The Dutch Surgical Aneurysm Audit

The first results of adjusted hospital comparisons of Abdominal Aortic Aneurysm Surgery in the Netherlands

Niki Lijftogt

Colofon

The Dutch Surgical Aneurysm Audit, The first results of adjusted hospital comparisons of Abdominal Aortic Aneurysm Surgery in the Netherlands

Thesis, Leiden University, the Netherlands, 2019

ISBN: 978-94-6361-330-9
Design of cover, chapter sheets and bookmarker: Caroline Cracco, www.cracco.nl
Lay-out, printing and application: Optima Grafische Vormgeving, Rotterdam, www.ogc.nl

Financial support for printing of this thesis was provided by:
Stichting DICA, Universiteitsbibliotheek Leiden, Stichting Wetenschappelijk Onderzoek Algemene Chirurgie OLVG

The Dutch Surgical Aneurysm Audit

The first results of adjusted hospital comparisons of Abdominal Aortic Aneurysm Surgery in the Netherlands

Proefschrift

Ter verkrijging van de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties,
te verdedigen op
donderdag 7 november 2019 klokke 11.15 uur.

door

Niki Lijftogt

Geboren te Soest in 1986

Promotor:	Prof. dr. J.F. Hamming
Co-promotores:	Dr. A.C. Vahl *Onze Lieve Vrouwe Gasthuis, Amsterdam* Dr. M.W.J.M. Wouters *Nederlands Kanker Instituut – Antoni van Leeuwenhoek ziekenhuis, Amsterdam*
Promotiecommissie:	Prof. dr. M.J. Schalij Prof. dr. Y. van der Graaf *Universitair Medisch Centrum Utrecht, Utrecht* Prof. dr. G.J. de Borst *Universitair Medisch Centrum Utrecht, Utrecht*

May your choices reflect your hopes not your fears
Nelson Mandela
18 juli 1918 – 5 december 2013

Contents

1

General introduction, aim and outline of the thesis

Ever since the publication of the Institute of Medicine reports entitled 'To Err is Human' in 1999 and 'Crossing the Quality Chasm' in 2001, improving the quality of healthcare has been high on the political agenda in many Western countries.[1, 2] In the Netherlands too, professional associations of medical specialists have initiated quality improvement initiatives. One of these initiatives is the nationwide registration of vascular surgical patients, initiated by the Association of Vascular Surgeons in the Netherlands (Nederlandse Vereniging voor Vaatchirurgie: NVVV), a subassociation of the Dutch Association of Surgeons (Nederlandse Vereniging voor Heelkunde: NVVH).[3, 4]

On the initiative of the NVVV and facilitated by the Dutch Institute for Clinical Auditing (DICA), the Dutch Surgical Aneurysm Audit (DSAA) started up in 2012 and has been fully operational since 2013.[5] From this audit, casemix adjusted outcomes of abdominal aortic aneurysm surgery can be calculated, with the objective of quality improvement by providing benchmarked feedback to the vascular teams in the participating hospitals. All vascular centres in the Netherlands participate. Registration of operated patients is mandatory and quality indicators of the Dutch Health Care Inspectorate (Inspectie Gezondheidszorg en Jeugd: IGJ) were generated directly from this audit.[6] Other stakeholders who contributed to the development of the DSAA were health care insurers and patient associations.[7] The previous quality indicators of the Inspectorate have been combined with new quality indicators, which can be seen in the transparency portal of DICA (https://dica.nl/dsaa/documenten). Insurers were interested in the comparison of hospital outcomes, for the purpose of the ''selective purchasing'' of care. Patient associations (Harteraad) intended to use the information to assist individual patients in choosing a hospital.[7] In addition, the Dutch National Health Care Institute (Zorginstituut Nederland: ZIN) was interested in the effectiveness of endovascular aneurysm treatment (EVAR) in relation to the classic open repair (OR), and the development and implementation of volume standards for ruptured abdominal aortic aneurysm (AAA) surgery.[8]

Measuring Health Care Quality in the Netherlands

Registries have become an inherent part of the quality policy of professional organizations in the Netherlands. Traditionally, the quality instruments used by the NVVH encompassed the initial and continuing training of surgeons, the development of evidence-based guidelines, and external peer-review of surgical departments. Over the past decade, quality standards, certification in surgical specialties (e.g. vascular surgery) and nationwide surgical audits have been added to the quality policy of the NVVH.[9]

Comparing the adjusted outcomes of hospital surgical departments may stimulate providers to improve their practice.[10, 11] The results of these comparisons can be used primarily for internal quality improvement projects, and subsequently, where appropriate, to improve transparency for the public.[12] It is suggested that they may also be used for negotiations with health insurance companies for the selective purchasing of hospital health care.[12]

The NVVV has collaborated with the vascular patient society and the IGJ to develop a set of outcome measures, i.e. the "quality indicators of care".[3, 6, 7] Structural and process parameters have also been developed (see Table 1). The aim was to develop quality indicators that described the minimum necessary qualifications for a surgical team to perform AAA surgery in the Netherlands.[13] The development of quality indicators is a continuous process, in which this set is now adapted and can be seen in the DICA portal. Only the outcome indicators are left for hospital comparisons.

Table 1. Yearly/Annual quality measurements of Dutch Surgical Aneurysm Audit [2014]

Structure indicators:
the number of patients who underwent primary elective AAA surgery per hospital.
the availability of an aneurysm intervention team. [Cardiologist specialised in cardiovascular risk management, internal vascular specialist, neurologist, two certified vascular surgeons, the number of certified endovascular specialists e.g. intervention radiologist or vascular surgeon]
certified specialists in vascular/endovascular surgery. [available 24/7]
the presence of a weekly multidisciplinary meeting to discuss vascular patients, the results of which are registered in the electronic patient file.
file documented information is handed over to every patient prior to an AAA operation
every abdominal aortic aneurysm patient may ask their question by phone during office hours and will receive an answer on that same day.
EVAR procedures are performed by an endovascular specialist
Process indicators:
percentage of primary elective AAA patients discussed in a multidisciplinary team preoperatively.
percentage of primary AAA patients of whom all necessary variables are fully registered in the Dutch Surgical Aneurysm Audit.
Outcome indicators:
percentage mortality of primary [elective or acute, EVAR or OSR] patients within 30 days or during hospital stay.
percentage complications of primary [elective or acute, EVAR or OSR] patients within 30 days or during hospital stay.
percentage reinterventions of primary elective AAA patients.
percentage readmissions of primary elective AAA patients within 30 days..

An important possible indicator is hospital volume. There is substantial evidence for an association between the volume of AAA surgery per hospital and postoperative mortality.[14] However, there is no clear volume cut-off, but outcomes tend to improve exponentially up to 20 operations per year.[15] In the Netherlands, a minimum of 20 elective operations per year has been set as one of the quality standards for elective AAA surgery.[6] The volume indicator results from vascular surgery departments in the Netherlands can be derived directly from the DSAA.

Challenges in measuring quality of health care

There are several potential problems in comparing the outcomes of health care providers; so providers may be mistakenly identified as underperforming hospitals or the discriminating ability of the outcome indicator is only minor and non-significant.

The first problem is the possible difference in hospital casemix. Therefore, risk adjustment for these casemix variables is necessary.[16]

The second problem is the statistical uncertainty (random error), if the number of patients or number of events is low.[16] Hopefully, a minimum annual volume for each hospital and the inclusion of several consecutive years in the analysis helps to overcome this issue.

Thirdly, the question arises whether the chosen outcome indicator (e.g. mortality) is a satisfactory indicator of hospital quality. Other factors, such as patient-reported outcome measurements or composite measures such as textbook outcome (TO), or failure to rescue (FTR) could be more informative.[17-19]

Fourthly, the chosen observation period of in- hospital and/or 30 days in the DSAA has its limitations. For example, the outcome of EndoVascular Aneurysm Repair (EVAR) may be less favorable than open surgical repair (OSR) on long-term follow -up.[20]

The aim of this thesis is to investigate the following questions:

1. What are the results of AAA surgery in the Netherlands in general and what are the results of a single hospital compared to the national mean?
2. How important is case mix correction? Which models or variables should we use?
3. What are the true advantages of EVAR and is there any relation with preference for EVAR and outcome of EVAR and OSR?
4. Should mortality be the main outcome variable? What is the value of composite outcome indicators?

Outline of the thesis

The DSAA registers preselected variables for quality indicators and casemix variables for risk-adjustment and has been developed to measure and compare the risk-adjusted outcomes of hospitals.[4]

In **Chapter 2** the design and the initial results of the first 2 years of registration in the DSAA by every hospital in the Netherlands performing AAA surgery are described.[21] Results were compared with other registries, e.g. SWEDVASC and Medicare.[22, 23] Collaboration for purposes of international comparison between several registries has already been initiated.[24] However, differences in definitions of elective or urgent care, type of reporting, starting point, follow-up, and the voluntary character of these registries hamper comparisons.

Casemix adjustment

In **Chapters 3** and **4** the importance of risk-adjustment on casemix variables when investigating hospital variation is addressed.[25, 26] Hospitals may have population differences regarding patient- and disease-specific variables that are of influence on mortality.[27] Examples of patient- and disease-specific variables are age, gender, co-morbidity, laboratory results, physiologic parameters, and planned or unplanned surgery.[26, 28] Only variables that influence outcome but are not associated with the hospital can be used for risk-adjustment for hospital comparisons.[16] One variable of special interest is the type of procedure performed (EVAR or OSR). Because of the specific aneurysm configuration and patient factors it is decided by the vascular team which treatment to use, but this also depends on local expertise and infrastructure. The preference of the vascular team can therefore not be used for risk-adjustment.

In **Chapter 3** existing mortality risk prediction models for AAA surgery are compared.[26] Predicting mortality after aneurysm surgery is complex and may vary depending on emergency setting (acute ruptured, acute symptomatic (non-ruptured) or planned surgery), type of procedure (EVAR or OSR), and differences related to the timeline, populations or regions that are under analysis.[24] For example, prediction models developed before the introduction of EVAR have become less suitable for clinical practice.[29] Generalisability and the number of variables used continue to be a subject of debate. In **Chapter 4** it is argued that only a limited number of casemix variables might be necessary for the prediction of mortality.[25] Currently, the DSAA includes the variables based on V-POSSUM.[28] This has resulted in a rather extensive set of variables with a major registration burden. In **Chapter 4** we aim to determine the minimum set of meaningful variables required to effect risk-adjusted hospital comparisons.

Comparing outcome of EVAR with OSR

EVAR in elective AAA surgery results in a lower perioperative mortality than OSR (DREAM trial[30], EVAR 1 trial[31], OVER trial[32]). However, one study (ACE 1 trial[33]) found similar results for OSR and EVAR in a group of low - to moderate risk patients only. Almost 80% of patients undergoing elective infrarenal AAA can be treated by EVAR, which is leading to a decline in operative experience of OSR in lower volume hospitals.[34, 35] This might become a confounding factor and it implies that the operative technique, often used for risk-stratification and risk-prediction, cannot be used for casemix correction. In observational research patient and/or treatment selection can bias the results. Randomised trials do not have this bias but often contain a selected group of patients group what does not take into account treatment preference and/or competence of the surgeon and does not completely reflect clinical practice. This affects the generalisability.[36] To overcome confounding factors affecting the choice of treatment, regression- or propensity score analyses have been proposed, in order to adjust for know confounders.[37] Another technique to deal with measurable and also unmeasurable confounders is the instrumental variable analysis or ecological analysis.[36, 37] This analysis, **Chapter 5**, works under the assumption that the choice of treatment is not determined by patient characteristics, prognosis or by differences in other aspects of surgical care.[36-38]

Composite outcome measures

Besides mortality, also other outcome measures are registered in the DSAA: complications, prolonged hospital stay, reinterventions and readmissions.

Failure to Rescue (FTR) is an example of a composite outcome measure, which is presented in **Chapter 6**.[39] FTR measures the consequences of complications after surgery.[40] Because FTR seems to be more closely related to hospital structure and processes, it could be a useful additional quality indicator to mortality and major complication rates.[41, 42]

Another example of a relatively new composite outcome measure in vascular surgery is "textbook outcome", which represents in short a hospital stay after surgery without any adverse event.[43] More comprehensive summative outcome measures such as textbook outcome have proven to be of additional value to the more detailed individual quality indicators.[44]

Moreover, composite outcome measures, also used in colorectal and gastro-esophageal cancer surgery, result in higher percentages of event rates, what usually leads to a better discriminatory ability in the evaluation of hospital performances.[18, 45]

REFERENCES

1 Institute of Medicine (US) Committee on Quality of Health Care in America 2000 doi:NBK225182 [bookaccession].
2 Institute of Medicine (US) Committee on Quality of Health Care in America 2001 doi:NBK222274 [bookaccession].
3 Nederlandse Vereniging voor Vaatchirurgie (NVVV). Available at: http://nvvv-vaatchirurgie.nl/.
4 Dutch Surgical Aneurysm Audit (DSAA). Available at: https://www.dica.nl/dsaa/home.
5 The Dutch Institute for Clinical Auditing (DICA). Available at: www.clinicalaudit.nl.
6 Inspectie Gezondheidszorg en Jeugd (IGJ). Available at: https://www.igz.nl/.
7 Harteraad. Available at: http://www.harteraad.nl/.
8 Zorginstituut Nederland (ZIN). Available at: https://www.zorginstituutnederland.nl/.
9 Nederlandse Vereniging voor Heelkunde (NVVH). Available at: https://heelkunde.nl/.
10 Van Leersum NJ, Snijders HS, Henneman D, et al. The Dutch surgical colorectal audit. Eur J Surg Oncol 2013;39:1063-70.
11 Fischer C, Lingsma H, Hardwick R, et al. Risk adjustment models for short-term outcomes after surgical resection for oesophagogastric cancer. Br J Surg 2016;103:105-16.
12 Wouters MWJM. Measuring and improving quality of care in surgical oncology. ISBN: 9789082065206. [bookaccession]
13 Indicatoren DSAA. Available at: http://nvvv-vaatchirurgie.nl/Indicatoren.
14 Holt PJ, Poloniecki JD, Gerrard D, et al. Meta-analysis and systematic review of the relationship between volume and outcome in abdominal aortic aneurysm surgery. Br J Surg 2007;94:395-403.
15 Holt PJ, Poloniecki JD, Loftus IM, et al. Epidemiological study of the relationship between volume and outcome after abdominal aortic aneurysm surgery in the UK from 2000 to 2005. Br J Surg 2007;94:441-8.
16 Steyerberg E. Clinical prediction models; A Practical Approach to Development, Validation, and Updating: Springer 2009. ISBN: 9780387772448. [bookaccession]
17 Duncan R, Essat M, Jones G, et al. Systematic review and qualitative evidence synthesis of patient-reported outcome measures for abdominal aortic aneurysm. Br J Surg 2017;104:317-27.
18 Busweiler LA, Schouwenburg MG, van Berge Henegouwen MI, et al. Textbook outcome as a composite measure in oesophagogastric cancer surgery. Br J Surg 2017;104:742-50.
19 Henneman D, Snijders HS, Fiocco M, et al. Hospital variation in failure to rescue after colorectal cancer surgery: results of the Dutch Surgical Colorectal Audit. Ann Surg Oncol 2013;20:2117-23.
20 de la Motte L, Falkenberg M, Koelemay MJ, et al. Is EVAR a durable solution? Indications for reinterventions. J Cardiovasc Surg (Torino) 2018;59:201-12.
21 Lijftogt N, Vahl AC, Wilschut ED, et al. Adjusted Hospital Outcomes of Abdominal Aortic Aneurysm Surgery Reported in the Dutch Surgical Aneurysm Audit. Eur J Vasc Endovasc Surg 2017;53:520-32.
22 Troëng T. The Swedish National Registry for Vascular Surgery; Swedvasc 2011.
23 Kathleen N. Lohr E, Committee to Design a Strategy for Quality Review and Assurance in Medicare, Institute of Medicine. Medicare: A Strategy for Quality Assurance, Volume II: Sources and Methods 1990, ISBN: 0-309-55992-8. [bookaccession].

24 Mani K, Lees T, Beiles B, et al. Treatment of abdominal aortic aneurysm in nine countries 2005-2009: a vascunet report. Eur J Vasc Endovasc Surg 2011;42:598-607.

25 Niki Lijftogt MD, Anco Vahl M, PhD, Esmee M van der Willik, MSc, et al. Towards Optimizing Risk Adjustment in the Dutch Surgical Aneurysm Audit (DSAA). Ann Vasc Surg 2019 May 7, doi: 10.1016/j.avsg.2019.02.032.

26 Lijftogt N, Luijnenburg TWF, Wilschut ED, et al. Systematic review of mortality risk prediction models in the era of endovascular abdominal aortic aneurysm surgery. Br J Surg 2017 Jul;104(8):964-976.

27 Mani K, Venermo M, Beiles B, et al. Regional Differences in Case Mix and Peri-operative Outcome After Elective Abdominal Aortic Aneurysm Repair in the Vascunet Database. Eur J Vasc Endovasc Surg 2015;49(6):646-652.

28 Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. Br J Surg 2001;88(7):958-63.

29 Boyle JR, Thompson MM, Vallabhaneni SR, et al. Pragmatic minimum reporting standards for endovascular abdominal aortic aneurysm repair. J Endovasc Ther 2011;18:263-71.

30 Prinssen M, Verhoeven EL, Buth J, et al. A randomized trial comparing conventional and endovascular repair of abdominal aortic aneurysms. N Engl J Med 2004;351:1607-18.

31 Greenhalgh RM, Brown LC, Kwong GP, et al. Comparison of endovascular aneurysm repair with open repair in patients with abdominal aortic aneurysm (EVAR trial 1), 30-day operative mortality results: randomised controlled trial. Lancet 2004;364:843-8.

32 Lederle FA, Freischlag JA, Kyriakides TC, et al. Outcomes following endovascular vs open repair of abdominal aortic aneurysm: a randomized trial. JAMA 2009;302:1535-42.

33 Becquemin JP, Pillet JC, Lescalie F, et al. A randomized controlled trial of endovascular aneurysm repair versus open surgery for abdominal aortic aneurysms in low- to moderate-risk patients. J Vasc Surg 2011;53:1167,1173.e1.

34 Lozano FS, Marinello J, Moreno RM, et al. Monitoring the practice of vascular surgery: findings from a national registry (1996-2011). World J Surg 2014;38:241-51.

35 Arya S, Kim SI, Duwayri Y, et al. Frailty increases the risk of 30-day mortality, morbidity, and failure to rescue after elective abdominal aortic aneurysm repair independent of age and comorbidities. J Vasc Surg 2015;61:324-31.

36 Johnston SC. Combining ecological and individual variables to reduce confounding by indication: case study--subarachnoid hemorrhage treatment. J Clin Epidemiol 2000;53:1236-41.

37 Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med 2014;33:2297-340.

38 Lijftogt, N., Vahl, A.C. , Karthaus, E.G., van der Willik, E.M., Amodio, S., van Zwet, E.W. Hamming, J.F., In collaboration with the Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing. High postoperative mortality after elective open surgical abdominal aneurysm repair in hospitals with EVAR preference and lessons from an instrumental variable analysis of the Dutch Surgical Abdominal Aneurysm Audit. Submitted Br J Surg 2019.

39 Lijftogt N, Karthaus EG, Vahl A, et al. Failure to Rescue - a Closer Look at Mortality Rates Has No Added Value for Hospital Comparisons but Is Useful for Team Quality Assessment in Abdominal Aortic Aneurysm Surgery in The Netherlands. Eur J Vasc Endovasc Surg 2018 Nov;56(5):652-661.

40 Silber JH, Williams SV, Krakauer H, et al. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. Med Care 1992;30:615-29.
41 Sinha S, Ata Ozdemir B, Khalid U, et al. Failure-to-rescue and interprovider comparisons after elective abdominal aortic aneurysm repair. Br J Surg 2014;101:1541-50.
42 Waits SA, Sheetz KH, Campbell DA, et al. Failure to rescue and mortality following repair of abdominal aortic aneurysm. J Vasc Surg 2014;59:909,914.e1.
43 Karthaus E, Lijftogt N, Busweiler L, et al. Textbook Outcome: a Composite Measure for Quality of Elective Aneurysm Surgery. Ann Surg. 2017 Nov;266(5):898-904.
44 Dijs-Elsinga J, Otten W, Versluijs MM, et al. Choosing a hospital for surgery: the importance of information on quality of care. Med Decis Making 2010;30:544-55.
45 Kolfschoten NE, Kievit J, Gooiker GA, et al. Focusing on desired outcomes of care after colon cancer resections; hospital variations in 'textbook outcome'. Eur J Surg Oncol 2013;39:156-63.

2

Adjusted Hospital Outcomes of Abdominal Aortic Aneurysm Surgery Reported in the Dutch Surgical Aneurysm Audit

N. Lijftogt
A. C. Vahl
E. D. Wilschut
B. H. P. Elsman
S. Amodio
E. W. van Zwet
V. J. Leijdekkers
M. W. J. M. Wouters
J. F. Hamming

in collaboration with the Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit the Dutch Institute for Clinical Auditing

European Journal for Vascular and Endovascular Surgery 2017

ABSTRACT

Background:

The Dutch Surgical Aneurysm Audit (DSAA) is mandatory for all patients with primary abdominal aortic aneurysms (AAAs) in the Netherlands. The aims are to present the observed outcomes of AAA surgery against the predicted outcomes by means of V-POSSUM (Vascular-Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity). Adjusted mortality was calculated by the original and re- estimated V(physiology)-POSSUM for hospital comparisons.

Material and Methods:

All patients operated on from January 2013 to December 2014 were included for analysis. Calibration and discrimination of V-POSSUM and V(p)-POSSUM was analysed. Mortality was benchmarked by means of the original V(p)-POSSUM formula and risk-adjusted by the re-estimated V(p)-POSSUM on the DSAA.

Results:

In total, 5898 patients were included for analysis: 4579 with elective AAA (EAAA) and 1319 with acute abdominal aortic aneurysm (AAAA), acute symptomatic (SAAA; n = 371) or ruptured (RAAA; n = 948). The percentage of endovascular aneurysm repair (EVAR) varied between hospitals but showed no relation to hospital volume (EAAA: p = .12; AAAA: p = .07). EAAA, SAAA, and RAAA mortality was, respectively, 1.9%, 7.5%, and 28.7%. Elective mortality was 0.9% after EVAR and 5.0% after open surgical repair versus 15.6% and 27.4%, respectively, after AAAA. V-POSSUM overestimated mortality in most EAAA risk groups ($p < .01$). The discriminative ability of V- POSSUM in EAAA was moderate (C-statistic: .719) and poor for V(p)-POSSUM (C-statistic: .665). V-POSSUM in AAAA repair overestimated in high risk groups, and underestimated in low risk groups ($p < .01$). The discriminative ability in AAAA of V-POSSUM was moderate (.713) and of V(p)-POSSUM poor (.688). Risk adjustment by the re-estimated V(p)-POSSUM did not have any effect on hospital variation in EAAA but did in AAAA.

Conclusion:

Mortality in the DSAA was in line with the literature but is not discriminative for hospital comparisons in EAAA. Adjusting for V(p)-POSSUM, revealed no association between hospital volume and treatment or outcome. Risk adjustment for case mix by V(p)-POSSUM in patients with AAAA has been shown to be important.

INTRODUCTION

Auditing hospital outcomes after surgery is a powerful tool with which to monitor healthcare quality.[1] In the Netherlands several audits for surgical outcomes have been developed in cooperation with the Dutch Institute for Clinical Auditing. These audits, meant to improve healthcare, are developed in agreement with several stakeholders, such as insurance companies and the health inspectorate of the ministry of healthcare. Complete registration of data with a minimum of missing values and a motivated administrative culture are essential for robust and accurate conclusions for healthcare quality.[2] Therefore, a reduced set of preoperative patient -or disease related variables, easy to register, is desirable, especially as not every variable registered and of influence on mortality, needs to be included for casemix adjustment.[3,4]

The web based Dutch Surgical Aneurysm Audit (DSAA), introduced in 2012 and mandatory since 2013, registers all primary abdominal aortic aneurysm (AAA) operations in the Netherlands.

Because baseline characteristics of populations may differ between hospitals, with concomitant differences in outcome, risk adjustment by patient and disease specific characteristics for outcome measurement is necessary.[5] This can be achieved by using pre-operative variables of influence on the outcome.[6] Numerous models predicting mortality by pre- or perioperative variables have been developed for aneurysm surgery. Only a few of them have been validated multiple times and are therefore considered as accurate, such as the Glasgow Aneurysm Score (GAS) or the Vascular Biochemistry and Haematology Outcome Model (VBHOM).[7,8]

The Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (V-POSSUM) is a well known peri-operative mortality risk prediction model.[9,10] However, the operative variables included in the model are not suitable for adjustment to compare hospitals because they are, to a large extent, dependent on surgical care, such as, for example, blood loss. The "physiology-only" score of V-POSSUM (V(p)-POSSUM) only contains patient and disease specific characteristics, which can be suitable as casemix information for hospital comparisons.

Since the introduction of endovascular aneurysm repair (EVAR) mortality has decreased in elective AAA surgery (EAAA); however, the advantage of EVAR over open surgical repair (OSR) in ruptured abdominal aortic aneurysm (RAAA) suggested in observational studies has not been confirmed in randomised trials.[11-18] An explanation for differences between observational research and randomised trials could be selection bias.[16,19]

Large registries, of consecutive patients undergoing surgery for acute aneurysms, might add insight to this issue. However, the results from national registries can be difficult to compare owing to differences in prevalence of RAAA in countries with screening programs, the percentage that refrains from operative repair of RAAA, and the variation in percentage of EVAR implemented.[20-22]

The aim of this study was to report the first results of auditing AAA surgery in the Netherlands. Post-operative mortality was the primary outcome parameter. As a secondary outcome parameter, variations in the implementation of EVAR and the possible association with volume were investigated. The performance of V-POSSUM, as prediction model, was assessed. For casemix correction hospital outcomes were compared and adjusted with the original V(p)- POSSUM and the re-estimated V(p)-POSSUM on the DSAA population.

MATERIAL AND METHODS

Clinical data

The DSAA is a mandatory, nationwide, population and web based database with detailed patient, diagnostic, procedural, and outcome data of all patients with a primary infra- or juxtarenal AAA operation in the Netherlands. Under Dutch law, no ethical approval or informed consent was required. In 2017 a project will be initiated to validate the existing data set. Patients prospectively registered in the DSAA, operated on for an AAA between 1 January 2013 and 31 December 2014 were included for analysis. Excluded were patients with secondary or revision surgery, surgery of highly complex aneurysm (suprarenal and thoraco-abdominal), and mycotic or infected aneurysms.[23] Furthermore, patients with incomplete data concerning date of birth, date of surgery, survival state, setting, or type of procedure (EVAR/OSR) were excluded (see "Results", subsection "Baseline characteristics"). Patient and treatment characteristics were described. Procedure for analysis, other than baseline, was calculated following "inten-

Table 1. Formula for the calculation of the POSSUM scores

	Model	Scoring algorithm + formula
Risk prediction	V-POSSUM	(ln(R/1-R))=-8.0616 + (0.1552 * physiologic score) + (0.1238*operative score)
		R=1/(1+e^-(-8.0616 + (0.1552 * physiologic score) + (0.1238*operative score)))
Risk adjustment	V(p)-POSSUM	(ln(R/1-R))=-6.0386 + (0.1539 * physiologic score)
		R=1/(1+e^-(-6.0386 + (0.1539 * physiologic score))

tion to treat" analysis and the percentage of EVAR (EVAR/(EVAR + OSR)) was tested for the association with hospital volume. For hospital comparisons two groups of patients were analysed: EAAA and AAAA.

AAAA was defined as either acute non-ruptured without extravasation needing surgery within 24h after presentation (SAAA), or ruptured with extravasation requiring immediate surgery (RAAA).

Clinical outcomes

The primary outcome measure was 30 day or in hospital mortality. A sub-analysis was performed, when appropriate, by year of registration. Other outcome measurements were peri- and post-operative complications, any re-interventions, and length of hospital stay. Peri-operative complications were cardiopulmonary resuscitation, unplanned closure of a hypogastric artery, and visceral and renal injury. Postoperative complications concerned bleeding defined as blood loss needing surgery or blood transfusion; colonic ischaemia; arterial occlusion; paralysis; prosthesis associated issues (migration, infection, any endovascular leakage); abscess, defined as an abscess of the inguinal wound; abdominal wound or intra-abdominal wounds; visceral complications (colonic or splenic); wound dehiscence; ileus; colostomy; major amputation; or profound wound infections and cardiopulmonary complications; renal insufficiency; neurological or thromboembolic complications; and infections other than surgical site or pulmonary infections not directly related to the surgical procedure. Because readmission could only be registered as an optional choice of the DSAA survey it was analysed when registered.

Prediction by V-POSSUM and adjustment by V(p)-POSSUM

The V-POSSUM (operative and physiological score) and V(p)- POSSUM (only physiological score) were calculated using the following variables: (i) physiological (age, cardiac comorbidity, pulmonary comorbidity, electrocardiogram status, systolic blood pressure, pulse rate, haemoglobin, leukocytes, urea (calculated from creatinine), sodium, potassium, Glasgow Coma Scale); (ii) operative (operation severity [severity of procedure] was calculated as "major" for every procedure -EVAR and OSR- in accordance with the available literature], number of procedures, perioperative blood loss, peritoneal contamination, malignancy status and setting [EAAA, SAAA, or RAAA]). Calculations for the V-POSSUM and V(p)-POSSUM were performed using the formulas shown in Table 1.[9,10,24,25] Predicted mortality was calculated using the exponent of the V-POSSUM in the following formula:[9]

Mortality = 1/1 + exp -(V-POSSUM or V(p)-POSSUM)

Mortality risk prediction

The observed mortality was compared with the expected (or predicted) mortality by V-POSSUM using the Hosmer and Lemeshow test,[9,10] which indicates a good calibration when not significant.[9,26] This goodness of fit statistic is computed as the Pearson chi-square from the contingency table of observed and expected (predicted) frequencies after having grouped the observations into deciles based on the predicted probabilities. The null hypothesis states that there are no systematic differences between observed and expected counts in different severity classes. The main idea behind this test statistic is the more closely the predicted and the observed frequencies match, the better the fit. Differences between observed and expected were shown in a bar plot in terms of percentages. The expected mortality was also calculated for the different procedures and compared with the observed mortality, tested according to the Fisher's combined probability test. As described earlier, two groups of patients were analysed: EAAA and AAAA. Combining the two patient groups having acute surgery was necessary in order to have an adequate sample size for the acute setting. When appropriate SAAA and RAAA were analysed separately.

Performance comparison

To compare the mortality between centres, an unadjusted funnel plot was constructed. Next, the adjusted mortality, based on the V(p)-POSSUM as casemix adjustment, was computed in a funnel plot to compare the performance of hospitals in the DSAA with the original British population (benchmark) on which V(p)-POSSUM was constructed. Note that V(p)-POSSUM was used rather than V-POSSUM, because the former is based on pre-operative patient characteristics (physiology parameters) only. Finally, V(p)-POSSUM was used as a casemix variable by fitting a logistic regression model on the DSAA data. This allowed a risk-adjusted comparison to be made between the centres in the DSAA. All results are shown in funnel plots as the (effective) hospital volume versus the standardised mortality rate (i.e., the ratio of observed to expected events), together with 95% confidence intervals (CIs).

In the funnel plots two 95% CIs are reported. The orange narrow one represents a 95% CI that can be used to test the performance of any particular hospital. A hospital that actually performs exactly according the national average will still have 5% probability of falling outside this funnel (i.e., false positive). The wider, red 95% CI is corrected for multiple comparisons by using the Bonferroni correction. This means that if, for example, all hospitals perform exactly according to the national average, then there is a 5% probability that at least one of them will fall outside the red, wider funnel.

Missing data presented as "missing values" in the baseline tables were allocated to the normal category in V- POSSUM.[27] Normality for continuous variables was tested by the Kolmogorov-Smirnov test and rejected when $p < .05$. Medians are presented with an interquartile range; means are presented with a SD. Analysis was performed in SPSS version 23.0 (IBM, Armonk, NY, USA).

RESULTS

Baseline characteristics

A total of 5979 patients had primary AAA surgery and were registered in the DSAA during the study period in 65 hospitals. Patients with specific missing data, as is specified above under "Clinical data" (n = 81; 1.4%), were excluded. Of the remaining 5898 patients, 4579 patients had EAAA surgery (77.6%) and 1319 patients had AAAA surgery (RAAA surgery [n = 948; 16.1%] and SAAA surgery [n = 371; 6.3%]). Almost three quarters of the EAAA patients (74.8%) were treated primarily by EVAR (74.5% were completed by EVAR [0.3% converted to OSR]). The majority of AAAA patients received OSR (60.7%). In the subgroup of patients with SAAA, 53.6% had EVAR (0.8% converted to OSR) versus 33.8% in patients with RAAA (1.4% converted to OSR). The converted EVAR were analysed as EVAR according the "intention to treat" principle. General baseline characteristics used for V-POSSUM are shown in Tables 2, 3 and 4.

Clinical outcomes

Procedure

The variation in percentage of EVAR performed was wide. In the majority of hospitals >50% of EVAR were performed in patients with EAAA (range 13-100%) There was no association between hospital volume and the percentage of EVAR performed ($p = .12$). High AAAA volume hospitals had a greater preference for EVAR compared with low volume hospitals, but this was not significant (range 0- 100%; $p = .07$).

Mortality

The overall 30 day or in hospital mortality after EAAA surgery was 1.9% versus 7.5% after SAAA and 28.7% after RAAA surgery. EAAA mortality in 2013 and 2014 was comparable (1.9% and 2.0%, respectively). Mortality for AAAA was higher in 2013 than in 2014 in both settings (8.6% vs. 6.8% after SAAA and 34.8% vs. 23.8% after RAAA). The overall mortality after AAAA surgery was 22.7% (15.6% after EVAR vs. 27.4% after OSR). EVAR in EAAA showed a mortality rate of 0.9% and OSR a mortality rate of 5.0%. Mortality by procedure and setting is presented in Table 5.

Table 2. Baseline clinical characteristics

Setting	EAAA	SAAA	RAAA
Patients (n)	4 579	371	948
Patient characteristics			
Gender. Male (%. 95%CI)	86.8 (83.4-87.8)	81.9 (78.0-85.8)	85.6 (85.8-87.8)
Age. Mean ± SD (years)	73 ± 7.7	73 ± 8.8	74 ± 8.4
Diameter. Median (IQR) (mm)	58 (55-64)	66 (55-80)	78 (65-90)
Missing n(%)	107 (2.3)	8 (2.2)	60 (6.3)
Heart frequency median (IQR) (BPM)	72 (63-81)	79 (69-87)	83 (70-100)
Missing n(%)	359 (7.8)	49 (13.2)	131 (13.8)
Systolic blood pressure median (IQR) (mmHg)	140 (127-152)	144 (127-160)	107 (84-135)
Missing n(%)	278 (6.1)	42 (11.3)	91 (9.6)
Comorbidity			
Cardiac comorbidity (%. 95%CI)			
None	46.2 (44.8-47.6)	44.5 (39.4-49.6)	40.1 (37.0-43.2)
Peripheral oedema	8.1 (7.4-9.0)	5.9 (3.2-7.8)	6.4 (4.8-8.0)
Elevated CVP	1.4 (1.1-1.7)	2.2 (0.7-3.7)	1.2 (0.5-1.9)
Antihypertensive medication	38.6 (37.2-10.0)	39.4 (34.4-44.4)	28.8 (26.0-31.7)
Missing	5.7 (5.0-6.4)	8.1 (5.3-10.9)	23.5 (20.8-26.2)
Pulmonary comorbidity (%. 95%CI)			
None	75.4 (74.2-76.6)	73.9 (69.4-78.4)	59.9 (56.8-63.0)
Dyspnoea during exercise	19.3 (18.2-20.4)	14.8 (11.2-18.4)	15.3 (13.0-17.6)
Invalidating dyspnoea	2.7 (2.2-3.2)	1.9 (0.5-3.3)	2.8 (1.8-3.9)
Dyspnoea during rest/fibrosis	1.1 (0.8-1.4)	1.6 (0.3-2.9)	2.1 (1.2-3.0)
Missing	1.4 (1.1-1.7)	7.8 (5.1-10.5)	19.8 (17.3-22.4)
Malignancy (%. 95%CI)			
None	80.4 (79.3-81.6)	88.9 (85.7-92.1)	87.1 (85.0-89.2)
Primary only	4.2 (3.6-4.8)	2.2 (0.7-3.7)	2.3 (1.4-3.3)
Lymph node metastasis	13.9 (12.9-14.9)	7 (4.4-9.6)	7.2 (5.6-8.9)
Distant metastasis	0.6 (0.4-0.8)	1.1 (0.0-2.2)	0.7 (0.2-1.2)
Missing	1 (0.7-1.3)	0.8 (-0.1-1.7)	2.6 (1.6-3.6)

95% CI: p +- (1.96*(SQRT(p*(1 - p))/n)), where p = proportion and n = sample size. EAAA = elective abdominal aortic aneurysm; SAAA = symptomatic abdominal aortic aneurysm; RAAA = ruptured abdominal aortic aneurysm; CI = confidence interval; IQR = interquartile range; bpm = beats per min; SBP = systolic blood pressure; CVP = central venous pressure.

Table 3. Baseline characteristics: diagnostics

	EAAA	SAAA	RAAA
Patients (n)	4 579	371	948
Diagnostics			
Laboratory results (median. IQR)			
haemoglobin (mmol/L)	8.8 (8.1-9.3)	8.4 (7.5-9.2)	7.4 (6.4-8.3)
missing n(%)	104 (2.3)	7 (1.9)	36 (4.0)
leucocytes (*10^9/L)	7.9 (6.6-9.6)	9.0 (7.4-12.0)	12.8 (9.9-16.4)
missing n(%)	1727 (37.7)	40 (10.8)	90 (9.5)
sodium (mmol/L)	140 (138-141)	138 (136-140)	138 (135-140)
missing n(%)	387 (8.4)	13 (3.5)	58 (6.1)
potassium (mmol/L)	4.2 (4.0-4.5)	4.1 (3.8-4.5)	4.0 (3.7-4.4)
missing n(%)	288 (6.3)	12 (3.2)	60 (6.3)
creatinin (micromol/L)	90 (77-108)	85 (70-110)	108 (86-133)
missing n(%)	121 (2.6)	14 (3.8)	56 (5.9)
GCS (%. 95%CI)			
15	90.8 (90.0-91.7)	92.2 (89.0-94.5)	60.9 (57.7-63.9)
12 - 14	0	1.9 (0.5-3.3)	15.3 (13.1-17.7)
9 - 11	0	0	3.7 (2.5-4.9)
<9	0	0.5 (-0.2 to 1.2)	6.9 (5.3-8.5)
missing	9.2 (8.4-10.0)	5.4 (3.5-8.2)	13.3 (11.3-15.6)
ECG (%. 95%CI)			
normal	60.7 (59.3-62.1)	50.1 (45.1-55.2)	32 (29.1-35.0)
atrial fibrillation 60-90 BPM	7.1 (6.4-7.9)	6.5 (4.4.9.4)	5.2 (3.9-6.8)
Ischaemia	21.8 (20.6-23.0)	26.4 (22.2-31.1)	17.6 (15.3-20.2)
missing	10.4 (9.6-11.3)	17 (13.5-21.1)	45.3 (42.1-48.3)

95% CI: p+- (1.96*(SQRT(p*(1 - p))/n)), where p 1= proportion, n = sample size, SQRT = square root. Data are median (IQR) unless otherwise indicated. EAAA = elective abdominal aortic aneurysm; SAAA 1= symptomatic abdominal aortic aneurysm; RAAA = ruptured abdominal aortic aneurysm; GCS = Glasgow Coma Scale; CI = confidence interval; ECG = electrocardiography; bpm = beats per min.

Table 4. Baseline characteristics (continued): operative

	EAAA	SAAA	RAAA
Patients (n)	4 579	371	948
Treatment			
Procedure (%. 95%CI)			
EVAR Completed	74.5 (73.2-75.8)	52.8 (47.7-57.9)	32.4 (29.4-35.4)
EVAR Converted	0.3 (0.1-0.5)	0.8 (-0.1 to 1.7)	1.4 (0.7-2.2)
Open	25.2 (23.9-26.5)	46.4 (41.3-51.5)	66.2 (63.2-69.2)
No. Procedures (%. 95%CI)			
>2	1.4 (1.1-1.7)	0.5 (-0.2 to 1.2)	2.1 (1.2-3.0)
Perioperative blood loss (%. 95%CI)			
≤ 100ml	22.0 (20.8-23.2)	13.7 (10.6-17.6)	7.1 (5.6-8.0)
101-500ml	23.8 (22.6-25.0)	24.3 (19.9-28.7)	12.1 (10.0-14.2)
501-999ml	6.1 (5.4-6.8)	12.1 (8.8-15.4)	4 (2.8-5.2)
≥ 1000ml	12.5 (11.5-13.5)	21 (16.9-25.2)	40.8 (37.7-43.9)
Missing	35.6 (34.2-37.0)	28.8 (24.5-33.7)	36 (33.0-39.1)
Peritoneal contamination (%. 95%CI)			
None	95.2 (94.6-95.8)	93 (90.4-95.6)	76.3 (73.6-79.0)
Fluid	0.5 (0.3-0.7)	3 (1.3-4.7)	5.3 (3.9-6.7)
Abscess	0	1.1 (0.0-2.2)	0.3 (-0.0 to 0.7)
Peritonitis	0.3 (0.1-0.5)	1.3 (0.2-2.5)	14.8 (12.5-17.1)
Missing	4 (3.4-4.6)	1.6 (0.3-2.9)	3.5 (2.3-4.7)

95% CI: p +- (1.96*(SQRT(p*(1 -p))/n)), where p = proportion, n = sample size, SQRT = square root. Data are % (95% CI).
EAAA = elective abdominal aortic aneurysm; SAAA = symptomatic abdominal aortic aneurysm; RAAA = ruptured abdominal aortic aneurysm; CI = confidence interval; EVAR = endovascular aneurysm repair.

Morbidity

Twenty-three percent (n = 1068) of patients with EAAA had a peri-operative and/or post-operative complication. Patients receiving EVAR had fewer complications than those undergoing OSR (16.1% vs. 44.8%). Almost 39% (n = 144) of the patients with SAAA had one or more peri- and/or post-operative complications versus 69.2% (n = 656) of the RAAA patients. Patients undergoing OSR had a higher percentage of complications than those undergoing EVAR (Tables 5 and 6).

In general, after OSR, there were more complications than after EVAR. Cardiopulmonary complications accounted for the most post-operative problems, especially with OSR.

Table 5. Outcome after AAA repair by procedure.

	EAAA		SAAA		RAAA	
	EVAR	OSR	EVAR	OSR	EVAR	OSR
Patients (n)	n=3 426	n=1 153	n=199	n=172	n=320	n=628
Outcome	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)
Mortality (in- hospital or <30-days)	0.9 (0.6-1.3)	5.0 (3.9-6.5)	5.0 (2.7-9.0)	10.5 (6.8-16.0)	22.2 (18.0-27.1)	32.0 (28.5-35.8)
Perioperative complications	4.1 (3.4-4.8)	6.5 (5.2-8.1)	5.5 (3.1-9.6)	9.9 (6.3-14.3)	13.8 (10.4-18.0)	21.8 (18.8-25.2)
Postoperative complications	12.5 (11.4-13.6)	42.9 (40.1-45.8)	27.6 (21.9-34.2)	44.2 (37.0-51.7)	49.5 (44.1-55.0)	72.6 (69.0-75.9)
Reinterventions	2.5 (2.0-3.1)	10.7 (9.0-12.6)	7.0 (4.2-13.9)	9.3 (5.8-14.6)	12.8 (9.6-16.9)	20.7 (17.7-24.0)
Hospitalstay >14 days	2.5	19.0	6.0	26.2	16.3	41.2
Hospitalstay >10 days	3.6	33.3	10.1	43.0	25.9	55.1
Hospitalstay >5 days	11.5	85.8	34.2	93.6	54.4	73.1
Hospital readmission[a]	6.2	7.1	11.3	9.1	10.3	4.5

95% CI: p+- (1.96*(SQRT(p*(1 -p))/n)), where p = proportion, n = sample size, SQRT = square root. EAAA = elective abdominal aortic aneurysm; SAAA = symptomatic abdominal aortic aneurysm; RAAA = ruptured abdominal aortic aneurysm; EVAR = endovascular aneurysm repair; OSR = open surgical repair; CI = confidence interval.
[a]Missing values excluded because not in short survey.

In OSR for EAAA 5.2% of the patients versus 0.2% in EVAR had renal failure; the majority of these patients (4.6% and 0.1%, respectively) were temporarily dialysed. In the AAAA group most patients had renal failure after RAAA and OSR (18.0%). Patients undergoing EVAR had the most unplanned occlusions of the hypogastric artery during RAAA surgery (3.2%).

Re-interventions occurred more frequently after OSR than after EVAR (EAAA 10.7% vs. 2.5%; SAAA 7.0% vs. 9.3%; RAAA 12.8% vs. 20.7%). A total of 88.5% of patients with EAAA, treated with EVAR, were discharged within 5 days, and 85.8% of patients undergoing OSR were discharged after >5 days; 19% of the patients undergoing OSR remained in hospital for >14 days. The majority of patients with RAAA and SAAA, treated by EVAR, were discharged within 14 days (12.3% remained in hospital), while 38% of the patients undergoing OSR remained in hospital for > 14 days.

