



Published in final edited form as:

J Vasc Surg. 2018 January ; 67(1): 151–156.e3. doi:10.1016/j.jvs.2017.05.104.

External validation of a 5-year survival prediction model after elective abdominal aortic aneurysm repair

Randall R. DeMartino, MD, MS^a, Ying Huang, MD^a, Jay Mandrekar, PhD^b, Philip P. Goodney, MD, MS^c, Gustavo S. Oderich, MD^a, Manju Kalra, MBBS^a, Thomas C. Bower, MD^a, Jack L. Cronenwett, MD^c, and Peter Gloviczki, MD^a

^aDivision of Vascular and Endovascular Surgery, Mayo Clinic, Rochester, Minn

^bDivision of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minn

^cDivision of Vascular Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, NH

Abstract

Objective—The benefit of prophylactic repair of abdominal aortic aneurysms (AAAs) is based on the risk of rupture exceeding the risk of death from other comorbidities. The purpose of this study was to validate a 5-year survival prediction model for patients undergoing elective repair of asymptomatic AAA <6.5 cm to assist in optimal selection of patients.

Methods—All patients undergoing elective repair for asymptomatic AAA <6.5 cm (open or endovascular) from 2002 to 2011 were identified from a single institutional database (validation group). We assessed the ability of a prior published Vascular Study Group of New England (VSGNE) model (derivation group) to predict survival in our cohort. The model was assessed for discrimination (concordance index), calibration (calibration slope and calibration in the large), and goodness of fit (score test).

Results—The VSGNE derivation group consisted of 2367 patients (70% endovascular). Major factors associated with survival in the derivation group were age, coronary disease, chronic obstructive pulmonary disease, renal function, and antiplatelet and statin medication use. Our validation group consisted of 1038 patients (59% endovascular). The validation group was slightly

Correspondence: Randall R. DeMartino, MD, MS, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (demartino.randall@mayo.edu).

Author conflict of interest: none.

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

Presented at the 2016 Vascular Annual Meeting of the Society for Vascular Surgery, National Harbor, Md, June 8–11, 2016.

Additional material for this article may be found online at www.jvascsurg.org.

AUTHOR CONTRIBUTIONS

Conception and design: RD, YH, JM, PPG, PG

Analysis and interpretation: JM, PPG, JC, PG

Data collection: RD, YH, GO, MK, JC, PG

Writing the article: RD, JM

Critical revision of the article: RD, YH, JM, PPG, GO, MK, JC, PG

Final approval of the article: RD, YH, JM, PPG, GO, MK, JC, PG

Statistical analysis: RD, JM

Obtained funding: Not applicable

Overall responsibility: RD

older (74 vs 72 years; $P < .01$) and had a higher proportion of men (76% vs 68%; $P < .01$). In addition, the derivation group had higher rates of advanced cardiac disease, chronic obstructive pulmonary disease, and baseline creatinine concentration (1.2 vs 1.1 mg/dL; $P < .01$). Despite slight differences in preoperative patient factors, 5-year survival was similar between validation and derivation groups (75% vs 77%; $P = .33$). The concordance index of the validation group was identical between derivation and validation groups at 0.659 (95% confidence interval, 0.63–0.69). Our validation calibration in the large value was 1.02 ($P = .62$, closer to 1 indicating better calibration), calibration slope of 0.84 (95% confidence interval, 0.71–0.97), and score test of $P = .57$ ($> .05$ indicating goodness of fit).

Conclusions—Across different populations of patients, assessment of age and level of cardiac, pulmonary, and renal disease can accurately predict 5-year survival in patients with AAA < 6.5 cm undergoing repair. This risk prediction model is a valid method to assess mortality risk in determining potential overall survival benefit from elective AAA repair.

The benefit of prophylactic abdominal aortic aneurysm (AAA) repair requires the patient to live long enough for the reduction in annual rupture risk to overcome the up-front risk of surgery. Rupture risk increases rapidly above AAA diameter of 5.5 cm, which is generally the accepted threshold for elective repair (in men). However, the Endovascular Aneurysm Repair trial 2 (EVAR 2) confirmed that patients with severe comorbidities that shorten their life expectancy gain little benefit in overall survival from prophylactic AAA repair.¹ Thus, proper decision-making for prophylactic AAA repair requires an assessment of not only rupture and operative risk but also long-term survival. Unfortunately, there are few validated risk scores against which to gauge the potential benefit of repair to assist clinicians in estimating the long-term survival of patients after AAA repair.² A previously developed risk prediction model for survival after AAA repair has been created using data from the Vascular Study Group of New England (VSGNE).³ The purpose of our current study was to externally validate this risk prediction model in an independent group of patients.

