



CLÍNICA E INVESTIGACIÓN EN ARTERIOSCLEROSIS

www.elsevier.es/arterio



ORIGINAL ARTICLE

Validation of the IberScore model in a primary care population



Carlos Fernández-Labandera Ramos^{a,*}, Irene Moral^{b,c}, Carlos Brotons^{b,c},
Luis Quevedo Aguado^a, Inmaculada Coca Prieto^d, Pedro Valdivielso^{d,e,f},
Miguel Ángel Sánchez Chaparro^{d,e,f}

^a Dirección de Gestión del Conocimiento y la Innovación, Ibermutua, MATEPSS 274, Madrid, Spain

^b Unidad de Investigación, Equipo de Atención Primaria Sardenya, Spain

^c Institut de Recerca Sant Pau (IR Sant Pau), Barcelona, Spain

^d Internal Medicine Unit, Hospital Universitario Virgen de la Victoria, Málaga, Spain

^e IBIMA-Plataforma BIONAND, Málaga, Spain

^f Departamento de Medicina y Dermatología, Universidad de Málaga, Málaga, Spain

Received 27 September 2023; accepted 14 December 2023

KEYWORDS

Risk prediction;
Cardiovascular
disease;
Primary prevention

Abstract

Background: This study aimed to validate the IberScore cardiovascular risk model in a population attended in the primary care setting.

Methods: A cohort of patients with no history of cardiovascular disease visited in a primary care center during the years 2008 and/or 2009 and followed up until 2018 was selected.

Cardiovascular risk was calculated with the IberScore formula for all the subjects of the cohort and the model was calibrated, graphically represented by risk deciles the proportion of expected events and proportion of observed events at 10-year follow-up, stratified by sex. The area under the ROC curve was calculated to assess the discrimination of the model.

Results: A total of 10,085 patients visited during the years 2008 and/or 2009 were included in the study. Men showed a mean 10-year risk of suffering a fatal or non-fatal cardiovascular events according to IberScore of 17.07% (SD 20.13), with a mean estimated vascular age of more than 4 years higher than the biological age; while women had a mean 10-year risk of 7.91% (SD 9.03), with an estimated vascular age of more than 2 years above the biological age.

The area under the ROC curve showed a discrimination index of the model of 0.86 (95% CI 0.84–0.88) in men and 0.82 (95% CI 0.79–0.85) in women.

* Corresponding author.

E-mail address: carlosfernandezlabandera@ibermutua.es (C. Fernández-Labandera Ramos).

Conclusion: IberScore model discriminates well in the population attended in primary care but the model overestimates the risk.

© 2023 Sociedad Española de Arteriosclerosis. Published by Elsevier España, S.L.U. All rights reserved.

PALABRAS CLAVE

Predicción del riesgo;
Enfermedad
cardiovascular;
Prevención primaria

Validación del modelo IberScore en una población de atención primaria

Resumen

Objetivo: Validar el modelo de riesgo cardiovascular IberScore en una población atendida en atención primaria.

Métodos: Se seleccionó una cohorte de pacientes sin antecedentes de enfermedad cardiovascular visitados en un centro de atención primaria durante los años 2008 y/o 2009 y seguidos hasta el año 2018.

El riesgo cardiovascular se calculó con la fórmula IberScore para todos los sujetos de la cohorte y se calibró el modelo, representado gráficamente por la proporción de eventos esperados y la proporción de eventos observados a los 10 años de seguimiento por deciles de riesgo, estratificando por sexo. Se calculó el área bajo la curva ROC para evaluar la discriminación del modelo.

Resultados: Se incluyeron en el estudio un total de 10.085 pacientes visitados durante los años 2008 y/o 2009. Los hombres presentaron un riesgo medio a los 10 años de sufrir un evento cardiovascular fatal o no fatal según IberScore del 17,07% (DE: 20,13), con una edad vascular media estimada superior en más de 4 años a su edad biológica. Las mujeres tenían un riesgo medio a 10 años del 7,91% (DE: 9,03), con una edad vascular estimada de más de 2 años por encima de su edad biológica.

El área bajo la curva ROC mostró un índice de discriminación del modelo de 0,86 (IC 95%: 0,84-0,88) en hombres y de 0,82 (IC 95%: 0,79-0,85) en mujeres.

Conclusión: El modelo IberScore discrimina bien en la población atendida en atención primaria pero sobreestima el riesgo.

