Comparison of Models for Predicting Outcomes in Patients with Coronary Artery Disease Focusing on Microsimulation

Masoud Amiri, Roya Kelishadi

ABSTRACT

Background: Physicians have difficulty subjectively estimating the cardiovascular risk of their patients. Using an estimate of global cardiovascular risk could be more relevant to guide decisions than using binary representation (presence or absence) of risk factors data. The main aim of the paper is to compare different models of predicting the progress of coronary artery diseases (CAD) to help the decision making of physician.

Methods: There are different standard models for predicting risk factors such as models based on logistic regression model, Cox regression model, dynamic logistic regression model, and simulation models such as Markov model and microsimulation model. Each model has its own application which can or cannot use by physicians to make a decision on treatment of each patient.

Results: There are five main common models for predicting outcomes, including models based on logistic regression model (for short-term outcomes), Cox regression model (for intermediate-term outcomes), dynamic logistic regression model, and simulation models such as Markov model and microsimulation model. The advantages and disadvantages of these models have been discussed and summarized.

Conclusion: Given the complex medical decisions that physicians face in everyday practice, the multiple interrelated factors that play a role in choosing the optimal treatment, and the continuously accumulating new evidence on determinants of outcome and treatment options for CAD, physicians may potentially benefit from a clinical decision support system that accounts for all these considerations. The microsimulation model could provide cardiologists, researchers, and medical students a user-friendly software, which can be used as an intelligent interventional simulator.

Key words: Coronary artery disease, microsimulation, prediction models

INTRODUCTION

Each cardiac disease can be treated in multiple ways according to multiple interrelated factors dependent on patient...
characteristics, severity and progression of disease, and patient and physician preference. Every day the cardiologist faces these complex clinical decision-making challenges. For example, which cardio-protective medication would he/she prefer to increase life expectancy of patients with chronic stable coronary artery disease (CAD)? What are the influences of medications, including aspirin, beta-blockers, statins, and ACE-inhibitors on prognosis of CAD patients? What are the effects of percutaneous coronary intervention (PCI) and/or coronary surgery on (event-free) life expectancy of patients with chronic CAD? It can be quite complicated to answer these questions for an individual patient, because multiple factors affect outcome in patients with a heart disease. At present, the selected choice is based on general medical knowledge, individual experience, and experience of the cardiologist.

Research has shown the difficulty that physicians have to subjectively estimate the cardiovascular risk of their patients.[1,2] Moreover, it has been suggested that using an estimate of global cardiovascular risk could be more relevant to guide decisions, than using binary representation (presence or absence) of risk factor data.[3] During the last decades, preventive and therapeutic strategies have been developed that contributed to improved management and outcome of patients with atherosclerosis. A major challenge for contemporary medicine is to rationally implement these strategies in clinical practice. By applying standard parametric and semiparametric models for risk factor assessment, it is possible to identify factors that may influence long-term outcome in that particular patient group. In recent years, there have been calls for a greater use of models in decision making.[4,5] After using these models, it is supposed to lead to a more accurate identification of patients who will most benefit from the treatment.[6] It is obvious that the access for the practitioners to a numerical expression of risk does modify their behavior.[7,8] There are different standard models for predicting risk factors such as models based on logistic regression model, Cox regression model, dynamic logistic regression model, and simulation models such as Markov and microsimulation models.[9,10]

The main aim of the paper is to compare different models of predicting the progress of a coronary artery disease, in order to help cardiologists to make a good decision, and also answer the above-mentioned questions. The microsimulation model should provide cardiologists, researchers, and medical students a user-friendly software package, which can be used as an intelligent interventional simulator. The model will be a valuable tool for health professionals to optimize their therapeutic plans. Hospitals, clinics, cardiological research centers, and other health units can use this model for medical care and accurate treatment of their patients. It can also be used as an advanced educational program for training specialists in the cardiovascular field. Patients will also greatly benefit since their treatment will become more efficient. There are two specific research questions that we would like to address in this paper. First, what is the usefulness of microsimulation models for the cardiology field? And second, how can we construct such a model? With developing this user-friendly, fast, and reliable tool that will provide access to heterogeneous health information sources and also introducing new methods for decision support and risk analysis, it might support health professionals in taking promptly the best possible decision for prevention and treatment.

