

Editor's Choice - External Validation of Models Predicting Survival After Ruptured Abdominal Aortic Aneurysm Repair **CME**

S.C. van Beek ^a, J.J. Reimerink ^a, A.C. Vahl ^b, W. Wisselink ^c, R.J.G. Peters ^d, D.A. Legemate ^a, R. Balm ^{a,*}, on behalf of the Amsterdam Acute Aneurysm Trial Collaborators ^e

^a Department of Vascular Surgery, Academic Medical Centre, Amsterdam, The Netherlands

^b Department of Vascular Surgery, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands

^c Department of Vascular Surgery, VU University Medical Centre, Amsterdam, The Netherlands

^d Department of Cardiology, Academic Medical Centre, Amsterdam, The Netherlands

WHAT THIS PAPER ADDS

In patients with a ruptured abdominal aortic aneurysm (RAAA), prediction models could support the decision to select surgical or conservative treatment. In the present study the prediction model “updated GAS” most accurately predicted death, and accuracy was further improved after recalibration. In future clinical practice, the predictions can be used for case mix comparison between hospitals and tailoring the prognosis for patients and relatives. However, the updated GAS was insufficiently accurate to identify patients who would die despite intervention. Therefore, future studies should aim to improve the identification of these high risk patients to support the decision to withhold intervention.

Objective: Prediction of survival after intervention for ruptured abdominal aortic aneurysms (RAAA) may support case mix comparison and tailor the prognosis for patients and relatives. The objective of this study was to assess the performance of four prediction models: the updated Glasgow Aneurysm Score (GAS), the Vancouver scoring system, the Edinburgh Ruptured Aneurysm Score (ERAS), and the Hardman index.

Design, materials, and methods: This was a retrospective cohort study in 449 patients in ten hospitals with a RAAA (intervention between 2004 and 2011). The primary endpoint was combined 30 day or in hospital death. The accuracy of the prediction models was assessed for discrimination (area under the curve [AUC]). An AUC >0.70 was considered sufficiently accurate. In studies with sufficiently accurate discrimination, correspondence between the predicted and observed outcomes (i.e. calibration) was recalculated.

Results: The AUC of the updated GAS was 0.71 (95% confidence interval [CI] 0.66–0.76), of the Vancouver score was 0.72 (95% CI 0.67–0.77), and of the ERAS was 0.58 (95% CI 0.52–0.65). After recalibration, predictions by the updated GAS slightly overestimated the death rate, with a predicted death rate 60% versus observed death rate 54% (95% CI 44–64%). After recalibration, predictions by the Vancouver score considerably overestimated the death rate, with a predicted death rate 82% versus observed death rate 62% (95% CI 52–71%). Performance of the Hardman index could not be assessed on discrimination and calibration, because in 57% of patients electrocardiograms were missing.

Conclusions: Concerning discrimination and calibration, the updated GAS most accurately predicted death after intervention for a RAAA. However, the updated GAS did not identify patients with a $\geq 95\%$ predicted death rate, and therefore cannot reliably support the decision to withhold intervention.

© 2014 European Society for Vascular Surgery. Published by Elsevier Ltd. All rights reserved.

Article history: Received 5 March 2014, Accepted 11 October 2014, Available online 3 December 2014

Keywords: Aortic aneurysm, Abdominal, Aortic rupture, Decision support techniques, Prediction model, Ruptured abdominal aortic aneurysm

CME To access continuing medical education questions on this paper, please go to www.vasculareducation.com and click on ‘CME’

^e The collaborators are listed in [Online supplement](#).

DOI of original article: <http://dx.doi.org/10.1016/j.ejvs.2014.10.019>

* Corresponding author. R. Balm, Department of Vascular Surgery, Academic Medical Centre, Meibergdreef 9, 1105 AZ Amsterdam, PO Box 22700, 1100 DE Amsterdam, The Netherlands.

E-mail address: r.balm@amc.nl (R. Balm).

1078-5884/\$ — see front matter © 2014 European Society for Vascular Surgery. Published by Elsevier Ltd. All rights reserved.

