Risk stratification and outcome assessment in cardiac surgery and transcatheter interventions

Menno van Gameren

ISBN: 978-90-8559-114-6

Cover design: M. van Gameren
Lay-out: M. van Gameren
Printed by Optima Grafische Communicatie, Rotterdam, The Netherlands

© 2010 M. van Gameren
All rights reserved. No part of this thesis may be reproduced, distributed, stored in a retrieval system or transmitted in any form or by any means, without permission of the author, or when appropriate, of the publishers of the publications.

Risk stratification and outcome assessment in cardiac surgery and transcatheter interventions

Risicostratificatie en evaluatie van resultaten van hartchirurgie en catheter-interventies

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.G. Schmidt

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
Woensdag 10 november 2010 om 9:30 uur

door

Menno van Gameren

geboren te Rotterdam

Promotiecommissie

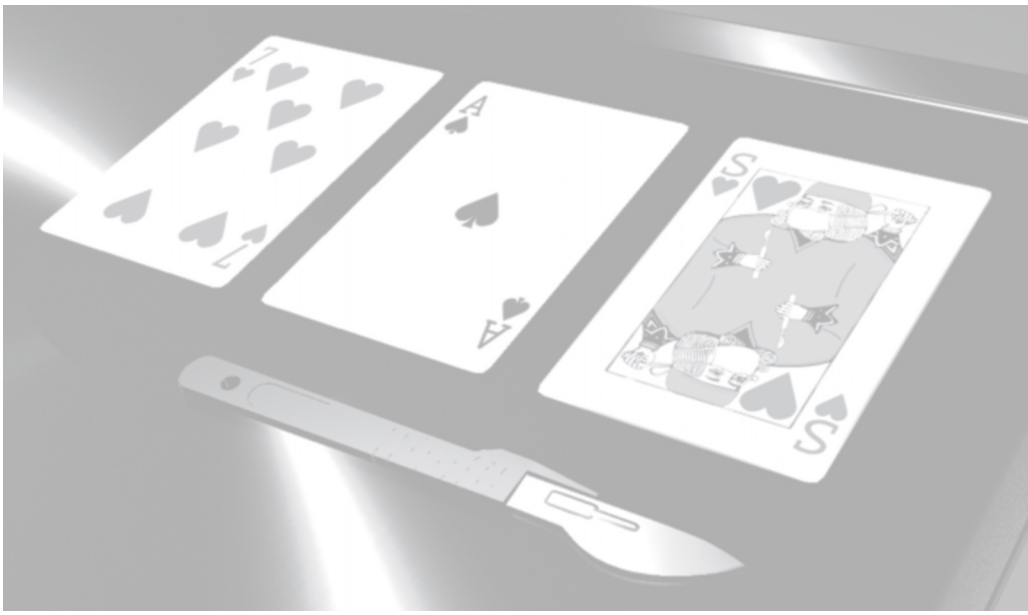
Promotor:	Prof.dr. A.J.J.C. Bogers
Overige leden:	Prof.dr.ir. H. Boersma Prof.dr. R.J.M. Klautz Prof.dr. E.W. Steyerberg
Copromotoren:	Dr. A.P. Kappetein Dr. J.J.M. Takkenberg

Financial support by the Netherlands Heart Foundation for the publication of this thesis is gratefully acknowledged.

Aan mijn ouders

CONTENTS

CHAPTER 1



General Introduction

There is a steady increase in the number of patients undergoing cardiac surgery in The Netherlands [1]. As can be appreciated from Figure 1, 16,877 adult surgical cardiac procedures were performed in 2008. In addition, the number of transcatheter procedures, including valve and coronary stent implantation, is also growing rapidly.

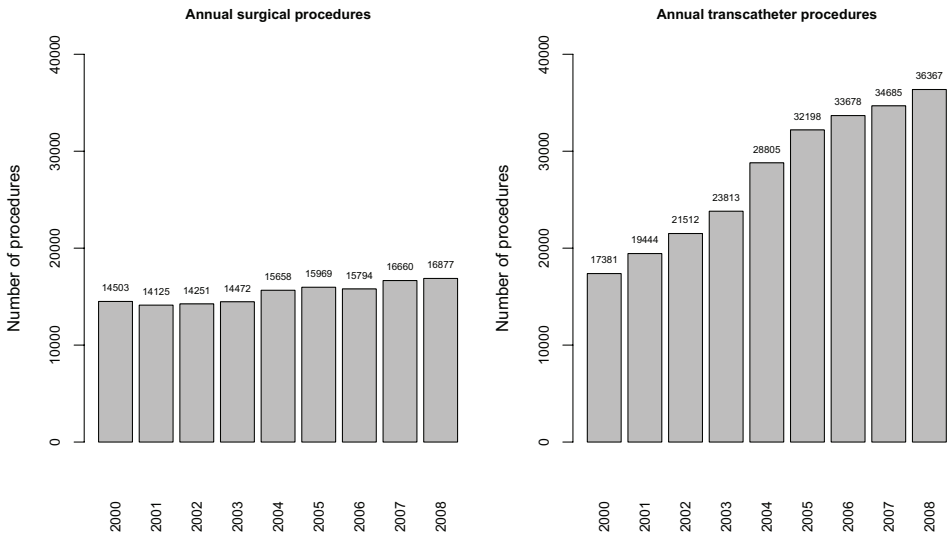


Figure 1. Annual cardiac procedures in The Netherlands [1].

Given the ageing of the population and the increasing number of patients with congenital heart disease that reaches adulthood, the number of surgical and transcatheter interventions is likely to increase even further [2, 3].

The growing population requiring these cardiovascular interventions will lead to an increase in health care expenditure. This calls for a cost-effective approach of health care, with constant attention for the relation between cost-effectiveness and quality of care. Quality assessment is an inherent component of this approach. In addition, by improving the quality of care (including optimizing treatment selection), fewer adverse outcomes are to be expected, with a subsequent restraint of costs.

QUALITY OF CARE

Quality assessment and improvement have emerged as a dominant theme in all fields of medicine. Traditional quality assessment focused mainly on outcomes. They are the most obvious, intuitive, and historically used measures of quality by both health care professionals and patients. Quality improvement based upon outcomes usually arises from identification and investigation of outliers and public accountability in the form of report cards of these outcomes [4]. A disadvantage of this report card system however is the risk of inducing avoidance of high-risk cases to ensure good outcome [5].

More recently, there is an increasing interest for continuous quality improvement (CQI) programs [4, 6, 7]. These quality improvement programs comprise a number of performance measures to evaluate and improve quality of care: outcome measures, process measures and structural measures [8]. CQI programs are usually confidential peer processes based upon benchmarking, exchange visits, identification and implementation of best practice, and education.

As both report cards and CQI programs have their own potential advantages, a combination of these approaches can be implemented as an alternative.

Whatever approach of quality assessment or improvement is used, outcome assessment and risk stratification are among the fundamental components.

