three developed models for the whole sample, as well as separately for boys and girls. The models I and III showed high sensitivity (82%), but limited specificity (63%–69%). The model II showed limited sensitivity (67%) and specificity (65%) in the whole sample; likewise, it showed large differences in diagnostic performance between girls and boys (sensitivity of 50% and 83%, specificity of 57% and 74%, respectively).

3.2 | Development of the prediction protocol and external validation

Figure 1 shows the developed decision protocol algorithm for the identification of children with overweight or obesity candidates to confirmatory diagnosis of MAFLD (HEPAKID prediction protocol). This algorithm includes two steps: 1st) pre-screening of children with high risk of having MAFLD using the short questionnaire of the HEPKID index (available on <https://bit.ly/2AQTUPa>)¹⁶ and 2nd) those children whose HEPKID index is ≥ 25 are derived to a blood test to confirm their risk using the

TABLE 2 Multiple logistic regressions analysis showing the association of biochemical parameters with metabolic associated fatty liver disease in the exploratory sample of children (MAFLD, dependent variable)

	MAFLD		
	OR (95% CI)	β	p
Model I (n = 109)			
Constant	-	-9.620	0.000
HOMA-IR	1.53 (0.93–2.52)	0.425	0.095
TG (mg/dl)	1.01 (0.99–1.02)	0.010	0.164
ALT (U/L)	1.03 (0.93–1.14)	0.030	0.558
AST (U/L)	1.14 (0.98–1.32)	0.133	0.078
GGT (U/L)	1.13 (0.98–1.30)	0.126	0.069
Ferritin (ng/ml)	1.02 (0.99–1.04)	0.020	0.070
Model II (n = 75)			
Constant	-	-1.547	0.000
<i>PPARG</i> (rs13081389)	4.10 (0.39–42.65)	1.411	0.238
<i>PPARG</i> (rs1801282)	1.74 (0.33–9.36)	0.556	0.516
<i>HFE</i> (rs1800562)	4.47 (0.87–22.90)	1.498	0.072
<i>PNPLA3</i> (rs4823173)	2.93 (0.95–9.02)	1.075	0.061
Model III (n = 72)			
Constant	-	-10.940	0.002
HOMA-IR	1.38 (0.73–2.60)	0.320	0.326
TG (mg/dl)	1.01 (0.99–1.03)	0.009	0.415
ALT (U/L)	1.00 (0.81–1.23)	0.001	0.996
AST (U/L)	1.10 (0.89–1.37)	0.098	0.372
GGT (U/L)	1.35 (1.00–1.82)	0.301	0.050
Ferritin (ng/ml)	1.01 (0.98–1.04)	0.013	0.360
<i>PPARG</i> (rs13081389)	1.03 (0.05–23.30)	0.029	0.985
<i>PPARG</i> (rs1801282)	5.76 (0.65–51.10)	1.751	0.116
<i>HFE</i> (rs1800562)	10.96 (1.09–109.87)	2.394	0.042
<i>PNPLA3</i> (rs4823173)	1.6 (0.282–9.08)	0.470	0.596

Note: Only participants with no missing data were included into the model. Missing data model I, GGT ($n = 2$), HOMA-IR ($n = 1$), ferritin ($n = 2$). Missing data model II, GGT ($n = 2$), HOMA-IR ($n = 1$), Ferritin ($n = 2$), SNPs analysing ($n = 37$). Missing data model III, SNPs analysing ($n = 37$). Abbreviations: β , standardized regression coefficient; ALT, alanine transaminase; AST, aspartate transaminase; CI, confidence interval; GGT, gamma-glutamyl-transferase; HFE, homeostatic iron regulator; HOMA-IR, homeostatic model assessment; OR, odds ratio; PNPLA3, patatin like phospholipase domain containing 3; PPARG, peroxisome proliferator activated receptor gamma; TG, triglycerides.

screening tool based on Model I equation (available on <https://acortar.link/1yeEyY>).

Figure S2 (panel A) shows the performance of the complete algorithm in the main sample. The complete algorithm showed high discriminatory capacity with 9 false negatives and 10 false positives in the original sample, reaching a sensitivity of 72% and a specificity of 84% ($N = 93$).