METHODS

This study is a retrospective review of the clinical data of all consecutive patients undergoing infrarenal AAA repair by either open aortic repair (OAR, transperitoneal or retroperitoneal) or endovascular aortic repair (EVAR) for AAA between 2003 and 2011 at Mayo Clinic, Rochester, Minnesota. This time period was selected to correspond to the time period of the original VSGNE derivation study. In addition, it permitted all patients to have 5-year follow-up. We restricted our analysis to patients with a diameter of AAA < 6.5 cm. This was done to match the prior VSGNE inclusion criterion that was established to select a group of patients whose rupture risk may be low enough to allow assessment of other factors in medical decision-making. This resulted in 1038 patients for analysis. This is a subset of patients whose outcomes have been previously reported.⁴ All patients had infrarenal aortic clamping (thus potentially eligible for either OAR or EVAR), and emergent/urgent cases were excluded. Further detail of this cohort of patients has been previously reported.⁴ Patient demographics, comorbidities (defined by Society for Vascular Surgery/American Association for Vascular Surgery [SVS/AAVS] comorbidity scores),⁵ and procedural

information were recorded. This study was approved by the Mayo Clinic Institutional Review Board and was deemed exempt from the need for patient consent.

The prior VSGNE risk prediction model for 5-year survival after AAA repair was based on consecutive repairs done within the VSGNE and contained several significant factors. These were used to create a risk score to identify low-, intermediate-, and high-risk patients for mortality within 5 years after AAA repair. These included (with respective points assigned) age 75 to 79 years (2), age ≥ 80 years (3), prior myocardial infarction (1), stable angina (1), unstable angina or recent myocardial infarction (4), oxygen-dependent chronic obstructive pulmonary disease (COPD, 3), renal dysfunction as measured by estimated glomerular filtration rate <30 mL/min/1.73 m² (3), and not taking a statin (1) or aspirin (1) at discharge. Based on these factors and risk score, patients were at low risk (≤ 3 points), intermediate risk (4–5 points), or high risk (≥ 6 points) for 5-year mortality.

Similar factors were obtained in our current study for patients from Mayo Clinic. Age, sex, and SVS/AAVS comorbidity scores were originally abstracted for our derivation data set.⁶ Aspirin and statin use were abstracted from preoperative and discharge medication lists. Discharge medication was used to assign medication use. The most recent preoperative serum creatinine value within 3 months of surgical date was abstracted from the medical record and used to determine renal function (estimated glomerular filtration rate) by the Modification of Diet in Renal Disease equation.⁷ However, patient data regarding coronary artery disease (CAD) and COPD from our derivation group (VSGNE) were recorded into categories adapted ultimately by the Vascular Quality Initiative (VQI; Supplementary Table, online only). The VQI definitions are not identical to the SVS/AAVS comorbidity grading for CAD and COPD (Supplementary Table, online only). Thus, we assigned our validation group SVS/AAVS scored to the closest comorbidity grade in VQI for analysis (Supplementary Table, online only). Vital status and death date information were queried using an institutionally approved fee-based Internet research location service (Accurint, accurint.com) queried on January 12, 2016.

Statistical analysis

Data on patient demographics were summarized as means with standard deviations and frequencies for continuous and categorical factors. These were compared between derivation and validation with the Student *t*-test and χ^2 test, respectively, with Fisher exact correction, as needed. Survival was analyzed as time to event data with life tables and displayed as Kaplan-Meier curves. Survival between groups was compared with the log-rank test. Comparison of discrimination and calibration of the derivation data with the prior risk model was done in several ways.⁸ First, discrimination was assessed with the concordance index. Then calibration was assessed through several tests. Calibration in the large is a basic measure of agreement between the observed and predicted risks. In the case of time to event analysis, assessment of calibration in the large is based on an estimate of a ratio of observed events in the new data set to the number predicted by the risk model. A value of the ratio closer to 1 indicates better calibration. The calibration slope is calculated by regressing the observed outcome on the predicted probabilities. The estimated slope obtained from the regression model provides a measure of effect size and a confidence interval (CI) in addition

to a *P* value. For these reasons, the calibration slope is more informative and is the preferred method for assessing calibration. A value closer to 1 indicates better calibration. Finally, the score test is a goodness-of-fit test, similar to the Hosmer-Lemeshow (where *P* > .05 indicates goodness of fit). Otherwise, a *P* value of < .05 was considered significant. Statistical analysis was performed with Stata SE (version 14.1; StataCorp LP, College Station, Tex) and SAS (version 9.4; SAS Institute Cary, NC).