The variable "readmission" was recorded in 3471 (75.8%) patients with EAAA and 994 (75.4%) patients with AAAA. Of those with EAAA, 6.5% were readmitted: 6.2% after EVAR and 7.1% after OSR. In the AAAA group the majority of patients were not

Table 6. Peri-and postoperative complications after AAA repair by procedure.

Procedure	EAAA				SAAA				RAAA			
	EVAR	n=3426	OSR	n=1153	EVAR	n=199	OSR	n=172	EVAR	n=320	OSR	n=628
	%. (95%CI)		%. (95%CI)		%. (95%CI)		%. (95%CI)		%. (95%CI)		%. (95%CI)	
One or more complications (peri- and postoperative)	16.1		44.8		30.7		48.3		52.5		77.7	30.7
CPR	0.1	(0.02-0.02)	0.4	(0.2-1.0)	0.5	(0.09-2.8)	0.6	(0.1-3.2)	5.0	(3.1-8.0)	8.0	(6.1-10.3)
unplanned closure of hypogastric artery	1.2	(0.9-1.6)	0.4	(0.2-1.0)	1.5	(0.5-4.3)	0.6	(0.1-3.2)	3.2	(1.7-5.7)	0.9	(0.4-2.0)
visceral injury perioperative	0	-	0.5	(0.2-1.1)	0	-	1.2	(0.3-4.1)	0.3	(0.06-1.8)	2.1	(1.2-3.5)
urethral damage	0	-	0.4	(0.2-1.0)	0	-	0.6	(0.1-3.2)	0	-	0.3	(0.09-1.2)
other perioperative	2.7	(2.2-3.3)	4.7	(3.6-6.1)	3.5	(1.7-7.1)	7.0	(4.0-11.8)	5.3	(3.3-8.3)	10.6	(8.4-13.2)
bleeding	1.0	(0.7-1.4)	2.7	(1.9-3.8)	2.5	(0.1-5.7)	1.2	(0.3-4.2)	2.5	(1.3-4.9)	4.9	(3.5-6.9)
colonic Ischaemia	0.2	(0.1-0.4)	3.7	(2.8-5.0)	1.0	(0.3-3.6)	2.9	(1.2-6.6)	5.3	(3.0-8.3)	9.4	(7.4-11.9)
arterial occlusion	1.7	(1.3-2.2)	4.0	(3.0-5.3)	1.0	(0.3-3.6)	5.8	(3.2-10.4)	4.4	(2.6-7.2)	5.4	(3.9-7.5)
paralysis	0	-	0	-	0	-	0	-	0	-	0	-
any prosthetic complications[a]	1.3	(1.0-1.7)	0.3	(0.1-0.8)	4.0	(2.0-7.7)	0.6	(0.01-3.3)	3.7	(2.1-6.4)	0.3	(0.08-1.1)
abscess	0	-	0.4	(0.2-1.0)	0.5	(0.09-2.8)	0.6	(0.01-3.3)	0	-	0.8	(0.3-1.9)
visceral injury postoperative	0	-	0.2	(0.06-0.7)	0	-	0.6	(0.01-3.3)	0.6	(0.2-2.2)	1.5	(0.7-2.5)
wound dehiscence	0	-	1.6	(1.0-2.5)	0	-	0.6	(0.01-3.3)	0	-	2.1	(1.2-3.6)
ileus	0.1	(0.04-0.3)	2.7	(1.9-3.8)	0	-	1.7	(0.6-4.9)	0.9	(0.3-2.7)	1.1	(0.5-2.3)
colostomy	0.1	(0.04-0.3)	3.0	(2.1-4.2)	0	-	1.7	(0.6-4.9)	3.1	(1.7-5.6)	9.2	(7.2-11.7)
major amputation	0	-	1.0	(0.5-1.8)	0.5	(0.09-2.8)	0.6	(0.01-3.3)	0.3	(0.05-1.7)	1.1	(0.5-2.3)
profound wound infection	0.8	(0.6-1.2)	0.7	(0.4-1.4)	0	-	0	-	0.6	(0.2-2.2)	0.2	(0.04-1.0)

Table 6. Peri-and postoperative complications after AAA repair by procedure. (*continued*)

Procedure	EAAA				SAAA				RAAA			
	EVAR	n=3426	OSR	n=1153	EVAR	n=199	OSR	n=172	EVAR	n=320	OSR	n=628
	%. (95%CI)		%. (95%CI)		%. (95%CI)		%. (95%CI)		%. (95%CI)		%. (95%CI)	
other surgical	4.5	(3.9-5.3)	14.6	(12.7-16.8)	11.1	(7.5-16.2)	16.3	(11.5-22.6)	13.4	(10.1-17.6)	23.1	(20.0-26.6)
cardiac	1.3	(1.0-1.7)	10.1	(8.5-12.0)	6.5	(3.8-10.8)	14.5	(10.0-20.5)	12.5	(9.3-16.6)	19.7	(16.8-23.0)
pulmonary	1.8	(1.4-2.3)	17.4	(15.3-19.7)	4.5	(2.4-8.4)	19.2	(14.0-25.7)	18.1	(14.3-22.7)	27.2	(23.9-30.8)
neurologic	0.6	(0.4-0.9)	3.6	(2.7-4.9)	1.5	(0.5-4.3)	3.5	(1.6-7.4)	3.4	(1.9-6.0)	7.3	(5.5-9.6)
thrombo-embolic	1.4	(1.1-1.9)	4.7	(3.6-6.1)	2.5	(0.1-5.7)	5.2	(2.8-9.6)	4.4	(2.6-7.2)	8.0	(6.1-10.4)
infections[b]	2.7	(2.2-3.3)	9.5	(7.9-11.3)	5.5	(3.2-9.6)	11.0	(7.2-16.6)	10.0	(7.2-13.8)	18.3	(15.5-21.5)
renal insufficiency	0.2	-	5.2	(4.1-6.6)	1.0	(2.8-3.6)	8.7	(5.4-13.9)	6.9	(4.6-10.2)	18.0	(15.2-26.9)

95% CI: p +- (1.96*(SQRT(p*(1 - p))/n)), where p = proportion, n = sample size, SQRT = square root. EAAA 1⁄4 elective abdominal aortic aneurysm; SAAA = symptomatic abdominal aortic aneurysm; RAAA = ruptured abdominal aortic aneurysm; EVAR = endovascular aneurysm repair; OSR = open surgical repair; CI = confidence interval; CPR = cardiopulmonary resuscitation.

[a] Migration, infection, any leakage. [b] Other than surgical or pulmonary.

readmitted to the hospital (92.5%). Readmissions occurred twice as often after EVAR as after OSR (10.7% vs. 5.5%).

Risk prediction V-POSSUM

Predicted or expected mortality for EAAA by V-POSSUM showed significant miscalibration with observed mortality (Hosmer-Lemeshow $p < .01$; as reported in Fig. 1A). The observed mean mortality for EVAR differed significantly from that predicted ($p < .01$): 0.9% (95% CI 0.6-1.3) and 3.5% (95% CI 2.9-4.1), respectively. Also, the mean predicted mortality for EVAR was lower than for OSR. Observed mortality after OSR was 5% and predicted by V-POSSUM to be 5.3% (95% CI 4.1-6.6; $p = .65$), as shown in Table 7. The overall p value calculated with the Fisher's combined probability test showed a significant difference in observed versus expected mortality ($p < .001$). The discriminative ability of V-POSSUM was moderate (C-statistic = .719).

Table 7 Observed- and predicted mortality (V-POSSUM) for AAA patients

Procedure	Setting	Observed	Predicted	Lower PI	Upper PI	P -value
EVAR	EAAA	0.88	3.52	2.91	4.12	< 0.01
OSR	EAAA	5.03	5.32	4.08	6.56	0.65
EVAR	SAAA	5.03	6.88	3.55	10.21	0.28
OSR	SAAA	10.47	9.06	5.00	13.13	0.50
EVAR	RAAA	22.19	21.39	17.44	25.34	0.70
OSR	RAAA	32.01	28.58	25.52	31.65	0.03

PI = prediction interval; EVAR = endovascular aneurysm repair; EAAA = elective abdominal aortic aneurysm; OSR = open surgical repair; SAAA = symptomatic abdominal aortic aneurysm; RAAA = ruptured abdominal aortic aneurysm.

The observed mortality for AAAA surgery by V-POSSUM showed significant miscalibration (Hosmer-Lemeshow $p < .01$) compared with the predicted mortality (Fig. 1B). The observed mortality for RAAA was 22.2% (EVAR) versus 32.0% (OSR) and for SAAA 5.0% (EVAR) versus 10.5% (OSR). As reported in Table 7, predicted mortality was 6.9% (95% CI 3.6-10.2) for EVAR and 9.1% (95% CI 5.0-13.1) for OSR in patients with SAAA, implying there were no significant differences between the observed and predicted percentages. However, the predicted mortality in RAAA was 21.4% (95% CI 17.4-25.3) for EVAR and 28.6% (95% CI 25.5-31.7) for OSR, which differed significantly from the observed mortality ($p = .03$). The overall p value by Fisher's combined probability test showed a non-significant difference in observed versus expected mortality ($p = .16$). The discriminative ability of V-POSSUM was moderate (C-statistic = .713).

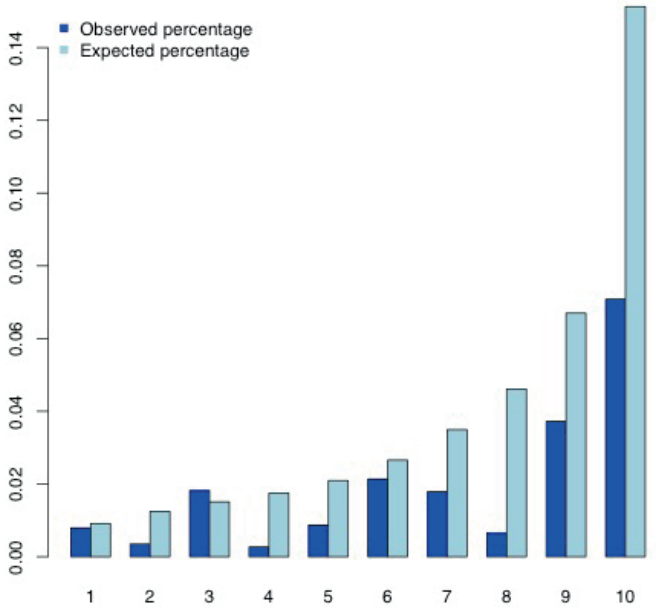


Figure 1. (A) The percentage observed mortality compared with the percentage expected mortality by V-POSSUM in deciles in elective AAA.

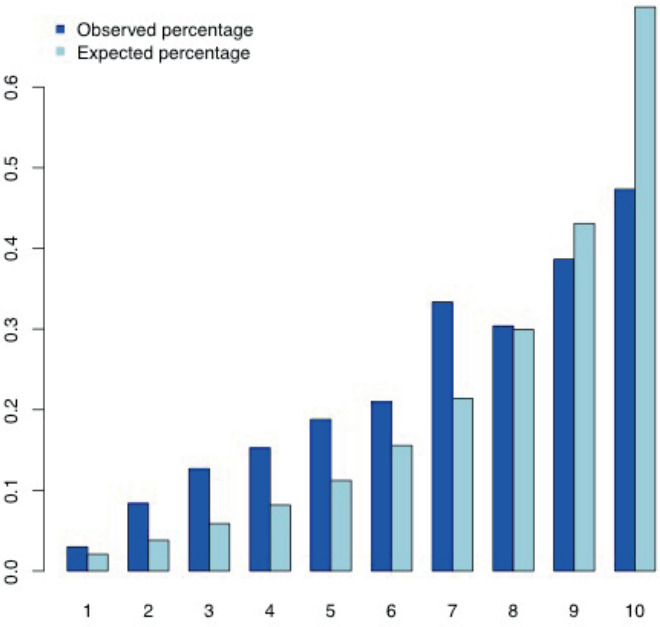


Figure 1. (B) The percentage observed mortality compared with the percentage expected mortality by V-POSSUM in deciles in acute AAA patients (e.g., symptomatic abdominal aortic aneurysm and ruptured abdominal aortic aneurysm patients).

Risk adjustment V(p)-POSSUM: hospital comparison

The V(p)-POSSUM showed a moderate discriminative ability of 0.665 in patients with EAAA and 0.688 in patients with AAAA. Unadjusted mortality is shown in Fig. 2A and B for patients with EAAA and AAAA, respectively.

In both EAAA and AAAA, mortality was low and there was no evidence of over or under-performance of certain centres. In Fig. 3A the EAAA DSAA population was compared with the reference population (i.e., UK) on which the V(p)- POSSUM was calibrated. A much lower mortality was seen in the DSAA population, especially in the EVAR group, as reported in Table 7. In Fig. 3B the AAAA DSAA population is compared with the reference population (i.e., UK) on which the V(p)-POSSUM was calibrated. There was a higher mortality in the DSAA population with respect to the reference population. Finally, in Fig. 4 (A, B) the risk adjusted comparison of all centres in the EAAA DSAA and AAAA DSAA is shown. While for patients with EAAA there is no under- or over performance, in AAAA there is no evidence of under performance for any centre either, except for one hospital, which showed a significantly better performance after multiple testing.

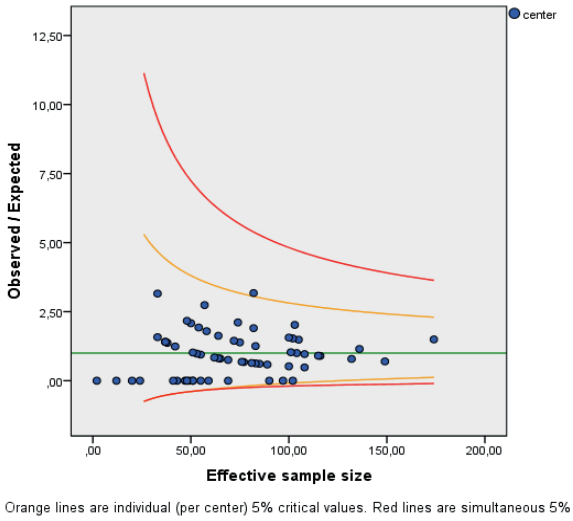


Figure 2. (A) The unadjusted standardised mortality ratio (SMR) (y-axis) of hospital mortality for elective abdominal aortic aneurysm patients. The x-axis describes the effective sample size, which takes into account the precision of the estimation of expected events, in this case the actual sample size. The expected number of events defined as the national average. The orange and red lines are 95% confidence intervals (CIs). The green line resembles 'SMR = 1' when observed divided by expected is the same.

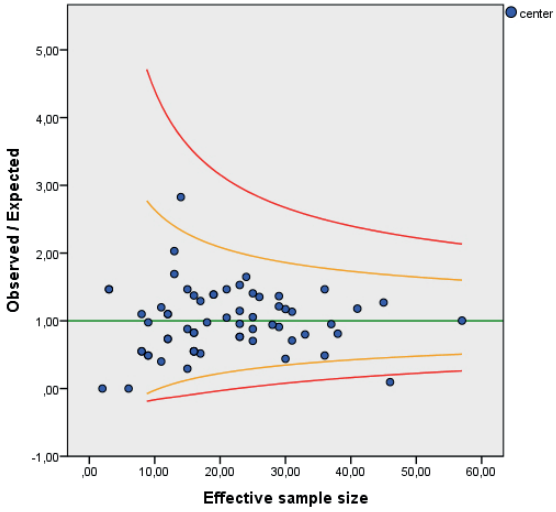


Figure 2. (B) The unadjusted SMR (y-axis) of hospital mortality for acute abdominal aortic aneurysm patients. The x-axis describes the effective sample size, which takes into account the precision of the estimation of expected events, in this case the actual sample size. The expected number of events defined as the national average. The orange and red lines are 95% CIs. The green line resembles 'SMR = 1' when observed divided by expected is the same.

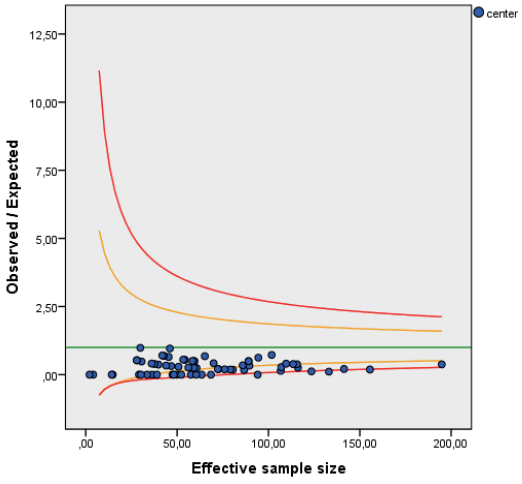


Figure 3. (A) The adjusted standardised mortality ratio (SMR) (y-axis) of hospital mortality for elective abdominal aortic aneurysm patients by V(p)-POSSUM benchmarked on the UK. The x-axis describes the effective sample size, which takes into account the precision of the estimation of expected events by V(p)-POSSUM. The expected numbers of patients are calculated by hospital based on the variables included in V(p)-POSSUM. The green line resembles 'SMR = 1' when observed divided by expected is the same.

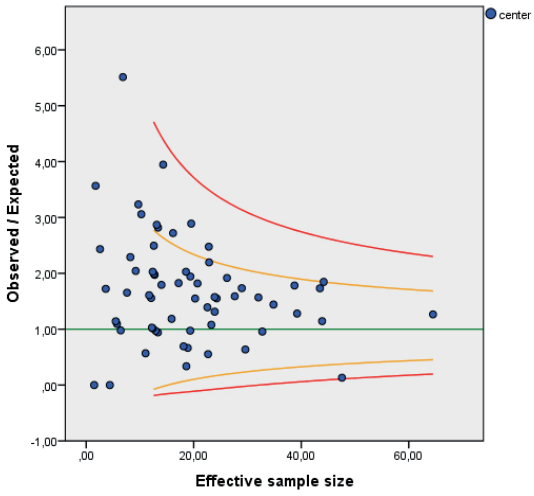


Figure 3. (B) The adjusted SMR (y- axis) of hospital mortality for acute abdominal aortic aneurysm patients by V(p)-POSSUM benchmarked on the UK. The x-axis describes the effective sample size, which takes into account the precision of the estimation of expected events by V(p)-POSSUM. The expected numbers of patients are calculated by hospital based on the variables included in V(p)-POSSUM. The orange and red lines are both 95% confidence intervals. The green line resembles 'SMR = 1' when observed divided by expected is the same.

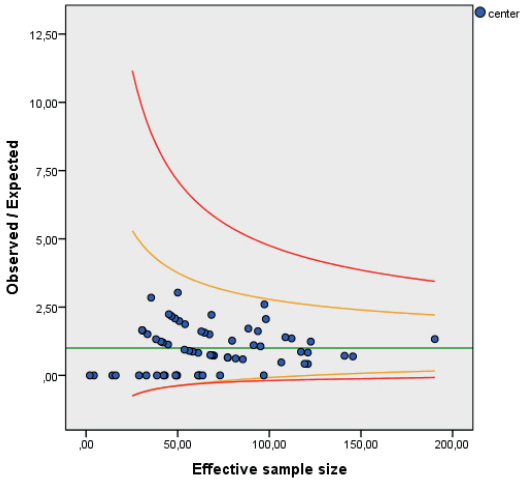


Figure 4. (A) The adjusted standardised mortality ratio (SMR) (y-axis) of hospital mortality for elective abdominal aortic aneurysm patients by V(p)-POSSUM re-estimated on the Dutch Surgical Aneurysm Audit (DSAA). The x-axis describes the effective sample size, which takes into account the precision of the estimation of expected events. The orange and red lines are both 95% confidence intervals (CIs). The green line resembles 'SMR = 1' when observed divided by expected is the same.

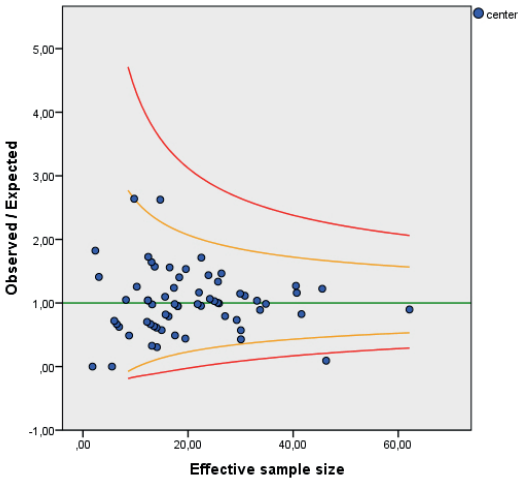


Figure 4. (B) The adjusted standardised mortality ratio (SMR) (y- axis) of hospital mortality for acute abdominal aortic aneurysm patients by V(p)-POSSUM re-estimated on the DSAA. The x-axis describes the effective sample size, which takes into account the precision of the estimation of expected events. The orange and red lines are both 95% CIs. The green line resembles 'SMR = 1' when observed divided by expected is the same.

DISCUSSION

The 30 day or in hospital mortality of 1.9% for elective AAA surgery in the Dutch Surgical Aneurysm Audit is comparable with other European registries. For example, the Swedish and UK elective populations reported mortality percentages of 1.5% and 2.4%, respectively.[21,23,28] With the Dutch mandatory minimum volume of 20 AAA operations per year per centre set by the Dutch Healthcare Inspectorate, mortality was not a discriminative outcome parameter between hospitals in the DSAA, as almost all unadjusted and adjusted observations were within the 95% CI. Patients with SAAA appear to be very different from those with EAAA, indicated by the mortality rate of 7.5%. The international reported mortality rate for acute symptomatic, non-ruptured aneurysms ranges between 11% and 18%.[29] Mortality after RAAA surgery in the DSAA was also comparable with mortality after RAAA in the Swedvasc (18% after EVAR, 32% after OSR in 2013).[30] The mortality after EAAA EVAR was lower than after OSR in the DSAA and comparable with the UK data.[28] However, mortality after OSR in the DSAA (5%) compared less favourable with other registries, such as Swedvasc, which reported 3.2% mortality after OSR in their yearly report.[30] Patients undergoing OSR had a higher predicted mortality than those undergoing EVAR, which might be an indication of more comorbidities and also of more peri-operative blood loss.

The mortality after EVAR in patients with RAAA was lower compared to OSR, while most RAAA patients were treated with OSR. The mortality differences between OSR and EVAR, as in other observational studies, indicate that selection bias (i.e., different case mix) and a weighed choice of treatment could be responsible for this observation. The lower predicted mortality in patients undergoing EVAR compared with those undergoing OSR might indicate that patients undergoing EVAR had fewer comorbidities, less perioperative blood loss, or both. So, when comparing the results after EVAR and OSR there is at least some selection bias.[16] Conclusions about whether EVAR is a better operative technique cannot be made from this analysis.

The V-POSSUM is one of the most frequently validated mortality risk prediction models in the literature. Because all V-POSSUM variables were implemented in the DSAA, mortality risk adjustment by V(p)-POSSUM, containing only the pre-operative variables, could be performed easily. Risk adjustment of outcomes in the DSAA by, for example, V(p)- POSSUM, in order to compare hospital performances is not performed by other registries, such as Swedvasc. They do not risk adjust their yearly outcomes by case mix, which makes comparisons between registries difficult. Interestingly, in the DSAA, risk adjustment by V(p)-POSSUM for EAAA did not influence hospital variation, even after re-estimation on the Dutch population. This might be caused by the relatively low event rate of the outcome "mortality". Perhaps compound measurements can be the key when comparing hospital outcomes. Examples are "failure to rescue", the number of patients that die as a result of complications, and "textbook outcome", the ideal healthcare pathway for every patient.[31,32] Risk adjustment for AAAA did change the position on the y-axis of every hospital, showing the effect of differences in case mix on mortality and the necessity for risk adjustment.

Missing data

Missing data are a well- known and common problem in registries.[2] To maintain data quality, there are several ways of dealing with missing data. It is possible for instance, to exclude patients that miss relevant data, to choose imputation of the mean or use multiple imputation.[27,33] Although missing values are an unwanted outcome, the effect on hospital outliers is only relevant in low volume hospitals.[34] Missing data in the DSAA were scarce and exceeded the 20% for leukocytes in EAAA, which may be a non-routinely measured variable in patients who undergo AAA surgery. For peri-operative blood loss, data were missing in >25% in every setting. Therefore, the percentage of missing values could indicate poor administrative performance, and a decrease in the number of missing values might therefore be used as a quality indicator when comparing hospitals.

Clinical outcomes

It has been suggested that only specialised centres with appropriate expertise should perform EVAR. However, there is no significant variation in the outcome of EAAA between hospitals in the DSAA. Furthermore, there was no relationship between the percentage of EVAR performed and hospital volume in the DSAA, as well as no association between hospital volume (minimum volume of 20 patients per year) and outcome mortality in both EAAA and AAAA. In the DSAA almost three quarters of the patients with EAAA were treated by EVAR. There is no reason to concentrate on EVAR for EAAA in the Netherlands in the current setting.

However, the volume per centre for primary elective surgery in the Netherlands is 20 to more than 100 procedures per year, indicating a volume of five to more than 20 OSRs per hospital. This number could be challenging for many hospitals, as several studies have proposed a minimum of 3-12 elective OSRs per surgeon per year, or at least 7-30 elective OSRs per hospital per year.[35-41] Moreover, as hospital experience in one procedure does not translate into expertise in the other, it is necessary to retain experience in both.[38] Potential bias in the outcome of the DSAA can also be caused by the selection by indication for operation dependent on patient or disease characteristics (aneurysm diameter, restriction to patients with comorbidities), and the concomitant choice for a certain operative technique (OSR or EVAR preference, or even fenestrated EVAR or chimneys). The choice of operative procedure influences mortality and depends on patient characteristics as well as on surgeon's preference. Therefore, operative variables cannot be used for casemix adjustment, because a correction for surgical skills is undesirable. Unfortunately, correction for this kind of bias is not possible. However, the overall mortality rate of 5% for OSR for EAAA is a matter of concern. The differences in outcome between EVAR and OSR for AAAA in the DSAA can be biased by "selection by indication" for surgery. Because, results are influenced by patient or disease characteristics (aneurysm diameter, restriction to patients with comorbidities), and the concomitant choice for a certain operative technique (liberal use of EVAR or conservative choice for OSR). Patients receiving EVAR in the DSAA seem to have less comorbidity. Identifying the best operative technique for the individual patient remains a challenge. Vascular units face the challenge of choosing the surgical technique while at the same time retaining experience in both open and endovascular techniques.[42]

Risk prediction V-POSSUM

Mortality risk prediction models like V-POSSUM aim to predict mortality for an individual patient. Ideally, a model is discriminative and calibrates well. Because discrimination and calibration are reversely dependent, this will never be the case.[26]

The observed miscalibration of V-POSSUM can be a sign of overfitting, which can be explained by several factors: the presence of too many variables compared with the number of events, the statistical procedure used for selection of the variables (e.g., forward or backward selection, or high p value for inclusion), the number of categories used per variable, the handling of missing data, and the degree to which a population differs from the original population in severity.[26] The significant miscalibration between observed and expected mortality after EAAA EVAR can be explained, in part, by the fact that V-POSSUM was developed before the introduction of EVAR.[43] However, in patients with AAAA mortality was underestimated for those undergoing OSR, but still higher compared with EVAR.

The discriminative ability was moderate in the DSAA, still resulting in false predictions compared with the observed outcomes. This might imply that there are variables lacking in the model that could lead to better predictions.[26] Moreover, according to the instructions of V-POSSUM, EVAR was scored in the same operative severity category as OSR (major surgery, 4 points as exponential in the regression coefficient). It is questionable whether EVAR has to be marked as major vascular surgery.

Risk adjustment V(p)-POSSUM: hospital comparison

Risk adjusted hospital outcomes are a prerequisite for meaningful hospital comparisons. Adjusting mortality in the DSAA with V(p)-POSSUM provides the effect of risk adjustment by case mix according to the population (UK) in which V(p)-POSSUM was developed. Therefore, and because it was built on an overall aneurysm population and on top of that the continuous predominance of EVAR procedures, the V(p)-POSSUM was re-estimated for the Dutch population by logistic regression. The POSSUM physiology- only models can be a useful tool for comparative outcome audits.[9] However, it might have become necessary to include more EVAR and outcome specific variables or to re-estimate the variables included on a mixed (EVAR and OSR) population. The POSSUM physiology-only models contain a significant number of variables compared with other pre-operative mortality-risk prediction models such as the GAS and VBHOM.[7,8] These latter models might be more suitable and easier to use. Suitability for clinical practice not only depends on the number of variables, but also on the administrative burden in clinical practice. An ideal model should contain clear and distinct variables, be suitable for both acute and elective surgery at a definite endpoint, and have well defined categories. Although, hospital mortality changed owing to the effect of casemix adjustment, it was still not possible to recognise underperforming hospitals. Most hospitals, except one, remained within the CIs.

Limitations

When registering data, coding and documentation errors (internal validity), or errors in the external validity of the data, occur. As the registry started in 2013 there were fewer patients than in 2014. This could have been the result of under-registration. However, a crude check of mortality between the two years revealed no differences in mortality in elective AAA or a registered lower mortality for acute AAA in 2014. As the data are not yet validated and hospitals were not audited for data verification, the results presented in the current overview should be interpreted with care. The presence of missing data does not necessarily indicate that a comparison between hospitals is unreliable providing the volume of AAA repair is large enough and comparable.[34] In the Netherlands external validation is difficult because all AAA operations, including revisions and suprarenal AAA surgery, are registered nationally in the national hospital statistics with the same code as primary AAA surgery. Visits to hospitals in order to validate the registered data will be the next best step in the verification process.

It was not possible to differentiate between referral centres and non-referral centres in the current DSAA, as there was no definition for referral centre for highly complex cases. The option of registering the referral of a patient was recently added to the updated dataset of the DSAA. Referral centres potentially have more complex aneurysm morphology with a greater risk of proximal aneurysm neck related complications and increased mortality.[44] Reported mortality for complex aneurysms is higher than average AAA, but published results for endovascular repair of difficult aortic necks look promising.[45]

CONCLUSION

Nearly all patients registered in the Dutch Surgical Aneurysm audit could be included for analysis. Operative mortality, adjusted and non-adjusted, after EAAA surgery was not a discriminative outcome parameter for hospital comparisons in the DSAA. The overall post-operative (EVAR and OSR) and, specifically, EVAR related mortality was low and there was no significant association between hospital volume and (risk adjusted) percentage of EVAR performed. Therefore, the Dutch minimum volume of 20 EAAA procedures appears to be sufficient for EVAR. However, the overall mortality after OSR was relatively high, resulting in concerns with regard to this low volume operation in the era of preference for EVAR. Also in patients with AAAA, the observed mortality of OSR for RAAA was significantly higher than the predicted mortality. Patients undergoing EVAR have a lower mortality, but this can be at least partly explained by the lower predicted mortality by V-POSSUM, indicating patient selection. In this

study, risk adjusted mortality for elective AAA surgery has limited capability for hospital comparison quality assessment.

ACKNOWLEDGEMENTS

The authors acknowledge V.C.E. van der Slot and van der Willik for their help in cross-checking the performed analysis of the data. The authors would like to thank the Dutch Society for Vascular Surgery, which registered their patients in the DSAA; the Dutch Institute for Clinical Auditing, which facilitated the registry; and the Dutch Surgical Aneurysm Audit group, which consists of the following members: Dr. B.H.P. Elsman, Prof. Dr. J.F. Hamming, Prof. Dr. D.A. Legemate, Dr. R.H. Geelkerken, Dr. E.D. Wilschut, and Dr. A.C. Vahl.

Conflicts of interest: None.

Funding: None.

REFERENCES

1 Van Leersum NJ, Snijders HS, Henneman D, Kolfschoten NE, Gooiker GA, ten Berge MG, et al. The Dutch surgical colorectal audit. Eur J Surg Oncol 2013;39:1063-70.

2 Atkins ER, Boyle JR. Accurate clinical data is vital in the age of surgeon level outcome reporting. Eur J Vasc Endovasc Surg 2015;50:809.

3 Dekker JW, Gooiker GA, van der Geest LG, Kolfschoten NE, Struikmans H, Putter H, et al. Use of different comorbidity scores for risk-adjustment in the evaluation of quality of colorectal cancer surgery: does it matter? Eur J Surg Oncol 2012;38:1071-8.

4 Meguid RA, Bronsert MR Juarez-Colunga E, Hammermeister KE, Henderson WG. Surgical risk preoperative assessment system (SURPAS): III. accurate preoperative prediction of 8 adverse outcomes using 8 predictor variables. Ann Surg 2016 Jun;263(6): 1042-8.

5 Fischer C, Lingsma H, Hardwick R, Cromwell DA, Steyerberg E, Groene O. Risk adjustment models for short-term outcomes after surgical resection for oesophagogastric cancer. Br J Surg 2016;103:105-16.

6 Kolfschoten NE, Marang van de Mheen PJ, Gooiker GA, Eddes EH, Kievit J, Tollenaar RA, et al. Variation in case-mix between hospitals treating colorectal cancer patients in The Netherlands. Eur J Surg Oncol 2011;37:956-63.

7 Samy AK, Murray G, MacBain G. Glasgow aneurysm score. Cardiovasc Surg 1994;2:41e4.

8 Tang T, Walsh SH, Prytherch DR, Lees T, Varty K, Boyle A. Audit Research Committee of the Vascular Society of Great Britain and Ireland. VBHOM, a data economic model for predicting the outcome after open abdominal aortic aneurysm surgery. Br J Surg 2007;94:717-21.

9 Prytherch DR, Ridler BM, Beard JD, Earnshaw JJ. Audit and Research Committee, The Vascular Surgical Society of Great Britain and Ireland. A model for national outcome audit in vascular surgery. Eur J Vasc Endovasc Surg 2001;21:477-83.

10 Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. Br J Surg 2001;88:958-63.

11 Pini R, Faggioli G, Longhi M, Mauro R, Freyrie A, Gargiulo M, et al. The influence of study design on the evaluation of ruptured abdominal aortic aneurysm treatment. Ann Vasc Surg 2014;28:1568-80.

12 Sweeting MJ, Ulug P, Powell JT, Desgranges P, Balm R. Ruptured Aneurysm Trialists. Ruptured aneurysm trials: the importance of longer-term outcomes and meta-analysis for 1-year mortality. Eur J Vasc Endovasc Surg 2015;50:297-302.

13 Hoornweg LL, Wisselink W, Vahl A, Balm R. Amsterdam Acute Aneurysm Trial Collaborators. The Amsterdam Acute Aneurysm Trial: suitability and application rate for endovascular repair of ruptured abdominal aortic aneurysms. Eur J Vasc Endovasc Surg 2007;33:679-83.

14 IMPROVE Trial Investigators, Powell JT, Sweeting MJ, Thompson MM, Ashleigh R, Bell R, et al. Endovascular or open repair strategy for ruptured abdominal aortic aneurysm: 30 day outcomes from IMPROVE randomised trial. BMJ 2014;348:f7661.

15 Desgranges P, Kobeiter H, Castier Y, Sénéchal M, Majewski M, Krimi A. The endovasculaire vs chirurgie dans les anevrysmes rompus PROTOCOL trial update. J Vasc Surg 2010;51:267-70.

16 Gunnarsson K, Wanhainen A, Djavani Gidlund K, Björk M, Mani K. Endovascular versus open repair as primary strategy for ruptured abdominal aortic aneurysm: a national population- based study. Eur J Vasc Endovasc Surg 2016;51:22-8.

17 Mastracci TM, Garrido-Olivares L, Cinà CS, Clase CM. Endovascular repair of ruptured abdominal aortic aneurysms: a systematic review and meta-analysis. J Vasc Surg 2008;47:214-21.
18 Harkin DW, Dillon M, Blair PH, Ellis PK, Kee F. Endovascular ruptured abdominal aortic aneurysm repair (EVRAR): a systematic review. Eur J Vasc Endovasc Surg 2007;34:673-81.
19 IMPROVE Trial Investigators. The effect of aortic morphology on peri-operative mortality of ruptured abdominal aortic aneurysm. Eur Heart J 2015;36:1328-34.
20 Karthikesalingam A, Holt PJ, Vidal-Diez A, Ozdemir BA, Poloniecki JD, Hinchliffe RJ, et al. Mortality from ruptured abdominal aortic aneurysms: clinical lessons from a comparison of outcomes in England and the USA. Lancet 2014;383:963-9.
21 Tröeng T. The Swedish National Registry for Vascular Surgery; Swedvasc. Available at: http://www.ucr.uu.se/swedvasc/arsrapporter/swedvasc-2011-english-version/download.
22 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 vascunet report. Eur J Vasc Endovasc Surg 2011;42:598-607.
23 The Dutch Surgical Aneurysm Audit. Available at: www.clinicalaudit.nl [last accessed 16 January 2017].
24 Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg 1991;78:355-60.
25 Hadjianastassiou VG, Tekkis PP, Athanasiou T, Young JD, Hands LJ. Comparison of mortality prediction models after open abdominal aortic aneurysm repair. Eur J Vasc Endovasc Surg 2007;33:536-43.
26 Steyerberg E. Clinical prediction models; a practical approach to development, validation, and updating. 1st ed. Springer; 2009.
27 Grant SW, Grayson AD, Mitchell DC, McCollum CN. Evaluation of five risk prediction models for elective abdominal aortic aneurysm repair using the UK national vascular database. Br J Surg 2012;99:673-9.
28 The Vascular Society of Great Britain and Ireland. Outcomes after elective repair of infra-renal abdominal aortic aneurysm, a report from the Vascular Society. Available at: https://www. vascularsociety.org.uk/_userfiles/pages/files/Document% 20Library/VSGBI-Mortality-Report.pdf [last accessed 12 November 2012].
29 Antonello M, Lepidi S, Kechagias A, Frigatti P, Tripepi A, Biancari F, et al. Glasgow aneurysm score predicts the outcome after emergency open repair of symptomatic, unruptured abdominal aortic aneurysms. Eur J Vasc Endovasc Surg 2007;33:272-6.
30 Kragsterman B. Nationella kvalitetsregistret för kärlkirurgi. Available at: http://www.ucr.uu.se/swedvasc/arsrapporter/ swedvasc-2014/download.
31 Henneman D, Snijders HS, Fiocco M, van Leersum NJ, Kolfschoten NE, Wiggers T, et al. Hospital variation in failure to rescue after colorectal cancer surgery: results of the Dutch Surgical Colorectal Audit. Ann Surg Oncol 2013;20:2117-23.
32 Kolfschoten NE, Kievit J, Gooiker GA, van Leersum NJ, Snijders HS, Eddes EH, et al. Focusing on desired outcomes of care after colon cancer resections; hospital variations in 'text- book outcome'. Eur J Surg Oncol 2013;39:156-63.
33 Ambler GK, Gohel MS, Mitchell DC, Loftus IM, Boyle JR. Audit and Quality Improvement Committee of the Vascular Society of Great Britain and Ireland. The abdominal aortic aneurysm statistically corrected operative risk evaluation (AAA SCORE) for predicting mortality after open and endovascular interventions. J Vasc Surg 2015;61:35-43.

34 Holt PJ, Poloniecki JD, Thompson MM. Multicentre study of the quality of a large administrative data set and implications for comparing death rates. Br J Surg 2012;99:58-65.
35 Holt PJ, Poloniecki JD, Khalid U, Hinchliffe RJ, Loftus IM, Thompson MM. Effect of endovascular aneurysm repair on the volume-outcome relationship in aneurysm repair. Circ Cardiovasc Qual Outcomes 2009;2:624-32.
36 McPhee JT, Robinson III WP, Eslami MH, Arous EJ, Messina LM, Schanzer A. Surgeon case volume, not institution case volume, is the primary determinant of in-hospital mortality after elective open abdominal aortic aneurysm repair. J Vasc Surg 2011;53:591-9.
37 Dua A, Furlough CL, Ray H, Sharma S, Upchurch GR, Desai SS. The effect of hospital factors on mortality rates after abdominal aortic aneurysm repair. J Vasc Surg 2014;60:1446-51.
38 Landon BE, O'Malley AJ, Giles K, Cotterill P, Schermerhorn ML. Volume-outcome relationships and abdominal aortic aneurysm repair. Circulation 2010;122:1290-7.
39 Ilonzo N, Egorova NN, McKinsey JF, Nowygrod R. Failure to rescue trends in elective abdominal aortic aneurysm repair between 1995 and 2011. J Vasc Surg 2014;60:1473-80.
40 Chen CK, Chang HT, Chen YC, Chen TJ, Chen IM, Shih CC. Surgeon elective abdominal aortic aneurysm repair volume and outcomes of ruptured abdominal aortic aneurysm repair: a 12- year nationwide study. World J Surg 2013;37:2360-71.
41 Young EL, Holt PJ, Poloniecki JD, Loftus IM, Thompson MM. Meta- analysis and systematic review of the relationship between surgeon annual caseload and mortality for elective open abdominal aortic aneurysm repairs. J Vasc Surg 2007;46:1287-94.
42 Haveman JW, Karliczek A, Verhoeven EL, Tielliu IF, de Vos R, Zwaveling JH. Results of streamlined regional ambulance transport and subsequent treatment of acute abdominal aortic aneurysms. Emerg Med J 2006;23:807-10.
43 Nelissen BG, Herwaarden JA, Pasterkamp G, Moll FL, Vaartjes I. Shifting abdominal aortic aneurysm mortality trends in The Netherlands. J Vasc Surg 2015;61:642-647.e2.
44 Bastos Goncalves F, Hoeks SE, Teijink JA, Moll FL, Castro JA, Stolker RJ, et al. Risk factors for proximal neck complications after endovascular aneurysm repair using the endurant stentgraft. Eur J Vasc Endovasc Surg 2015;49:156-62.
45 Bastos Goncalves F, de Vries JP, van Keulen JW, Dekker H, Moll FL, van Herwaarden JA, et al. Severe proximal aneurysm neck angulation: early results using the endurant stentgraft system. Eur J Vasc Endovasc Surg 2011;41:193-200.

3

Systematic review of mortality risk prediction models in the era of endovascular abdominal aortic aneurysm surgery

N. Lijftogt
T. W. F. Luijnenburg
A. C. Vahl
E. D. Wilschut
V. J. Leijdekkers
M. F. Fiocco
M. W. J. M. Wouters
J. F. Hamming

British Journal of Surgery 2017

ABSTRACT

Background:

The introduction of endovascular aneurysm repair (EVAR) has reduced perioperative mortality after abdominal aortic aneurysm (AAA) surgery. The objective of this systematic review was to assess existing mortality risk prediction models, and identify which are most useful for patients undergoing AAA repair by either EVAR or open surgical repair.

Material and Methods:

A systematic search of the literature was conducted for perioperative mortality risk prediction models for patients with AAA published since 2006. PRISMA guidelines were used; quality was appraised, and data were extracted and interpreted following the CHARMS guidelines.

Results:

Some 3903 studies were identified, of which 27 were selected. A total of 13 risk prediction models have been developed and directly validated. Most models were based on a UK or US population. The best performing models regarding both applicability and discrimination were the perioperative British Aneurysm Repair score (C-statistic 0.83) and the preoperative Vascular Biochemistry and Haematology Outcome Model (C-statistic 0.85), but both lacked substantial external validation.

Conclusion:

Mortality risk prediction in AAA surgery has been modelled extensively, but many of these models are weak methodologically and have highly variable performance across different populations. New models are unlikely to be helpful; instead casemix correction should be modelled and adapted to the population of interest using the relevant mortality predictors.

INTRODUCTION

Abdominal aortic aneurysm (AAA) surgery is a common vascular operation performed in both elective and emergency situations that is intended to prevent AAA-related death.[1–3]

The 30-day or in-hospital mortality of patients with AAA is an important outcome measure reported in studies and by national registries.[4] Perioperative mortality is dependent on many factors, including the operative technique – either endovascular aneurysm repair (EVAR) or open surgical repair (OSR).[5]

A number of risk prediction scores were developed in the era of OSR, with the intention of aiding the comparison of outcomes. These have been used widely to provide risk-adjusted hospital and surgeon outcomes, and even to improve risk prediction for individual patients undergoing AAA surgery.[6] However, the recent widespread introduction of EVAR has resulted in a marked reduction in perioperative mortality in many countries.[4,7] Consequently, new prediction models have been developed. It is therefore important to determine which models are most appropriate for use in contemporary clinical practice.

The objective of the present study was to perform a systematic review of mortality risk prediction models used in AAA surgery since the implementation of EVAR for elective and emergency AAA.

MATERIAL AND METHODS

A systematic research was performed using the PRISMA guidelines.[8]

Literature search

A systematic search was performed on 6 July 2015 in PubMed, Embase and the Cochrane Library, by two researchers and an academic librarian. A limit to publication dates from 2006 to the current date in 2015 was applied to include only the most contemporary risk prediction scores. Relevant articles were identified by title and abstract by two reviewers. Reference lists of the relevant articles were snowballed, and cited articles were checked using Web of Science to identify additional articles. Another reviewer joined for further consensus over the studies selected. The full search strategies are shown in Appendix S1 (supporting information).

