Advances in Abdominal Aortic Aneurysm Care - Towards personalized, centralized and endovascular care

van Beek, S.C.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

Download date: 05 Apr 2020
Chapter 3

Validation of three models predicting in-hospital death in patients with an abdominal aortic aneurysm eligible for both endovascular and open repair

Sytse C. van Beek
Jan D. Blankensteijn
Ron Balm
on behalf of the DREAM Trial Collaborators

CHAPTER 3

Abstract

Background
The Medicare, the Vascular Governance North West (VGNW), and the British Aneurysm Repair (BAR) models can be used to predict in-hospital death after an intervention for an asymptomatic abdominal aortic aneurysm (AAA). Validation of these models in patients with suitable aortic anatomy for endovascular repair and a general condition fit for open repair is lacking. We validated the Medicare, VGNW, and BAR models in patients from a randomized controlled trial comparing open and endovascular AAA repair.

Methods
A per-protocol analysis was done of 345 Dutch and Belgian patients with in-hospital death as the primary end point. The prediction models were validated taking into account discrimination (the ability to distinguish between death and survival) and calibration (the agreement between predicted and observed death rates). Discrimination was assessed using the area under the receiver-operating characteristics curve (AUC). An AUC >0.70 was considered to be sufficiently accurate. Calibration was assessed using the Hosmer and Lemeshow (HL) test, and P>.05 was considered to be sufficiently accurate.

Results
The AUC was 0.77 (95% confidence interval (CI) 0.64 to 0.90, HL test P=.52) for the Medicare model, 0.88 (95% CI 0.81 to 0.95, HL test P=.31) for the VGNW model, and 0.79 (95% CI 0.67 to 0.91, HL test P=.15) for the BAR model.

Conclusion
In AAA patients eligible for endovascular and open repair, the predictions of in-hospital death by the Medicare, VGNW, and BAR models were sufficiently accurate. Therefore, these models can be used to support deciding between endovascular and open repair.
Introduction

Patients with an abdominal aortic aneurysm (AAA) can be treated with endovascular repair, open repair, or conservatively. When deciding between endovascular and open repair, the clinical decision-making process is predominantly based on estimates of the incidence of mortality, reinterventions, and complications as well as on the expected quality-adjusted life-years. The challenge in current practice is to determine which patients will benefit the most from endovascular repair and which from open repair. Randomized trials have shown long-term survival and quality-adjusted life-years are equal after both interventions. However, the incidence of complications and reinterventions is higher after endovascular repair, whereas the incidence of in-hospital death is higher after open repair. Therefore, predicting in-hospital death and long-term reinterventions before the intervention could support clinical decision making.

A prediction model is a standardized and objective way to assess individual outcomes after an intervention. Several models predicting in-hospital death after aortic repair have been developed; for example, the prediction model most frequently used is the well-validated Glasgow Aneurysm Score (GAS), which was developed in 1994. The results of validation studies evaluating the accuracy of the GAS in endovascular repair are conflicting. Three new prediction models have recently been developed: the Medicare model in the United States and the Vascular Governance North West (VGNW) and the British Aneurysm Repair (BAR) models in the United Kingdom.

Two validation studies have reported sufficiently accurate predictions of in-hospital death for the Medicare and VGNW models. These validations included patients whose aortic anatomy was unsuitable for endovascular repair and whose general condition was unfit for open repair. As such, these studies included patients where there was no choice between endovascular and open repair. To our knowledge, the BAR model has not yet been validated externally.

To apply the Medicare, VGNW, and BAR models to support the decision between endovascular and open repair in individual patients, validation is needed in patients eligible for both options. Therefore, patients enrolled in randomized trials comparing endovascular and open repair, such as the Dutch Randomized Endovascular Aneurysm Management (DREAM) trial, can be used to validate the models.
As well as supporting decision making, prediction models have additional value in improving patient education. In the consulting room, prediction models can be used to advise patients and relatives of the short-term risk of dying after the intervention.

