Applying engineering in healthcare: a proposed computer-assisted mathematical model for atherosclerotic cardiovascular risk assessment

ALINA MIHAELA PASCU¹, DANIELA MARIANA BARBU², ION BARBU²
LIGIA NEICA¹, ANDREEA FLEANCU¹
¹Faculty of Medicine, ²Faculty of Mechanical Engineering
Transilvania University of Brasov, ROMANIA
alina.pascu@unitbv.ro

Abstract: Despite major advances in the diagnosis and treatment of atherosclerotic cardiovascular disease (CVD) in the past century, it remains a serious clinical and public health problem. There is a need for a new cardiovascular disease model that includes a wider range of relevant risk factors, in particular lifestyle factors, to aid targeting of interventions and improve population models of the impact of cardiovascular disease and preventive strategies. The model needs to be applicable to a wider population including different ethnic groups, different countries and to those with and without cardiovascular disease. Separate multivariable risk algorithms are commonly used to assess risk of specific atherosclerotic cardiovascular disease events, i.e., coronary heart disease (CHD), cerebrovascular disease, peripheral vascular disease, and heart failure. In recent years a number of algorithms for cardiovascular risk assessment have been proposed to the medical community. These algorithms consider a number of variables and express their results as the percentage risk of developing a major fatal or non-fatal cardiovascular event in the following 10 to 20 years. Decades of evaluation of CVD risk factors by the Framingham Study led to the conclusion that CVD risk evaluation is most fruitfully appraised from the multivariable risk posed by a set of established risk factors. Such assessment is essential because risk factors seldom occur in isolation, and the risk associated with each varies widely depending on the burden of associated risk factors. Multivariable risk stratification is now recognized as essential in efficiently identifying likely candidates for CVD and quantifying the hazard.

The present paper aims to propose a computer-assisted model for estimating short-term (10-years) risk for CHD or CHD risk-equivalents based on the steps proposed in the most validated risk-score algorithm, i.e., Third Report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III).

Key Words: medical engineering, computer-assisted mathematical model, logistic regression, Cox proportional hazards regression, accelerated time failure analysis, artificial neural networks, cardiovascular risk assessment, Adult Treatment Panel III, Framingham Point Scores, atherogenic dyslipidemia

1. Introduction

Cardiovascular diseases (CVD) still represent the leading cause of mortality and morbidity, despite major advances in its diagnosis and development of modern therapeutical techniques. It is why prevention strategies have become major approaches for reducing CVD and its unfavourable consequences. Prevention strategies relay on risk score for CVD estimation, based on identifying, assessing, and treating the risk factors associated with CVD. Presently, many of these strategies involve focusing on a component of CVD such as hard coronary disease consisting of myocardial infarction and coronary death, assessing the risk by mathematical risk functions or scoring functions, and designing treatment (behavioural and/or drug) according to the level of risk. [9]

Some valuable algorithms for cardiovascular risk assessment have been proposed to the medical community in the last few years. [2, 6, 14, 15, 24, 25, 37] Their purpose is to assist physicians in defining the risk level of an individual patient with regard to developing major cardiovascular events in the following years. These algorithms have been drawn from statistical analyses performed on longitudinal study cohorts. These analyses have taken into account events occurring in general populations undergoing adequate follow-up for a sufficient length of time. These algorithms consider a number of variables and express their results as the percentage risk of developing a major fatal or non-fatal cardiovascular event in the following 10 to 20 years. [15]

There are several reasons for calculating the risk of cardiovascular disease in an individual or a population. Health care providers need to model future patterns of need for health services, and to identify the cost effectiveness of different intervention strategies. [20-22] Insurance companies and pension funds have to evaluate risk in both individuals and populations when assessing portfolio risks. In clinical medicine, cardiovascular risk is increasingly accepted as the appropriate criterion to use to identify the patients who will most benefit from interventions designed to prevent cardiovascular disease and death. [27, 35] Another (and perhaps overlooked) requirement is to inform shared decision-making with patients. [11, 29]
2. Methods and Discussion

It is widely accepted that age, gender, high blood pressure, smoking, dyslipidemia, and family history of premature atherosclerotic CVD are the major risk factors for developing CVD. [7] It also is recognized that CVD risk factors cluster and interact multiplicatively to promote vascular risk. [17] This knowledge led to the development of multivariable risk prediction algorithms incorporating these risk factors that can be used by primary care physicians to assess in individual patients the risk of developing all atherosclerotic CVD [4, 10, 12, 13, 16, 23, 28] or specific components of CVD, ie coronary heart disease, [12, 13, 36] stroke, [38] peripheral vascular disease, [26] or heart failure. [18] Multivariable assessment has been advocated to estimate absolute CVD risk and to guide treatment of risk factors. [10, 17]