Study selection

Articles were excluded when they selected suprarenal or thoracoabdominal aortic aneurysms or where the population included only specific subgroups of patients, such as those with a small aneurysm, octogenarians or those referred to an ICU after surgery. The primary outcome measure was mortality. Studies that investigated only the predictive ability of observed versus expected mortality for prediction models, and provided no information about model performance, were excluded. Reviews, editorials and conference abstracts were not included. Papers in languages other than English or Dutch were also excluded.

Data selection

Results were sorted in a data extraction table and appraised for a representative population, outcome and type of study in accordance with the CHARMS guidelines for prognostic systematic reviews.[9] The definitions used for mortality made a distinction between short-term mortality (30-day, in-hospital, early, perioperative, at discharge) and long-term mortality (60-day, 90-day, long term, survival analysis). Two study types were included: prediction model development studies with internal 'non-apparent' validation in 'independent' data, meaning that model performance was evaluated in a subset of data by two-sampling testing techniques, bootstrap (multiple samples testing withdrawn from the same population), cross-validation or other techniques; and external model validation studies for models developed since 2006, with or without model updating, reporting performance by either a C-statistic and/or calibration test (see model performance below).[9,10] Studies not included for data extraction were: external validation studies containing only models developed before 2006, prediction model development studies without internal validation or with internal 'apparent' validation, and prediction studies exploring which predictors independently contribute to the prediction of a particular prognostic or diagnostic outcome.[9]Articles were appraised for model performance and type of statistical analysis performed.

Data extraction

Outcomes were extracted using a preformatted ExcelTM (Microsoft, Redmond, Washington, USA) spreadsheet using the CHARMS checklist.[9]

Model performance

The developed models were assessed for the presence of randomization to create a modelling and validation data set. The selection of variables was evaluated for forward or backward selection, or univariable or multivariable analysis, and criteria used (P value) for selecting variables.

Adjustment of the developed or existing model was evaluated either by the description of shrinkage for internal validation or by the description of updating for external validation. This is defined by adjustment of the coefficients and/or the intercept to reduce optimism, and/or recalibration of the model (observed versus expected outcome). Model performance was defined by discrimination, a measure of how well the model can separate those who do and those who do not have the disease of interest, and calibration, a measure of how well predicted probabilities agree with the actual observed risk. The discrimination 'C-statistic' (balance between negative and positive predictive value) was defined as low or poor (below 0.70), moderate (0.70 – 0.79) or good (at least 0.80). The correlation between observed and expected (calibration) outcome was measured by the Hosmer – Lemeshow (H-L) test ($P > 0.050$ indicates good fit), or assessed using a $\chi 2$ test. The number of events per variable (EPV; preferably more than 10) was also calculated if possible to give insight into the magnitude of overfitting (especially the case when more variables than events are present).

RESULTS

Some 27 studies were included in the final analysis (Fig. 1). Studies excluded based on critical appraisal for type of prognostic study and analytical performance are summarised in Table S1 (supporting information).

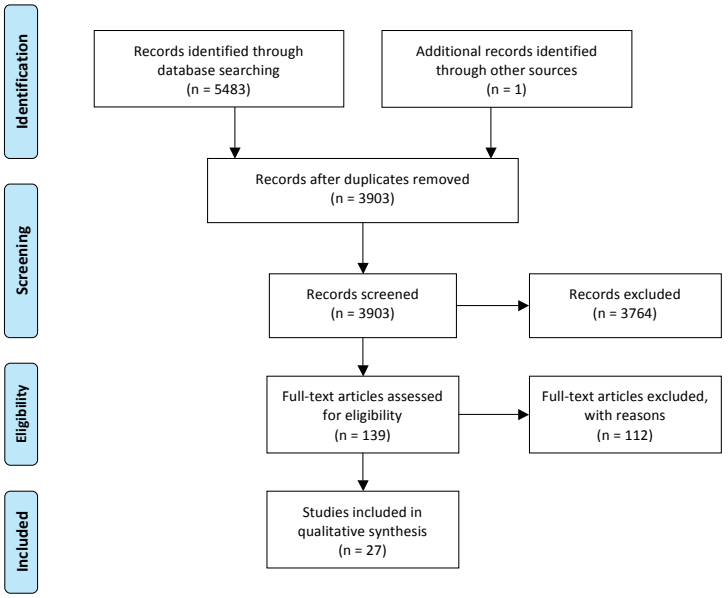


Figure 1. Flow chart showing selection of studies for review

A total of 13 studies developed their own mortality risk prediction model, of which 12[11 – 22] described internal non-apparent validation. Seven[11,13 – 16,20,22] of these performed validation by two-sample testing and five[12,17 – 19,21] used a bootstrap method for internal validation. One study[23] used artificial neural network (ANN) methodology, with a computational model to predict a model by pattern recognition and continually adapting to new input data, resulting in a multilayer model of several variables (Table 1). Two studies[18,22] performed geographical external validation. The remaining studies[24-37] externally validated one or more existing models, which consisted of elective and emergency ruptured or non-ruptured AAA surgery, either in the group as a whole or separately (Table 2).

Data extraction

Many model development studies were performed in the UK (6)[11,13,16,17,20,22] or USA (5).[14,15,18,19,23] Two UK studies[20,22] used nearly the same Cambridge population for the development and internal validation of a new Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (Cambridge POSSUM) and for validation of the Vascular Biochemistry and Haematology Outcome Model (VBHOM). Another study[11] used the UK National Vascular Database (NVD) for the development and internal validation of the Abdominal Aortic Aneurysm Statistically Corrected Operative Risk Evaluation (AAA SCORE).

There was overlap between data sets used for either development or validation studies.[14,15,20,22,26,31,33] One of these[22] was used for external or geographical validation after building a model on another population. One external validation study[34] derived its data from the Dutch Randomised Endovascular Aneurysm Management (DREAM) trial, an RCT recently performed in Belgium and the Netherlands.

Sample size varied between the studies, ranging from 21 patients in a validation data set[23] to 22 830 in both a modelling and validation data set[15], and 44 630 in a modelling set.[14] The number of events and mortality rates by setting (emergency versus elective) and procedure also varied widely.

Studies differed in their selection of patients and the percentages of procedures performed. One study[11] included patients with symptomatic AAAs. Some studies[11,12,22] included patients who also received an infrainguinal bypass, had an iliac aneurysm or underwent revisional surgery.

Table 1 Data extracted from studies that developed a new model since 2006

Reference	Follow-up	Model	n	Randomi-zation	Cohort	Setting	Procedure	Study type*	Primary outcome	Statistics
Ambler et al.[11]	2010–2011	Preop. AAA SCORE†	8086	+/–‡	P: NVD, UK	EAAA SAAA RAAA Redo	Both	A/B	In-hospital mortality	Schwartz– Bayes criterion/ AIC Net reclassification index
		Periop. AAA SCORE†								
		VBHOM								
		VGNW								
		Medicare								
		POSSUM								
		(p)-POSSUM								
Barnes et al.[12]	1999–2001	ERA	938 Bootstrap 100 ×	n.a.	P: Australian audit	EAAA SAAA	EVAR	A/B	30-day mortality	FW AIC
Choke et al.[13]	2000–2010	Leicester score†	1153	+	P: NHS Trust, UK	EAAA	Both, (51.7% EVAR)	A/B	30-day/in-hospital mortality	Univariable FW $P < 0.100$ BW $P < 0.050$
		Leicester score (OSR)†	564	+			OSR			
		GAS								
		CPI m-CPI								
		VGNW								
		Medicare								
Egorova et al.[14]	2000–2006	Mount Sinai score†	44 630	+	P: Medicare, USA	EAAA	EVAR	A	30-day mortality	Univariable $P < 0.250$ Multivariable $P < 0.050$

Table 1 Data extracted from studies that developed a new model since 2006 (*continued*)

Reference	Follow-up	Model	n	Randomi-zation	Cohort	Setting	Procedure	Study type*	Primary outcome	Statistics
Giles et al. [15]	2001–2004	Giles score†	45 660	+	P: Medicare, USA	EAAA	Both (50% EVAR)	A	30-day/in-hospital mortality	Univariable P < 0.100 BW P < 0.050 (+ interaction)
Grant et al.[16]	1999–2009	VGNW†	2765	+	P: VGNW, UK	EAAA	Both (18.9% EVAR)	A	30-day/in-hospital mortality	FW P < 0.100 BW P < 0.050
Grant et al.[17]	2008–2011	BAR†	11 423 Bootstrap 40 ×	n.a.	P: NVD, UK	EAAA	Both (56.8% EVAR)	A	In-hospital mortality	BW AIC
Mastracci et al.[18]	1998–2009	Cleveland Clinic experience†	412 Bootstrap 200 ×		P: Ohio, USA	EAAA	EVAR	A	Long-term mortality/30-day mortality	Univariable P < 0.200 Multivariable P < 0.050 Shrinkage
			697 (geographical validation)		P: Zenith multicenrte trial, USA			A		
Robinson et al.[19]	2003–2009	VSGNE†	242	n.a.	P: VSGNE, USA	RAAA	OSR	A/B	In-hospital mortality	–
		GAS	Bootstrap							
		Hardman index	1000 ×							
		Vancouver score								
		ERAS								
Tang et al.[20]	1998–2006	Cambridge POSSUM†	452	+/–§	P: Cambridge, UK	EAAA (57.1%) RAAA SAAA	OSR	A/B	30-day/in- hospital mortality	–
		POSSUM								
		P-POSSUM								
		V-POSSUM								
		V(p)-POSSUM								

Table 1 Data extracted from studies that developed a new model since 2006 (*continued*)

Reference	Follow-up	Model	n	Randomi-zation	Cohort	Setting	Procedure	Study type*	Primary outcome	Statistics
Tang et al.[22]	2002–2004	VBHOM†	2718		P: NVD, UK	EAAA Non-elective EAAA RAAA	OSR/ IB	A	Mortality at discharge	–
	1998–2005	VBHOM†	327 (geographical validation)		P: Cambridge, UK			A		
Visser et al.[21]	2002–2006	GAS†	201 Bootstrap 200 ×	n.a.	P: The Netherlands	RAAA	Both (28.9% EVAR)	A/B	30-day mortality	–
		GAS (updated)†								
Wise et al.[23]	1998–2013	ANN (digital)†	86 ANN method	n.a.	R: van der Bilt,	RAAA	Both (13.6%	A/B	In-hospital mortality	–
		ANN (digital)	21		Tennessee, USA		EVAR)			
		GAS	83							
		Multiple regression	107							

*A, prediction model development study with validation; B, external model validation study with or without model updating. †New models published since 2006. ‡Every third patient moved to the modelling set. §Alternating ruptured and non-ruptured. +, Good; +/–, moderate; –, not good; preop., preoperative; AAA SCORE, Abdominal Aortic Aneurysm Statistically Corrected Operative Risk Evaluation; periop., perioperative; VBHOM, Vascular Biochemistry and Haematology Outcome Model; VGNW, Vascular Governance North West; (P) (V) POSSUM, (Portsmouth) (Vascular) Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity, with (p) denoting physiological only; P, Prospective; NVD, National Vascular Database; EAAA, elective aortic abdominal aneurysm repair; SAAA, symptomatic aortic abdominal aneurysm repair; RAAA, ruptured aortic abdominal aneurysm repair; AIC; Akaike's information criterion; ERA, Endovascular aneurysm repair Risk Assessment; n.a., not applicable; EVAR, endovascular aneurysm repair; FW, forward; OSR, open surgical repair; GAS, Glasgow Aneurysm Score; (m-)CPI, (modified) Customized Probability Index; NHS, National Health Service; BW, backward; BAR, British Aneurysm Repair; VSGNE, Vascular Study Group of New England; ERAS, Edinburgh Ruptured Aneurysm Score; IB, infraguinal bypass; ANN, artificial neural network; R, retrospective.

Table 2 Data extracted from studies investigating an existing model

Reference	Follow-up	Model	n	Cohort	Setting	Procedure	Study type*	Primary outcome
Ali et al.[24]	2003–2013	VSGNE†	1165	R: VQI	RAAA	Both (44.1% EVAR)	B	In-hospital mortality
Antonopoulos et al.[25]	2006–2012	GAS updated†	418	P: Athens, Thessaloniki, Larissa, Greece	RAAA	Both (27% EVAR)	B	In-hospital mortality
Barnes et al.[26]	2001–2007	ERA†	310	R: St George's, London, UK	EAAA	EVAR	B	Early death
		ERA†		P: Adelaide, Australia	EAAA SAAA	EVAR		
Bryce et al.[27]	2005–2007	GAS	106	P: Glasgow, UK	EAAA	OSR	B	30-day/in-hospital mortality Major adverse cardiac events
		V(p)-POSSUM						
		VBHOM†						
		Lee Index						
		E-PASS						
Grant et al.[28]	2008–2010	VBHOM†	10 891	P: NVD, UK	EAAA	Both (54.5% EVAR)	B	In-hospital mortality
		VGNW†						
		V-POSSUM						
		GAS						
		Medicare†						
Grant et al.[29]	2011–2013	BAR†	1124	P: VGNW, UK	EAAA (7.7%) SAAA	Both (67.5% EVAR)	B	In-hospital mortality
		Medicare†						
		VGNW†						
Nevala et al.[30]	2001–2005	GAS	205	R: Uolu, Kuopio, Helsinki	EAAA	EVAR	B	30-day mortality

Table 2 Data extracted from studies investigating an existing model (*continued*)

Reference	Follow-up	Model	n	Cohort	Setting	Procedure	Study type*	Primary outcome
		Giles score†		Hospitals, Finland				
		Modified Leiden score						
Patterson et al.[31]	2004–2008	GAS	846	R: South UK	EAAA	EVAR	B	30-day mortality
		CPI						
		Leiden score						
		m-CPI						
		VBHOM†						
Tambyraja et al.[32]	2002–2004	GAS	84	P: Edinburgh, UK	RAAA	OSR	B	30-day/in-hospital mortality
		Hardman index						
		V(p)-POSSUM						
		RAAA (p)-POSSUM						
		ERAS†						
Tang et al.[33]	1998–2004	GAS	204	P: Cambridge, UK (same cohort)	EAAA	OSR	B	30-day mortality
		VBHOM†						
		E-PASS						
van Beek et al.[34]	2000–2003	Medicare†	345	RCT: Belgium and the Netherlands	EAAA	Both (49.3% EVAR)	B	30-day/in-hospital mortality
		VGNW†						
		BAR†						
van Beek et al.[35]	1997–2010	ERA†	433	R: Amsterdam, The Netherlands	EAAA	EVAR	B	30-day mortality

Table 2 Data extracted from studies investigating an existing model (*continued*)

Reference	Follow-up	Model	n	Cohort	Setting	Procedure	Study type*	Primary outcome
van Beek et al.[36]	2004–2011	GAS updated†	449	R: Amsterdam, The Netherlands	RAAA	Both (18.2% EVAR)	B	30-day/in-hospital mortality
		Vancouver score						
		ERAS†						
		Hardman index						
Wisniowski et al.[37]	Until 2009	ERA†	197	P: Royal Brisbane Womens Hospital, Australia	EAAA	EVAR	B	Early death

*B, external model validation study with or without model updating. †New models published since 2006. VSGNE, Vascular Study Group of New England; R, retrospective; VQI, Vascular Quality Initiative; RAAA, ruptured aortic abdominal aneurysm repair; EVAR, endovascular aneurysm repair; GAS, Glasgow Aneurysm Score; P, Prospective; ERA, Endovascular aneurysm repair Risk Assessment; EAAA, elective aortic abdominal aneurysm repair; SAAA, symptomatic aortic abdominal aneurysm repair; (P) (V) POSSUM, (Portsmouth) (Vascular) Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity, with (p) denoting physiological only; VBHOM, Vascular Biochemistry and Haematology Outcome Model; E-PASS, estimating the physiological ability and surgical stress; OSR, open surgical repair; VGNW, Vascular Governance North West; NVD, National Vascular Database; BAR, British Aneurysm Repair; (m-)CPI, (modified) Customized Probability Index; ERAS, Edinburgh Ruptured Aneurysm Score.

Model performance

Five[12,13,15 – 17] of the model development studies performed a forward or backward selection with an accompanying P value for inclusion or exclusion of the variable. Two studies[14,18] included variables according to a prespecified P value by hand using primarily a univariable and then a multivariable analysis to select variables relevant to the model. Five studies[19 – 23] did not mention a method of inclusion. Ambler and colleagues[11] performed a reclassification procedure in which patients were reassigned to different risk categories and cross-tabulated to show the changes in categories when changing models. Mastracci and co-workers[18] shrank their data to reduce overfitting of their final model.

A calibration plot or measurement of the correlation between observed and expected outcome by the H-L test was not often performed.[10] All studies, except that of Barnes and colleagues[12], who mentioned a P value only for the H-L test, calculated a C-statistic to describe the discriminative ability as a measure of goodness of fit.

Discriminative performance of every model reported in each article is shown in Figs 2 – 4 by procedure (EVAR, OSR or both) and setting: elective aneurysm repair, symptomatic aneurysm repair, ruptured aneurysm repair, or all three combined.

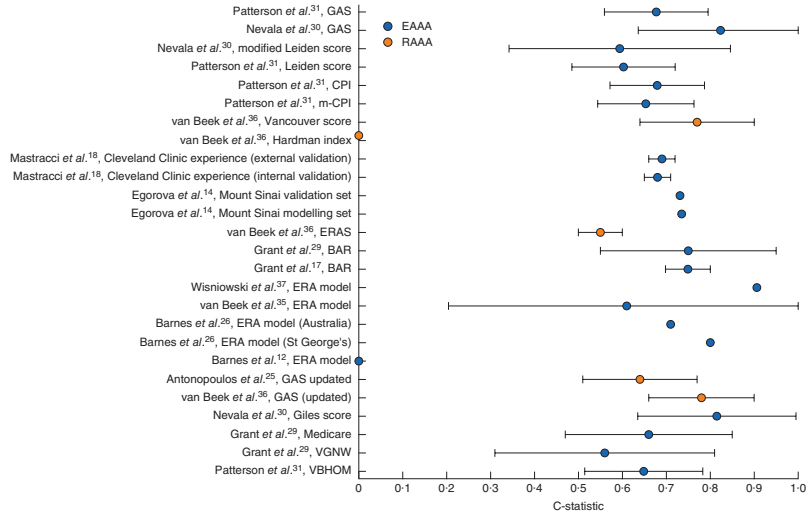


Figure 2. C-statistic values, with 95% confidence intervals, for the mortality risk prediction models evaluating endovascular aneurysm repair. EAAA, elective aortic abdominal aneurysm repair; RAAA, ruptured aortic abdominal aneurysm repair; GAS, Glasgow Aneurysm Score; (m-)CPI, (modified) Customized Probability Index; ERAS, Edinburgh Ruptured Aneurysm Score; BAR, British Aneurysm Repair; ERA, Endovascular aneurysm repair Risk Assessment; VGNW, Vascular Governance North West; VBHOM, Vascular Biochemistry and Haematology Outcome Model

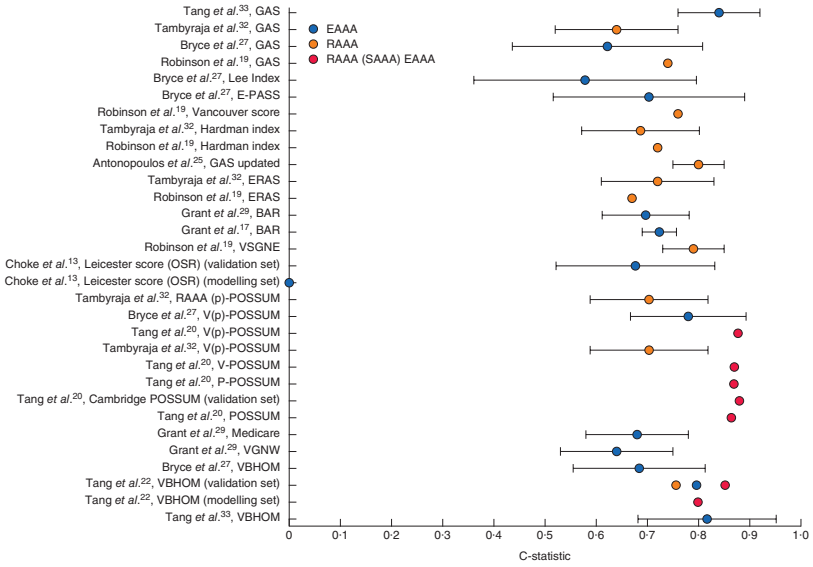


Figure 3. C-statistic values, with 95% confidence intervals, for the mortality risk prediction models evaluating open surgical repair (OSR). EAAA, elective aortic abdominal aneurysm repair; RAAA, ruptured aortic abdominal aneurysm repair; SAAA, symptomatic aortic abdominal aneurysm repair; GAS, Glasgow Aneurysm Score; E-PASS, estimating the physiological ability and surgical stress; ERAS, Edinburgh Ruptured Aneurysm Score; BAR, British Aneurysm Repair; VSGNE, Vascular Study Group of New England; (P) (V) POSSUM, (Portsmouth) (Vascular) Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity, with (p) denoting physiological only; VGNW, Vascular Governance North West; VBHOM, Vascular Biochemistry and Haematology Outcome Model

Risk prediction models

Detailed information about model formulas is provided in Table S2 (supporting information).

Cambridge POSSUM

The Cambridge POSSUM was developed in 2007 on part of a population in Cambridge and validated on another part of that population for elective and ruptured AAA undergoing OSR (C-statistic 0.88)[20]. It has not been validated externally.

Vascular Biochemistry and Haematology Outcome Model

VBHOM was developed in 2007 using the UK NVD of some 3045 patients with either elective or ruptured AAA undergoing OSR.[22] It demonstrated good performance by external geographical validation (C-statistic 0.852, H-L P = 0.59, EPV 6.6).[22] External (geographical) validation is the standard when assessing a model's capabilities and its applicability.[10]

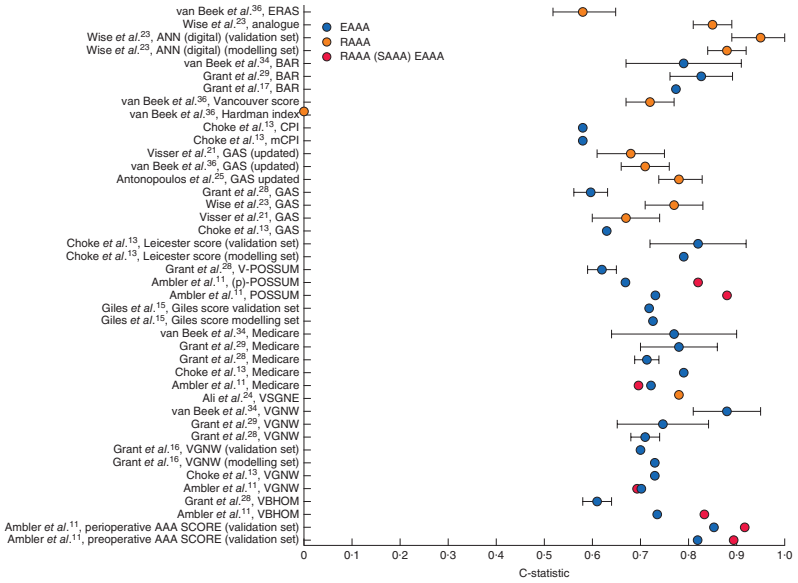


Figure 4. C-statistic values, with 95% confidence intervals, for the mortality risk prediction models evaluating both endovascular and open aneurysm repairs. EAAA, elective aortic abdominal aneurysm repair; RAAA, ruptured aortic abdominal aneurysm repair; SAAA, symptomatic aortic abdominal aneurysm repair; ERAS, Edinburgh Ruptured Aneurysm Score; ANN, artificial neural network; BAR, British Aneurysm Repair; (m-)CPI, (modified) Customized Probability Index; GAS, Glasgow Aneurysm Score; (V) POSSUM, (Vascular) Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity, with (p) denoting physiological only; VSGNE, Vascular Study Group of New England; VGNW, Vascular Governance North West; VBHOM, Vascular Biochemistry and Haematology Outcome Model; AAA SCORE, Abdominal Aortic Aneurysm Statistically Corrected Operative Risk Evaluation

Another study[27] performed the same validation on an elective AAA OSR population but with a low discriminative ability (C-statistic 0.684, 95% CI 0.555 to 0.813; EPV 1.1). Similar results were reported by Patterson and colleagues[31], when performing the validation on an elective EVAR population (C-statistic 0.649, 95% CI 0.514 to 0.783). In 2012, Grant et al.[28] validated VBHOM on a population of patients having elective EVAR or OSR, but performance was poor (C-statistic 0.61, 95% CI 0.58 to 0.64; EPV 39.4). Ambler and colleagues[11] also externally validated the model in the UK. It performed well by discrimination (C-statistic 0.833) but with very poor calibration on a data set with ruptured AAA, symptomatic AAA and elective AAA treated by either EVAR or OSR. For elective AAA surgery alone, it had a C-statistic of 0.735 and also poor calibration.

Edinburgh Ruptured Aneurysm Score

The Edinburgh Ruptured Aneurysm Score (ERAS) was developed in 2007 on a population undergoing OSR for ruptured AAA (2000 – 2002) by Tambyraja and co-workers[38],

but not directly validated (Table S1, supporting information). In 2008, they validated their model on patients having OSR for ruptured aneurysm between 2002 and 2004 with a moderate performance (C-statistic 0.72).[32] Five years later, Robinson and co-workers[19] validated a similar population (OSR for ruptured AAA), but performance was poorer (C-statistic 0.67). Van Beek et al.[36] were the only ones who tested the ERAS on a ruptured AAA group for both procedures and for a subselection of EVAR, with poor results (C-statistic 0.58 for both procedures and 0.55 for EVAR only).

Endovascular aneurysm repair Risk Assessment model

Barnes and colleagues[12] developed the Endovascular aneurysm repair Risk Assessment model in 2008 for EVAR specifically for patients undergoing elective and symptomatic AAA surgery. They bootstrapped their data set 100 times and assessed goodness of fit by the Cessie – van Houwelingen – Copas – Hosmer unweighted sum of squares test statistic. It showed very good calibration of the model with a P value of 0.92 (the same as with the normal H-L test). No C-statistic was presented. In 2010, ERA was validated externally by the same research group with a moderate to good performance.[26] The exact formula was not available, but the model can be used for the individual patient online.

Updated Glasgow Aneurysm Score

The updated Glasgow Aneurysm Score (GAS) was introduced in 2009, with a new variable: the procedure performed (EVAR or OSR).[21] It was validated in two studies (C-statistic 0.710, 95% CI 0.66 to 0.76, H-L P = 0.01[36]; C-statistic 0.780, 0.740 to 0.830[25]). Because of the paucity and continuous nature of the variables in the model, the EPV was larger than 1.

Giles score

The Giles score was developed on elective AAA repair for both procedures in 2009 by Giles and colleagues.[15] It showed only a minimal difference between the modelling and validation set with a moderate performance. When Nevala and co-workers[30] validated this score for EVAR only, it performed well (C-statistic 0.815, 95% CI 0.635 to 0.995). In recent studies, the Giles score has been referred to as the Medicare model.[13,28,29] These studies all showed C-statistic values of 0.66 or higher for patients undergoing elective AAA surgery and a population based on only EVAR (C-statistic 0.66, 0.47 to 0.85)[29] or OSR (C-statistic 0.68, 0.85 to 0.87)[29] or both (C-statistic 0.79[13]; C-statistic 0.71, 95% CI 0.69 to 0.74[28] ; C-statistic 0.78, 0.70 to 0.86[29]). When validated for OSR or EVAR only, the C-statistic values decreased. However, Ambler and colleagues[11] showed an increase in discriminative ability despite a less heterogeneous population for elective AAA repair only (C-statistic 0.722) compared with the overall AAA group (C-statistic

0.696). The only Dutch study[34] that validated the Medicare model for both procedures in patients having elective AAA surgery had a moderate performance and good calibration (H-L P = 0.52; C-statistic 0.77, 95% CI 0.64 to 0.90).

Mount Sinai score

Egorova and colleagues[14] developed a model for patients undergoing elective surgery by EVAR based on Medicare data for 30-day mortality, with moderate to good results considering discrimination and calibration (C-statistic 0.731; H-L P = 0.24). One study[39] described risk stratification using the Mount Sinai score, but without reporting a C-statistic or calibration test.

Cleveland Clinic experience

Mastracci et al.[18] also validated their model on an external population. The different elective EVAR population used for external validation were participants in a multicentre trial for EVAR, but performance of the model was poor (C-statistic 0.69, 95% CI 0.66 to 0.72).

Leicester score

The Leicester score[13] for elective AAA surgery was developed for both procedures, but also for OSR only. The internally validated model performed well in the validation data set (C-statistic 0.82, 95% CI 0.72 to 0.92; EPV 3.3). However, in analyses for OSR only it performed poorly (C-statistic 0.68, 0.52 to 0.83).

Vascular Governance New West model

The Vascular Governance New West model is the only one of the newer models that has been validated five times, although all in a UK population, between 1999 and 2013. It was developed in 2011 for both procedures in patients undergoing elective AAA surgery.[16] However, only 18.9% of patients had EVAR, resulting in a C-statistic of 0.71 with a H-L P value of 0.853 and EPV of 7.1. In 2012, external validation in patients having elective surgery (EVAR 54.5%) yielded a C-statistic of 0.71, but the H-L value almost became significant (P = 0.066).[28] In 2014, in a population with 67.5% EVAR, both the C-statistic and the H-L value were similar to those in 2011 and 2012 (C-statistic 0.75, 95% CI 0.65 to 0.84).[29] Validation in two studies for all patients with AAA and only elective repair with both procedures resulted in a poor overall performance (C-statistic 0.693)[11] and a moderate performance (C-statistic 0.702[11] and 0.73[13]) for elective procedures only. It was also validated by van Beek and colleagues[34] in an elective Dutch population with a C-statistic of 0.88 (H-L P = 0.31).

Vascular Study Group of New England model

A US population with ruptured AAA undergoing OSR was used to create a new mortality risk prediction model, the Vascular Study Group of New England (VSGNE) model, in 2013.[19] After bootstrapping the model (multiple sampling), it showed a good calibration (H-L P = 0.85) and good discriminative ability (C-statistic 0.79).[19] It has been validated externally in one study[24], with a C-statistic of 0.78 and H-L P=0.1.

British Aneurysm Repair score

The British Aneurysm Repair (BAR) score[17] was developed in patients undergoing either elective EVAR or OSR and, in external validation, had a C-statistic of 0.83 (95% CI 0.76 to 0.89) and H-L P value of 0.581, with similar values for EVAR and OSR only.[29] When validated by van Beek and colleagues[34] in 2013, the performance remained good (C-statistic 0.79, 95% CI 0.67 to 0.91), with a H-L P value of 0.15, but the EPV was less than 1.

Abdominal Aortic Aneurysm Statistically Corrected Operative Risk Evaluation Score

Ambler and colleagues[11] showed some promising results for the AAA SCORE in a population consisting of both acute and elective AAA. The preoperative model performed well (C-statistic 0.894 for all patients and 0.819 for elective aneurysm repair only; H-L P = 0.33). The H-L test result for the perioperative model was good (P = 0.35) with a good C-statistic (0.917) and remained good for patients having elective AAA repair (C-statistic 0.853). The preoperative model also contained procedural details, which makes it a perioperative model and not one that can be applied before surgery.

Artificial neural network

Instead of calculating the model by hand, Wise and colleagues[23] developed an ANN model digitally for patients undergoing ruptured AAA surgery with both EVAR and OSR. Performance was good; the C-statistic was 0.88 in the modelling set, 0.95 in the validation set and 0.85 for the analogue multiple logistic regression data set.

DISCUSSION

This systematic review of contemporary mortality prediction models for AAA surgery reveals the difficulties in developing, validating and using a mortality risk prediction rule. The large number of models developed is the result of low optimal performance, possibly influenced by the introduction of EVAR and the concomitant altered prognosis for the individual patient.[40] Instead of updating existing models, most studies have

presented a new model. This has resulted in better performance in their population compared with existing models that were developed in another population and validated externally. At present, the ideal model does not exist.

External validation studies give an indication of model performance, but the large number of subgroups by setting (acute, elective or both) and procedure (EVAR, OSR or both) and the highly variable number of events, compared with the number of variables and categories, hampers comparison between models and generalisability for the individual patient with AAA.

Although nearly every study calculated a C-statistic to describe the discriminative abilities of the model, calibration measures varied widely. Only a few studies reported a H-L analysis for calibration, whereas others used $\chi 2$ or alternative measurements, or must have compared observed with expected but did not report the comparison statistic.

Overall, discrimination is easy to measure compared with calibration, and the latter is easily improved by updating methods when applied to another population. However, good calibration is necessary for calculating predictions, independent of a high or low C-statistic. The clinical usefulness of a model can only be determined when both discrimination and calibration are available, and a cut-off value is defined for the sensitivity and specificity.[10]

The BAR score was found to be one of the best performing perioperative mortality risk prediction models, with good calibration, clinical practicality (variables easy to retrieve) and generalisability.[17] The large number of variables with one to three categories might become a problem in surgery with low volumes and few events (mortality), for example for elective AAA surgery, especially EVAR. The choice of repair is a strong predictive variable for mortality and therefore a valuable determinant of an individual patient's prognosis. VBHOM was found to be one of the best performing preoperative models, considering applicability and a few merely continuous variables.[22]

Whether a model is suitable for clinical practice is highly dependent on its purpose: to predict mortality, stratify mortality risk or perform risk adjustment. Prediction of mortality and the model of choice is dependent on the moment of prediction. It has become clear that, despite multiple attempts, no one model is used primarily in clinical practice. This is probably because of differences in setting and concomitant definitions, and different implementation of procedures, but also differences between countries in, for example, screening or the proportion of patients with ruptured AAA turned down for surgery.[41,42]

Since the introduction of EVAR, patients who would have been turned down for OSR previously, owing to their co-morbidity, might now receive intervention.[43] Therefore, recently developed models are promising. However, the effect of AAA morphology on outcome is critical. In the era of EVAR, perioperative mortality rates are very low. For this reason, a more suitable outcome parameter, such as reintervention, handling of complications or long-term follow-up, is required.

The objective of this systematic review was to find one model suitable for clinical practice, with a small number of variables, ease of interpretation, and utility for both emergency and elective AAA, with good statistical performance. There are too many models, and no single one that is applicable universally, for either elective or acute aneurysm surgery. None of the current scoring systems appears ideal.

Validation of existing studies continues, as does the development of newer or updated models, such as a recently new VSGNE model.[44,45] A recently published tripartite study[46] has shown that laboratory values do not necessarily lead to better prediction of mortality after several surgical procedures. It was also found that including more variables does not improve model performance. There is thus a need for a model consisting of only a few important variables.[47] A model should be re-estimated on a preferred population before being used.

ACKNOWLEDGEMENTS

The authors thank J. W. M. Plevier, of the Walaeus Library at Leiden University Medical Centre, for her collaboration in this work.

Disclosure: The authors declare no conflict of interest.

REFERENCES

1 Chaikof EL, Brewster DC, Dalman RL, Makaroun MS, Illig KA, Sicard GA et al. SVS practice guidelines for the care of patients with an abdominal aortic aneurysm: executive summary. J Vasc Surg 2009;50:880 – 896.

2 Moll FL, Powell JT, Fraedrich G, Verzini F, Haulon S, Waltham M et al.; European Society for Vascular Surgery. Management of abdominal aortic aneurysms clinical practice guidelines of the European Society for Vascular Surgery. Eur J Vasc Endovasc Surg 2011;41(Suppl 1):S1 – S58.

3 Lederle FA, Wilson SE, Johnson GR, Reinke DB, Littooy FN, Acher CW et al. Immediate repair compared with surveillance of small abdominal aortic aneurysms. N Engl J Med 2002;346:1437 – 1444.

4 Lijftogt N, Vahl AC, Wilschut ED, Elsman BH, Amodio S, van Zwet E et al.; Dutch Society of Vascular Surgery, Steering Committee of the Dutch Surgical Aneurysm Audit, and Dutch Institute for Clinical Auditing. Adjusted hospital outcomes of abdominal aortic aneurysm surgery reported in the Dutch Surgical Aneurysm Audit. Eur J Vasc Endovasc Surg 2017;53:520 – 532.

5 Arko FR III, Murphy EH, Boyes C, Nussbaum T, Lalka SG, Holleman J et al. Current status of endovascular aneurysm repair: 20 years of learning. Semin Vasc Surg 2012;25:131 – 135.

6 Patterson BO, Holt PJE, Hinchliffe R, Loftus IM, Thompson MM. Predicting risk in elective abdominal aortic aneurysm repair: a systematic review of current evidence. Eur J Vasc Endovasc Surg 2008;36:637 – 645.

7 Kragsterman B. Nationella kvalitetsregistret för kärlkirurgi; 2014. http://www.ucr.uu.se/swedvasc/arsrapporter/ swedvasc-2014/download [accessed 30 April 2017].

8 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8:336 – 341.

9 Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11:e1001744.

10 Steyerberg E. Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating (1st edn). Springer, Rotterdam, 2009.

11 Ambler GK, Gohel MS, Mitchell DC, Loftus IM, Boyle JR; Audit and Quality Improvement Committee of the Vascular Society of Great Britain and Ireland. The Abdominal Aortic Aneurysm Statistically Corrected Operative Risk Evaluation (AAA SCORE) for predicting mortality after open and endovascular interventions. J Vasc Surg 2015;61:35–43.

12 Barnes M, Boult M, Maddern G, Fitridge R. A model to predict outcomes for endovascular aneurysm repair using preoperative variables. Eur J Vasc Endovasc Surg 2008;35:571 – 579.

13 Choke E, Lee K, McCarthy M, Nasim A, Naylor AR, Bown M et al. Risk models for mortality following elective open and endovascular abdominal aortic aneurysm repair: a single institution experience. Eur J Vasc Endovasc Surg 2012;44:549 – 554.

14 Egorova N, Giacovelli JK, Gelijns A, Greco G, Moskowitz A, McKinsey J et al. De ning high-risk patients for endovascular aneurysm repair. J Vasc Surg 2009;50:1271 – 1279.e1.

15 Giles KA, Schermerhorn ML, O'Malley AJ, Cotterill P, Jhaveri A, Pomposelli FB et al. Risk prediction for perioperative mortality of endovascular vs open repair of abdominal aortic aneurysms using the Medicare population. J Vasc Surg 2009;50:256 – 262.

16 Grant SW, Grayson AD, Purkayastha D, Wilson SD, McCollum C. Logistic risk model for mortality following elective abdominal aortic aneurysm repair. Br J Surg 2011;98:652 – 658.
17 Grant SW, Hickey GL, Grayson AD, Mitchell DC, McCollum CN. National risk prediction model for elective abdominal aortic aneurysm repair. Br J Surg 2013;100:645 – 653.
18 Mastracci TM, Greenberg RK, Hernandez AV, Morales C. De ning high risk in endovascular aneurysm repair. J Vasc Surg 2010;51:1088 – 1095.
19 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;57:354 – 361.
20 Tang TY, Walsh SR, Prytherch DR, Wijewardena C, Gaunt ME, Varty K et al. POSSUM models in open abdominal aortic aneurysm surgery. Eur J Vasc Endovasc Surg 2007;34:499 – 504.
21 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;49:1093 – 1099.
22 Tang T, Walsh SR, Prytherch DR, Lees T, Varty K, Boyle JR et al.; Audit Research Committee of the Vascular Society of Great Britain and Ireland. VBHOM, a data economic model for predicting the outcome after open abdominal aortic aneurysm surgery. Br J Surg 2007;94:717 – 721.
23 Wise ES, Hocking KM, Brophy CM. Prediction of in-hospital mortality after ruptured abdominal aortic aneurysm repair using an arti cial neural network. J Vasc Surg 2015;62:8 – 15.
24 Ali MM, Flahive J, Schanzer A, Simons JP, Aiello FA, Doucet DR et al. In patients strati ed by preoperative risk, endovascular repair of ruptured abdominal aortic aneurysms has a lower in-hospital mortality and morbidity than open repair. J Vasc Surg 2015;61:1399 – 1407.
25 Antonopoulos CN, Kakisis JD, Andrikopoulos V, Dervisis K, Georgopoulos S, Giannoukas A et al. Predictors affecting in-hospital mortality of ruptured abdominal aortic aneurysms: a Greek multicenter study. Ann Vasc Surg 2014;28:1384 – 1390.
26 Barnes M, Boult M, Thompson MM, Holt PJ, Fitridge RA. Personalised predictions of endovascular aneurysm repair success rates: validating the ERA model with UK Vascular Institute data. Eur J Vasc Endovasc Surg 2010;40:436 – 441.
27 Bryce GJ, Payne CJ, Gibson SC, Kingsmore DB, Byrne DS, Delles C. Risk strati cation scores in elective open abdominal aortic aneurysm repair: are they suitable for preoperative decision making? Eur J Vasc Endovasc Surg 2012;44:55 – 61.
28 Grant SW, Grayson AD, Mitchell DC, McCollum CN. Evaluation of ve risk prediction models for elective abdominal aortic aneurysm repair using the UK National Vascular Database. Br J Surg 2012;99:673 – 679.
29 Grant SW, Hickey GL, Carlson ED, McCollum CN. Comparison of three contemporary risk scores for mortality following elective abdominal aortic aneurysm repair. Eur J Vasc Endovasc Surg 2014;48:38–44.
30 Nevala T, Biancari F, Perälä J, Manninen H, Aho P-S, Matsi P et al. Risk prediction in patients undergoing elective endovascular repair of an abdominal aortic aneurysm. Scand Cardiovasc J 2010;44:125–128.

31 Patterson BO, Holt PJ, Hinchliffe R, Nordon IM, Loftus IM, Thompson MM. Existing risk prediction methods for elective abdominal aortic aneurysm repair do not predict short-term outcome following endovascular repair. J Vasc Surg 2010;52:25–30.
32 Tambyraja AL, Lee AJ, Murie JA, Chalmers RTA. Prognostic scoring in ruptured abdominal aortic aneurysm: a prospective evaluation. J Vasc Surg 2008;47:282–286.
33 Tang TY, Walsh SR, Fanshawe TR, Seppi V, Sadat U, Hayes PD et al. Comparison of risk-scoring methods in predicting the immediate outcome after elective open abdominal aortic aneurysm surgery. Eur J Vasc Endovasc Surg 2007;34:505 – 513.
34 van Beek SC, Blankensteijn JD, Balm R; Dutch Randomised Endovascular Aneurysm Management (DREAM) trial collaborators. Validation of three models predicting in-hospital death in patients with an abdominal aortic aneurysm eligible for both endovascular and open repair. J Vasc Surg 2013;58:1452–1457.e1.
35 van Beek SC, Legemate DA, Vahl A, Wisselink W, Barnes M, Fitridge RA et al. External validation of the Endovascular aneurysm repair Risk Assessment model in predicting survival, reinterventions, and endoleaks after endovascular aneurysm repair. J Vasc Surg 2014;59:1555 – 1561.
36 van Beek SC, Reimerink JJ, Vahl AC, Wisselink W, Peters RJG, Legemate DA et al.; Amsterdam Acute Aneurysm Trial Collaborators. Editor's choice – External validation of models predicting survival after ruptured abdominal aortic aneurysm repair. Eur J Vasc Endovasc Surg 2015;49:10–16.
37 Wisniowski B, Barnes M, Jenkins J, Boyne N, Kruger A, Walker PJ. Predictors of outcome after elective endovascular abdominal aortic aneurysm repair and external validation of a risk prediction model. J Vasc Surg 2011;54:644 – 653.
38 Tambyraja A, Murie J, Chalmers R. Predictors of outcome after abdominal aortic aneurysm rupture: Edinburgh Ruptured Aneurysm Score. World J Surg 2007;31:2243 – 2247.
39 Haji Mohd Yasin NA, Thomson IA, Leon de la Barra S. Endovascular aortic repair: can we predict who will not get long-term bene t? N Z Med J 2013;126:44 – 52.
40 Best WB, Ahanchi SS, Larion S, Ammar CP, Lavingia KS, Panneton JM. Abdominal aortic aneurysm anatomic severity grading score predicts implant-related complications, systemic complications, and mortality. J Vasc Surg 2016;63:577 – 584.
41 Karthikesalingam A, Holt PJ, Vidal-Diez A, Ozdemir BA, Poloniecki JD, Hinchliffe RJ et al. Mortality from ruptured abdominal aortic aneurysms: clinical lessons from a comparison of outcomes in England and the USA. Lancet 2014;383:963–969.
42 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 Vascunet report. Eur J Vasc Endovasc Surg 2011;42:598–607.
43 Badger SA, Harkin DW, Blair PH, Ellis PK, Kee F, Forster R. Endovascular repair or open repair for ruptured abdominal aortic aneurysm: a Cochrane systematic review. BMJ Open 2016;6:e008391.
44 Krenzien F, Wiltberger G, Hau HM, Matia I, Benzing C, Atanasov G et al. Risk strati cation of ruptured abdominal aortic aneurysms in patients treated by open surgical repair. Eur J Vasc Endovasc Surg 2016;51:30–36.
45 Eslami MH, Rybin D, Doros G, Kalish JA, Farber A; Vascular Study Group of New England. Comparison of a Vascular Study Group of New England risk prediction model with established risk prediction models of in-hospital mortality after elective abdominal aortic aneurysm repair. J Vasc Surg 2015;62:1125–1133.e2.