<http://dx.doi.org/10.1016/j.ejvs.2014.10.012>

INTRODUCTION

The overall death rate in patients with a ruptured aneurysm of the abdominal aorta (RAAA) is approximately 74% (95% confidence interval [CI] 72–77%).¹ In patients reaching the hospital and undergoing intervention, the death rate ranges between 24% and 49%.^{2–4} Surgeons have proposed distinguishing between those who would potentially benefit from surgery, and those in whom it might be better to withhold intervention, for example after cardiopulmonary resuscitation.^{5–9} In current clinical practice, the decision to

start surgical or conservative treatment is based on a fast evaluation of the patients' clinical condition, the surgeon's experience, and the wishes of the patient. It is a subjective interpretation of a harsh reality by the doctor, the patient, and the relatives. A prediction model is a more standardized and objective way to evaluate the chances of successful intervention and might be helpful at these moments of vital choices. Further benefits of prediction models lie in case mix comparison between hospitals and a tailored prognosis for patients and relatives.

Several models have been developed to predict death after intervention in patients with a RAAA: the Glasgow Aneurysm Score (GAS),¹⁰ the Vancouver scoring system,¹¹ the Edinburgh Ruptured Aneurysm Score (ERAS),¹² and the Hardman index.⁶ These scoring systems were initially designed before the introduction of endovascular aneurysm repair (EVAR). Nowadays, EVAR is being carried out increasingly.² Only the GAS has been updated to the era of EVAR by the addition of a variable for type of intervention.¹³

The primary objective of the present study was to assess the accuracy of the updated GAS (the model including differentiation between EVAR and OR), the Vancouver score, the ERAS, and the Hardman index in predicting death. Only extremely reliable models, those predicting death accurately in more than 95% of cases, may be useful in clinical decision making. A secondary objective was the assessment of accuracy in patients with a predicted death rate of $\geq 95\%$ in whom withholding intervention might be considered.

MATERIALS AND METHODS

A retrospective study was conducted in all consecutive surgically treated patients with a RAAA in the Amsterdam ambulance region between May 2004 and February 2011. The present study was carried out as a sequel to the previously published Amsterdam Acute Aneurysm Trial.¹⁴ Other details and analyses of this cohort have been published previously.^{14,15} None of these previous studies aimed to validate prediction models for patients with a RAAA. The Amsterdam ambulance region covers an area of 1025 km² with 1.38 million inhabitants.¹⁶ During the inclusion period, care for patients with a RAAA was centralized in two university hospitals and one teaching hospital in cooperation with seven regional hospitals. All patients with a RAAA in all ten hospitals of the region were registered prospectively by the vascular surgeons, and included in the present study. Patients with a previous aortic reconstruction, a RAAA with associated trauma or aortoenteric fistula, were excluded. The primary endpoint was the combined 30 day or in hospital death rate. Compared with some previous validation studies of the prediction models, in hospital death was added to the definition; from a patients' perspective the ultimate goal is survival and being discharged. Approval from a medical ethics committee was not needed because of the observational design. This study adhered to the STrengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines.¹⁷

Updated GAS

The updated GAS score was calculated with the formula: age (years) + 7 for cardiac comorbidity (defined as previous history of myocardial infarction, cardiac surgery, angina pectoris or arrhythmia) + 10 for cerebrovascular comorbidity (defined as previous history of stroke or transient ischemic attack) + 17 for shock (defined as an in hospital systolic blood pressure <80 mmHg) + 14 for renal insufficiency (defined as a pre-operative serum creatinine >160 $\mu\text{mol/L}$) + 7 for OR (Fig. S1, online supplement).

Vancouver score

The Vancouver score was calculated with the formula: age (years)*0.062 + loss of consciousness (yes = 1/no = -1)*1.14 + cardiac arrest (yes = 1/no = -1)*0.6 (Fig. S2, online supplement).

ERAS

The ERAS score was calculated with the formula: +1 for best recorded in hospital Glasgow coma scale (GCS) <15, +1 for in hospital systolic blood pressure <90 mmHg, +1 for pre-operative hemoglobin level <5.6 mmol/L. A score of 0 or 1 corresponded with a predicted death rate of 30%, a score of 2 with a predicted death rate of 50%, and a score of 3 with a predicted death rate of 80%.

Hardman index

The Hardman index was calculated with the formula: +1 for age >76 years, +1 for in hospital loss of consciousness, +1 for a pre-operative serum creatinine >190 $\mu\text{mol/L}$, +1 for pre-operative serum hemoglobin level <5.6 mmol/L, +1 for electrocardiographic (ECG) ischemia (defined as ST segment depression greater than 1 millimeter or an associated T wave change determined by a senior cardiologist [RJGP]). A score of 3 or more corresponded with a predicted death rate of 100%.

