Received: 17 April 2019 | Revised: 17 June 2019 | Accepted: 27 June 2019

DOI: 10.1111/ijcp.13389

ORIGINAL PAPER

THE INTERNATIONAL JOURNAL OF CLINICAL PRACTICE WILEY

Machine learning to predict cardiovascular risk

Jose A. Quesada[1] | Adriana Lopez-Pineda[1] | Vicente F. Gil-Guillén[1] | Ramón Durazo-Arvizu[2] | Domingo Orozco-Beltrán[1] | Angela López-Domenech[1] | Concepción Carratalá-Munuera[1]

[1]Clinical Medicine Department, Miguel Hernandez University, San Juan de Alicante, Spain

[2]Public Health Department, Stritch School of Medicine, Universidad Loyola Chicago, Maywood, IL, USA

Correspondence
Dr. Adriana Lopez-Pineda, Clinical Medicine Department, Miguel Hernandez University, Ctra. Nnal. 332 Alicante-Valencia s/n 03550 San Juan de Alicante, Alicante, Spain.
Email: adriannalp@hotmail.com

Abstract

Aims: To analyse the predictive capacity of 15 machine learning methods for estimating cardiovascular risk in a cohort and to compare them with other risk scales.

Methods: We calculated cardiovascular risk by means of 15 machine-learning methods and using the SCORE and REGICOR scales and in 38 527 patients in the Spanish ESCARVAL RISK cohort, with 5-year follow-up. We considered patients to be at high risk when the risk of a cardiovascular event was over 5% (according to SCORE and machine learning methods) or over 10% (using REGICOR). The area under the receiver operating curve (AUC) and the C-index were calculated, as well as the diagnostic accuracy rate, error rate, sensitivity, specificity, positive and negative predictive values, positive likelihood ratio, and number needed to treat to prevent a harmful outcome.

Results: The method with the greatest predictive capacity was quadratic discriminant analysis, with an AUC of 0.7086, followed by Naive Bayes and neural networks, with AUCs of 0.7084 and 0.7042, respectively. REGICOR and SCORE ranked 11th and 12th, respectively, in predictive capacity, with AUCs of 0.63. Seven machine learning methods showed a 7% higher predictive capacity (AUC) as well as higher sensitivity and specificity than the REGICOR and SCORE scales.

Conclusions: Ten of the 15 machine learning methods tested have a better predictive capacity for cardiovascular events and better classification indicators than the SCORE and REGICOR risk assessment scales commonly used in clinical practice in Spain. Machine learning methods should be considered in the development of future cardiovascular risk scales.

1 | INTRODUCTION

Cardiovascular diseases (CVD) are the main cause of morbidity and mortality worldwide and, thus, a high public health priority. About 17.7 million people died of CVD in 2015, representing around 31% of all deaths globally.[1] Making correct decisions on optimal treatment is important to address cardiac risk. Over the past few decades, several cardiovascular risk-stratification tools have been developed and widely used to help physicians to identify high-risk individuals for treatment.[2]

Attempts to predict cardiovascular events based on their main determinants, such as biomarkers and health-related behaviours, date back to the Framingham study, which began in the United States in 1948[3] and published its first results in 1959.[4-6] Today, a number of different risk tables are available, usually employing a 10-year outlook. In the United States, these include the Framingham scale,[7] the REYNOLDS risk score,[8] and the American College of Cardiology/America Heart Association (ACC/AHA) risk score.[9] In Europe, the SCORE[10] and QRISK[11] scales predominate, and some countries are using population-specific calibrated models.

