This editorial refers to 'Combined anatomical and clinical factors for the long-term risk stratification of patients undergoing percutaneous coronary intervention: the Logistic Clinical SYNTAX score', by V. Farooq et al., on page 3098.

Cardiothoracic surgery pioneered the use of risk models in the reporting of operative results of patients referred to myocardial revascularization. Recently, interventional cardiology has rekindled its interest in constructing specific scores for long-term risk prediction of percutaneous coronary intervention (PCI). One of them, dubbed the SYNTAX score, is a successful model with website dissemination and extensive validation, in which the final score is calculated after coronary angiography by summing the scorings assigned to each individual lesion.1 Inherent limitations of the SYNTAX score, such as the lack of clinical or ischaemia information, have lately prompted its integration into more sophisticated algorithms with increased prognostic accuracy.2

There are two ways in which a prognostic model may be useful in clinical decision-making. First, it may be collapsed by tertiles or arbitrary cut-offs to classify patients into risk categories. This knowledge assists in defining the threshold for a procedure and contributes to better resource allocation. Secondly, it can be used to estimate the prognosis of individual patients. These methods may be perceived as two ways to look at the same information, but they differ significantly. For instance, consider a category-based score (i.e. the SYNTAX score) which predicts that 19.2% of patients in the high-risk group will be deceased at 5 years from complex PCI. This model implies that among a large population of patients undergoing complex PCI, ~19% of those categorized in the high-risk group will be deceased after 5 years. Owing to the typical stochastic and time-dependent nature of prognostic models, classification into risk strata is frequently the best that can be achieved. However, while death is a binary outcome (no/yes, 0/1), the predictions are probabilities ranging between these extremes. Therefore, categorizing patients into schemes may work well in distinguishing between high and low risk, but the ability to give a prediction at the individual level is intrinsically limited.

In this issue of the Journal, Farooq and colleagues introduce the Logistic Clinical SYNTAX score, an attempt to tailor risk prediction of PCI patients to individual clinical and angiographic characteristics.3 Patient-level data from seven contemporary trials (n > 6000) were pooled together and logistic regression analyses performed to explore the independent association of 1-year death with a minimum relevant set of prognostic predictors including age, creatinine clearance, left ventricular ejection fraction, and SYNTAX score. Adding to this core model, an extended model encompassing more clinical variables was investigated. The core model was associated with superior discrimination compared with the SYNTAX score in isolation (areas under the curve 0.75 vs. 0.66) and slightly less discrimination than the extended model (0.79) for the outcome of 1-year death, while no relevant differences were noted among the three models for the combined outcome of 1-year major adverse cardiac events. The study was pragmatically complemented by simple additive score charts for the bedside calculation of the core and extended versions of the logistic Clinical SYNTAX score, in which an additional variable ("SYNTAX-like") was included to address the need for recalibrating the risk of 1-year death to patients presenting with complex angiographic presentations (i.e. left main and/or three-vessel disease).

The authors should be commended for a nice addition to the field of risk stratification, which indeed introduces a novel tool for the personalized characterization of patients undergoing PCI. Their study adds meaningfully to the evolving understanding that prognostic prediction in PCI should encompass both angiographic and clinical variables.4 However, a number of unsolved questions remain. First, a key aspect of risk prediction is to consider whether a model coming from a derivation data set is generalizable to similar patients from alternative cohorts. A model will have no clinical value unless there is a demonstration that it works acceptably for patients other than those from whose data the model was obtained. This step is called validation. Importantly, validating a
score is semantically as well as conceptually different from evaluating its performance with metrics such as discrimination (i.e. distinguishing between patients who do or do not experience the event of interest) or calibration (i.e. comparing observed and predicted event rates for groups of patients). In the study of Farooq and colleagues, the predictive performance of the Logistic Clinical SYNTAX score was cross-validated by the sequential omission of each of the trials pooled in the derivation cohort. This kind of internal validation suffers from the well-known problem of over-optimistic prediction when a model is fitted into and evaluated with part of the same data set. Therefore, external validation is necessary for the Logistic Clinical SYNTAX score to be considered reproducible and ideally transportable to populations with a different case mix.

Another consideration deals with the way the final model is presented. The authors used logistic regression to produce an equation that provides each patient with a custom-made prediction of subsequent mortality. Because logistic regression requires complex calculations, they converted the logistic model to a simpler, additive model based on rounding of the regression coefficients, which can then be summed together without any computing assistance. In other words, while the Logistic Clinical SYNTAX score is the probability of 1-year mortality derived from a logistic regression model, the additive Clinical SYNTAX score is intended to work as a simplified approximation to the logistic probability. Indeed, this hypothesis also needs independent verification and a dedicated comparative appraisal. A simple model may have undeniable merits, but it should not sacrifice accuracy. For instance, whether the broadly accepted logistic EuroSCORE offers a distinct advantage over the additive EuroSCORE, or whether it is more accurate in estimating the probability of death for high-risk patients undergoing cardiac surgery, still raises some controversy, and the new EuroSCORE II is available only in the logistic form. Duplication of the Clinical SYNTAX score into a simpler additive and a more complex logistic algorithm probably makes less sense in the Internet era, in which the clinical variables of the core model, after external validation, could be easily integrated in the SYNTAX score web calculator, allowing for the automatic computation of the logistic formula by means of few more additional steps.

Figure 1 Evolving risk score algorithms for percutaneous coronary intervention. ACEF, age, creatinine, ejection fraction; CABG, coronary artery bypass grafting; Compos, compositional; CrCl, creatinine clearance; CSS, Clinical SYNTAX score; FSS, Functional SYNTAX score; GRC, Global Risk Classification; MI, myocardial infarction; MDRD, Modification of Diet in Renal Disease; SrCr, serum creatinine; SYNTAX, Synergy between PCI with Taxus and Cardiac Surgery.
Finally, any scoring system, be it additive or logistic, should be proved to be a workable tool before it is used for prognostic purposes. Usefulness is a function of how well the model works and is accepted in practice, not only a problem of statistical performance. As a matter of fact, prognostic indices for PCI are still underused and, in this context, simpler models may have some advantage. By merging the SYNTAX score with a parsimonious selection of clinical predictors relevant to PCI, the Logistic Clinical SYNTAX score fulfils the philosophical principle that ‘plurality is not to be postulated without necessity’ (i.e. the Occam’s razor). However, this new model enters a crowded scene of emerging risk models for PCI, many of which were found to be associated with superior accuracy compared with the SYNTAX score8–10 (Figure 1). A simple treatment algorithm based on the Global Risk Classification, a combination of the SYNTAX score and the additive EuroSCORE,4 has been recently externally validated and shown to enhance the identification of low-risk patients who could be safely and efficaciously be treated with cardiac surgery or PCI.11 Without direct comparisons in independent training data sets, whether the Logistic Clinical SYNTAX score represents a true advancement over previously developed models for PCI is unknown. Studies aimed at head-to-head comparisons of these models are warranted to enable a consensus, drive the implementation of the most accurate into existing practice guidelines, and push its acceptance into daily routine. The ideal prognostic index should be reliable, accurate, validated elsewhere (i.e. reproducible), effective in patients from different but plausibly related populations (i.e. transportable), and, most importantly, it should provide useful additional information for clinical decision-making. Without any of these features, the chance exists to lose sight of the medical prerogatives, be charmed by the appeal of risk modelling, and get lost in calculation.

Conflict of interest: none declared.

References