**BMJ Open**

External validation and clinical usefulness of three commonly used cardiovascular risk prediction scores in an Emirati population: a retrospective longitudinal cohort study

Saif Al-Shamsi, Romona Devi Govender, Jeffrey King

**ABSTRACT**

Objectives Cardiovascular disease (CVD) risk prediction models are useful tools for identifying those at high risk of cardiovascular events in a population. No studies have evaluated the performance of such risk models in an Arab population. Therefore, in this study, the accuracy and clinical usefulness of two commonly used Framingham-based risk models and the 2013 Pooled Cohort Risk Equation (PCE) were assessed in a United Arab Emirates (UAE) national population.

Design A 10-year retrospective cohort study.

Setting Outpatient clinics at a tertiary care hospital, Al-Ain, UAE.

Participants The study cohort included 1041 UAE nationals aged 30–79 who had no history of CVD at baseline. Patients were followed until 31 December 2019. Eligible patients were grouped into the PCE and the Framingham validation cohorts.

Exposure The 10-year predicted risk for CVD for each patient was calculated using the 2008 Framingham risk model, the 2008 office-based Framingham risk model, and the 2013 PCE model.

Primary outcome measure The discrimination, calibration and clinical usefulness of the three models for predicting 10-year cardiovascular risk were assessed.

Results In women, the 2013 PCE model showed marginally better discrimination (C-statistic: 0.77) than the 2008 Framingham models (C-statistic: 0.74–0.75), whereas all three models showed moderate discrimination in men (C-statistic: 0.69–0.70). All three models overestimated CVD risk in both men and women, with higher levels of predicted risk. The 2008 Framingham risk model (high-risk threshold of 20%) classified only 46% of women who subsequently developed incident CVD within 10 years as high risk. The 2013 PCE risk model (high-risk threshold of 7.5%) classified 74% of men who did not develop a cardiovascular event as high risk.

Conclusions None of the three models is accurate for predicting cardiovascular risk in UAE nationals. The performance of the models could potentially be improved by recalibration.

**Strengths and limitations of this study**

- This longitudinal study is the first to assess the validity and applicability of different cardiovascular risk prediction models in a predominantly Arab population.
- This study’s follow-up period of 10 years and monitoring of outcome events is in line with the Framingham and Pooled Cohort Risk Equation models.
- More than half of the patients were receiving lipid-lowering medication, which could potentially have affected the study results.
- The study population was from outpatient clinics of a single large medical centre, and the results may not be applicable to the general United Arab Emirates population.

**INTRODUCTION**

Cardiovascular diseases (CVD), which include coronary heart disease and stroke, account for approximately 17.8 million deaths worldwide in 2017,1 with this number expected to increase to more than 23.6 million deaths by 2030.2 In the United Arab Emirates (UAE), an affluent developing country, CVD is an increasing public health concern and is responsible for more than a quarter of all deaths.3

The burden of CVD can be lessened using primary prevention and by effectively targeting preventive strategies for high-risk individuals. Risk prediction models are important tools that have been developed to help in preventing CVD. They can aid clinicians in identifying individuals who would benefit from effective treatment using measurements of conventional risk factors for CVD to estimate individuals’ 10-year risk.4 5 Currently, there are three CVD risk prediction models recommended in UAE...
clinical practice guidelines for risk assessment of developing CVD\textsuperscript{6–8}: the 2008 Framingham risk model,\textsuperscript{9} the 2008 office-based Framingham risk model,\textsuperscript{9} and the 2013 American College of Cardiology/American Heart Association Pooled Cohort Risk Equation (PCE).\textsuperscript{10} However, the Framingham models are based on individuals aged 30–74 years of predominantly European descent, whereas the PCE model was obtained from several US population-based cohort studies that included Caucasian and African American individuals aged 40–79 years.