These scales assess cardiovascular risk based on survival analyses, with Cox proportional hazards models[7-9,11,12] that follow the Weibull distribution.[10] The models assume a linear relationship between predictors and outcomes, as well as risk predictors that are constant over time, making it difficult to detect complex interactions between them.[13]

Machine learning (ML) is a set of statistical techniques used to predict a quantitative (regression) or categorical (classification) variable. In general, these are supervised techniques used to predict one variable in a group of new observations. There is no intent to "explain" the behaviour of an outcome in the presence of different factors, but only to make the best prediction possible in the face of non-linear relationships and complex interactions. Thus, these techniques are applicable to high-dimensional problems, but some of them can also be considered black box methods in which they tend to conceal relationships between predictors and outcomes.[14,15]

ML encompasses common regression methods such as linear regression, logistic regression and Cox regression, but also lesser known models, including penalized regression, principal component regression, cluster analysis, discriminant analysis, nearest neighbour methods, neural networks, support vector machines, classification trees, bagging, boosting, and random forest regression, among others. These techniques are not new, but their use has seen an uptick because of recent advances in calculation capacity.[16]

Researchers have now begun to apply some of these techniques to predicting morbimortality in CVD. In Canada, one study compared five ML methods for predicting 30-day mortality in patients admitted for myocardial infarction or congestive heart failure (CHF),[17] while investigators in the United States have compared neural networks versus the Framingham risk score for CVD risk prediction,[18] assessed five methods for predicting mortality in patients with CHF,[19] compared the power of two methods for predicting any cardiovascular event,[20] and assessed diverse clinical risk scales for predicting coronary heart disease.[21] There have also been studies elsewhere: in South Korea, comparing six risk prediction methods for CVD[22]; in China, examining six methods for predicting intrahospital mortality in patients with myocardial infarction[23]; and in the UK, comparing four methods versus the AHA/ACC scale.[24]

In Spain, clinicians tend to use the SCORE risk charts, on recommendation from the Spanish Society of Cardiology and the Spanish Committee for Cardiovascular Prevention.[25] In some regions of Spain, the REGICOR[12] risk score—a local adaptation of the Framingham scale—is used. Clinicians are advised to assess risk at 5-year intervals in patients with at least one important risk factor, such as exposure to tobacco, diabetes, hypertension, or dyslipidaemia.[26] However, SCORE overestimates 10-year cardiovascular mortality in the Spanish population,[27] resulting in false positives that lead to unnecessary treatment as well as false negatives that fail to reflect individuals' real risk.[24] In addition, SCORE for low-risk countries seems to provide acceptable results when applied in a high-risk country, whereas the version adapted for high-risk countries overestimates true risk[28]. With this in mind, it is important to obtain cardiovascular risk predictions with both high sensitivity and high specificity, as well as to perform risk estimations with a shorter outlook, such as 5 years. Thus, this study aimed to compare the predictive capacity of 15 machine learning methods versus the SCORE scale and REGICOR calibration of Framingham scale in a clinical practice cohort followed up for 5 years in Spanish primary care.

What's known

- Current cardiovascular risk scores make difficult to detect complex interactions between predictors.
- Researchers have now begun to apply machine learning to predict morbimortality in cardiovascular disease.

What's new

- There are machine learning methods which showed better predictive capacity than the most commonly used methods for estimating cardiovascular risk in Spain.
- Machine learning methods should be considered in the development of future cardiovascular risk scales.

2 | MATERIALS AND METHODS

We designed an analytical observational study of the ESCARVAL RISK clinical practice cohort,[29] following patients in primary care from January 2008 to December 2012. The study cohort consisted of patients aged 40 years or older and diagnosed with hypertension, dyslipidaemia, or diabetes. Study variables were collected from patients' e-health records during the inclusion phase. The primary outcome of interest was all-cause mortality or hospital admission because of stroke (International Classification of Diseases 9 [ICD-9] codes 430-438, 444) or ischemic cardiopathy (ICD-9 410-414). Hospitalisations and CVD mortality were assessed by reviewing hospital records and annual mortality data during follow-up.

To calculate risk using all the ML methods and for the SCORE assessment, we included only the variables of age (years), sex (male/female), total cholesterol (mg/dL), systolic blood pressure (mm Hg), and tobacco use (yes/no); for the REGICOR scale, we additionally included diastolic blood pressure (mm Hg), high-density lipoprotein (HDL) cholesterol (mg/dL), and the presence of diabetes (yes/no). We did not consider other predictors, as our intent was not to build the best predictive model but to compare different techniques with the SCORE and REGICOR scales under the same conditions.