TABLE 3 Diagnostic performance of the two developed models for magnetic resonance imaging-diagnosed paediatric metabolic associated fatty liver disease identification in the exploratory sample of children

Cut-off points	SN, % (95% CI)	SP, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Model I				
Cut-point ≥ 25				
Whole sample ($n = 109$)	82 (69–95)	63 (51–75)	55 (42–69)	86 (76–97)
Girls ($n = 61$)	75 (54–96)	63 (47–79)	50 (30–70)	84 (69–98)
Boys ($n = 48$)	89 (73–100)	62 (43–81)	61 (41–81)	90 (74–100)
Model II				
Cut-point ≥ 22				
Whole sample ($n = 75$)	67 (46–88)	65 (51–79)	47 (29–65)	80 (67–94)
Girls ($n = 61$)	50 (18–82)	57 (37–77)	33 (9–58)	73 (52–94)
Boys ($n = 48$)	83 (58–100)	74 (54–94)	63 (36–89)	89 (73–100)
Model III				
Cut-point ≥ 24				
Whole sample ($n = 72$)	82 (65–100)	69 (55–83)	56 (38–74)	89 (78–100)
Girls ($n = 40$)	83 (58–100)	75 (57–93)	58 (32–85)	91 (78–100)
Boys ($n = 32$)	82 (54–100)	62 (39–85)	53 (26–80)	87 (66–100)

Abbreviations: CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; SN, sensitivity; SP, specificity.

The discriminatory capacity of the algorithm in the external validation sample can be found in Figure S2 (panel B). The validation algorithm showed high discriminatory capacity (Table S5), with 4 false negatives and 5 false positives, reaching a sensitivity of 70% and a specificity of 85% ($N = 45$).

Finally, the two steps comprised in the prediction protocol, the HEPAKID index and the biochemical screening, were also independently validated with the external sample, and the results showed similar diagnostic performance in the two samples (Table S5).

4 | DISCUSSION

The most important contribution of this study is the development of an easy to perform and minimally invasive prediction protocol for the identification of MAFLD among children with overweight/obesity, which encompasses elevated sensitivity, specificity and high accuracy. This algorithm, based on a short questionnaire and easy to measure biochemical parameters, may be useful in routine Primary Care clinical practice to identify early those children who should be referred to perform a confirmatory diagnosis.

We developed three different models in order to identify the most appropriate model to serve as a second-step screening tool for MAFLD in children with overweight/obesity. However, the exclusive application of these models, based on biochemical and/or SNPs data, showed moderate accuracy (sensitivity 67%–82% and specificity 63%–69%) to detect MAFLD.

These findings are in concordance with previous reports.^{12,15,16,36,37} In this way, the application of the model I,

based exclusively on biochemical parameters, showed limited applicability. Thus, the prioritization of high sensitivity (82%) with a cut point of ≥ 25 showed low specificity (63%), while the prioritization of high specificity (94%) with a cut-point of ≥ 60 showed very low sensitivity (49%). Indeed, although biochemical parameters such as HOMA-IR, TG, ALT, AST, GGT or ferritin levels are increased in children with MAFLD,^{38,39} their prediction capacity is not enough for the screening of MAFLD.^{12,13,15,40}

The genetic risk score (model II) based on four SNPs associated with MAFLD (*PPARG* rs13081389, *PPARG* rs1801282, *HFE* rs1800562 and *PNLPLA3* rs4823173) also showed limited discriminatory capacity (67% sensitivity and 65% specificity). In turn, the combination of the biochemical and genetic variables (model III) did not improve the accuracy enough (82% sensitivity and 69% specificity) in our study sample. Thus, considering the necessary technological resources for the analysis of the SNPs, the minimal specificity improvement of the prediction tool and its high economic cost, this model becomes non-eligible for the routine clinical practice. These results concur with previous reports in children and adults, where the addition of genetic information to clinical parameters in the calculations of the risk scores resulted in minimal improvements of sensitivity and specificity.^{41,42} In a cohort of Italian children with obesity and adolescent aged 6–18 years, the addition of three, four or eleven SNPs only slightly improved the AUC-ROC from 0.77 to 0.80, 0.80 and 0.81, respectively.⁴² In older Chinese adults, the addition of genetic variants to different models improved their sensitivity but worsened the specificity.⁴¹

Genetic susceptibility seems to play a crucial role in the development and progression of MAFLD.⁴³ Therefore, genetic variants have been proposed as potential biomarkers of MAFLD

