The Role of Dietary and Socioeconomic Status Assessment on the Predictive Ability of the HellenicSCORE

Konstantinos Vlismas, Demosthenes B. Panagiotakos, Christos Pitsavos, Christina Chrysohoou, Yannis Skoumas, Vassilios Stavrinos, Christodoulos Stefanadis

1Harokopio University, Department of Nutrition Science-Dietetics, 2First Cardiology Department, School of Medicine, University of Athens, Athens, Greece

Introduction: The optimal performance of cardiovascular disease (CVD) risk models in various populations (such as the Framingham Heart Sheet or the ESC SCORE) is of major interest in risk prediction modeling nowadays. We evaluated whether the inclusion of socioeconomic status (SES) in the HellenicSCORE would increase the accuracy of prediction, irrespectively of dietary information and the classical CVD risk factors.

Methods: Data from 1514 men and 1528 women (age >18 years), who were free of known CVD on enrolment in 2001-02, were studied (the ATTICA study). Five years later a follow up was performed and the development of CVD was defined (WHO-ICD-10 criteria). As SES indicators, education status and mean annual income were recorded, and a special SES 3-class index was calculated (low, moderate and high). The MedDietScore, which incorporates the inherent characteristics of the Mediterranean diet, was used as a dietary assessment tool, while the HellenicSCORE, which reflects the level of CVD risk factors, was also calculated. Additive logistic regression models were used to test the additive effect of SES and dietary assessment on the predictive ability of the HellenicSCORE.

Results: SES assessment did not improve the predictive ability of the estimated risk model compared to the model that included the HellenicSCORE, physical activity status, waist-to-hip ratio, diabetes and family history of CVD. Additionally, SES did not improve the predictive ability of the estimated risk model even when dietary assessment was added to the above model.

Conclusions: Socioeconomic status does not improve the predictive ability of a CVD risk model, even when dietary information is also taken into account.

The cardiovascular disease (CVD) epidemic still constitutes the leading cause of death worldwide. It is a general belief nowadays that prevention strategy should not only focus on people who are at high risk for developing CVD, but needs to be used in a wider perspective in terms of reduction of risk factors, if CVD is to be substantially reduced. Effective prevention is primarily based on the accurate identification of individuals at risk, through the global assessment of risk factors. Many risk prediction models have been developed over the years in order to provide an overall assessment of CVD risk. Two of the most well-known CVD predictive risk models are the Framingham risk score and the ESC SCORE (European Society of Cardiology: Systematic Coronary Risk Evaluation) system. Other predictive models to measure the risk of future CVD events are: the PROCAM (Prospective Cardiovascular Münster) study, which uses neural net-
works; the ASSIGN risk score, involving family history of CVD, developed in Scotland; the QRISK1 and QRISK2 which use a population-based clinical research database in the UK; and the accompanying World Health Organization/International Society of Hypertension (WHO/ISH) risk prediction charts. These risk prediction models use the classical CVD risk factors (i.e. cholesterol and systolic blood pressure levels, age, gender, and smoking habits), but very few attempts have been made to include either more lifestyle-related CVD risk factors, such as diet, or other strong predictors of health, such as socioeconomic status (SES).

Recent recommendations by the Fourth Joint Task Force of the ESC for the primary prevention of CVD in clinical practice have emphasized the importance of a “multifactorial approach in the assessment and management of CVD, including dietary habits, as dietetics is considered an integral part of patient risk management.” Moreover, according to the World Health Organization, emphasis should be placed on lifestyle modification, including physical activity, moderation of alcohol intake, increased fresh fruit and vegetables and reduced saturated fat in the diet, weight loss in the overweight, reduction of dietary sodium intake, and increased potassium intake. Recently, Panagiotakos et al showed that the inclusion of dietary habits in the Hellenic version of the ESC SCORE model, the HellenicSCORE, increases accuracy and reduces bias in the estimations of future cardiac events.