© 2023 Sociedad Española de Arteriosclerosis. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Introduction

The European guidelines for the management of vascular risk factors and cardiovascular prevention recommend the use of the European SCORE model (the most recent, SCORE2 and SCORE2 OP) to estimate the risk and establish preventive priorities (modification of lifestyles, initiation of pharmacological treatment)¹⁻³; although only from the age of 40 it is possible to estimate the risk with these scales. Identifying people at high cardiovascular risk is challenging, specially at a younger age. The estimated ten-year risk with these scales in young individuals is low or moderate even in the presence of several cardiovascular risk factors. This is relevant, as it has been shown how CVR factors are significantly involved in the prevalence and progression of subclinical arteriosclerosis in individuals under 50 years of age.⁴ In this regard, the latest European Guidelines for Cardiovascular Disease Prevention already recommend including and assessing the calculation of lifetime risk or vascular age.¹ For subjects under 40 years of age, several models have been developed that estimate lifetime cardiovascular risk (CVR), such as the QRISK-LTR equation, which estimates lifetime CVR between 30 and 95 years of age⁵ and the "IBERLIFERISK"

model,⁶ recently recalibrated,⁷ developed in a large sample of the Spanish working population with a mean age of 35 years, which estimates the lifetime risk from 18 to 75 years of age. However, as acknowledged by the authors, there is no evidence regarding the present threshold for considering a patient at high risk according to the lifetime risk.⁶ Also, the evidence of the benefit of applying these lifetime models in the long term is limited and highly variable depending on the model applied. The percentage of patients who were reclassified as high-risk patients in the long-term calculation was 1.61% (95% CI 1.55-1.66) for the QRISK-LTR,⁵ while for the model developed by the American College of Cardiology/American Heart Association (ACC/AHA) for subjects over 50 years was 27.1% (95% CI 27.11-27.70).⁸

Taking into account the above-mentioned points, our group has also developed a model to estimate the 10-year risk in a large sample of the working population (almost one million workers included with a mean age of 34 years), which improves risk prediction for fatal and non-fatal events compared to the European SCORE model.⁹ The aim of this paper is to validate this model in a population attended in the primary care setting.

Methods

The ICARIA^{10,11} study evaluates the cardiovascular health of the Spanish working population by analyzing a cohort of workers who underwent an occupational health examination at the Ibermutuamur Prevention Society (now Ibermutua) during the years 2004–2018. The methodology of the ICARIA study has already been described previously.^{10,11}

The IberScore Project^{9,11} included in the ICARIA study,^{10,11} systematically evaluates the cardiovascular risk factors of the study cohort and aims to develop and validate a mathematical model for cardiovascular risk prediction for fatal and non-fatal events at 10 years for the working population.

In the IberScore equation, coronary disease is classified as a non-fatal cardiovascular event, codes 410–414 according to the International Classification of Diseases, 9th revision, clinical modification (ICD-9-CM) in 2002, heart failure (ICD-9-CM 428), cerebrovascular diseases (ICD-9-CM 431–438) and peripheral vascular disease (ICD-9-CM 440–444). Events that were not of atherosclerotic origin, corresponding according to the ICD-9-CM to codes 432.1, 437.2, 437.3, 437.7, 442, 443.0 and 443.1, were not considered a cardiovascular event. Fatal cardiovascular events include, in addition to non-fatal events codes, different cardiovascular conditions (coded according to the ICD-10-CM as I20–I26, I46, I50, I60–I89, I95, I97–I99), hypertension (I10–I13, I15) and arrhythmias (I44, I45, I47–I49).

The IberScore model computes an estimated risk exposure index based on exposure to risk factors as a weighted average of standardized abnormal values stratifying by age and sex. The model assumes that there is an association between cardiovascular disease and cardiovascular aging, a latent variable whose values may differ from the patient's chronological age. Cardiovascular age might be higher than chronological age when the cardiovascular risk was greater than the baseline risk expected for the patient's age and sex. Exposure to cardiovascular risk factors might influence said deviation. Following these assumptions, the model estimates a cardiovascular biological age from the patient's chronological age modified by the exposure index (to the cardiovascular risk factors), independently by sex. The model assumes that there is a correlation between the estimated cardiovascular biological age and the cardiovascular aging (or cardiovascular health status).⁹ In the internal validation of the equation carried out on a sample of 366,008 workers, the equation showed good discrimination (AUC = 0.8) and correct calibration.⁹

The present study aims to assess an external validation of the IberScore equation in a cohort of patients with no history of cardiovascular disease visited in a primary care center during the years 2008 and/or 2009 and followed up until 2018. Baseline data were recorded during clinical visits made to medical and/or nursing professionals during the study period. The blood samples were processed in the center's reference laboratory, while the rest of the measurements were carried out in the primary care center's consultations.