STANDARD MODELS FOR PREDICTING OF OUTCOMES

There are five main common models for predicting of outcomes, including models based on logistic regression model (for short-term outcomes), Cox regression model (for intermediate-term outcomes), dynamic logistic regression model, and simulation models such as Markov and microsimulation models (for long-term outcomes).[9,10]

Logistic regression is part of the category of statistical models called generalized linear models (GLM). This broad class of models includes ordinary regression and analysis of variance (ANOVA), as well as multivariate statistics such as analysis of covariance (ANCOVA) and log-linear regression. Logistic regression can be used to predict a dependent variable on the basis of independents and to determine the percent of variance in the dependent variable explained by the independents; to rank the relative importance of independents; to assess interaction effects; and to understand the impact of covariate control variables. This model
allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these. Logistic regression estimates the probability of a certain event occurring.

The Cox (proportional hazard) regression model allows analyzing risk factors on survival. A Cox model is a well-recognized statistical technique for exploring the relationship between the survival of a patient and several explanatory variables. This model provides an estimate of the treatment effect on survival after adjustment for other explanatory variables. It allows us to estimate the hazard (risk) of death, or other event of interest, for individuals, given their prognostic variables. The probability of the endpoint (death, or any other event of interest, disease) is called the hazard. Even if the treatment groups are similar with respect to the variables known to affect survival, using the Cox model with these prognostic variables may produce a more precise estimate of the treatment effect (for example, by narrowing the confidence interval).

The dynamic logistic regression model is developed for application in nonlinear, non-normal time series, and regression problems, providing dynamic extensions of standard generalized linear models. In other words, the dynamic logistic regression model is an extension of logistic regression to modeling nonstationary data that combines the strengths of classical statistical models with the higher expressivity of features generated from nonstationarity of data. Nonstationary data are a sequence of input-output pairs that may vary with time.

The Oxford English Dictionary describes “simulation” as: “The technique of imitating the behavior of some situation or system by means of an analogous model, situation, or apparatus, either to gain information more conveniently or to train personnel.” Since the early 1960s, simulation has been one of many methods used to aid strategic decision-making within industry. Its main strength lies in the ability to imitate complex real-world problems and to analyze the behavior of the system as time progresses. In other words, simulation is an attempt to solve a problem or to work out the consequences of doing something by representing the problem or possible course of events mathematically, often using a computer. Simulation is also defined as the process of imitating a real phenomenon with a set of mathematical formulas. The use of a mathematical model to recreate a situation, often repeatedly, so that the likelihood of various outcomes can become more actual estimates, is the third definition of simulation. In fact, we simulate a set of conditions, to produce specified conditions in order to conduct an experiment artificially. Simulation models can be divided into two main groups: macrosimulation, such as Markov model, and microsimulation.

The Markov model provides a framework to model recurring events and to extend models to encompass the lifetime of a patient. This model is useful when the decision problem involves risk over time, when the timing of events is important, and when events may happen more than once. In a Markov model uncertain events are modeled as transitions during defined time intervals (cycles) between defined health states. Defining a Markov model process requires several steps: define the states, determine the cycle length, consider possible transitions among states, assess transition probabilities, and assess utilities and the costs (in a case of cost-effectiveness analysis) associated with being in each state for one cycle. The schematic concept of the Markov model in provided in Figure 1.

Microsimulation models are an important tool for estimating the health consequences and economic costs associated with alternative clinical strategies in a population of interest. Microsimulation provides a powerful technique to account for variability across subjects. Discrete-event simulation is a technique developed in industrial engineering to model chains of stochastic events and to model competition for the available resources. Microsimulation can also be performed using discrete-event simulation. Sets of equations can be used to model directly the demographic characteristics and risk factors.

---

**Figure 1:** A simple Markov model
of the subjects. The life history of each subject is simulated and the development of disease, symptoms, and disease progression can be modeled with equations that track the time to the next event. Instead of dividing time into intervals during which events may or may not occur, as happens in the Markov model, in discrete-event simulation the time to the next event is estimated based on the probability distribution of that event. The variables are then updated to the point in time at which that next event occurs. In other words, discrete-event simulation (DES) concerns the modeling of a system as it evolves over time by representing the changes as separate events. Figure 2 shows a simple microsimulation model.

