proceedings

MDPI

Abstract

Leveraging Machine Learning and Genetic Risk Scores for the Prediction of Metabolic Syndrome in Children with Obesity †

Concepción M. Aguilera [1,2,3,*], Mireia Bustos-Aibar [1,2,4], Augusto Anguita-Ruiz [2,5], Álvaro Torres-Martos [1,2,3], Gloria Bueno [2,4,6], Rosaura Leis [7,8] and Jesús Alcalá-Fernández [9]

1 Department of Biochemistry and Molecular Biology II, Institute of Nutrition and Food Technology "José Mataix", Center of Biomedical Research, University of Granada, 18071 Granada, Spain; mbustos@iisaragon.es (M.B.-A.); alvarotorres@ugr.es (Á.T.-M.)
2 The Center for Biomedical Research Network Physiopathology of Obesity and Nutrition (CIBEROBN), Institute of Health Carlos III (ISCIII), 28029 Madrid, Spain; augusto.anguita@isglobal.org (A.A.-R.); mgbuenol@unizar.es (G.B.)
3 Instituto de Investigación Biosanitaria ibs.GRANADA, 18012 Granada, Spain
4 GENUD (Growth, Exercise, NUtrition and Development) Research Group, Agri-Food Institute of Aragon (IA2), Aragon Health Research Institute (IIS Aragón), University of Zaragoza, 50018 Zaragoza, Spain
5 Institute for Global Health (ISGlobal), 08003 Barcelona, Spain
6 Unit of Pediatric Endocrinology, University Clinical Hospital Lozano Blesa, 50009 Zaragoza, Spain
7 Unit of Pediatric Gastroenterology, Hepatology and Nutrition, Pediatric Service, Hospital Clínico Universitario de Santiago, 15706 Santiago de Compostela, Spain; mariarosaura.leis@usc.es
8 Unit of Investigation in Nutrition, Growth and Human Development of Galicia-USC, Pediatric Nutrition Research Group-Health Research Institute of Santiago de Compostela (IDIS), 15706 Santiago de Compostela, Spain
9 Department of Computer Science and Artificial Intelligence, Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI), University of Granada, 18071 Granada, Spain; jalcala@decsai.ugr.es
* Correspondence: caguiler@ugr.es
† Presented at the 14th European Nutrition Conference FENS 2023, Belgrade, Serbia, 14–17 November 2023.

Abstract: Background and objectives: Obesity is a growing global epidemic, associated with increased cardiometabolic disorders. Metabolic syndrome (MS) is defined by altered insulin, blood pressure, glucose, and lipid levels. Pubertal children with obesity are highly susceptible to developing MS, necessitating its early identification. This study aims to compute phenotype-specific genetic risk scores for MS-related biochemical markers and evaluate their clinical utility using machine learning-based models. Methods: Longitudinal data from the PUBMEP Spanish cohort were analyzed, including 138 children (71 girls and 67 boys) at two time points, spanning from prepuberty to puberty. Clinical, endogenous, environmental, and omics variables were measured. Genetic risk scores were generated using GWAS data and PRSice-2 software. These scores, alongside non-genetic prepubertal data (e.g., biochemical, anthropometric, and physical activity data), were integrated into predictive models using machine learning techniques to forecast the MS status during puberty. Linear models explored interactions between environmental factors, genetic risk scores, and disease risk. Results: Strong associations were observed between each genetic risk score and its corresponding phenotypic biomarker. Notably, certain scores related to obesity and high-density lipoprotein levels exhibited significant interactions with environmental factors, such as sedentary lifestyle, modulating disease effects. The predictive machine learning models incorporating prepubertal genetics, high-density lipoprotein, and sedentary lifestyle achieved reasonable performance in predicting pubertal obesity (AUC, accuracy, and sensitivity of 0.89). These models strike a favorable balance between risk scores derived from genetic factors and clinical variables. However, when individual risk scores were considered in isolation, limited predictive results were observed for MS and associated altered components. Discussion: This study demonstrates the importance of the early identification of at-risk children for MS. The integration of genetic risk scores, clinical variables, and machine learning techniques offers promising avenues for predicting pubertal MS. While individual risk scores have limitations in isolation, polygenic risk scores serve as valuable tools for investigating gene–environment interactions. Following our results, polygenic risk scores lacked sufficient predictive

check for updates

Citation: Aguilera, C.M.; Bustos-Aibar, M.; Anguita-Ruiz, A.; Torres-Martos, Á.; Bueno, G.; Leis, R.; Alcalá-Fernández, J. Leveraging Machine Learning and Genetic Risk Scores for the Prediction of Metabolic Syndrome in Children with Obesity. *Proceedings* **2023**, *91*, 377. https://doi.org/10.3390/proceedings2023091377

Academic Editors: Sladjana Sobajic and Philip Calder

Published: 27 February 2024

ability in most clinical traits, limiting their clinical application. Nevertheless, they remain valuable analytical tools for exploring the association with the environment, by consolidating the effects of multiple single nucleotide polymorphisms into a single variable.

Keywords: machine learning; genetics; gene-environment interaction; genetic markers; childhood obesity; polygenetic risk scores

---

Author Contributions: Conceptualization, C.M.A., A.A.-R. and J.A.-F.; methodology, M.B.-A., A.A.-R. and J.A.-F.; software, M.B.-A., Á.T.-M. and A.A.-R.; validation, M.B.-A., A.A.-R. and J.A.-F.; formal analysis, M.B.-A., A.A.-R. and J.A.-F.; investigation, C.M.A., A.A.-R. and J.A.-F.; resources, C.M.A., G.B., R.L. and J.A.-F.; data curation, M.B.-A., A.A.-R. and Á.T.-M.; writing—original draft preparation, C.M.A., M.B.-A., A.A.-R. and J.A.-F.; writing—review and editing, C.M.A., M.B.-A., A.A.-R., Á.T.-M., G.B., R.L. and J.A.-F.; visualization, M.B.-A., A.A.-R. and J.A.-F.; supervision, C.M.A., A.A.-R. and J.A.-F.; project administration, C.M.A., G.B., R.L. and J.A.-F.; funding acquisition, C.M.A., G.B., R.L. and J.A.-F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Instituto de Salud Carlos III co-funded by the European Union and ERDF A way of making Europe (grant numbers PI20/00563, PI20/00711, PI20/00924, P20/00988, PI23/00028, PI23/00129, PI23/01032, PI23/00165 and also PI23/00191), and by the European Union through the Horizon Europe Framework Programme (eprObes project, grant number GA 101080219). The authors also acknowledge Instituto de Salud Carlos III for personal funding of A.A.R, A.T.M and M.B.A.: i-PFIS and PFIS contracts: IIS doctorates—company in health sciences and technologies of the Strategic Health Action (IFI17/00048, IFI22/00013 and FI23/00042). We also thank the support from the grant FJC2021-046952-I by Ministerio de Ciencia, Innovacion y Universidades y Agencia Estatal de Investigacion.

Institutional Review Board Statement: The study was authorized by the Ethics Committee of the Reina Sofia Hospital, in accordance with the recommendations of the European Union's Good Clinical Practice (Document 111/3976/88 July 1990) and following the Revised Declaration of Helsinki. It was also approved by current Spanish legislation regulating clinical studies with human subjects (RD 223/04 on clinical trials).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Due to stringent privacy and ethical considerations, the raw data supporting the conclusions of this article cannot be made available by the authors, but further information about the data and how it was collected can be provided upon request with the understanding that the privacy of all participants will be preserved.

Conflicts of Interest: The authors declare no conflict of interest.