Review
Predictors of mortality after aortic valve replacement

Yanto Sandy Tjang a,b,c,*, Yvonne van Hees d, Reiner Körfer b, Diederick E. Grobbee a,c, Geert J.M.G. van der Heijden a,e

a Julius Center of Health Sciences and Primary Care, University Medical Center, Utrecht, The Netherlands
b Department of Thoracic & Cardiovascular Surgery, Heart & Diabetes Center NRW, Bad Oeynhausen, Germany
c Netherlands Institutes for Health Sciences, Erasmus Medical Center, Rotterdam, The Netherlands
d Hartenzorg Foundation, St. Antonius Hospital, Nieuwegein, The Netherlands
e Department of Cardio-Thoracic Surgery/Heart Lung Center Utrecht, University Medical Center, Utrecht, The Netherlands

Received 20 March 2007; received in revised form 1 June 2007; accepted 11 June 2007; Available online 19 July 2007

Summary
Aortic valve replacement (AVR) is recommended as a standard surgical procedure for aortic valve disease. Still the evidence for commonly claimed predictors of post-AVR prognosis, in particular mortality, appears scant. This systematic review reports on the evidence for predictors of post-AVR mortality, and may be helpful in pre-surgical risk-stratification. In PubMed, we searched for original reports of post-AVR follow-up studies. We assessed the quality of study design and methods with a standardized checklist. Data of the reported predictors of mortality and outcomes were extracted. Twenty-eight studies met our inclusion criteria. Sixteen studies were considered of high quality. There is strong evidence that the risk of early mortality is increased by emergency surgery, while the risk of late mortality is increased with older age and preoperative atrial fibrillation. There is moderate evidence that the risk of early mortality is increased by older age, aortic insufficiency, coronary artery disease, longer cardiopulmonary bypass time, reduced left ventricular ejection fraction (LVEF), infective endocarditis, hypertension, mechanical valves, preoperative pacing, dialysis-dependent renal failure and valve size; and that the risk for late mortality is increased by emergency surgery and urgency of the operation. There is little evidence for high New York Heart Association class, concomitant coronary artery bypass graft and many other commonly claimed risk factors for post-AVR mortality. The reported evidence on predictors of post-AVR mortality will help for pre-surgical risk-stratification, i.e. to discern patients at high or low risk for early and late post-AVR mortality. Future prognostic studies should take the evidence from this review into account and should focus on derivation of a predictive model for post-AVR survival.

© 2007 European Association for Cardio-Thoracic Surgery. Published by Elsevier B.V. All rights reserved.

Keywords: Aortic valve replacement; Mortality; Prognosis; Systematic review

1. Introduction
Aortic valve replacement (AVR) is recommended as a standard surgical procedure for most patients with symptomatic aortic valve disease [1]. Currently, AVR accounts for 13% of all adult cardiac surgery [2], and remains the most common procedure among all cardiac valve operations in the United States [3]. The 5-year mortality of aortic valve disease without surgery is estimated to range from 50% to 80% [4]. After surgery, patients with severe aortic valve disease show dramatic improvement in their cardiovascular symptoms and survival [5,6].

Although patients undergoing AVR are at relatively high risk for post-procedural mortality, evidence for commonly claimed predictors of post-AVR prognosis, in particular mortality, appears scant [7,8]. To date, a systematic synthesis of evidence for predictors of post-AVR prognosis is lacking. Knowledge of predictors for post-AVR prognosis could be helpful in pre-surgical risk-stratification. Therefore, we systematically reviewed the available evidence for predictors of early and late post-AVR prognosis.

2. Material and methods
2.1. Identification and selection of publications
studies were categorized according to early mortality, i.e. considered statistical significant.

2.2. Appraisal of study design

With a standardized checklist for study design, two reviewers (YST, YvH) independently appraised the study design reported in selected publications. This checklist (Table 1) [9] includes each of the three items on validity and precision of methods (V1, V2, V3 and P1, P2, P3) and seven items on clinical aspects of the study design (criteria C1, C2, C3, C4, C5, C6, C7). The legend of Table 1 provides an explanation of each item. Disagreements between the two reviewers on checklist items were resolved during a consensus discussion. In case of persisting disagreement, appraisal by a third reviewer (GvdH) resulted in a decision.