Statistical analyses were undertaken using the R statistical program v.3.5.0 (R project). With regard to the predictive procedures, we used the SCORE - European Low Risk Chart, which is most appropriate for populations, such as the Spanish one, that are at low cardiovascular risk,[10] and the REGICOR risk score, as a local adaptation of the Framingham risk score.[12] We also tested 15 ML methods (R function used): Cox regression (coxph), penalized Cox regression (glmnet), Naive Bayes (naiveBayes), linear discriminant analysis (lda), quadratic

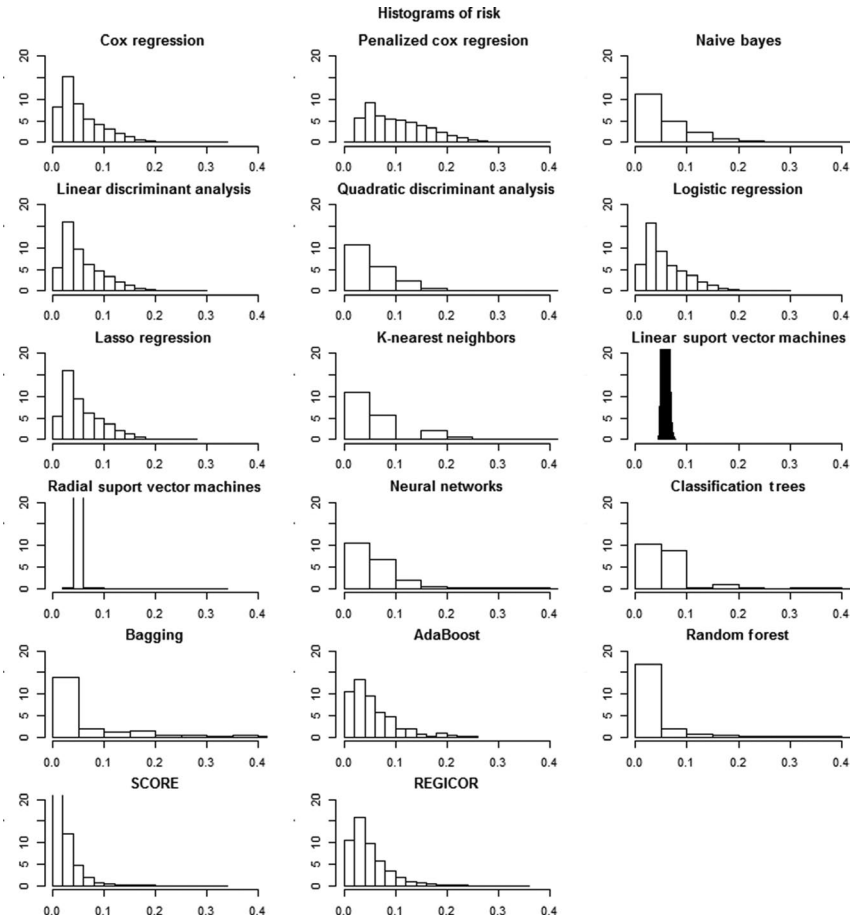

FIGURE 1 Histograms for 5-year risk of cardiovascular event by prediction method

discriminant analysis (qda), logistic regression (glm), penalized Logistic regression (glmnet), K-nearest neighbours (kknn), linear support vector machines (svm), radial support vector machines (svm), single-hidden-layer neural network (nnet), classification trees (rpart), bagging (bagging), AdaBoost (AdaBoostM1), and random forest (randomForest). We also carried out tests with random survival forest (rfsrc) but ran into problems of convergence. Where necessary, we standardised the variables and performed prior optimisations of control parameters as follows: for penalized Cox regression with the function cv.glmnet (nfolds = 10 and alpha = 1), for penalized logistic regression with the function cv.glmnet (nfolds = 10 and alpha = 0.9), for K-nearest neighbours with the function train.kknn (kernel = "rectangular"), for linear and radial support vector machines with the function tune.svm, for single-hidden-layer neural network with the function train (method = "cv", number = 5) and for classification trees with the function rpart (cp=0.0003). These methods are described in detail elsewhere.[14,16,30]

To estimate the predictive and classification capacity rigorously, we randomly divided the sample into two parts: 70% were used for the estimation of the models (training) and 30% for validation (testing). We assessed predictive capacity by calculating the C-index for the Cox and penalized Cox regressions, as the time to the event is considered in the model, and by analysing the area under the receiver operating curve (AUC) in the rest of the analyses, including SCORE and REGICOR. The 95% confidence interval (CI) was determined in all cases.

