How many times do we hear from our patients, “Doctor, what is going to happen to me?” Patients, physicians, and other providers would like to be able to predict what will happen after a major event to more rationally plan future care. Few health-related events are as dramatic and have as much impact on a patient as an acute myocardial infarction. There have been a number of algorithms for the prediction of outcomes after an acute myocardial infarction, including Thrombolysis In Myocardial Infarction (TIMI), Global Utilization of Streptokinase and tPA for Occluded arteries (GUSTO), Platelet glycoprotein IIb/IIIa in Unstable angina: Receptor Suppression Using Integrilin Therapy (PURSUIT), Predicting Risk of Death in Cardiac Disease Tool (PREDICT), and the Cooperative Cardiovascular Project (CCP).1–6

In the present issue of Circulation, Singh et al7 assessed outcomes of myocardial infarctions occurring in Olmsted County, Minn, between 1983 and 1994. The outcome measures were death and the composite of death plus recurrent myocardial infarction at 1 month and 1 year. The total number of incident myocardial infarctions was 1279, with 562 ST-segment elevation myocardial infarctions (STEMI) and 717 non–ST-elevation (NSTEMI). The MIs were found by searching for ICD-9 codes in the Olmsted County epidemiological database, and were confirmed by chart abstraction. The electrocardiograms were read according to the Minnesota code. Deaths are routinely tracked in Olmsted County, and recurrent MIs were diagnosed from admissions to Mayo Clinic or Olmsted Medical Center. The prediction scores compared were PURSUIT and TIMI (see Table 1 in the article by Singh et al7). For STEMI, the TIMI score is based on 8 clinical indicators available on admission, with scores ranging from 0 to 14. For NSTEMI, the TIMI score is based on 7 clinical indicators, with scores ranging from 0 to 7. PREDICT is based on 7 clinical indicators and includes measures of co-morbidity, in particular renal function and the Charlson co-morbidity index,8 and ranges from 0 to 24.

The statistical models from which scores such as these are derived are often evaluated in three ways, using indices of discrimination, calibration, and validation. The scores were developed in populations other than Olmsted County, and thus by testing them within Olmsted County, this study is itself a validation of PREDICT and TIMI. The scores were tested for discrimination by the \( c \) index, which is equivalent to the area under a receiver operating characteristic (ROC) curve.9 The \( c \) index tests, for any pair of patients with and without the outcome variable (death or death plus MI), calculates the fraction of pairs in which the patient with the outcome of interest has the higher score. Thus, a model with a \( c \) index of 0.5 is no better at predicting the event of interest than chance alone, while a \( c \) index of 1 offers perfect prediction. A model which, when applied to a new population, yields a \( c \) index as high as that obtained using the population from which it was developed is considered to have good validation. Calibration refers to how well the score predicts outcome over a range of scores. This can be examined graphically and may also be tested with the Hosmer-Lemeshow statistic, where a high probability value corresponds to good fit.10

So how did PREDICT and TIMI do? The ability of TIMI to discriminate death or death plus MI at 1 month and 1 year after STEMI was good, with \( c \) indices of 0.71 to 0.73 (Singh et al7, Table 3). The PREDICT score offered better \( c \) indices of 0.77 to 0.81. The ability of TIMI to discriminate death or death plus MI at 1 month and 1 year after NSTEMI was not as good, with \( c \) indices of 0.59 to 0.62. The PREDICT score again offered better \( c \) indices of 0.73 to 0.81. Adding co-morbidity to the TIMI score did improve the \( c \) indices, but they remained below that of the PREDICT score. The addition of a measure of ejection fraction significantly improved the discrimination of PREDICT (Singh et al7, Table 4). The measures of goodness-of-fit were generally quite good, except for a somewhat marginal Hosmer-Lemeshow probability value of 0.11 for TIMI to predict death at 1 year after NSTEMI.

That the PREDICT score offered better discrimination than TIMI is not entirely surprising, as PREDICT was developed from a community-based cohort, while TIMI was developed from data from clinical trials. If the variables chosen were sufficiently predictive, then it should not matter. However, results from trial-based populations may not be entirely generalizable to the community, and those characteristics that make the population within trials different may not be adequately measured. Thus, a community prediction tool should be developed within that community.

Are prediction tools useful in stratifying patients and making clinical decisions? Prediction tools may not only offer the probability of an event, but may also help identify the appropriate form of therapy if the optimal therapy has
been shown to depend on the score. For instance, in the Treat angina with Aggrastat and determine Costs of Therapy with Invasive or Conservative Strategies (TACTICS)-TIMI 18 trial for the treatment of patients with unstable angina or NSTEMI, an invasive strategy with early catheterization and revascularization as appropriate was shown to offer improved outcome compared with a conservative strategy of medical therapy and catheterization and revascularization for recurrent ischemia or a positive stress test. The benefit of an invasive strategy was not noted at a low TIMI score, was modest at an intermediate TIMI score, and was dramatic at a high TIMI score.

Most of the scores that have been proposed are quite simple, permitting bedside calculation even in one’s head if one were so disposed. There is a disadvantage to this, as the ability to predict the probability of an event is limited. With a range of 0 to 7 and only integer values, the precision of the prediction using the TIMI score for NSTEMI to predict the probability of an event is limited to broad ranges. This is clearly better with PREDICT with a range of 0 to 24, but it also remains somewhat limited. In addition, another step is needed to turn the scores into probabilities, generally with a table listing scores and associated probabilities. The type of modeling that was used to develop TIMI, PREDICT, and other scores can also be used to predict the probability of an event. The probability can then be calculated from nomograms or from tables. However, with hand-held computer devices, scores can be developed that predict the probability of an event directly, and even offer 95% confidence intervals for the prediction. In such a case, the scores can be slightly more complicated. For instance, co-morbidity improved the discrimination of TIMI and ejection fraction improved the discrimination of PREDICT. Since the ability of a prediction model to be used to in the clinical arena depends on all variables included in the model being available, models can be developed with and without co-morbidity and with and without ejection fraction. Also, scores can be developed with continuous variables expressed either continuously or in a number of groups. For instance, in TIMI-NSTEMI, age is categorized as above or below 65. The discrimination might be improved by considering age as a continuous variable or by decade. A simple program for calculating the TIMI score is already available for the Palm Pilot (Palm, Inc) operating system. Somewhat more sophisticated models to predict the probability of events with 95% confidence intervals could similarly be programmed and would be just marginally more difficult to use.

Finally, scores to predict events do not tell physicians or patients all they want to know. Patients do not just want to know their probability of another event. They want to know about their ability to function and how they will feel. They want to know about their ability to return to work or play golf. At present, physicians are largely guessing when they answer questions that patients often pose on these subjects.

Thus, the scores that have been developed are useful, but better statistical approaches, better ways of offering the calculations of scores to clinicians, and scores offering consideration of wider ranges of outcomes should be developed. As these scores and approaches are developed, they will need to be validated and tested for discrimination and calibration. This should be a fruitful area of research. The article by Singh et al offers a very good start.

Acknowledgments

The author thanks Elizabeth M. Mahoney for her careful review of the manuscript.

References


Key Words: Editorials ■ myocardial infarction ■ epidemiology ■ prognosis
Prediction Scores After Myocardial Infarction: Value, Limitations, and Future Directions
William S. Weintraub

_Circulation_. 2002;106:2292-2293
doi: 10.1161/01.CIR.0000035925.23046.0B
_Circulation_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2002 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/106/18/2292

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation_ is online at:
http://circ.ahajournals.org/subscriptions/