46 Meguid RA, Bronsert MR, Juarez-Colunga E, Hammermeister KE, Henderson WG. Surgical Risk Preoperative Assessment System (SURPAS): II. Parsimonious risk models for postoperative adverse outcomes addressing need for laboratory variables and surgeon specialty-speci c models. Ann Surg 2016;264:10–22.

47 Meguid RA, Bronsert MR, Juarez-Colunga E, Hammermeister KE, Henderson WG. Surgical Risk Preoperative Assessment System (SURPAS): III. Accurate preoperative prediction of 8 adverse outcomes using 8 predictor variables. Ann Surg 2016;264:23–31.

4

Towards Optimizing Risk Adjustment in the Dutch Surgical Aneurysm Audit (DSAA)

N. Lijftogt
A. C. Vahl
E. M. van der Willik
V. J. Leijdekkers
M. W. J. M. Wouters
J. F. Hamming

in collaboration with the Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing

Annals of Vascular Surgery 2019

ABSTRACT

Background:

To compare hospital outcomes of aortic aneurysm surgery, casemix correction for preoperative variables is essential. Most of these variables can be deduced from mortality risk prediction models. Our aim was to identify the optimal set of preoperative variables associated with mortality in order to establish a relevant and efficient casemix model.

Material and Methods:

All patients prospectively registered between 2013 - 2016 in the Dutch Surgical Aneurysm Audit (DSAA) were included for analysis. After multiple imputation for missing variables, predictors for mortality following univariable logistic regression were analysed in a manual backward multivariable logistic regression model and compared with three standard mortality risk prediction models: Glasgow Aneurysm Score (GAS, mainly clinical parameters), Vascular Biochemical and Haematological Outcome Model (VBHOM, mainly laboratory parameters) and Dutch Aneurysm Score (DAS, both clinical and laboratory parameters). Discrimination and calibration were tested and considered good with a C-statistic > 0.8 and Hosmer- Lemeshow (H-L) $p > 0.05$.

Results:

There were 12 401 patients: 9537 (76.9%) elective patients (EAAA), 913 (7.4%) acute symptomatic patients (SAAA), and 1951 (15.7%) patients with acute rupture (RAAA). Overall postoperative mortality was 6.5%; 1.8% after EAAA surgery, 6.6% after SAAA and 29.6% after RAAA surgery. The optimal set of independent variables associated with mortality were a mix of clinical and laboratory parameters: gender, age, pulmonary comorbidity, operative setting, creatinine, aneurysm size, hemoglobin, Glasgow coma scale, ECG and systolic blood pressure (C-statistic 0.871). External validation overall of VBHOM, DAS and GAS revealed C-statistics of 0.836, 0.782, 0.761, with an H-L of 0.028, 0.00 and 0.128, respectively.

Conclusion:

The optimal set of variables for casemix correction in the DSAA comprises both clinical and laboratory parameters which can be collected easily from electronic patient files and will lead to an efficient casemix model.

INTRODUCTION

Background

Since 2013, it has been mandatory for all patients undergoing surgery for an abdominal aortic aneurysm (AAA) to be registered in the Dutch Surgical Aneurysm Audit (DSAA).[1] For a true interpretation of hospital outcomes, casemix risk adjustment has to be performed in order to level those differences in preoperative patient- and disease related variables that influence outcome and which vary between hospitals.[2] Many of the variables present in casemix models are also represented in mortality risk prediction models, as was summarised in a recent systematic review.[3] Despite this multiplicity of models, no standard mortality risk prediction model in AAA surgery has been broadly implemented in clinical practice, because every model has been developed for a certain population during a certain time period, which makes them less generalisable to other populations.[3]

Prediction models are based on physiological parameters, e.g. the Glasgow Aneurysm Score (GAS)[4], on laboratory parameters e.g. the Vascular Biochemical and Haematological Outcome Model (VBHOM)[5] or mixed models such as the Dutch Aneurysm Score (DAS)[6]. The physiology-only Vascular Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (V(p) - POSSUM), contains one of the largest numbers of variables and has been extensively investigated.[7] All these variables were included in the original Dutch Surgical Aneurysm Audit (DSAA) dataset to calculate the V(p)-POSSUM score. However, preoperative surgical risk assessment has recently been shown to be possible using a maximum of eight, easily-retrievable variables.[8] Due to the substantial registration burden imposed by the large number of variables, a validated optimal dataset with a minimum number of risk adjustment parameters was needed, one in which the number of parameters is proportional to the number of events.[9]

Objectives

The aim of this study was to identify the optimal set for risk adjustment with a minimum number of casemix variables by means of an interhospital comparison of postoperative mortality for every AAA patient and for both elective and acute AAA patients, and to internally validate this set. External validation was performed using previously developed preoperative risk prediction models for variable comparisons.

MATERIAL AND METHODS

Study design and setting

This study was set up in accordance with the STROBE statement for reporting of cohort studies.[10] It was designed to extract a minimum set of casemix variables and validate them internally. Subsequently, variables included in the casemix model were compared between hospitals and the casemix model and its variables were compared with existing mortality risk prediction models externally validated in the DSAA.

Patients and data source

Patients who had undergone surgery between 2013 and 2016 for primary juxtarenal or infrarenal abdominal aortic aneurysm, both elective and acute, were prospectively registered in the DSAA and included for analysis. Details of the DSAA have been published previously.[11, 12] The DSAA data of 2015 were verified over a randomly selected group of hospitals.[13] Where data regarding date of surgery, date of birth, operative setting/urgency (elective - or acute, symptomatic - or ruptured aneurysm), type of procedure (EVAR: endovascular aneurysm repair or OSR: open surgical repair) or mortality were missing, patients were not included for further analysis. In the Netherlands, the minimum volume per hospital was set at 20 operations per year and hospitals where fewer than 60 patients had been registered over a four-year period were excluded from analysis.

Primary outcome

The primary outcome was in-hospital or 30-day mortality. Subgroups of elective (EAAA) and acute operations (AAAA), based on symptomatic (SAAA) or ruptured AAA (RAAA) patients, were analysed separately.

Statistical analysis

Patients in whom EVAR had been converted to OSR were analysed following intention-to-treat and included with EVAR. First, baseline characteristics were analysed for the overall group (AAA) for EAAA and AAAA surgery. Continuous variables were tested for normality and linearity. Subsequently, if not normal or linear, variables were analysed in categories. Missing or unknown values for categorical variables were estimated using multiple imputations. Multiple imputations were performed by an iteration of 10 datasets using the automatic imputation method in SPSS (version 23.0) for the following variables; cardiac status, pulmonary status, malignant comorbidity, Glasgow Coma Scale (GCS), electrocardiography (ECG), sodium, potassium, creatinine, hemoglobin, white blood count (WBC), pulse, aneurysm size, age, gender, blood pressure, and three indicator variables; year of surgery, hospital and setting.

Univariable analysis was performed to identify variables associated with mortality ($p<0.05$). Casemix variables were analysed in a multivariable logistic regression enter model with backward manual selection to reduce the chance of overfitting.[14] A selection p-value of $p<0.1$ was used to reduce the set of variables to as few as possible.[9]

Hospital variation

Those casemix variables selected for multivariable analysis were also studied for between-hospital variation as if no variation is present, casemix correction would not be of great importance. By means of calculating continuous variables into dichotomous variables by the mean, and dichotomising categorical data into the presence or absence of a certain patient characteristic, the percentages were analysed by hospital. Significant variation was reached if hospital percentages extended beyond the 95% confidence intervals (CIs).

Model validation

Internal validation was performed by 100% apparent validation in which the population used for the development of the model is also used for internal model validation.[9]

External validation of the overall AAA group and the EAAA and AAAA subgroups was performed with three standard mortality risk prediction models (VBHOM, GAS and DAS [4-6]; Table 1). Two models were selected from an earlier systematic review, one based on laboratory values (VBHOM), and one based on clinical parameters (GAS). The third model was a newly-validated Dutch model for RAAA surgery (DAS). Model performance was analysed using the C-statistic and Hosmer – Lemeshow (H-L) tests for both the discrimination and calibration of these models. An AUC (C-statistic) of ≥0.7 described a moderate discriminative ability and ≥0.8 a good discriminative ability.

Table 1. Arithmetic formulas of mortality risk scores according to VBHOM, GAS and DAS

Model	Model formula
VBHOM[#]	−2·257 + (0·1511 * male) + (0·9940 * mode of admission) + (0·05923 * age (continuous in years)) + (0·001401 * serum urea (continuous mmol/l)) – (0·01303 * sodium (continuous mmol/l)) – (0·03585 * potassium (continuous mmol/l)) – (0·2278 * haemoglobin (continuous g/dl)) + (0·02059 * white cell count (continuous * 10^9 /l))
GAS	GAS: age (years) + (17 for shock) + (7 for myocardial disease) + (10 for cerebrovascular disease) + (14 for renal disease).
DAS[#]	-4.73 + (age * 0.074) + (systolic blood pressure [mm Hg]/10 * -0.12) + (1 for cardiopulmonary resuscitation) + (((haemoglobin [g/dL]/10) ^{3}) * -1.27).

[#] To calculate mortality risk use exp(model)/1+exp(model), Legend: GAS; Glasgow Aneurysm Score, VBHOM; Vascular Biochemistry and Haematology Outcome Models, DAS; Dutch Aneurysm Score

P-values ≥0.05 for the H-L showed sufficient calibration; the expected outcome did not significantly differ from the observed outcome.

RESULTS

Patients

A total of 13 417 patients with an AAA were registered in the DSAA, of which 12 524 (93.3%) had a primary AAA for which either open surgical repair (OSR) or endovascular repair (EVAR) was performed. In total, 99.1% (n=12416) of these patients were analysed. Two hospitals that had only performed 12 and 3 operations, respectively, over a 4-year period were excluded. Of the remaining 12 401 patients there were 9537 elective EAAA patients (76.9%), 913 (7.4%) acute symptomatic AAA (SAAA) patients and 1951 (15.7%) patients with an acute ruptured AAA (RAAA). (Figure 1)

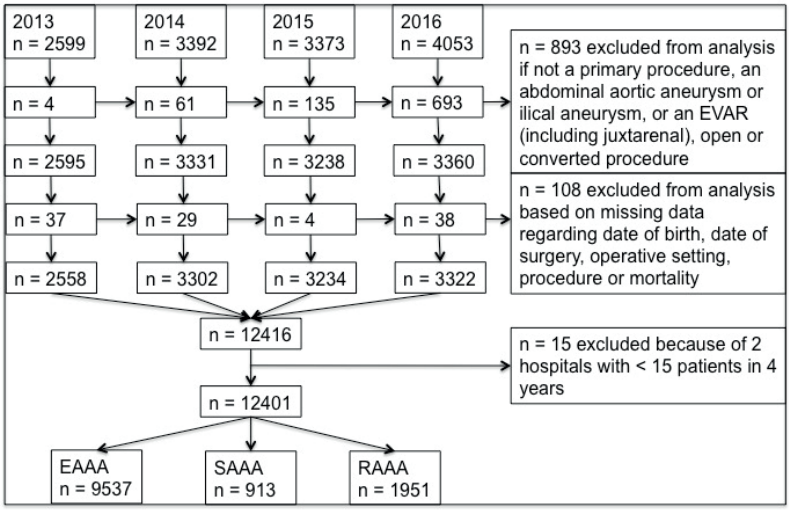


Figure 1. Flowchart of selected data

Overall, 8614 patients (69.5%) underwent EVAR, compared with 3787 (30.5%) undergoing OSR. The percentage of EVAR in EAAA patients was 77.1%, in SAAA patients 60.2% and in RAAA patients this was 36.4%. Mean age was 73.2 (7.9SD) years and the majority were male (85.5%, n=10 596). Detailed information about baseline characteristics, disease specifics and interventions can be found in Table 2 (both the original and the imputed dataset).

Table 2. Baseline variables in the original dataset, in the imputed dataset, and the Odds Ratio of variables in the imputed dataset that are associated with mortality. (n=12401 patients)

		Original dataset N (%)	Imputed dataset N (%)	Imputed OR for mortality (95%CI)
Age (mean (SD))		73.2 (7.9)	73.2	**1.07 (1.06-1.08)**
Sex:	Male	10596 (85.4)	10604 (85.5)	Ref.
	Female	1795 (14.5)	1797 (14.5)	**1.50 (1.26-1.80)**
	Missing	10 (0.1)	-	-
Card.:	No cardiac problems	5699 (46.0)	6112.1 (49.3)	Ref.
	Peripheral oedema	958 (7.7)	1024.9 (8.3)	1.12 (0.95-1.31)
	Elevated CVP	190 (1.5)	206.4 (1.7)	**1.56 (1.18-2.05)**
	Medication*	4769 (38.5)	5057.6 (40.8)	**1.83 (1.05-3.20)**
	Unknown	785 (6.3)	-	-
Pulm.:	No dyspnoea	8882 (71.6)	9333.2 (75.3)	Ref.
	Dyspnoea exercise	2432 (19.6)	2544.7 (20.5)	**1.45 (1.20-1.76)**
	Invalidating dyspnoea	357 (2.9)	375.7 (3.0)	**1.85 (1.29-2.66)**
	Dyspnoea rest	136 (1.1)	147.4 (1.2)	**3.84 (2.44-6.04)**
	Unknown	594 (4.8)	-	-
Mal:	No malignancy	10127 (81.7)	10236.3 (82.5)	Ref.
	Malignancy	2143 (17.3)	2164.7 (17.5)	0.86 (0.71-1.05)
	Unknown	131 (1.1)	-	-
GCS:	15	11076 (89.3)	11375.8 (91.7)	Ref.
	12-14	290 (2.3)	442.8 (3.6)	**9.81 (2.81-34.26)**
	9-11	65 (0.5)	330.3 (2.7)	7.90 (0.48-130.8)
	<9	119 (1.0)	252.1 (2.0)	**15.26 (1.87-124.47)**
	Unknown/missing	851 (6.9)	-	-
Aneurysm size (mean (SD)(mm))		63.14 (14.2)	63.30	**1.04 (1.04-1.05)**
Setting:	EAAA	9537 (76.9)	-	Ref.
	SAAA	913 (7.4)	-	**3.83 (2.83-5.18)**
	RAAA	1951 (15.7)	-	**22.87 (19.11-27.36)**
ECG:	No abnormalities	6260 (50.5)	7329.9 (59.1)	Ref.
	Atrial Fibrillation	802 (6.5)	965.6 (7.8)	**2.23 (1.70-2.93)**
	MI or other	3432 (27.7)	4105.5 (33.1)	**2.01 (1.63-2.48)**
	Unknown	1907 (15.4)	-	-
Creatinine	[normal, 45-100]	7579 (61.1)	7707.4 (62.2)	Ref.
	[not normal, <45 or >100]	4455 (35.9)	4693.6 (37.8)	**3.18 (2.73-3.70)**
	Unknown	367 (3.0)	-	-

Table 2. Baseline variables in the original dataset, in the imputed dataset, and the Odds Ratio of variables in the imputed dataset that are associated with mortality. (n=12401 patients) (*continued*)

		Original dataset N (%)	Imputed dataset N (%)	Imputed OR for mortality (95%CI)
Sodium:	[normal, 135-145]	10348 (83.4)	11227.6 (90.5)	Ref.
	[not normal, <135 or >145]	1035 (8.3)	1173.4 (9.5)	**2.77 (2.29-3.35)**
	Unknown	1018 (8.2)	-	-
Potassium:	[normal, 3.5-5.0]	10593 (85.4)	11279.1 (91.0)	Ref.
	[not normal, <3.5 or >5.0]	1025 (8.3)	1121.9 (9.0)	**2.39 (1.97-2.91)**
	Unknown	783 (6.3)	-	-
WBC (*10^9) (mean (SD))		9.21 (3.24)	9.06	**1.23 (1.20-1.25)**
Haemoglobin (mmol/l) (mean (SD))		8.43 (1.20)	8.43	**0.50 (0.48-0.53)**
SBP (mmHg)		135.95 (26.5)	135.77	**0.97 (0.97-0.97)**
Pulse:	[normal, 60-100 bpm]	9206 (74.2)	9905 (79.9)	Ref.
	[not normal, <60 or >100 bpm]	2279 (18.4)	2496 (20.1)	**1.81 (1.54-2.13)**
	Unknown	916 (7.4)	-	-

*Hypertension, angina pectoris, diuretics or digoxin

Outcome

Overall mortality was 6.5% (n=809): EAAA surgery 1.8% (n=172), 6.6% (n=60) for SAAA surgery and 29.6% (n=577) for RAAA surgery. Mortality for AAAA surgery (combined SAAA and RAAA) was 22.2%. By procedure, elective procedures had the lowest mortality (0.7% EVAR, 5.4% OSR) compared with mortality following symptomatic procedures, (4.5% EVAR and 9.6% OSR) and acute ruptures (22.2% and 33.8%).

Main results

Univariable analysis

Table 2 shows the association of baseline characteristics with the outcome mortality for the imputed dataset. Overall, age, gender, cardiac comorbidity, pulmonary comorbidity, GCS, aneurysm size, setting, electrocardiography (ECG), creatinine, sodium, potassium, white blood count (WBC), systolic blood pressure (SBP), pulse rate and hemoglobin were associated with mortality. Subgroups of special interest and of influence on mortality were a decreased GCS 12-14, (OR 9.81 (95% CI;2.81-34.26)) and <9 (OR 15.26 (95% CI;1.87-124.47)) and urgent setting RAAA (OR 22.87 (95% CI;19.11-27.36)).

Multivariable analysis

Overall, independent variables associated with mortality were age, gender, pulmonary comorbidity, operative setting, GCS, systolic blood pressure, ECG, hemoglobin and creatinine. (Table 3). The strongest overall predictors of mortality were increased pulmonary comorbidity, GCS and setting (RAAA). Tables 3b and 3c show subgroup analyses for EAAA and AAAA patients. Potassium, aneurysm size and malignancy were additional independent factors associated with mortality.

Table 3. Final multivariable analysis model with the Odds Ratio for the entire AAA group and elective and acute AAA separately.

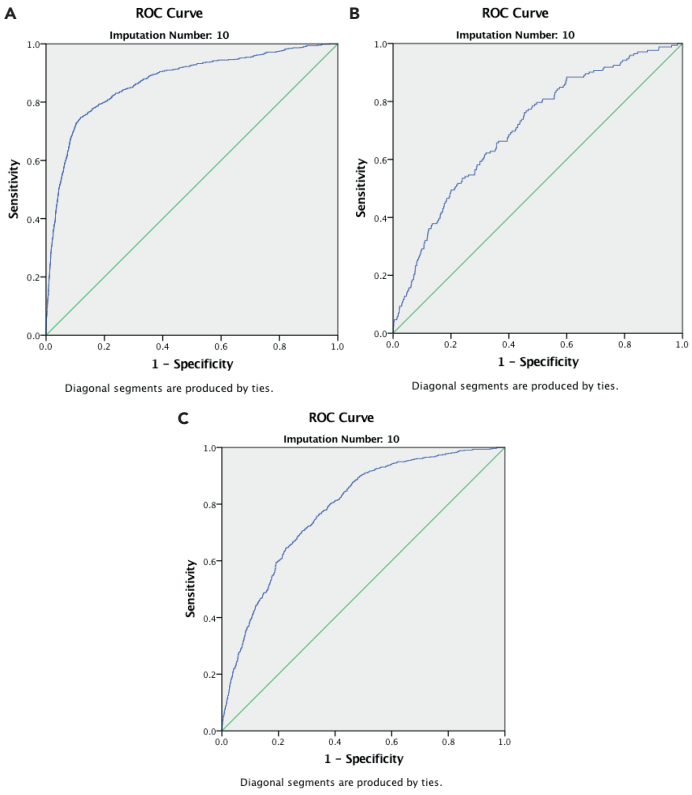
	3a. AAA overall	3b. EAAA subgroup	3c. AAAA subgroup
	Beta-coefficients	Beta-coefficients	Beta-coefficients
Age (years)	.045		.055
Gender [male]	-.398	-.677	-.323
Pulmonary comorbidity [dyspnoea during exercise]	.509	.707	.364
Pulmonary comorbidity [invalidating dyspnoea]	.587	.338	.681
Pulmonary comorbidity [dyspnoea in rest]	.970	1.420	.702
Ruptured AAA	2.519		
Symptomatic AAA	1.279		-1.285
Glasgow Coma Scale [12-14]	.589		.583
Glasgow Coma Scale [9-11]	1.123		1.239
Glasgow Coma Scale [<9]	1.182		1.296
Haemoglobin	-.113	-.289	
ECG [Atrial fibrillation]	.387	.393	.397
ECG [ischemia or other]	.322	.391	.294
Creatinine	.398	.304	0.469
Systolic Blood Pressure	-.005		-.007
Potassium		.518	
Aneurysm Size		.019	
Malignancy			.340
Intercept	-5.910	-2.748	-4.909
AUC	**0.871**	**0.703**	**0.785**
H-L	**0.198**	**0.476**	**0.109**

Hospital variation

Significant interhospital variation was observed for all variables included in the model, with the exception of gender.

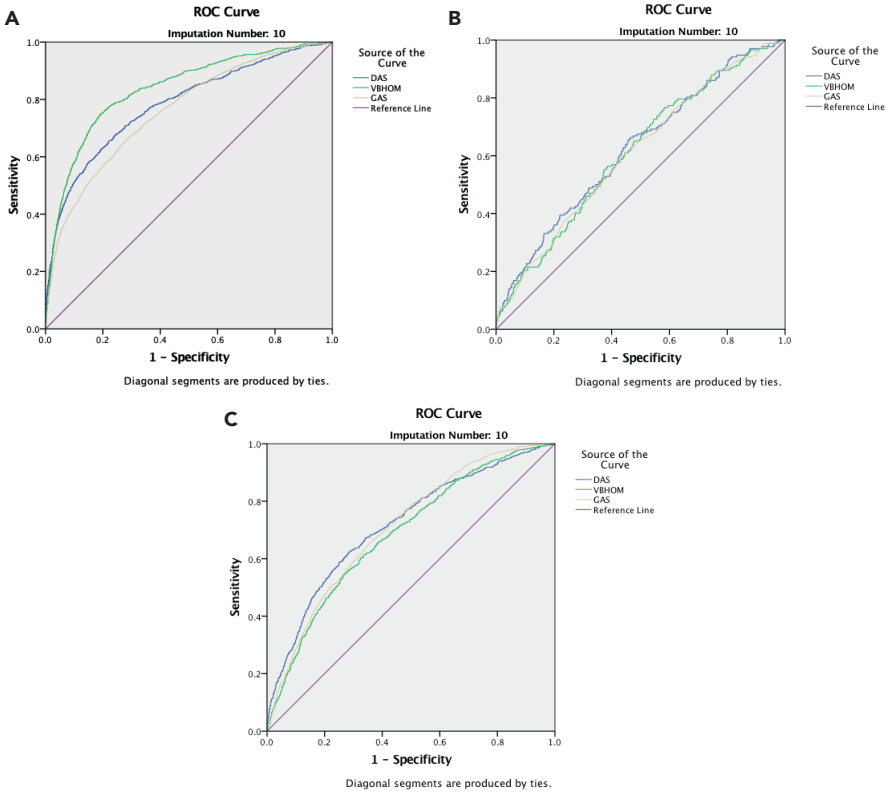
Model validation

Internal validation, for the overall casemix model, showed a good pooled calibration (H-L of 0.198) and good C-statistic of 0.871. For the elective submodel, the pooled C-statistic was 0.703 with an H-L of 0.476. For the acute submodel the pooled C-statistic was 0.785 with an H-L of 0.109. (Table 3 and Figures 2a-c)



Figures 2A-C. Tenth iteration of ROC curves following internal validation of the DSAA casemix model. (A) overall AAA, (B) EAAA, (C) AAAA

External validation revealed that overall the VBHOM had the highest discriminative performance with a C-statistic of 0.836 and H-L of 0.028. Followed by the DAS with a C-statistic of 0.782 and H-L of 0.000. Validation of the GAS resulted in a pooled C-statistic of 0.761 and a H-L of 0.128. For the subgroup of EAAA, pooled performance of GAS was 0.608 and H-L 0.645, of VBHOM 0.612 and H-L 0.614 and of DAS 0.622 with a H-L of 0.456. Performance of these models for AAAA surgery was 0.711 and H-L 0.218, 0.687 and H-L 0.448, 0.716 and H-L of 0.004, respectively. See figures 3a-c for the AUC curves by setting.



Figures 3A-C.Tenth iteration of ROC curves following external validation of VBHOM, GAS, DAS. (A) overall AAA, (B) EAAA, (C) AAAA

DISCUSSION

Key results

The current DSAA data set was based on V-POSSUM, of which the V(p)-POSSUM was regarded as being the casemix adjustment model for outcome comparison between hospitals. Following thorough investigation, a limited number of casemix variables to

decrease the registration burden was arrived at. A mix of easily collectible variables was identified including patient identifiers (age and sex), physiological variables (cardiac comorbidity represented by ECG, pulmonary comorbidity, GCS (in RAAA patients) and SBP), setting (EAAA, SAAA and RAAA), anatomical findings (AAA diameter in EAAA patients) and laboratory results (creatinine, hemoglobin). Calibration for all three models varied widely for the population of the DSAA. However, DAS calibrated well for EAAA patients only compared with VBHOM that did not calibrate well overall and GAS that had a good calibration for all groups of patients.

Overall, many variables had a significant association with mortality after both uni- and multivariable logistic regression. However, previous studies have shown that casemix adjustment has a limited effect on the observed difference between hospitals.[11, 15] While, patient casemix seemed to influence outcome, it did not explain - or only in part - the observed differences between hospitals.[15, 16] These interhospital differences could also be related to differences in structural - and process factors, by patient selection or by the proficiency of the surgical team.[15-17] Therefore, risk adjustment by casemix seems to have little impact on outcome differences between hospitals when compared with no risk adjustment.[17] A more complex casemix, which includes older patients with more comorbidity, may be counterbalanced by the continuous improvement in quality of health care.[18] However, maintaining risk adjustment by means of a limited set of patient casemix variables will remain necessary to moderate potential discussion among hospital stakeholders regarding differences in outcomes.

Some variables were more predictive for mortality by operative setting. For example, hemodynamic parameters and GCS were particularly important factors in acute AAA surgery, while comorbidity and AAA morphology were more associated with mortality in elective AAA surgery. Consequently, the observed differences between the mortality risk prediction models seemed to be partially related to the population analysed and the period of development. For example, DAS was built recently on a Dutch RAAA population, while GAS and VBHOM were developed over 10 years ago in an overall AAA population, when EVAR had just started to become common practice.[3, 19] Consequently, DAS had the best discriminative performance in AAAA surgery on comparison with VBHOM and GAS. However, calibration by H-L of DAS was very significant indicating a low generalisability of the population analysed. This could be due to the fact that SAAA patients were also included in the AAAA cohort of the DSAA, resulting in a lower mortality than that in the RAAA population in which DAS was developed. Moreover, mortality risk prediction only explains the association of the variables with mortality and need to be included in the model, whereas casemix variables that are associated

with mortality and which do not differ between hospitals not necessarily need to be adjusted for.

Missing data in our study were resolved by multiple imputations. Another option to handle missing data would be to allocate missing values towards the 'normal' category under the assumption 'if not registered then it may not be present at all' as in V-POSSUM for example.[7, 20] Proper handling of missing values is important in prediction models. Automatic transfer (IT links to the electronic patient file) of hospital data and of expense claims from other specialists treating co-morbidities, to the web – based vascular registry (DSAA) will improve registry compliance and the validity of the data.

Limitations

In the Netherlands, the DSAA is mandatory and in 2015 was validated in 15 randomly selected hospitals.[13] There were no significant registration flaws. Only some minor complications were not registered. Important factors for casemix correction and mortality were not missed. However, there could have been under-registration of patients or other data could have been missed per individual patient. Another limitation of this study is that the DSAA contains a limited set of casemix variables, based on V-POSSUM, and therefore it is possible that other relevant variables were disregarded. However, risk adjustment will always be limited to a fixed set of variables leaving immeasurable confounders unadjusted for.[21]

CONCLUSION

It was possible to establish a compact set of 10 variables, i.e. age, sex, cardiac co-morbidity, pulmonary comorbidity, GCS, SBP, setting, aneurysm size, creatinine and hemoglobin for casemix correction in AAA surgery in the DSAA. Preoperative casemix variables associated with mortality can be found in existing mortality risk prediction models, such as GAS and VBHOM, but when performing casemix correction they should be extracted from the dataset under analysis and ideally differ between hospitals.

ACKNOWLEDGEMENTS

The authors would like to thank the members of the Dutch Society for Vascular Surgery who registered their patients in the DSAA, the Dutch Institute for Clinical Auditing that facilitated the registry and the Dutch Surgical Aneurysm Audit group that consists of the following members: Dr. B.H.P. Elsman, Prof. J.F. Hamming, Dr. R.H. Geelkerken,.

E.D. Wilschut, Dr. A.C. Vahl, Prof. G.J. de Borst, Dr. J.A. Van Herwaarden, Dr. J.W. Elshof, Dr. M.C.M. Willems.

Conflicts of interest: None.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Lijftogt N, Vahl AC, Wilschut ED, et al. Adjusted Hospital Outcomes reported in the Dutch Surgical Aneurysm Audit after 2 years of Abdominal Aortic Aneurysm surgery. Eur J Vasc Endovasc Surg. 2017 Apr;53(4):520-532.
2. Beck N, Hoeijmakers F, van der Willik EM, et al. National Comparison of Hospital Performances in Lung Cancer Surgery: The Role of Case Mix Adjustment. Ann Thorac Surg 2018;106:412-420.
3. Lijftogt N, Luijnenburg TWF, Wilschut ED, et al. Systematic review of mortality risk prediction models in the era of endovascular abdominal aortic aneurysm surgery. Br J Surg. 2017 Jul;104(8):964-976.
4. Samy AK, Murray G, MacBain G. Glasgow aneurysm score. Cardiovasc Surg 1994;2:41-44.
5. Tang T, Walsh SR, Prytherch DR, et al. VBHOM, a data economic model for predicting the outcome after open abdominal aortic aneurysm surgery. Br J Surg 2007;94:717-721.
6. Von Meijenfeldt GCI, van Beek SC, Bastos Gonçalves F, et al. Development and external validation of a model predicting death after surgery in patients with a ruptured abdominal aortic aneurysm; the Dutch Aneurysm Score. Eur J Vasc Endovasc Surg 2017;53:168-174.
7. Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. Br J Surg 2001;88:958-963.
8. Meguid RA, Bronsert MR, Juarez-Colunga E, et al. Surgical Risk Preoperative Assessment System (SURPAS): III. Accurate Preoperative Prediction of 8 Adverse Outcomes Using 8 Predictor Variables. Ann Surg 2016 Jun;263(6):1042-8.
9. Steyerberg E. Clinical prediction models; A Practical Approach to Development, Validation, and Updating. : Springer; 2009.
10. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg 2014;12:1495-1499.
11. Lijftogt N, Vahl AC, Wilschut ED, et al. Adjusted Hospital Outcomes of Abdominal Aortic Aneurysm Surgery Reported in the Dutch Surgical Aneurysm Audit. Eur J Vasc Endovasc Surg 2017;53:520-532.
12. Karthaus E, Lijftogt N, Busweiler L, et al. Textbook Outcome: a Composite Measure for Quality of Elective Aneurysm Surgery. Ann Surg 2017 Nov;266(5):898-904.
13. The Dutch Institute for Clinical Auditing. Available at: http://dica.nl/media/660/Eindrapport_dataverificatie_DSAA_2016.pdf.
14. Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11:e1001744.
15. Mani K, Venermo M, Beiles B, et al. Regional Differences in Case Mix and Peri-operative Outcome After Elective Abdominal Aortic Aneurysm Repair in the Vascunet Database. Eur J Vasc Endovasc Surg 2015 Jun;49(6):646-652.
16. Lekander I, Willers C, Ekstrand E, et al. Hospital comparison of stroke care in Sweden: a register-based study. BMJ Open 2017;7:e015244-2016-015244.
17. Mayer EK, Bottle A, Aylin P, et al. What is the role of risk-adjusted funnel plots in the analysis of radical cystectomy volume-outcome relationships? BJU Int 2011;108:844-850.
18. Mani K, Bjorck M, Lundkvist J, et al. Improved long-term survival after abdominal aortic aneurysm repair. Circulation 2009;120:201-211.

19. Patterson B.O., Karthikesalingam A., Hinchliffe R.J., et al. The Glasgow Aneurysm Score does not predict mortality after open abdominal aortic aneurysm in the era of endovascular aneurysm repair. J Vasc Surg. 2011 Aug;54(2):353-7.
20. Prytherch DR, Ridler BM, Beard JD, et al. A model for national outcome audit in vascular surgery. Eur J Vasc Endovasc Surg 2001;21:477-483.
21. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med 2014;33:2297-2340.

5

High postoperative mortality after elective open surgical abdominal aortic aneurysm repair in hospitals with EVAR preference and lessons from an instrumental variable analysis of the Dutch Surgical Aneurysm Audit

N. Lijftogt
A. C. Vahl
E. G. Karthaus
E. M. van der Willik
S. Amodio
E. W. van Zwet
J. F. Hamming

in collaboration with the Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing

Submitted

ABSTRACT

Introduction:

Mortality following AAA surgery may have changed due to the increased application of endovascular aneurysm repair (EVAR) at the expense of open surgical repair (OSR), along with hospital preference. Aim of this study was to investigate the effect of hospital preference for EVAR on mortality by using an instrumental variable (IV) analysis.

Material and Methods:

Primary elective infra-or juxtarenal aneurysm (EAAA) repairs registered in the Dutch Surgical Aneurysm Audit (DSAA) (2013-2017) were analysed. Mortality in hospitals with higher preference (high%EVAR) for EVAR was compared with hospitals with lower preference (low%EVAR) divided by the median percentage of EVAR. The RD between EVAR and OSR was determined by: unadjusted and adjusted linear regression analysis, propensity score (PS) analysis, and IV analysis, adjusting for unobserved confounders, using EVAR percentage by hospital as IV-instrument.

Results:

11997 EAAA patients were analysed. The overall mortality RD between high% and low%EVAR hospitals was 0.1% (95% CI-0.5-0.4). OSR mortality was significantly higher in high%EVAR hospitals compared to low%EVAR hospitals: 7.3% versus 4.0% (RD 3.3%; 95% CI1.4-5.3). EVAR mortality was 0.9% versus 0.7% (RD 0.2%; 95% CI-0.02-0.6). The RD following unadjusted, adjusted and PS analysis was 4.2% (95% CI3.7-4.8), 4.4% (95% CI3.8-5.0) and 4.7% (95% CI 4.1-5.3) in favour of EVAR. RD following IV analysis was 1.3 (95% CI-0.9-3.6).

Conclusion:

Overall, in high% and low%EVAR hospitals mortality was not significantly different. Adjusting for observed confounders EVAR had significantly lower postoperative mortality compared to OSR. However, RD was low and non-significant following IV analysis. High- compared to low%EVAR hospitals had significantly higher postoperative mortality following OSR.

INTRODUCTION

Background

Postoperative mortality in elective abdominal aortic aneurysm (EAAA) surgery has decreased significantly since the introduction of EndoVascular Aneurysm Repair (EVAR).[1] A meta-analysis of four historic randomised trials reported the odds ratio for mortality following EVAR compared with Open Surgical Repair (OSR) as low as 0.40.[2] Also in the mandatory registry Dutch Surgical Aneurysm Audit (DSAA) there is a difference in mortality between EVAR and OSR (0.9% and 5.0%, respectively)[1, 3] When compared with the earlier DREAM (Dutch Randomised Endovascular Management) trial in the Netherlands, this risk difference (RD) between EVAR and OSR was 3.4%, i.e. EVAR (1.2%) and OSR (4.6%).[4] The question arises whether in the Netherlands the mortality of OSR is changed due to extension of indications for EVAR and how the mortality difference in clinical trials relates to the difference in observational studies reflecting practice in general. Trials include a selected group of patients and therefore might not reflect the real world.[5] However, in observational data from national registries, comparisons may be biased due to both measured and unmeasured confounders.[6] Confounding by indication for instance occurs when the choice for a specific treatment is influenced by the characteristics and comorbidity of patients and preferences of both patients and surgeons.[7] Change in patient selection and technical skills over time could explain the observed increased RD in mortality between EVAR and OSR in the DSAA, compared with the randomised trials.

Standard statistical methods for the adjustment of measured confounders are multivariable regression analysis and propensity score analysis. However, these methods do not adjust for variables that are not or cannot be measured, such as the interpretation of the anatomical characteristics of the aneurysm or the preference for one surgical procedure or the other. Possible techniques to adjust for unmeasured confounders are the instrumental variable (IV) analysis or ecological analysis.[6, 8]

Objective

In this study the mortality in hospitals with a high preference for EVAR (high%EVAR) was compared with low EVAR preference hospitals (low%EVAR). First, the overall RD between EVAR and OSR was determined by conventional univariate and multivariate linear regression analysis and propensity score analysis, all adjusting for observed confounders. Then, the current RD in postoperative mortality between EVAR and OSR, adjusted for observed and unobserved confounders, was determined following IV analysis.

MATERIAL AND METHODS

Study design

This observational study was conducted in accordance with the STROBE checklist (www.strobe-statement.org). First, the effect of preference for EVAR on EAAA mortality at hospital level was examined and, secondly, the RD in postoperative mortality between EVAR and OSR at patient level was calculated.

Data source and participants

Consecutive patients registered in the DSAA, who underwent operation for a primary infra- or juxtarenal EAAA between 2013 and 2017, were included for analysis.[3] The DSAA is a mandatory audit and registers all patients with an aortic aneurysm undergoing surgical treatment in the Netherlands. Data verification was performed in 2015.[9] Patients were non- eligible for and excluded from analysis if data on date of birth, date of surgery, survival state, emergency setting or type of procedure (EVAR or OSR) were missing. Hospitals performing less than 15 procedures in five years were excluded from analysis.

Outcomes

The primary outcome of this study was postoperative mortality, which was defined as mortality within 30 days after surgery or during the initial admission period (30-day/ in-hospital).

Statistical analysis

In order to investigate the effect of hospital preference for EVAR on postoperative mortality, hospitals were divided into two groups: hospitals with a relatively high percentage(%) of EVAR (> median %) and hospitals with a relatively low% EVAR (< median %). The RD on postoperative mortality (%) between patients treated with EVAR and OSR was determined in 4 ways: a linear model unadjusted for confounders, a linear model adjusted for observed confounders, a propensity score and an IV analysis adjusted for unobserved confounders. Patient characteristics and hospital-related factors were compared using the t-test and chi-square test.

High%EVAR hospitals versus low%EVAR hospitals

Mortality rates in hospitals with a high%EVAR were compared with low%EVAR hospitals. The percentage of AAA patients that was treated with EVAR per hospital (%EVAR) (i.e. ''treatment preference'' of the hospital) was used as instrumental variable for further analysis. We divided all hospitals into two groups with the median %EVAR per hospital as a cut-off point, i.e. those with a low%EVAR and those with a high%EVAR in

AAA patients. We demonstrated the distribution of measured possible confounders between these two groups. Then, results of these patient groups were used in a calculation model, based on the postoperative results in our patient groups. In the model we hypothesised that certain patients will always get EVAR or OSR in every hospital. The remaining third ''marginal'' group were there is a choice, depends on hospital preference: all EVAR or all OSR.

Unadjusted linear regression analysis

Crude mortality rates in patients treated with EVAR and OSR were compared using a linear regression model. When considering a binary outcome, it is standard practice to use logistic regression. The effect of EVAR versus OSR will then be estimated as an odds ratio (OR). As we preferred to estimate the effect as an RD, we used linear regression.

Adjusted linear regression analysis

In order to correct for observed confounders, we used a linear regression model to compare adjusted mortality rates in patients treated with EVAR and OSR. Patient characteristics of influence on mortality were selected by univariable logistic regression (see Table 1). Then, the adjusted RD for mortality was analysed by multivariable linear regression analysis.

Propensity score risk adjustment

This was carried out in two successive steps. In the first step, a multivariable logistic regression analysis with every variable associated with 'choice of treatment' by univariable analysis was performed. In the second step the RD was estimated, using multivariable linear regression analysis for the primary outcome 'postoperative mortality', adjusted for the propensity score obtained in step 1 and the choice of treatment as predictors.

Instrumental variable analysis

For the IV analysis the proportion of patients treated with EVAR at each centre was used as an instrumental variable to adjust for unobserved confounders with the Two-Stage Least Squares (2SLS) method. We started by computing the proportion of EVAR in each hospital from the hospital identifier. Next, we fitted a model for mortality with the predicted probability of EVAR as the only co-variable.

An IV analysis can be used to estimate the effect of a treatment in observational data, corrected for unobserved confounders. An IV is a factor that strongly influences the choice of treatment, but which has no independent influence on patient outcome. Thus, an IV is not related to the prognosis of the patient. When carrying out IV analysis

individual patients with differing treatments are not compared, but rather the outcomes of patients with a different chance of getting a certain treatment. The methods of IV analysis are described in detail elsewhere.[6]

When using an IV analysis to compare mortality after OSR and EVAR in patients with an AAA, we made two assumptions, based on earlier results from the DSAA:

1. Patients with an AAA are randomly divided over all hospitals that perform AAA surgery in the Netherlands[1]
2. Quality of AAA-related care is equal in each hospital[1]

The strength of the IV was tested with the partial F-statistic. The co-variables used in this model were the same as in the first step of the propensity score analysis, except that the actual treatment is not in the model. Outcome was reported as an RD between EVAR and OSR.

All statistical analyses were performed using R statistical software (version 3.5.1) and SPSS (version 23.0).

RESULTS

Participants

A total of 12 350 patients were registered. Following application of the exclusion criteria, 12 009 (97.6%) of these patients were analysed. One hospital that registered 12 patients in 5 years was excluded, leaving 11 997 EAAA patients. (Figure 1) Verification in 2015 showed that the data in the DSAA are virtually complete and that no major events have been missed.[9]

Descriptive Data

Of these patients, 77.1% (n=9255) were treated with EVAR without conversion, 0.2% (n=24) were converted from EVAR to OSR and analysed in the EVAR group, and 22.7% (n=2718) were treated with OSR. The percentage of EVAR varied between hospitals (range 53.4-100) with a mean of 77.3% and median of 76.6%. There were 5961 patients in the high%EVAR group and 6036 in the low%EVAR group. The mean percentages of EVAR were 85.7% versus 69.1%, respectively; mean difference 16.6%. There were 28 high%EVAR hospitals (8 university/large teaching hospitals: 29%) and 34 low%EVAR hospitals (9 university/large teaching hospitals: 26%).

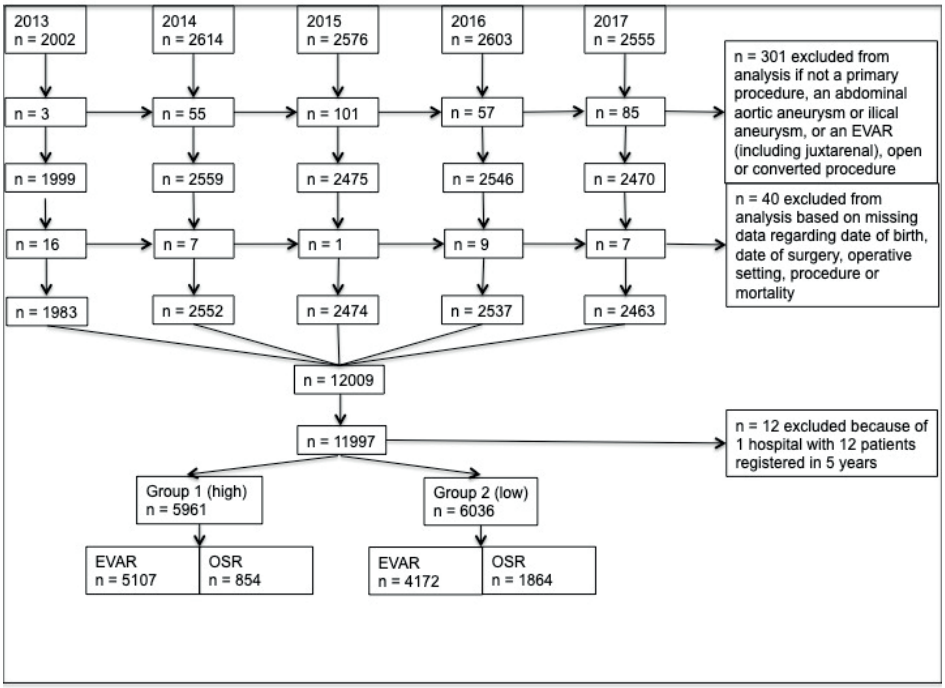


Figure 1. Flowchart of analysed and excluded patients.

Table 1 shows the baseline characteristics. Age, gender, cardiopulmonary comorbidity, malignancy, aneurysm size, electrocardiography (ECG) abnormalities, sodium, potassium, white blood count (WBC), systolic blood pressure (SBP) and haemoglobin differed significantly between EVAR and OSR patients. Summarizing, EVAR patients were older, more often men, had smaller AAA diameters and less comorbidities.