Non-fatal cardiovascular events included those that appeared in the patient's medical history or reported by the patients themselves, providing the corresponding hospital discharge report to verify the diagnosis after the date

of inclusion in the study. Mortality data were obtained from the Spanish National Institute of Statistics.

Statistical analysis

The missing values in the main variables that integrate the IberScore equation were iteratively and stochastically imputed using the distribution of the observed data to reflect the uncertainty of the real values. For this, a data augmentation algorithm is applied in an iterative way of the Monte Carlo method of Markov chains,¹² combining the effect of a total of 20 imputations following Rubin's rules.¹³

Once the cardiovascular risk was calculated with the IberScore formula for all the subjects of the cohort and imputation techniques were used for replacing the missing values, descriptive analysis of the data was carried out and then the model was calibrated, graphically represented by risk deciles the proportion of expected events and proportion of observed events at 10-year follow-up, stratified by sex. The percentage of observed events was calculated using the non-parametric estimator of the accumulated risk incidence of Nelson-Aalen,¹⁴ which assumes in the calculation the subjects with censored data by taking into account the presence of competitive events (deaths from any other cause). The area under the ROC curve was calculated to assess the discrimination of the model.

Statistical analysis was performed using the STATA/MP 14.2 program.

Results

Characteristics of the study population

A total of 10,085 patients visited during the years 2008 and/or 2009 were included in the study. 55.3% of the subjects were women and the overall mean age was 41.6 (SD 12.96) years. 0.18% of the participants had a diagnosis of type 1 diabetes mellitus and 3.3% of type 2, 9.3% of hypertension and 9.4% of dyslipidemia (Table 1). 300 (2.98%) cardiovascular events (fatal and non-fatal) and 207 (2.05%) competitive events were recorded. In Table 1 it is shown the original mean values within the total number of valid cases, and the imputed mean values for each variable. The two variables with more missing values were glycaemia and CHDL.

The total follow-up of cardiovascular events was 45,562.67 and 57,941.77 person-years, for men and women, respectively.

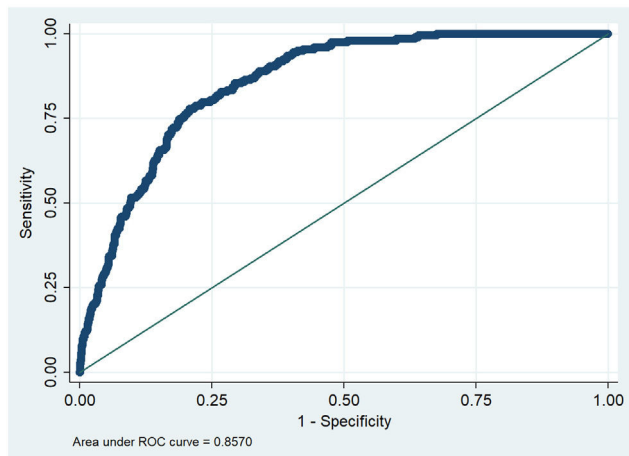
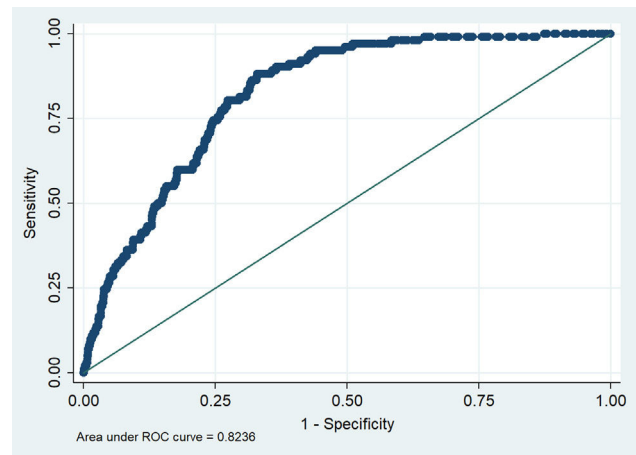
Men showed a mean risk of suffering a fatal or non-fatal cardiovascular events according to IberScore of 17.07% (SD 20.13), with a mean estimated vascular age of more than 4 years higher than the biological age; while women had a mean risk of 7.91% (SD 9.03), with an estimated vascular age of more than 2 years above the biological age.