Checklist items were based on theoretical considerations from the guidelines for reporting morbidity and mortality after cardiac valvular operations [10], and methods for prognostic studies [11]. The score on validity items (V1, V2, V3) was summed and used as methodological quality score. With a range from 0 to 6, low scores reflect poor methodological quality, i.e. they are likely to report biased results due to deficiencies in study methods. A priori, we graded a score of three points or more as high quality and a score of 1 or 2 points as low quality.

2.3. Data extraction

Details on study population, response, follow-up, outcome measurement and predictors were extracted from the included studies. Detailed characteristics of the included studies are presented on the Internet.1 For each study the reported data on the univariate, and if available, multivariate association between predictors and outcome were tabulated. Risks ratios, odds ratios or hazard ratio above 2 or below 0.5 was considered clinically relevant, while 95% confidence interval excluding unity or a p-value <0.05 was considered statistical significant.

To facilitate comparison and interpretation of results, the studies were categorized according to early mortality, i.e.

---

30-days (operative) mortality or hospital mortality, and late mortality (otherwise).

2.4. Data analysis

To date, methods for statistical pooling of data from prognostic studies are lacking. Therefore, we classified evidence according to four levels of evidence. The classification table included the number of studies reporting an increased or reduced risk of a poor prognosis for a particular predictor, the methodological quality of these studies, and the consistency of the evidence across these studies [12,13].

2.4.1. Strong evidence

Of at least two high quality studies on the same predictor, at least 75% report an effect in the same direction, i.e. the risk increase or decrease depending on predictor status, that is clinically important, statistically significant, or both. Strong evidence should always be incorporated in decision-making.

2.4.2. Moderate evidence

Of at least one high quality study and one low quality study on the same predictor, at least 75% report an effect in the same direction, i.e. the risk increase or decrease depending on predictor status, that is clinically important, statistically significant, or both. Moderate evidence should always be considered in decision-making.

2.4.3. Weak evidence

At least one high quality study reports an effect that is clinically important, statistically significant or both; or at least 75% of at least three low quality studies on the same predictor report an effect in the same direction that is clinically important, statistically significant, or both. Weak evidence may be considered in decision-making.

2.4.4. Inconclusive evidence

Less than three low quality cohorts, or irrespective of study quality inconsistent effects. Inconclusive evidence indicates areas where research may be needed.

3. Results

3.1. Identification and selection of the literatures

With our search we identified 1920 citations, 1711 with our PubMed search filter and 209 by screening citations and related articles. After assessment of title and abstracts 55 titles were selected for assessment of the full publications. Subsequently, 28 articles met our selection criteria and remained for critical appraisal [2,6—8,14—37]. For duplicate publications, only the most recent or most complete report was included.

3.2. Methodological quality assessment

Table 1 presents the results of the critical appraisal. Their methodological quality score ranked the studies: the sum of internal validity criteria $V_1$, $V_2$, $V_3$ (min—max: 0—6). The quality score ranged from 0 to 6 points, with a median score of 4 points. Of the 28 studies, 16 (57%) were of high quality. Among the included 28 studies, 21 (75%) did not describe blinding (item $V_3$), clearly; 17 (61%) did not fully describe (standardized) surgical procedures (item $V_2$). An additional 10 (36%) did not report on follow-up over 12 months (item $C_3$); 6 (21%) did not report standardized assessment of outcome (item $P_1$); and 6 (21%) did not include information about non-responders and responders (item $C_3$).

3.3. Study characteristics

In total 11 studies [2,6,14—16,19,21,24,33—35] reported on early and late mortality; 14 studies exclusively reported on early mortality [7,8,20,22,23,25,26,28—32,36,37] another 3 studies exclusively reported on late mortality [17,18,27]. There was considerable variation among the studies with respect to the length of follow-up (range 30 days to 187 months). The outcome was denoted as operative mortality in 13 studies [2,8,16,18,20,21,23,25,29,32—35], as either in-hospital mortality or hospital mortality in 7 studies [7,14,15,19,22,24,31], and as early mortality [6,17,27,30,36,37], or short-term mortality [26,28] in 8 studies. For 9 (32%) of the 28 studies [7,15,16,21,22,24,30—32] effects were not adequately reported.