RESULTS

Comparison of derivation and validation cohorts and 5-year survival

As previously published, the VSGNE derivation cohort consisted of 2367 patients undergoing elective AAA repair by OAR (*n* = 714 [30%]) and EVAR (*n* = 1653 [70%]) within the VSGNE.³ The validation cohort consisted of 1038 patients undergoing elective AAA repair during the same time interval. Compared with the derivation cohort, a higher proportion of patients in the validation group underwent OAR (*n* = 430 [41%]) vs EVAR (*n* = 608 [59%]; *P* < .001 for comparison of procedures between derivation and validation). Overall, there were slight differences between the derivation and validation groups. The validation cohort was slightly older, was more male predominant, and had higher severity of CAD and COPD. It had more patients who were ever smokers, were more likely to be taking a discharge aspirin or statin, and had a slightly smaller aneurysm size. Length of stay and in-hospital mortality were similar between derivation and validation groups (Table).

Despite the differences in the patients' characteristics and choice of operation between the derivation and validation groups, survival at 5 years was similar. Overall survival at 5 years was 77% (95% CI, 74.7%–79.5%) for the derivation group and 75% (95% CI, 72.3%–77.7%; log-rank, *P* = .33) for the validation group (Fig 1). The previously developed VSGNE risk score was then applied to the validation cohort (low, intermediate, and high risk for 5-year mortality) to separate patients by mortality risk. Compared with the 5-year survival risk groups of the derivation set, survival was nearly identical in the risk groups of the validation patient set (Fig 2).

Discrimination and validation of survival model

Overall, the concordance index for the survival model on the derivation data set was 0.659 (95% CI, 0.626–0.69) and was nearly identical to the concordance index of the model on the validation group (0.659; 95% CI, 0.63–0.69), demonstrating similar discriminatory ability of the model. The calibration in the large assessment was 1.02 (95% CI, 0.94–1.12; *P* = .62). This measure of agreement between the observed and predicted mortality risks indicates excellent calibration, whereas the nonsignificant *P* value indicated that the ratio is not different from 1 (ie, proposed model calibrates well). The calibration slope was 0.84 (95% CI, 0.71–0.97; *P* < .0001). This regression of the observed outcome on the predicted probabilities indicates excellent calibration, and the significant *P* value suggests that slope is different from 0. Finally, the *P* value for the score test (goodness-of-fit test) was .57, which indicates that the expected and observed numbers of events were similar in the validation group as predicted by the derivation model.

DISCUSSION

Proper selection of patients is essential for ensuring that patients receive a benefit from prophylactic AAA repair. The decision to repair a patient’s AAA requires an assessment of the aneurysm rupture risk vs the operative risk plus the patient’s life expectancy. In this study, we have demonstrated that the VSGNE risk prediction model and risk score can accurately predict 5-year mortality risk for patients undergoing open and endovascular AAA repair. Information regarding age of the patient, severity of CAD, COPD, renal function, and medication use can accurately stratify patients into low, intermediate, and high risk for 5-year mortality. This works equally well for patients who are planned to undergo both open and endovascular repair of their AAA. Our validation group of patients was noted to have differences in baseline characteristics. Despite this, the survival model was able to accurately discriminate survival with good calibration, further strengthening our validation of this model. Now validated, the VSGNE survival model can serve as a bedside tool to educate physicians and to counsel patients on risks and benefits of surgery.

Our model is unique from other nonvalidated survival models in several ways. In the model developed by Mousa et al, they studied patients undergoing EVAR who were enrolled in the VQI database from 2003 to 2014.⁹ They used our prior VSGNE risk model and adjusted it for application to EVAR patients in the VQI data set by adding body mass index, AAA diameter, and high surgical risk (a data element tracked only in the EVAR module of the VQI). We intentionally did not include AAA diameter in our model and included only those with AAA <6.5 cm as this size offers a low enough rupture risk that one may justly consider patient factors before proceeding with repair. Also, we included both OAR and EVAR, given that 5-year survival has been shown to be similar between repair methods and validation with both groups of patients is the most useful for clinicians. Finally, given that their data set spans 2003 to 2014,⁹ many of their patients with long-term follow-up (5–10 years) are likely to be from the VSGNE cohort that was used in our original survival study (VGSNE was initiated in 2003). The VQI and other regions were not generally formalized until 2011.¹⁰ Thus, our study represents a truly new set of patients for which to validate our risk prediction model.