The area under the ROC curve showed a discrimination index of the model of 0.86 (95% CI 0.84–0.88) in men (Fig. 1) and 0.82 (95% CI 0.79–0.85) in women (Fig. 2). The calibration is presented plotting predicted versus observed risks, taking into account death from causes other than cardio-

Table 1 Clinical characteristics of the study sample.

	Men <i>n</i> = 4511 (44.7%)	Women <i>n</i> = 5574 (55.3%)	Total <i>n</i> = 10,085
Age, mean (SD)	40.95 (14.65)	42.06 (13.07)	41.56 (12.96)
Diabetes mellitus 1 or 2, <i>n</i> (%)	203 (4.50%)	149 (2.67%)	352 (3.49%)
Hypertension, <i>n</i> (%)	630 (13.97%)	696 (12.49%)	1.326 (13.15%)
Dyslipidaemia, <i>n</i> (%)	541 (11.59%)	654 (11.73%)	1.195 (11.85%)
Smoker (active), <i>n</i> (%)	865 (36.84%)	843 (27.10%)	1.708 (16.94%)
BMI [<i>n</i>]	[<i>n</i> = 1691]	[<i>n</i> = 2319]	[<i>n</i> = 4010]
Mean (SD)	27.36 (4.55)	25.97 (5.30)	26.55 (5.04)
Imputed, mean (SD)	26.90 (3.06)	25.57 (3.72)	26.16 (3.50)
Total cholesterol [<i>n</i>]	[<i>n</i> = 2152]	[<i>n</i> = 3129]	[<i>n</i> = 5281]
Mean (SD)	199.02 (40.87)	199.28 (39.84)	199.17 (40.26)
Imputed, mean (SD)	193.37 (30.58)	192.63 (33.13)	192.96 (32.01)
HDL [<i>n</i>]	[<i>n</i> = 1528]	[<i>n</i> = 2124]	[<i>n</i> = 3652]
Mean (SD)	47.92 (11.15)	59.06 (13.99)	54.40 (14.00)
Imputed, mean (DE)	47.08 (7.21)	58.45 (9.34)	53.36 (10.17)
Glycemia [<i>n</i>]	[<i>n</i> = 864]	[<i>n</i> = 1283]	[<i>n</i> = 2147]
Mean (SD)	96.45 (29.16)	89.43 (21.15)	92.25 (24.92)
Imputed, mean (DE)	90.37 (16.46)	86.95 (12.58)	88.48 (14.54)
Non-fatal CV events, <i>n</i> (%)	170 (3.77%)	79 (1.42%)	249 (2.47%)
Fatal CV events, <i>n</i> (%)	28 (0.62%)	23 (0.41%)	51 (0.51%)
Deaths from other causes <i>n</i> (%)	116 (2.57%)	91 (1.63%)	207 (2.05%)

SD: standard deviation; BMI: body mass index; HDL: high density lipoprotein; CV: cardiovascular.

**Figure 1** ROC curve of the IberScore scores for men.**Figure 2** ROC curve of the IberScore scores for women.

vascular causes as competing events and stratified by risk deciles for men (Fig. 3) and for women (Fig. 4).

Discussion

The present study shows the results of the external validation of the IberScore equation in a sample of the population visited in the primary care setting, concluding that the aforementioned model discriminates well, with ROC and Harrel's C indices greater than 0.80 for men and women (values greater than 0.7 indicates correct discrimination). However, this model overestimates the risk of events, more

markedly in women. Therefore, it would be necessary to recalibrate the IberScore equation to enable its use in primary care.