OUTCOME ASSESSMENT

As soon as unadjusted outcome data were made publicly available (by Health Care Financing Administration in USA 1986 [9]), the interest in data collection, measurement, and analysis to allow for risk adjustment, increased [10]. Cardiac surgeons started collecting data to enable outcome corrections. This subsequently led to efforts to decrease the rate of less acceptable outcomes. Health care regulators and insurance companies then also started using outcome assessment to increase the quality and decrease the costs of cardiac surgery. The most assessed outcome measure is mortality; an easy to define, readily measured outcome, with an undeniable value for patients.

In addition to mortality, other endpoints can also be used to reflect outcome. Post operative complications as renal failure, stroke or reoperation, resource utilization, and quality of life are all considered important outcome mea-

sures for cardiovascular interventions [7].

Both short and long term outcomes can be assessed to reflect quality. An example of a short term outcome is 30 day mortality. Examples of long term outcomes are five year survival or complications that occur over time, for example related to an implanted device.

When used appropriately, a composite endpoint can also be used in outcome research [11].

Outcome measurements should be well and consistently defined to allow for valid assessment or comparison. For example, short term mortality after cardiac interventions can be defined as in-hospital mortality, mortality within a certain time period or a combination of both. Current guidelines recommend that early mortality after valve procedures should be reported as all cause mortality at 30, 60, or 90 days, regardless of the patient's location [12].

RISK STRATIFICATION

The risks involved with a given procedure differ from patient to patient and from procedure to procedure. For example, performing aortic valve surgery on an 80 year old hospitalized patient requiring dialysis carries a higher risk compared to the same procedure on a 60 year old patient without serious comorbidity. To properly assess outcomes, risk adjustment using patient and procedural characteristics is essential. Risk stratification is defined as the process of arranging patients according to their severity of illness.

Over the last couple of decades, many risk stratification systems have been developed and many are used in daily practice. Today's risk stratification methods are based upon the use of clinical prediction models, the risk stratification models [13-16].

Risk stratification models are generally used to provide outcome predictions for new patients. It is therefore important to know if a model provides valid predictions outside the dataset it was derived from.

The aspect of validity that is most frequently assessed is predictive validity. This measures how well a risk stratification model performs in another dataset. Predictive validity can be tested internally or externally. When a model is validated internally, a part of the population that was not used to develop the

model is used to validate the model. For example, a model that is developed in one center based upon patient data from 2005 and 2006, can be validated internally by using data from 2007 from the same center. This approach is called temporal validation. Other approaches to internally validate a model comprise randomly splitting the dataset, performing a cross-validation or drawing bootstrap samples. External validation is a more informative way of assessing predictive validity. A population that has no connection with the original dataset is then used to assess validity. Numerous validation studies have been conducted, mostly investigating whether geographical differences influence the predictive performance of risk stratification models. Many surgical models appear to perform well when used outside their original populations [17-20].

Estimation of the predictive validity of a model involves the evaluation of both discriminatory performance and calibration. Discrimination refers to the ability to differentiate patients with and without events. Calibration reflects the relation between the observed number of events and predicted probability of events. As models can be recalibrated to better fit a specific population, discriminatory performance is the most important measure of predictive validity.

Sometimes, the explained variation (R^2) is used to reflect the predictive ability of a model.

In addition to ensuring that a model has a good predictive ability and is clinically relevant, several other aspects of risk stratification models are of interest prior to implementation. Model variables and endpoints should be well defined and the model should ideally be developed using patient and procedural characteristics that are comparable to those in the population for which predictions should be made.

Risk stratification models are often based upon regression analyses. Other approaches include additive score systems or neural networks [21]. Regardless of the complexity of a risk stratification model, its accuracy will ultimately be limited by the quality of the data used; a model is only as good as the database it was derived from.

The field of cardiovascular interventions is dynamic, with continuous developments resulting in new approaches and improved outcomes of existing approaches. A change in treatment possibilities may affect outcome and risk stratification models. For example, patients that in the past were considered too high risk for surgical aortic valve replacement are nowadays treated [22]. It

is self-explanatory that these patients have different characteristics than those treated previously. In addition, different risk factors are associated with emerging interventions and the patient group undergoing them.

This might restrict general applicability or optimal performance of currently available models.

APPLICATIONS OF RISK STRATIFICATION

Risk stratification models can be implemented in several ways to assess or improve quality of care.

First, risk stratification models can be used for center and physician benchmarking. Results of individual physicians or centers are compared with results from others to provide a point of reference. Comparing predicted outcomes with observed outcomes allows for calculation of risk adjusted outcomes. As crude mortality or morbidity rates provide an inadequate measure for quality of care, adjustments of procedural complexity are used by both score card and CQI programs [4]. Examples are the publicly available New York State Report Cards [23], containing detailed results on performance of all surgeons and interventional cardiologists employed in that state, and the Gold Standards Reports and associated bubble charts provided by the European Association for Cardio-Thoracic Surgery (EACTS) as part of a feedback report of a CQI program (see Figure 2) [24].

Second, risk models can help educate patients and improve patient informed consent. Awareness of risks associated with available treatment options can be increased. Results from models should be interpreted and explained properly: obtained results are valid for a group of patients that –with the limitations of included model variables- have similar characteristics as the patient in question. Current models are not accurate enough to provide appropriate individual patient predictions. They can however help clinicians in treatment selection by adding an objective risk assessment to their clinical judgment. Risk models can also be incorporated into guidelines to identify high risk patients who may benefit from a specific treatment strategy.

Several applications of risk stratification models are available for cost containment. An example is the use of risk models to predict prolonged length of stay [14-16, 25, 26]. Finally, risk stratification models can be used in the field of

research. The relative importance of specific risk factors can be retrieved from model variables to evaluate the effects of preventive measures [27, 28]. In addition, clinical trials commonly incorporate risk models into their protocol [29].

CURRENT SITUATION

Until recently, the emphasis in measuring cardiovascular procedural outcome was on mortality. Nowadays, with improving procedural techniques and decreas-