In a review of 24 years of AAA repair, Kashram et al defined factors associated with long-term survival.¹¹ Although our study is limited to overall mortality, they delineated that cardiac, respiratory, and oncologic causes remain the leading causes of death for patients after AAA repair (accounting for two-thirds of late deaths). They identified age, female sex, American Society of Anesthesiologists class, COPD, peripheral arterial disease, renal impairment, and heart failure as factors associated with late death.¹¹ This is similar to a review of 247 patients treated with EVAR reported by Lim et al, in which they identified advanced age, renal disease, and oxygen-dependent COPD as predictors of overall mortality. In addition, advanced heart failure (ejection fraction <40%) and current cancer treatment were major predictors.¹² A review by Sevilla et al also demonstrated how current cancer treatment negatively affects survival after EVAR.¹³ These reports are similar but contain slightly different factors from our findings. In contrast to our work, Kashram et al incorporated intra-operative and postoperative factors in their model for late death, whereas we chose risk factors that are available preoperatively to allow patient education and

counseling. The study by Lim included only EVAR cases with a large number of high-risk patients (30%), which may have led to their slightly different findings.¹² Finally, our findings apply to both open and endovascular repair.

For our validation study, we chose to use several validation measures. As opposed to a logistic model where the outcome is dichotomous, a Cox proportional hazards model includes the element of time to the event, and distinct methods are needed for these models. We performed several measures to assess our model's strength as previously proposed.⁸ The score test, similar to the Hosmer-Lemeshow goodness-of-fit test, compares the observed to expected number of events within k groups based on their predicted probability of risk from the model (usually $k = 10$ groups). This provides a P value to identify whether the observed to expected number of events in each group is different. However, it does not provide an estimate with CIs, which can improve interpretation. Both the calibration in the large and calibration slope provide an estimate and CIs. Although our calibration slope was 0.84 (95% CI, 0.71–0.97, a value of 1 would be perfect calibration), because of the number of factors in our model, this remains a significant slope and we think it is sufficient for calibration of risk. In addition, our other calibration methods demonstrate excellent calibration to strengthen the validation of our model. As opposed to the validation study by Carlisle et al, who internally validated their survival model after AAA repair, our study is truly an external validation on a separate group of patients. In addition, they included factors such as cardiopulmonary exercise testing and aerobic fitness, both commonly used in preoperative assessments.² Our model uses a fairly simple risk score based on nearly all clinical factors.

Our study has limitations. Importantly, neither the derivation nor the validation group included patients who did not undergo surgery. All patients were deemed fit for either OAR or EVAR by the treating physician. Thus, our model cannot be applied to patients who are not thought to be candidates for either approach. It will be important to extend this work to those not thought to benefit from surgery because of comorbidities to improve our model for all patients with AAA. Second, both our derivation and validation cohorts are from northern regions of the country that have a predominantly white population. Validation of our model in regions with a higher degree of ethnic diversity will further strengthen our model. As the VQI matures and different regions attain sufficient 5-year follow-up on AAA repairs, this type of validation will be possible. In addition, we do not know the cause of death for patients and cannot determine aneurysm-related mortality. We do not have long-term medication assessment, and patients may have changed their medication profile; however, prior work has shown that >80% continue to take antiplatelet and statin medications at 1 year after surgery.¹⁴ Last, there are likely to be factors, such as a frailty index, active cancer, or complete assessment of aneurysm anatomy, that are not currently captured in the VQI and may further improve the discriminatory ability of our model. The impact of other factors, such as suprarenal clamping for OAR, and the impact of dialysis will also require further study.