This overestimation was expected, since significant differences were found in the characteristics of the referral population (mostly young, mean age 35 years, and "healthy" active workers: workers on sick leave were excluded from occupational medical examinations) and the external population (attended in primary care, mean age 41.5 years, which includes subjects from age above working age range). In this regard, the effects of the risk factors or their distributions will be different. Other reasons could be that there were some potential predictors are not included

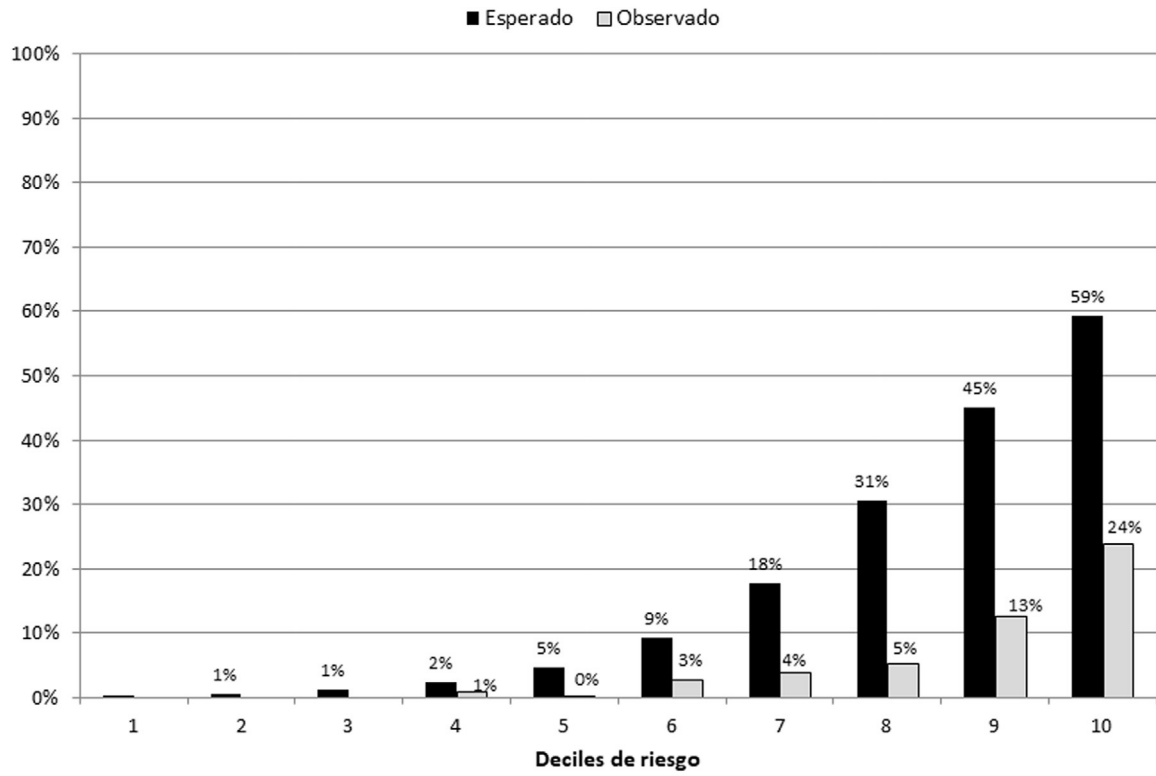


Figure 3 Calibration of the IberScore model based on risk deciles in men.