In addition, it has now been well established that SES is an important determinant of health. Moreover, the association between the classic SES indicators (i.e. education, income and occupation) and CVD is well established, but the predictive ability of such indicators has not been thoroughly examined. One of the few attempts was made by the QRISK scores, where SES was included in the form of social deprivation (area-based index) to avoid social gradients in health outcomes. Recently, Fiscella et al included SES in the Framingham risk score in order to reduce bias through the underestimation of coronary heart disease (CHD) risk in people with a low SES. Moreover, it is well known that SES can potentially modify people’s dietary habits. In general, people with a lower SES tend to practice less healthy behaviors, including unhealthy dietary habits, consequently leading to the development of chronic diseases such as CVD. Finally, there is strong evidence of an interaction between SES and diet in relation to CVD risk factors and CVD outcome.

Therefore, in this study it was assessed whether the inclusion of SES in a CVD risk prediction model would increase the accuracy of prediction, irrespectively of dietary information and the classical CVD risk factors as evaluated through the HellenicSCORE.

Methods

Study design and participants

To test the research hypothesis of this work, the “ATTICA” epidemiological study database was used. The “ATTICA” study started as a nutrition and health survey of the Greek population (during 2001-2002), and in 2006 performed the first follow up. The sampling was carried out in the region of Attica, which includes 78% urban and 22% rural areas, where Athens is the major metropolis. During the enrolment period, 4056 inhabitants from the above area were selected. Of these, 3042 consented to participate (75% participation rate); 1514 of the participants were men and 1528 were women. Further details about the aims, design and methods of the ATTICA epidemiological study may be found elsewhere in the literature.

Measurements

SES was assessed through years of education and mean annual income of the family (through self reports), during the last three years. For people in the family who were not working, we used the average family income, while for unemployed individuals we used the basic monthly allowance they received from the Social Service Office. To provide a more composite estimator of people’s SES, a special “socioeconomic” index was developed by multiplying the years of school in each individual by their mean annual income. The range of the index was from 5500 to 880,000: 1st tertile, bad SES,<172,800; 2nd tertile, moderate SES, 172,800 to 240,000; and 3rd tertile, good SES >240,000. This index was normally distributed according to the Kolmogorov-Smirnov criterion.

In addition, the baseline evaluation of the ATTICA study included information about clinical and biological measurements, personal and family history of hypertension, hypercholesterolemia and diabetes, family history of CVD, dietary and other lifestyle habits, such as smoking and physical activity status.
Participants’ baseline risk of CVD risk was calculated globally using the HellenicSCORE model. Annual CVD mortality rates for Greece were obtained from the World Health Organization mortality database for 2002, while the average age- and sex-specific levels of systolic blood pressure, total cholesterol and smoking prevalence were predicted using data from the ATTICA study.

The evaluation of the nutritional habits of the ATTICA study participants was based on a validated semi-quantitative food-frequency questionnaire, the EPIC-Greek questionnaire, which was kindly provided by the Unit of Nutrition of Athens Medical School. All participants were asked to report the average intake (per week or day) of several food items that they had consumed during the last 12 months. Then, the frequency of consumption was quantified approximately in terms of the number of times a month a food was consumed. Any type of alcohol consumption was measured in wineglasses (100 ml) and quantified by ethanol intake (in g per day). In order to describe overall diet the MedDietScore (range 0-55) was used. Higher values of this score indicate better adherence to the Mediterranean diet.

**Follow-up evaluation**

During 2006, the ATTICA study’s investigators performed the 5-year follow up. Of the 3042 participants initially enrolled, 1012 men and 1035 women were still alive at the time of the follow up, while 32 (2.1%) men and 22 (1.4%) women died during the 5-year period. The rest of the participants (941) were lost to follow-up (69% participation rate). Of the individuals who did not participate in the re-examination, 75% were not found because of missing or wrong addresses and telephone numbers, and the rest refused to be re-examined. No significant differences were observed in the baseline characteristics between those who participated in the follow up and those who did not, regarding the distribution of age (p=0.78), sex (p=0.99), years of school (p=0.67), presence of hypertension (p=0.12), diabetes (p=0.27), hypercholesterolemia (p=0.10), or obesity (p=0.54).