CONCLUSIONS

We have been able to validate a survival model after OAR and EVAR. We have confirmed that age, severe CAD, COPD, renal insufficiency, and use of appropriate cardiovascular

medications can help identify patients at risk for reduced 5-year survival after AAA repair. Whereas medications are a modifiable factor, the remaining factors represent severe levels of organ dysfunction that place patients at high risk for overall death. Although EVAR can reduce perioperative mortality, these factors must remain important components of patient selection to ensure that optimal benefit is achieved from prophylactic AAA repair. This model can be easily used in the office to assist providers and patients in treatment decision-making.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. EVAR Trial Participants. Endovascular aneurysm repair and outcome in patients unfit for open repair of abdominal aortic aneurysm (EVAR trial 2): randomised controlled trial. *Lancet*. 2005; 365:2187–92. [PubMed: 15978926]
2. Carlisle JB, Danjoux G, Kerr K, Snowden C, Swart M. Validation of long-term survival prediction for scheduled abdominal aortic aneurysm repair with an independent calculator using only pre-operative variables. *Anaesthesia*. 2015; 70:654–65. [PubMed: 25959175]
3. De Martino RR, Goodney PP, Nolan BW, Robinson WP, Farber A, Patel VI, et al. Optimal selection of patients for elective abdominal aortic aneurysm repair based on life expectancy. *J Vasc Surg*. 2013; 58:589–95. [PubMed: 23642927]
4. Huang Y, Gloviczki P, Oderich GS, Duncan AA, Kalra M, Fleming MD, et al. Outcome after open and endovascular repairs of abdominal aortic aneurysms in matched cohorts using propensity score modeling. *J Vasc Surg*. 2015; 62:304–11.e2. [PubMed: 25943452]
5. Chaikof EL, Blankensteijn JD, Harris PL, White GH, Zarins CK, Bernhard VM, et al. Reporting standards for endovascular aortic aneurysm repair. *J Vasc Surg*. 2002; 35:1048–60. [PubMed: 12021727]
6. Chaikof EL, Fillinger MF, Matsumura JS, Rutherford RB, White GH, Blankensteijn JD, et al. Identifying and grading factors that modify the outcome of endovascular aortic aneurysm repair. *J Vasc Surg*. 2002; 35:1061–6. [PubMed: 12021728]
7. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al. Using standardized serum creatinine values in the Modification of Diet in Renal Disease study equation for estimating glomerular filtration rate. *Ann Intern Med*. 2006; 145:247–54. [PubMed: 16908915]
8. Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res*. 2016; 25:1692–706. [PubMed: 23907781]
9. Mousa AY, Bozzay J, Broce M, Yacoub M, Stone PA, Najundappa A, et al. Novel risk score model for prediction of survival following elective endovascular abdominal aortic aneurysm repair. *Vasc Endovascular Surg*. 2016; 50:261–9. [PubMed: 27114446]
10. Cronenwett JL, Kraiss LW, Cambria RP. The Society for Vascular Surgery Vascular Quality Initiative. *J Vasc Surg*. 2012; 55:1529–37. [PubMed: 22542349]
11. Khashram M, Jenkins JS, Jenkins J, Kruger AJ, Boyne NS, Foster WJ, et al. Long-term outcomes and factors influencing late survival following elective abdominal aortic aneurysm repair: a 24-year experience. *Vascular*. 2016; 24:115–25. [PubMed: 25972030]
12. Lim S, Halandras PM, Park T, Lee Y, Crisostomo P, Hershberger R, et al. Outcomes of endovascular abdominal aortic aneurysm repair in high-risk patients. *J Vasc Surg*. 2015; 61:862–8. [PubMed: 25704411]
13. Sevilla N, Clara A, Diaz-Duran C, Ruiz-Carmona C, Ibanez S. Survival after endovascular abdominal aortic aneurysm repair in a population with a low incidence of coronary artery disease. *World J Surg*. 2016; 40:1272–8. [PubMed: 26711643]

14. De Martino RR, Eldrup-Jorgensen J, Nolan BW, Stone DH, Adams J, Bertges DJ, et al. Perioperative management with antiplatelet and statin medication is associated with reduced mortality following vascular surgery. *J Vasc Surg.* 2014; 59:1615–21.e1. [PubMed: 24439325]