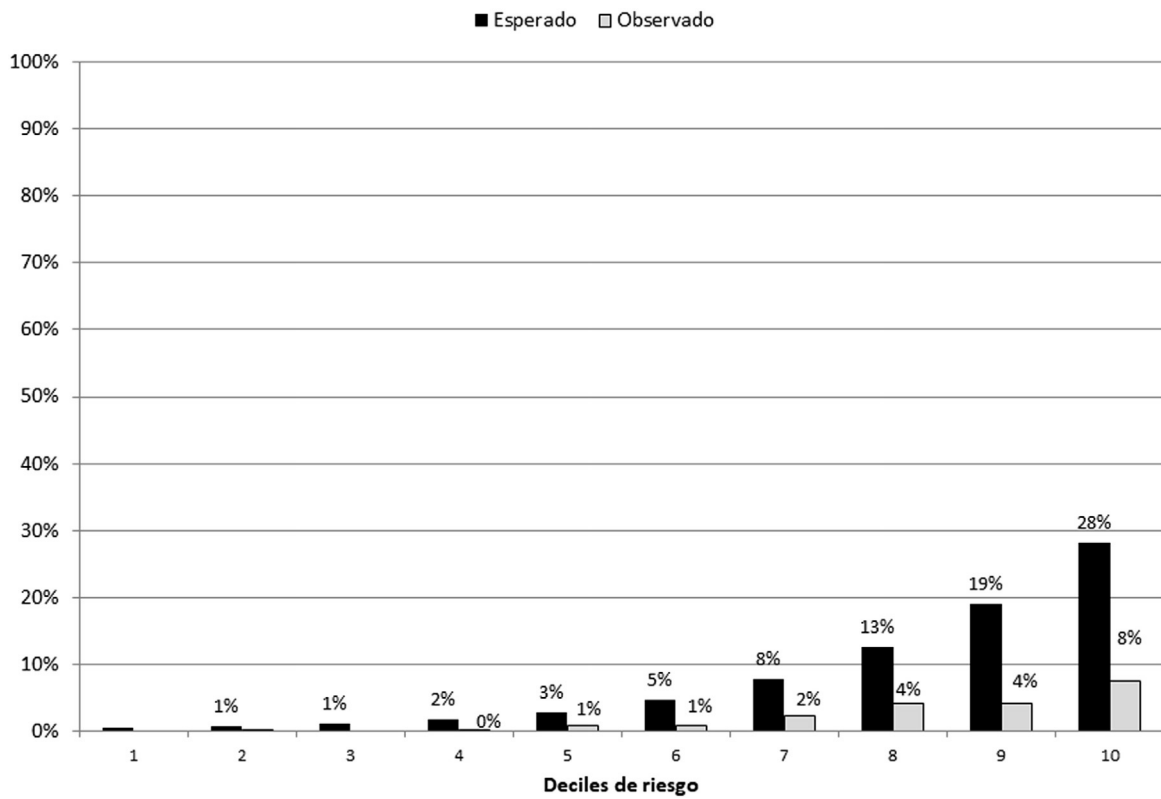


Figure 4 Calibration of the IberScore model based on risk deciles in women.

or to overfitting of the model in the derivation sample, all of which can contribute to an inadequate calibration.

Regarding the high percentage of missing values, although multiple imputation has been applied, it may happen that the distributions, after imputation, do not correctly reflect the real distributions in the population. Although in general, the imputed distributions will resemble those of the non-missing data, the model could not be validated for the population group without missing values.

On the other hand, it is common for a model derived in one population, even having a good discrimination, do not calibrate well in an external population. This was the case of the original Framingham tables, which overestimated the real risk in the Spanish population by more than 2.5 times¹⁵; after calibrating the equation, replacing the prevalence of risk factors and the incidence rates of coronary events for those of Spain,¹⁶ the equation was validated for the Spanish population.¹⁷ This is also the case with the tables of the European SCORE model; the Third European Task Force recommended adapting the model taking into account the level of risk and mortality rates in each country. In that case the calibration was carried out using the Spanish mortality rates and the risk factors of the MONICA-Catalonia cohort, assuming, among other aspects, that the prevalence of risk factors in that cohort could be extrapolated to all of Spain.¹⁸

Regarding strengths, the validation test is carried out in a representative sample of the population attended in primary care as it has been used in previous studies⁹ and with a high rate of long-term follow-up. The IberScore model was derived from a sample of more than 600,000 workers, and it was validated in a sample of more than 360,000 workers, other than those included in the referral cohort. Both cohorts had a 10-year follow-up, showing greater precision in the estimates compared to the SCORE model.¹⁹

To overcome the overestimation of the model, the easiest way to do it would be to recalibrate the IberScore model. If still the recalibration do not fully work, the model parameters will have to be re-estimated in a new population; that is, to derive a new model for primary care but using the same specifications of the IberScore model.

In conclusion, the IberScore model discriminates well in the population attended in primary care but the model overestimates the risk.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: this work was supported by Instituto de Salud Carlos III and Fondo Europeo de Desarrollo Regional-FEDER (grant number PI18/01809), in 2018.

Conflict of interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Acknowledgements

This study is part of a larger research project in cardiovascular risk, the ICARIA (Ibermutua Cardiovascular Risk Assessment) Project, which has been active since 2004. A broad team of general practitioners, specialists and nurses contributed to the collection of demographic and clinical data. The authors wish to thank all the members of the ICARIA Study Group.