**Statistical analysis**

Incidence rates of CVD were calculated as the ratio of new cases developed during the preceding years to the number of people participating in the follow up. Continuous variables are presented as mean values ± standard deviation and categorical variables are presented as frequencies. Associations between categorical variables were tested using the chi-square test. Comparisons of mean values of normally distributed variables between those who developed an event and the rest of the participants were performed using Student’s t-test. For some continuous variables that were not normally distributed the Mann-Whitney non-parametric test was applied to evaluate the differences in the distributions of the skewed variables between those who developed a CVD event and the rest of the participants. The associations between participants’ baseline CVD risk factor levels (as assessed through the calibrated HellenicSCORE equations), dietary habits (as assessed through the MedDietScore), and socioeconomic status (as assessed through the SES index), and the development of CVD were evaluated using nested logistic regression models. Then the \(-2\log (\text{Likelihood})\) ratio between the full and reduced models and the chi-square of each model were calculated (model’s chi-square measures the improvement in fit that the explanatory variables make compared to the model without this variable). The Nagelkerke “pseudo” \(R^2\) was also used to interpret the level of explained “variance” of each model. Nagelkerke \(R^2\) is a an attempt to imitate the interpretation of the common multiple \(R^2\) based on the log (Likelihood) of the final model vs. log (Likelihood) for the baseline model, with a modification in order to ensure its values lie within the interval [0, 1]; however, it could better be described as an attempt to measure strength of association. The Hosmer-Lemeshow statistic also evaluated the model’s goodness-of-fit (fit was considered adequate if computed \(p>0.05\), indicating model prediction is not significantly different from observed values). The correct classification rate was calculated as the ratio of the predicted events divided by the true events, and the ratio of the predicted event-free to the true event-free participants. All reported \(p\)-values are based on two-sided tests. SPSS version 14 software (Statistical Package for Social Sciences, SPSS Inc, Chicago, IL, USA) was used for all the statistical calculations.

**Results**

Table 1 presents the demographic, clinical and behavioral characteristics of the participants according to 5-year CVD status, and shows that people who developed CVD exhibited a lower adherence to the Mediterranean diet (i.e. lower MedDietScore, \(p=0.001\)).
Regarding the analysis of SES, there was a strong association between CVD events and years of school (p<0.001) as well as SES status (p<0.001), meaning that people who developed CVD were more likely to be less well educated and to appear in the lower SES index level. No significant association was found between CVD events and financial status (p=0.44). In addition, when the HellenicSCORE was divided into 7 classes (i.e. <1%, 1%, 2%, 3%-4%, 5%-9%, 10%-14% and >14% risk of developing fatal CVD events during the next 10 years), there was a strong association between those who developed a CVD event and classes of HellenicSCORE (p<0.001; Figure 1).

Table 2 shows the results from the main research hypothesis tested (i.e. that the inclusion of SES in a CVD risk prediction model would increase the accuracy in prediction, irrespectively of dietary information and the classical CVD risk factors). In particular, a model was estimated (core model, model 0) with only the HellenicSCORE as a measure of global CVD risk24 and, as expected, it was highly significant in predicting future CVD events, with an overall correct classification rate equal to 91.4%. When the previous model was adjusted for physical activity status, waist-to-hip ratio, diabetes and family history of CVD (model 1) the HellenicSCORE remained highly significant, while the explained variability (Nagelkerke $R^2$) increased by 22.5% (i.e. from $R^2=15.5\%$ for model 0 to $R^2=19.0\%$ for model 1). However, when SES was added (model 2) it did not significantly improve the predictive ability of the model (odds ratio, OR=0.48, 95% confidence interval, CI: 0.16-1.42, p=0.18). In contrast, when MedDietScore was added to model 1 (model 3), the additive predictive ability of the model exhibited borderline significance (OR=0.96, CI: 0.91-1.00, p=0.06), while the explained variability of the model (Nagelkerke $R^2$) increased by 3.5% (i.e. from $R^2=19.0\%$ for model 1 to $R^2=19.7\%$ for model 3). Finally, when SES and...
CVD Risk Models, Diet and Socioeconomic Status

Table 2. Results from multiple logistic regression analysis that evaluated the effect of socioeconomic status (SES), in addition to the HellenicSCORE, the MedDietScore and other covariates* in predicting 5-year cardiovascular disease events.