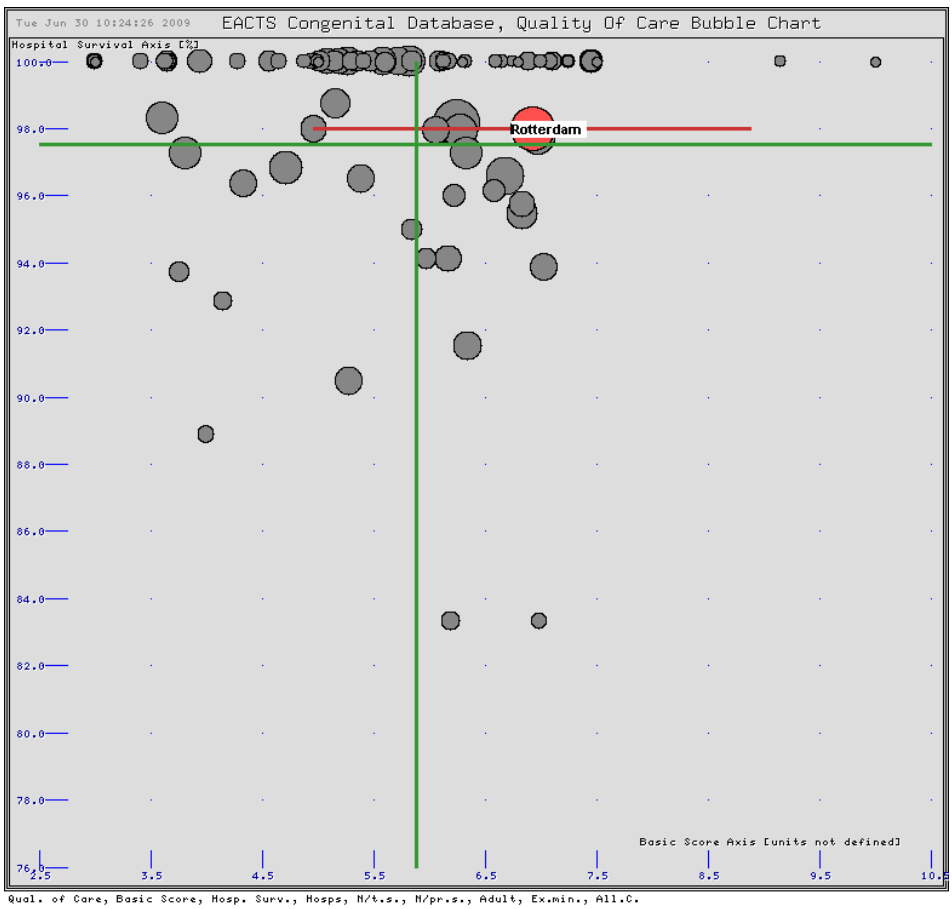


Figure 2. EACTS bubble chart example (hospital survival for adult congenital surgery patients)
This figure reveals that the Rotterdam population has an above average hospital survival while patient and procedural complexity is also above average. The size of the bubble furthermore suggests that the population is of considerable size in comparison to other centers.

ing mortality rates, there is a shift in focus towards measuring quality of care and morbidity rates are increasingly becoming an important way to describe procedural outcome.

With the aforementioned increasing concern about the quality, safety and cost-effectiveness of hospital care, and the growing demand for systematic reporting of mortality and morbidity after cardiothoracic procedures, the importance of adequate measurement of risk adjusted adverse events has further increased [7].

Currently the largest cardiac surgery database is that of the Society of Thoracic Surgeons (STS) [14-16]. Although this is a voluntary database, its size is unrivalled and models derived from this database are therefore generally highly valued. Currently, three main models are available; one for isolated valve surgery, one for coronary surgery and one for a combination of these procedures [14-16, 30]. A variety of submodels for different outcomes is available, ranging from mortality and serious morbidity to length of hospital stay.

In Europe, including the Netherlands, the EuroSCORE is the most commonly used risk stratification model. Data for this model was collected in 1995, comprises 60% coronary surgery procedures, and short term mortality is the sole model outcome measure [13]. A number of studies however demonstrated that the EuroSCORE might be used for other outcome measures as well [31-33].

In the Netherlands, each hospital performing cardiac surgery is committed to submit outcome and EuroSCORE prediction for each procedure to a national quality assessment registry.

Although the current approach of risk stratification is the result of decades of research, there is still room for improvement and updates are required to reflect current clinical practice. The STS models are updated on a regular basis. Recently, the EuroSCORE developers released plans to start data collection for a new model: EuroSCORE 2010 [34].

For pediatric congenital cardiac surgery, dedicated databases such as the EACTS Congenital Database [35] are available to assess outcomes. Procedure complexity is corrected by using the consensus based Aristotle Score [36]. Although adult congenital procedures are entered into this database as well, no separate risk stratification system is available for this population.

Several available risk stratification models are available online. Websites have been developed that enable both health care professionals and patients to calculate predicted mortality or morbidity rates [37-39]. Examples of these

online calculators are provided in Figure 3.

Until now, no transcatheter database of importance or adequate size is available. Subsequently, no reliable risk stratification model has yet been developed.