To assess the quality of the classification, we used a 5% risk cutoff (indicating high cardiovascular risk) for SCORE and for all the ML methods, and a 10% cutoff for REGICOR.[26,31] Applying this cutoff on the testing sample, we calculated the diagnostic accuracy rate, the error rate, sensitivity and specificity with 95% CIs, positive and negative predictive values, the positive likelihood ratio with 95% CI, and the number needed to treat to prevent a harmful outcome. Histograms are presented for cardiovascular risk according to each method used for calculation.

3 | RESULTS

The study sample included 38 527 patients (18 778 men, 48.7%) with a mean age of 55.8 years. The training sample consisted of 26 853 (70%) patients, and the testing sample, 11 674 (30%) patients.

Figure 1 shows the histograms for estimated cardiovascular risk for each method; the worst-performing methods in the distribution of risks were the linear and radial support vector machines and bagging. SCORE and REGICOR present similar risk histograms, although the density of risks is higher in the latter.

Table 1 shows the predictive and classification indicators. This table is organised in descending order of AUC values, so the method with the greatest predictive capacity is Quadratic Discriminant Analysis, with an AUC of 0.7086, followed by Naive Bayes and

TABLE 1 Estimated predictive and classification indicators by prediction method

Method	AUC	95% CI (AUC)	AR	ER	SE.	95% CI (SE)	SP.	95% CI (SP)	PPV	NPV	LR+	95% CI (LR)	NNT
QDA	0.7086	(0.6885-0.7288)	0.570	0.430	0.736	(0.703-0.769)	0.559	(0.550-0.568)	0.093	0.972	1.669	(1.588-1.754)	10.7
Naive Bayes	0.7084	(0.6884-0.7284)	0.591	0.409	0.718	(0.684-0.752)	0.583	(0.574-0.592)	0.096	0.971	1.722	(1.635-1.814)	10.4
Neural networks	0.7042	(0.6840-0.7245)	0.559	0.441	0.743	(0.710-0.776)	0.548	(0.539-0.557)	0.092	0.972	1.644	(1.566-1.726)	10.9
AdaBoost	0.7006	(0.6801-0.7212)	0.580	0.420	0.721	(0.687-0.755)	0.571	(0.562-0.580)	0.094	0.971	1.681	(1.596-1.770)	10.7
LDA	0.7004	(0.6801-0.7207)	0.568	0.432	0.728	(0.694-0.762)	0.558	(0.549-0.567)	0.092	0.971	1.647	(1.566-1.732)	10.9
Logistic regression	0.7001	(0.6798-0.7204)	0.570	0.430	0.728	(0.694-0.762)	0.561	(0.552-0.570)	0.093	0.971	1.658	(1.576-1.744)	10.8
Penalized Logistic regression	0.7001	(0.6797-0.7204)	0.566	0.434	0.731	(0.698-0.764)	0.556	(0.547-0.565)	0.092	0.971	1.646	(1.565-1.731)	10.9
Cox regression	0.6911	(0.6692-0.7131)	0.560	0.440	0.347	(0.311-0.383)	0.573	(0.564-0.582)	0.048	0.934	0.813	(0.732-0.904)	21.0
Penalized Cox regression	0.6911	(0.6513-0.7309)	0.215	0.785	0.527	(0.489-0.565)	0.196	(0.189-0.203)	0.039	0.871	0.655	(0.610-0.704)	25.8
Classification trees	0.6494	(0.6300-0.6689)	0.540	0.460	0.725	(0.691-0.759)	0.528	(0.519-0.537)	0.086	0.969	1.536	(1.460-1.615)	11.6
REGICOR	0.6342	(0.6142-0.6542)	0.876	0.124	0.155	(0.128-0.182)	0.920	(0.915-0.925)	0.107	0.947	1.938	(1.608-2.336)	9.4
SCORE	0.6333	(0.6131-0.6536)	0.858	0.142	0.182	(0.153-0.211)	0.900	(0.894-0.906)	0.101	0.947	1.820	(1.536-2.156)	9.9
Random forest	0.6281	(0.6059-0.6503)	0.809	0.191	0.307	(0.272-0.342)	0.840	(0.833-0.847)	0.106	0.952	1.919	(1.700-2.166)	9.5
K-nearest neighbours	0.6002	(0.5791-0.6214)	0.557	0.443	0.612	(0.575-0.649)	0.554	(0.545-0.563)	0.078	0.959	1.372	(1.288-1.462)	12.8
Bagging	0.5886	(0.5671-0.6102)	0.682	0.318	0.443	(0.406-0.480)	0.696	(0.687-0.705)	0.082	0.953	1.457	(1.333-1.593)	12.1
Radial SVM	0.5209	(0.4980-0.5439)	0.062	0.938	0.994	(0.988-1.000)	0.004	(0.003-0.005)	0.058	0.925	0.998	(0.992-1.004)	17.3
Linear SVM	0.4096	(0.3884-0.4308)	0.109	0.891	0.885	(0.861-0.909)	0.061	(0.057-0.065)	0.055	0.896	0.942	(0.916-0.968)	18.2