The objective of this study was to validate the Medicare, VGNW, and BAR prediction models in Dutch and Belgian patients with an AAA who were eligible for both open and endovascular repair.

**Methods**

This study retrospectively analyzed 345 patients included in the DREAM trial. Details of the DREAM trial (registration number clinicaltrials.gov NCT00421330) are described in detail elsewhere. Briefly, the DREAM trial was a multicenter, randomized trial conducted at 26 hospitals in The Netherlands and in four hospitals in Belgium. Inclusion criteria for the DREAM trial were informed consent, an AAA sized at least 5 cm, and suitability for both endovascular and open repair. Excluded were patients with an inflammatory aneurysm, anatomic variations, connective tissue disease, a history of organ transplantation, or a life expectancy of <2 years.

The study was conducted in accordance with the principles of the Declaration of Helsinki. Approval for the DREAM trial was given by the Institutional Review Board of all hospitals.

**Statistical analysis**

The analysis was done per-protocol. The primary end point was the combination of 30-day and in-hospital death. Statistical analysis was done using IBM SPSS 19.0 software (IBM, Armonk, New York) and R software (The R Foundation for Statistical Computing, Vienna, Austria). The predictions of death by the Medicare, VGNW, and BAR models were calculated with the formulas presented in Table 1.

The accuracy of the predictions was assessed taking discrimination and calibration into account. Discrimination is the ability of a model to distinguish between those patients who die and those who survive. Discrimination was assessed using the area under the receiver-operating characteristics curve (AUC). An AUC >0.70 is generally considered to be sufficiently accurate. Calibration refers to the agreement between the predicted and observed death rates and was
Predictions of short-term survival after elective EVAR and OR

Table 1. Formula and definitions of the Medicare, the Vascular Governance North West (VGNW), and the British Aneurysm Repair (BAR) prediction models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Score</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare</td>
<td>-5.02 + age &lt;75 years x 0.15 + age 75-80 years x 0.63 + age &gt;80 years x 1.14 + female sex x 0.42 + chronic renal insufficiency x 0.71 + end-stage renal disease x 0.95 + congestive heart failure x 0.55 + vascular disease x 0.30 + open repair x 1.17</td>
<td></td>
</tr>
<tr>
<td>VGNW</td>
<td>-9.3431 + age (years) x 0.0486 + female sex x 0.7322 + diabetes x 0.6620 + creatinine (µmol/L) x 0.0073 + respiratory disease x 0.4718 + antiplatelet medication x 0.7762 + open repair x 1.3130</td>
<td></td>
</tr>
<tr>
<td>BAR</td>
<td>-10.9187 + open repair x 1.6466 + age (years) x 0.0568 + female sex x 0.7062 + creatinine &gt;120 µmol/L x 0.5979 + abnormal ECG x 0.3033 + previous aortic surgery or stent x 0.8812 + abnormal white cell count x 0.3697 + abnormal sodium level x 0.3099 + AAA diameter (cm) x 0.1285 + ASA 2 x 0.2292 + ASA 3 x 0.7334 + ASA 4 x 1.6775</td>
<td></td>
</tr>
</tbody>
</table>

1 SVS/ISCS renal status ≥1 (equal or worse condition, then moderately elevated creatinine level as high as 220 µmol/L)
2 Need for dialysis
3 SVS/ISCS cardiac status ≥2 (equal or worse condition, then stable angina, ejection fraction between 25% and 45%, asymptomatic arrhythmia, or history of congestive heart failure)
4 SVS/ISCS carotid disease ≥2 (equal or worse condition, then transient or temporary stroke), ankle-brachial index <0.90, or previous history of peripheral artery surgery
5 SVS/ISCS diabetes ≥1 (equal or worse condition, then adult-onset diabetes controlled by diet or oral agents)
6 SVS/ISCS pulmonary status ≥1 (equal or worse condition, then mild dyspnea on exertion, parenchymal X-ray changes, or pulmonary function tests between 65% and 85% of predicted)
7 SVS/ISCS cardiac status ≥1 (remote myocardial infarction by history of >6 months, occult myocardial infarction by electrocardiogram, or fixed defect on dipyridamole thallium or similar scan)
8 White cell count <3.0 x 10^9/L or >11.0 x 10^9/L
9 Sodium level <135 mmol/L or >145 mmol/L