Data collection and statistical analysis

Data were collected from the medical records by the first and second authors. Data entry was done using Microsoft Access 2003 (Microsoft Corporation, Redmond, WA, USA) using field limits, univariate and multivariate checks. A valid way of coping with missing values is by imputation.¹⁸ Missing data were imputed for the variables blood pressure, hemoglobin, creatinine, cardiac comorbidity, cerebrovascular comorbidity, resuscitation, loss of consciousness, and GCS. Multiple imputation was done creating ten datasets. Age, sex, renal and pulmonary comorbidity, death, and the above mentioned imputed variables were used as predictors in the imputation model. Baseline characteristics and prediction model scores are reported in both the original dataset and in the imputed datasets (Tables 1 and 2).

The statistical analysis and the imputation procedure were done using IBM SPSS Statistics 19.0 (SPSS Inc.,

Table 1. Baseline pre-operative characteristics.

Pre-operative variable	Original data		Imputed data
	Median (IQR)	Missing data %	Median (IQR)
Age in years	76 (69–80)	0	—
Lowest in hospital SBP in mmHg	90 (70–125)	11 (48/449)	90 (70–125)
Hemoglobin at ER in mmol/L	7.0 (5.9–8.0)	1 (5/449)	7 (5.9–8.0)
Creatinine at ER in μ mol/L	106 (86–133)	3 (14/449)	107 (87–134)
	% (Number)	Missing data %	% (Number)
Male:female	80:20 (360:89)	0	—
Cardiac comorbidity	42 (184/435)	3 (14/449)	43 (191/449)
Cerebrovascular comorbidity	15 (67/433)	4 (16/449)	15 (69/449)
In hospital cardiopulmonary resuscitation	11 (46/429)	4 (20/449)	12 (52/449)
In hospital loss of consciousness	21 (81/388)	14 (61/449)	21 (96/449)
Best recorded Glasgow Coma Scale <15	17 (63/372)	17 (77/449)	18 (82/449)
ECG ischemia	21 (40/192)	57 (257/449)	—
EVAR:OR	15:85 (69:380)	0	—

RAAA = ruptured abdominal aortic aneurysm; SD = standard deviation; SBP = systolic blood pressure; IQR = interquartile range; ER = emergency room; EVAR = endovascular aneurysm repair; OR = open repair, ECG = electrocardiogram.

Armonk, NY, USA) and R (The R Foundation for Statistical Computing, Boston, MA, USA). Continuous data were described by the mean with corresponding standard deviation (SD) for data normally distributed, and by the median with corresponding interquartile range (IQR) for data with skewed distribution. The statistical analysis comprised five steps. First, the accuracy of the predictions was determined with regard to overall performance and discrimination.¹⁹

Table 2. The outcomes and distribution in the original dataset and the imputed dataset of the GAS, the updated GAS, the Vancouver score and the ERAS.

Prediction model	Original dataset		Imputed dataset
		Missing data % (number)	
Updated GAS, mean (\pm SD)	93 (\pm 15)	14 (64/449)	93 (\pm 15)
Vancouver, median (IQR)	3.10 (2.66–3.72)	14 (64/449)	3.10 (2.66–3.74)
	% (number)	Missing data % (number)	% (number)
ERAS score \leq 1	79 (274/349)	22 (100/449)	77 (345/449)
2	17 (61/349)		19 (85/449)
3	4 (14/349)		4 (19/449)

SD = standard deviation; GAS = Glasgow Aneurysm Score; ERAS = Edinburgh Ruptured Aneurysm Score.

Overall performance represents the squared difference between the predicted outcome and actual outcome, and was assessed using the Brier Score. The Brier Score should be as close to 0 as possible and the threshold for a non-informative model was calculated to be at 0.23. Discrimination is the ability of a model to distinguish between dying and surviving patients and was assessed using the area under the receiver operating characteristics curve (AUC). An AUC >0.70 was considered sufficiently accurate. Second, in the models with an AUC >0.70, the calibration of the predictions was determined. Calibration refers to the agreement between the predicted and observed death rate. Calibration was assessed by dividing all patients into five comparable quintiles: 0–20%, >20–40%, >40–60%, >60–80% and >80–100%. Because patients with equal predictions were categorized in the same quintile, the sizes of the quintiles differed slightly between the several prediction models. Subsequently, the mean predicted death rate per quintile was plotted with the corresponding mean observed death rate. In addition, the Hosmer-Lemeshow (HL) chi-square test was done to compare the observed and predicted death rates. In the HL test $p < .05$ reflects a significant difference between the predicted and observed death rate which is a poor calibration. Third, the models with an AUC >0.70 and an HL test $p < .05$ were recalibrated using the ‘calibration intercept method’.²⁰ Fourth, a subgroup analysis was done in patients with a predicted death rate of \geq 95% to assess the accuracy in high risk patients in whom withholding intervention might be considered. Fifth, the accuracy of the prediction models was determined in patients treated with EVAR. In this second subgroup analysis, the threshold of the Brier score for a non-informative model was calculated to be at 0.20.