The microsimulation model is comparable to a flight simulator in the aviation industry; a flight simulator simulates flights of a particular airplane on a particular route, taking into account several conditions like the type of weather and possible changes of malfunction of parts of the plane. Microsimulation simulates the lives of particular patients with coronary artery disease (CAD), taking into account several related events that may occur during a particular remaining life expectancy. If microsimulation of a particular patient is repeated numerous times, a “virtual” patient population is created, consisting of patients with identical characteristics and with all possible outcomes one can think of. This is the main power of microsimulation; it actually simulates individual life histories of numerous (for example 100,000) virtual patients with the same characteristics, allowing insight into all probable outcomes for that particular patient and the importance of the individual-related events. A typical run would be for 10,000 to 1,000,000 individuals. From this large dataset with identical patients the average prognosis of an individual patient with those characteristics can be calculated.

COMPARISON OF THE MODELS

Logistic and Cox regression methods are practical tools used to model the relationships between certain outcomes and their relevant explanatory variables. Making no assumption about the distribution of the independent variables is the advantage of the logistic regression model; but its limitation is that it calculates changes in the log odds of the independent, not changes in the dependent itself. On the other hand, the advantage of the Cox model is investigating several variables at a time (simultaneously); however, Cox’s method does not assume a particular distribution for the survival times, but rather assumes that the effects of the different variables on survival are constant over time and are additive in a particular scale.

A key feature of the dynamic logistic regression model is the use of conjugate prior and posterior distributions for the exponential family parameters. This leads to the calculation of closed, standard-form predictive distributions for forecasting and model criticism. The structure of model depends on the time evolution of underlying state variables, and the feedback of observational information to these variables is achieved using linear Bayesian prediction methods. However, the distributions are not always in standard-form and linear.

In a Markov model, the uncertain events are modeled as transitions between defined health states, depending on individual patient characteristics and histories. However, the Markov model has “no memory” implying that subjects in a particular state are a homogeneous group without variability. It can get very complex when dealing with extensive variability within a population.

In contrast to the Markov model, which simulates the outcome of large cohorts of patients, microsimulation directly simulates the life histories of individual patients. This is done by aging the individual, modeling progression of the disease, and by updating the individual’s disease status, as he or she passes through the model. Instead of dividing the time horizon into cycles during which events may occur or not, as done in the Markov models, microsimulation uses the probability distributions of time-to-events. The time-to-events for each individual patient are obtained by computer-generated random numbers. This process of drawing numbers resembles a casino game, hence the name Monte Carlo simulation after the casinos in Monte Carlo, Monaco.
Basic principles of the microsimulation model for treatment of patients with CAD

Microsimulation, a technique developed in the field of operational research, has been used to model disease screening programs\cite{19,20} and in health economics.\cite{21} The essential ingredients are the use of computer-based sampling, and an analysis that is conducted at the maximally disaggregated level, that is, that of the individual (which might be a person, a couple, a firm or organization – whatever is the fundamental analytic unit at hand). Microsimulation techniques have also been used sporadically in clinical medicine studies.\cite{16} These techniques may be more appreciate compared to standard methods because they allow modeling of complex outcome paths resulting from many simultaneous risks. In order to make predictions using microsimulation it is necessary to obtain real-life estimates of the occurrence of related events and the effect they have on prognosis.\cite{16} In other words, the clinical evidence from real-life practice is used to feed the model with information on outcomes. At first, with systematic review (meta-analysis) or primary data on survival, occurrence of events and their consequences, we design a general microsimulation model. The model will give us estimates of outcomes such as: life expectancy (LE), event free life expectancy (EFLE), and lifetime risk of events. The model proposed this paper should assist cardiologists in making the optimal treatment choice for a particular patient.

Figure 3 represents an example of schematic representation of the basic principles of stochastic microsimulation of treatment for a patient with stable coronary artery disease. After adjusting the age of death because of background mortality (BM) based on general population mortality, there are probabilities for death or life of patients after treatment. Alive patients might be with or without symptoms. Alive patients can die with or without event. Sometimes we need to change treatment.