Abbreviations: 95% CI: 95% confidence interval; AR: accuracy rate; AUC: area under the receiver operating curve; ER: error rate; LDA: linear discriminant analysis; LR+: positive likelihood ratio; NNT: number needed to treat; NPV: negative predictive value; PPV: positive predictive value; QDA: quadratic discriminant analysis; SE: sensitivity; SP: specificity; SVM: support vector machines.

Neural networks, with AUCs of 0.7084 and 0.7042, respectively. Of the 15 ML methods analysed, three present AUC values of less than 0.6. In descending order, these are: Bagging, Radial Support Vector Machines, and Linear Support Vector Machines; the latter two did not show significant predictive capacity with our data. REGICOR ranked 11th and SCORE 12th, with AUCs of 0.6342 and 0.6333, respectively.

With regard to classification indicators, the method garnering the highest diagnostic accuracy rate was REGICOR, at 87.6%, followed by SCORE at 85.8%. Although specificity was high with both methods (≥ 90%), sensitivity was very low (< 20%). The seven methods with the highest AUC values (all > 0.70) presented similar sensitivities (> 70%) and specificities (> 55%). These seven methods showed a 7% higher predictive capacity (AUC) than SCORE and REGICOR.

4 | DISCUSSION

The results of this study suggest that 10 of the 15 machine learning methods tested have a better predictive capacity for cardiovascular events and better classification indicators than the SCORE and REGICOR risk assessment scales commonly used in clinical practice in Spain.

REGICOR and SCORE showed similar predictive and classification capacities, even though REGICOR considers three additional variables (diabetes, diastolic blood pressure, and HDL cholesterol). However, there are 10 ML techniques with better predictive capacity than either of these risk assessment tools. Quadratic Discriminant Analysis was the best technique according to our analyses, outperforming SCORE and REGICOR with a 7.5% difference, while the top seven methods showed about a 7% improvement over the usual risk scores (Table 1).

To our knowledge, this is the first study comparing the predictive capacity of ML methods with SCORE and REGICOR in a routine clinical practice cohort in Spain. Another study compared the algorithm for risk estimation proposed by the ACC/AHA versus four ML methods in a British cohort, showing that ML procedures improved cardiovascular risk prediction by up to 3.6% (for neural networks)[24]. In another investigation using data from the MESA study, researchers found that ML methods had a 7.7% higher predictive capacity than standard methods for assessing the 10-year risk of atherosclerotic cardiovascular disease[21]. There is also evidence that ML methods improve predictive capacity for intrahospital mortality[23] and CHF mortality[19] compared with standard tools. With regard to the worst prediction methods in this context, our results are consistent with those reported by Kim et al.[22] in finding that support vector machines do not produce a significant predictive value. Moreover, they come at a high computational cost, requiring prior optimisation of control parameters.