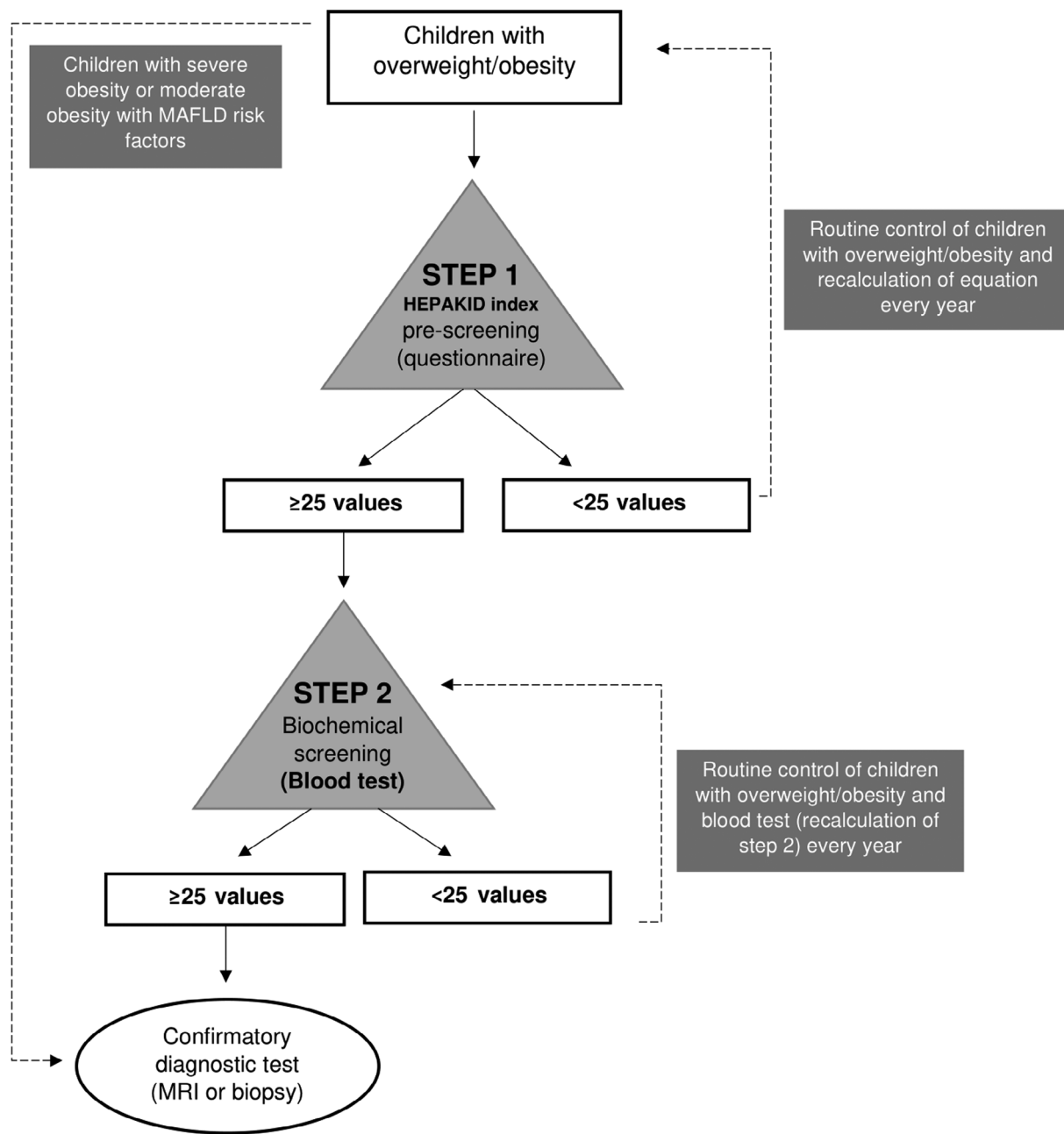


FIGURE 1 HEPAKID prediction protocol algorithm for screening paediatric metabolic associated fatty liver disease (MAFLD); MRI, magnetic resonance imaging

in adults⁴⁴ and children.^{22,45,46} However, MAFLD is a polygenetic disease where dynamic interactions between genes and environmental factors can modulate the development and progression of the disease.⁴³ Therefore, we probably need more genetic information of MAFLD-susceptible genes, as well as studies examining the gene-environmental factors interactions, rather than just several SNPs, to establish accurate predictive models. In addition, to date, genetic variables are not easily available in routine clinical practice, which limits its application as a massive screening tool. Thus, the model I is the most appropriate model to serve as a second-step of the proposed screening protocol.