<table>
<thead>
<tr>
<th>Model 0 (unadjusted)</th>
<th>Odds ratio</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>HellenicSCORE (per 1% increase)</td>
<td>1.21</td>
<td>1.17-1.25</td>
</tr>
<tr>
<td>Model's chi-square</td>
<td>141.6, df=1, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Nagelkerke R² and -2LogL</td>
<td>15.5% and 1020</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 1 (adjusted*)</th>
<th>Odds ratio</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>HellenicSCORE (per 1% increase)</td>
<td>1.19</td>
<td>1.12-1.26</td>
</tr>
<tr>
<td>Model's chi-square</td>
<td>83.9, df=5, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Nagelkerke R² and -2LogL</td>
<td>19.0% and 440</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 2 (adjusted*)</th>
<th>Odds ratio</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>HellenicSCORE (per 1% increase)</td>
<td>1.15</td>
<td>1.05-1.25</td>
</tr>
<tr>
<td>SES status (high vs. low)</td>
<td>0.48</td>
<td>0.16-1.42</td>
</tr>
<tr>
<td>Model's chi-square</td>
<td>34.4, df=7, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Nagelkerke R² and -2LogL</td>
<td>15.1% and 225</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 3 (adjusted*)</th>
<th>Odds ratio</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>HellenicSCORE (per 1% increase)</td>
<td>1.17</td>
<td>1.10-1.24</td>
</tr>
<tr>
<td>MedDietScore (per 1/55 unit)</td>
<td>0.96</td>
<td>0.91-1.00</td>
</tr>
<tr>
<td>Model's chi-square</td>
<td>87.2, df=6, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Nagelkerke R² and -2LogL</td>
<td>19.7% and 437</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 4 (adjusted*)</th>
<th>Odds ratio</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>HellenicSCORE (per 1% increase)</td>
<td>1.12</td>
<td>1.03-1.23</td>
</tr>
<tr>
<td>MedDietScore (per 1/55 unit)</td>
<td>0.95</td>
<td>0.89-1.00</td>
</tr>
<tr>
<td>SES status (high vs. low)</td>
<td>0.51</td>
<td>0.17-1.54</td>
</tr>
<tr>
<td>Model's chi-square</td>
<td>37.2, df=8, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Nagelkerke R² and -2LogL</td>
<td>16.3% and 222</td>
<td></td>
</tr>
</tbody>
</table>

*Adjusted for physical activity status, waist-to-hip ratio, diabetes and family history of cardiovascular disease.

MedDietScore were modeled along with the HellenicSCORE, adjusted for the previously mentioned risk factors (model 4), the explained variability of the model (Nagelkerke R²) decreased by 20% (i.e. from R²=19.7% for model 3 to R²=16.3% for model 4). This was mostly attributed to the insignificant SES (p=0.23), showing that SES does not improve the predictive ability of the model, even when MedDietScore is already present.

Discussion

This study examined whether SES would improve the predictive ability of a CVD risk prediction model that also included dietary information (expressed through the MedDietScore), the HellenicSCORE (a marker of global CVD risk based on age, gender, smoking habits, total serum cholesterol and systolic blood pressure levels), and other risk factors (i.e. physical activity, waist-to-hip ratio, diabetes and family history of CVD). Additive logistic models showed that SES information does not improve the predictive ability of a CVD risk model, even when dietary data have already been taken into account.

The area of CVD risk prediction has received much interest in recent years. The use of such a model is a very quick and simple way to assess risk, without failing to address the multifactorial nature of CVD, and the optimal performance of such models is of high importance in order to identify more effectively the people who are at risk of developing CVD.11 Although it seems that existing models lack fundamental information, resulting in under-treatment or over-treatment of the people at risk, increasing the number of variables is often viewed with skepticism by experts, as many do not observe significant gains once the basic factors have been included. Furthermore, issues of increased model complexity and cost have been raised.27-30 A recent scientific statement by the American Heart Association suggested that “the cost, safety, and acceptability of a novel risk marker are additional important practical issues in its evaluation.”31 Thus, there is an increased interest in searching for simpler CVD risk prediction models using non-laboratory predictors.