To determine a risk prediction model's clinical applicability, it should be externally validated in the local patient setting where it is to be applied. To the best of our knowledge, there have been no studies that have evaluated the performance of any long-term CVD risk prediction models in an Arab population. Therefore, the aim of this study is to assess the predictive values of the two most commonly used Framingham models and the 2013 PCE model in predicting 10-year CVD risk in an Emirati population.

METHODS

Study population and setting
This 10-year retrospective cohort study was conducted at outpatient clinics of Tawam Hospital, a large, government-subsidised tertiary care hospital in Al-Ain, UAE, which provides medical care exclusively for UAE nationals. Baseline ambulatory electronic medical records of patients were manually extracted from 1 April to 31 December 2008. Follow-up data were reviewed annually until 31 December 2019. The study included 1048 UAE nationals, aged 30–79 years, with no history of CVD at baseline. A history of CVD was defined as a previous diagnosis established by specialist physicians of a transient ischaemic attack, a stroke, angina, myocardial infarction, coronary artery surgery, coronary angioplasty, heart failure or peripheral arterial disease. We excluded three patients for whom endpoint information was missing and four patients who were lost to follow-up, defined as last clinic visit prior to 12 months from the baseline visit and were not known to have a cardiovascular outcome. The final sample included 1041 individuals.

Risk score calculations
The 10-year predicted risk for CVD for each study participant was calculated using their baseline data. The three CVD risk prediction models that were used included the 2008 Framingham model,\textsuperscript{9} the 2008 office-based Framingham model\textsuperscript{9} and the 2013 PCE model.\textsuperscript{10}

The baseline variables used for the risk scores were age; sex; body mass index (BMI), which was calculated as the patient’s weight divided by the square of their height (kg/m\textsuperscript{2}); serum total cholesterol; serum high-density lipoprotein-cholesterol (HDL-C); history of diabetes (defined as glycated haemoglobin (HbA1c) level \geq 6.5% or receiving antihyperglycaemic treatment\textsuperscript{11}); history of smoking (defined as previous or current history of smoking tobacco); systolic blood pressure; and the use of blood pressure-lowering medications (online supplemental table S1). For the variable of ethnicity in the PCE risk model, ‘White or others’ was used for our study cohort. Blood pressure was assessed as part of the routine practice of trained nurses using an automated oscillometric blood pressure measuring device. Lipid-lowering treatment included the use of statins, fibrates or bile acid sequestrants. All laboratory tests were performed at Tawam Hospital’s medical laboratory department.

Cardiovascular endpoints
During the follow-up, the primary endpoints were defined for each model and assessed separately based on the original Framingham and PCE cohorts. The primary endpoints for the 2008 Framingham models were coronary death, acute myocardial infarction, angina, fatal or non-fatal stroke, transient ischaemic attack, congestive heart failure and peripheral artery disease. For the 2013 PCE model, the endpoints were defined as coronary death, acute myocardial infarction, and fatal or non-fatal stroke. Coronary death or fatal stroke was determined from the manual review of electronic medical records or death certificates. In addition, sudden death outside the hospital was considered a coronary death unless documented otherwise. The occurrence of the first event was used to define the onset of the disease if there was more than one endpoint recorded for a patient. The study endpoints were adjudicated by a clinical expert committee that manually reviewed the hospital records and death certificates.

Statistical analysis
Single imputation using linear regression with a random error term was performed for missing data: HbA1c (4.0% missing), BMI (<1% missing), total cholesterol (<1% missing), and HDL-C (<1% missing).

The 1041 eligible patients were grouped into two overlapping cohorts: the PCE validation cohort (patients aged 40–79 years) and the Framingham validation cohort (patients aged 30–74 years) (figure 1).

Continuous variables that are normally distributed are presented as means (SD). Medians (quantiles percentiles) are used for continuous variables that are not normally distributed. Categorical variables are presented as percentages. Differences in baseline characteristics by sex were analysed using the Mann-Whitney U test for continuous variables that are not normally distributed, the independent samples t-test for continuous variables that are normally distributed, and Fisher’s exact test (two tailed) for categorical variables.