As shown by our analyses, ML techniques represent an improvement over standard risk assessment scales in terms of predictive capacity when using the same variables. We have thus entered a new technological era in which calculation capacity is sufficient to apply complex ML methods in large populations, based on multiple and interacting predictors, using data collected from e-health records[13]. It is therefore possible to fit ML models that achieve high predictive and classification capacity, improving the low sensitivity of traditional scales that fail to identify many people at high risk of cardiovascular events and enabling better control of their individual risk factors, with positive implications for the prevention of possible cardiovascular events.

4.1 | Strengths and limitations

The use of the same predictors and the same population for both the ML methods and the SCORE risk chart ensures the comparability of the predictive and classification indicators. The validation of the methods in a random testing sample—uninvolved in the estimation of the parameters—was used to ensure the comparability of the AUC.

One limitation could be that the SCORE and REGICOR endpoints are not exactly the same, as SCORE is used to predict cardiovascular mortality and REGICOR, cardiovascular morbimortality. However, both scales are of great interest, as these are the most frequently used instruments to estimate general cardiovascular risk in Spain. To calculate classification indicators, we used a 5% risk cutoff for SCORE and a 10% risk cutoff for REGICOR, as recommended by the authors and minimising the conceptual difference between the endpoints. The endpoint of the ESCARVAL RISK study is cardiovascular morbidity and all-cause mortality, which is closer to the REGICOR endpoint.

On the other hand, some of these methods might not provide an explicit relationship between predictors and outcomes and could be black box methods: Neural Networks, Support Vector Machine, K-Nearest Neighbours, Bagging, AdaBoost y Random Forest. This might be a limitation for clinical practice. However, the rest of methods (discriminant analysis, Naive Bayes, logistic and Cox regression, classification) provide this relationship, therefore, the risk score for each predictor could be calculated for clinical practice. Only two of the top seven top machine learning methods could be considered black box methods.

5 | CONCLUSIONS

The top seven machine learning methods were Quadratic Discriminant Analysis, Naive Bayes, Neural Networks, AdaBoost, Linear Discriminant Analysis, Logistic Regression and Penalized Logistic Regression, which showed better predictive capacity than the usual risk scores. Thus, ML methods should be considered in the development of future cardiovascular risk scales.

DISCLOSURE

None.