In the case of SES (assessed through education and income), costly laboratory measurements are not required and it exhibits several advantages, notably
that it can be quickly and accurately measured.\textsuperscript{32} The present study is one of the first to attempt to examine the additive effect of SES on a risk prediction model; previously, SES has been considered as an additional variable in the prediction models, on the rationale that it is likely to influence CVD risk independently.\textsuperscript{15,33,34} Specifically, the QRISK\textsuperscript{8} and the ASSIGN\textsuperscript{7} scores include SES in the form of social deprivation (area-based index) to avoid social gradients in health outcomes. In a recent study by Fiscella et al.,\textsuperscript{17} SES (assessed through education and income) was added to the Framingham risk score in a US sample. The results demonstrated that the Framingham risk score underestimates CHD risk for those with a low SES, suggesting that treatment decisions that ignored SES could magnify SES inequalities. This bias seemed to be reduced by adding SES to Framingham risk score.

As a confirmation of the main finding it was revealed that people who developed CVD were more likely not to adhere to the Mediterranean diet, were less well educated, in the lower SES index level, older, men, and had greater prevalence of the common CVD risk factors (Table 1). Additionally, the fact that the inclusion of dietary habits in a risk-prediction model significantly increases the predictive ability of the estimated CVD model, as reported in a previous work by Panagiotakos et al.,\textsuperscript{12} was confirmed, although a different set of covariates was used. In that work, it was demonstrated that inclusion of dietary habits in the Hellenic version of the ESC SCORE model (HellenicSCORE) increases accuracy and reduces the bias of the estimations.

Mozaffarian et al.\textsuperscript{35} recently highlighted that lifestyle risk factors (such as dietary habits) should be major concerns when looking at CVD from any perspective (i.e. public, patients, clinicians and policy makers). Therefore, dietary assessment should be considered as an additional variable in risk prediction, but to the best of our knowledge, almost none of the commonly used models have included dietary assessment in predicting the risk of future CVD events.

This study revealed that SES does not improve the predictive ability of a CVD risk model, even when dietary data have already been taken into account. As SES is associated directly with CVD, and also seems to be interacting with dietary habits to influence CVD outcome,\textsuperscript{20,21} one can speculate that SES does not function in an exclusively independent manner, as do traditional CVD risk factors, but is also acting as a mediator, increasing CVD risk through an indirect effect on other risk factors.\textsuperscript{36}

**Limitations**

There are some limitations that need to be acknowledged and addressed regarding this study. First, the baseline and follow-up examinations were performed only once, and there is the possibility of measurement error. In addition, the measurement of dietary intake was based on memory and self-reports, both of which are characterized by measurement error. Low-income individuals are generally not well represented because data from the homeless or unemployed are difficult to obtain. As the HellenicSCORE calculates CVD risk for a 10-year period, comparisons during the shorter time period of 5 years may not be so accurate. Finally, logistic regression was preferred to Cox proportional hazard models in this case, because the former class of tests provides more information about the predictive characteristics of the risk model (i.e. Nagelkerke “pseudo” $R^2$) and the duration of follow up was only 5 years.

**Conclusions**

Our results indicate that SES does not improve the predictive ability of a CVD risk model, even when dietary information is already taken into account. This means that the estimated models did not improve in accuracy when SES was included. This extension of scope will provide clinicians and public health policy makers with additional information in order to reduce the burden and health inequalities of CVD.

**Acknowledgements**

The ATTICA study was supported by research grants from the Hellenic Cardiological Society (HCS2002) and the Hellenic Atherosclerosis Society (HAS2003). The authors would like to thank the investigators of the “ATTICA” study: A. Zeimbekis, N. Papaioannou, E. Tsetsekou, L. Papadimitriou, N. Massoura, S. Vellas, A. Katinioti (physical examination), M. Toutouza (data management), M. Kambaxis (dietary assessment), M. Toutouza, C. Tselika and S. Poulopoulou (technical support).

**References**


33. Vasan RS, Kannel WB. Strategies for cardiovascular risk as-

(Hellenic Journal of Cardiology) HJC • 397
K. Vlismas, et al