For the Framingham risk models, patients in the Framingham validation cohort were classified as high risk if their 10-year predicted CVD risk was \geq 20%. For the PCE validation cohort, participants with a 10-year predicted CVD risk of \geq 7.5% were considered high risk.
The clinical performance of each of the three models was determined by analysing their discrimination, calibration and clinical usefulness for predicting 10-year CVD risk in men and women separately.

**Discrimination**
Discrimination is defined as the probability that a model can correctly discriminate between patients who develop and those who do not develop the endpoint. Discrimination was assessed using the C-statistic, proposed by Harrell et al., which is comparable to the area under the receiver-operating characteristic curve. A C-statistic value >0.75 was considered to represent good discrimination.

**Calibration**
Calibration, which assesses the agreement between observed and predicted 10-year CVD risks, was plotted using the Kaplan-Meier estimator to determine the observed incidence of CVD events, which was then compared with the 10-year CVD risk predicted by each model. A 45° line in the calibration curve indicates perfect calibration; the closer the risk model calibration curve to this 45° line, the better the calibration.

**Decision curve analysis**
The clinical usefulness of the three risk prediction models was assessed by applying decision curve analysis, which plots the expected net benefit relative to ‘intervention for none’ and compares it to ‘intervention for all’ across different probability thresholds. Net benefit takes into account both discrimination and calibration, as well as clinically relevant factors, such as the consequences of the decisions made based on the results provided by the model. A higher net benefit is provided by those models that are farthest away from the ‘intervention for none’ and ‘intervention for all’ lines.

A two-tailed p<0.05 was considered statistically significant. All statistical analyses were performed using IBM SPSS software, V.26 (IBM) and R software, V.3.5.2 (The R Foundation, Vienna, Austria).

**Patient and public involvement**
There was no patient or public involvement in the design and conduct of the study.

**RESULTS**

**Baseline characteristics**
In the Framingham validation cohort, a total of 1001 patients, aged 30–74 years, with no history of CVD were identified, and 810 patients, aged 40–79 years, were included for analysis in the PCE validation cohort (figure 1). The median follow-up times in the Framingham and PCE validation cohorts were 10.3 years (IQR, 8.4–11.1 years) and 10.6 years (IQR, 9.1–11.2 years), respectively.

In both cohorts, men were more likely to have a history of diabetes and smoking, a higher systolic blood pressure, and lower BMI and HDL-C compared with women (tables 1 and 2). In addition, in the PCE validation cohort, men were older and had lower total cholesterol than women (table 2).
Of the 484 men in the Framingham validation cohort, 49.6% and 53.7% were determined to be at high risk of CVD using the Framingham and office-based Framingham prediction models, respectively, and of the 378 men in the PCE validation cohort, 77% were identified as high risk. Among women, 18.2%, 26.1% and...
35% were classified as high risk using the Framingham, office-based Framingham, and PCE models, respectively. During the follow-up, there were 175 and 102 incident cardiovascular events recorded in the Framingham validation cohort and the PCE validation cohort, respectively.

Online supplemental table S2 depicts the percentages of men and women who were categorised by the three risk prediction tools as high risk and who experienced a cardiovascular event. Among men who experienced a cardiovascular event, 93% (57/61) were classified as high risk by the 2013 PCE risk model. However, of the men who did not experience a cardiovascular event, 74% (234/317) were classified as high risk. In women, the 2008 Framingham risk model classified less than half (38/82) of the patients who did go on to develop incident CVD as high risk.

Discrimination
The C-statistic value indicated marginally better discrimination by the PCE model than by the Framingham models in Emirati women. In Emirati men, all three models showed moderate discrimination (table 3). Overall, the discriminatory power was higher for women (C-statistic: 0.77) than for men (C-statistic: 0.69) based on the 2013 PCE model.