The mean overall OSR volume was 52 patients per hospital (range 0-118): 35 in high%EVAR hospitals (range 0-68) and 69 in low%EVAR hospitals (range 17-118). The mean overall EVAR volume was 185 patients per hospital (range 32-387): 205 in high%EVAR hospitals (range 101-382) and 164 (range 32-387) in the low%EVAR hospitals. Information about suprarenal clamping was registered in 1545 consecutive cases since 2015: 1011 patients in low%EVAR hospitals and 534 patients in high%EVAR hospitals. Of these patients, 256 (25.3%) and 167 (31.3%), respectively, underwent suprarenal clamping (OR 1.34, 95% CI 1.07-1.69, p=0.013).

Table 1. Baseline characteristics by type of procedure and hospital use of EVAR.

Variable	Type of procedure			Hospital use of EVAR		
	EVAR (n=9279)	OSR (n=2718)	Chi-square/ t-test*(p-value)	High percentage EVAR >median% (n=5961)	Low percentage EVAR <median% (n=6036)	Chi-square/ t-test* (p-value)
Age (mean (SD))	73.8 (7.5)	70.5 (7.6)	<0.001	73.3 (7.7)	72.8 (7.7)	0.001
Sex: Male	8086 (87.1)	2186 (80.4)	<0.001	5123 (85.9)	5149 (85.3)	0.567
Female	1190 (12.8)	530 (19.5)		836 (14.0)	884 (14.6)	
Missing	3 (0%)	2 (0.1)		2 (0)	3 (0.1)	
Card.: No cardiac problems	4191 (45.2)	1205 (44.3)	0.002	2952 (49.5)	2444 (40.5)	<0.001
Peripheral oedema	811 (8.7)	191 (7.0)		464 (7.8)	538 (8.9)	
Elevated CVP	153 (1.6)	35 (1.3)		59 (1.0)	129(2.1)	
Medication[#]	3797 (40.9)	1205 (44.3)		2250 (37.7)	2752 (45.6)	
Unknown	327 (3.5)	82 (3.0)		236 (4.0)	173 (2.9)	
Pulm.: No dyspnoea	6840 (73.7)	2015 (74.1)	0.035	4416 (74.1)	4439 (73.5)	<0.001
Dyspnoea exercise	1927 (20.8)	573 (21.1)		1265 (21.2)	1235 (20.5)	
Invalidating dyspnoea	302 (3.3)	60 (2.2)		188 (3.2)	174 (2.9)	
Dyspnoea rest	90 (1.0)	24 (0.9)		52 (0.9)	62 (1.0)	
Unknown/missing	120 (1.3)	46 (1.7)		40 (0.7)	126 (2.1)	
Mal.: No malignancy	7313 (78.8)	2286 (84.1)	<0.001	4777 (80.1)	4822 (79.9)	0.438
Malignancy	1904 (20.5)	412 (15.2)		1138 (19.1)	1178 (19.5)	
Unknown	62 (0.7)	20 (0.7)		46 (0.8)	36 (0.6)	
Aneurysm diameter (mm) (mean (SD))	59.2 (10.2)	61.7 (13.0)	<0.001	59.5 (10.8)	60.0 (11.0)	0.012

Table 1. Baseline characteristics by type of procedure and hospital use of EVAR. (*continued*)

Variable	Type of procedure			Hospital use of EVAR		
	EVAR (n=9279)	OSR (n=2718)	Chi-square/ t-test*(p-value)	High percentage EVAR >median% (n=5961)	Low percentage EVAR <median% (n=6036)	Chi-square/ t-test* (p-value)
ECG: No abnormalities	4902 (52.8)	1519 (55.9)	<0.001	3314 (55.6)	3107 (51.5)	<0.001
Atrial Fibrillation	684 (7.4)	149 (5.5)		399 (6.7)	434 (7.2)	
MI or other	2753 (29.7)	862 (31.7)		1656 (27.8)	1959 (32.5)	
Unknown	940 (10.1)	188 (6.9)		592 (9.9)	536 (8.9)	
Creatinin: [normal, 45-100]	6070 (65.4)	1766 (65.0)	0.206	3943 (66.1)	3892 (64.5)	<0.001
[not normal, <45 or >100]	3007 (32.4)	877 (32.3)		1926 (32.3)	1958 (32.4)	
Unknown	202 (2.2)	75 (2.8)		92 (1.5)	185 (3.1)	
Sodium: [normal, 135-145]	7903 (85.2)	2371 (87.2)	<0.001	5031 (84.4)	5243 (86.9)	<0.001
[not normal, <135 or >145,0]	464 (5.0)	148 (5.4)		304 (5.1)	308 (5.1)	
Unknown	912 (9.8)	199 (7.3)		626 (10.5)	485 (8.0)	
Potassium: [normal, 3.5-5.0]	8113 (87.4)	2374 (87.3)	0.001	5208 (87.4)	5279 (87.5)	0.983
[not normal, <3.5 or >5.0]	520 (5.6)	194 (7.1)		355 (6.0)	359 (5.9)	
Unknown	646 (7.0)	150 (5.5)		398 (6.7)	398 (6.6)	
WBC (*10^9) (mean (SD))	8.3 (1.9)	8.6 (2.1)	<0.001	8.4 (2.0)	8.4 (2.0)	0.501
Syst. Blood pressure (mmHg) (mean (SD))	140.0 (20.1)	141.1 (20.0)	0.007	139.6 (20.4)	140.9 (19.8)	<0.001
Pulse: [normal, 60-100 bpm]	7207 (77.7)	2129 (78.3)	0.478	4724 (79.2)	4612 (76.4)	<0.001
[not normal, <60 or >100 bpm]	1554 (16.7)	430 (15.8)		1013 (17.0)	971 (16.1)	
Unknown	518 (5.6)	159 (5.8)		224 (3.8)	453 (7.5)	
Haemoglobin (g/dl) (mean (SD))	8.7 (1.0)	8.6 (1.0)	0.006	8.7 (1.0)	8.7 (1.0)	0.089

[#] Hypertension, angina pectoris, diuretics or digoxin

Outcome data and main results

Outcome by hospital preference; EVAR versus OSR.

Table 2 describes the postoperative mortality in high% and low%EVAR hospitals with a RD of 0.1% (95% CI -0.5-0.4) which was not statistically significant. The mortality of 7.3% after OSR in high%EVAR hospitals was significantly higher than in low%EVAR hospitals: 4.0%. In order to understand differences in outcome related to hospital preference for EVAR, we analysed confounding variables for both OSR and EVAR by high and low%EVAR hospitals. (Table 3). We can conclude that in high%EVAR hospitals patients in both treatment groups have significantly less comorbidity. In other words, patients in high%EVAR hospitals seemed to be healthier.

Table 2. Mortality by high and low percentage EVAR hospitals for EVAR. Univariate analysis observed mortality in the DSAA by procedure by hospital preference; low% versus high% EVAR hospitals.

Crude analysis	Mortality					
	OSR		EVAR		Overall	
	Number	%	Number	%	Number	
High percentage EVAR	62 / 854	7.3%	46 / 5107	0.9%	108 / 5961	1.8%
Low percentage EVAR	75 / 1864	4.0%	29 / 4172	0.7%	104 / 6036	1.7%
	OR 1.87 (95%CI; 1.32-2.64)		OR 1.30 (95%CI; 0.81-2.07)		OR 1.05 (95%CI; 0.80-1.38)	

Using the mortality results in our cohort, a model is described in the appendix. In this model we suppose that in high%EVAR hospitals there are no patients with OSR, who could have been treated with EVAR. In low%EVAR hospitals there are some patients with OSR who would have been treated with EVAR in high%EVAR hospitals. According to this calculation the mortality of OSR in patients, where there is discussion of preference for EVAR or OSR, is 1.3% and, in case of EVAR the mortality is 1.7%. Patients in the ''marginal'' group seem to be better off with OSR.

Unadjusted analysis

The overall crude mortality was 1.8% (n=212): 0.8% (n=75) after EVAR and 5.0% after OSR (n=137): RD 4.2% (95% CI 3.7-4.8). The RD in high% compared with low%EVAR hospitals was 0.1% (95% CI -0.5-0.4): mortality 1.8% and 1.7%, respectively (Table 2).

Table 3. Comparison of baseline characteristics of EVAR and OSR in hospitals with a high %EVAR preference compared to a low % EVAR preference.

	OSR					EVAR				
	Low% EVAR n=1864		High% EVAR n=854		P-value	Low% EVAR n=4172		High% EVAR n=5107		P-value
Age (mean (SD))	70.7	7.7	70,3	7.4	0.225	73,8	7.5	73.8	7.6	0.928
Sex: Male	1510	81.0%	676	79.2%	0.314	3639	87.2%	4447	87.1%	0.904
Female	352	18.9%	178	20.8%		532	12.8%	658	12.9%	
Missing	2	0.1%	0	0.0%		1	0.0%	2	0.0%	
Card.: No cardiac problems	765	41.0%	440	51.5%	<0.001	1679	40.2%	2512	49.2%	<0.001
Peripheral oedema	147	7.9%	44	5.2%		391	9.4%	420	8.2%	
Elevated CVP	31	1.7%	4	0.5%		98	2.3%	55	1.1%	
Medication[#]	865	46.4%	340	39.8%		1887	45.2%	1910	37.4%	
Unknown	56	3.0%	26	3.0%		117	2.8%	210	4.1%	
Pulm,: No dyspnoea	1394	74.8%	621	72.7%	0.139	3045	73.0%	3795	74.3%	<0.001
Dyspnoea exercise	385	20.7%	188	22.0%		850	20.4%	1077	21.1%	
Invalidating dyspnoea	35	1.9%	25	2.9%		139	3.3%	163	3.2%	
Dyspnoea rest	14	0.8%	10	1.2%		48	1.2%	42	0.8%	
Unknown/missing	36	1.9%	10	1.2%		90	2.2%	30	0.6%	
Mal: No malignancy	1572	84.3%	714	83.6%	0.862	3250	77.9%	4063	79.6%	0.044
Malignancy	279	15.0%	133	15.6%		899	21.5%	1005	19.7%	
Unknown	13	0.7%	7	0.8%		23	0.6%	39	0.8%	
Aneurysm size (mm)	61,6	12.7	61,8	13.7	0.757	59,3	10.1	59,1	10.2	0.449

Table 3. Comparison of baseline characteristics of EVAR and OSR in hospitals with a high %EVAR preference compared to a low % EVAR preference. (*continued*)

	OSR					EVAR				
	Low% EVAR n=1864		High% EVAR n=854		P-value	Low% EVAR n=4172		High% EVAR n=5107		P-value
ECG: No abnormalities	994	53.3%	525	61.5%	<0.001	2113	50.6%	2789	54.6%	<0.001
Atrial Fibrillation	115	6.2%	34	4.0%		319	7.6%	365	7.1%	
Myocardial Infarction	612	32.8%	250	29.3%		1347	32.3%	1406	27.5%	
Unknown	143	7.7%	45	5.3%		393	9.4%	547	10.7%	
Creatinine [normal, 45-100]	1200	64.4%	566	66.3%	0.140	2693	64.5%	3377	66.1%	<0.001
[not normal, <45 or >100]	605	32.5%	272	31.9%		1353	32.4%	1654	32.4%	
Unknown	59	3.2%	16	1.9%		126	3.0%	76	1.5%	
Sodium: [normal, 135-145]	1640	88.0%	731	85.6%	0.222	3603	86.4%	4300	84.2%	0.001
[not normal, <135 or >145,0]	95	5.1%	53	6.2%		213	5.1%	251	4.9%	
Unknown	129	6.9%	70	8.2%		356	8.5%	556	10.9%	
Potassium: [normal, 3.5-5.0]	1633	87.6%	741	86.8%	0.503	3646	87.4%	4467	87.5%	0.977
[not normal, <3.5 or >5.0]	126	6.8%	68	8.0%		233	5.6%	287	5.6%	
Unknown	105	5.6%	45	5.3%		293	7.0%	353	6.9%	
WBC (*109)	8,5	2.1	8.7	2.2	0.169	8.3	1.9	8.4	1.9	0.168
Syst blood pressure (mmHg)	141.4	19.6	140.5	20.8	0.291	140.7	19.9	139.4	20.3	0.002
Pulse: [normal, 60-100 bpm]	1446	77.6%	683	80.0%	0.031	3166	75.9%	4041	79.1%	<0.001
[not normal, <60 or >100 bpm]	294	15.8%	136	15.9%		677	16.2%	877	17.2%	
Unknown	124	6.7%	35	4.1%		329	7.9%	189	3.7%	
Haemoglobin (g/dl)	8.6	1.0	8.6	1.0	0.879	8.7	1.0	8.7	1.0	0.162

Hypertension, angina pectoris, diuretics or digoxin

Adjusting for observed confounders

Potential confounding variables following univariable analysis for the outcome mortality were sex, age, cardiopulmonary state, ECG, AAA diameter, sodium, potassium, creatinine, haemoglobin and year of surgery. This resulted in an RD of 4.4% (95% CI 3.8-5.0).

Table 4. Propensity scores for treatment with EVAR.

Variables	Odds Ratio	95% Confidence Interval	
Female sex	.497	.438	.564
Age per year	1.069	1.062	1.076
2013	Ref.		
2014	1.215	1.050	1.406
2015	1.518	1.305	1.765
2016	1.369	1.180	1.588
2017	1.357	1.167	1.578
Percentage EVAR	1.064	1.059	1.069
Aneurysm size per mm	.971	.967	.975
Systolic Blood Pressure	.996	.994	.999
Haemoglobin (g/dl)	1.099	1.047	1.154
White Blood Count	.959	.937	.981
Normal Sodium (135-145)	Ref.		
High- or low Sodium (<135 or >145)	1.012	.822	1.246
Sodium Unknown	1.091	.811	1.468
Normal Potassium (3.5-5.0)	Ref.		
High- or low Potassium (<3,5 or >5,0)	.775	.644	.934
Potassium Unknown	1.262	.899	1.772
No Malignancy	Ref.		
Any Malignancy	1.298	1.146	1.471
Malignancy Unknown	.767	.442	1.331
Normal Electrocardiography	Ref.		
Atrial Fibrillation	1.085	.881	1.337
MI or any other deviating result	.877	.787	.977
Electrocardiography Unknown	1.349	1.127	1.615
No Cardiac Comorbidity	Ref.		
Oedema	1.145	.945	1.386
Elevated CVP	1.737	1.163	2.596
Medication for Hypertension	.950	.859	1.050
Cardiac Comorbidity Unknown	.874	.663	1.153

Table 4. Propensity scores for treatment with EVAR. (*continued*)

Variables	Odds Ratio	95% Confidence Interval	
No Pulmonary Comorbidity	Ref.		
Dyspnoea exercise	1.000	.891	1.122
Invalidating dyspnoea	1.568	1.159	2.120
Dyspnoea rest	1.467	.901	2.390
Pulmonary comorbidity Unknown	.909	.624	1.323

Propensity score

We fitted two models. First, a model to estimate the probability of EVAR, given the covariates selected by univariate analysis: the propensity score. (Table 4) Second, a model estimating the EVAR/OSR effect, adjusted for the propensity score. The EVAR/OSR effect was estimated at 4.7% (95% CI 4.1-5.3) in favor of treatment with EVAR.

Adjusting for unobserved confounders

Concluding from these results, the observed clinical variables hardly confound the effect of EVAR versus OSR. However, there are still unobserved confounders. To explain the principle of instrumental variables hospitals are divided into two categories according to the percentage of patients treated with a particular treatment: high%EVAR hospitals versus low%EVAR hospitals. This (dichotomised) percentage can then be used as an instrumental variable. We do note that this is much more crude (less informative) than using the actual percentage per hospital. The mean difference of EVAR treatment is 16.6% in high%EVAR hospitals versus low%EVAR hospitals (see descriptive data). The overall RD in postoperative mortality is 0.1% (table 2). So, the mortality advantage of EVAR compared to OSR is 0.6% (0.1%/0.166). This is a crude number, but much less then the RDs calculated from unadjusted and adjusted analysis.

Following IV analysis (2SLS with the %EVAR per hospital as the instrument) the EVAR versus OSR effect was now estimated at a RD of 1.3% (95% CI -0.9-3.6) (p=0.26), which is not significant anymore.

DISCUSSION

This study shows that hospitals with a preference for EVAR provide no benefit in postoperative mortality compared to hospitals with a lower preference for EVAR. The postoperative mortality after OSR is 7.3% in high%EVAR hospitals with a RD of 3.3% in favour of low%EVAR hospitals. The estimated RD of postoperative mortality

after EVAR compared with OSR is dependent on the analytical method used in large observational studies. After adjustment for known confounders, the RD was 4.4%. Following IV analysis we found a RD of 1.3% in favour of EVAR, which is not significant anymore. This seems to be a paradox. Does this mean that there is really no difference in postoperative mortality between EVAR and OSR?

This RD between EVAR and OSR with IV analysis in our cohort is less good then the RD in the Dutch DREAM trial: 3.4% (95% CI -0.3-7.7).[4] The results of IV analysis are expected to be somewhere between those of randomised controlled trials and observational data studies.[10] This is due to the fact that randomised controlled trials are optimised by patient selection and tend to overestimate the effect of clinical practice, while observational data might be subject to bias because of confounders that need to be adjusted for.[10] When the instrumental variable points are strong and valid, the RD would ideally be comparable to those presented in RCTs.[10] However, in the Netherlands, after correction for non-observed confounders in the IV analysis, EVAR results have a non-significant lower postoperative mortality compared to OSR. There may be 3 reasons for this observation: (1) by correcting for unobserved confounding, bias is removed and the effect is smaller, (2) by replacement of the actual treatment by the expected treatment, there is a major loss of information and therefore a loss of power and (3) choosing an IV, explained in the methods section, means that the hospital is a proxy for the choice of the treatment and that the outcome only depends on the (choice) of treatment and not by the hospital or practitioner. However, in IV analysis there may still be hidden bias at hospital level with regard to degree of surgical skill and the hospital infrastructure.[8] So, are there anyway hospitals with worse results?

The results from OSR were less good in high%EVAR hospitals compared to low%EVAR hospitals, while comorbidities were less, also in OSR patients. Unfortunately, the question whether a lower threshold for performing EVAR results in relatively more complex OSR procedures cannot be answered as this was not registered in the DSAA. However, the poorer result of OSR in hospitals with a relative preference for EVAR is not in line with the literature. A recent meta-analysis concluded that postoperative mortality after OSR in the pre-EVAR and post-EVAR eras is almost the same at around 2%.[11] Patients who underwent OSR in the post-EVAR era have a more complex anatomy, but were shown to be fitter, resulting in an unchanged overall postoperative mortality.[11] This increased complexity of OSR caused by more difficult anatomy, despite similar comorbidities, has occurred since more hospitals have chosen the endovascular procedure.[12] High%EVAR hospitals had a higher percentage of suprarenal clamping as compared to low%EVAR hospitals in our study with a difference of 6%. However, according to the literature, suprarenal clamping may result in increased morbidity but not increased

mortality.[13] Where suprarenal clamping is necessary, mortality is comparable with those patients undergoing OSR with infrarenal clamping.[14] On the basis of the IV analysis (rough calculation as well as 2SLS) and our calculation model we conclude that to apply a low threshold to EVAR does not lead to mortality profit for the entire patient group. Should we choose more often for OSR?

In the SWEDVASC[15] with relatively fewer patients undergoing EVAR (65%), the overall mortality (1.9%) is the same; the EVAR mortality is slightly higher (1.7%) and OSR mortality is much lower (2.4%) than our results. A low threshold for EVAR may also result in EVAR being carried out in relatively more complex cases, more chimneys (ChEVAR) and fenestrations (FEVAR).[11] This strategy also probably leaves relatively more complex cases in the OSR group, but unfortunately this could not be analysed in the DSAA. These more complex procedures approximate the results of OSR.[14, 16] The appendix explains the mathematical scenario in which patients are divided into 3 groups: patients in whom EVAR is always possible (N1) or is always impossible (N2) and a third, marginal, group where there is no agreement (N3). From this scenario, based on DSAA results, it may be concluded when in doubt, OSR might be the preferred treatment.

Limitations

This study has the same limitations as other observational studies. First, patient selection. The turndown rate may differ between hospitals, and there may be referral selection, meaning that choices of treatment might be different. However, in our data set the university and large teaching hospitals are equally divided between the groups of high and low%EVAR hospitals. High%EVAR centres do not have more observed comorbidities, however the severity of comorbidities is difficult to capture, even with the items included in the V-POSSUM.[17, 18] It is possible that patients with a difficult anatomy are referred, leading to higher mortality in the OSR group in high%EVAR hospitals. Another possible limitation are the missing values, but there were few of these in our study.

CONCLUSION

The mortality after AAA surgery is similar between hospitals with a higher preference for EVAR (high%EVAR hospitals) when compared with hospitals with a lower preference for EVAR (low%EVAR hospitals). In hospitals with a higher preference for EVAR the mortality after OSR is high and significantly higher than in hospitals with a relatively low preference for EVAR. We consider that a major concern. In classical adjusted analyses, with adjustment for known confounders, EVAR (on the patient level) results in a

significant lower in hospital and 30-day mortality, compared to OSR. However, after instrumental variable analysis, with adjustment for unknown variables, the risk difference is less good than in the randomised trials and not significant anymore. This means that the assumption for IV analysis ''the outcome of the chosen treatment should not depend on the hospital'' is subject to debate in The Netherlands.

ACKNOWLEDGEMENTS

The authors would like to thank the members of the Dutch Society for Vascular Surgery who registered their patients in the DSAA, the Dutch Institute for Clinical Auditing that facilitated the registry and the Dutch Surgical Aneurysm Audit group

Conflicts of interest: None.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1 Lijftogt N, Vahl AC, Wilschut ED, et al. Adjusted Hospital Outcomes of Abdominal Aortic Aneurysm Surgery Reported in the Dutch Surgical Aneurysm Audit. Eur J Vasc Endovasc Surg 2017;53:520-32.

2 Powell JT, Sweeting MJ, Ulug P, et al. Meta-analysis of individual-patient data from EVAR-1, DREAM, OVER and ACE trials comparing outcomes of endovascular or open repair for abdominal aortic aneurysm over 5 years. Br J Surg 2017;104:166-78.

3 The Dutch Institute for Clinical Auditing. Available at: www.clinicalaudit.nl.

4 Prinssen M, Verhoeven EL, Buth J, et al. A randomized trial comparing conventional and endovascular repair of abdominal aortic aneurysms. N Engl J Med 2004;351:1607-18.

5 Stansby G, Earnshaw J. Open repair versus endovascular stent graft repair of abdominal aortic aneurysms: do we need more randomized clinical trials? Angiology 2014;65:457-8.

6 Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med 2014;33:2297-340.

7 Steyerberg E. Clinical prediction models; A Practical Approach to Development, Validation, and Updating: Springer 2009.

8 Johnston SC. Combining ecological and individual variables to reduce confounding by indication: case study--subarachnoid hemorrhage treatment. J Clin Epidemiol 2000;53:1236-41

9 Available at: http://dica.nl/media/660/Eindrapport_dataverificatie_DSAA_2016.pdf.

10 Stukel TA, Fisher ES, Wennberg DE, et al. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. JAMA 2007;297:278-85.

11 Kontopodis N, Tavlas E, Georgakarakos E, et al. Has Anatomic Complexity of Abdominal Aortic Aneurysms Undergoing Open Surgical Repair Changed after the Introduction of Endovascular Treatment? Systematic Review and Meta-analysis of Comparative Studies. Ann Vasc Surg 2018 Oct;52:292-301.

12 Barshes NR, McPhee J, Ozaki CK, et al. Increasing complexity in the open surgical repair of abdominal aortic aneurysms. Ann Vasc Surg 2012;26:10-7.

13 Jongkind V, Yeung KK, Akkersdijk GJ, et al. Juxtarenal aortic aneurysm repair. J Vasc Surg 2010;52:760-7.

14 Dubois L, Durant C, Harrington DM, et al. Technical factors are strongest predictors of postoperative renal dysfunction after open transperitoneal juxtarenal abdominal aortic aneurysm repair. J Vasc Surg 2013;57:648-54.

15 Available at: http://www.ucr.uu.se/swedvasc/arsrapporter.

16 de Beaufort HWL, Cellitti E, de Ruiter QMB, et al. Midterm outcomes and evolution of gutter area after endovascular aneurysm repair with the chimney graft procedure. J Vasc Surg 2018;67:104,112.e3.

17 Prytherch DR, Ridler BM, Beard JD, et al. A model for national outcome audit in vascular surgery. Eur J Vasc Endovasc Surg 2001;21:477-83.

18 Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. Br J Surg 2001;88:958-63.

19. Patterson BO, Karthikesalingam A, Hinchliffe RJ, et al. The Glasgow Aneurysm Score does not predict mortality after open abdominal aortic aneurysm in the era of endovascular aneurysm repair. J Vasc Surg 2011;54:353e7.

20. Prytherch DR, Ridler BM, Beard JD, et al. A model for na- tional outcome audit in vascular surgery. Eur J Vasc Endo- vasc Surg 2001;21:477e83.
21. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med 2014;33: 2297e340.

6

Failure To Rescue - a closer look at mortality rates - has no added value for hospital comparisons but is useful for team quality assessment in Abdominal Aortic Aneurysm Surgery in the Netherlands

N. Lijftogt
E. G. Karthaus
A. Vahl
E. W. van Zwet
E. M. van der Willik
R. A. E. M. Tollenaar
J. F. Hamming
M. W. J. M. Wouters

in collaboration with the Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing

European Journal for Vascular and Endovascular Surgery 2018

ABSTRACT

Objectives:

Failure to rescue (FTR) is a composite quality indicator, defined as the proportion of deceased patients following major complications. The aims of this study were to compare FTR with mortality for hospital comparisons in abdominal aortic aneurysm (AAA) surgery in The Netherlands and investigate hospital volume and associated factors.

Material and Methods:

Patients prospectively registered between 2013 and 2015 in the Dutch Surgical Aneurysm Audit (DSAA) were analysed. FTR was analysed for AAA patients and subgroups elective (EAAA) and acute (AAAA; symptomatic or ruptured) aneurysms. Variables and hospital volume were analysed by uni- and multivariable regression analysis. Adjusted hospital comparisons for mortality, major complications, and FTR were presented in funnel plots. Isomortality lines were constructed when presenting FTR and major complication rates.

Results:

A total of 9258 patients were analysed in 61 hospitals: 7149 EAAA patients (77.2%) and 2109 AAAA patients (22.8%). There were 2785 (30.1%) patients with complications (unadjusted range 5-65% per hospital): 2161 (77.6%) with major and 624 (28.4%) patients with minor complications. Overall mortality was 6.6% (adjusted range 0-16% per hospital) and FTR was 28.4% (n = 613) (adjusted range 0-60% per hospital). Glasgow Coma Scale, age, pulse, creatinine, electrocardiography, and operative setting were independently associated with FTR. Hospital volume was not associated with FTR. In AAAA patients hospital volume was significantly associated with a lower adjusted major complication and mortality rate (OR 0.62, 95% CI 0.49-0.78; and 0.64, 95% CI 0.48-0.87). Four hospitals had a significant lower adjusted FTR with different major complication rates on different isomortality lines.

Conclusion:

There was more variation in FTR than in mortality between hospitals. FTR identified the same best performing hospitals as for mortality and therefore was of limited additional value in measuring quality of care for AAA surgery. FTR can be used for internal quality improvement with major complications in funnel plots and diagrams with isomortality lines.

INTRODUCTION

Clinical audits have become increasingly appreciated as a tool for quality improvement in vascular surgical care.[1] Medical improvement can be achieved using feedback on the hospital's structure, processes and outcomes of care. Traditionally, outcome parameters such as post-operative mortality and complication rates are used for this purpose.[2] However, statistical uncertainties associated with low hospital volume or low event rates cause difficulties in the interpretation of the observed variation in outcome between hospitals.[3] In abdominal aortic aneurysm (AAA) surgery the introduction of endovascular aneurysm repair (EVAR) has resulted in a marked decrease in post-operative mortality and complications.[4,5] With a mean mortality of 0.9% in EVAR patients in registries such as the Dutch Surgical Aneurysm Audit (DSAA) or 0.6% in the Swedish vascular registry (SWEDVASC), post-operative mortality as a single quality indicator seems to be of limited discriminative value for hospital comparisons.[6,7] Moreover, mortality is dependent on both casemix and hospital performance.[8]

To focus on hospital processes, composite outcome measures have been developed that tend to be more related to hospital processes and are less sensitive to errors in risk adjustment.[8-12] Examples of composite outcome measures are "textbook outcome" (TO), a measure for the percentage of patients with full achievement of desired outcomes, and "failure to rescue" (FTR), which represents the ability to treat complications effectively and therefore prevent death.[10,13-15]

The primary objective of this study was to compare FTR between hospitals performing AAA surgery. The second aim was to investigate whether FTR is a more discriminative outcome parameter than mortality or major complications. Additionally, variables contributing to FTR and the consequent adjusted association between FTR and hospital volume were investigated with the aim of comparing quality of AAA surgery between hospitals.

MATERIAL AND METHODS

Dataset

Variables and outcomes were retrieved from the DSAA, a mandatory national vascular audit in which every vascular unit has registered all primary AAA repairs in the Netherlands since 2013. Registration is performed according to a protocol approved by the scientific board, a group of vascular surgeons representing the interests of Dutch

hospitals. For the year 2015 data verification was performed by at random selection of 14 hospitals for review (Supplementary material A).

Patients

All patients undergoing primary, infra- or juxtarenal AAA surgery in 2013, 2014, and 2015 registered in the DSAA were evaluated. Analysis was performed on a patient level. The minimal complete dataset to consider a patient eligible for analysis included date of birth, date of surgery, operative setting/urgency (elective/acute symptomatic/ruptured aneurysm), type of procedure (EVAR or open surgical repair; OSR), and mortality. Patients undergoing surgery in hospitals that stopped performing AAA surgery during the study period and fewer than 15 patients in three years registered in the DSAA were excluded from analysis.

Procedures

Procedures were divided into EVAR or OSR and the setting into elective (EAAA) or acute (AAAA) surgery. The AAAA group was a composite group of ruptured (RAAA), a patient needing surgery within 2 hours, or as soon as possible if extravasation was seen on computed tomography angiography (CTA), and acute non-ruptured symptomatic aneurysms (SAAA), a patient needing surgery within 24 hours if no extravasation was present on CTA. EVAR procedures converted to OSR were categorised by intention to treat.

Outcome definitions

Mortality was defined as death within 30 days after surgery or within the same hospital admission (in hospital mortality). A complication was defined as death or any peri-operative- or post-operative complication.[11] A major complication was defined as post-operative death or a peri- or post-operative complication leading to a re-intervention or prolonged hospital stay. A minor complication was defined as a complication not resulting in a re-intervention, prolonged hospital stay, or mortality. A prolonged hospital stay was defined as the length of hospital stay (LOS) exceeding the 75th percentile of the LOS per subgroup of living patients registered in the DSAA between 2013 and 2015: EVAR or OSR stratified by EAAA, SAAA, or RAAA surgery.

Failure to rescue was defined as the number of patients that died within 30 days after surgery or in the same hospital admission, divided by the number of patients with major complications.

Statistical analysis

Baseline characteristics, surgical treatment and 30 day mortality were analysed for the overall group of AAA patients and for the subgroups of EAAA and AAAA patients when appropriate. Baseline analysis was performed for three outcomes on a patient level: no complications, only minor complications, or at least one major complication. Continuous variables were analysed to test normality and linearity. Missing values for continuous variables were imputed with the overall mean in the case of linearity and normality. When no linearity or normality was found for continuous variables, these were categorised. Categorical variables were dichotomised and missing values were ana- lysed as the group unknown for further analysis. Univariable and multivariable logistic regression analysis were performed in order to identify independent casemix and operative setting variables associated with FTR. A p value < .05 was considered as statistically significant.

Adjusted mortality, major complication rates, and FTR were compared between hospital volume tertiles. To ensure casemix corrected comparison between hospitals, a multivariable logistic regression analysis was used to adjust outcomes for patient characteristics by variables measured on admission in part based on the re-estimated V(p)-POSSUM variables on the Dutch population,[6,16,17] including age, gender, systolic blood pressure, heart rate, pulmonary status, cardiac status, preoperative electrocardiography (ECG), creatinine, Glasgow Coma Scale (GCS), haemoglobin, and operative setting: EAAA, SAAA or RAAA. Hospital comparisons were displayed in funnel plots with 95% confidence intervals: hospital volume versus mortality and major complications, as well as volume of major complications versus FTR. Finally, per hospital the percentage of adjusted FTR was shown in relation to the percentage of adjusted major complications represented by isomortality lines. Consequently, if mortality is considered a major complication, then the mortality rate equals the major complication rate multiplied by the FTR. If the major complication rate is plotted on the x-axis and the FTR on the y-axis, then all points lying on the line $y = c/x$ correspond to the same mortality rate c.

Adjusted outcome was calculated as the percentage observed events divided by the percentage expected events times the mean observed percentage of events, represented in funnel plots with 95% confidence intervals. Iso- mortality lines reflect the same mortality percentage across these lines for every hospital.

Additionally hospitals were divided into tertiles based on their procedural volume after three years of surgery: low, medium, and high volume hospitals. Mortality, major

complications, and FTR were compared between these tertiles, both adjusted and unadjusted. Statistical analysis was performed in SPSS version 23.0.

RESULTS

Patients

A total of 9353 patients were registered by 63 hospitals and 9273 (99.1%) of these patients met the inclusion criteria of this study. After exclusion of two hospitals (registration of three and 12 patients), 9258 patients were included for analysis (Figure 1). In Table 1 patient and treatment characteristics are shown as input for adjustment. The mean age was 73.2 years and 85.8% were male. Cardiac comorbidities and pulmonary comorbidities were most frequently seen, respectively 47.5% and 22.8%. There were 7149 EAAA patients (77.2%) and 2109 AAAA patients: 641 SAAA (6.9%) and 1468 RAAA (15.9%). The majority of patients were treated by EVAR (n = 6,317, 68.2%; and 0.5%, n = 47 EVAR converted to OSR). In the subgroup of EAAA patients 76.6% (n = 5473) were treated by EVAR compared with 58.0% (n = 372) of the SAAA and 35.4% (n = 519) of the RAAA patients.

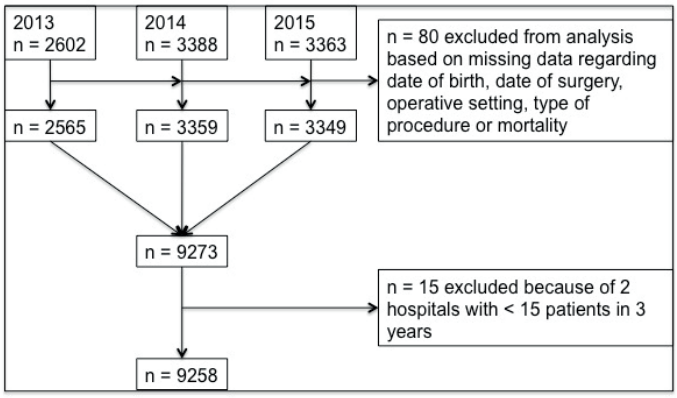


Figure 1. Flowchart of patient selection.

Outcomes: complications and mortality

Overall, a total of 6473 (69.9%) patients had no complications. There were 2785 (30.1%) patients with one or more complications (range 5-65% per hospital). There were differences in baseline characteristics between patients with no, only minor, or at least one major complication. Patients with major complications had more comorbidities. Overall 6.7% (n = 624) of the patients had only minor complications. The overall percentage of major complications was 23.3% (n = 2161). The overall mortality was 6.6% (n = 613)

Table 1. Patient characteristics: case mix variables, including operative technique and setting, and outcome variables in patients with and without complications

		Total patient group		Patients without complications		Patients with only minor complications		Patients with major complications	
Variables		No. Patients	Percentage	No. Patients	Percentage	No. Patients	Percentage	No. Patients	Percentage
Total No. Patients		9 258	100%	6473	69.9%	624	6.7%	2161	23.3%
Age	mean ± SD	73.22 ± 7.85		72.87 ± 7.84		73.57 ± 7.97		74.19 ± 7.75	
Sex	Male	7939	85.8%	5627	87.0%	553	88.8%	1759	81.4%
Procedure	EVAR	6317	68.2%	5142	79.4%	487	78.0%	688	31.8%
	EVAR converted	47	0.5%	16	0.2%	2	0.3%	29	1.3%
	OSR	2894	31.3%	1315	20.3%	135	21.6%	1444	66.8%
	Total	100%		100%		100%		100%	
Setting	RAAA	1468	15.9%	459	7.1%	193	30.9%	816	37.8%
	SAAA	641	6.9%	420	6.5%	35	5.6%	186	8.6%
	EAAA	7149	77.2%	5594	86.4%	396	63.5%	1159	53.6%
	Total	100%		100%		100%		100%	
Heart	No cardiac problems	4229	45.7%	3043	47.0%	302	48.4%	884	40.9%
	Peripheral edema	708	7.6%	492	7.6%	42	6.7%	174	8.1%
	Elevated CVP	129	1.4%	91	1.4%	6	1.0%	32	1.5%
	Medication for hypertension	3566	38.5%	2522	39.0%	224	35.9%	820	37.9%
	Unknown	626	6.8%	325	5.0%	50	8.0%	251	11.6%
	Total	100%		100%		100%		100%	
Lungs	No dyspnea	6712	72.5%	4920	76.0%	448	71.8%	1344	62.2%
	Dyspnea exercise	1741	18.8%	1155	17.8%	120	19.2%	466	21.6%
	Invalidating dyspnea	263	2.8%	166	2.6%	21	3.4%	76	3.5%
	Dyspnea in rest	107	1.2%	60	0.9%	6	1.0%	41	1.9%

Table 1. Patient characteristics: case mix variables, including operative technique and setting, and outcome variables in patients with and without complications (*continued*)

		Total patient group		Patients without complications		Patients with only minor complications		Patients with major complications	
Variables		No. Patients	Percentage	No. Patients	Percentage	No. Patients	Percentage	No. Patients	Percentage
	Unknown	435	4.7%	172	2.7%	29	4.6%	234	10.8%
	Total	100%		100%		100%		100%	
Malignancy	None	7576	81.8%	5231	80.8%	528	84.6%	1817	84.1%
	Current or history	1596	17.2%	1188	18.4%	93	14.9%	315	14.6%
	Unknown	86	0.9%	54	0.8%	3	0.5%	29	1.3%
	Total	100%		100%		100%		100%	
GCS	15	8175	88.3%	5931	91.6%	546	87.5%	1698	78.6%
	12-14	219	2.4%	56	0.9%	23	3.7%	140	6.5%
	9-11	58	0.6%	7	0.1%	4	0.6%	47	2.2%
	<9	93	1.0%	8	0.1%	4	0.6%	81	3.7%
	Unknown	713	7.7%	471	7.3%	47	7.5%	195	9.0%
	Total	100%		100%		100%		100%	
ECG	No abnormalities	4836	52.2%	3634	56.1%	297	47.6%	905	41.9%
	Atrial Fibrillation	589	6.4%	406	6.3%	42	6.7%	141	6.5%
	MI or other	2350	25.4%	1604	24.8%	161	25.8%	585	27.1%
	Unknown	1483	16.0%	829	12.8%	124	19.9%	530	24.5%
	Total	100%		100%		100%		100%	
Sodium	Normal (135-145)	7754	83.8%	5470	84.5%	532	85.3%	1752	81.1%
	Divergent (<135 or >145)	761	8.2%	445	6.9%	55	8.8%	261	12.1%
	Unknown	743	8.0%	558	8.6%	37	5.9%	148	6.8%
	Total	100%		100%		100%		100%	

Table 1. Patient characteristics: case mix variables, including operative technique and setting, and outcome variables in patients with and without complications (*continued*)

		Total patient group		Patients without complications		Patients with only minor complications		Patients with major complications	
Variables		No. Patients	Percentage	No. Patients	Percentage	No. Patients	Percentage	No. Patients	Percentage
Potassium	Normal (3,5-5,0)	7915	85.5%	5612	86.7%	539	86.4%	1764	81.6%
	Divergent (<3,5 of >5.0)	769	8.3%	440	6.8%	54	8.7%	275	12.7%
	Unknown	574	6.2%	421	6.5%	31	5.0%	122	5.6%
	Total	100%		100%		100%		100%	
Creatinin	mean ± SD	101.52 ± 45.11		98.66 ± 42.72		106.24 ± 51.49		108.74 ± 49.05	
WBC	mean ± SD	9.01 ± 2.73		8.63 ± 2.36		9.56 ± 3.20		9.97 ± 3.29	
Hemoglobin	mean ± SD	8.43 ± 1.18		8.60 ± 1.05		8.26 ± 1.20		7.96 ± 1.38	
SBP	mean ± SD	135.87 ± 25.92		138.78 ± 22.43		133.85 ± 27.44		127.71 ± 32.58	
Heart rate	mean ± SD	75.50 ± 15.72		73.97 ± 14.15		77.55 ± 17.65		79.48 ± 18.59	
Aneurysm size	mean ± SD	63.31 ± 14.17		61.39 ± 12.62		64.64 ± 15.67		68.69 ± 16.50	

Hypertension, angina pectoris, diuretics or digoxin. SD = standard deviation; CVP = central venous pressure; heart = cardiac comorbidity; lungs = pulmonary comorbidity; GCS = Glasgow Coma Scale; ECG = electrocardiography; MI = myocardial Infarction; IQR = interquartile range; WBC = white blood count; SBP = systolic blood pressure; LOS = length of stay.

with an adjusted range of 0-16% per hospital. In Table 2, more detailed information is shown for operative setting and technique performed.

Table 2. Summary of mortality, minor and major complications by procedure and setting

	Overall									
	No. Patients	%								
			EAAA		AAAA					
			No. Patients	%	No. Patients	%				
							SAAA		RAAA	
							No. Patients	%	No. Patients	%
Mortality										
EVAR	175	2.7%	42	0.8%	133	14.9%	18	4.8%	115	22.2%
OSR	438	15.1%	93	5.5%	345	28.3%	28	10.4%	317	33.4%
All patients	613	6.6%	135	1.9%	478	22.7%	46	7.2%	432	29.4%
Minor complications										
EVAR	489	7.7%	378	6.9%	111	12.5%	32	8.6%	79	15.2%
OSR	135	4.7%	18	1.1%	117	9.6%	3	1.1%	114	12.0%
All patients	624	6.7%	396	5.5%	228	10.8%	35	5.5%	193	13.1%
Major complications										
EVAR	717	11.3%	456	8.3%	261	29.3%	64	17.2%	197	38.0%
OSR	1444	49.9%	703	41.9%	741	60.8%	122	45.4%	619	65.2%
All patients	2161	23.3%	1159	16.2%	1002	47.5%	186	29.0%	816	55.6%

The median length of hospital stay for EAAA, SAAA, and RAAA patients after EVAR was respectively 2 (IQR 2-4), 4 (IQR 2-6), and 7 (IQR 5-13) days. After OSR this was respectively 8 (IQR 7-12), 10 (IQR 7-15), and 15 (IQR 10-24) days.

Failure to rescue

In Table 3 the odds ratio (OR), unadjusted and adjusted, for FTR is shown after analysing the overall patient group, and for the subgroups of EAAA and AAAA patients. In 613 patients (28.4%) FTR was observed. The adjusted variation in FTR between hospitals ranged from 0 to 60%. The number of patients with a major complication in EAAA patients was 1159 with a mean percentage of FTR of 11.6% (n = 135): 9.2% after EVAR and 13.2% after OSR. In AAAA patients the number of patients with a major complication was 1002 with a percentage FTR of 47.7% (n = 478): 51.0% after

Table 3. Uni- and multivariable analysis of variables associated with FTR for the whole group and subgroups of EAAA and AAAA patients.