assessed using a graph plotting the mean predicted death rates in tertiles with the corresponding observed death rates. The tertiles were created by sorting the predictions in ascending order and categorizing the patients in three subgroups of comparable size accordingly. The subgroups included 126, 119, and 100 patients for the calibration of the Medicare model, 114, 116, and 115 patients for the VGNW model, and 106, 118, and 121 patients for the BAR model. The sizes of these subgroups differed slightly because patients with equal predictions were categorized in the same tertile. Calibration was also assessed using the Hosmer and Lemeshow (HL) test. The HL test compares predicted and observed outcomes in a χ² distribution. An HL test P value <.05 reflects statistically significant differences between predicted and observed outcomes. Hence, a P value >.05 indicates sufficiently accurate calibration of the model.
In two patients, the preoperative serum creatinine was missing and was imputed as the mean creatinine. The preoperative aneurysm diameter was missing in two other patients and was imputed as the mean diameter. For the BAR model, the white cell count and the sodium level were unknown and assumed to be within normal reference ranges.

Results

In the study, 351 patients were randomized: 173 were assigned to endovascular repair and 178 to open repair. The treatment allocation flowchart is published elsewhere. Six patients were excluded because they did not undergo aneurysm

<table>
<thead>
<tr>
<th>Variable</th>
<th>EVAR n = 175</th>
<th>OR n = 170</th>
<th>Total n = 345</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>71 (67-75)</td>
<td>70 (66-75)</td>
<td>70 (66-75)</td>
</tr>
<tr>
<td>Male : Female</td>
<td>93% : 7% (163 : 12)</td>
<td>90% : 10% (153 : 17)</td>
<td>92% : 8% (316 : 29)</td>
</tr>
<tr>
<td>Previous aortic surgery or stent</td>
<td>0 (0/175)</td>
<td>1% (1/170)</td>
<td>1% (1/345)</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>8% (14/175)</td>
<td>7% (12/170)</td>
<td>8% (26/345)</td>
</tr>
<tr>
<td>Cardiac disease</td>
<td>42% (74/175)</td>
<td>46% (78/170)</td>
<td>44% (151/345)</td>
</tr>
<tr>
<td>Abnormal ECG</td>
<td>42% (74/175)</td>
<td>45% (77/170)</td>
<td>44% (152/345)</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>27% (47/175)</td>
<td>18% (31/170)</td>
<td>23% (78/345)</td>
</tr>
<tr>
<td>Serum creatinine (µmol/L)</td>
<td>96 (83-109)</td>
<td>95 (84-107)</td>
<td>95 (84-108)</td>
</tr>
<tr>
<td>Chronic renal insufficiency</td>
<td>7% (13/175)</td>
<td>7% (12/170)</td>
<td>7% (25/345)</td>
</tr>
<tr>
<td>End-stage renal disease</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Serum creatinine &gt;120 µmol/L</td>
<td>13% (22/175)</td>
<td>12% (20/170)</td>
<td>12% (42/345)</td>
</tr>
<tr>
<td>Vascular disease</td>
<td>32% (56/175)</td>
<td>27% (46/170)</td>
<td>30% (102/345)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>10% (18/175)</td>
<td>9% (16/170)</td>
<td>10% (34/345)</td>
</tr>
<tr>
<td>Antiplatelet medication</td>
<td>40% (70/175)</td>
<td>41% (69/170)</td>
<td>40% (139/345)</td>
</tr>
<tr>
<td>AAA diameter (cm)</td>
<td>5.8 (5.5-6.5)</td>
<td>5.8 (5.4-6.4)</td>
<td>5.8 (5.4-6.5)</td>
</tr>
<tr>
<td>ASA 2</td>
<td>92% (161/175)</td>
<td>85% (146/170)</td>
<td>89% (307/345)</td>
</tr>
<tr>
<td>ASA 3</td>
<td>8% (14/175)</td>
<td>14% (24/170)</td>
<td>11% (38/345)</td>
</tr>
<tr>
<td>ASA 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Continuous data are presented as median (inter-quartile range) and categorical data as percentage (number).
Predictions of short-term survival after elective EVAR and OR repair. In the intention-to-treat analysis, 171 patients were included in the endovascular repair group and 174 in the open repair group. One patient randomized to endovascular repair underwent open repair, and five patients randomized to open repair crossed over to endovascular repair. Ultimately, the per-protocol analysis included 175 patients in the endovascular repair group and 170 patients in the open repair group. Three of the patients treated with endovascular repair were converted to open repair perioperatively, and one procedure was aborted. The baseline characteristics of the patients included in the per-protocol analysis are reported in Table 2. The death rate was 1% (2/175, 95% confidence interval (CI) 0 to 4%) after endovascular repair and 5% (8/170, 95% CI 2 to 9%) after open repair.