RESULTS

Of 539 patients with a RAAA in the greater Amsterdam region, 66 did not have an intervention and 24 had to be excluded because of other reasons (Fig. 1). The reasons to refrain from intervention were predominantly shock or resuscitation with an expected low chance of survival ($n = 20$), patient or patient’s family decision ($n = 17$), or unknown ($n = 17$). The updated GAS, the Vancouver score, and the ERAS of these patients without intervention is shown in the online supplement (Table S3, online supplement). The baseline characteristics of the 449 patients included in the analysis are shown in Table 1. Sixty-nine patients were treated with EVAR and 380 patients were treated with OR. The death rate was 36% (160/449, 95% CI 31–40).

Updated GAS

The mean updated GAS score was 93 (standard deviation (SD) \pm 15) (Table 2). The Brier Score was 0.21 and the AUC was 0.71 (95% CI 0.66–0.76). The calibration plot showed an overestimation of the death rate in patients with a predicted death rate >50% (HL test $p = .01$) (Fig. 2). In the quintile of patients with a mean predicted death rate of

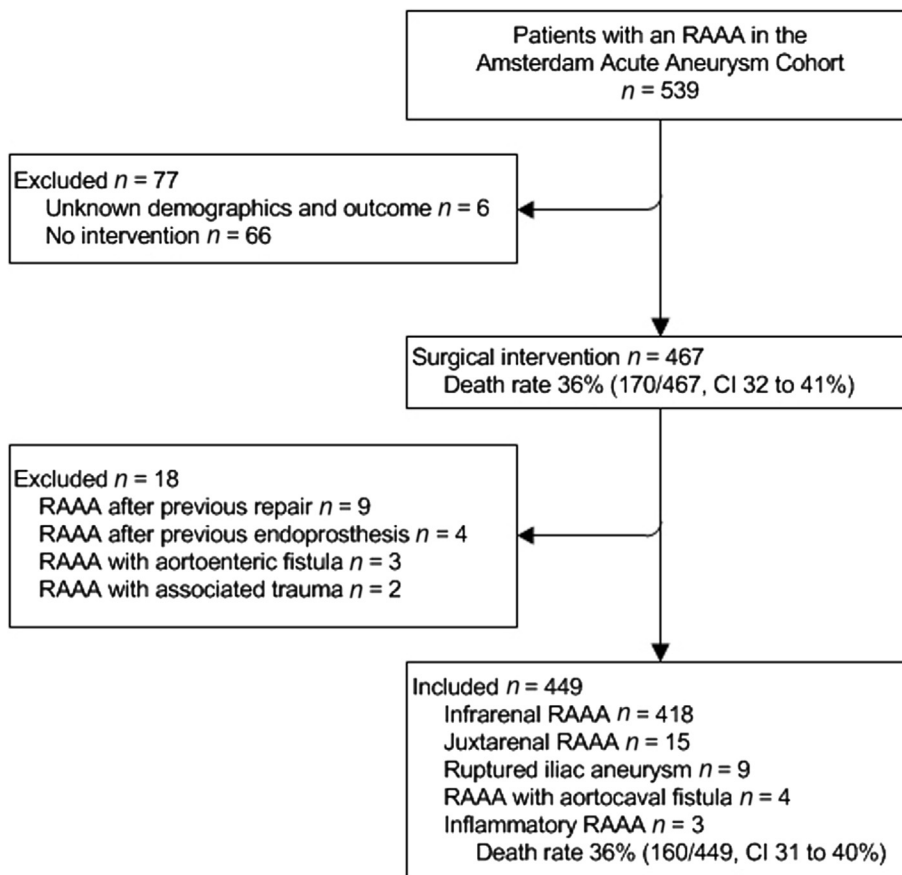


Figure 1. Flowchart of inclusion and exclusion in the analysis. RAAA = ruptured abdominal aortic aneurysm; CI = 95% confidence interval.