ORCID

Adriana Lopez-Pineda https://orcid.org/0000-0002-2117-0178

Concepción Carratalá-Munuera https://orcid.org/0000-0002-1303-6294

REFERENCES

1. World Health Organization. World Health Statistics 2017: monitoring health for the SDGs. http://www.who.int/gho/publications/world_health_statistics/2017/en/. Accessed September 18, 2018.
2. Ker JA, Oosthuizen H, Rheeder P. Decision-making using absolute cardiovascular risk reduction and incremental cost-effectiveness ratios: a case study. *Cardiovasc J Afr*. 2008;19(2):97–101.
3. Framingham Heart Study. 2018. https://www.framinghamheartstudy.org/. Accessed September 18, 2018.
4. Dawber TR, Kannel WB, Revotskie N, Stokes JI, Kagan A, Gordon T. Some factors associated with the development of coronary heart disease. Six years' follow-up experience in the Framingham Study. *Am J Public Health*. 1959;49:1349–1356.
5. Gordon T, Moore FE, Shurtleff D, Dawber TR. Some methodologic problems in the long-term study of cardiovascular disease: Observations on the Framingham Study. *J Chronic Dis*. 1959;10:186–206.
6. Stokes JI, Dawber TR. The silent coronary: the frequency and clinical characteristics of unrecognized myocardial infarction in the Framingham study. *Ann Intern Med*. 1959;50:1359–1369.
7. D'Agostino RB, Ramachandran S, Pencina MJ, et al. General Cardiovascular Risk Porfile for use in primary care. The Framingham Heart Study. *Circulation*. 2008;117:743–753.
8. Ridker P, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: The Reynolds risk score. *JAMA*. 2007;297:611–619.
9. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129:S49–S73.
10. Conroy RM, Pyorala K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart. J*. 2003;24:987–1003.
11. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ*. 2007;335:136.
12. Marrugat J, Solanas P, D'Agostino R, et al. Estimación del riesgo coronario en España mediante la ecuación de Framingham calibrada. *Rev Esp Cardiol*. 2003;56:253–261.
13. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J*. 2017;38:1805–1814.
14. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data mining, Inference and Prediction*, 2nd edn. California/USA: Springer; 2008.
15. Obermeyer Z, Emanuel EJ. Predicting the future big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375:1216–1219.
16. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. California/USA: Springer; 2013.
17. Austin P, Lee DS, Steyerberg EW, Tu JV. Regression trees for predicting mortality in patients with cardiovascular disease: What improvement is achieved by using ensemble-based methods? *Biom J*. 2012;54(5):657–673.
18. Narain R, Saxena S, Goyal AK. Cardiovascular risk prediction: a comparative study of Framingham and quantum neural network based approach. *Patient Prefer Adherence*. 2016;10:1259–1270.
19. Panahizar M, Taslimitehrani V, Pereira N, Pathak J. Using EHRs and machine learning for heart failure survival analysis. *Stud Health Technol Inform*. 2015;216:40–44.
20. Ambale-Venkatesh B, Yang X, Wu CO, et al. Cardiovascular event prediction by machine learning. The multi-ethnic study of atherosclerosis. *Circ Res*. 2017;121:1092–1101.
21. Nakanishi R, Day D, Commandeur F, et al. Machine Learning in predicting coronary heart disease and cardiovascular disease events: results from the multi-ethnic study of atherosclerosis (MESA). *JACC*. 2018;71:11.
22. Kim J, Kang U, Lee Y. Statistics and deep belief network-based cardiovascular risk prediction. *Health Inform Res*. 2017;23:169–175.
23. Li X, Liu J, Yang J, Xie G, Xu M, Yang Y. Using Machine Learning Models to Predict In- ospital Mortality for ST-Elevation Myocardial Infarction Patients. *Stud Health Technol Inform*. 2017;245:476–480.
24. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE*. 2017;12:e0174944.
25. Jiménez Navarro MF. Comentarios a la guía ESC 2016 sobre prevención de la enfermedad cardiovascular en la práctica clínica. *Rev Esp Cardiol (Engl Ed)*. 2016;69:894–899.
26. Royo-Bordonada MI, Armario P, Lobos JM, et al. Adaptación española de las guías europeas 2016 sobre prevención de la enfermedad cardiovascular en la práctica clínica. *Rev Esp Salud Pública*. 2016;90:24.
27. Baena-Díez JM, Subirana I, Ramos R, et al. Validity assessment of low-risk SCORE function and SCORE function calibrated to the Spanish population in the FRESCO cohorts. *Rev Esp Cardiol*. 2018;71:274–282.
28. Baena-Diez JM, Ramos R, Marrugat J. Predictive value of cardiovascular risk functions: limitations and future potential. *Rev Esp Cardiol*. 2009;9(Suppl):4B–13B.
29. Orozco-Beltran D, Gil-Guillen VF, Redon J, et al. Lipid profile, cardiovasculardisease and mortality in a Mediterranean high-riskpopulation: The ESCARVAL-RISK study. *PLoS ONE*. 2017;12:e0186196.
30. Kuhn M, Johnson K. *Applied Predictive Modeling*, 1st edn. Springer; 2013.
31. Sans S, Fitzgerald AP, Royo D, Conroy R, Graham I. Calibracion de la tabla SCORE de riesgo cardiovascular para España. *Rev Esp Cardiol*. 2007;60:476–485.

How to cite this article: Quesada JA, Lopez-Pineda A, Gil-Guillén VF, et al. Machine learning to predict cardiovascular risk. *Int J Clin Pract*. 2019;73:e13389. https://doi.org/10.1111/ijcp.13389