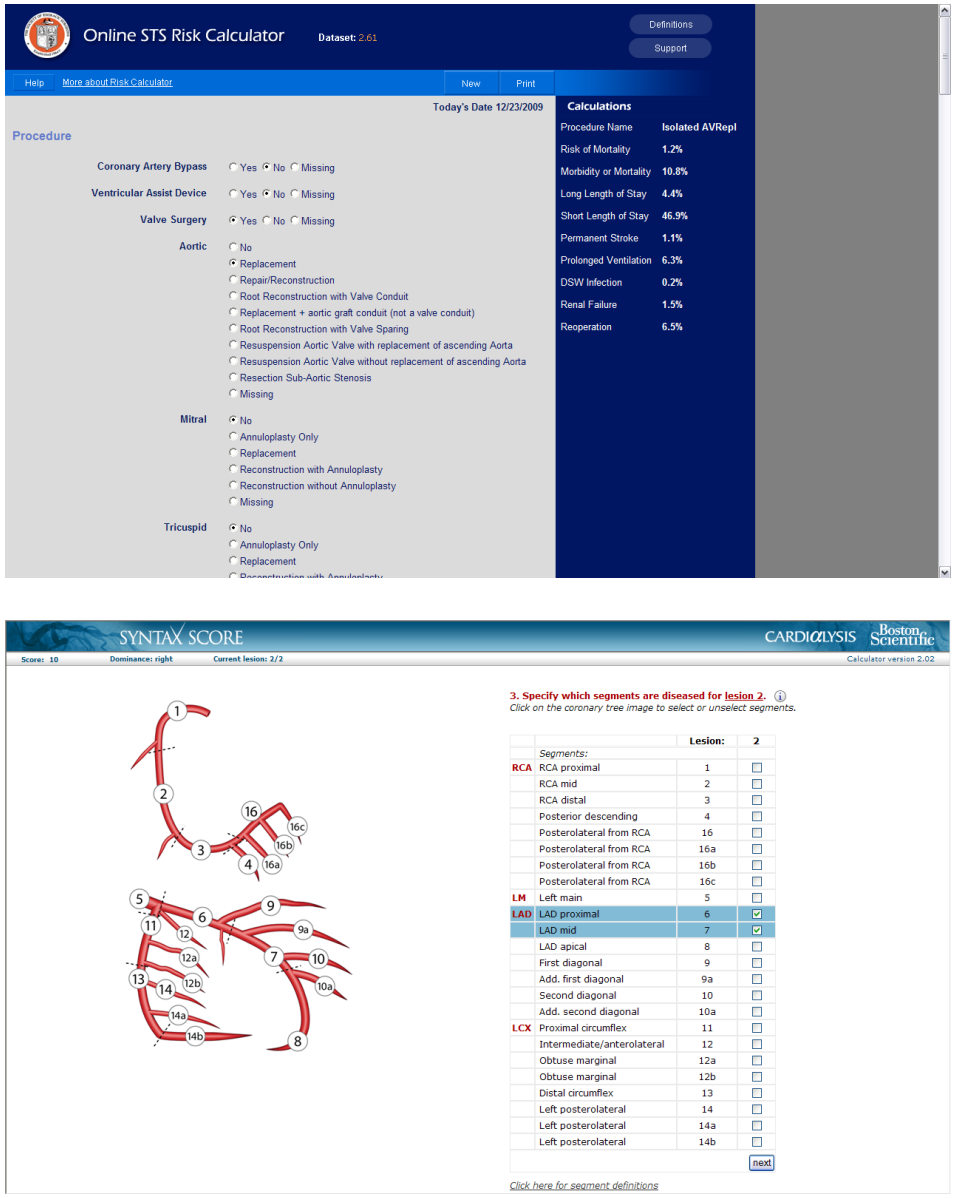


Figure 3. STS Score and SYNTAX Score online calculators (from top to bottom).

With increasing numbers of hospital admissions for procedures that lack appropriate risk stratification models, there is a need for further expansion and updates of current risk stratification models as well as development of new risk stratification models.

AIMS AND OUTLINE OF THE THESIS

The goal of this thesis is to study aspects of morbidity and mortality after cardiovascular procedures in current clinical practice, with special emphasis on the role of risk stratification models to improve outcome assessment.

To achieve this goal, the following research questions are addressed:

1. Can risk stratification models be optimized further for different types of cardiac surgery and transcatheter interventions?
2. What are the risks of procedures in patients with congenital heart disease and how should this heterogeneous patient group be stratified?
3. How can potential predictors of outcome be optimally used in risk stratification models for cardiac surgery?
4. Can clinical prediction models and quality improvement programs increase the quality and decrease the costs of cardiac surgery?

REFERENCES

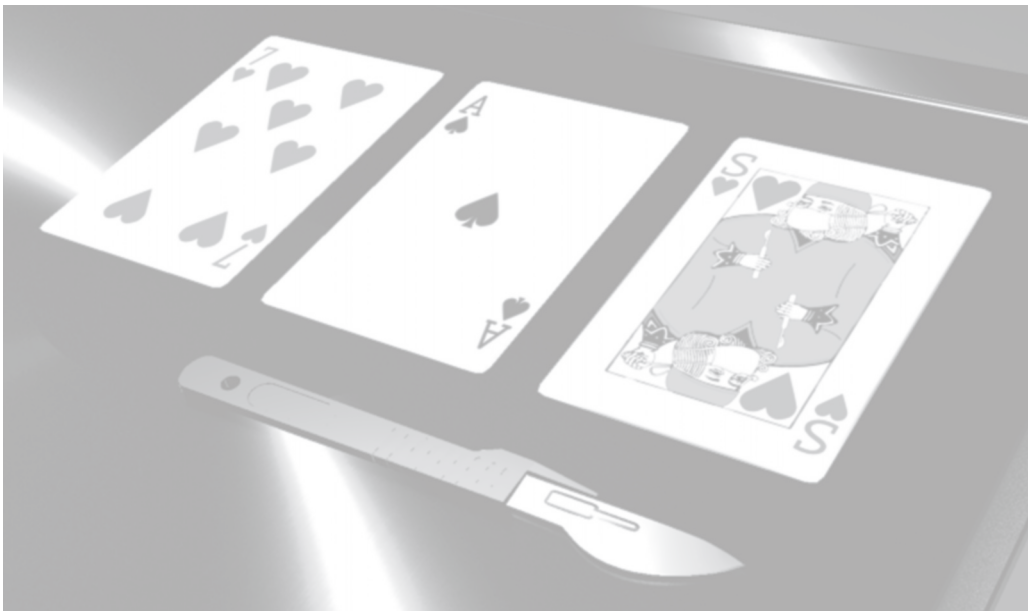
1. Nederlandse Hartstichting (BHN data). Available at: http://www.hartstichting.nl. Accessed 10-12-2009.
2. CBS. Available at: http://www.cbs.nl. Accessed 10-12-2009.
3. Warnes CA, Williams RG, Bashore TM, Child JS, Connolly HM, Dearani JA, del Nido P, Fasules JW, Graham TP, Jr., Hijazi ZM, Hunt SA, King ME, Landzberg MJ, Miner PD, Radford MJ, Walsh EP, Webb GD, Smith SC, Jr., Jacobs AK, Adams CD, Anderson JL, Antman EM, Buller CE, Creager MA, Ettinger SM, Halperin JL, Hunt SA, Krumholz HM, Kushner FG, Lytle BW, Nishimura RA, Page RL, Riegel B, Tarkington LG and Yancy CW. ACC/AHA 2008 guidelines for the management of adults with congenital heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Develop Guidelines on the Management of Adults With Congenital Heart Disease). Developed in Collaboration With the American Society of Echocardiography, Heart Rhythm Society, International Society for Adult Congenital Heart Disease, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol. 2008;52(23):e1-121.
4. Shahian DM, Normand SL, Torchiana DF, Lewis SM, Pastore JO, Kuntz RE and Dreyer PI. Cardiac surgery report cards: comprehensive review and statistical critique. Ann Thorac Surg. 2001;72(6):2155-68.
5. Omoigui NA, Miller DP, Brown KJ, Annan K, Cosgrove D, 3rd, Lytle B, Loop F and Topol EJ. Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. Circulation. 1996;93(1):27-33.
6. O'Brien SM, Shahian DM, DeLong ER, Normand SL, Edwards FH, Ferraris VA, Haan CK, Rich JB, Shewan CM, Dokholyan RS, Anderson RP and Peterson ED. Quality measurement in adult cardiac surgery: part 2--Statistical considerations in composite measure scoring and provider rating. Ann Thorac Surg. 2007;83(4 Suppl):S13-26.
7. Shahian DM, Edwards FH, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, O'Brien SM, Shewan CM, Dokholyan RS and Peterson ED. Quality measurement in adult cardiac surgery: part 1--Conceptual framework and measure selection. Ann Thorac Surg. 2007;83(4 Suppl):S3-12.
8. Donabedian A. Methods for deriving criteria for assessing the quality of medical care. Med Care Rev. 1980;37(7):653-98.
9. United States. Health Care Financing Administration. and United States. Dept. of Health and Human Services. Medicare hospital mortality information. HCFA pub. no. 01-002. U.S. Dept. of Health & Human Services For sale by the Supt. of Doc., U.S.G.. Washington, D.C., 1987.
10. Kouchoukos NT, Ebert PA, Grover FL and Lindesmith GG. Report of the Ad Hoc Committee on Risk Factors for Coronary Artery Bypass Surgery. Ann Thorac Surg. 1988;45(3):348-9.
11. Kip KE, Hollabaugh K, Marroquin OC and Williams DO. The problem with composite end points in cardiovascular studies: the story of major adverse cardiac events and percutaneous coronary intervention. J Am Coll Cardiol. 2008;51(7):701-7.
12. Akins CW, Miller DC, Turina MI, Kouchoukos NT, Blackstone EH, Grunkemeier GL, Takkenberg JJ, David TE, Butchart EG, Adams DH, Shahian DM, Hagl S, Mayer JE and Lytle BW.