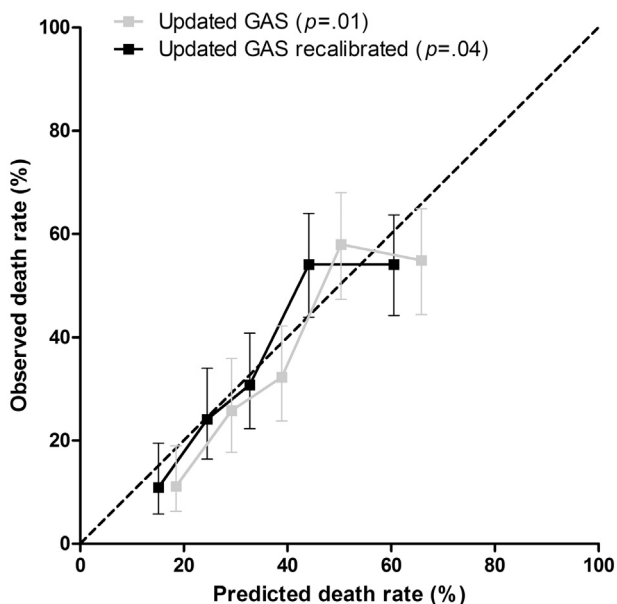


Figure 2. The calibration plots of the updated GAS before and after recalibration. The predicted death rate is plotted with the corresponding death rate and surrounding 95% confidence interval. The interrupted black line indicates ideal calibration. The *p* corresponds to the Hosmer-Lemeshow (HL) chi-square test. GAS = Glasgow Aneurysm Score.

66%, the observed death rate was 55% (95% CI 44–65). After recalibration, the plot slightly improved, although there was still a statistically significant deviation between the predicted and observed risks (HL test *p* = .04, Fig. 3). In the quintile of patients with a mean predicted death rate of 60%, the observed death rate was 54% (95% CI 44–64) after recalibration.

Subgroup analysis to assess the accuracy in high risk patients showed that no patients had a predicted death rate $\geq 95\%$. In patients treated with EVAR, the Brier Score was 0.17, the AUC was 0.78 (95% CI 0.66–0.90), and in the HL test *p* was .18.

Vancouver score

The median Vancouver score was 3.10 (interquartile range 2.66–3.72) (Table 2). The Brier Score was 0.22 and the AUC was 0.72 (95% CI 0.67–0.77). With regard to calibration, in the quintile of patients with a mean predicted death rate of 33%, the observed death rate was 21% (95% CI 14–31%), and in the quintile of patients with a mean predicted death

$$1 - \frac{1}{1 + (\text{EXP} (-5.486 + 0.052 * \text{updated GAS score}))}$$

Figure 3. The formula to calculate the predicted death rate using the updated GAS after recalibration. GAS = Glasgow Aneurysm Score.

rate of 89%, the observed death rate was 62% (95% CI 52–71%). Hence, the calibration plot showed an overestimation of death (HL test $p < .01$) (Fig. 3). After recalibration, this overestimation decreased minimally (HL test $p < .01$, Fig. 4). In high risk patients there was a significant overestimation of the observed risk by the recalibrated model. In the quintile of patients with a mean predicted death rate of 82%, the observed death rate was 62% (95% CI 52–71%).

Subgroup analysis to assess the accuracy in high risk patients showed that of 21 patients with a predicted death rate $\geq 95\%$, 18 patients died. In patients treated with EVAR, the Brier Score was 0.19, the AUC was 0.77 (95% CI 0.64–0.90), and the HL test p was .03.

ERAS

The distribution of patients per ERAS outcome is shown in Table 2. The Brier Score was 0.23 and the AUC was 0.58 (95% CI 0.52–0.64). Calibration was not assessed because of an AUC < 0.70 . Subgroup analysis to assess the accuracy in high risk patients showed that no patients had a predicted death rate $\geq 95\%$. In patients treated with EVAR, the Brier Score was 0.20 and the AUC was 0.55 (95% CI 0.50–0.60).

Hardman index

In 57% (257/449), the pre-operative ECGs were missing. Therefore, the Hardman index was excluded from the analysis.

DISCUSSION

The present study shows that following intervention for a RAAA, the updated GAS predicted death most accurately for