Calibration
The calibration curves in figure 2 show that all three models generally overestimated risk in Emirati men and women, with higher levels of predicted risk and underestimated risk in lower levels of predicted risk, particularly in women.

Decision curve analysis
The clinical usefulness of the three risk prediction models was assessed graphically by performing a decision curve analysis (online supplemental figure S1). On the basis of this analysis, in both Emirati men and women, the two Framingham models were slightly superior relative to the PCE model for risk threshold probability. In men, the curves are further away from the ‘treat none’ and ‘treat all’ lines at threshold probabilities of approximately 22%–26% and 20%–22% for the 2008 Framingham models and the 2013 PCE model, respectively. In women, the corresponding threshold probabilities are 17%–20% and 11%–14%, respectively.

DISCUSSION
This study evaluated the predictive values and applicability in the UAE of three commonly used CVD risk assessment tools by assessing their discrimination, calibration

<table>
<thead>
<tr>
<th></th>
<th>2008 Framingham risk</th>
<th>2008 office-based Framingham risk</th>
<th>2013 PCE risk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-statistic (95% CI)</td>
<td>0.70 (0.65 to 0.75)</td>
<td>0.69 (0.64 to 0.74)</td>
<td>0.69 (0.63 to 0.75)</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-statistic (95% CI)</td>
<td>0.74 (0.69 to 0.80)</td>
<td>0.75 (0.70 to 0.80)</td>
<td>0.77 (0.71 to 0.83)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-statistic (95% CI)</td>
<td>0.72 (0.69 to 0.76)</td>
<td>0.72 (0.69 to 0.76)</td>
<td>0.74 (0.69 to 0.78)</td>
</tr>
</tbody>
</table>

PCE, Pooled Cohort Risk Equation.
and clinical usefulness. The external validation of the two 2008 Framingham models and the 2013 PCE model suggests that none is ideal for predicting cardiovascular events in an Emirati population. Furthermore, all three models overestimated cardiovascular risk in both Emirati men and women at high risk of CVD. In terms of discrimination, the 2013 PCE model’s performance was marginally better in women compared with the performance of either of the Framingham models, but it did not show good clinical usefulness at levels higher than the 25% threshold.

The moderate discrimination observed in our study of the three risk models, particularly in men, could be explained by the high-risk cardiovascular profile of the cohorts as most of the patients had hypertension, diabetes or dyslipidaemia at baseline. Therefore, it is possible that these risk prediction models could provide better discrimination if applied to the general population, among whom there is a wider variation in cardiovascular risk. However, most studies that have validated these models in different ethnic groups have also shown similar low to moderate discrimination.15–18

Our finding that all three models overestimate cardiovascular risk in higher-risk UAE nationals is consistent with numerous previous studies that have externally validated these models in different populations.18–22 This discrepancy is to be expected as calibration, to a large extent, relies on the similarity between the validation cohort and the development cohort. The three prediction models were developed using data from an era that preceded the introduction of statins.23 Therefore, the extended use of effective preventive measures targeting CVD events over time would have resulted in the overestimation of risk seen in our and other studies. However, external validation cohort studies that adjusted for lipid-lowering treatment have still reported discrepancies between observed and predicted CVD risks.20,24

In our study, the two Framingham models showed slightly better clinical utility in both sexes compared with that of the PCE model. For adults ≤39 years of age, there are sparse data relating to the utility and performance of the 2013 PCE risk prediction model in the UAE national population, age has been shown to be an independent predictor of incident CVD,25 with several studies in the region having reported a high incidence of atherosclerotic CVD in younger age groups.26–28 A study conducted in six Middle Eastern countries, including the UAE, found that 50% of patients diagnosed with an acute coronary syndrome were aged <55 years, a decade younger than those being diagnosed in North America and Europe.29 Therefore, the exclusion of younger patients from the PCE validation cohort could potentially explain the lower clinical utility of the 2013 PCE risk prediction model relative to the Framingham models seen in our study. Given the implications of the earlier onset of CVD among UAE nationals, we believe that having validated risk assessment tools that include young adults may be of greater clinical relevance in our population.