Guidelines for reporting mortality and morbidity after cardiac valve interventions. Eur J Cardiothorac Surg. 2008;33(4):523-8.

13. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S and Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. 1999;16(1):9-13.
14. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1--coronary artery bypass grafting surgery. Ann Thorac Surg. 2009;88(1 Suppl):S2-22.
15. O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2--isolated valve surgery. Ann Thorac Surg. 2009;88(1 Suppl):S23-42.
16. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg. 2009;88(1 Suppl):S43-62.
17. Geissler HJ, Holzl P, Marohl S, Kuhn-Regnier F, Mehlhorn U, Sudkamp M and de Vivie ER. Risk stratification in heart surgery: comparison of six score systems. Eur J Cardiothorac Surg. 2000;17(4):400-6.
18. Nilsson J, Algotsson L, Hoglund P, Luhrs C and Brandt J. Comparison of 19 pre-operative risk stratification models in open-heart surgery. Eur Heart J. 2006;27(7):867-74.
19. Nashef SA, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, Wyse RK and Ferguson TB. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. Eur J Cardiothorac Surg. 2002;22(1):101-5.
20. Yap CH, Mohajeri M, Ihle BU, Wilson AC, Goyal S and Yii M. Validation of Euroscore model in an Australian patient population. ANZ J Surg. 2005;75(7):508-12.
21. Peng SY and Peng SK. Predicting adverse outcomes of cardiac surgery with the application of artificial neural networks. Anaesthesia. 2008;63(7):705-13.
22. Kappetein AP, van Geldorp M, Takkenberg JJ and Bogers AJ. Optimum management of elderly patients with calcified aortic stenosis. Expert Rev Cardiovasc Ther. 2008;6(4):491-501.
23. New York State Department of Health reports. Available at: http://www.health.state.ny.us/statistics/diseases/cardiovascular/. Accessed 21-12-2009.
24. EACTS Congenital Database Gold Standards Report. Available at: http://www.eactscongenitaldb.org/index.php?LANG=en&level=1&struct=14. Accessed March 2010.
25. Janssen DP, Noyez L, Wouters C and Brouwer RM. Preoperative prediction of prolonged stay in the intensive care unit for coronary bypass surgery. Eur J Cardiothorac Surg. 2004;25(2):203-7.
26. Kurki TS, Hakkinen U, Lauharanta J, Ramo J and Leijala M. Evaluation of the relationship between preoperative risk scores, postoperative and total length of stays and hospital costs in coronary bypass surgery. Eur J Cardiothorac Surg. 2001;20(6):1183-7.
27. Cleveland JC, Jr., Shroyer AL, Chen AY, Peterson E and Grover FL. Off-pump coronary artery bypass grafting decreases risk-adjusted mortality and morbidity. Ann Thorac Surg.

2001;72(4):1282-8; discussion 1288-9.
28. Ferguson TB, Jr., Coombs LP and Peterson ED. Preoperative beta-blocker use and mortality and morbidity following CABG surgery in North America. Jama. 2002;287(17):2221-7.
29. PARTNER Trial. Available at: http://clinicaltrials.gov/ct2/show/study/NCT00530894. Accessed May 2010.
30. Shahian DM and Edwards FH. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: introduction. Ann Thorac Surg. 2009;88(1 Suppl):S1.
31. Hirose H, Inaba H, Noguchi C, Tambara K, Yamamoto T, Yamasaki M, Kikuchi K and Amano A. EuroSCORE predicts postoperative mortality, certain morbidities, and recovery time. Interact Cardiovasc Thorac Surg. 2009;9(4):613-7.
32. Toumpoulis IK, Anagnostopoulos CE, Swistel DG and DeRose JJ, Jr. Does EuroSCORE predict length of stay and specific postoperative complications after cardiac surgery? Eur J Cardiothorac Surg. 2005;27(1):128-33.
33. Nilsson J, Algotsson L, Hoglund P, Luhrs C and Brandt J. EuroSCORE predicts intensive care unit stay and costs of open heart surgery. Ann Thorac Surg. 2004;78(5):1528-34.
34. EuroSCORE 2010. Available at: http://www.euroscore2010.org. Accessed May 2010.
35. EACTS Congenital Database. Available at: http://www.eactscongenitaldb.org. Accessed March 2010.
36. Lacour-Gayet F, Clarke D, Jacobs J, Comas J, Daebritz S, Daenen W, Gaynor W, Hamilton L, Jacobs M, Maruszsewski B, Pozzi M, Spray T, Stellin G, Tchervenkov C and Mavroudis And C. The Aristotle score: a complexity-adjusted method to evaluate surgical results. Eur J Cardiothorac Surg. 2004;25(6):911-24.
37. STS Score Online. Available at: http://www.sts.org/sections/stsnationaldatabase/riskcalculator/. Accessed December 2009.
38. SYNTAX Score Online. Available at: http://www.syntaxscore.com. Accessed December 2009.
39. EuroSCORE Online. Available at: http://www.euroscore.org. Accessed March 2010.

CHAPTER 2



How to assess risks of valve surgery: quality, implementation and future of risk models

Van Gameren M
Piazza N
Bogers AJJC
Takkenberg JJM
Kappetein AP

Heart. 2009; 95: 1958-63

KEY POINTS

- Calibration reflects the relation between the observed number of events and predicted probability of events.
- Discrimination is the model's ability to distinguish between patients who will suffer from an event and those that will not.
- Recalibration adjusts the weights of the independent variables to the population one intends to use the model for, allowing for extrapolation and increased precision of the model when used in that population.
- External validation is performed by using a different population (for example, from another centre) to assess discrimination and calibration.
- Internal validation is done by using patients from the same population the model was derived from.
- Risk models using in-hospital mortality may underestimate the true nature of the short term risk.
- Differences in epidemiology of disease, risk profiles, surgical strategies, decision making, selection bias, and referral bias can all influence the applicability and performance of a model.
- Current risk models used in heart valve surgery cannot be applied to a population considered for transcatheter heart valve implantation.

INTRODUCTION

Risk stratification models for operative mortality have gained widespread acceptance in cardiac surgery [1]. These models, however, are not 100% accurate. A number of factors can influence their performance. The introduction of transcatheter aortic valve technology and the need to identify more precisely those patients at high or prohibitive risk for surgery has created awareness among the cardiology and surgical community of the importance to understand the applications and limitations of these risk models.