	Univariate analysis			Multivariate analysis		
	Overall	EAAA	AAAA	Overall	EAAA	AAAA
	Odds Ratio (95% CI)	Odds Ratio (95% CI)	Odds Ratio (95% CI)	Odds Ratio (95% CI)	Odds Ratio (95% CI)	Odds Ratio (95% CI)
Age (year)	**1.06 (1.05-1.08)**	**1.03 (1.00-1.06)**	**1.06 (1.04-1.08)**	**1.05 (1.03-1.07)**	1.02 (0.99-1.04)	**1.06 (1.04-1.08)**
Female gender	1.05 (0.82-1.33)	1.14 (0.74-1.77)	1.13 (0.82-1.57)	-	-	-
Cardiac co-morbidity	1.23 (1.00-1.51)	1.28 (0.88-1.86)	**1.42 (1.07-1.88)**	-	-	1.06 (0.77-1.46)
Pulmonary co-morbidity	**1.38 (1.11-1.72)**	**1.62 (1.11-2.35)**	**1.48 (1.09-2.02)**	1.34 (1.04-1.62)	**1.50 (1.03-2.20)**	1.26 (0.89-1.77)
Malignancies	0.99 (0.77-1.30)	1.30 (0.82-2.04)	1.14 (0.78-1.67)	-	-	-
GCS <15	**6.89 (5.23-9.08)**	-	**3.04 (2.25-4.10)**	**2.22 (1.61-3.07)**	-	**2.21 (1.59-3.07)**
AAA size (mm)	**1.02 (1.02-1.03)**	1.00 (0.99-1.02)	1.00 (0.99-1.01)	1.00 (0.99-1.00)	-	-
RAAA	**8.53 (6.81-10.7)**	-	**3.24 (2.39-4.91)**	**3.91 (2.73-5.61)**	-	**1.89 (1.24-2.90)**
SAAA	**2.49 (1.71-3.64)**	-	-	**1.99 (1.32-2.99)**	-	-
ECG irregular	**1.86 (1.47-2.35)**	1.36 (0.94-1.99)	**1.97 (1.40-2.76)**	**1.52 (1.16-1.98)**	-	**1.84 (1.26-2.69)**
Creatinin level	**1.01 (1.01-1.01)**	**1.01 (1.00-1.01)**	**1.01 (1.00-1.01)**	**1.00 (1.00-1.01)**	1.00 (1.00-1.01)	**1.00 (1.00-1.01)**
Abnormal Sodium	**2.02 (1.55-2.64)**	1.37 (0.70-2.68)	1.27 (0.92-1.75)	1.17 (0.85-1.60)	-	-
Abnormal Potassium	**1.71 (1.31-2.23)**	1.34 (0.73-2.44)	1.16 (0.84-1.60)	1.01 (0.74-1.38)	-	-
WBC	**1.17 (1.13-1.20)**	1.03 (0.95-1.12)	1.02 (0.99-1.06)	1.01 (0.97-1.04)	-	-
SBP	**0.98 (0.98-0.98)**	1.00 (0.99-1.01)	**0.99 (0.99-0.99)**	1.00 (0.99-1.00)	-	**1.00 (0.99-1.01)**
Pulse rate	**1.02 (1.02-1.03)**	1.00 (0.99-1.02)	**1.01 (1.00-1.01)**	**1.01 (1.00-1.01)**	-	1.01 (1.00-1.01)
Hemoglobin	**0.64 (0.60-0.69)**	**0.75 (0.63-0.88)**	**0.83 (0.76-0.90)**	0.97 (0.89-1.06)	**0.80 (0.67-0.95)**	1.02 (0.92-1.13)
OSR	**1.35 (1.10-1.65)**	**1.50 (1.02-2.21)**	0.84 (0.63-1.11)	-	-	-

Bold: statistically significant

EVAR and 46.6% after OSR. Preoperative variables independently associated with FTR were GCS, age, pulse rate, creatinine, ECG, and operative setting (acute operation). Furthermore, for the subgroup of EAAA patients, pulmonary comorbidity and preoperative haemoglobin were significantly associated with FTR; for AAAA preoperative systolic blood pressure was also independently associated with FTR.

FTR and hospital volume

Total hospital volume, as for the subgroups of EAAA and AAAA patients, were split into tertiles. Low overall volume was defined as up to 149 patients in three years of AAA surgery and high volume at 198 patients or more. For subgroup analysis the EAAA volume was split into the following groups: < 110, 110-156, and > 156 patients per hospital. AAAA volume was split in groups of <36, 36-49, >49 patients per hospital. Nine of 12 high volume EAAA hospitals were also high volume AAAA hospitals. For unadjusted only, there was an association between lower FTR and medium or high volume hospitals rather than with low volume hospitals (OR for high volume 0.79; 95% CI 0.63-0.99, and for medium volume 0.79; 95% CI 0.63-0.98). There was a statistically significant independent adjusted association between high hospital volume and mortality (OR 0.79; 95% CI 0.62-1.00), but not for the percentage of major complications and FTR (Table 4). The subgroup of high volume AAAA hospitals treating 50 patients or more in three years performed significantly better regarding mortality (OR 0.64; 95% CI 0.48-0.87) and major complications (OR 0.62; 95% CI 0.49-0.78) than low volume hospitals. However, the adjusted OR for FTR was not significant.

Hospital comparisons

For hospital comparisons, the first step was the overall adjusted mortality rate per hospital volume (Figure 2). There were four hospitals (green) with a statistically significant lower adjusted mortality than the national mean. There was only one low volume hospital with a significantly higher mortality (red). The second step was the percentage of patients with one or more major complications per hospital volume (Figure 3). There was a wide variation in the number of complications and the (adjusted) complication rate including several hospitals performing significantly better or worse than the national average. However, the four green hospitals (low mortality) and one red (high mortality) performed within the confidence limits of the national average of major complication rates. However, in the third step (Figure 4) it can be concluded that these four green hospitals, next to a lower mortality, also have significant lower FTR. In contrast, those hospitals with more complications than the national average scored within the confidence limits for FTR and for mortality, including the one hospital performing significantly worse for mortality than the national average.

Table 4. Three-year hospital volume tertiles for EAAA and AAAA surgery and the association with mortality, major complications and FTR

	Overall			EAAA			AAAA		
	Low volume <149	Medium volume 149-197	High volume >197	Low volume <110	Medium volume 110-156	High Volume >156	Low volume <36	Medium volume 36-49	High Volume >49
	Reference	Odds Ratio (95% CI)	Odds Ratio (95% CI)	Reference	Odds Ratio (95% CI)	Odds Ratio (95% CI)	Reference	Odds Ratio (95% CI)	Odds Ratio (95% CI)
No. hospitals	30	19	12	31	18	12	33	17	11
Mortality (unadjusted)	Ref.	**0.82 (0.67-0.99)**	**0.78 (0.64-0.95)**	Ref.	0.95 (0.63-1.45)	1.08 (0.72-1.64)	Ref.	1.09 (0.86-1.39)	**0.75 (0.58-0.97)**
Mortality (adjusted)	Ref.	0.90 (0.71-1.13)	**0.79 (0.62-1.00)**	Ref.	0.96 (0.62-1.47)	1.08 (0.70-1.66)	Ref.	1.14 (0.86-1.51)	**0.64 (0.48-0.87)**
Major complications (unadjusted)	Ref.	0.97 (0.87-1.10)	0.92 (0.81-1.03)	Ref.	1.01 (0.87-1.18)	1.03 (0.89-1.21)	Ref.	0.93 (0.76-1.15)	**0.72 (0.58-0.89)**
Major complications (adjusted)	Ref.	1.04 (0.92-1.18)	0.91 (0.79-1.04)	Ref.	1.01 (0.86-1.18)	1.01 (0.86-1.18)	Ref.	0.92 (0.73-1.15)	**0.62 (0.49-0.78)**
FTR (unadjusted)	Ref.	**0.79 (0.63-0.98)**	**0.79 (0.63-0.99)**	Ref.	0.94 (0.60-1.46)	1.06 (0.69-1.64)	Ref.	1.22 (0.91-1.64)	0.90 (0.66-1.23)
FTR (adjusted)	Ref.	0.88 (0.67-1.14)	0.91 (0.69-1.20)	Ref.	0.97 (0.62-1.53)	1.11 (0.70-1.76)	Ref.	1.30 (0.93-1.82)	0.82 (0.58-1.17)

Bold indicates statistically significant. Adjusted for variables: age, systolic blood pressure, heart rate, pulmonary status, cardiac status, pre-operative electrocardiogram and pre-operative creatinine, GCS, haemoglobin, gender, setting/urgency, and year of surgery, OR 1/4 odds ratio (95% confidence interval).

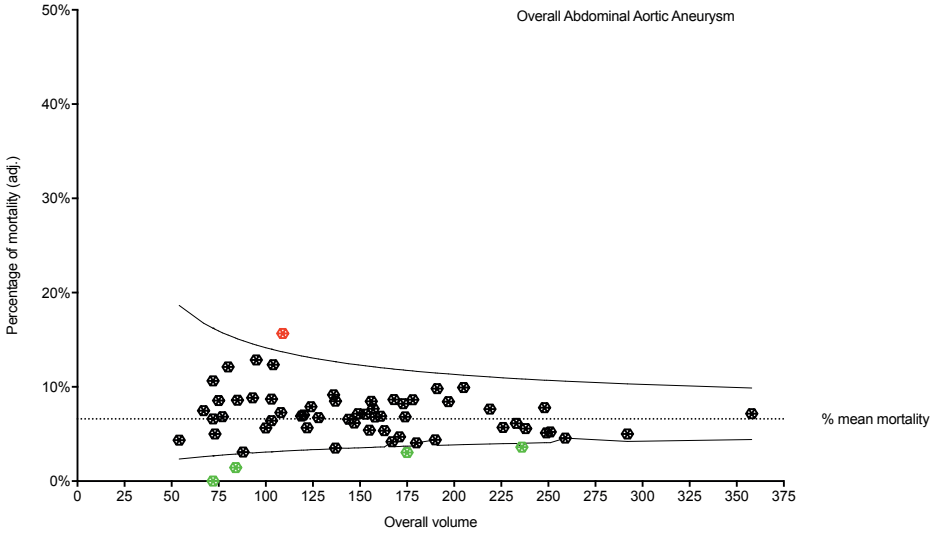


Figure 2. The adjusted mortality rate per hospital (volume) in patients after AAA surgery (95% CI).

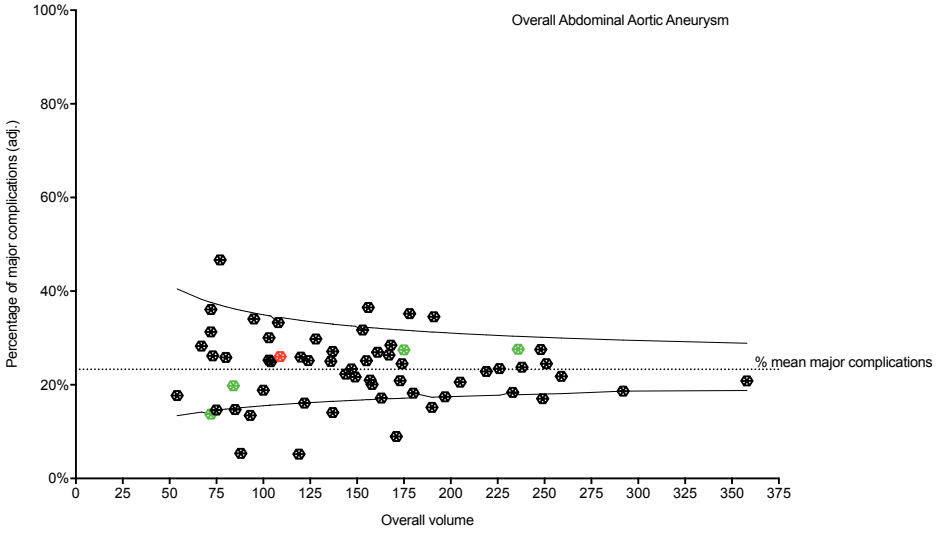


Figure 3. The adjusted major complication rate per hospital (volume) in patients after AAA surgery (95% CI).

In the final step (Figure 5) adjusted FTR is plotted against the adjusted complication rate. By adding the isomortality lines, hospitals can get insight into their performance, compared with other hospitals, for three parameters together in one plot: mortality, complication rate, and FTR.

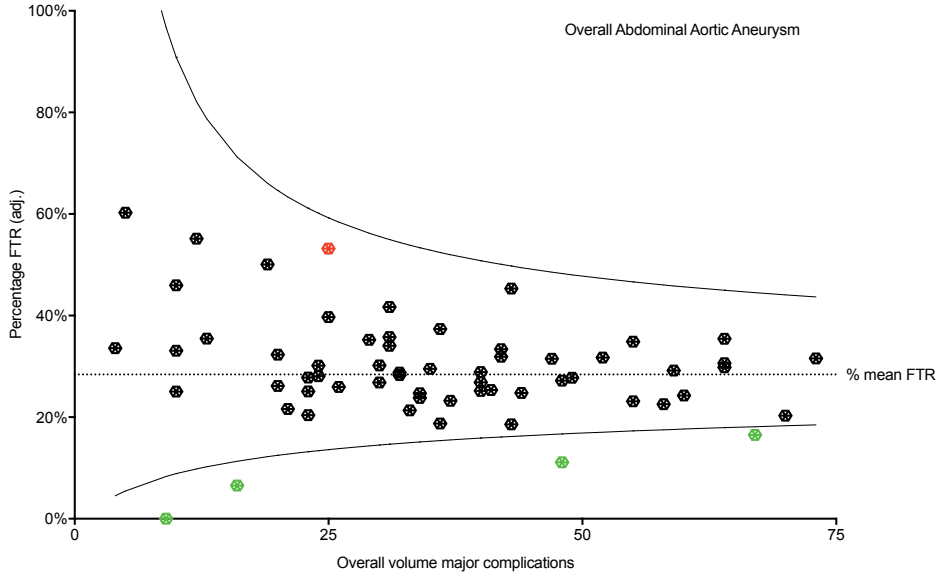


Figure 4. The adjusted FTR per volume of complications by hospital after AAA surgery (95% CI).

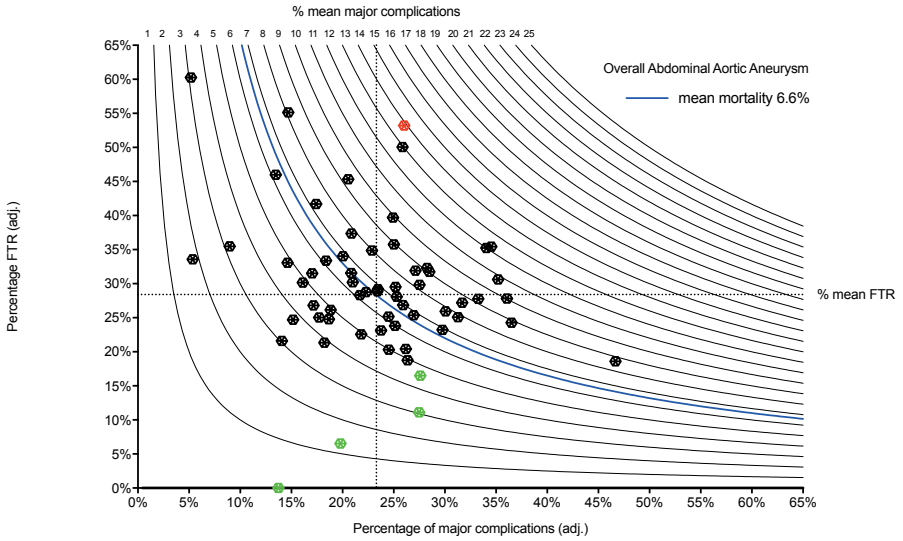


Figure 5. The adjusted percentages of FTR and major complications after AAA surgery. The hyperbolic lines project the mortality in relation to the percentage major complications and the mortality after major complications. From left to right each line has a 1% higher mortality for every hospital crossing that line starting at 1%, with the overall mean percentage mortality of 6.6% projected with the blue line.

Additional analysis for EAAA and AAAA can be found in the Supplementary material B and C.

DISCUSSION

This study has shown an increased variation between hospitals in The Netherlands performing AAA surgery for the composite outcome measure FTR compared with the single measure mortality. But FTR did not identify those hospitals with a significantly higher mortality following major complications. Besides, best performing hospitals regarding mortality were the same hospitals for FTR, independent of the low power in the denominator. Therefore, FTR can be used for internal quality purposes, but is of limited use for hospital comparisons in the AAA treatment evaluation next to mortality and complication rates. By combining FTR and complication rate in one plot together with isomortality lines, each hospital gets insight in these three parameters for quality assessment in one plot.

FTR reflects the ability of a surgical team to recognise and treat complications adequately. However, hospitals with undesirably high complication rates can have a low FTR while a high FTR can be observed next to a low complication rate. Since FTR is merely dependent on the number of major complications related to overall hospital volume and death rate, it is important to report this outcome measure together with mortality and major complications by hospital volume. Mortality as a single indicator is highly dependent on the type of admission and operative setting and separate analyses are necessary for EAAA and AAAA surgery for hospital comparisons. The advantage of FTR as an indicator is that all AAA patients can be analysed together. The ability to recognise and treat major complications and avoid mortality is unrelated to the type of admission and operative setting. In other words, complications associated with acute surgery must be just as appropriately treated as complications in the elective setting.

One of the difficulties regarding FTR is the definition used in which every patient that dies following a major complication is graded as preventable death.[18] However, death may be an unwanted outcome but cannot always be prevented, which can be the case in emergency surgery, such as RAAA patients.[19,20] Another difficulty is that the definition used for FTR varies in the types of complications that are included in the denominator or whether or not to exclude those patients that die without a complication.[19] The authors support the view of the developers of FTR, that death is considered as a result of a complication and therefore all deaths should be included in FTR, in the numerator and in the denominator.[11] Although alternatives have been proposed, the inclusion of mortality in the denominator also overcomes the problem of under, or different, registration of complications and statistical uncertainties following low volume and wide confidence intervals.[9,18,20,21] These discrepancies make international comparisons difficult.

Primarily, FTR was investigated because of the low discriminative ability of the low mortality rates in elective AAA surgery.[6] Indeed, the mean FTR in this study is higher (9.2% EVAR and 13.2% OSR), with more variation, than the mean mortality as described for EAAA patients and was similar to a study in 2015: 9.6% FTR for EVAR and 11.1% FTR for OSR.[22] However hospital variation regarding outliers is disappointing. A study investigating FTR in EAAA surgery revealed percentages even lower than earlier described: 0.6% for EVAR and 2.7% for OSR.[23] This can be caused by the inclusion of different complications in the FTR denominator. The standardisation of definitions, assessment tools, clinically relevant endpoints, and adherence to national reporting guidelines would help improve the investigation of system factors that influence vascular outcomes.[24] In addition, combining FTR with other composite outcome measures, for example TO, could improve quality assessment and create awareness regarding the performance of individual hospitals.[15]

An advantage of FTR is that it reflects the ability of the surgical team to treat complications and avoid (consequential) death. It is therefore more dependent on hospital factors than on patient (casemix) factors.[10,25] However, the influence of age on outcomes like FTR cannot be underestimated.[26] In this study there were only a few baseline characteristics of influence on FTR with the majority for overall and acute AAA surgery indicating that in general, FTR is merely influenced by hospital characteristics instead of casemix. Hospital volume was not found to be associated with FTR. However several studies have underlined the importance and influence of hospital volume on FTR.[14,23,27,28] One study found a significant volume FTR association for OSR but not for EVAR.[25] Consequently, as expected, AAAA surgery resulted in a less favourable FTR than EAAA and OSR had a worse FTR than EVAR. Therefore, operative setting was included for adjustment as it is an important casemix variable. Owing to the minimum number of 20 EAAA operations in The Netherlands the volume differences between hospitals have become smaller and have probably reached the goal by improving Dutch AAA healthcare. However, in AAAA surgery a significant association was observed between the three year volume tertiles and the outcomes major complications and mortality.

Limitations

There are some limitations of this study. First, patients requiring more specialised care being referred to a tertiary referral centre, could not be extracted from the data. However, adjusting for casemix will in part solve this problem of treating more morbid patients in the case of a tertiary referral centre. Second, no other hospital characteristics than procedural volume were registered in this dataset. Therefore, the observed variation of FTR cannot be easily attributed to specific differences in (infra) structure or other processes that influence FTR. Third there is a risk of registration bias due to miss-

ing data, which is a common problem of observational data. It has been reported that complications are often underreported.[11] However, every hospital in The Netherlands participates in the mandatory DSAA and every compulsory variable needs to be filled before finalising a patient's registration. Therefore, missing or inaccurate data are kept as low as possible and exceeded 10% in only a few cases. Moreover, data verification of randomly selected patients of 14 hospitals revealed that only minor complications were missing in 7.1% of 298 patients (Supplementary material A). However, no re-interventions or mortality, included in the definition of a major complication, were missed. Still variation in the percentage of complications registered by hospitals was wide. Though FTR and complication rate were analysed over all AAA (EAAA, SAAA and RAAA) patients, with the assumption that the outcome severe complications, like bleeding or colonic ischaemia can be recognised and treated in every setting, and that adjustment for operative setting will correct for the difference in incidence of these outcomes between settings. Under this assumption complete surgical care can be analysed, which would be a great advantage when comparing hospitals. However, adequate adjustment remains important, but the difference between the percentages of EAAA and AAAA by hospital could result in bias. Additional figures are presented as Supplementary material to compare FTR versus major complication rates with isomortality lines for EAAA and AAAA surgery.

Last, patients turned down for surgery are not included in this registry. It is possible that this may cause selection bias, especially in the group of RAAA patients where turning down a patient for surgery is highly variable between countries.

CONCLUSION

FTR reflects the ability of the vascular team to recognise and treat complications after AAA surgery in order to prevent consequent mortality. Hospitals with a significant adjusted difference from the mean for mortality, major complications and FTR could be identified. However, there were only a few significant outliers that were all performing better than the national average regarding FTR, despite a wide variation of FTR rates between hospitals. These hospitals corresponded to the hospitals that also scored better on the single variable mortality on different isomortality lines without significant differences in major complication rates. This shows FTR alone to be of limited use when comparing hospitals. There was also no association between FTR and hospital volume. To get a useful interpretation of FTR for internal quality improvement it needs to be combined with the percentage of major complications related to mortality reflected in diagrams by isomortality lines.

ACKNOWLEDGEMENTS

The authors would like to thank the members of the Dutch Society for Vascular Surgery who registered their patients in the DSAA, the Dutch Institute for Clinical Auditing that facilitated the registry and the Dutch Surgical Aneurysm Audit group.

Conflicts of interest: None.

Funding: None.

REFERENCES

1 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 Vascunet report. Eur J Vasc Endovasc Surg 2011;42:598-607.

2 Van Leersum NJ, Snijders HS, Henneman D, Kolfschoten NE, Gooiker GA, ten Berge MG, et al. The Dutch surgical colorectal audit. Eur J Surg Oncol 2013;39:1063-70.

3 Fischer C, Lingsma H, Hardwick R, Cromwell DA, Steyerberg E, Groene O. Risk adjustment models for short-term outcomes after surgical resection for oesophagogastric cancer. Br J Surg 2016;103:105-16.

4 Prinssen M, Verhoeven EL, Buth J, Cuypers PW, van Sambeek MR, Balm R, et al. A randomized trial comparing conventional and endovascular repair of abdominal aortic an- eurysms. N Engl J Med 2004;351:1607-18.

5 Greenhalgh RM, Brown LC, Kwong GP, Powell JT, Thompson SG. EVAR trial participants. Comparison of endo- vascular aneurysm repair with open repair in patients with abdominal aortic aneurysm (EVAR trial 1), 30-day operative mortality results: randomised controlled trial. Lancet 2004;364:843-8.

6 Lijftogt N, Vahl AC, Wilschut ED, Elsman BH, Amodio S, van Zwet EW, et al. Adjusted hospital outcomes of abdominal aortic aneurysm surgery reported in the Dutch surgical aneurysm audit. Eur J Vasc Endovasc Surg 2017;53:520-32.

7 Kragsterman B. Nationella kvalitetsregistret för kärlkirurgi. Årsrapport 2014. Retrieved July 8, 2018, from http://docplayer. se/4398776-Nationella-kvalitetsregistret-for-karlkirurgi-arsrapport-2014.html.

8 Holt PJ, Poloniecki JD, Hofman D, Hinchliffe RJ, Loftus IM, Thompson MM. Re-interventions, readmissions and discharge destination: modern metrics for the assessment of the quality of care. Eur J Vasc Endovasc Surg 2010;39:49-54.

9 Sinha S, Ata Ozdemir B, Khalid U, Karthikesalingam A, Poloniecki JD, Thompson MM, et al. Failure-to-rescue and interprovider comparisons after elective abdominal aortic aneurysm repair. Br J Surg 2014;101:1541-50.

10 Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. Med Care 1992;30:615-29.

11 Silber JH, Romano PS, Rosen AK, Wang Y, Even-Shoshan O, Volpp KG. Failure-to-rescue: comparing definitions to measure quality of care. Med Care 2007;45:918-25.

12 Rigby KA, Palfreyman S, Michaels JA. Performance indicators from routine hospital data: death following aortic surgery as a potential measure of quality of care. Br J Surg 2001;88:964-8.

13 Kolfschoten NE, Kievit J, Gooiker GA, van Leersum NJ, Snijders HS, Eddes EH, et al. Focusing on desired outcomes of care after colon cancer resections; hospital variations in 'text- book outcome'. Eur J Surg Oncol 2013;39:156-63

14 Henneman D, Snijders HS, Fiocco M, van Leersum NJ, Kolfschoten NE, Wiggers T, et al. Hospital variation in failure to rescue after colorectal cancer surgery: results of the Dutch Surgical Colorectal Audit. Ann Surg Oncol 2013;20:2117-23.

15 Karthaus E, Lijftogt N, Busweiler L, Elsman B, Wouters M, Vahl A, et al. Textbook outcome: a composite measure for quality of elective aneurysm surgery. Ann Surg 2017;266:898-904

16 Prytherch DR, Ridler BM, Beard JD, Earnshaw JJ, the Audit and Research Committee, The Vascular Surgical Society of Great Britian and Ireland. A model for national outcome audit in vascular surgery. Eur J Vasc Endovasc Surg 2001;21:477-83.
17 Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. Br J Surg 2001;88:958-63.
18 Hornor MA, Bilimoria KY. Moving beyond failure to rescue. Surgery 2017;161:791-2.
19 Kuo LE, Kaufman E, Hoffman RL, Pascual JL, Martin ND, Kelz RR. Failure-to-rescue after injury is associated with preventability: the results of mortality panel review of failure-to-rescue cases in trauma. Surgery 2017;161:782-90.
20 Ingraham AM, Greenberg CC. Failure to rescue and prevent- ability: striving for the impossible? Surgery 2017;161:793-4
21 Holena DN, Kaufman EJ, Delgado MK, Wiebe DJ, Carr BG, Christie JD, et al. A metric of our own: failure to rescue after trauma. J Trauma Acute Care Surg 2017;83:698-704.
22 Arya S, Kim SI, Duwayri Y, Brewster LP, Veeraswamy R, Salam A, et al. Frailty increases the risk of 30-day mortality, morbidity, and failure to rescue after elective abdominal aortic aneurysm repair independent of age and comorbidities. J Vasc Surg 2015;61:324-31.
23 Ilonzo N, Egorova NN, McKinsey JF, Nowygrod R. Failure to rescue trends in elective abdominal aortic aneurysm repair between 1995 and 2011. J Vasc Surg 2014;60:1473-80.
24 Lear R, Godfrey AD, Riga C, Norton C, Vincent C, Bicknell CD. The Impact of system factors on quality and safety in arterial surgery: a systematic review. Eur J Vasc Endovasc Surg 2017;54:79-93.
25 Waits SA, Sheetz KH, Campbell DA, Ghaferi AA, Englesbe MJ, Eliason JL, et al. Failure to rescue and mortality following repair of abdominal aortic aneurysm. J Vasc Surg 2014;59:909-914.e1.
26 Hicks CW, O'Kelly A, Obeid T, Locham S, Malas MB. Predicting failure to rescue after abdominal aortic aneurysm repair in elderly patients. J Surg Res 2017;217:265-70.
27 Ghaferi AA, Birkmeyer JD, Dimick JB. Hospital volume and failure to rescue with high-risk surgery. Med Care 2011;49:1076-81.
28 Gonzalez AA, Dimick JB, Birkmeyer JD, Ghaferi AA. Under- standing the volume-outcome effect in cardiovascular surgery: the role of failure to rescue. JAMA Surg 2014;149:119-23.

7

Summary, General Discussion and Future Perspectives

THE DUTCH SURGICAL ANEURYSM AUDIT

In the Netherlands several surgical audits have been developed in cooperation with the Dutch Institute for Clinical Auditing.[1] Although, this started as a surgical initiative only, there are 22 clinical audits currently being carried out on behalf of a number of different specialisations. The web-based DSAA, introduced in 2012 and mandatory since 2013, registers all primary abdominal aortic aneurysm (AAA) operations in the Netherlands.[2] It currently contains the data of patients in 60 hospitals and registers both electively planned AAA (EAAA) patients and patients with acute symptomatic AAA (SAAA) or acute ruptured AAA (RAAA) who undergo either open surgical repair (OSR) or endovascular aneurysm repair (EVAR).

The main outcome is 30-day or in-hospital short-term mortality, however, other outcomes such as processes of care and structural features are also registered.

The Dutch Society of Vascular Surgery (Nederlandse Vereniging voor Vaatchirurgie: NVVV) initiated the development of a set of quality measures in collaboration with several stakeholders.[3-7]

In **Chapter 1**, a brief overview is given of the development of the DSAA and the importance of audits. When measuring outcome indicators, differences between hospitals need to be casemix adjusted. Examples of variables associated with the outcome are often included in mortality risk prediction models such as V-POSSUM - a model with many variables or the V(p)-POSSUM with only physiological variables.[8] Since the value of single outcome parameters may be limited, alternative outcome measures might be of added value for comparing outcome between hospitals.

Chapter 2 presents the first results of the DSAA.[9] The 30 day or in-hospital mortality of 1.9% for elective AAA (EAAA) was comparable with other European registries. Mortality after elective EVAR was 0.9% and comparable to other registries. However, mortality following open surgical repair (OSR) in the DSAA was higher compared to the registries of other countries, e.g. the Swedish registry SWEDVASC (5.0% vs 3.2%). In ruptured aneurysms the mortality after EVAR and OSR was 22.2% and 32.0%, respectively and comparable with the results reported in SWEDVASC.

Because of the low mortality after EAAA, the adjusted and unadjusted variation between hospitals - shown in funnels - was not sufficient to discriminate between hospitals. Risk adjustment by V(p)- POSSUM for EAAA did not influence hospital variation. The discriminative ability of V-POSSUM to predict mortality was moderate for both

EAAA and AAAA, with C-statistics of 0.719 and 0.713, respectively. Also, three more recent years of follow-up did not result in significant differences between hospitals. (See Figures 1a-b)

Searching for case mix variables associated with mortality, in **Chapter 3** we searched for the best performing mortality risk prediction models by performing a systematic review.[10] No model has been used as a standard in clinical practice. However, the best performing models regarding both applicability and discrimination were the perioperative British Aneurysm Repair score (C-statistic 0.83) and the preoperative Vascular Biochemistry and Haematology Outcome Model (C-statistic 0.85). However, both lacked substantial external validation.

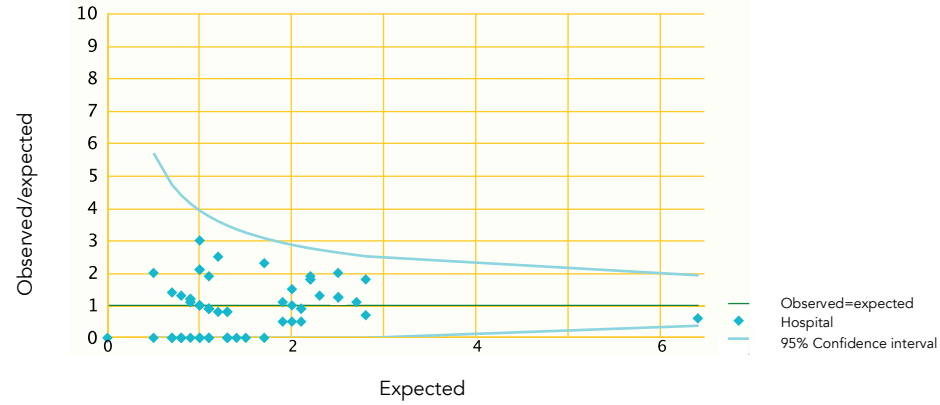


Figure 1A. Funnelplot for mortality by hospital in primary infra- or juxtarenal elective AAA patients (2015-2017).

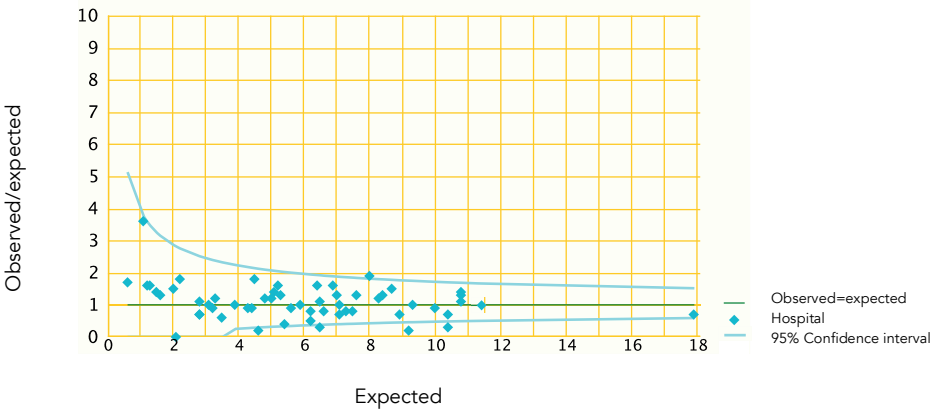


Figure 1B. Funnelplot for mortality by hospital in primary infra- or juxtarenal acute AAA patients (2015-2017).

The relevance of using casemix variables for comparing outcome might be limited according to the DSAA, because the variation in outcome between hospitals is low.[11] Also, most mortality risk prediction models have been developed in the past and have given varying results in different populations. In **Chapter 4** the development of the more simple Dutch casemix model is further described and has now replaced the more extensive V(p)-POSSUM in the DSAA.[11] The optimal set of independent variables associated with mortality were a mix of clinical and laboratory parameters: gender, age, pulmonary comorbidity, operative setting, creatinine, aneurysm size, hemoglobin, Glasgow Coma Scale, ECG and systolic blood pressure (C- statistic 0.871). However, as previously mentioned, the influence of casemix becomes less important if high quality of health care leads to very little variation between hospitals as in EAAA surgery in the Netherlands.

In **Chapter 5** the possible influence on the mortality of a higher or lower preference of a hospital for EVAR or OSR is examined, while also examining the current difference in patient level mortality between EVAR and OSR.[12] Hospitals can have different preferences, motivated by anatomical characteristics of the aneurysm, comorbidity of the patient or preference of the practitioner. Since, as previously concluded, mortality after EAAA is more favourable in patients who have undergone EVAR, the question arises whether a high hospital preference for EVAR would also result in a lower total mortality compared to hospitals with a lower preference for EVAR. If two treatments are compared in observational research, a reliable estimation should be corrected for (known) confounders. In order to approximate the estimation of a randomised trial, it should also be corrected for unknown confounders. The latter can be done by means of instrumental variable (IV) analysis.[13] If it can be assumed that all hospitals have about the same patients mix and the outcomes of EVAR or OSR are not dependent on the (skills of the) practitioner or hospital, then from the treatment difference/preference between hospitals for EVAR or OSR the difference in mortality between EVAR and OSR are calculated. We found that a stronger hospital preference for EVAR does not lead to a lower total mortality. But also that after IV analysis no difference in mortality between EVAR and OSR seemed to exist, while classic multivariable analysis-with correction for known casemix factors-resulted in a significantly lower mortality rate for EVAR compared to OSR. This seems a paradox and has been further investigated. On closer examination, the mortality rate after OSR in hospitals with more preference for EVAR (high% EVAR hospitals) was significantly higher than in low% EVAR hospitals: 7.3% versus 4.0%. This may have disrupted the IV analysis. This is a great concern and will need to be discussed within the NVVV.[3]

As stated in **Chapter 2**, mortality as single outcome indicator is no longer useful to discriminate between hospitals since the widespread use of EVAR.[9] Moreover, indicators reflecting the processes or structure of a hospital also provide internal information for hospital providers that is valuable. A hospital might perform well on one quality indicator and worse on another, reflecting one of the problems of single quality indicators.[14]

In **Chapter 6** we describe an example of a composite outcome measure called Failure to Rescue (FTR).[15] It reflects the ability to overcome major complications and avoid consequent death. Using FTR the variation between hospitals performing AAA surgery in the Netherlands is higher than by using the single outcome indicator mortality. Mortality of elective and acute aneurysm surgery together was 6.6% (range 0-16%) while FTR was 28.4% (range 0-60%). However, a high or low complication rate may have a low of high FTR rate. Therefore, we constructed so-called iso-mortality lines that can be implemented in the funnel plot of FTR, to create a three dimensional figure. These lines show that mortality is in conjunction with the major complication rate and the FTR and provides information necessary for the interpretation of FTR. This is essential due to the fact that mortality rates are calculated by overall hospital volume while FTR is calculated over a limited number of major complications by hospital. A great advantage of FTR compared to mortality is that it can be calculated for the overall group of patients since the main issue of this indicator refers to the efficient treatment of major complications. Unfortunately, statistical comparisons are not possible, given the composite nature of this measure, but this indicator can be used for internal quality improvement.

INTERNATIONAL PERSPECTIVES

With Sweden as one of the leading registry countries collecting data by medical specialists on quality of care for a broad spectrum of diseases, the Swedish vascular registry 'SWEDVASC' was developed in 1987, and registers more than 90% of all aneurysm operations in Sweden.[16, 17] Together with other international registries they are unified into the International Consortium for Vascular Registries (ICVR).[18] Most European vascular audits also provide data to Vascunet, an initiative of the European Society for Vascular Surgery (ESVS).[19, 20] Coverage of every patient undergoing surgery is a problem for most audits. A lower coverage may lead to possible selection bias. Data verification of the DSAA for the year 2015 revealed coverage of 98.4%. Mortality was not missed.[21] However, other difficulties remain regarding differences in registration of type of aneurysm or the classification of symptomatic AAA patients as either RAAA or EAAA patients. Recently, there seems to be an international agreement that AAA sur-

gery is classified in intact aneurysm versus ruptured aneurysm. An acute symptomatic aneurysm is then classified as intact.[19]

FUTURE PERSPECTIVES

EVAR versus OSR

From the DSAA can be concluded that an '' EVAR preference policy'' may lead to inferior results for OSR, resulting in equal overall mortalities in low% and high% EVAR hospitals. Moreover, the results following OSR especially in high% EVAR hospitals are rather worrying and need further attention. Considering the higher complexity of OSR, not only with regard to the operative technique, but also the perioperative care, further concentration of care is warranted. But the additional advantages of concentration of care would be, improved quality measuring and an improved quality cycle time, both due to the effect of larger hospital volumes.

Patient reported experience and outcome measurements

Quality of care viewed from the patients' and social perspective has become increasingly important. Patient reported experience measurements (PREMs) and patient reported outcome measurements (PROMs) measure quality from a patient's perspective. However, in AAA surgery, only very few qualitative studies have reported PROMs.[22] The development of PROMs in AAA surgery will require the exploration of the patients pathway in several manners like for example the speed of revalidation, the experience of pain or the presence of scarring tissue, physical activity or food intake and questions about quality of life inside the hospital but also outside the hospital evaluating long term effects. Most of these elements are present in quality of life questionnaires and could be validated in order to find relevant PROMs.[22] In addition to the identification of the right PROMs, future research should focus on the way PROMs can be used in daily patient care, how individual results can be benchmarked, interpreted in a valid way and lead to evidence-based interventions for individual patients to improve outcomes important to them.

Composite quality measures

The focus on health care quality by participation in a national clinical audit has shown to eventually lead to medical cost reduction in some audits.[23] Next to FTR other composite measures can be investigated. FTR reflects non-desired outcomes, but with a composite quality measure like 'Textbook Outcome', desired patient outcome is reflected by combining key elements, like duration of hospital admission and an uneventful course.[14] Such composite measures, encompassing the entire care process,

can be very useful in monitoring all important aspects of a complex care path. For example, when the number of patients with a 'Textbook outcome' is relatively low, clinicians can evaluate which steps in their care process need improvement and plan interventions.

CONCLUSION

The outcomes following aneurysm surgery in the Netherlands are good and comparable with international registries. There are no significant differences between the Dutch hospitals. Results following EVAR are especially favourable, but mortality following OSR is high (5%), especially in hospitals with a preference for EVAR. Further research and probable specific measures will be necessary. Risk adjustment for casemix will remain necessary for surgery with large variation in outcome, but a minimal set of variables will do. FTR is a complex quality indicator and may only be interpreted by means of iso-mortality lines but its value for quality improvement purposes remains unclear.

REFERENCES

1 Dutch Institute for Clinical Auditing. Available at: www.clinicalaudit.nl.
2 Dutch Surgical Aneurysm Audit. Available at: https://www.dica.nl/dsaa/home.
3 Dutch Society of Vascular Surgery. Available at: http://nvvv-vaatchirurgie.nl/.
4 NVvH. Available at: https://heelkunde.nl/.
5 Vascular patient society. Available at: http://www.harteraad.nl/.
6 Zorginstituut Nederland. Available at: https://www.zorginstituutnederland.nl/.
7 Healthcare inspectorate. Available at: https://www.igz.nl/.
8 Risk Prediction in Surgery - Vascular POSSUM scoring. Available at: http://www.riskprediction.org.uk/vasc-index.php.
9 Lijftogt N, Vahl AC, Wilschut ED, et al. Adjusted Hospital Outcomes of Abdominal Aortic Aneurysm Surgery Reported in the Dutch Surgical Aneurysm Audit. Eur J Vasc Endovasc Surg 2017;53:520-32.
10 Lijftogt N, Luijnenburg TWF, Wilschut ED, et al. Systematic review of mortality risk prediction models in the era of endovascular abdominal aortic aneurysm surgery. Br J Surg 2017 Jul;104(8):964-976.
11 Niki Lijftogt MD, Anco Vahl M, PhD, Esmee M van der Willik, MSc, et al. Towards Optimizing Risk Adjustment in the Dutch Surgical Aneurysm Audit (DSAA). Ann Vasc Surg 2019 May 7. pii: S0890-5096(19)30287-0. doi: 10.1016/j.avsg.2019.02.032.
12 Lijftogt, N., Vahl, A.C. , Karthaus, E.G., van der Willik, E.M., Amodio, S., van Zwet, E.W. Hamming, J.F., In collaboration with the Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing. High postoperative mortality after elective open surgical abdominal aneurysm repair in hospitals with EVAR preference and lessons from an instrumental variable analysis of the Dutch Surgical Aneurysm Audit. Br J Surg 2019 Submitted.
13 Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med 2014;33:2297-340
14 Karthaus E, Lijftogt N, Busweiler L, et al. Textbook Outcome: a Composite Measure for Quality of Elective Aneurysm Surgery. Ann Surg 2017 Nov;266(5):898-904.
15 Lijftogt N, Karthaus EG, Vahl A, et al. Failure to Rescue - a Closer Look at Mortality Rates Has No Added Value for Hospital Comparisons but Is Useful for Team Quality Assessment in Abdominal Aortic Aneurysm Surgery in The Netherlands. Eur J Vasc Endovasc Surg 2018 Nov;56(5):652-661.
16 Emilsson L, Lindahl B, Koster M, et al. Review of 103 Swedish Healthcare Quality Registries. J Intern Med 2015;277:94-136.
17 SWEDVASC. Available at: http://www.ucr.uu.se/swedvasc/arsrapporter.
18 International Consortium of Vascular Registries. Available at: http://icvr-initiative.org/about-us.aspx.
19 VASCUNET. Available at: http://vascunet.org/.
20 European Society for Vascular Surgery. Available at: https://www.esvs.org/.
21 Dataverification DSAA. Available at: http://dica.nl/media/660/Eindrapport_dataverificatie_DSAA_2016.pdf.
22 Duncan R, Essat M, Jones G, et al. Systematic review and qualitative evidence synthesis of patient-reported outcome measures for abdominal aortic aneurysm. Br J Surg 2017;104:317-27.

23 Govaert JA, van Dijk WA, Fiocco M, et al. Nationwide Outcomes Measurement in Colorectal Cancer Surgery: Improving Quality and Reducing Costs. J Am Coll Surg 2016;222:19,29.e2.

8

Dutch summary

DUTCH SURGICAL ANEURYSM AUDIT

In Nederland zijn diverse chirurgische audits ontwikkeld in samenwerking met de Dutch Institute for Clinical Auditing. Er zijn momenteel 22 audits die worden uitgevoerd in opdracht van verschillende chirurgische en ook niet-chirurgische specialismen.

In de web-based DSAA, geïntroduceerd in 2012 en verplicht sinds 2013, worden de gegevens geregistreerd van alle patiënten die in Nederland zijn geopereerd aan een primair aneurysma van de aorta abdominalis (AAA). Sinds 2016 worden ook de thoracale aneurysmata, de dissecties en de chirurgische revisies geregistreerd. Momenteel bevat de DSAA de gegevens van patiënten uit 60 ziekenhuizen met zowel electief geplande AAA (EAAA) patiënten als ook patiënten met een acuut symptomatisch AAA (SAAA) of acuut geruptureerd AAA (RAAA), die een open chirurgische operatie (OSR) of endovasculaire aneurysma ingreep (EVAR) hebben ondergaan.

De belangrijkste kwaliteitsindicatoren die uit de DSAA gehaald worden zijn onder andere uitkomsten zoals de postoperatieve mortaliteit binnen 30 dagen of tijdens de ziekenhuisopname en complicaties, maar ook processen van zorg en de structurele kenmerken van het ziekenhuis.

De Nederlandse Vereniging voor Vaatchirurgie (NVVV) heeft in samenwerking met diverse stakeholders de ontwikkeling van de set kwaliteitsindicatoren geïnitieerd, die rechtstreeks uit de DSAA kunnen worden gegenereerd.

In **Hoofdstuk 1**, wordt een kort overzicht gegeven van de ontwikkeling van de DSAA en het belang van kwaliteitsregistraties. Bij het meten van uitkomst indicatoren moeten de verschillen tussen ziekenhuizen gecorrigeerd worden voor de zogenaamde casemix, die de complexiteit van de patiënten weergeeft. Variabelen die geassocieerd zijn met de postoperatieve mortaliteit zijn vaak opgenomen in mortaliteitspredictiemodellen zoals V-POSSUM - een model met veel variabelen of de V(p)-POSSUM met alleen fysiologische variabelen. Aangezien het gebruik van enkelvoudige uitkomst parameters zoals mortaliteit of postoperatieve complicaties kan leiden tot een beperking in het inzicht van de kwaliteit van de geleverde zorg, zouden alternatieve of samengestelde uitkomst parameters van toegevoegde waarde kunnen zijn om de resultaten tussen ziekenhuizen te vergelijken.