This study adds to the current knowledge the development of an accurate, sensitive (72%), specific (84%), simple and minimally invasive screening protocol for the identification of MRI-diagnosed MAFLD among children with overweight/obesity: the HEPAKID prediction protocol. This algorithm combines two consecutive steps without genetic information and/or difficult to measure biochemical parameters in routine clinical practice. In the first step, children are classified as 'at risk of having MAFLD' or 'not' depending on the score achieved in the HEPAKID index pre-screening tool,¹⁶ which is derived from a questionnaire based on the recording of sociodemographic factors (ethnic minority status and gestational age at birth), anthropometric data (WHtR) and lifestyle variables

(SSB consumption and screen time). In the second step, those children identified in the previous step as 'at risk' (HEPAKID index ≥ 25) have to be referred for a blood test to perform a second screening using common blood biochemical parameters (glucose and insulin to calculate HOMA-IR, TG, ALT, AST, GGT and ferritin). Those children with a score ≥ 25 in this second step should be sent to a medical specialist to confirm the diagnosis. In addition, the proposed protocol was validated in an external sample ($N = 45$) showing similar results (sensitivity 70% and specificity 85%), which strengthens its prediction capacity.

In adults, several prediction scores showed the elevated capacity of anthropometric and clinical parameters to predict the risk of suffering fatty liver disease.^{47,48,49} Nevertheless, in children, these scores have very limited accuracy (AUC-ROC between 0.68 and 0.75).³⁶ Previously proposed prediction scores or algorithms for the screening of paediatric MAFLD^{15,36,42} showed reasonable accuracy (between 0.81 and 0.88) and sensitivity (between 77% and 89%) but very limited application in external validations. For instance, the Ped-NAFLD score¹⁵ was tested in a cohort of 119 children, showing 75% of sensitivity and 68% of specificity.³⁶ In another study with 113 children, its sensitivity dropped to 33%, while the specificity was 95%.¹⁶ However, these models include non-easy-to-measure parameters such as blood leptin and adiponectin¹⁵ or genetic information⁴² that limits their routine applicability. Similarly, the algorithms and the ALT-level-based cut-off points proposed by either NASPGHAN¹⁰ or ESPGHAN¹¹ show high specificity (between 88% and 94%) but very low sensitivity (between 26% and 48%) compromising their utility as screening tools.

The combination of sociodemographic, anthropometric, lifestyle and clinical information within the same algorithm seems to be the key to achieve high sensitivity ($>70\%$), specificity ($>80\%$) and elevated discriminatory capacity to identify children with MAFLD among those with overweight or obesity. Likewise, the high specificity achieved after performing the two steps makes this tool useful for clinical practice avoiding unnecessary costly or invasive testing in patients without the disease, allowing its application in the entire child population with overweight or obesity. The proposed decision tree also contemplates the possibility of direct derivation to confirmatory diagnostic tests of children with moderate or severe obesity with MAFLD risk factors (such as family history of MAFLD, very high hepatic enzyme levels or hepatic symptomatology). Moreover, those children who maintain their overweight/obesity status, but who were not classified as children at risk of MAFLD in the first or in the second step, should be monitored and assessed yearly to avoid leaving any patient untreated in the future. A simple guide explaining the application of the proposed protocol can be found in Table S6.

The proposal of the current study complements our previous sensitive pre-screening tool (the HEPKID index),¹⁶ adding the necessary specificity of a medical screening tool, but maintaining its simplicity, easiness and low economic cost. In any case, although the results were consistent in the validation sample, the proposed protocol should be externally validated in larger, multiethnic and representative cohorts of children with overweight/obesity before its implementation in clinical settings. In this line, in our study, ethnic minority

condition was defined as belonging to a recognized ethnic minority for Spain or as having a foreign-born mother from a low- or middle-income country. In Spain, these groups share social disadvantages, and the results were consistent in the external validation sample: However, it should be also tested in other multiethnic cohorts from other European and non-European countries.

In conclusion, the HEPKID prediction protocol identifies with high sensitivity, specificity and accuracy, as well as low time-consuming and economic cost children with overweight/obesity who likely suffer MAFLD, and who should be referred for confirmatory diagnosis.