This review aims to give an overview of currently available risk stratification models intended for patients undergoing cardiac surgery, in particular heart valve surgery, and describe the potential applications of risk models in the clinical and research setting. Model quality will be discussed through critical appraisal of currently available heart valve surgery risk models. Finally future directions are given for the development of heart valve surgery risk models in the context of emerging transcatheter valve therapies.

AVAILABLE RISK MODELS

Generally speaking, available risk models for cardiac surgery can be divided into three categories: (1) general cardiac surgery models — that is, coronary artery bypass surgery, valve surgery or other related cardiac surgery [2,3]; (2) general valve surgery models [4–10]; and (3) specific aortic valve surgery risk models [11,12].

More recently, investigators have developed dedicated risk models for patients undergoing valve surgery with or without concomitant coronary artery bypass surgery [4–12]. An overview of these risk models is provided in Table 1.

APPLICATIONS OF RISK STRATIFICATION MODELS

Risk models can serve multiple purposes if used correctly. Firstly, risk models can be used for benchmarking; they may allow for control of procedural complexity when analysing hospital and surgeon performance. Secondly, risk models can help educate patients and improve informed patient consent.

Table 1. Quality Aspects of currently used risk stratification models to predict early mortality after valve surgery.

	Reference/ dataset	Region	Model design	Year of publi-cation
1	Nashef et al. [3] - EuroSCORE	European	CABG, valve surgery	1999
2	Edwards et al. [5] - STS	USA+Canada	Ao/M without CABG	2000
			Ao/M with CABG	2000
3	Florath et al. [11] - Single center	Germany	Ao	2003
4	Nowicki et al. [9] -NNE	USA	Ao	2004
			M	2004
5	Gardner et al. [6] - VA	USA	Ao	2004
			M	2004
6	Jin et al. [8] - PHS	USA	Ao	2005
			M	2005
7	Ambler et al. [4] -SCTSGBI	UK	Ao/M/Combined Ao and M	2005
8	Rankin et al. [10] - STS score	USA	Ao/M/T/P/combined	2006
9	Hannan et al. [7] - New York State	USA	Ao/M/T/P/comb	2007
			Ao/M/T/P/comb + CABG	2007
10	Kuduvalli et al. [12] - Multi-center	UK	Ao	2007

Ao = aortic valve; CABG = coronary artery bypass graft surgery; H-L = Hosmer-Lemeshow; ROC = receiver operator characteristics; M = mitral valve; NNE = Northern New England medical centers; PHS = Providence Health System Cardiovascular Study Group; P = pulmonary valve; STS = Society

Risk models can also be incorporated into guidelines to help identify high risk patients who may benefit from additional work-up or alternative treatment strategies. If risk models can accurately identify high risk patients with expected longer lengths of stay in hospital, they may be useful for administrative logistic and budget planning.

Research and clinical trials commonly incorporate risk models into their protocol. For example, the CoreValve safety and efficacy trials and the PARTNER

Table 1. Continued.

Study period	Sample size	Sample size validation	Validation method	Discrimination results (ROC)	Calibration results (H-L)
1995	13302	1479	Internal (time-split)	0,79	0,68
1994-1997	49073	25460	Internal (time split)	0,77	0,23
1994-1997	43463	25852	Internal (time split)	0,74	0,14
1996-2001	1400	1400	Internal	0,73	N/A
1991-2001	5793	Bootstrapping	Internal validtion (Bootstrapping)	0,75	0,16
1991-2001	3150	Bootstrapping	Internal validtion (Bootstrapping)	0,79	0,71
1991-2001	7450	Not validated	Not validated	N/A	N/A
1991-2001	1850	Not validated	Not validated	N/A	N/A
1997-2004	3324	Development population	Internal	0,79	0,94
1997-2004	1596	Development population	Internal	0,84	0,14
1995-2003	16679	16160	Internal, time split	0,77	0,78
1994-2003	409100	Development population	Internal	0,74	N/A
2001-2003	10702	9662	Internal (time-split)	0,79	0,52
2001-2003	8823	8463	Internal (time-split)	0,75	0,04
1997-2004	4550	816+ Bootstrapping	Split+bootstrap	0,78	0,73

of Thoracic Surgeons database ; SCTSGBI = Society of Cardiothoracic Surgeons of Great Britain and Irelland; T = tricuspid valve; VA = Veteran Affairs records

USA trial, evaluating the Sapien valve, used the logistic EuroSCORE and the STS (Society of Thoracic Surgeon) score, respectively, as part of the patient inclusion criteria. In addition, the relative importance of specific risk factors can be retrieved from models to evaluate the effects of preventive measures.

ASPECTS OF MODEL QUALITY

How to test model performance

The fit of a (logistic regression) model is generally measured in terms of its discrimination and calibration [13]. The process of measuring these terms is called validation.

Discrimination

Discrimination, which is measured using the c-index (area under the receiver operating characteristic curve) with 95% confidence limits (CI), captures the model's ability to distinguish between patients who will suffer from an event and those that will not — thereby measuring the trade-off between the specificity and sensitivity. It is defined as the proportion of the time that an event-free patient is assigned a higher probability of not suffering from an event than a patient who suffers from an event. A value of 1.0 is perfect, and a value of 0.5 denotes only random ability to distinguish between patients with and without events. In general, a value above 0.7 is considered good and above 0.8 excellent.

Calibration

Calibration reflects the relation between the observed number of events and predicted probability of events and is evaluated by the Hosmer-Lemeshow goodness-of-fit test (H-L). Models with H-L p values above 0.05 are generally considered to be well calibrated. Calibration plots can be used to give a graphical representation of model calibration.

Comparing mean event rates (predicted versus observed) to assess calibration is not sufficient. Calibration is less crucial than discrimination in the evaluation process because models can be recalibrated. Recalibration is done by adjusting the weights of the independent variables to the population one intends to use the model for, allowing for extrapolation and increased precision of the model when used in that population.

Validation cohort

The cohort used to validate a model greatly influences results. A model is said to be ''internally'' validated when validation is performed using patients from the same population that the model was derived from.

Several approaches of internal validation are available: data splitting, cross

validation, or bootstrapping. Temporal validation is a form of data splitting and evaluates the performance of a model on subsequent patients within the same centre(s).

Complementary to internal validation is external validation. This method uses a different population (for example, from another centre) to assess discrimination and calibration and lends the model to extrapolation.

Determinants of model quality

Factors such as data quality, cohort characteristics, end point definitions, and types of risk factor variables play an important role in the performance and general applicability of risk models.

Data quality

The quality of the dataset is crucial — a model is only as good as the dataset it was derived from (''the garbage in-garbage out principle''). Data collection can be accomplished in a prospective or retrospective manner. Prospective data collection is the preferred approach. It is less vulnerable to missing data and allows for easier data auditing (source documentation, data source verification). Single centre and voluntary databases (for example, the STS database) might be vulnerable to reporting bias—that is, centres with above average clinical results might be more inclined to report their results.