**Figure 1.** The discrimination of the Medicare, the Vascular Governance North West (VGNW), and the British Aneurysm Repair (BAR) prediction models indicated by the area under curve and the surrounding 95% confidence intervals. An area under the curve >0.70 was considered as sufficiently accurate (indicated by the dashed line).

**Medicare model**

The median predicted death rate of the Medicare model was 1% (inter-quartile range (IQR) 1-1%, range 1-4%) in patients treated with endovascular repair and 3% (IQR 2-4%, range 2-12%) in patients treated with open repair. The AUC of the predictions was 0.77 (95% CI 0.64 to 0.90, Figure 1). The plot showed close to ideal calibration (Figure 2) and a P=.47 for the HL test. In the tertile of patients with the highest predictions, the mean predicted death rate was 5%, and the corresponding observed death rate was 6% (95% CI 3 to 13%).
Figure 2. Calibration plots of the Medicare, the Vascular Governance North West (VGNW), and the British Aneurysm Repair (BAR) prediction models. The mean predicted death rates in tertiles are plotted with the corresponding observed death rates. The range bars indicate the 95% confidence interval and the diagonal dashed line corresponds with ideal calibration.

VGNW model
The median predicted death rate of the VGNW model was 1% (IQR 1-2%, range 0-6%) in patients treated with endovascular repair and 4% (IQR 2-5%, range 1-14%) in patients treated with open repair. The AUC of the predictions was 0.88 (95% CI 0.81 to 0.95, Figure 1). The plot showed close to ideal calibration (Figure 2) and a $P=.24$ for HL test. In the tertile of patients with the highest predictions, the mean predicted death rate was 5%, and the corresponding observed death rate was 8% (95% CI 4 to 14%).

BAR model
The median predicted death rate of the BAR model was 0% (IQR 0-1%, range 0-2%) in patients treated with endovascular repair and 2% (IQR 1-4%, range 0-13%) in patients treated with open repair. The AUC of the predictions was 0.79 (95% CI 0.67 to 0.91, Figure 1). The plot showed close to ideal calibration (Figure 2), with $P=.15$ for the HL test. In the tertile of patients with the highest predictions, the mean predicted death rate was 3%, and the corresponding observed death rate was 6% (95% CI 3 to 12%).
Discussion

The predictions of death by the Medicare, VGNW, and BAR prediction models were sufficiently accurate in Dutch and Belgian patients with an AAA eligible for both open and endovascular repair.

The discrimination of the Medicare model (AUC=0.77) was comparable with two previous validation studies from the United Kingdom. In these validations, the AUC of the predictions was 0.71 (95% CI 0.69 to 0.74) and 0.79 (95% CI 0.73 to 0.86). The discrimination of the VGNW model (AUC=0.88) was higher than in two previous validations reporting an AUC of 0.71 (95% CI 0.68 to 0.74) and 0.73 (95% CI 0.65 to 0.81). To our knowledge, no previous external validation of the BAR model has been done.