The model simulates one individual life at a time, from their age at diagnosis or treatment to death; the process is repeated until all simulated lives in the run have been completed. In other words, the microsimulation is designed to easily implement “what-if” scenarios. A simple change in input parameters provides a new estimate of the summary measures of health. For instance, we could estimate the population-attributable fraction of heart disease arising from obesity, by eliminating obesity in a “what-if” scenario. More realistic policy relevant scenarios can also be done to evaluate which interventions would lead to the greatest impact on health outcomes. The model takes into account characteristics of the patient, cardiac treatment, nondisease-related events, and disease-related events which can influence the outcome, allows defining the outcome useful to a cardiologist, incorporating the possibility to indicate how uncertain the output is, due to uncertain input, and is being as adjustable as possible. There is also included a method to consider statistically uncertain input, and to reflect its influence on the output of the model. This feature can be used for sensitivity analysis. The model is able to give required predictions with an indication of the precision of the output, even if one or more input parameters are uncertain.

CAD health states could be one of these options: no angina, stable CAD, unstable CAD, heart failure, etc. Patients with CAD can be treated with different cardiac treatments: no treatment, medication (including aspirin, beta-blockers, statins, and angiotensin converting enzyme (ACE) inhibitors), percutaneous coronary intervention (PCI), coronary artery bypass graft (CABG), and heart transplantation (possible only for patients without heart failure). The events that can happen with a patient with CAD consist of death (disease-related and non disease-related), disease-related events (including: getting symptoms like angina, worsening the symptoms, stroke, myocardial infarction (MI), and heart failure or HF), medication-related events (side effects), and non-disease-related events (accidents, cancer, and diabetes mellitus or DM).
DISCUSSION

Table 1 briefly presents the comparison of the above-mentioned methods by their advantages and disadvantages.

To provide physicians with tools for clinical decision-making, cardiovascular risk stratification models have been created for various categories of patients, such as patients with acute coronary syndromes. Microsimulation models play an increasingly important role in the prediction of outcomes for the cardiovascular field. Although some studies have developed patient-based microsimulation models in the field of aortic valve diseases, there is no study on the microsimulation model focused on prediction of outcomes in individual patients with coronary artery disease (CAD). Nevertheless, cardiac risk models are still lacking for sizeable groups of patients, including those with chronic stable angina pectoris. Furthermore, available risk stratification models are restricted to specific outcome events, and have a limited time horizon.

<table>
<thead>
<tr>
<th>Model</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| Logistic regression | • Making no assumption about the distribution  
• First choice for limited follow-up  
• Suitable for events with incidence <10% | • Calculation of changes in the log odds of the independent, not changes in the dependent itself |
| Cox regression | • Investigating several variables at the same time (simultaneously)  
• Expression by the hazard function | • Not assuming a particular distribution for the survival times |
| Simulation | • Involving risk over time  
• Importance of the timing of events  
• Happening of events more than once  
• Simulate the outcome of large cohorts of patients | • No memory  
• Very complex when dealing with extensive variability within a population  
• Not simulating the outcome of individual patients |
| Markov | • Easy estimating the impacts of a new scheme on a wide range of measures of effectiveness often difficult to measure in the field  
• Taking into account life expectancy of the patient  
• Considering of changing in hazards over time,  
• Allowing events to occur repeatedly over time  
• Adjusting hazards depending on events that occurred in the past  
• Allowing detailed insight into the life history of each virtual patient  
• Integrating multiple, complex and interrelated factors that determine outcome after disease  
• Modeling two or more diseases that occur together and allocating portions of the health impact to each  
• Taking into account the various health states that individuals experience just prior to diagnosis  
• Making projections in future years | • Simplification of real life |
| Microsimulation | • Requiring several assumptions regarding mortality and the occurrence of events  
• Limited by the quality of the input | • Calculation of changes in the log odds of the independent, not changes in the dependent itself |

Table 1: Comparison of common predicting models
There are seven advantages of microsimulation:

1. Microsimulation can be used to develop new systems and optimize their effectiveness. They can easily estimate the impacts of a new scheme by producing outputs on a wide range of measures of effectiveness. Many of these impacts are often difficult to measure in the field.

2. Microsimulation is capable of:
   - Taking into account life expectancy of the patient,
   - Considering of changing in hazards over time,
   - Allowing events to occur repeatedly over time, and
   - Adjusting hazards depending on events that occurred in the past.\[16\]

3. It allows detailed insight into the life history of each virtual patient, including the duration of the event-free period, the total number of years lived and the numbers of each of the events per patient. Standard statistical techniques for outcome analysis also address each of these issues, individually.

4. Microsimulation integrates multiple, complex, and interrelated factors that determine outcome after CAD.

5. Microsimulation can model two or more diseases that occur together and allocate portions of the health impact to each.