The Weqaya cardiovascular screening programme was established in 2008, and it is available to all UAE nationals residing in Abu Dhabi, which includes Al-Ain city.6,30 Every adult Emirati enrolled in the programme is risk stratified using the 2008 Framingham prediction model. However, in our study, only 46% of Emirati women who experienced a cardiovascular event within 10 years, were categorised as high risk (20% cut-off threshold) using the 2008 Framingham model. This underestimation of risk using the 20% cut-off threshold could result in undertreatment using appropriate therapies in Emirati women, who would benefit from a reduction in CVD.

Alternatively, the 2013 PCE risk prediction model has been recommended for screening CVD risk in our population,7 yet the 7.5% cardiovascular risk threshold to determine the need for lipid-lowering treatment in our study cohort appears to be too low, particularly in men. Therefore, this current American College of Cardiology/American Heart Association recommended cut-off threshold could potentially result in inappropriate lipid-lowering therapy in our population.

Our results showing the performance of three commonly used models in the UAE highlight the clinical dilemma faced by local healthcare providers when conducting CVD risk assessments in Emirati men and women. The accurate estimation of cardiac risk underpins the need to either recalibrate existing models or develop new prediction tools that will effectively balance the risks and benefits of preventive therapies in our population.

Strengths and limitations
This study is the first to assess the validity and applicability of different cardiovascular risk prediction models in a predominantly Arab population. Furthermore, our study’s long follow-up period of 10 years and monitoring of outcome events is in line with the Framingham and PCE risk models.

There are, however, several limitations to our study. First, more than half of our patients were already receiving lipid-lowering medication, which could potentially have affected the study results. However, as discussed earlier, the use of medication does not fully explain the overestimation of risk seen among the risk models. Second, the retrospective nature of our study presents an inherent limitation, possibly rendering it prone to selection bias. In addition, misclassification or underascertainment of cardiovascular endpoints may have led to the overestimation of risk observed by the risk models. Finally, our study population was from outpatient clinics of a single large medical centre, and the results may not be applicable to the general Emirati population. However, patients visiting outpatient clinics would be the ideal target for risk stratification and subsequent preventive therapies.

CONCLUSION
Our study demonstrated that neither the 2008 Framingham risk models nor the 2013 PCE risk model can
accurately predict cardiovascular risk in Emirati men and women. Furthermore, using the recommended cut-off thresholds, the 2013 PCE risk model overestimates cardiovascular risk in men, whereas the 2008 Framingham model underestimates cardiovascular risk in women. The performance of the models could potentially be improved by recalibration. The need to accurately estimate CVD risk in our population is essential to avoid missed opportunities for disease prevention and to balance the risks and benefits of treatment.

Contributors Conception and design of the study: SA-S, RDG and JK. Data collection: SA-S. Statistical analysis: SA-S. Interpretation of the data: SA-S, RDG, JK. Drafting of the manuscript: SA-S. Critical review of the manuscript: SA-S, RDG and JK. All authors have read and approved the final version of the manuscript.

Funding This work was supported by the College of Medicine and Health Sciences, UAE University (No. 31M325).

Disclaimer The funders had no role in the study design, data collection, or analysis, decision to publish, or preparation of the manuscript.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study protocol was approved by Tawam Hospital and the UAE University Research and Ethics Board (CRD 239/13). The requirement for informed consent was waived because the patient records and information were anonymized and de-identified prior to the analyses conducted.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The data that support the findings of this study are available from the corresponding author, SA-S, on reasonable request.

Supplemental material This content has been supplied by the author(s). It has been peer reviewed and has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Saif Al-Shamsi http://orcid.org/0000-0001-9755-3493

REFERENCES