References

1. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J*. 2021;42:3227–337.
2. Mancia G, Kreutz R, Brunström M, Burnier M, Grassi G, Januszewicz A, et al. 2023 ESH Guidelines for the management of arterial hypertension The Task Force for the management of arterial hypertension of the European Society of Hypertension: endorsed by the International Society of Hypertension (ISH) and the European Renal Association (ERA). *J Hypertens*. 2023;41:1874–2071. Available from: https://journals.lww.com/jhypertension/fulltext/2023/12000/2023_esh_guidelines_for_the_management_of_arterial.2.aspx [cited 30.11.23].
3. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *Eur Heart J*. 2020;41:111–88.
4. Berry JD, Liu K, Folsom AR, Lewis CE, Carr JJ, Polak JF, et al. Prevalence and progression of subclinical atherosclerosis in younger adults with low short-term but high lifetime estimated risk for cardiovascular disease: the coronary artery risk development in young adults study and multi-ethnic study of atherosclerosis. *Circulation*. 2009;119:382–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/19139385/> [cited 22.11.21].
5. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. *BMJ*. 2011;342:93.
6. Brotons C, Moral I, Fernández D, Puig M, Calvo Bonacho E, Martínez Muñoz P, et al. Estimation of lifetime risk of cardiovascular disease (IBERLIFERISK): a new tool for cardiovascular disease prevention in primary care. *Rev Esp Cardiol*. 2019;72:562–8. Available from: <https://www.revescardiol.org/en-linkresolver-estimation-lifetime-risk-cardiovascular-disease-S1885585718302962> [cited 20.01.23].
7. Brotons C, Moral-Peláez I, Vicuña J, Ameixiras C, Fernández-Labandera C, Sánchez-Chaparro MÁ. Update and validation of the lifetime cardiovascular risk in Spain: IBERLIFERISK2. *Clin Investig Arterioscler*. 2023;35:115–22. <http://dx.doi.org/10.1016/j.arteri.2022.09.003>.
8. Lloyd-Jones DM, Leip EP, Larson MG, D'Agostino RB, Beiser A, Wilson PWF, et al. Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. *Circulation*. 2006;113:791–8.
9. Fernández-Labandera C, Calvo-Bonacho E, Valdivielso P, Quevedo-Aguado L, Martínez-Munoz P, Catalina-Romero C, et al. Prediction of fatal and non-fatal cardiovascular events in young and middle-aged healthy workers: the IberScore model. *Eur J Prev Cardiol*. 2021;28:177–86.
10. Sánchez-Chaparro MA, Román-García J, Calvo-Bonacho E, Gómez-Larios T, Fernández-Meseguer A, Sáinz-Gutiérrez JC, et al. Prevalencia de factores de riesgo vascular en la población laboral española. *Rev Esp Cardiol*. 2006;59:421–30.

11. Calvo Bonacho E, Catalina Romero C, Fernández-Lavandera Ramos C, Cortés Arcas MV, García Margallo MT, Sánchez-Chaparro MA. Clinical usefulness of a predictive model of cardiovascular risk among workers [Utilidad clínica de un modelo predictivo de riesgo cardiovascular en trabajadores]. *Rev Asoc Esp Med Trab.* 2018;27:133–40.
12. Gao S. Analysis of incomplete multivariate data. *Statistical methods in medical research*, vol. 8; 1999. p. 88–9.
13. Campion WM, Rubin DB. Multiple imputation for nonresponse in surveys. *J Mark Res.* 1989;26:485.
14. Aalen O. Nonparametric inference for a family of counting processes. *Ann Stat.* 1978;6:701–26. Available from: <http://www.jstor.org/stable/2958850>
15. Marrugat J, D'Agostino R, Sullivan L, Elosua R, Wilson P, Ordovas J, et al. An adaptation of the Framingham coronary heart disease risk function to European Mediterranean areas. *J Epidemiol Community Health.* 2003;57:634–8.
16. Marrugat J, Solanas P, D'Agostino R, Sullivan L, Ordovas J, Cordon F, et al. Estimación del riesgo coronario en España mediante la ecuación de Framingham calibrada. *Rev Esp Cardiol.* 2003;56:253–61.
17. Marrugat J, Subirana I, Comín E, Cabezas C, Vila J, Elosua R, et al. Validity of an adaptation of the Framingham cardiovascular risk function: the VERIFICA Study. *J Epidemiol Community Health.* 2007;61:40–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/17183014/> [cited 19.01.23].
18. Sans S, Fitzgerald AP, Royo D, Conroy R, Graham I. Calibración de la tabla SCORE de riesgo cardiovascular para España. *Rev Esp Cardiol.* 2007;60:476–85.
19. Brotons C, Moral I, Soriano N, Cuixart L, Osorio D, Bottaro D, et al. Impact of using different SCORE tables for estimating cardiovascular risk. *Rev Esp Cardiol (Engl Ed).* 2014;67:94–100, <http://dx.doi.org/10.1016/j.rec.2013.06.018>.