There are a variety of CVD risk estimators available, the best known are summarized in Table 1. Each has strengths and weaknesses. [1, 3, 6, 8, 16, 40] The principal problems include limited applicability to different geographic areas or ethnic groups, application to men but not women, and the omission of important risk factors. [22]

<table>
<thead>
<tr>
<th>Risk equation</th>
<th>Risk factors included</th>
<th>Risks evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framingham (Anderson) [1]</td>
<td>Age, Gender, Smoking, Blood pressure (BP), Total cholesterol (TC)/high density lipoprotein (HDL) ratio, Diabetes, Left ventricular hypertrophy (LVH)</td>
<td>4 to 12-year risk of CHD events and death, All CVD events and death, All cerebrovascular events, Myocardial infarction events</td>
</tr>
<tr>
<td>Framingham (D’Agostino) [8]</td>
<td>Age, Gender, Smoking, BP, TC/HDL ratio, Alcohol, Existing CVD, Menopausal status women, Triglycerides in women</td>
<td>2-year risk of CHD events</td>
</tr>
<tr>
<td>Whitehall equation [33]</td>
<td>Age, Gender, TC, BP, Cigarettes per day</td>
<td>5 or 10-year risk of CHD event</td>
</tr>
<tr>
<td>Systematic Coronary Risk Evaluation (SCORE) [6]</td>
<td>Age, Gender, Smoking, BP, TC, Residence in a ‘high’ or ‘low’ risk country</td>
<td>10-year risk of death from CHD or CVD</td>
</tr>
<tr>
<td>Munster Heart Study (PROCAM)</td>
<td>Age, Smoking, BP</td>
<td>Major coronary event</td>
</tr>
</tbody>
</table>

All the models in Table 1 include age, gender, blood pressure, cholesterol, cigarette consumption and diabetes as risk factors. All omit some important independent risk factors such as family history, existing CVD, obesity but also diet, alcohol consumption and exercise. We are particularly interested in risk factors related to lifestyle: if an estimate of risk is to be used in consultations as part of discussions with patients about lifestyle modification, it is important that the estimate should include the fullest possible range of risk factors relating to lifestyle. [22]

3. Mathematical predictive models

Because of the multifactorial predisposition to CVD, and the need to determine and quantify the net and joint contribution of predisposing risk factors, multivariable risk formulations were needed. The first of these was devised in the 1960s and subsequently followed by risk formulations devised on the basis of longer periods of follow-up, better predictive variables, and increasingly sophisticated statistical methods, including logistic regression, Cox proportional hazards regression, and accelerated time failure analysis. [10, 19] Cornfield et al used the logistic model in 1961, using logistic coefficients derived from a linear discriminant function.
[19] This was the least costly approach for computer time, which in those days was quite expensive. This was followed in 1967 by a quintile of risk approach that used discriminant function analysis. In the same year, Walker and Duncan suggested an iterative solution comprising maximum likelihood equations that did not require making any assumptions about the distribution of the variables in the population. The Walker and Duncan approach is what today is called “logistic regression”. [19]

Logistic regression analysis became established, and quantitative synthesis of a number of major risk factors into a composite score on the basis of Framingham Study data was accomplished, first for coronary disease, then for stroke, peripheral artery disease, and finally heart failure. [19] The risk factors selected were not highly intercorrelated, made a contribution in the presence of other risk factors (i.e., “were independent” contributors), and were obtainable with ordinary office procedures and readily available laboratory tests. In time, the addition of data from more extended follow-up facilitated a closer quantitative examination of risk factor interrelationships, and new variables were incorporated in statistical models either to provide pathogenic insights or improve risk estimation. [19]

Risk profiles composed of the standard risk factors are just as useful for predicting CVD events in elderly patients as in middle-aged patients, despite a lesser prevalence which can be used to recalculate national mortality statistics in this way. [22, 31, 32]

The best known estimators are the Framingham equations. These have been criticized for their inaccuracy in some countries, in particular Southern Europe where they tend to over-estimate risk significantly. [22] This variation is an inevitable consequence of the exclusion of significant risk factors from the model. If a model is derived in a particular population, the prevalence and impact of any missing risk factors is tacitly embedded in coefficients of the risk equations. When applied to a population with different prevalences or one in which risk factors have different impacts, the model's predictions will be less accurate. Attempts have been made to recalibrate the Framingham equations for different ethnic groups in the United States and the United Kingdom. [3] However, the recalibrated equations have not been validated and questions about their applicability to other geographic areas remain unanswered. [22]