ACKNOWLEDGEMENTS

This project was funded by the Spanish Ministry of Health 'Fondos de Investigación Sanitaria del Instituto de Salud Carlos III' (PI13/O1335), the Spanish Ministry of Industry and Competitiveness (DEP2016-78377-R), by EU Fondos Estructurales de la Unión Europea (FEDER) funds ('Una manera de hacer Europa') and Department of Economic Development of Government of Navarra (0011-1365-2019-000000, 0011-1365-2020-00140). M.O. is supported by a grant from the Spanish Ministry of Economy and Competitiveness, grant number; BES-2017-080770. We thank Open Access funding provided by Universidad Pública de Navarra. A special thank for the families, both children and parents, for their participation and collaboration in the study.

CONFLICT OF INTEREST

No conflict of interest was declared.

AUTHOR CONTRIBUTIONS

Maddi Oses analysed the data, drafted the manuscript and takes full responsibility for the integrity of the data analyses and collected data. Likewise, generated the figures, participated in the interpretation of the results and critically revised the manuscript for important intellectual content. Cristina Cadenas-Sanchez, María Medrano and Emiliano Miranda-Ferrua collected the data, participated in the interpretation of the results and critically revised the manuscript for its intellectual content. Arkaitz Galbete participated in the statistical analysed, interpretation of the results and critically revised the manuscript for its intellectual content. Rafael Cabeza, Arantxa Villanueva and Fernando Idoate participated in medical imaging processing and critically revised the manuscript. Jonatan Ruiz, Felix Sánchez-Valverde. and Francisco B. Ortega critically revised the manuscript for its intellectual content. Idoia Labayen designed the study, coordinated and supervised data collection, drafted the manuscript, participated in the interpretation of the results and critically revised the manuscript for important intellectual content. All authors were involved in writing the article and had final approval of the submitted and published versions.