6. It can also take into account the various health states that individuals experience just prior to diagnosis, rather than assuming that all individuals of a given age experience the same health state. It is important to consider this heterogeneity of health states because individuals with certain risk profiles will be at higher risk of being diagnosed with more than one disease, and thus their functional health score will be lower than that of the general population for that age group. For example, smokers are at higher risk for respiratory disease and lung cancer. The microsimulation takes into account that, among those diagnosed with lung cancer, a certain proportion also have respiratory disease and thus a preference score lower than the average score for their age group.

7. Finally, the microsimulation is suited to making projections in future years: It monitors changes until death of the patient. Thus the simulation can report health outcomes in any future year and at any level of detail (by age group, gender, disease, high risk groups).\[34\] Coupled with “what-if” scenarios, it provides a powerful tool to evaluate the impact of various intervention strategies.

Three major disadvantages of microsimulation model are as follows:

1. That is a simplification of real life. By structuring the clinical problem, simplification of reality cannot be avoided.\[16\] To date, the microsimulation model only considers age and gender when calculating prognosis, while a number of other factors are also important determinants of outcome. Therefore, it is yet unable to make predictions taking into account all these important additional risk factors. However, by enabling us to make age- and gender-specific estimates of outcome, the microsimulation model is already helping us to get some rudimentary insights into overall patient prognosis. As long as the model remains an adequate representation of reality for the aims of the study, this is no problem. On the other hand, one can test whether this is the case. For example, mortality as calculated with the microsimulation model should correspond to mortality in a large real life dataset of similar patients.\[16\]

2. It requires several assumptions regarding mortality and the occurrence of events.\[16\] First of all, an extra mortality hazard after CAD is assumed. This is done because the additional mortality that is observed after CAD can only in part be explained by CAD-related events. Also, several assumptions are made with regard to the occurrence of CAD events, which is implemented in the background mortality of the model that is based on the general population.

3. It is limited by the quality of the input. Since most of the input of the model is currently obtained from pooled reported clinical evidence, the quality of the input may be adversely affected by heterogeneity between the studies and publication bias.\[16\] By means of sensitivity analysis, one can investigate the magnitude of the effect that this uncertainty may have on the outcome of the microsimulation model.

In comparison to the other common deterministic population models, which allow
proportions of a patient cohort to transition among health states over time, microsimulation (probabilistic) models replicate individual patient histories in a random fashion according to specified probability distributions. The results of a microsimulation model can inform decision makers by estimating not only mean outcomes in a large number of patients, but also the variance of outcomes in finite panels of patients, such as those that might be seen in a mid-sized medical practice, hospital, or managed care organization. It will explore key factors for consideration when choosing a microsimulation design, including the audience of interest, “memory” requirements of the model, the use of dynamic structures, the need for sensitivity analyses, and computing speed. For example, a microsimulation model can be developed to assess the cost-effectiveness of alternative drug therapies for a chronic degenerative disease.

The simulation tool, which is described in this paper, uses real statistical data reflecting the consequences of cardiac treatments applied to patients with various characteristics. Based on these data, the tool simulates the consequences of a chosen primary cardiac treatment for a particular patient. The mathematical model describes disease-related events (or complications) that can happen to a patient after a primary cardiac treatment, and possible progress of CAD. The number of the events can be arbitrary. Together with a wide range of methods to describe the statistical data, it makes the model very realistic and flexible. Simulation modeling will be used to combine information on the incidence, fatality, and (long-term) complications and their consequences of (first) manifestations of atherosclerotic diseases, as well as information on the effectiveness and safety of preventive and therapeutic interventions.

Given the complex medical decisions cardiologists face in everyday practice, the multiple interrelated factors that play a role in choosing the optimal treatment, and the continuously accumulating new evidence on determinants of outcome and treatment options for CAD, the cardiologist may potentially benefit from a clinical decision support system that accounts for all these considerations and allows for evidence-based objective selection of the optimal treatment for the patient that is sitting in his/her office.

REFERENCES
15. Law AM, Kelton WD. Simulation modeling and analysis.
Amiri and Kelishadi: Comparison of predicting models for CAD patients

17. Puvimanasinghe JP. Prognosis after aortic valve replacement with mechanical valves and bioprostheses. Rotterdam: Erasmus Medical Center; 2005.

Source of Support: Nil. Conflict of Interest: None declared.