In **Hoofdstuk 2** worden de eerste resultaten van de DSAA gepresenteerd. De mortaliteit binnen 30 dagen en/of in het ziekenhuis was 1,9% voor electieve AAA (EAAA) en daarmee vergelijkbaar met andere Europese registraties. Mortaliteit na electieve

EVAR was 0,9% en tevens vergelijkbaar met andere registraties. Echter, de mortaliteit na open chirurgie (OSR) in de DSAA was hoger vergeleken met enkele andere landen, zoals bijvoorbeeld de Zweedse registratie SWEDVASC (5,0% versus 3,2%). Bij RAAA was de sterfte na EVAR en OSR respectievelijk 22,2% en 32,0% en vergelijkbaar met de resultaten van de SWEDVASC. Internationale vergelijkingen zijn echter nog lastig gezien de verschillen in registratie en het al dan niet verplichte karakter.

Vanwege de lage mortaliteit na EAAA was de voor casemix gecorrigeerde en niet-gecorrigeerde variatie tussen ziekenhuizen onvoldoende om onderscheid te kunnen maken tussen ziekenhuizen. Correctie aan de hand van V(p) - POSSUM had geen invloed op de variatie tussen ziekenhuizen. Het discriminerende vermogen van V-POSSUM voor het voorspellen van de mortaliteit was redelijk voor zowel EAAA als AAAA, met respectievelijk een C-statistic van 0,719 en 0,713. Ook in de drie meest recente jaren van registratie bleken er nog steeds geen statistisch significante verschillen tussen ziekenhuizen te bestaan.

Op zoek naar de juiste casemix variabelen, hebben we in **Hoofdstuk 3** de best voorspellende mortaliteitspredictie modellen besproken met behulp van een systematic review. Geen enkel model blijkt standaard gebruikt te worden in de klinische praktijk. De beste voorspellende modellen op het gebied van de toepasbaarheid en het discriminerend vermogen waren de perioperatieve British Aneurysm Repair score (C-statistic 0,83) en de preoperatieve Vascular Biochemistry and Haematology Outcome Model (C-statistic 0,85). De mate van externe validatie van beide modellen bleek echter matig te zijn.

Wegens de minimale variatie tussen de ziekenhuizen in de DSAA is de waarde van casemix correctie voor het vergelijken van de ziekenhuis resultaten beperkt. Bovendien zijn de meeste mortaliteitspredictiemodellen ontwikkeld in het verleden en niet goed toepasbaar op het huidige beleid. Daarnaast verschillen de resultaten in verschillende populaties. In **Hoofdstuk 4** wordt de ontwikkeling van een vereenvoudigd Nederlands casemix model voor de DSAA beschreven als vervanger voor de meer uitgebreide V(p)-POSSUM. Dit model wordt vanaf 2019 toegepast in de DSAA. Het optimale aantal onafhankelijke variabelen geassocieerd met mortaliteit waren een mix van klinische -en laboratoriumparameters: geslacht, leeftijd, pulmonale co-morbiditeit, operatieve setting, kreatinine, aneurysma diameter, hemoglobine, Glasgow Coma Scale, ECG en systolische bloeddruk (C-statistic 0,871). De invloed van casemix blijkt echter minder belangrijk geworden wegens de beperkte ziekenhuisvariatie in mortaliteit na EAAA chirurgie in Nederland.

In **Hoofdstuk 5** wordt de invloed op de mortaliteit van de voorkeur voor EVAR of OSR onderzocht. De keuze voor behandeling hangt af van de anatomische kenmerken van het aneurysma, de co-morbiditeit van de patiënt maar ook van de voorkeur van de behandelaar in het ziekenhuis. Aangezien de mortaliteit na EAAA gunstiger is bij patiënten die EVAR hebben ondergaan, rijst de vraag of een hoge ziekenhuisvoorkeur voor EVAR ook zou resulteren in een lagere totale mortaliteit in vergelijking met ziekenhuizen met een lagere voorkeur voor EVAR. Als in observationeel onderzoek twee behandelingen worden vergeleken dan dient voor een betrouwbare schatting er gecorrigeerd te worden voor (bekende) confounders. Om de schatting die van een gerandomiseerde trial te laten benaderen dient ook gecorrigeerd te worden voor onbekende confounders. Dit laatste kan worden gedaan door middel van Instrumentele Variabele (IV) analyse. Indien er vanuit kan worden gegaan dat alle ziekenhuizen ongeveer dezelfde patiënten mix hebben en de uitkomsten van EVAR of OSR niet afhankelijk zijn van de (kunde van) behandelaar of ziekenhuis, dan kan door het behandelverschil/voorkeur tussen ziekenhuizen voor EVAR of OSR het verschil in mortaliteit tussen EVAR en OSR worden berekend aan de hand van pseudorandomisatie. Uit deze analyse bleek dat een sterkere voorkeur van een ziekenhuis voor EVAR niet leidt tot een lagere totale mortaliteit. Daarnaast bleek dat na IV analyse er geen verschil meer was in mortaliteit tussen EVAR en OSR, terwijl de klassieke multivariabele regressie analyse - met correctie voor bekende casemix factoren – wederom resulteerde in een significant lagere mortaliteit voor EVAR vergeleken met OSR. Dit lijkt een paradox en is nader onderzocht.

Bij nader onderzoek bleek de mortaliteit na OSR in ziekenhuizen met meer voorkeur voor EVAR (hoog % EVAR ziekenhuizen) significant hoger dan in laag % EVAR ziekenhuizen: 7,3% versus 4,0%. Het verschil in kwaliteit tussen de ziekenhuizen op het gebied van OSR kan een verstoring geven van de IV analyse. Echter, de significant hogere mortaliteit van OSR in ziekenhuizen met een voorkeur voor EVAR is een grote zorg en dient nader onderzocht te worden.

Zoals aangegeven in **Hoofdstuk 2** is de mortaliteit van AAA chirurgie sinds de invoer van de EVAR procedure zodanig afgenomen dat er geen significante verschillen meer bestaan tussen de ziekenhuizen. Daarnaast worden indicatoren die tevens informatie geven over de processen of structuur van een ziekenhuis steeds belangrijker voor interne ziekenhuis evaluaties. Bovendien kan een ziekenhuis goed presteren op de ene kwaliteitsindicator en slechter op een andere, wat een van de problemen van een enkelvoudige kwaliteitsindicator weergeeft. Daarom werd onderzocht of met een samengestelde uitkomstmaat wel een verschil tussen ziekenhuizen kon worden gevonden.

In **Hoofdstuk 6** beschrijven we een voorbeeld van een samengestelde uitkomstparameter 'Failure to Rescue' (FTR). Deze parameter geeft het vermogen weer om een ernstige complicaties op te lossen en zo de gerelateerde mortaliteit te voorkomen. FTR zou een gevoeligere uitkomstmaat kunnen zijn en zo de mogelijke variatie tussen ziekenhuizen in Nederland na AAA chirurgie inzichtelijk kunnen maken. De mortaliteit van electieve en acute AAA chirurgie, samen, bedroeg 6,6% (bereik 0-16%), maar de FTR was 28,4% (bereik 0-60%). Echter een laag of hoog complicatiepercentage kan zowel een laag als hoog FTR hebben. Daarom hebben we iso-mortaliteitslijnen ontwikkeld die in de grafiek van de FTR kunnen worden geïmplementeerd, zodat een driedimensionale figuur ontstaat. Deze lijnen geven de mortaliteit weer in relatie tot het percentage ernstige complicaties en FTR. Dit is noodzakelijk doordat de mortaliteit wordt berekend over het totale ziekenhuis volume terwijl FTR wordt berekend over een beperkt aantal ernstige complicaties. Een groot voordeel van FTR ten opzichte van mortaliteit is dat het in theorie berekend kan worden over de gehele groep patiënten, ongeacht operatieve setting, omdat het gaat om de effectiviteit van de behandeling van ernstige complicaties.

Omdat het een variabele is die lastig is te interpreteren en ook statistisch gezien nadelen heeft is deze parameter vooral geschikt om te gebruiken voor interne kwaliteitsmetingen maar niet voor ziekenhuisvergelijkingen.

INTERNATIONALE PERSPECTIEVEN

Zweden is één van de toonaangevende landen op het gebied van het aantal nationale registraties, ontwikkeld en gecontroleerd door medische specialisten, waarbij gegevens worden verzameld over de kwaliteit van zorg voor een breed spectrum van ziekten. De Zweedse vaatregistratie 'SWEDVASC' werd ontwikkeld in 1987, en registreert nu meer dan 90% van alle aneurysma operaties in Zweden. Samen met andere internationale registraties, zijn ze verenigd in het International Consortium voor Vascular Registries (ICVR). De meeste Europese vaatregistraties verstrekken ook hun gegevens aan Vascunet, een initiatief van de European Society for Vascular Surgery (ESVS). Het volledig registreren van elke patiënt die een operatie heeft ondergaan is een probleem voor de meeste registraties. Bij een niet volledige registratie van alle patiënten kan er sprake zijn van selectie bias. Bij verificatie van de DSAA bleek er sprake van een registratie van 98,4% van de patiënten. Mortaliteit werd niet gemist. Echter, volledigheid van een registratie is niet het enige probleem voor de vergelijking tussen de registraties. Zo zijn er registraties die alleen infrarenale aneurysmata registreren in plaats van ook juxtarenale aneurysmata. Daarnaast blijkt ook dat er nog discussie is over de analyse van

symptomatische aneurysmata en of die kan worden samengevoegd met ofwel RAAA ofwel EAAA. Recent lijkt er echter internationaal overeenstemming te zijn bereikt over de indeling van aneurysmata gedefinieerd als intact aneurysma versus geruptureerd aneurysma. Een acuut symptomatisch aneurysma wordt dan geclassificeerd als intact.

TOEKOMSTPERSPECTIEVEN

EVAR versus OSR

Uit de DSAA blijkt dat een voorkeur voor EVAR kan leiden tot minder goede resultaten voor OSR, wat resulteert in gelijke mortaliteit bij ziekenhuizen met een hoge versus een lage preferentie voor EVAR. Vooral de resultaten na OSR en dan vooral in hoog % EVAR ziekenhuizen zijn zorgelijk en moeten nader worden onderzocht. De hogere complexiteit van OSR kan leiden tot de noodzaak van verdere centralisatie van zorg.

Patient reported experience and outcome measurements

De kwaliteit van de vaatchirurgie in Nederland is hoog en dat is belangrijk voor vaatchirurgisch Nederland. Maar kwaliteit gemeten vanuit een patiënten perspectief is eveneens steeds belangrijker geworden. Patient reported experience measurements (PREMs) en de Patient reported outcome measurements (PROMs) meten kwaliteit vanuit een patiënten perspectief. In de aneurysma chirurgie zijn er tot nu toe slechts een aantal studies die de PROMs bespreken. De ontwikkeling van de PROMs voor AAA patiënten kan plaatsvinden aan de hand van kwaliteit van leven vragenlijsten en verder worden gevalideerd om relevante PROMs te ontdekken.

Samengestelde kwaliteitsparameters

De focus op de kwaliteit van de zorg door het opzetten van kwaliteitsregistraties heeft geleid tot een vermindering van de medische kosten bij andere vormen van chirurgie. Naast FTR kunnen ook andere samengestelde kwaliteitsparameters worden onderzocht. FTR toont niet-gewenste resultaten, maar een samengestelde kwaliteitsparameter zoals 'Textbook Outcome', geeft juist het gewenste patiënten resultaat weer door de belangrijkste elementen in het vaatchirurgische proces te combineren in relatie tot het ziekenhuisvolume, aan de hand van bijvoorbeeld de duur van de ziekenhuisopname en een ongecompliceerd beloop. Deze samengestelde parameters tezamen omvatten het gehele zorg proces en zijn daarom erg handig om alle belangrijke aspecten in een complex zorg pad te kunnen monitoren.

CONCLUSIE

De uitkomsten van de aneurysmachirurgie in Nederland zijn goed en vergelijkbaar met internationale registraties. Er zijn geen significante verschillen tussen de Nederlandse ziekenhuizen, de resultaten na EVAR zijn beter, maar tegelijkertijd zijn de resultaten na OSR zorgelijk, vooral in ziekenhuizen met een voorkeur voor EVAR. Verder onderzoek en specifieke maatregelen zijn hiervoor noodzakelijk. Correctie voor casemix blijft nodig, maar een minimale set van variabelen is daarvoor voldoende. FTR als uitkomstmaat is complex, alleen te interpreteren met behulp van iso-mortaliteitslijnen en de werkelijke waarde als uitkomstmaat voor kwaliteit van deze variabele is nog onduidelijk.

DANKWOORD

Bij het schrijven van een proefschrift zijn het de mensen om je heen die het verschil maken. Ik wil graag **alle mensen** bedanken die mij de draagkracht hebben gegeven dit intensieve project af te ronden en een aantal in het bijzonder.

Prof. dr. Hamming, beste Jaap, hartelijk dank voor al het vertrouwen en het geduld in de afgelopen jaren. Jij hebt de definitieve richting aan mijn promotie gegeven.

Dr. Vahl, beste Anco, met engelen geduld wist jij mijn ongeleid enthousiasme te structureren. Met een reactie snelheid gelijk aan het licht stond je al die tijd vol overtuiging klaar. Onvermoeibaar in de zoektocht naar antwoorden heb je mij altijd gemotiveerd door te gaan.

Dr. Wouters, beste Michel, veel dank voor jouw begeleiding. Jouw rustige manier van argumenten overwegen en bespreken hebben me enorm veel geleerd. Daarnaast heb jij samen met **Rob, prof. dr. Tollenaar**, mij de kans te geven deel uit te maken van een van de belangrijkste zorginitiatieven in Nederland, DICA. **Beste Rob**, dankjewel dat ook jij dit voor me mogelijk hebt gemaakt.

Leden van de wetenschappelijke commissie van de DSAA en de **registrerende vaatchirurgen en medewerkers** in de ziekenhuizen. Jullie ervaring en de geregistreerde data hebben de kwaliteit van de aneurysmachirurgie in Nederland in beeld gebracht.

Mede - auteurs, zonder jullie waren deze publicaties niet mogelijk geweest.

Arts onderzoekers bij DICA en het LUMC, congressen, hoogtepunten en dieptepunten. Die herinneringen blijven!

Collega's van de SOA poli, in het OLVG, Dijklander ziekenhuis en nu het Erasmus MC, langs deze tocht heeft mijn verdere loopbaan zich uitgestippeld. Dank voor jullie geduld en steun op goede en minder goede momenten.

Lieve paranimfen en Cracco, jullie hebben de laatste loodjes een stralende glans gegeven. Dat was jullie dan ook geheel toevertrouwd.

Voor altijd VINO, geneeskunde vriendinnen voor altijd.

Allerliefste clubbies, roeimaten, de schatjes en alle vrienden die mijn leven meerwaarde geven op zoveel vlakken, dank jullie wel dat ook een afspraak over 3 maanden gewoon mogelijk mocht zijn.

Lieve schoonfamilie, dankjewel dat ik jullie erbij heb gekregen.

Lieve familie, klein maar zo fijn dat jullie er zijn.

Lieve pap, ik heb er wat van gemaakt.

Lieve mam, ik weet niet wat ik zonder je had gemoeten.

Lieve Haroen, mijn maatje, nu we dit hebben doorstaan kunnen we de wereld aan!

CURRICULUM VITAE

Niki Lijftogt was born in Soest, in the Netherlands, on April 7th 1986. She graduated from the Baarnsch Lyceum in Baarn in 2004 and started studying medicine at Utrecht University that same year.

At the start of her study Niki became a member of the students rowing club U.S.R. Triton, where she was a fresh-years cox swan and had the privilege of winning the Varsity, an important national rowing regatta. In the following years being a cox swan in a rowing team was replaced for leading sports classes as one of her favourite activities. Next to her internships the sports classes were an ultimate performance but rendered energy the most.

When Niki received her doctoral degree in 2011 she started working as an intern, not in training, at the department of surgery at the Sint Antonius Hospital in Nieuwegein. Following, she started as a PhD student at the Dutch Institute for Clinical Auditing, working on Quality of Care in Abdominal Aortic Aneurysm Surgery. Meanwhile, epidemiology became one of her special interests with an always-present curiosity leading to the current chapters in this manuscript.

Then, after several months of working again in the surgical clinical field of the OLVG in Amsterdam, Niki decided to switch to dermatology and started working at the STI clinic at the GGD in Amsterdam. Since the end of 2017 she has been working as an intern, not in training, at the department of dermatology, and now resides in Rotterdam, which is the perfect setting to further develop her skills in dermatology.

More will be announced in the case of interviews for a residency in dermatology in the upcoming period.

LIST OF PUBLICATIONS

Lijftogt N, Cancrinus E, Hoogervorst EL, van de Mortel RH, de Vries JP. Median nerve neuropraxia by a large false brachial artery aneurysm. Vascular. 2014 Oct;22(5):378-80.

Geubbels N, Lijftogt N, Fiocco M, van Leersum NJ, Wouters MW, de Brauw LM. Meta-analysis of internal herniation after gastric bypass surgery. Br J Surg. 2015 Apr;102(5):451-60.

Govaert JA, Lijftogt N, van Dijk WA, Tseng LN, Liem RS, Tollenaar RA, Fiocco M, Wouters MW; Dutch Value Based Healthcare Study Group. Colorectal cancer surgery for obese patients: Financial and clinical outcomes of a Dutch population-based registry. J Surg Oncol. 2016 Apr;113(5):489-95.

Karthaus EG, Lijftogt N, Busweiler LAD, Elsman BHP, Wouters MWJM, Vahl AC, Hamming JF; Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit, the Dutch Institute for Clinical Auditing. Textbook Outcome: A Composite Measure for Quality of Elective Aneurysm Surgery. Ann Surg. 2017 Nov;266(5):898-904.

Tomee SM, Lijftogt N, Vahl A, Hamming JF, Lindeman JHN. A registry-based rationale for discrete intervention thresholds for open and endovascular elective abdominal aortic aneurysm repair in female patients. J Vasc Surg. 2018 Mar;67(3):735-739.

Poelemeijer YQM, Lijftogt N, Detering R, Fiocco M, Tollenaar RAEM, Wouters MWJM. Obesity as a determinant of perioperative and postoperative outcome in patients following colorectal cancer surgery: A population-based study (2009-2016). Eur J Surg Oncol. 2018 Dec;44(12):1849-1857.

Lijftogt N, Vahl AC, Wilschut ED, Elsman BHP, Amodio S, van Zwet EW, Leijdekkers VJ, Wouters MWJM, Hamming JF; Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit, and the Dutch Institute for Clinical Auditing.
Adjusted Hospital Outcomes of Abdominal Aortic Aneurysm Surgery Reported in the Dutch Surgical Aneurysm Audit. Eur J Vasc Endovasc Surg. 2017 Apr;53(4):520-532.

Lijftogt N, Luijnenburg TWF, Vahl AC, Wilschut ED, Leijdekkers VJ, Fiocco MF, Wouters MWJM, Hamming JF. Systematic review of mortality risk prediction models in the era of endovascular abdominal aortic aneurysm surgery. Br J Surg. 2017 Jul;104(8):964-976.

Lijftogt N, Karthaus EG, Vahl A, van Zwet EW, van der Willik EM, Tollenaar RAEM, Hamming JF, Wouters MWJM; Dutch Society of Vascular Surgery; Steering Committee of the Dutch Surgical Aneurysm Audit; Dutch Institute for Clinical Auditing. Failure To Rescue - a closer look at mortality rates - has no added value for hospital comparisons but is useful for team quality assessment in Abdominal Aortic Aneurysm Surgery in the Netherlands. Eur J Vasc Endovasc Surg. 2018 Nov;56(5):652-661.

Lijftogt N, Vahl AC, van der Willik EM, Leijdekkers VJ, Wouters MWJM, Hamming JF; Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing.
Towards Optimizing Risk Adjustment in the Dutch Surgical Aneurysm Audit (DSAA). Ann Vasc Surg. 2019 May 7. pii: S0890-5096(19)30287-0. doi: 10.1016/j.avsg.2019.02.032. [Epub ahead of print]

APPENDIX CHAPTER 3

Appendix S1 Search strategy

Pubmed: ("Aortic Aneurysm, Abdominal"[Mesh] OR (("aneurysm"[MeSH] OR "aneurysm"[All Fields] OR "aneurysms"[All Fields] OR "aneurysmatic"[All Fields] OR "aneurysmata"[All Fields]) AND "abdominal"[all fields])) AND ("Mortality"[Mesh] OR "mortality" [Subheading] OR mortal*[all fields] OR "death rate"[all fields] OR "death rates"[all fields] OR "survival"[all fields] OR "Survival"[Mesh]) AND ("2006/01/01"[PDAT] : "3000/12/31"[PDAT])

6th July 2015: 3274

Embase: (exp *abdominal aorta aneurysm/ OR ((exp *aneurysm/ OR "aneurysm".ti. OR "aneurysms".ti. OR "aneurysmatic".ti. OR "aneurysmata".ti.) AND "abdominal".ti.)) AND (exp Mortality/ OR Mortal*.ti. OR "death rate".ti. OR "death rates".ti. OR "survival".ti. OR exp survival/)

limit 2006-current NOT conference abstracts.

6th July 2015: 2073

Cochrane Library: ("Aortic Aneurysm, Abdominal" OR (("aneurysm" OR "aneurysms" OR "aneurysmatic" OR "aneurysmata") AND "abdominal")) AND ("Mortality" OR mortal* OR "death rate" OR "death rates" OR "Survival")

2006-2015 trials only

6th July 2015:136

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance.

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Tang et al.[1]	2007	UK	1998–2004	PRS (E-PASS)	204	P: Cambridge, UK	EAAA	OSR	(b)	n.a.	n.a.	+	In-hospital mortality
				SSS (E-PASS)									
				CRS (E-PASS)									
Patterson et al.[2]	2011	UK	2005–2007	GAS	330	R: South UK (same cohort)	EAAA	OSR	(b)	n.a.	n.a.	+	30-day mortality
Sandford et al.[3]	2007	UK	18 months	P-POSSUM	136	P: Leicester, UK	EAAA RAAA	OSR	(b)	n.a.	n.a.	+	In-hospital mortality
				RAAA-POSSUM									
				V-POSSUM									
				SAPS-II									
				APACHE-II									
Supsamutchai et al.[4]	2008	Thailand	1996–2007	P-POSSUM	146	R: Ramathibodi, Thailand	EAAA RAAA SAAA	both (6% EVAR)	(b)	n.a.	n.a.	+	Hospital mortality
				GAS									
				Hardman index									
Menezes et al.[5]	2011	Brazil	2000–2008	PRS	214	R: Campinas Academic Hospital Brazil	EAAA	OSR	(b)	n.a.	n.a.	+	Mortality
				CRS									
				PRS + renal imp.									
				PSI + cardiopulm.									
				cardiopulmren.									
Leo et al.[6]	2006	Italy	1996–2004	GAS	114	R: St. Camillo-Forlanini, Rome, Italy	RAAA	OSR + other type of vascular reconstruction	(b)	n.a.	n.a.	+	Mortality
				Hardman Index									
				Chen calculated risk									

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Mani et al.[7]	2015	VASCUNET	2005–2009	GAS	5895	P: VASCUNET	EAAA	both (54% EVAR)	(b) (c)	n.a.	n.a.	+	Perioperative mortality
Kodama et al.[8]	2011	Japan	2001–2007	POSSUM	225	R:	EAAA	OSR	(b)	n.a.	n.a.	+	In-hospital mortality
Kurc et al.[9]	2012	Turkey	2000–2010	GAS	101	R: Istanbul, Turkey	RAAA	OSR	(b) (d)	n.a.	n.a.	+	In-hospital mortality
				Hardman Index	82								
Hirzalla et al.[10]	2006	The Netherlands	1994–2003	GAS	229	R: AMC Amsterdam, The Netherlands	EAAA	OSR	(b)	n.a.	n.a.	+	In-hospital Mortality
Kapma et al.[11]	2014	The Netherlands	1994–2010	GAS	206	R: OLVG Amsterdam, The Netherlands	RAAA	both (13.1% EVAR)	(b)	n.a.	n.a.	+	30-day mortality
				V-POSSUM									
Gatt et al.[12]	2009	UK	1998–2003	GAS	59	R: Scarborough, UK	RAAA	both (% EVAR Unknown)	(b)	n.a.	n.a.	+	30-day mortality
				Hardman index									
Faizer et al.[13]	2007	Canada	1999–2004	GAS	862	R: Ontario, Canada	EAAA	both (34.8% EVAR)	(b)	n.a.	n.a.	+	30-day mortality
				M-LS									
				SVS (M-CSS)									
Bohm et al.[14]	2008	UK	2001–2007	GAS	266	P: St Georges, London, UK	EAAA	EVAR	(b)	n.a.	n.a.	+	30-day Mortality
				V-POSSUM									
				m-CPI									
				CPI									

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Biancari et al.[15]	2006	Finland	1996–2005	GAS	5498	P: EUROSTAR registry	EAAA	EVAR	(b)	n.a.	n.a.	+	30-day Mortality
Baas et al.[16]	2008	The Netherlands	2000–2003	GAS	345	RCT/ Belgium and The Netherlands	EAAA	both (49.6% EVAR)	(b)	n.a.	n.a.	+	30-day mortality
Archan et al.[17]	2010	USA/Austria	1999–2006	Lee Index	225	R: n.a.	EAAA	OSR	b)	n.a.	n.a.	+	1–2 year all-cause mortality/MACE
Antonello et al.[18]	2007	Italy	1988–2005	GAS	42	R: Padua, Italy	SAAA	OSR	(b)	n.a.	n.a.	+	30-day or in-hospital mortality
Haji et al.[19]	2013	New Zealand	2000–2009	ASA/ Mount Sinai	54	R:	EAAA	EVAR	(b)	n.a.	n.a.	–	30-day mortality
Hicks et al.[20]	2015	USA	2003–2012	Medicare	297	R: Johns Hopkins	EAAA SAAA	both	(b)	n.a.	n.a.	–	Mortality
Sajid et al.[21]	2007	UK	1996–2005	GAS/Hardman index	71	R:	EAAA	EVAR	(b)	n.a.	n.a.	–	Mortality
Mani et al.[22]	2011	VASCUNET	2005–2009	AAA/RAAA	38 467	P: VASCUNET	RAAA AAA	both	(c)	+	+	+	30-day or in-hospital mortality
Acosta et al.[23]	2006	Sweden	2000–2004	Hardman index	162	R	RAAA	both	(d)	+	+	–	In- hospital mortality
Conroy et al.[24]	2011	UK	1994–2008	Hardman index	95	R	RAAA	EVAR	(d)	+	–	–	30-day mortality
Sharif et al.[25]	2007	Ireland	2001–2006	Hardman index	126	R	RAAA	both	(c)	+	+/–	–	Mortality
Anderson et al.[26]	2015	USA	2005–2010	Objective/ subjective NSQIP variables	14 042	P	RAAA SAAA EAAA	both	(c)	+	+	+	In-hospital mortality
Bhutta et al.[27]	2011	UK/Ireland	2003–2007	n.a.	1021	P	EAAA	both	(c)	+	+	+	2-year mortality
Carlisle et al.[28]	2015	UK	1999–2013	External Survival Calculator	1096	P	EAAA	both	(c)	+	+	+	1-5 year mortality

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Cho et al.[29]	2008	USA	2001–2007	n.a.	170	R	RAAA	OSR	(c)	+	+	+/–	Survival
Giordano et al.[30]	2009	Finland	1992–2008	GAS	68	R	RAAA	OSR	(c)	+	+	+/–	Mortality
Gupta et al.[31]	2014	USA	2005–2010	n.a.	11 229	P	EAAA	EVAR	(c)	+	+	+	30-day mortality (post-discharge)
Hawkins et al.[32]	2014	USA	2005–2010	n.a.	1893	P	RAAA	Both	(c)	+	+	+	30-day mortality
Ramanan et al.[33]	2013	USA	2007–2009	n.a.	2845	P	EAAA	OSR	(c)	+	+	+	30-day mortality
Davenport et al.[34]	2010	USA	2005–2007	n.a.	427	P	RAAA	both	(c)	+	+	+	30-day mortality
O'Malley et al.[35]	2011	USA	2001–2004	n.a.	61 414	R	EAAA	both	(c)	+	+/–	+/–	In-hospital or 30-day mortality
Scarcello et al.[36]	2010	Italy	2002–2007	TBS	94	R	RAAA	OSR	(c)	+	+	+	Mortality
Sharif et al.[37]	2007	UK	1999–2004	Hardman Index	178	R	RAAA	OSR	(c)	+	+/–	–	30-day mortality
Tambyraja et al.[38]	2007	UK	2000–2002	ERAS	105	P	RAAA	OSR	(c)	+	+/–	–	Mortality
Soong et al.[39]	2008	UK	1998–2005	n.a.	155	R	EAAA	EVAR	(c)	+	+/–	+	Perioperative or long term Mortality
Treska et al.[40]	2006	Czech Republic	1992–2005	n.a.	182	R	RAAA	both	(c)	+	+	–	30-day mortality
Haveman et al.[41]	2008	The Netherlands	1990–2003	APACHE -II	290	R	RAAA	OSR	(c)	+	+	–	In-hospital mortality
Stone et al.[42]	2013	USA	2003–2011	n.a.	3455	P	EAAA	both	(c)	+	+	–	In-hospital mortality
Powell et al.[43]	2014	UK/Canada	2009–2013	Hardman Index	613	RCT	RAAA SAAA	both	(d)	+	–	–	30-day mortality
Patel et al.[44]	2011	USA	2005–2008	SVS	3569	R	EAAA non-ruptured	both	(d)	+	–	–	30-day mortality

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Mofidi et al.[45]	2008	Scotland (UK)	1991–2006	n.a.	12 706	R	EAAA RAAA SAAA	both	(c)	+	+	–	In-hospital or 30-day mortality
Mehta et al.[46]	2013	USA	2002–2011	n.a.	283	P	RAAA	both	(c)	+	+	–	30-day mortality, survival
Lomazzi et al.[47]	2011	Italy	2000–2008	SVS/ASA	235	P	EAAA	EVAR	(c)	+	+	–	In-hospital or 30-day mortality
Lapar et al.[48]	2012	USA	2008	n.a.	261 412	R	EAAA majority	NA	(c)	+	+/–	+	Mortality
Giles et al.[49]	2010	USA	2005–2007	n.a.	5455	R	EAAA	both	(c)	+	+	–	Mortality
Gloviczki et al.[50]	2015	USA	1997–2011	SVS	934	R	RAAA SAAA EAAA	EVAR	(c)	+	+	+/–	Survival
Grant et al.[51]	2015	UK	2007–2012	CPET	506	P	EAAA	both	(c)	+	+	–	Survival
Hartley et al.[52]	2012	UK	2005–2011	CPET	415	P	EAAA	both	(c)	+	+	–	30/90-day mortality
Eckstein et al.[53]	2007	Germany	1999–2004	n.a.	10 163	P	EAAA	OSR	(c)	+	+	–	Mortality
Arya et al.[54]	2015	USA	2005–2012	n.a.	23 207	P	EAAA	both	(c)	+	+	+	30-day mortality
Azizzadeh et al.[55]	2006	USA	1999–2004	n.a.	398	R	EAAA not specified	EVAR	(d)	+	–	–	30-day mortality
Beck et al.[56]	2009	USA	2003–2007	n.a.	1387	P	EAAA	both	(c)	+	+	–	1- year mortality
Bonardelli et al.[57]	2011	Italy	n.a.	n.a.	137	P	RAAA	OSR	(c)/(d)	+	+/–	–	30-day mortality
Antonello et al.[58]	2009	Italy	1998–2006	GAS/APACHE -II	103	R	RAAA	OSR	(c)	+	–	–	In-hospital mortality
Robinson et al.[59]	2015	USA	2007–2011	AHRQ IQI #11	187 773	P	RAAA SAAA EAAA	both	(c)	+	+	+/–	Mortality
Acosta et al.[60]	2007	Sweden	2000–2004	n.a.	162	R	RAAA	both	(d)	+	–	–	In- hospital Mortality/ survival

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Chagpar et al.[61]	2010	Canada	2003–2008	n.a.	167	R	RAAA	both	(d)	+	–	–	30-day mortality
Chen et al.[62]	2013	Taiwan	1998–2009	n.a.	537	R	RAAA	both	(d)	+	–	–	Survival/in-hospital mortality
Bruin et al.[63]	2014	The Netherlands	2000–2003	n.a.	351	RCT	EAAA	both	(d)	+	–	–	Survival
Virgilio et al.[64]	2006	USA	1996–2005	n.a.	468	P	EAAA	EVAR	(d)	+	–	–	Survival
Diehm et al.[65]	2007	Switzerland	1994–2006	n.a.	711	P	EAAA	EVAR	(d)	+	+/–	–	Survival
Feeney et al.[66]	2009	USA	2000–2008	POSSUM	121	R	RAAA	both	(d)	+	–	–	Mortality
Hashimoto et al.[67]	2013	Japan	2000–2009	Hardman Index/ GAS	55	n.a.	RAAA	both	(d)	+	–	–	30-day mortality
Hughes et al.[68]	2013	USA	2005–2007	n.a.	2110	P	EAAA	OSR	(d)	+	–	–	30-day mortality
Hultgren et al.[69]	2007	Sweden	1987–2002	n.a.	12 917	R	EAAA RAAA SAAA	both	(d)	+	–	–	60-day mortality, survival
Karkos et al.[70]	2008	Greece	1998–2006	Hardman Index	41	R	RAAA	EVAR	(d)	+	–	–	30-day mortality
Kim et al.[71]	2012	South-Korea	1999–2008	n.a.	34	R	RAAA	OSR	(d)	+	–	–	2-day and 30-day mortality
Koga et al.[72]	2011	Japan	2005–2010	n.a.	42	R	RAAA SAAA	OSR	(d)	+	–	–	30-day mortality
Kunishige et al.[73]	2013	Japan	1984–2012	Rutherford and Fitzgerald - classification	105	R	RAAA Non-ruptured	OSR	(d)	+	–	–	Operative and in hospital mortality
Law et al.[74]	2015	China	1990–2012	n.a.	383	P	EAAA RAAA SAAA	OSR	(d)	+	–	–	30-day mortaiity
Lee et al.[75]	2011	USA	2000–2008	n.a.	479	R	EAAA	OSR	(d)	+	–	–	Mortality

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Lo et al.[76]	2013	USA	2003–2011	n.a.	4026	P	EAAA RAAA SAAA	both	(d)	+	–	–	in-hospital, 30-day mortality
Matsumara et al.[77]	2008	USA	1999–2000	n.a.	334	P	asymptom-atic	both	(d)	+	–	–	Mortality
Montan et al.[78]	2013	Sweden	2008–2011	n.a.	525	R	EAAA RAAA SAAA	EVAR	(d)	+	–	–	30-day mortality
Nakayama et al.[79]	2014	Japan	2003–2011	n.a.	1055	R	EAAA RAAA	both	(c)	+	–	+/–	30-day mortality
Ohrlander et al.[80]	2012	Sweden	1998–2006	n.a.	304	P	EAAA	EVAR	(d)	+	–	–	all cause long term mortality
Ohrlander et al.[81]	2012	Sweden	1998–2006	n.a.	304	P	EAAA	EVAR	(d)	+	–	–	long term mortality
Ohrlander et al.[82]	2011	Sweden	1998–2006	n.a.	304	P	EAAA	EVAR	(d)	+	–	–	long term mortality
Oyague et al.[83]	2015	USA	2009–2013	n.a.	49	R	RAAA SAAA	both	(d)	+	–	–	Mortality
Richards et al.[84]	2009	UK	1994–2007	Hardman Index	142	R	RAAA SAAA	EVAR	(d)	+	–	–	24h mortality, 30-day mortality
Rigberg et al.[85]	2006	USA	1995–1999	Charlson score	12 406	R	EAAA RAAA SAAA	OSR	(d)	+	–	–	In-hospital, 30-day mortality, 1-year mortality
Sahal et al.[86]	2008	Austria	1995–2006	n.a.	895	P	EAAA	both	(d)	+	–	–	Mortality
Sarac et al.[87]	2011	USA	1990–2008	APACHE-II	160	R	RAAA	both	(d)	+	–	–	Mortality
Saratzis et al.[88]	2013	UK/Greece	2008–2011	n.a.	338	P	EAAA	EVAR	(d)	+	–	–	Mid term mortality
Saratzis et al.[89]	2014	UK/Greece	2008–2011	n.a.	159	P, nested case–control	EAAA	EVAR	(d)	+	–	–	All cause mortality

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Schlosser et al.[90]	2010	USA/The Netherlands	1997–2000	n.a.	3457	R	EAAA	both	(d)	+	–	–	Mortality
Schlosser et al.[91]	2010	USA/The Netherlands	1997–2000	n.a.	1463	R	RAAA	both	(d)	+	–	–	Mortality
Statius et al.[92]	2007	UK/The Netherlands	1996–2005	SVS	5167	P	AAA	EVAR	(d)	+	–	–	30-day mortality
Tsilimparis et al.[93]	2012	Germany	4–years	n.a.	119	P	EAAA	EVAR	(d)	+	–	–	Survival
Von Meijenfeldt et al.[94]	2014	The Netherlands	2000–2013	n.a.	221	R	RAAA	both	(d)	+	–	–	In-hospital 30-day Mortality
Wanhainen et al.[95]	2008	Sweden	1994–2005	n.a.	10 691	P	EAAA RAAA SAAA	both	(d)	+	–	–	30-day Mortality
Khashram et al.[96]	2015	Australia	1990–2013	n.a.	1340	P	EAAA	both	(d)	+	–	–	Survival
Mureebe et al.[97]	2010	USA	1995–2006	n.a.	74 000	R	RAAA	both	(d)	+	–	–	30-day mortality
Walsh et al.[98]	2007	UK	1998–2004	POSSUM	159	P	EAAA	OSR	(d)	+	–	–	Mortality
Takahashi et al.[99]	2011	Japan	2004–2009	Index (rupture)	43	R, case–control	RAAA	both	(d)	–	–	–	Mortality
Reed et al.[100]	2013	UK	2006–2011	n.a.	119	R:	RAAA	both	(d)	–	–	–	Survival
Molnar et al.[101]	2014	Romania	2003–2011	n.a.	138	R:	EAAA RAAA	OSR plus	(d)	–	–	–	Mortality at discharge
Kennedy et al.[102]	2010	USA	2004–2008	n.a.	143	R:	EAAA SAAA	both	(d)	–	–	–	30-day mortality
Ho et al.[103]	2006	China	1993–2004	n.a.	63	P:	RAAA	OSR	(d)	–	–	–	Hospital mortality (<2d, >2d)
Bryce et al.[104]	2013	UK	2005–2007	n.a.	106	P:	EAAA	OSR	(d)	–	–	+	Mortality mace
Boult et al.[105]	2007	Australia	1999–2001	n.a.	961	P:	EAAA semi-urgengt	EVAR	(d)	–	–	–	Survival

Table S1 Excluded studies based on critical appraisal for type of prognostic study and analytical performance. (*continued*)

Reference	Year	Country	Follow-up	Model name if applicable	No. of patients	Source of data	Domain	Determinant	Model type*	Multivariable analysis	Final model	Predictive performance	Outcome
Appleton et al.[106]	2014	UK	NA	GAS	350	P:	EAAA urgent	OSR	(d)	–	–	–	30-day mortality
Ahn et al.[107]	2012	South-Korea	2007–2011	n.a.	30	R:	RAAA	both	(d)	–	–	–	Mortality
Grant et al.[108]	2008	New Zealand	1993–2005	n.a.	740	P:	RAAA	EVAR	(d)	+	–	–	30-day mortality
Khandanpour et al.[109]		UK	1996–2006	n.a.	1 105	P:	AAA	OSR	(d)	+/–	–	–	Mortality
Kordzadehet al.[110]	2015	UK	2007–2014	n.a.	80	R:	RAAA	OSR	(d)	+	–	–	30-day mortality
Lemaire et al.[111]	2008	USA	1990–2003	n.a.	9 169	P:	RAAA EAAA	EVAR	(d)	–	–	–	Mortality
Matyal et al.[112]	2015	USA	2003–2011	n.a.	3 738	P:	EAAA	both	(d)	–	–	–	30-day mortality

*(a), prediction model development studies with validation; (b), external model validation studies with or without model updating; (c), prediction model development studies without internal validation or with 'apparent' internal validation; (d), prediction studies exploring which predictors independently contribute to the prediction of a particular prognostic or diagnostic outcome. PRS, Preoperative Risk Score; SSS, Surgical Stress Score; CRS, Comprehensive Risk Score; E-PASS, Estimation of Physiologic Ability and Surgical Stress; P, Prospective; OSR, open surgical repair; n.a., not applicable; +, good; +/–, moderate; –, not good; R, Retrospective; GAS, Glasgow Aneurysm Score; POSSUM, Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity; P-POSSUM, Portsmouth POSSUM; RAAA-POSSUM, ruptured abdominal aortic aneurysm POSSUM; V-POSSUM, vascular POSSUM; SAPS-II, Simplified Acute Physiology Score II; APACHE-II, Acute Physiology and Chronic Health Evaluation II; EAAA, elective AAA; SAAA, symptomatic AAA; RAAA, Ruptured AAA; EVAR, endovascular aneurysm repair; PSI, Performance Status Index; AMC, Amsterdam Medical Centre; OLVG, Onze Lieve Vrouwe Gasthuis; M-LS, modified Leiden Score; SVS (M-CSS), Society for Vascular Surgery (modified Comorbidity Severity Score); m-CPI, modified Customized Probability Index; CPI, Customized Probability Index; MACE, major adverse cardiac events; ASA, American Society of Anesthesiologists; VQI, Vascular Quality Initiative; NSQIP, National Surgical Quality Improvement Program; CPET, cardiopulmonary exercise testing; AHRQ, Agency for Healthcare Research and Quality; IQI, Inpatient Quality Indicators.

REFERENCES APPENDIX TABLE S1

1 Tang T., Walsh S.R., Fanshawe T.R., Gillard J.H., Sadat U., Varty K., et al. Estimation of physiologic ability and surgical stress (E-PASS) as a predictor of immediate outcome after elective abdominal aortic aneurysm surgery. Am J Surg 2007 Aug;194(2):176-82.

2 Patterson B.O., Karthikesalingam A., Hinchliffe R.J., Loftus I.M., Thompson M.M., Holt P.J.E. The Glasgow Aneurysm Score does not predict mortality after open abdominal aortic aneurysm in the era of endovascular aneurysm repair. J Vasc Surg 2011;54(2):353-357.

3 Sandford R.M., Bown M.J., Sayers R.D. Scoring systems do not accurately predict outcome following abdominal aortic aneurysm repair. ANZ J Surg 2007;77(4):275-282.

4 Supsamutchai C, Wilasrusmee C, Lertsithichai P, Proprom N, Kittur DS. Comparison of risk-scoring systems in predicting hospital mortality after abdominal aortic aneurysm repair. Int J Angiol 2008;17(4):181-185.

5 Menezes F.H., Gomes De Souza V.M. Physiologic component of the estimation of physiologic ability and surgical stress scoring system as a predictor of immediate outcome after elective open abdominal aortic aneurysm repair. Ann Vasc Surg 2011;25(4):485-495.

6 Leo E., Biancari F., Nesi F., Pogany G., Bartolucci R., De Pasquale F., 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.

7 Mani K, Venermo M, Beiles B, Menyhei G, Altreuther M, Loftus I, et al. Regional Differences in Case Mix and Peri-operative Outcome After Elective Abdominal Aortic Aneurysm Repair in the Vascunet Database. Eur J Vasc Endovasc Surg 2015 Jun;49(6):646-652.

8 Kodama A, Narita H, Kobayashi M, Yamamoto K, Komori K. Usefulness of POSSUM physiological score for the estimation of morbidity and mortality risk after elective abdominal aortic aneurysm repair in Japan. Circ J 2011;75(3):550-556.

9 Kurc E., Sanioglu S., Ozgen A., Aka S.A., 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):150-155.

10 Hirzalla O., Emous M., Ubbink D.T., Legemate D. External validation of the Glasgow Aneurysm Score to predict outcome in elective open abdominal aortic aneurysm repair. J Vasc Surg 2006;44(4):712-716.

11 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) 2017 Jun;58(3):439-445.

12 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):123-127.

13 Faizer R., DeRose G., Lawlor D.K., Harris K.A., Forbes T.L. Objective scoring systems of medical risk: A clinical tool for selecting patients for open or endovascular abdominal aortic aneurysm repair. J Vasc Surg 2007;45(6):1102.

14 Bohm N., Wales L., Dunckley M., Morgan R., Loftus I., Thompson M. Objective Risk-scoring Systems for Repair of Abdominal Aortic Aneurysms: Applicability in Endovascular Repair? Eur J Vasc Endovasc Surg 2008;36(2):172-177.

15 Biancari F., Hobo R., Juvonen T. Glasgow Aneurysm Score predicts survival after endovascular stenting of abdominal aortic aneurysm in patients from the EUROSTAR registry. Br J Surg 2006;93(2):191-194.