ORCID

Maddi Oses  <https://orcid.org/0000-0001-7815-7583>

Cristina Cadenas-Sanchez  <https://orcid.org/0000-0002-4513-9108>

María Medrano  <https://orcid.org/0000-0001-7048-642X>

Arkaitz Galbete  <https://orcid.org/0000-0001-5622-5418>

Arantxa Villanueva  <https://orcid.org/0000-0001-9822-2530>

REFERENCES

- Draijer L, Benninga M, Koot B. Pediatric NAFLD: an overview and recent developments in diagnostics and treatment. *Expert Rev Gastroenterol Hepatol* [Internet]. 2019;13:447-461. doi:10.1080/17474124.2019.1595589
- Eslam M, Newsome PN, Sarin SK, et al. A new definition for metabolic dysfunction-associated fatty liver disease: an international expert consensus statement. *J Hepatol* [Internet]. 2020;73(1):202-209. doi:10.1016/j.jhep.2020.03.039
- Eslam M, Alkhoury N, Vajro P, et al. Defining paediatric metabolic (dysfunction)-associated fatty liver disease: an international expert consensus statement. *Lancet Gastroenterol Hepatol*. 2021;6(10):864-873. doi:10.1016/S2468-1253(21)00183-7
- Newton KP, Hou J, Crimmins NA, et al. Prevalence of prediabetes and type 2 diabetes in children with nonalcoholic fatty liver disease. *JAMA Pediatr*. 2016;170(10):e161971.
- Schwimmer JB, Deutsch R, Kahen T, Lavine JE, Stanley C, Behling C. Prevalence of fatty liver in children and adolescents. *Pediatrics*. 2006;118(4):1388-1393.
- Holterman AXL, Guzman G, Fantuzzi G, et al. Nonalcoholic fatty liver disease in severely obese adolescent and adult patients. *Obesity*. 2013;21(3):591-597.
- Labayen I, Medrano M, Arenaza L, et al. Effects of exercise in addition to a family-based lifestyle intervention program on hepatic fat in children with overweight. *Diabetes Care*. 2020;43(2):306-313.
- Luttikhuis HO, Baur L, Jansen H, et al. Interventions for treating obesity in children. *Cochrane Database Syst Rev*. 2008.
- Clemente MG, Mandato C, Poeta M, Vajro P. Pediatric non-alcoholic fatty liver disease: recent solutions, unresolved issues, and future research directions. *World J Gastroenterol*. 2016;22(36):8078-8093.
- Vos MB, Abrams SH, Barlow SE, et al. NASPGHAN clinical practice guideline for the diagnosis and treatment of nonalcoholic fatty liver disease in children: recommendations from the Expert Committee on NAFLD (ECON) and the North American Society of Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN). *J Pediatr Gastroenterol Nutr*. 2017;64(2):319-334.
- Vajro P, Lenta S, Socha P, et al. Diagnosis of nonalcoholic fatty liver disease in children and adolescents: position paper of the ESPGHAN Hepatology Committee. *J Pediatr Gastroenterol Nutr*. 2012;54(5):700-713.
- Koot BGP, Nobili V. Screening for non-alcoholic fatty liver disease in children: do guidelines provide enough guidance? *Obes Rev*. 2017;18(9):1050-1060.
- Ma X, Liu S, Zhang J, et al. Proportion of NAFLD patients with normal ALT value in overall NAFLD patients: a systematic review and meta-analysis. *BMC Gastroenterol*. 2020;20(1):10.
- Schwimmer JB, Dunn W, Norman GJ, et al. SAFETY study: alanine aminotransferase cutoff values are set too high for reliable detection of pediatric chronic liver disease. *Gastroenterology*. 2010;138(4):1357-1364.
- Maffei C, Banzato C, Rigotti F, et al. Biochemical parameters and anthropometry predict NAFLD in obese children. *J Pediatr Gastroenterol Nutr*. 2011;53(6):590-593.
- Oses M, Medrano M, Galbete A, et al. A sociodemographic, anthropometric and lifestyle-based prediction score for screening children with overweight and obesity for hepatic steatosis: the HEPAKID index. *Pediatr Obes*. 2021;1-9.
- Buzzetti E, Pinzani M, Tsochatzis EA. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Metabolism*. 2016;65(8):1038-1048. doi:10.1016/j.metabol.2015.12.012
- Prokopowicz Z, Malecka-Tendera E, Matusik P. Predictive value of adiposity level, metabolic syndrome, and insulin resistance for the risk of nonalcoholic fatty liver disease diagnosis in obese children. *Can J Gastroenterol Hepatol*. 2018;2018:9465784.
- Yang HR, Kim HR, Kim MJ, Ko JS, Seo JK. Noninvasive parameters and hepatic fibrosis scores in children with nonalcoholic fatty liver disease. *World J Gastroenterol*. 2012;18(13):1525-1530.
- Stanislawski MA, Shaw J, Litkowski E, et al. Genetic risk for hepatic fat among an ethnically diverse cohort of youth: the exploring perinatal outcomes among children study. *J Pediatr*. 2020;220:146-153.e2. doi:10.1016/j.jpeds.2020.01.031
- Goran MI, Walker R, Le K, et al. Effects of PNPLA3 on liver fat and metabolic profile in Hispanic children and adolescents. *Diabetes*. 2010;59:3127-3130.
- Lin YC, Wu CC, Ni YH. New perspectives on genetic prediction for pediatric metabolic associated fatty liver disease. *Front Pediatrics*. 2020;8:603654.
- Medrano M, Maiz E, Maldonado-Martín S, et al. The effect of a multidisciplinary intervention program on hepatic adiposity in overweight-obese children: protocol of the EFIGRO study. *Contemp Clin Trials*. 2015;45:346-355. doi:10.1016/j.cct.2015.09.017
- Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatr Obes*. 2012;7(4):284-294.
- Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-W73.
- Browning JD, Szczepaniak LS, Dobbins R, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology*. 2004;40(6):1387-1395.
- Lo K, Wong M, Khalechelvam P, Tam W. Waist-to-height ratio, body mass index and waist circumference for screening paediatric cardio-metabolic risk factors: a meta-analysis. *Obes Rev*. 2016;17(12):1258-1275.
- European Commission against Racism and Intolerance. ECRI Report on Spain (fifth monitoring cycle) 2018; December 2017:46.
- Arenaza L, Medrano M, Oses M, et al. Dietary determinants of hepatic fat content and insulin resistance in overweight/obese children: a cross-sectional analysis of the Prevention of Diabetes in Kids (PREDIKID) study. *Br J Nutr*. 2019;121(10):1158-1165.
- Moreno LA, Gottrand F, Huybrechts I, Ruiz JR, González-Gross M, DeHenauf S. Nutrition and lifestyle in European adolescents: the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) study. *Adv Nutr*. 2014;5(5):615S-623S.
- American Academy of Pediatrics. Children, adolescents, and television. *Comm Public Educ Child*. 2001;107:423-426.
- Matthews DR, Hosker JR, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. 1985;28:412-419.
- Eslam M, Valenti L, Romeo S. Genetics and epigenetics of NAFLD and NASH: clinical impact. *J Hepatol*. 2018;68(2):268-279. doi:10.1016/j.jhep.2017.09.003
- Del Campo JA, Gallego-Durán R, Gallego P, Grande L. Genetic and epigenetic regulation in nonalcoholic fatty liver disease (NAFLD). *Int J Mol Sci*. 2018;19(3):5-10.
- Liu X. Classification accuracy and cut point selection. *Stat Med*. 2012;31(23):2676-2686.
- Koot BGP, van der Baan-Slootweg OH, Bohte AE, et al. Accuracy of prediction scores and novel biomarkers for predicting nonalcoholic fatty liver disease in obese children. *Obesity*. 2013;21(3):583-590.
- Yu EL, Golshan S, Harlow KE, et al. Prevalence of nonalcoholic fatty liver disease in children with obesity. *J Pediatr*. 2019;207:64-70. doi:10.1016/j.jpeds.2018.11.021