A striking observation was the high accuracy of the VGNW model. First, the AUC (0.88) in our validation was high compared with other surgical prediction models. For example, the European System for Cardiac Operative Risk Evaluation (EuroSCORE) is a reliable prediction model widely used in daily practice in cardiac surgery. A validation study showed an AUC of 0.79 for the EuroSCORE. Second, the AUC of the predictions by the VGNW model was higher than the AUC of the Medicare and of the BAR models in our validation. This indicates that the predictions by the VGNW model were more accurate. However, given the limitations of our validation and the equivalent AUCs in previous validations, no definite conclusions can be drawn.

From a practical perspective, the Medicare and VGNW models require only a few patient characteristics, and the predictions can be calculated within a minute. The BAR model, however, requires several more patient characteristics and is thereby more complex compared with the Medicare and VGNW models. The BAR model was primarily developed for risk adjustment in mortality outcome analyses in the United Kingdom, which explains the higher complexity. Extra diagnostic assessments are required, including an electrocardiogram, the serum sodium level, and the white cell count. The latter is not routinely measured before intervention in The Netherlands and Belgium. Therefore, the Medicare and VGNW models have a clear practical advantage over the BAR model in our clinical practice.

The variables included in the three models correspond largely, which is suggestive for an accurate representation of a patient’s risk profile. Age, female sex, renal comorbidity, generalized atherosclerosis, and open repair increase the
prediction of death. In the Medicare model, atherosclerosis is represented by the variables ‘congestive heart failure’ and ‘vascular disease.’ In the VGNW model, ‘antiplatelet medication’ is used as a surrogate marker for atherosclerosis. In the BAR model, ‘cardiac disease,’ an ‘abnormal electrocardiogram,’ and ‘previous aortic surgery or stent’ are used to represent atherosclerosis.

Decision making
Our validation was done in patients in whom a decision between endovascular and open repair was relevant; that is, patients with aortic anatomy suitable for endovascular repair and in a general condition fit for open repair. Table 3 provides an example of six imaginary patients in whom the models could support decision making. The BAR model was not included in Table 3 because of the previously discussed lower applicability in Dutch clinical practice.

Table 3. Predictions of in-hospital death by the Medicare and Vascular Governance North West (VGNW) prediction model in examples of six imaginary patients.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Characteristics</th>
<th>Medicare</th>
<th>VGNW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66 years, male, previous history of diabetes, creatinine of 100 µmol/L, no antiplatelet medication</td>
<td>1% 2%</td>
<td>1% 3%</td>
</tr>
<tr>
<td>2</td>
<td>69 years, male, no previous history, creatinine of 85 µmol/L, no antiplatelet medication</td>
<td>1% 2%</td>
<td>1% 2%</td>
</tr>
<tr>
<td>3</td>
<td>71 years, male, previous history of chronic renal insufficiency, heart failure, and peripheral arterial occlusive disorder; creatinine of 190 µmol/L, antiplatelet medication</td>
<td>4% 11%</td>
<td>2% 8%</td>
</tr>
<tr>
<td>4</td>
<td>75 years, female, previous history of transient ischemic attack, creatinine of 76 µmol/L, antiplatelet medication</td>
<td>3% 8%</td>
<td>3% 9%</td>
</tr>
<tr>
<td>5</td>
<td>81 years, female, previous history of heart failure, diabetes, and peripheral arterial occlusive disorder, creatinine of 100 µmol/L, antiplatelet medication</td>
<td>7% 19%</td>
<td>8% 23%</td>
</tr>
<tr>
<td>6</td>
<td>82 years, male, previous history of vascular disease and COPD GOLD II, creatinine of 110 µmol/L, antiplatelet medication</td>
<td>3% 8%</td>
<td>4% 12%</td>
</tr>
</tbody>
</table>
Patients 1 and 2 are relatively young, which means open repair can be considered to prevent intensive yearly follow-up. The models support a choice for open repair by a relatively low predicted in-hospital death rate of between 2% and 3%. Patients 3 and 4 are somewhat older, at 71 and 75 years, and on the basis of their ages, open repair can be considered. However, given the relatively high predicted in-hospital death rate of between 8% and 11% after open repair, this might not be the best choice. Patients 5 and 6 are relatively old and, considering the results of the United Kingdom EndoVascular Aneurysm Repair 2 (EVAR-2) trial, conservative treatment is a reasonable option in these patients. The predicted in-hospital death rate for patient 5 is between 7% and 8% after endovascular repair and could support a choice for conservative treatment. The predicted death rate for patient 6 is between 3% and 4% after endovascular repair, which could justify an intervention.