16 Baas A.F., Janssen K.J.M., Prinssen M., Buskens E., Blankensteijn J.D. The Glasgow Aneurysm Score as a tool to predict 30-day and 2-year mortality in the patients from the Dutch Randomised Endovascular Aneurysm Management trial. J Vasc Surg 2008;47(2):277-281.
17 Archan S., Roscher C.R., Fairman R.M., Fleisher L.A. Revised Cardiac Risk Index (Lee) and Perioperative Cardiac Events as Predictors of Long-term Mortality in Patients Undergoing Endovascular Abdominal Aortic Aneurysm Repair. J Cardiothorac Vasc Anesth 2010;24(1):84-90.
18 Antonello M., Lepidi S., Kechagias A., Frigatti P., Tripepi A., Biancari F., et al. Glasgow Aneurysm Score Predicts the Outcome after Emergency Open Repair of Symptomatic, Unruptured Abdominal Aortic Aneurysms. Eur J Vasc Endovasc Surg 2007 Mar;33(3):272-6;.
19 Haji Mohd Yasin NA, Thomson IA, Leon de la Barra S. Endovascular aortic repair: can we predict who will not get long-term benefit? N Z Med J 2013;126(1369):44-52.
20 Hicks C.W., Black J.H., Arhuidese I., Asanova L., Qazi U., Perler B.A., et al. Mortality variability after endovascular versus open abdominal aortic aneurysm repair in a large tertiary vascular center using a Medicare-derived risk prediction model. J Vasc Surg 2015;61(2):291-297.
21 Sajid M.S., Tai N., Goli G., Platts A., Baker D.M., Hamilton G. Applicability of Glasgow aneurysm score and hardman index to elective endovascular abdominal aortic aneurysm repair. Asian J Surg 2007;30(2):113-117.
22 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 vascunet report. Eur J Vasc Endovasc Surg 2011;42(5):598-607.
23 Acosta S., Ogren M., Bergqvist D., Lindblad B., Dencker M., Zdanowski Z. The Hardman index in patients operated on for ruptured abdominal aortic aneurysm: a systematic review. J Vasc Surg 2006;44(5):949-954.
24 Conroy D.M., Altaf N., Goode S.D., Braithwaite B.D., MacSweeney S.T., Richards T. Use of the Hardman index in predicting mortality in endovascular repair of ruptured abdominal aortic aneurysms. Perspect Vasc Surg Endovasc Ther 2011;23(4):274-279.
25 Sharif M.A., Lee B., Makar R.R., Loan W., Chee V.S. Role of the Hardman index in predicting mortality for open and endovascular repair of ruptured abdominal aortic aneurysm. J Endovasc Ther 2007;14(4):528-535.
26 Anderson J.E., Chang D.C. Using electronic health records for surgical quality improvement in the era of big data. JAMA Surg 2015;150(1):24-29.
27 Bhutta H., Agha R., Wong J., Tang T.Y., Wilson Y.G., Walsh S.R. Neutrophil-lymphocyte ratio predicts medium-term survival following elective major vascular surgery: A cross-sectional study. Vasc Endovasc Surg 2011;45(3):227-231.
28 Carlisle J.B., 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(6):654-665.
29 Cho J.-S., Kim J.Y., Rhee R.Y., Gupta N., Marone L.K., Dillavou E.D., et al. Contemporary results of open repair of ruptured abdominal aortoiliac aneurysms: Effect of surgeon volume on mortality. J Vasc Surg 2008;48(1):10-18.
30 Giordano S., Biancari F., Loponen P., Wistbacka J.-O., Luther M. Preoperative haemodynamic parameters and the immediate outcome after open repair of ruptured abdominal aortic aneurysms. Interact Cardiovasc Thorac Surg 2009;9(3):491-493.

31 Gupta PK, Engelbert TL, Ramanan B, Fang X, Yamanouchi D, Hoch JR, et al. Postdischarge outcomes after endovascular abdominal aortic aneurysm repair. J Vasc Surg 2014;59(4):903-908.
32 Hawkins A.T., Smith A.D., Schaumeier M.J., De Vos M.S., Hevelone N.D., Nguyen L.L. The effect of surgeon specialization on outcomes after ruptured abdominal aortic aneurysm repair. J Vasc Surg 2014;60(3):590-596.
33 Ramanan B., Gupta P.K., Sundaram A., Gupta H., Johanning J.M., Lynch T.G., et al. Development of a risk index for prediction of mortality after open aortic aneurysm repair. J Vasc Surg 2013;58(4):871-878.
34 Davenport D.L., O'Keeffe S.D., Minion D.J., Sorial E.E., Endean E.D., Xenos E.S. Thirty-day NSQIP database outcomes of open versus endoluminal repair of ruptured abdominal aortic aneurysms. J Vasc Surg 2010;51(2):305.
35 O'Malley A.J., Cotterill P., Schermerhorn M.L., Landon B.E. Improving observational study estimates of treatment effects using joint modeling of selection effects and outcomes: The case of AAA repair. Med Care 2011;49(12):1126-1132.
36 Scarcello E., Ferrari M., Rossi G., Berchiolli R., Adami D., Romagnani F., et al. A New Preoperative Predictor of Outcome in Ruptured Abdominal Aortic Aneurysms: The Time Before Shock (TBS). Ann Vasc Surg 2010;24(3):315-320.
37 Sharif M.A., Arya N., Soong C.V., Lau L.L., O'Donnell M.E., Blair P.H., et al. Validity of the Hardman Index to Predict Outcome in Ruptured Abdominal Aortic Aneurysm. Ann Vasc Surg 2007;21(1):34-38.
38 Tambyraja A., Murie J., Chalmers R. Predictors of outcome after abdominal aortic aneurysm rupture: Edinburgh ruptured aneurysm score. World J Surg 2007;31(11):2243-2247.
39 Soong C.V., Makar R.R., O'Donnell M.E., Badger S.A., Lee B., Sharif M.A. Effect of preoperative renal dysfunction on mortality and postoperative renal failure following endovascular abdominal aortic aneurysm repair. Vasc Endovasc Surg 2008;42(5):427-432.
40 Treska V, Certik B, Cechura M, Novak M. Ruptured abdominal aortic aneurysms - university center experience. Interact Cardiovasc Thorac Surg 2006;5(6):721-723.
41 Haveman J.W., Zeebregts C.J., Verhoeven E.L.G., Van Den Berg A.P., Van Den Dungen J.J.A.M., Zwaveling J.H., et al. Changes in laboratory values and their relationship with time after rupture of an abdominal aortic aneurysm. Surg Today 2008;38(12):1091-1101.
42 Stone D.H., Goodney P.P., Kalish J., Schanzer A., Indes J., Walsh D.B., et al. Severity of chronic obstructive pulmonary disease is associated with adverse outcomes in patients undergoing elective abdominal aortic aneurysm repair. J Vasc Surg 2013;57(6):1531-1536.
43 IMPROVE trial investigators, Powell JT, Hinchliffe RJ, Thompson MM, Sweeting MJ, Ashleigh R, et al. Observations from the IMPROVE trial concerning the clinical care of patients with ruptured abdominal aortic aneurysm. Br J Surg 2014;101(3): 216-24; discussion 224.
44 Patel V.I., Lancaster R.T., Conrad M.F., Lamuraglia G.M., Kwolek C.J., Brewster D.C., et al. Comparable mortality with open repair of complex and infrarenal aortic aneurysm. J Vasc Surg 2011;54(4):952-959.
45 Mofidi R., Suttie S.A., Howd A., Dawson A.R.W., Griffiths G.D., Stonebridge P.A. Outcome from abdominal aortic aneurysms in Scotland, 1991-2006. Br J Surg 2008;95(12):1475-1479.
46 Mehta M., Byrne J., Darling III R.C., Paty P.S.K., Roddy S.P., Kreienberg P.B., et al. Endovascular repair of ruptured infrarenal abdominal aortic aneurysm is associated with lower

30-day mortality and better 5-year survival rates than open surgical repair. J Vasc Surg 2013;57(2):368-375.
47 Lomazzi C., Mariscalco G., Piffaretti G., Bacuzzi A., Tozzi M., Carrafiello G., et al. Endovascular treatment of elective abdominal aortic aneurysms: Independent predictors of early and late mortality. Ann Vasc Surg 2011;25(3):299-305.
48 LaPar DJ, Kron IL, Jones DR, Stukenborg GJ, Kozower BD. Hospital procedure volume should not be used as a measure of surgical quality. Ann Surg 2012;256(4):606-615.
49 Giles K.A., Wyers M.C., Pomposelli F.B., Hamdan A.D., Avery Ching Y., Schermerhorn M.L. The impact of body mass index on perioperative outcomes of open and endovascular abdominal aortic aneurysm repair from the National Surgical Quality Improvement Program, 2005-2007. J Vasc Surg 2010;52(6):1471-1477.
50 Gloviczki P., Huang Y., Oderich G.S., Duncan A.A., Kalra M., Fleming M.D., et al. Clinical presentation, comorbidities, and age but not female gender predict survival after endovascular repair of abdominal aortic aneurysm. J Vasc Surg 2015;61(4):853.
51 Grant S.W., Hickey G.L., Wisely N.A., Carlson E.D., Hartley R.A., Pichel A.C., et al. Cardiopulmonary exercise testing and survival after elective abdominal aortic aneurysm repair+. Br J Anaesth 2015;114(3):430-436.
52 Hartley R.A., Pichel A.C., Grant S.W., Hickey G.L., Lancaster P.S., Wisely N.A., et al. Preoperative cardiopulmonary exercise testing and risk of early mortality following abdominal aortic aneurysm repair. Br J Surg 2012;99(11):1539-1546.
53 Eckstein H.-H., Bruckner T., Heider P., Wolf O., Hanke M., Niedermeier H.-P., et al. The Relationship Between Volume and Outcome Following Elective Open Repair of Abdominal Aortic Aneurysms (AAA) in 131 German Hospitals. Eur J Vasc Endovasc Surg 2007;34(3):260-266.
54 Arya S, Kim SI, Duwayri Y, Brewster LP, Veeraswamy R, Salam A, et al. Frailty increases the risk of 30-day mortality, morbidity, and failure to rescue after elective abdominal aortic aneurysm repair independent of age and comorbidities. J Vasc Surg 2015;61(2):324-331.
55 Azizzadeh A., Sanchez L.A., Miller III C.C., Marine L., Rubin B.G., Safi H.J., et al. Glomerular filtration rate is a predictor of mortality after endovascular abdominal aortic aneurysm repair. J Vasc Surg 2006;43(1):14-18.
56 Beck A.W., Goodney P.P., Nolan B.W., Likosky D.S., Eldrup-Jorgensen J., Cronenwett J.L. Predicting 1-year mortality after elective abdominal aortic aneurysm repair. J Vasc Surg 2009;49(4):838-844.
57 Bonardelli S., Cervi E., Maffeis R., Nodari F., De Lucia M., Guadrini C., et al. Open surgery in endovascular aneurysm repair era: simplified classification in two risk groups owing to factors affecting mortality in 137 ruptured abdominal aortic aneurysms (RAAAs). Updates Surg 2011;63(1):39-44.
58 Antonello M, Frigatti P, Maturi C, Lepidi S, Noventa F, Pittoni G, et al. Open repair for ruptured abdominal aortic aneurysm: is it possible to predict survival? Ann Vasc Surg 2009;23(2):159-166.
59 Robinson WP, Huang W, Rosen A, Schanzer A, Fang H, Anderson FA, et al. The Agency for Healthcare Research and Quality Inpatient Quality Indicator #11 overall mortality rate does not accurately assess mortality risk after abdominal aortic aneurysm repair. J Vasc Surg 2015;61(1):44-49.

60 Acosta S., Lindblad B., Zdanowski Z. Predictors for Outcome after Open and Endovascular Repair of Ruptured Abdominal Aortic Aneurysms. Eur J Vasc Endovasc Surg 2007;33(3):277-284.
61 Chagpar RB, Harris JR, Lawlor DK, DeRose G, Forbes TL. Early mortality following endovascular versus open repair of ruptured abdominal aortic aneurysms. Vasc Endovascular Surg 2010;44(8):645-649.
62 Chen CK, Chang HT, Chen YC, Chen TJ, Chen IM, Shih CC. Surgeon elective abdominal aortic aneurysm repair volume and outcomes of ruptured abdominal aortic aneurysm repair: a 12-year nationwide study. World J Surg 2013;37(10):2360-2371.
63 de Bruin JL, Baas AF, Heymans MW, Buimer MG, Prinssen M, Grobbee DE, et al. Statin therapy is associated with improved survival after endovascular and open aneurysm repair. J Vasc Surg 2014;59(1):39-44.e1.
64 De Virgilio C., Tran J., Lewis R., Donayre C., Dauphine C., White R., et al. Factors affecting long-term mortality after endovascular repair of abdominal aortic aneurysms. Arch Surg 2006;141(9):905-909.
65 Diehm N., Benenati J.F., Becker G.J., Quesada R., Tsoukas A.I., Katzen B.T., et al. Anemia is associated with abdominal aortic aneurysm (AAA) size and decreased long-term survival after endovascular AAA repair. J Vasc Surg 2007;46(4):676-681.
66 Feeney JM, Burns K, Staff I, Bai J, Rodrigues N, Fortier J, et al. Prehospital HMG Co-A reductase inhibitor use and reduced mortality in ruptured abdominal aortic aneurysm. J Am Coll Surg 2009;209(1):41-46.
67 Hashimoto M., Ito T., Kurimoto Y., Harada R., Kawaharada N., Higami T. Preoperative arterial blood lactate levels as a predictor of hospital mortality in patients with a ruptured abdominal aortic aneurysm. Surg Today 2013;43(2):136-140.
68 Hughes K, Jackson JD, Prendergast TI, Rose DA, Bolorunduro O, Obirieze A, et al. Diabetes mellitus is not associated with major morbidity following open abdominal aortic aneurysm repair. J Surg Res 2013;184(2):751-754.
69 Hultgren R, Granath F, Swedenborg J. Different disease profiles for women and men with abdominal aortic aneurysms. Eur J Vasc Endovasc Surg 2007;33(5):556-560.
70 Karkos C.D., Karamanos D., Papazoglou K.O., Kantas A.S., Theochari E.G., Kamparoudis A.G., et al. Usefulness of the Hardman index in predicting outcome after endovascular repair of ruptured abdominal aortic aneurysms. J Vasc Surg 2008;48(4):788-794.
71 Kim SD, Hwang JK, Park SC, Kim JI, Moon IS, Park JS, et al. Predictors of postoperative mortality of ruptured abdominal aortic aneurysm: a retrospective clinical study. Yonsei Med J 2012;53(4):772-780.
72 Koga Y, Mishima Y, Hara M, Hiraki T, Ushijima K. Prognostic factors associated with mortality in patients undergoing emergency surgery for abdominal aortic aneurysms. J Anesth 2011;25(5):666-671.
73 Kunishige H, Ishibashi Y, Kawasaki M, Morimoto K, Inoue N. Risk factors affecting survival after surgical repair of ruptured abdominal aortic aneurysm. Ann Vasc Dis 2013;6(3):631-636.
74 Law Y, Chan YC, Cheung GC, Ting AC, Cheng SW. Outcome and risk factor analysis of patients who underwent open infrarenal aortic aneurysm repair. Asian J Surg 2016;39(3):164-171.

75 Lee JS, He K, Harbaugh CM, Schaubel DE, Sonnenday CJ, Wang SC, et al. Frailty, core muscle size, and mortality in patients undergoing open abdominal aortic aneurysm repair. J Vasc Surg 2011;53(4):912-917.
76 Lo RC, Bensley RP, Hamdan AD, Wyers M, Adams JE, Schermerhorn ML, et al. Gender differences in abdominal aortic aneurysm presentation, repair, and mortality in the Vascular Study Group of New England. J Vasc Surg 2013;57(5):1261-8, 1268.e1-5.
77 Matsumura JS, Katzen BT, Sullivan TM, Dake MD, Naftel DC, Excluder Bifurcated Endoprosthesis Investigators. Predictors of survival following open and endovascular repair of abdominal aortic aneurysms. Ann Vasc Surg 2009;23(2):153-158.
78 Montan C, Wannberg M, Holst J, Wahlgren CM. Perioperative haemorrhage in endovascular abdominal aneurysm repair affects outcome. Eur J Vasc Endovasc Surg 2013;46(1):87-92.
79 Nakayama A., Morita H., Miyata T., Hoshina K., Nagayama M., Takanashi S., et al. Predictors of mortality after emergency or elective repair of abdominal aortic aneurysm in a Japanese population. Heart Vessels 2014;29(1):65-70.
80 Ohrlander T, Dencker M, Acosta S. Preoperative spirometry results as a determinant for long-term mortality after EVAR for AAA. Eur J Vasc Endovasc Surg 2012;43(1):43-47.
81 Ohrlander T, Dencker M, Acosta S. Morphological state as a predictor for reintervention and mortality after EVAR for AAA. Cardiovasc Intervent Radiol 2012;35(5):1009-1015.
82 Ohrlander T., Dencker M., Dias N.V., Gottsater A., Acosta S. Cardiovascular predictors for long-term mortality after EVAR for AAA. Vasc Med 2011;16(6):422-427.
83 Oyague K.S., Mubarak O.A., Nowak L.R., Gainer J.G., Rehring T.F., O'Brien M.M., et al. Endovascular repair of ruptured and symptomatic abdominal aortic aneurysms using a structured protocol in a community teaching hospital. Ann Vasc Surg 2015;29(1):76-83.
84 Richards T, Goode SD, Hinchliffe R, Altaf N, Macsweeney S, Braithwaite B. The importance of anatomical suitability and fitness for the outcome of endovascular repair of ruptured abdominal aortic aneurysm. Eur J Vasc Endovasc Surg 2009;38(3):285-290.
85 Rigberg D.A., Zingmond D.S., McGory M.L., Maggard M.A., Agustin M., Lawrence P.F., et al. Age stratified, perioperative, and one-year mortality after abdominal aortic aneurysm repair: A statewide experience. J Vasc Surg 2006;43(2):224-229.
86 Sahal M, Prusa AM, Wibmer A, Wolff KS, Lammer J, Polterauer P, et al. Elective abdominal aortic aneurysm repair: does the aneurysm diameter influence long-term survival? Eur J Vasc Endovasc Surg 2008;35(3):288-294.
87 Sarac T.P., Bannazadeh M., Rowan A.F., Bena J., Srivastava S., Eagleton M., et al. Comparative predictors of mortality for endovascular and open repair of ruptured infrarenal abdominal aortic aneurysms. Ann Vasc Surg 2011;25(4):461-468.
88 Saratzis A, Sarafidis P, Melas N, Saratzis N, Kitas G. Impaired renal function is associated with mortality and morbidity after endovascular abdominal aortic aneurysm repair. J Vasc Surg 2013;58(4):879-885.
89 Saratzis A., Saedon M., Melas N., Kitas G.D., Mahmood A. Obesity as an independent predictor of outcome after endovascular abdominal aortic aneurysm repair. Ann Vasc Surg 2014;28(4):816-822.
90 Schlosser FJ, Vaartjes I, van der Heijden GJ, Moll FL, Verhagen HJ, Muhs BE, et al. Mortality after elective abdominal aortic aneurysm repair. Ann Surg 2010;251(1):158-164.

91 Schlosser F.J.V., Vaartjes I., Van Der Heijden G.J.M.G., Moll F.L., Verhagen H.J.M., Muhs B.E., et al. Mortality after hospital admission for ruptured abdominal aortic aneurysm. Ann Vasc Surg 2010;24(8):1125-1132.

92 van Eps RG, Leurs LJ, Hobo R, Harris PL, Buth J, EUROSTAR Collaborators. Impact of renal dysfunction on operative mortality following endovascular abdominal aortic aneurysm surgery. Br J Surg 2007;94(2):174-178.

93 Tsilimparis N, Mitakidou D, Hanack U, Deussing A, Yousefi S, Ruckert RI. Effect of preoperative aneurysm diameter on long-term survival after endovascular aortic aneurysm repair. Vasc Endovascular Surg 2012;46(7):530-535.

94 Von Meijenfeldt G.C.I., Ultee K.H.J., Eefting D., Hoeks S.E., Ten Raa S., Rouwet E.V., et al. Differences in mortality, risk factors, and complications after open and endovascular repair of ruptured abdominal aortic aneurysms. Eur J Vasc Endovasc Surg 2014;47(5):479-486.

95 Wanhainen A, Bylund N, Bjorck M. Outcome after abdominal aortic aneurysm repair in Sweden 1994-2005. Br J Surg 2008;95(5):564-570.

96 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(2):115-125.

97 Mureebe L, Egorova N, McKinsey JF, Kent KC. Gender trends in the repair of ruptured abdominal aortic aneurysms and outcomes. J Vasc Surg 2010;51(4 Suppl):9S-13S.

98 Walsh SR, Tang T, Sadat U, Varty K, Boyle JR, Gaunt ME. Preoperative glomerular filtration rate and outcome following open abdominal aortic aneurysm repair. Vasc Endovascular Surg 2007;41(3):225-229.

99 Takahashi K, Fukui D, Wada Y, Terasaki T, Ohtsu Y, Komatsu K, et al. Indicators of survival after open repair of ruptured abdominal aortic aneurysms and an index for predicting aneurysmal rupture potential. Ann Vasc Dis 2011;4(3):209-217.

100 Reed M.J., Burfield L.C. Initial emergency department coagulation profile does not predict survival in ruptured abdominal aortic aneurysm. Eur J Emerg Med 2013;20(6):397-401.

101 Molnar A, Sacui D, Scridon T. Risk factors influencing the surgical outcome in 138 consecutive patients with infrarenal aortic aneurysm: experience at the Cluj-Napoca Cardiovascular Surgery Center. Chirurgia (Bucur) 2014;109(2):223-228.

102 Kennedy N.A., Flynn L.M., Berg R.M., Lorelli D.R., Rama K., Rizk Y. The evaluation of morbidity and mortality in abdominal aortic aneurysm repair patients as related to body mass index. Am J Surg 2010;199(3):369-371.

103 Ho P, Cheng SW, Ting AC, Poon JT. Improvement of mortality of ruptured abdominal aortic aneurysm patients over 12 years and its relationship with tracheostomy. Ann Vasc Surg 2006; 20(2):175-182.

104 Bryce G.J., Payne C.J., Gibson S.C., Byrne D.S., Delles C., McClure J., et al. B-type natriuretic peptide predicts postoperative cardiac events and mortality after elective open abdominal aortic aneurysm repair. J Vasc Surg 2013;57(2):345-353.

105 Boult M, Maddern G, Barnes M, Fitridge R. Factors affecting survival after endovascular aneurysm repair: results from a population based audit. Eur J Vasc Endovasc Surg 2007;34(2):156-162.

106 Appleton ND, Bailey DM, Morris-Stiff G, Lewis MH. Neutrophil to lymphocyte ratio predicts perioperative mortality following open elective repair of abdominal aortic aneurysms. Vasc Endovascular Surg 2014;48(4):311-316.

107 Ahn HY, Chung SW, Lee CW, Kim MS, Kim S, Kim CW. Factors affecting the postoperative mortality in the ruptured abdominal aortic aneurysm. Korean J Thorac Cardiovasc Surg 2012;45(4):230-235.
108 Grant M.W., Thomson I.A., Van Rij A.M. In-hospital mortality of ruptured abdominal aortic aneurysm. ANZ J Surg 2008;78(8):698-704.
109 Khandanpour N, Armon MP, Foxall R, Meyer FJ. The effects of increasing obesity on outcomes of vascular surgery. Ann Vasc Surg 2009;23(3):310-316.
110 Kordzadeh A, Malietzis G, Browne T, Prionidis I, Panayiotopoulos YP. Neutrophil to lymphocyte ratio (NLR) of five predicts 30-day morbidity in ruptured abdominal aortic aneurysms (rAAA): a retrospective cohort study. Int J Surg 2015;15:45-48.
111 Lemaire A, Cook C, Tackett S, Mendes DM, Shortell CK. The impact of race and insurance type on the outcome of endovascular abdominal aortic aneurysm (AAA) repair. J Vasc Surg 2008;47(6):1172-1180.
112 Matyal R, Shakil O, Hess PE, Lo R, Jainandunsing JS, Mahmood B, et al. Impact of gender and body surface area on outcome after abdominal aortic aneurysm repair. Am J Surg 2015;209(2):315-323.

Table S2 Formulas and specifications by model

Model	Specifications	Formula
(New) Simplified Acute Physiology Score (SAPS-II)[1,2]	SAPS II = age + HF + systolic BP + temperature + ventilated or CPAP pulmonary artery pressure (mmHg) + ventilated or CPAP pulmonary artery pressure (kPa) + urinary output + serum urea (mmol/l) + serum urea nitrogen (mg/dl) + WBC (10^3 mm^3) + potassium (mmol/day) + sodium (mmol/l) + bicarbonate (mEq/l) + bilirubin (mol/l, mg/dl) + GCS + chronic diseases + type of admission	–7.7631 + 0.0737 (SAPSII) + 0.9971 (ln(SAPSII + 1))
Acute Physiology and Chronic Health Evaluation Score- II (APACHE-II)[2,3]	APACHE-II= APS + age (≤ 44 years = 0; 45–54 years= 2; 55–64 years = 3; 65–74 years = 5; ≥ 75 years = 6) + chronic health points (if history of severe organ system insufficiency or immunocompetent: 5 for emergency or 2 for elective patients) APS = temperature + MAP + HF + respiratory rate + ventilatory pulmonary artery pressure (mmHg) + ventilatory pulmonary artery pressure (kPa) + arterial pH + sodium + potassium + creatinine + haematocrit + WBC + GCS	–3517 + (APACHE II × 0.146) + (0.603 only if postemergency surgery) + (diagnostic category weight)
Society for Vascular Surgery Score (SVS) modified Co-morbidity Severity Score (m-CSS)[4,5]		4 × cardiac status + 2 × pulmonary status + 2 × renal status + 1 × hypertension +1 × age
POSSUM models[6,7]	Physiological score (p) = age + cardiac co-morbidity + pulmonary co-morbidity + systolic BP + HF + GCS + haemoglobin + WBC + creatinine + sodium + potassium + ECG Operative score = severity of procedure + no. of procedures + blood loss + peritoneal soiling + malignancy + urgency setting	POSSUM: − 9·065 (0·1692 × physiological score) (0·1550 × operative score) V-POSSUM: –8.0616 + (0.1552 × physiological score) + (0.1238 × operative score) V(p)-POSSUM: –6.0386 + (0.1539 × physiological score) RAAA-POSSUM: –4.9795 + (0.0913 × physiological score) + (0.0958 × operative score) RAAA(p)-POSSUM: –2.7569 + (0.0968 × physiological score) Cambridge POSSUM: –7.5365 + (0.1632 × physiological score) + (0.0518 × operative score)

Table S2 Formulas and specifications by model (*continued*)

Model	Specifications	Formula
Glasgow Aneurysm Score (GAS)[8–10]		GAS: age + (17)shock + (7)myocardial disease + (10)CVD + (14)renal disease Modified GAS: age + (7)myocardial disease + (10)CVD + (14)renal disease Updated GAS: age + (17)shock + (7) myocardial disease + (7)CVD + (14)renal disease + (7)OSR
Leiden score (and modified)[11,12]		Original Leiden: centre-specific score + age + cardiac score + respiratory impairment + renal impairment + female Modified Leiden: age + cardiac score (without ECG) + respiratory impairment + renal impairment + female
Vancouver score[13,14]		–3.44 + (1.14 × unconsciousness) + (–1.14 × conscious) + (0.062xage) + (0.60xcardiac arrest) + (–0.60 × no cardiac arrest)
Hardman index[15,16]		Age > 76 years, serum creatinine level >190 mol/l, haemoglobin level < 9g/dl (or 2.1 mg/dl), myocardial ischaemia on ECG and a history of loss of consciousness after hospital arrival
Estimation of Physiologic Ability and Surgical Stress (E-PASS)[17,18]	PRS = – 0.0686 + (0.00345 × age) + (0.323 × severe heart disease) + (0.205 × severe pulmonary disease) + (0.153 × diabetes) + (0.148 × performance status index (0–4)) + (0.0666 × ASA (I–V)) SSS = – 0.342 + (0.0139 × blood loss (g)/bodyweight(kg)) + (0.0392 × operating time (h)) + (0.352 × minor/moderate/major skin incision)	Comprehensive risk score = –0.328 + (0.936 × PRS) + (0.976 × SSS)
(Modified) Lee Index[19, 20]		Lee: high-risk surgery + IHD + CHF + stroke + hypertension + renal failure + respiratory disease + beta-blockers + statin Modified Lee: high-risk surgery + IHD + CHF + stroke + use of insulin + creatinine > 2 mg/dl 1 point for every parameter when present. In the case of MACE perioperatively add 1 point preoperatively
(Modified)-Customized Probability Index (CPI)[21,22]		CPI: (13)CAD + (14)CHF + (10)CVD + (7)BP + (7)COPD + (16)RF + (-15) beta-blocker + (-10)statin Modified CPI: (13)CAD + (14)CHF + (7)BP + (7)COPD + (16)RF + (-15)beta-blocker + (-10) statin

Table S2 Formulas and specifications by model (*continued*)

Model	Specifications	Formula
Vascular Biochemistry and Haematology Outcome Models (VBHOM)[23]		−2·257 + (0·1511 × male) + (0·9940 × mode of admission) + (0·05923 × age) + (0·001401 × serum urea (continuous mmol/l)) − (0·01303 × sodium (continuous mmol/l)) − (0·03585 × potassium (continuous mmol/l)) − (0·2278 × haemoglobin (continuous g/dl)) + (0·02059 × white cell count (continuous × 10^9/l))
Edinburgh Ruptured Aneurysm Score (ERAS)[24,25]		Haemoglobin level 9 g/dl, GCS 15, BP 90 mmHg. Score 0–3. These bands of risk correspond to a predicted mortality of no risk, 30% , 50%, 80%.
Endovascular aneurysm repair Risk Assessment model (ERA)[26]		Author e-mailed
Giles score/ Medicare[27]		−5·02 + (0·42 × female sex) + (0·15 × age (70–75 years)) + (0·63 × age (75–80 years)) + (1·14 × age > 80 years) + (0·71 × chronic renal insufficiency) + (0·95 × end-stage renal disease) + (0·55 × congestive heart failure) + (0·30 × vascular disease) + (1·17 × OSR)
Mount Sinai score[28]		Renal failure dialysis + renal failure without dialysis + clinical sign. Lower extremity ischaemia + liver disease + CHF + neurological disorders + chronic pulmonary diseases + age ≥ 85 + 80–84 + 75–79 years + female sex + hospital volume (<7 EVARs/ year) + surgeon experience (<3 EVARs/year)
Cleveland Clinic experience[29]		Age (per 5-year increase from 50 years) + aortic diameter (per 5 mm increase from 30 mm) + previous chronic heart failure + COPD + use of oxygen + use of aspirin
Leicester score[21]		AAA OSR/EVAR: –12.093 + (age × 0.080) + (MI within 10 years × 1.339) + (creatinine × 0.005) + (open surgery × 2.370) AAA OSR-only model: –9.848 + (age × 0.095) + (creatinine × 0.007).
Vascular Governance North West model (VGNW)[30]		−9·3431 + (0.0486 × age) + (0·7322 × female sex) + (0·6620 × diabetes) + (0·0073 × creatinine) + (0·4718 × respiratory disease) + (0·7762 × antiplatelet medication) + (1·3130 × open surgery)
Vascular Study Group of New England model (VSGNE)[16]		Age >76 years + cardiac arrest (+2) + loss of consciousness (+1) + suprarenal clamp (+1) VSGNE RAAA risk score 0–6: >4 80% mortality risk, ≥5 87% mortality risk

Table S2 Formulas and specifications by model (*continued*)

Model	Specifications	Formula
British Aneurysm Repair (BAR)[31]		–10.9187 + (OSR ×1.6466) + (0·0568 × age (continuous in years)) + (0·7062 × female sex) + (0·5979 × creatinine > 120 mmol/l) + (0·3422 × cardiac disease) + (0·3033 × abnormal ECG) + (0·8812 × previous aortic surgery or stent) + (0·3697 × abnormal WBC) + (0·3099 × abnormal sodium) + (0·1285 × AAA diameter (continuous in cm)) + (0·2292 × ASA grade II) + (0·7334 × ASA grade III) + (1·6775 × ASA grade IV)
Abdominal Aortic Aneurysm Statistically Corrected Operative Risk Evaluation Score (AAA SCORE)[32]		Preoperative AAA SCORE= –7.1026 + (reoperation ×1.7691) + (admission mode unplanned × 0.6877) + (admission mode emergency × 1.4731) + (age (years) × 0.0493) + (creatinine (mol/l) × 0.0035) + (lowest preoperative BP linear term × –0.0307) + (lowest preoperative BP quadratic term 1.01 × 10^{-4}) + (cardiac history × 0.4649) + (EVAR × –0.9526) + (ASA grade 0.5102) + (WBC 10^{9} × 0.0279) Perioperative AAA SCORE= –7.4339 + (reoperation × 1.6319) + (admission mode unplanned × 0.6284) + (admission mode emergency ×1.3532) + (age (years) ×0.0467) + (creatinine (mol/l) × 0.0036) + (lowest intraoperative BP linear term ×–0.0197) + (cardiac history × 0.4601) + (EVAR × –0.9526) + (ASA grade 0.3636) + (albumin × –0.0286)
Artificial neural network (ANN)[33]		(Age × 0.27) + (loss of consciousness × 0.36) + (shock × 0.27) + (cardiopulmonary rescucitation/cardiac arrest × 0.32)

REFERENCES APPENDIX TABLE S2

1 Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA 1993;270(24):2957-2963.

2 Sandford R.M., Bown M.J., Sayers R.D. Scoring systems do not accurately predict outcome following abdominal aortic aneurysm repair. ANZ J Surg 2007;77(4):275-282.

3 Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13(10):818-829.

4 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(5):1061-1066.

5 Faizer R., DeRose G., Lawlor D.K., Harris K.A., Forbes T.L. Objective scoring systems of medical risk: A clinical tool for selecting patients for open or endovascular abdominal aortic aneurysm repair. J Vasc Surg 2007;45(6):1102.

6 Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg 1991;78(3):355-360.

7 Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. Br J Surg 2001;88(7):958-963.

8 Samy A.K., Murray G., MacBain G. Prospective evaluation of the Glasgow Aneurysm Score. J R COLL SURG EDINBURGH 1996;41(2):105-107.

9 Samy AK, Murray G, MacBain G. Glasgow aneurysm score. Cardiovasc Surg 1994;2(1):41-44.

10 Visser J.J., Williams M., Kievit J., Bosch J.L. Prediction of 30-day mortality after endovascular repair or open surgery in patients with ruptured abdominal aortic aneurysms. J Vasc Surg 2009;49(5):1093-1099.

11 Steyerberg E.W., Kievit J., De Mol Van Otterloo J.C.A., Van Bockel J.H., Eijkemans M.J.C., Habbema J.D.F. Perioperative mortality of elective abdominal aortic aneurysm surgery: A clinical prediction rule based on literature and individual patient data. Arch Intern Med 1995;155(18):1998-2004.

12 Nevala T., Biancari F., Perala J., Manninen H., Aho P.-S., Matsi P., et al. Risk prediction in patients undergoing elective endovascular repair of an abdominal aortic aneurysm. Scand Cardiovasc J 2010;44(2):125-128.

13 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; discussion 621-3.

14 Leo E., Biancari F., Nesi F., Pogany G., Bartolucci R., De Pasquale F., 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.

15 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):123-129.

16 Robinson W.P., Schanzer A., Li Y., Goodney P.P., Nolan B.W., Eslami M.H., 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):354-361.

17 Tang T., Walsh S.R., Fanshawe T.R., Gillard J.H., Sadat U., Varty K., et al. Estimation of physiologic ability and surgical stress (E-PASS) as a predictor of immediate outcome after elective abdominal aortic aneurysm surgery. Am J Surg 2007 Aug;194(2):176-82.
18 Haga Y, Ikei S, Ogawa M. Estimation of Physiologic Ability and Surgical Stress (E-PASS) as a new prediction scoring system for postoperative morbidity and mortality following elective gastrointestinal surgery. Surg Today 1999;29(3):219-225.
19 Lee TH, Marcantonio ER, Mangione CM, Thomas EJ, Polanczyk CA, Cook EF, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation 1999;100(10):1043-1049.
20 Archan S., Roscher C.R., Fairman R.M., Fleisher L.A. Revised Cardiac Risk Index (Lee) and Perioperative Cardiac Events as Predictors of Long-term Mortality in Patients Undergoing Endovascular Abdominal Aortic Aneurysm Repair. J Cardiothorac Vasc Anesth 2010;24(1):84-90.
21 Choke E., Lee K., McCarthy M., Nasim A., Naylor A.R., Bown M., et al. Risk models for mortality following elective open and endovascular abdominal aortic aneurysm repair: A single institution experience. Eur J Vasc Endovasc Surg 2012;44(6):549-554.
22 Kertai MD, Boersma E, Klein J, van Sambeek M, Schouten O, van Urk H, et al. Optimizing the prediction of perioperative mortality in vascular surgery by using a customized probability model. Arch Intern Med 2005;165(8):898-904.
23 Tang T, Walsh SR, Prytherch DR, Lees T, Varty K, Boyle JR, et al. VBHOM, a data economic model for predicting the outcome after open abdominal aortic aneurysm surgery. Br J Surg 2007;94(6):717-721.
24 Tambyraja A., Murie J., Chalmers R. Predictors of outcome after abdominal aortic aneurysm rupture: Edinburgh ruptured aneurysm score. World J Surg 2007;31(11):2243-2247.
25 Tambyraja A.L., Lee A.J., Murie J.A., Chalmers R.T.A. Prognostic scoring in ruptured abdominal aortic aneurysm: A prospective evaluation. J Vasc Surg 2008;47(2):282-286.
26 Barnes M, Boult M, Maddern G, Fitridge R. A model to predict outcomes for endovascular aneurysm repair using preoperative variables. Eur J Vasc Endovasc Surg 2008;35(5):571-579.
27 Giles KA, Schermerhorn ML, O'Malley AJ, Cotterill P, Jhaveri A, Pomposelli FB, et al. Risk prediction for perioperative mortality of endovascular vs open repair of abdominal aortic aneurysms using the Medicare population. J Vasc Surg 2009;50(2):256-262.
28 Egorova N, Giacovelli JK, Gelijns A, Greco G, Moskowitz A, McKinsey J, et al. Defining high-risk patients for endovascular aneurysm repair. J Vasc Surg 2009;50(6):1271-9.e1.
29 Mastracci T.M., Greenberg R.K., Hernandez A.V., Morales C. Defining high risk in endovascular aneurysm repair. J Vasc Surg 2010;51(5):1088-1095.
30 Grant S.W., Grayson A.D., Purkayastha D., Wilson S.D., McCollum C. Logistic risk model for mortality following elective abdominal aortic aneurysm repair. Br J Surg 2011;98(5):652-658.
31 Grant S.W., Hickey G.L., Grayson A.D., Mitchell D.C., McCollum C.N. National risk prediction model for elective abdominal aortic aneurysm repair. Br J Surg 2013;100(5):645-653.
32 Ambler G.K., Gohel M.S., Mitchell D.C., Loftus I.M., Boyle J.R. The abdominal aortic aneurysm statistically corrected operative risk evaluation (AAA SCORE) for predicting mortality after open and endovascular interventions. J Vasc Surg 2015 Jan;61(1):35-43.;61(1):35-43.

33 Wise ES, Hocking KM, Brophy CM. Prediction of in-hospital mortality after ruptured abdominal aortic aneurysm repair using an artificial neural network. J Vasc Surg 2015;62(1):8-15.

APPENDIX CHAPTER 5

Appendix A.

The choice of treatment is based on the interpretation of anatomical and patient related factors. Suppose: certain patients will always get EVAR or OSR in every hospital. There is a ''marginal'' group were there is a choice, depending on hospital preference: all EVAR or all OSR.

Roughly, patients can be divided in 3 groups:

- N1: patients who will have OSR in every hospital: mortality A
- N2: patients with discussion (''marginal population'').
  - o In case of OSR: mortality B
  - o in case of EVAR: mortality C
- N3: patients who will have EVAR in every hospital: mortality D

Suppose, All N2 patients get OSR in low%EVAR hospitals and EVAR in high%EVAR hospitals. Then we know from the result in the outcome data and table 3 in the manuscript:

- Low%EVAR hospitals:
  - o N1+N2 will get OSR: Mortality (A*N1+B*N2)/(N1+N2)=4.0%
  - o N3 will get EVAR: Mortality is D=0.7%
- High%EVAR hospitals:
  - o N1 will get OSR: Mortality is A=7.3%
  - o N2+N3 will get EVAR: Mortality is (C*N2+D*N3)/(N2=N3)=0.9%

Further, we know from the DSAA

- The proportion of EVAR in low%EVAR hospitals is N3/(N1+N2+N3)=0.69
- The proportion of EVAR in high%EVAR hospitals is (N2+N3)/(N1+N2+N3)=0.86

Suppose, N1+N2+N3=100, we find:

- N3=69
- N2=86-69=17
- N1=100-69-17=14

We can calculate:

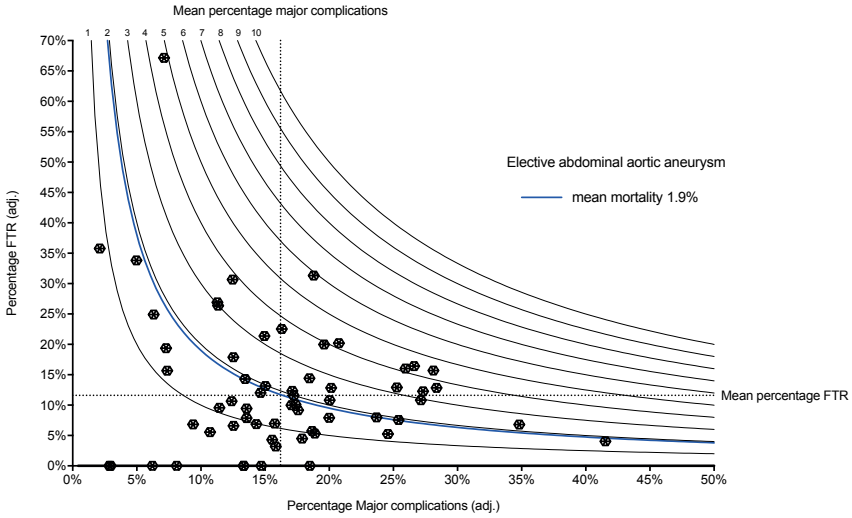
- A=7.3%
- B=((4.0*N1+N2) – 7.3*N1)/N2=1.3%
- D=0.7%
- C=((0.9*N2+N3) – 0.7*N3)/N2=1.7%

According to this calculation the mortality of OSR in N2 patients is 1.3% and in case of EVAR in N2 the mortality is 1.7%. Patients in the marginal group are better off with OSR.

APPENDIX CHAPTER 6

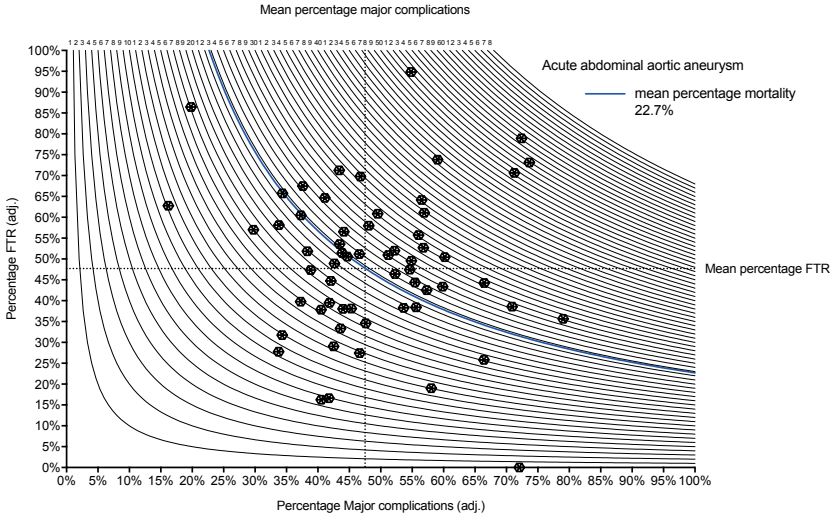
Supplement A.

For the year 2015 data verification was performed on a random selection of 304 patients who had AAA surgery in 14 Dutch hospitals. Five patients (1.7%) were not registered in the DSAA. Of those 5 patient, 3 had minor complications and there was no mortality. In the 298 registered patients no mortality was missed, however 21 complications (7.1%) were missed, all minor: fever and/or intestinal obstruction. No reinterventions were missed. Four readmissions were not registered (1.3%).



Supplemental material B:

The adjusted percentages of FTR and major complications after EAAA surgery. The hyperbolic lines project the mortality in relation to the percentage major complications and the mortality after major complications. From left to right each line has a 1% higher mortality for every hospital crossing that line starting at 1%, with the overall mean percentage mortality of 1.9% projected with the blue line.



Supplemental material C:

The adjusted percentages of FTR and major complications after AAAA surgery. The hyperbolic lines project the mortality in relation to the percentage major complications and the mortality after major complications. From left to right each line has a 1% higher mortality for every hospital crossing that line starting at 1%, with the overall mean percentage mortality of 22.7% projected with the blue line. There is one hospital with a mean percentage FTR of 116.8% that is not shown in this figure.