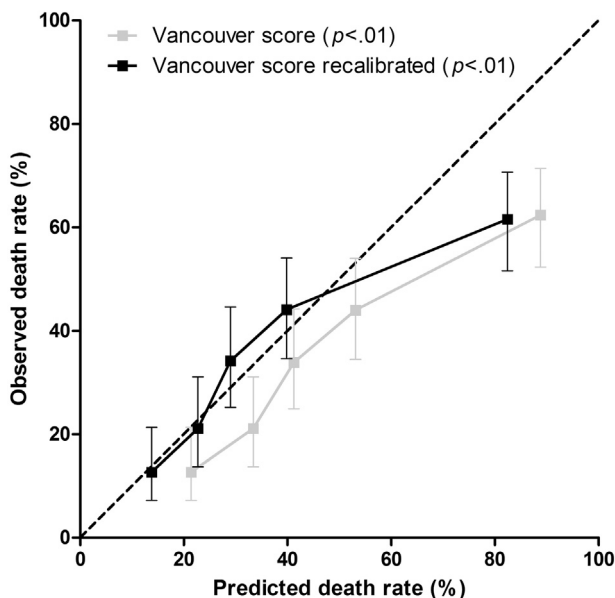


Figure 4. The calibration plots of the Vancouver score before and after recalibration. The predicted death rate is plotted with the corresponding death rate and surrounding 95% confidence interval. The interrupted black line indicates ideal calibration. The p corresponds to the Hosmer-Lemeshow (HL) chi-square test.

both discrimination and calibration. The present study expands on previous studies externally validating the updated GAS, the Vancouver score, and the ERAS. First, a cut-off value of patients was set in whom withholding intervention might be considered. In this way, it was aimed to assess the additional value of the prediction models in clinical practice. Second, the number of patients included ($n = 449$) was higher than the previous largest study ($n = 201$).¹³ Finally, the updated GAS and Vancouver score were recalibrated to improve accuracy in the era of EVAR. Because of the large number of missing ECGs, no definite conclusions could be drawn for the Hardman index.

Decision making

The decision to withhold intervention in patients with a RAAA can be very difficult. Only extremely reliable models can be useful in clinical decision making and in identifying patients in whom withholding intervention might be considered. For this purpose, a cut-off value for the predicted death rate was set at $\geq 95\%$. If the death rate was to be predicted accurately at 95%, the number needed to treat (NNT) would be 20. This cut-off value is arbitrary and could also have been 90% (NNT of 10) or 99% (NNT of 100). Different cut-off values can be used depending on the clinical situation. None of the prediction models met the criterion of identifying patients in whom to withhold intervention. This disappointing conclusion is in agreement with previous validation studies.^{21–23} Currently, the prediction models have insufficient accuracy to evaluate the chances of successful intervention and future studies should focus on improvement towards this aim. The usefulness of current prediction models lies in case mix comparisons between hospitals, and in a tailored prognosis for patients and relatives.

Updated GAS

The updated GAS predicted death most accurately for both discrimination and calibration. Several other studies have validated the GAS.^{13,22–26} In the only previous study including patients treated with EVAR, the AUC was 0.70 (95% CI 0.62–0.77).¹³ The calibration of the updated GAS was not assessed in this previous study. The strength of the previous validation (201 patients included, multicenter, prospective, including EVAR and OR)¹³ confirms the conclusion that the updated GAS is the most accurate in predicting death after intervention for a RAAA. If clinicians consider their patients to be comparable with those included in the present study, the model as shown in Fig. 3 can be used to predict the risk of dying after intervention.

Vancouver score

The Vancouver score discriminated sufficiently accurately, but even after recalibration its predictions still overestimated the death rate considerably. These results are in accordance with previous disappointing results on discrimination,²⁷ but in conflict with previous fairly accurate results on calibration.²³ Therefore, the accuracy of the Vancouver

score has not yet been proven and the present authors prefer the updated GAS.

ERAS

The prediction of death by the ERAS was insufficiently accurate. These results are in conflict with one validation study with sufficiently accurate discrimination,²⁵ but in accordance with another validation study with insufficiently accurate discrimination.²³ Concerning calibration, one previous validation reported an observed death rate of 50% in patients with a predicted death rate of 80% (estimated from figure).²³ Because results regarding the ERAS are conflicting, the present authors question its precision.

Hardman index

The Hardman index was excluded from the analysis because data on one variable, presence of electrocardiographic (ECG) ischemia, were missing in 57% of patients. The missing ECGs are a drawback of the present study, and also of the scoring system. Most surgical trainees and vascular surgeons do not know how to interpret an ECG with sufficient precision to use it as a variable in a prediction model. From a cardiac perspective, acute ischemia defined by ST segment depression greater than 1 millimeter or an associated T wave change is an oversimplification of the great diagnostic value of an ECG. Based on these considerations, the present authors are convinced that in their clinical practice the contribution of a pre-operative ECG is limited and, consequently, that the Hardman index is not a useful prediction model.