The Framingham formulation for predicting coronary heart disease (CHD) was incorporated into the Third Report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III – ATP III). [12] The Framingham CHD risk assessment tool has been validated in whites and blacks in the United States [12, 22, 28] and are transportable (with calibration) to culturally diverse populations in Europe, the Mediterranean region, and Asia. [12, 22, 28] Similar CHD risk prediction algorithms have been developed by other investigators worldwide and have been demonstrated to perform well. [13, 22]

We designed a computer-assisted model for estimating short-term (10-years) risk for CHD or CHD risk equivalents based on the steps proposed in the most validated risk-score algorithm, i.e., Third Report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). [12]

The main objective of this proposed model is to be easy to use both in medical practice, and in training students, residents, fellows and practitioners in estimating the atherosclerotic CVD risk-score, enabling all to follow rigorously the ATP III steps, from risk factors detection to patient assessment, therapeutic decision strategies, monitoring, and result-evaluating programmes.

In this aim, a medical doctors and engineers team joined efforts to create a user-friendly programme for CVD risk evaluation.

The programme enables the user to follow the ATP III guidelines step-by-step, providing also automatic calculation of serum LDL-cholesterol level from lipoproteins profile, and useful links to risk-category score, optimal correspondent LDL-cholesterol level to
achieve, therapeutic approach encompassing therapeutic lifestyle changing measures, food and dietary supplements recommended, and also medication – the most suitable lipid-lowering drugs to choose for each type of dyslipidemia, ideal doses, precaution to take into account, side-effects, and so on. Figure 1 illustrates some steps from the proposed atherosclerotic CVD risk assessment model proposed.

### 3. Conclusion

The use of a general CVD risk score is an attractive option, especially in office-based primary care practices. Serial assessment of global CVD risk could be used to monitor progress of patients on treatment and improvement in their multivariable risk scores.

This attempt to create a step-by-step algorithm for atherosclerotic CVD 10-year risk assessment demonstrates how published information can be used to construct a mathematical model of cardiovascular risk. The method should be applicable to other disease groups where there is sufficient information available.

It is now recognized that atherosclerotic CVD is attributable to a variety of factors and has several clinical manifestations. In every instance, the hazard of a particular risk factor varies widely depending on the burden of associated accompanying risk factors. Almost half of CVD events occur in the tenth of the population at highest multivariate risk. Single risk factor detection and correction may be worthwhile for prevention of CVD on a population basis, but is inefficient on an individual basis. Individual candidates for CVD can best be detected and targeted for treatment from a multivariable risk profile. [19]

Correction of dyslipidemia or hypertension is best directed at patients with a high multivariable risk for CVD, because they stand to benefit the most. Multivariable risk assessment also avoids overlooking patients who are at high risk for CVD with multiple marginal risk factors and avoids needlessly alarming patients with only one isolated risk factor. Much of CVD mortality attributed to individual risk factors is actually caused by the risk factor in combination with other risk factors. Relatively little CHD mortality is attributable to each risk factor in isolation. Analysis that fails to examine risk factors in combinations usually greatly overestimates the population-attributable risks associated with individual risk factors. [5, 19]

Global risk assessment is also useful in motivating...
patients and quantifying the hazard they face. [30] How to motivate more physicians to adopt multivariable risk assessment in their practice is an unresolved problem.

An approach worth considering is having the clinical laboratory request the values for the standard CVD risk factors with the requested tests of blood lipid or blood sugar levels, so they can report a multivariable risk estimate with the requested tests.

The use of predictive algorithms to assess individual absolute risk of cardiovascular future events is currently hampered by methodological and mathematical flaws. [15] The use of newer approaches, such as fuzzy logic and artificial neural networks, linked to artificial intelligence, seems to better address both the challenge of increasing complexity resulting from a correlation between predisposing factors, data on the occurrence of cardiovascular events, and the prediction of future events on an individual level.

The model proposed could be useful in modelling a broad range of disease areas. Further research and needs to be done to evaluate the accuracy of the model in different population groups using historical cohort data. Multiprofessional working groups encompassing medical doctors, engineers, statisticians, ethics and informatics specialists, should join efforts to develop algorithms and models to facilitate medical practice, both in accurate diagnosis and therapeutic decision.

References:

[15]. Grossi E. How artificial intelligence tools can be used to assess individual patient risk in cardiovascular disease: problems with the current methods. BMC Cardiovascular Disorders.