Altogether, 106,660 patients were included. Sample size varied between 83 patients [33] and 46,397 patients [31], with age ranging from 18 to 93 years. In total, 6036 patients died during early follow-up period, giving an overall early mortality risk of 0.06, ranging from 0.02 [27] to 0.18 [36]. Studies reporting on late mortality [2,6,14—19,24,33—36] accumulated 48,682 patient-years of follow-up, while 2109 patients died. Thereby, the late mortality rate is 4.3 per 100 patient-years.

The 28 included studies reported on 91 predictors of early mortality and 38 predictors of late mortality. Table 2 shows the most frequently reported predictors of early and late mortality. Although most of the 28 studies reported outcomes

coronary artery bypass graft (CABG) were examined in a large number of studies, but the evidence remains inconclusive as to whether they increase the risk for early or late mortality. Other predictors were only reported once. Hence, the evidence on their predictive value remains inconclusive.

### 4. Discussion

This systematic review shows that there is considerable evidence for several patient and procedural characteristics as risk factors for post-AVR mortality. The risk for early post-AVR mortality is increased with emergency operation, older age, aortic insufficiency, coronary artery disease, longer cardiopulmonary bypass time, reduced LV-EF, infective endocarditis, hypertension, mechanical valve, preoperative pacing, dialysis-dependent renal failure and with increasing valve size. In addition, the risk for late post-AVR mortality increases with emergency operation, older age and preoperative atrial fibrillation. There is insufficient evidence for many other commonly claimed putative predictors of post-AVR mortality, notably NYHA class and concomitant CABG.

### Table 3

Level of evidence for important predictors of early mortality in aortic valve replacement

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Cohort assessed</th>
<th>Positive findings</th>
<th>Positive HQ</th>
<th>Positive LQ</th>
<th>Negative findings</th>
<th>Negative HQ</th>
<th>Negative LQ</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergency</td>
<td>7</td>
<td>6/7 (86%)</td>
<td>2</td>
<td>4</td>
<td>1/7 (14%)</td>
<td>0</td>
<td>1</td>
<td>Strong</td>
</tr>
<tr>
<td>Age &gt;70 years</td>
<td>5</td>
<td>4/5 (80%)</td>
<td>1</td>
<td>3</td>
<td>1/5 (20%)</td>
<td>1</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Age &gt;80 years</td>
<td>3</td>
<td>3/3 (100%)</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Aortic insufficiency</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>3</td>
<td>3/3 (100%)</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>CPB time &gt;120 min</td>
<td>4</td>
<td>3/4 (75%)</td>
<td>1</td>
<td>2</td>
<td>1/4 (25%)</td>
<td>0</td>
<td>1</td>
<td>Moderate</td>
</tr>
<tr>
<td>EF &lt; 35%</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Infective endocarditis</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Hypertension</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Mechanical valve</td>
<td>3</td>
<td>3/3 (100%)</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Preoperative pacing</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Dialysis-dependent RF</td>
<td>3</td>
<td>3/3 (100%)</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Valve size</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

**Predictors with weak evidence:** concomitant CABG ≥ 3 vessel, creatinine > 110 μmol/l, digitalis use, unstable hemodynamic, hypercholesterolemia, LV-aorta gradient > 60 mmHg, male gender, renal dysfunction, respiratory failure, smoking, systolic PAP > 50 mmHg, valve size 19 mm, and valve type.

**Predictors with inconclusive evidence:** age 76–79 years, age >71 years + height < 1.57 m, aortic cross-clamping, BSA < 1.70 m², BUN > 25 mg/100 ml, cardioplegia use, cerebrovascular disease, concomitant CABG, concomitant MVR, concomitant surgery, COPD, creatinine gradient > 110 mmHg, renal failure (post-operative) and valve type.