As mentioned before, the randomized trials have shown that the in-hospital death rate and long-term reintervention rate differ after endovascular and open repair. Our validation shows that the Medicare, VGNW, and BAR models are useful tools to predict in-hospital death and support decision making on this outcome. Other models are needed to predict reinterventions and adverse events, such as aneurysm-related death and endograft-related complications, to further support decision making. An example of such a model is the Endovascular Aneurysm Repair Risk Assessment (ERA) model, which is designed to predict survival, endograft-related complications, and reinterventions after endovascular repair. However, the predictions of adverse events by the ERA model have not been as accurate as hoped. Possibly, a ‘meta-regression’ of the finalized randomized trials with uniform measurements of aortic anatomy can provide sufficiently accurate predictions of adverse events. An important characteristic of such a model would be the accuracy in patients in whom risk assessment is most needed and might support decision making. For example, the risk of adverse events is higher in patients with hostile aortic anatomy, and open repair might be a reasonable alternative. Moreover, in patients with severe comorbidity (patients 5 and 6 in Table 3), conservative treatment instead of EVAR is defendable based on results of the EVAR-2 trial.

Limitations
An important limitation of the validation of the BAR model was the unknown white cell count and sodium level. The effect on the discrimination is unknown,
and the calibration might be underestimated. However, in the developing cohort of the BAR model, the prevalence of abnormal outcomes of white cell count and sodium level was only 10%. Moreover, the contribution of these variables to the predictions are relatively small compared with the other included variables, shown by the lower coefficients and Wald Z statistics in the model.\textsuperscript{12} Therefore, we expect a limited effect of the unknown white cell count and sodium level on our conclusions. However, more studies are needed to confirm the external validity of the BAR model.

The primary end point of our validation was in-hospital death. Originally, the VGNW model was designed to predict 30-day death. We used in-hospital death because from a patient’s perspective, dying more than 30 days after the intervention but during the same hospital admission period cannot be considered a success.

Another limitation of our validation is that the DREAM trial excluded patients with severe comorbidity. All three models were developed in cohorts that included patients with severe comorbidity; therefore, we expect that these models take a patient’s severe comorbidity into account. Moreover, in previous validation studies, the Medicare and the VGNW models showed accurate predictions.\textsuperscript{7, 13} Our study focuses on patients eligible for both interventions. Patients with severe comorbidity are usually not eligible for both interventions; therefore, we expect the exclusion of patients with severe comorbidity had a limited effect on our conclusions.

Another limitation of our validation is the small sample size leading to a low event rate. As a consequence, the calibration plots showed large CIs surrounding the point estimates of the observed death rates (Figure 2).

One final limitation is that the inclusion period of the DREAM trial was about a decade ago, and intensive care unit and anesthetic care have improved since then. Moreover, all patients included in the validation were treated with an early-generation endograft. Although the type of endograft does not seem to have a major effect on the in-hospital death rate, the influence of these differences on the validity of our results in current practice is unknown.

Conclusions
The predictions of in-hospital death by the Medicare, VGNW, and BAR models were sufficiently accurate in patients eligible for both endovascular and open repair. The BAR model is more complex, has limited additional value in Dutch
Predictions of short-term survival after elective EVAR and OR

clinical practice, and needs further external validation. Therefore, the Medicare and VGNW models can be used to support deciding between endovascular and open repair in The Netherlands and Belgium and to advise patients and relatives about the risk of death after the intervention.
Reference List


Predictions of short-term survival after elective EVAR and OR