Cohort characteristics

It is crucial to appreciate the characteristics of the patient cohort (for example, demographics, coronary and/or valve surgery patients) used to develop the risk model, in order to best extrapolate the model. Depending on the intended purpose of the model, its extrapolation based on a single centre experience may be less desirable than a model based on multicentre experience.

The sample size of the original cohort has a direct influence on the number of independent variables that can be identified in the model. As a statistical rule of thumb, at least five events are required for each model variable [14]. Although models derived from large cohorts may have the ability to identify a large number of independent variables, the clinical significance of variables always needs to be assessed.

End point definition

End point definition should be taken into account when interpreting and comparing model results. It is not uncommon that risk models intended for ''similar'' purposes use different definitions for short term mortality (for example, in-hospital vs 30 day mortality) (Table 1).

Variable types

Some models have converted continuous variables (for example, serum creatinine value) to binary or categorical variables using (arbitrary) cut-off points. Although this may improve the usability of the model, a degree of ''predictability'' may be lost.

QUALITY OF CURRENT MODELS

End point definition

All models listed in Table 1 use short term mortality as an end point. The definition of short term mortality, however, may differ across models. For example, whereas Gardner and Florath use 30 day mortality regardless of patient location, Ambler, Jin, Nowicki, Hannan, Kuduvalli and the EuroSCORE only look at in-hospital mortality regardless of time period. In contrast, Edwards and Rankin look at 30 day mortality and beyond that period if the patient still resides in-hospital

Table 2. Commonly used variables in risk models described in table 1 .

Variable	Used in	OR range
Age	100 % (9 of 9)	1.03-1.06 per year
Reoperation	100 % (9 of 9)	1.5-6.4
Valve position	100 % (7 of 7)*	1.7-2.8
Renal failure	89 % (8 of 9)	1.7-3.4
Urgent procedure	78 % (7 of 9)	1.3-1.8
Concomitant CABG	67 % (6 of 9)	1.5-1.9
NYHA class 4	56 % (5 of 9)	1.4-1.8

CABG = coronary artery bypass grafting; NYHA = New York Heart Association; OR = odds ratio. Whenever a model consists of two submodels, the risk factor used to base the split upon counts as a variable.

* Two models cover aortic valve procedures only.

(as defined by the STS database). It is important to realise that the risk of surgery may continue beyond the date of discharge. It is not unusual nowadays for cardiac surgery patients to be discharged 7–10 days after the index procedure. Thus risk models using in-hospital mortality may underestimate the true nature of the short term risk. There is a growing consensus that time since operation rather than location (hospital) should be used as an end point. According to the 2008 valve reporting guidelines, early mortality after valve procedures should be reported as all cause mortality at 30, 60, or 90 days, regardless of the patient's location, instead of in-hospital mortality [15].

Performance/validation

The EuroSCORE model for cardiac surgery has been validated in a number of countries including a North American population [16].

While other models use temporal validation [3–5,7], Nowicki and Kuduvalli also used bootstrapping techniques [9,12], Rankin [10] and Florath [11] evaluated discrimination only, thus lacking assessment of calibration and resulting in an incomplete validation. The Gardner model lacks validation of any kind [6].

Included variables

Most valve risk models include variables relating to patient demographics and cardiac status—some models include procedural details. Anatomical parameters (for example, valve morphology, left ventricular dimensions, degree of calcification) or intraprocedural factors known to influence outcome are not included. Thus, it is not uncommon that risk factors known to predict mortality are sometimes missing from risk models. The most commonly used variables in valve risk models are presented in Table 2.

Most of these variables are also identified as predictors of mortality after aortic valve replacement by a systematic review [17]. To illustrate how the predicted mortality may vary between the various models, Figure 1 displays the predict mortality by age for the STS score, the EuroSCORE and the NY State score.

Model applicability

Differences in epidemiology of disease, risk profiles, surgical strategies, decision making, selection bias, and referral bias can all influence the applicability and performance of a model.

The EuroSCORE was based on a cohort of patients in whom 60% of patients

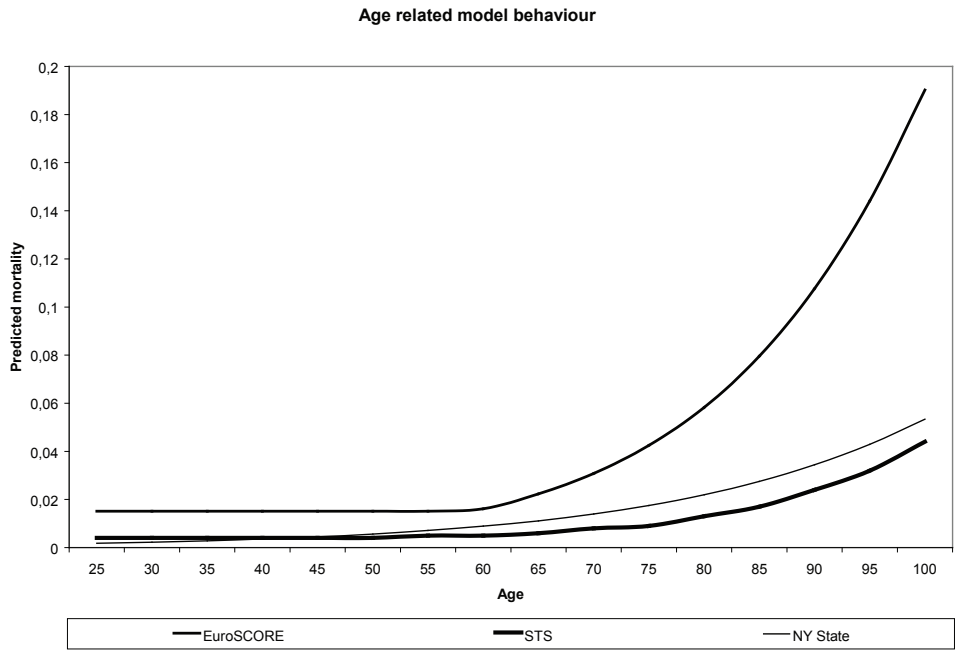


Figure 1. Predicted mortality by age for the STS (Society of Thoracic Surgeons) score, the EuroSCORE and the NY State score.

underwent coronary artery bypass graft (CABG) surgery, 30% underwent valve surgery with or without CABG, and 10% underwent ''other'' related cardiac surgery [3]. This may be one explanation why the EuroSCORE performs less well than dedicated valve models when used to stratify patients undergoing valve surgery [18].

Sustained improvements in cardiac surgery render older risk scores obsolete—it is expected that risk scores will overestimate the mortality rate as they become ''older''. As can be gleaned from Table 1, more ''recent'' risk scores are based on ''vintage'' data from 2004.

Inherent referral and selection bias can lead to the inclusion of ''healthier patients'' in risk models. Applying these risk models to patients considered high risk or inoperable may be associated with limitations.