### Table 4

Level of evidence for important predictors of late mortality in aortic valve replacement

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Cohort assessed</th>
<th>Positive findings</th>
<th>Positive HQ</th>
<th>Positive LQ</th>
<th>Negative findings</th>
<th>Negative HQ</th>
<th>Negative LQ</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt;71 years</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Strong</td>
</tr>
<tr>
<td>Preoperative AF</td>
<td>3</td>
<td>3/3 (100%)</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Strong</td>
</tr>
<tr>
<td>Emergency</td>
<td>2</td>
<td>2/2 (100%)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
<tr>
<td>Urgency</td>
<td>3</td>
<td>3/3 (100%)</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

**Predictors with weak evidence:** concomitant CABG, concomitant cardiac operation, COPD, coronary artery disease, LV-dysfunction, male gender, mechanical valve (no warfarin), myocardial infarction, >2 comorbidity, NYHA IIIB/IV, peak gradient > 75 mmHg, renal failure (post-operative) and valve type.

AF: atrial fibrillation, HQ: high quality, LQ: low quality.

**Predictors with weak evidence:** acute pulmonary edema, age 51–60 years, age 61–70 years, age >75 years, age >80 years, diabetes mellitus, EF, EF < 35%, previous carcinoma, and race (African-American).
To our knowledge, this is the first systematic review on predictors of mortality after aortic valve replacement. We identified and scrutinized 28 prognostic cohort studies. The majority suffered from flaws in the design and conduct. The methods of only 16 of the 28 studies were of sufficient quality. There was considerable heterogeneity regarding study design; notably study population, definition of predictors, definition of outcome measures and length of follow-up. Many studies focused on univariate analyses, while some reported outcomes of multivariate analyses. Only 4 studies [22, 26, 29, 32] reported on a developed prognostic model. For a relatively large number of studies outcome measures were not clearly defined or outcome data were not clearly reported.

### 4.1. Review methods

Because publication bias is likely in every systematic review, the reported risk estimates turn away from a null result, and so are likely to be optimistic [38, 39]. After comprehensively searching PubMed, we thoroughly checked related articles and reference lists of included publications. Since we only included full English language publications, we might have missed studies. Therefore, possible language and retrieval bias that may contribute to overestimation of risk cannot be excluded [40–42]. This strengthens the conclusion that there is little evidence for NYHA class and concomitant CABG and many other commonly claimed putative predictors of post-AVR mortality.

We used an explicit and transparent method to summarize the available evidence. With a cut-off point of 40% for high quality study methods, instead of 50%, the evidence for increased early mortality risk with older age, aortic insufficiency, coronary artery disease, valve size, and for an increased late mortality risk with emergency operation, changed from moderate to strong.

### 4.2. Implications for care

The reported evidence on predictors of post-AVR mortality will help pre-surgical risk-stratification, i.e. to discern patients at high or low risk of early and late post-AVR mortality. The results from this systematic review may serve on a developed prognostic model. For a relatively large number of studies outcome measures were not clearly defined or outcome data were not clearly reported.

### 4.3. Implications for research

Sufficiently large and well-designed prognostic studies on post-AVR mortality are needed. Such future studies should take the evidence from this review into account. They should in particular focus on adequate definition of outcome measures, assert blinding of subjective endpoints, either or both standardize or accurately describe the surgical procedures and assure completeness of follow-up at sufficient length. Instead of risk ratios, odds ratios or hazard ratios, both the numerator (number of events) and the denominator (total number of patients included, their accrued person-time of follow-up, or both) should be reported accurately [44]. A multivariable analysis of predictors may contribute to stratification of patients at high or low risk for early or late post-AVR mortality. Therefore, future prognostic studies should focus on derivation of a predictive model for post-AVR survival. Moreover, as the long-term consequences of AVR extend beyond post-operative complications or survival, improvement in quality of life is receiving more attention as an important goal of surgery. Relatively little is known about changes in post-AVR quality of life. Therefore, quality of life assessment should be included in the future post-AVR outcome studies.

### References