EVAR

The predictions by the updated GAS and the Vancouver score appeared slightly more accurate in patients treated with EVAR compared with all patients. The accuracy of the ERAS appeared similar for patients treated with EVAR compared with all patients. Recently published randomized clinical trials reported comparable death rates after EVAR and OR.^{28,29} This indicates that the risk profiles are based on the same pre-operative variables and that the accuracy of the prediction models probably does not differ substantially between both interventions. Also, because of a low event rate (19/69) and wider confidence intervals in patients treated with EVAR, the present authors are reluctant to draw definite conclusions regarding the accuracy in patients treated with EVAR separately.

Limitations

An important limitation of the present study was the exclusion of the Hardman index. Two other prediction models have been described in the literature, the RAAA-physiological and operative severity score for enumeration of mortality and morbidity (RAAA-POSSUM)³⁰ and the Vascular Study Group of New England (VSGNE) RAAA score.²³ The RAAA-POSSUM was not included in the present study because of its complexity including chest X-ray

examination, and hence low clinical applicability. The VSGNE RAAA score was not included in the present study because of the use of an intra-operative variable, thereby making predictions prior to the intervention impossible. In short, the authors consider the variables of the three excluded prediction models unsuitable for fast evaluation of the chances of successful intervention.

Another limitation of the present study was retrospective data collection. Probably, the variables 'best recorded in hospital Glasgow coma scale' for the ERAS and 'loss of consciousness' for the Vancouver score contain imprecise data. Another limitation was the amount of missing data (Tables 1 and 2). This is a consequence of the acute character of the disease. The authors coped with this problem by multiple imputation. Death was included as a predictor in the imputation model to correct for the bias that the most missing data were in patients who died.

Conclusions

The updated GAS most accurately predicted death after intervention for a RAAA. However, the updated GAS did not identify patients with a predicted death rate $\geq 95\%$, and therefore cannot reliably support the decision to withhold intervention.

ACKNOWLEDGEMENT

The authors wish to thank Susan van Dieren for statistical support.

CONFLICT OF INTEREST

None.

FUNDING

Partial funding was provided by the Dutch Heart Foundation (project: 2002B197) and the AMC Foundation. The sponsor had no involvement in the study design, in the collection, analysis and interpretation of data; in the writing of the manuscript; and in the decision to submit the manuscript for publication.

APPENDIX A. SUPPLEMENTARY MATERIAL

Supplementary material related to this article can be found online at <http://dx.doi.org/10.1016/j.ejvs.2014.10.012>.

REFERENCES

- 1 Reimerink JJ, van der Laan MJ, Koelemay MJ, Balm R, Legemate DA. Systematic review and meta-analysis of population-based mortality from ruptured abdominal aortic aneurysm. *Br J Surg* 2013;**100**(11):1405–13.
- 2 Mani K, Lees T, Beiles B, Jensen LP, Venermo M, Simo G, et al. Treatment of abdominal aortic aneurysm in nine countries 2005–2009: a vasconet report. *Eur J Vasc Endovasc Surg* 2011;**42**(5):598–607.
- 3 Hoornweg LL, Storm-Versloot MN, Ubbink DT, Koelemay MJ, Legemate DA, Balm R. Meta analysis on mortality of ruptured abdominal aortic aneurysms. *Eur J Vasc Endovasc Surg* 2008;**35**(5):558–70.