Author Manuscript

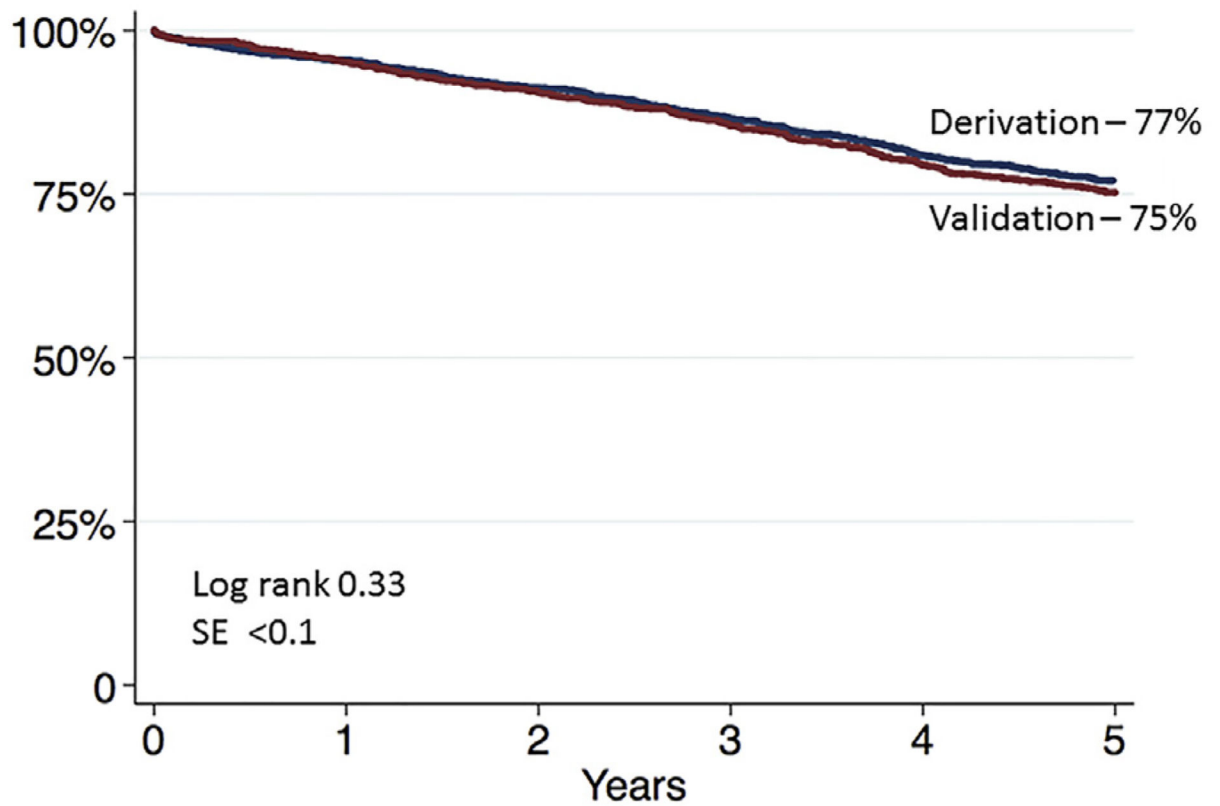
Author Manuscript

Author Manuscript

Author Manuscript

HIGHLIGHTS

- **Type of Research:** Single-center retrospective review and a comparison with Vascular Study Group of New England (VSGNE) database
- **Take Home Message:** Five-year survival of 75% in 1038 abdominal aortic aneurysm patients with aneurysms <6.5 cm after elective repair was accurately predicted using a model derived from data of 2367 similar patients in the VSGNE database.
- **Recommendation:** This study suggests that 5-year survival after elective abdominal aortic aneurysm repair can be accurately predicted using the VSGNE prediction tool that includes age and a measure of cardiac, pulmonary, and renal comorbidities.



At Risk						
Derivation	2367	1744	1311	1000	740	512
Validation	1037	986	936	885	792	708cc

Fig 1. Five-year survival after abdominal aortic aneurysm (AAA) repair for derivation and validation groups. *SE*, Standard error.

