Predicting risk in procedures for aortic stenosis: the next step forward

Michael J. Mack

Department of Cardiovascular Surgery, The Heart Hospital Baylor Plano, Plano, TX, USA

Corresponding author. Department of Cardiovascular Surgery, The Heart Hospital Baylor Plano, 1100 Allied Drive, Plano, TX 75093, USA. Tel: +1-469-8144105; fax: +1-469-8144132; e-mail: michael.mack@baylorhealth.edu (M.J. Mack).

Keywords: Aortic stenosis • Aortic valve replacement • Risk algorithms • Transcatheter aortic valve implantation

Proﬁling risk in patients undergoing medical procedures serves many purposes [1]. First, it allows outcomes prediction for individual patients, so that both patient and caregiver can have better-informed decision-making regarding the advisability and risks of undergoing a speciﬁc medical procedure. Secondly, patients undergoing medical procedures frequently have co-morbidities that cause varying levels of risk and therefore may adversely impact the outcomes of a procedure. When comparing different modalities of treatment or different caregivers, risk adjustment allows for a balanced analysis of outcomes by accounting for the risk-factor variation between different patient cohorts. This correction allows for a more ‘level playing ﬁeld’ of outcomes assessment, and this ‘apples-to-apples’ comparison is one of the beneﬁts of clinical outcomes databases over administrative databases that have limited ability to ‘risk adjust’. Risk adjustment, therefore, allows for a more meaningful analysis of hospitals or therapies for comparative safety and/or effectiveness of treatment. One could, for example, compare two standard procedures, e.g. coronary bypass surgery vs percutaneous coronary intervention, or a new procedure with an existing standard, transcatheter aortic valve implantation (TAVI) vs surgical aortic valve replacement (SAVR) or outcomes between different centres. Public reporting of surgical outcomes in the USA is done by risk-adjusted results in which the observed outcome divided by the expected outcome is based on known patient risk factors. Without the risk adjustment that takes into account these patient-speciﬁc factors that may adversely affect outcomes, meaningful comparison is not possible.

LIMITATIONS OF RISK ALGORITHMS

When employing risk adjustment, there are important limitations that have to be taken into account in order to obtain valid information and not misinformation from the ‘correction’ [2]. First, risk algorithms are accurate only for the population and in the timeframe in which they were developed. The implications being that although both TAVI and SAVR are used in treating patients with aortic stenosis, AVR risk algorithms are based on ‘surgical’ AVR outcomes and therefore may not to be directly applicable to TAVI. Fourthly, risk algorithms cannot account for variables not collected or analysed. This lack of accounting is due to either the fact that the occurrence of the factor or condition is so small that it’s impact cannot be measured, e.g. porcelain aorta, or that it was not previously known to be a factor that was causal or able to be accurately measured or quantified. The role of frailty and its impact on outcomes of treatment is a case in point. Fifthly, all risk predictors fall prey to the phenomenon of ‘garbage in equals garbage out’. Unless, the factors on which the algorithm is formulated are based on complete and accurate data, it is more likely that an inaccurate predictor will result. A corollary of this is that the risk predictor must be ‘user friendly’. The greater the number of variables collected in the formulation of the risk algorithm, the more accurate the prediction of risk; however, the more burdensome the collection of data required, then the less complete and accurate will be the information. So there needs to be a balance between including all information that is likely to be a factor in causing risk and ‘user-friendliness’ in order to facilitate complete and accurate collection and to assure that the tool is routinely employed in decision-making. Indeed, one risk algorithm for aortic stenosis, the ACEF score, provides reasonable prediction using only three factors, such as age, serum creatinine and ejection fraction [5].

AVAILABLE RISK ALGORITHMS FOR AORTIC STENOSIS

There are at least 12 risk algorithms that have been constructed in various populations and differing time periods to predict outcomes after SAVR (Tables 1 and 2). The two most widely used are the Logistic EuroSCORE (LES) and the Society of Thoracic Surgeons (STS) predicted risk of mortality [6–8]. Both the LES and STS have been repeatedly demonstrated to over-predict the actual risk when assessing patients at high risk for surgery [3, 4]. This is due to the factors mentioned above, including too few patients at high risk to be accurately analysed and that they were operated on an earlier time period. To address some of these shortcomings, both the LES and STS risk algorithms recently have been revised and updated to reflect more accurately...
current surgical outcomes and practice. The STS is now in version 2.73 and is readily available as an online calculator at [sts.org](http://sts.org). EuroSCORE was also recently revised and updated and can also be calculated online ([euroscore.org](http://euroscore.org)) [9]. The new EuroSCORE II has been demonstrated to more accurately predict observed outcomes in patients undergoing SAVR and possibly TAVI, although still not with optimal accuracy [10–12].

There is currently intense interest in predictive modelling for the management of patients with aortic stenosis because of the introduction of TAVI. This has unfortunately led to the overuse, and indeed abuse, of the risk algorithms for applications for which they were not developed or intended [13]. Overenthusiastic zealots of TAVI have touted the benefits of outcomes of the procedure because of the better-observed outcomes compared with the expected based on the predictive model. This abuse has been particularly prevalent with the use of the LES to estimate risk and report outcomes in patients undergoing TAVI. The reasons why current risk predictive models are inaccurate, not applicable and in fact yield misinformation is that the patients assessed for TAVI are at the extremes of risk where the current risk models fail. The risk predictors are being used for a procedure for which the risk models are not developed or validated and do not take into account the variables that may play a role in risk, including porcelain aorta, previous radiation therapy, liver disease and frailty.