38. Zhao K, Ju H, Wang H. Metabolic characteristics of obese children with fatty liver: a STROBE-compliant article. *Medicine (Baltimore)*. 2019;98(16):e14939.
39. Jimenez-Rivera C, Hadjiyannakis S, Davila J, et al. Prevalence and risk factors for non-alcoholic fatty liver in children and youth with obesity. *BMC Pediatr*. 2017;17(1):1-7.
40. Schwimmer JB, Newton KP, Awai HI, et al. Paediatric gastroenterology evaluation of overweight and obese children referred from primary care for suspected non-alcoholic fatty liver disease. *Aliment Pharmacol Ther*. 2013;38(10):1267-1277.
41. Yang H, Chen G, Song C, et al. A novel index including SNPs for the screening of nonalcoholic fatty liver disease among elder Chinese: a population-based study. *Medicine (Baltimore)*. 2018;97(13):e0272. doi:[10.1097/MD.00000000000010272](https://doi.org/10.1097/MD.00000000000010272)
42. Zusi C, Mantovani A, Olivieri F, et al. Contribution of a genetic risk score to clinical prediction of hepatic steatosis in obese children and adolescents. *Dig Liver Dis*. 2019;51(11):1586-1592. doi:[10.1016/j.dld.2019.05.029](https://doi.org/10.1016/j.dld.2019.05.029)
43. Macaluso FS, Maida M, Petta S. Genetic background in nonalcoholic fatty liver disease: a comprehensive review. *World J Gastroenterol*. 2015;21(39):11088-11111.
44. Severson TJ, Besur S, Bonkovsky HL. Genetic factors that affect non-alcoholic fatty liver disease: a systematic clinical review. *World J Gastroenterol*. 2016;22(29):6742-6756.
45. Castillo-Leon E, Cioffi CE, Vos MB. Perspectives on youth-onset non-alcoholic fatty liver disease. *Endocrinol Diabetes Metab*. 2020;3(4):1-12.
46. Hudert CA, Selinski S, Rudolph B, et al. Genetic determinants of steatosis and fibrosis progression in paediatric non-alcoholic fatty liver disease. *Liver Int*. 2019;39(3):540-556.
47. Kotronen A, Peltonen M, Hakkarainen A, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. *Gastroenterology*. 2009;137(3):865-872. doi:[10.1053/j.gastro.2009.06.005](https://doi.org/10.1053/j.gastro.2009.06.005)
48. Bedogni G, Bellentani S, Miglioli L, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterol*. 2006;6:1-7.
49. Lee JH, Kim D, Kim HJ, et al. Hepatic steatosis index: a simple screening tool reflecting nonalcoholic fatty liver disease. *Dig Liver Dis*. 2010;42(7):503-508. doi:[10.1016/j.dld.2009.08.002](https://doi.org/10.1016/j.dld.2009.08.002)

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Oses M, Cadenas-Sanchez C, Medrano M, et al. Development of a prediction protocol for the screening of metabolic associated fatty liver disease in children with overweight or obesity. *Pediatric Obesity*. 2022; 17(9):e12917. doi:[10.1111/ijpo.12917](https://doi.org/10.1111/ijpo.12917)