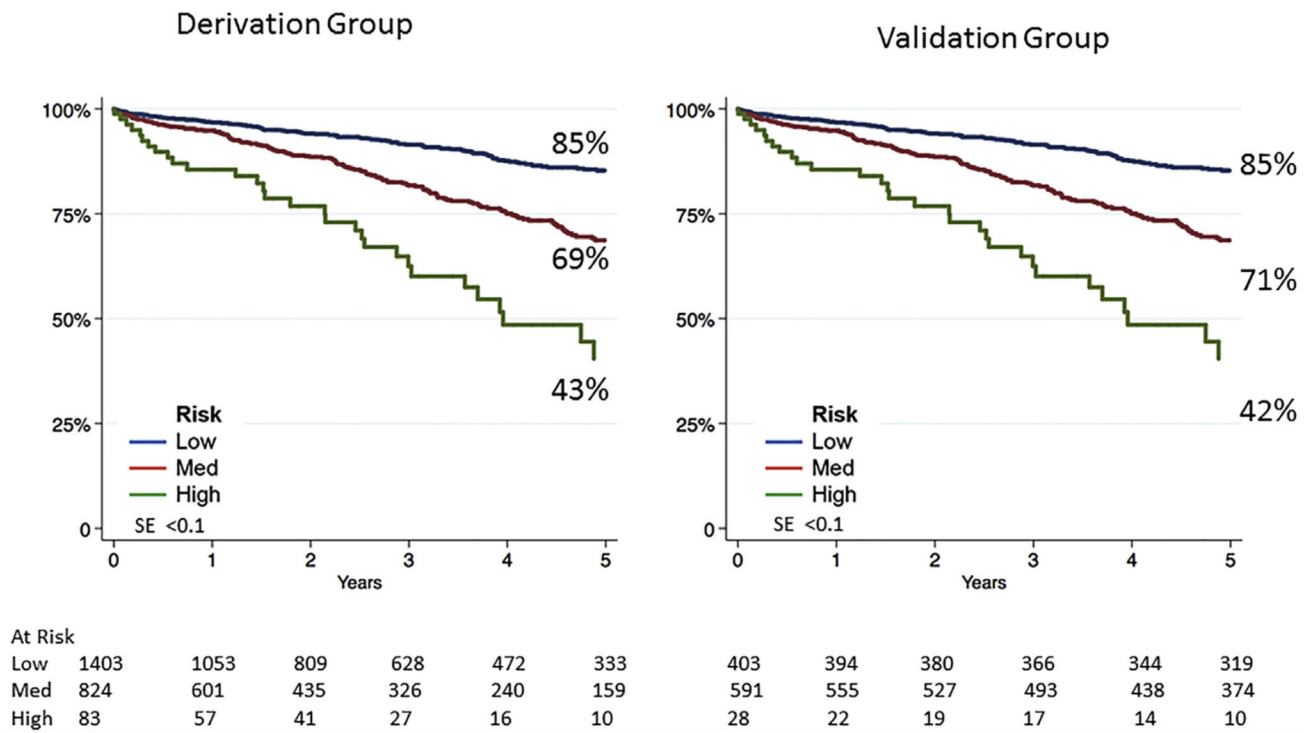


Fig 2. Five-year survival after abdominal aortic aneurysm (AAA) repair by risk category for derivation and validation groups. *SE*, Standard error.

Table

Comparison of patient and treatment factors between derivation and validation cohorts

Patient factors	Derivation (n = 2367)	Validation (n = 1038)	P
Age, years, mean (SD)	72.2 (8.2)	73.8 (7.9)	<.001
Male	67.8	76.3	<.001
Hypertension	84.7	85.1	.464
Diabetes	18.7	17.3	.352
CAD			<.001
None	66.2	50.7	
Prior MI/SVS cardiac 1	23.2	19.2	
Stable angina/SVS cardiac 2	9.8	26.7	
Recent MI/unstable angina/SVS cardiac 3	0.8	34.7	
COPD			<.001
None	66.0	67.2	
Not treated/SVS pulmonary 1	14.2	16.5	
On medication/SVS pulmonary 2	15.8	9.7	
On oxygen/SVS pulmonary 3	4.1	6.7	
eGFR			<.001
>60	65.3	50.5	
40–59	25.0	39.5	
30–39	7.2	6.9	
<30	2.6	3.1	
Prior or current smoker	88.3	82	<.001
Discharge aspirin	82	71.1	<.001
Discharge statin	74.4	60.1	<.001
Length of hospital stay, days, mean (SD)	4 (8.5)	4.6 (5.9)	.071
In-hospital mortality	0.8	0.8	1
% EVAR (vs open)	69.8	58.6	<.001
AAA diameter, mm, mean (SD)	53.7 (6.4)	54.2 (6.1)	.024

AAA, Abdominal aortic aneurysm; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate (Modification of Diet in Renal Disease formula); EVAR, endovascular aneurysm repair; MI, myocardial infarction; SD, standard deviation; SVS, Society for Vascular Surgery.

Values are reported as percentages unless otherwise indicated.