**GERMAN AORTIC VALVE SCORE**

To help address some of the inadequacies of the current risk-prediction models in adults undergoing aortic valve procedures, the German Aortic Valve Registry (GARY) has developed the German Aortic Valve Score [14]. It is based on 11 794 patients undergoing SAVR or TAVI in Germany in 2008. Using multiple logistic regression, 15 risk factors influencing in-hospital mortality were identified. Among the most important factors determined to predict risk were age, body mass index, renal disease, urgent status and left-ventricular function. There was a high degree of discrimination of the risk model with an area under the ROC curve of 0.808, with 1.0 being perfect and 0.5 being ‘a flip of the coin’ [15].

The authors are to be applauded for developing a risk algorithm that can finally be applied with some degree of accuracy in patients who are currently undergoing aortic valve procedures. However, rather than being viewed as the final answer, it is rather the first step in a long journey of risk prediction in TAVR. As the authors acknowledge, there are many limitations of this model. These include the fact that it was developed based on patients treated in 2008 and may already be not applicable to current treatment since the field is evolving so rapidly with new devices and techniques continuously being introduced. Secondly, patients undergoing TAVI constituted only 5.1% (573/11 147) of the study population, limiting the application to TAVI. TAVI was also performed in only 25 of the 81 participating institutions, again limiting the generalizability of the score. It should also be noted that the model was developed for interhospital comparisons and therefore can only predict overall outcomes in German hospitals. Comparisons can be made of overall programme outcomes between various centres, but it cannot be used to discriminate between different procedures, approaches or devices. One cannot as yet determine from this risk score whether an individual patient should undergo SAVR or TAVI or whether a specific device or approach could be preferable.

Another limitation of the German Aortic Valve Score is the methodology by which the risk model was constructed. Most of the risk models are developed using a portion of the overall analysed population, usually 50–60% of the study population, to construct a weighted risk model, and then use the remaining sample to validate it. Because of the small sample size of TAVI procedures in the study, this validation was not done, thus the model needs to be validated externally in other populations.

Since the model includes overall outcomes of surgical AVR and TAVI, the value is in comparing the overall outcomes of a programme and not necessarily surgical or TAVR outcomes per se. Indeed, it is also likely that there are different factors that constitute different risk profiles for different procedures. For example, frailty may be weighted more when considering SAVR compared with TAVI. The risk may not be the same when considering the different approaches for TAVR, as severe lung disease may be a significant factor impacting outcomes with the transapical, but not the transfemoral, approach.

It should also be noted that this model is based on in-hospital mortality, which is lower than the 30-day definition of mortality used by the STS algorithm. In the development of EuroSCORE II,

---

**Table 1:**

<table>
<thead>
<tr>
<th>Covariate</th>
<th>STS (Society of Thoracic Surgeons)</th>
<th>EuroSCORE (logistic)</th>
<th>EuroSCORE (additive)</th>
<th>Ambler (UK)</th>
<th>NNE—Northern New England</th>
<th>New York State</th>
<th>Providence Health System</th>
<th>VA Risk Score</th>
<th>ACF score</th>
<th>Australian—AVR score</th>
<th>EuroSCORE II</th>
<th>German Aortic Valve Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place</td>
<td>Europe (eight countries)</td>
<td>USA</td>
<td>May–July 2010</td>
<td></td>
<td></td>
<td>USA</td>
<td>2002–2006</td>
<td></td>
<td>43 countries worldwide</td>
<td>22 381</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of operations</td>
<td>14 799</td>
<td>67 292</td>
<td>All cardiac</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22 381</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of operations</td>
<td>All cardiac</td>
<td>Aortic valve only</td>
<td>All cardiac</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22 381</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariates for aortic-valve mortality</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22 381</td>
<td>24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LES: logistic EuroSCORE; STS: Society of Thoracic Surgeons.
in-hospital mortality was 4.0%, 30-day mortality 4.6% and 90-day mortality 5.5% [9].

Indeed, it is likely that TAVI-specific risk algorithms will be developed that predict both short- and long-term results and outcomes other than mortality. The linkage of clinical databases, which capture early outcomes with administrative databases that capture long-term outcomes, will allow the development of models predictive of long-term survival. The current trials in intermediate-risk patients have a primary endpoint of death and stroke at 2 years. One can envision the construction, eventually, of risk models that will predict composite outcomes, including mortality, stroke and functional quality of life. The current STS model predicts not only 30-day mortality, but individually and as a composite, six components of major morbidity including stroke.

A true TAVR-specific risk model needs to be developed, and at this time, there are at least two efforts to do so. The European registries of the Sapien Valve (Edwards Lifesciences, Irvine, CA, USA) and the USA Partner Trial and Continued Access patients are being collated and analysed to develop a TAVR-specific algorithm, while the Society of Thoracic Surgeons/American College of Cardiology Transcatheter Valve Therapy (STS/ACC TVT) Registry in the USA now has sufficient patients enrolled for a risk algorithm to be developed. The validation of a TAVR-specific risk algorithm between these two populations is planned.

Analysis of adult patients with aortic stenosis undergoing procedures is a rich area of outcomes and comparative effectiveness research. The GARY Registry has given us the first glimpse into how to accurately assess the outcomes of current therapy. As the authors envisage, they have established a baseline from which we can evolve to build even more current and robust predictive models that can serve individual patient decision-making and make accurate outcomes assessment and comparative effectiveness research possible.

REFERENCES