- 4 Rayt HS, Sutton AJ, London NJ, Sayers RD, Bown MJ. A systematic review and meta-analysis of endovascular repair (EVAR) for ruptured abdominal aortic aneurysm. *Eur J Vasc Endovasc Surg* 2008;**36**(5):536–44.
- 5 Hewin DF, Campbell WB. Ruptured aortic aneurysm: the decision not to operate. *Ann R Coll Surg Engl* 1998;**80**(3):221–5.
- 6 Hardman DT, Fisher CM, Patel MI, Neale M, Chambers J, Lane R, et al. Ruptured abdominal aortic aneurysms: who should be offered surgery? *J Vasc Surg* 1996;**23**(1):123–9.
- 7 Biancari F, Venermo M. Open repair of ruptured abdominal aortic aneurysm in patients aged 80 years and older. *Br J Surg* 2011;**98**(12):1713–8.
- 8 Prance SE, Wilson YG, Cosgrove CM, Walker AJ, Wilkins DC, Ashley S. Ruptured abdominal aortic aneurysms: selecting patients for surgery. *Eur J Vasc Endovasc Surg* 1999;**17**(2):129–32.
- 9 Greeven AP, Bouwman LH, Smeets HJ, van Baalen JM, Hamming JF. Outcome of patients with ruptured abdominal aortic aneurysm after cardiopulmonary resuscitation. *Acta Chir Belg* 2011;**111**(2):78–82.
- 10 Samy AK, Murray G, MacBain G. Glasgow aneurysm score. *Cardiovasc Surg* 1994;**2**(1):41–4.
- 11 Chen JC, Hildebrand HD, Salvian AJ, Taylor DC, Strandberg S, Myckatyn TM, et al. Predictors of death in nonruptured and ruptured abdominal aortic aneurysms. *J Vasc Surg* 1996;**24**(4):614–20.
- 12 Tambyraja A, Murie J, Chalmers R. Predictors of outcome after abdominal aortic aneurysm rupture: Edinburgh Ruptured Aneurysm Score. *World J Surg* 2007;**31**(11):2243–7.
- 13 Visser JJ, Williams M, Kievit J, Bosch JL. Prediction of 30 day mortality after endovascular repair or open surgery in patients with ruptured abdominal aortic aneurysms. *J Vasc Surg* 2009;**49**(5):1093–9.
- 14 van Beek SC, Reimerink JJ, Vahl A, Wisselink W, van Geloven N, Legemate DA, et al. Effect of regional cooperation on outcomes from ruptured abdominal aortic aneurysm. *Br J Surg* 2014;**101**(7):794–801.
- 15 van Beek SC, Reimerink JJ, Vahl AC, Wisselink W, Reekers JA, Legemate DA, et al. Outcomes after open repair for ruptured abdominal aortic aneurysms in patients with friendly versus hostile aortoiliac anatomy. *Eur J Vasc Endovasc Surg* 2014;**47**(4):380–7.
- 16 Statistics Netherlands. <http://www.cbs.nl/en-GB/menu/home/default.htm>; 2010.
- 17 Von EE, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007;**4**(10):e296.
- 18 Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;**59**(10):1087–91.
- 19 Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;**21**(1):128–38.
- 20 Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 2008;**61**(1):76–86.
- 21 Gatt M, Goldsmith P, Martinez M, Barandiaran J, Grover K, El-Barghouti N, et al. Do scoring systems help in predicting survival following ruptured abdominal aortic aneurysm surgery? *Ann R Coll Surg Engl* 2009;**91**(2):123–7.
- 22 Kurc E, Sanioglu S, Ozgen A, Aka SA, Yekeler I. Preoperative risk factors for in hospital mortality and validity of the Glasgow aneurysm score and Hardman index in patients with ruptured abdominal aortic aneurysm. *Vascular* 2012;**20**(3):150–5.
- 23 Robinson WP, Schanzer A, Li Y, Goodney PP, Nolan BW, Eslami MH, et al. Derivation and validation of a practical risk score for prediction of mortality after open repair of ruptured abdominal aortic aneurysms in a U.S. regional cohort and comparison to existing scoring systems. *J Vasc Surg* 2013;**57**(2):354–61.
- 24 Kapma M, Kahmann O, Van Stijn I, Zeebregts CJ, Vahl A. Evaluation of risk prediction models, V-POSSUM and GAS, in patients with acute abdominal aortic rupture treated with EVAR or an open procedure. *J Cardiovasc Surg (Torino)* 2014 Jan 16 [Epub ahead of print].
- 25 Tambyraja AL, Lee AJ, Murie JA, Chalmers RT. Prognostic scoring in ruptured abdominal aortic aneurysm: a prospective evaluation. *J Vasc Surg* 2008;**47**(2):282–6.
- 26 Korhonen SJ, Ylonen K, Biancari F, Heikkinen M, Salenius JP, Lepantalo M. Glasgow Aneurysm Score as a predictor of immediate outcome after surgery for ruptured abdominal aortic aneurysm. *Br J Surg* 2004;**91**(11):1449–52.
- 27 Leo E, Biancari F, Nesi F, Pogany G, Bartolucci R, De PF, et al. Risk-scoring methods in predicting the immediate outcome after emergency open repair of ruptured abdominal aortic aneurysm. *Am J Surg* 2006;**192**(1):19–23.
- 28 Reimerink JJ, Hoornweg LL, Vahl AC, Wisselink W, van den Broek TA, Legemate DA, et al. Endovascular repair versus open repair of ruptured abdominal aortic aneurysms: a multicenter randomized controlled trial. *Ann Surg* 2013;**258**(2):248–56.
- 29 Powell JT, Sweeting MJ, Thompson MM, Ashleigh R, Bell R, Gomes M, et al. Endovascular or open repair strategy for ruptured abdominal aortic aneurysm: 30 day outcomes from IMPROVE randomised trial. *BMJ* 2014;**348**:f7661.
- 30 Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. *Br J Surg* 2001;**88**(7):958–63.