To demonstrate potential differences in predicted mortality among risk models, the estimated mortality rate of two hypothetical patients undergoing isolated aortic valve surgery was calculated, and the results are presented in Table 3.

The first patient is a 75-year-old man with an ejection fraction of 55% and

no other comorbidities. This is a low risk patient. The second patient is an 85-year-old woman with an ejection fraction of 30%, renal dysfunction and pulmonary hypertension. This is a high risk patient.

Note that predicted mortality rates for a particular patient can vary depending on the risk model used. In our example, the predicted mortality for the low risk patient varied between 1–4%. In the high risk patient, this variation was more pronounced (4–60%). These variations can be explained by the different relative weights assigned to similar variables in the models and also by the fact that not all models include similar variables (that is, capture similar comorbidities).

THE WAY FORWARD

Sustained improvements in cardiac surgery and the emergence of transcatheter valve therapies necessitates that we reappraise existing cardiac surgery risk models and possibly develop new ones. Patients who are currently evaluated for transcatheter or transapical aortic valve replacement (AVR) form a small select population that is not representative of the populations that were used to develop the existing valve risk models. These models will therefore not perform well when applied to this ''outlier'' population.

Table 3. Estimated operative mortality in a low and high risk patient based on several models.

	Low risk patient 75 year-old male with ejection fraction 55% and no other comorbidities	**High risk patient** 85 year-old female with ejection fraction 30%, renal dysfunction and pulmonary hypertension
Nashef et al. [3] - EuroSCORE	4.3%	59.8%
STS online calculator, dataset 2.61 [23]	1.0%	15.4%
Nowicki et al. [9] -NNE	1.3%	4.3%
Jin et al. [8] - PHS	1.0%	6.6%
Rankin et al. [10]- STS score	3.4%	7.2%
Hannan et al. [7] - NY State	1.8%	8.5%
Kuduvalli et al. [12] - Multicenter	2.6%	27.4%

What is the role of risk modelling in the field of transcatheter aortic valve implantation? Firstly, a surgically based risk model that can accurately identify high or prohibitive surgical risk patients for transcatheter aortic valve implantation (TAVI) is needed for both clinical and research purposes; current risk models are not sufficient. The development of such risk models may be impractical given the extreme differences in characteristics of patients undergoing TAVI versus surgical aortic valve replacement (SAVR). It is axiomatic that clinical judgment will play an integral role in the selection process. Secondly, dedicated risk stratification models for TAVI should be investigated. Estimated mortality rates obtained from surgical models (for example, EuroSCORE) are currently being used as benchmark performance measures for TAVI. This practice can lead to complacency on the part of the cardiologist — accepting current results when there is definite room for improvement.

Ideally, a multicentre prospective database that aims to collect data on all valve interventions (that is, surgical and transcatheter) needs to be strongly considered. This endeavour may be endorsed by government or professional organisations (for example, European Association for Cardio- Thoracic Surgery, European Society of Cardiology). It is essential that data collection be comprehensive and include patient demographics as well as anatomic factors known to influence outcomes. Having said that, there needs to be a balance between being ''comprehensive'' and ''parsimonious''. Collection of a large number of variables might even be unnecessary. One could stratify a patient before treatment selection is made based upon a limited number of general risk factors (for example, age) that reflect general ''frailty'' by using a Charlson comorbidity scale [19] or Lee Score [20] equivalent. In a reaction to the Lee Score [20], De Craen even states that the discriminative power of a model with age and sex as sole variables is almost comparable to that of the same model with 12 extra variables [21]. Unfortunately, the effect on calibration was not discussed, while the extra variables might have made a substantial difference in calibration. Furthermore, incorporating too many variables into model calculations can result in ''overfitting'' — that is, the model follows characteristics of the population it was derived from to the extent that it only ''fits'' that population. Although some short term risk models have proven successful in the estimation of 1 year mortality [22], more attention is needed in long term risk assessment.

Most models are currently derived using a multivariate logistic regression analysis. New models should be updated (by redefining variable inclusion and

weightings) on a regular basis to ensure they reflect clinical practice at all times. While this approach may suffice for the future, other approaches using algorithmic models as, for example, artificial neural networks, should be investigated as well.

CONCLUSION

Currently, several risk models are available to predict early mortality after cardiac surgery. Many of these models, however, are applied inappropriately in the clinical or research setting. Our goal was to increase the awareness among the cardiology and cardiac surgery community of the quality, differences and possible applications of currently available heart valve risk models.

Importantly, risk models need to be continuously ''renewed''—that is, recalibration and refinement of important covariables. There should be an assessment of the need for, and usefulness of, more specific models aimed at particular subsets of patients.

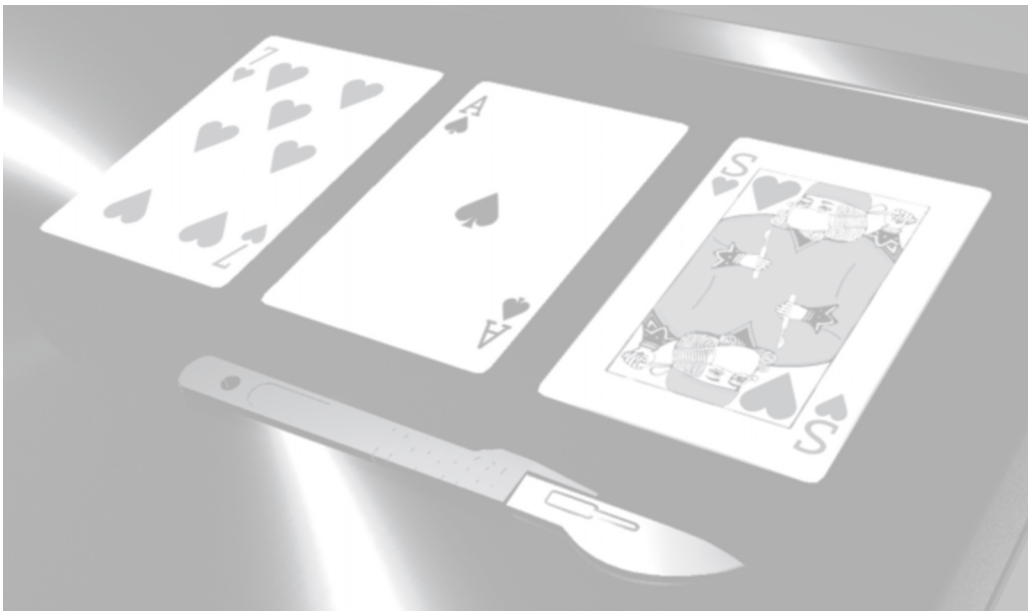
A large and complete database is required for decent modelling. This may be accomplished via a prospective collaborative effort including multiple centres and include valve interventions of all types (surgical and transcatheter).

REFERENCES

1. Nilsson J, Algotsson L, Hoglund P, et al. Comparison of 19 preoperative risk stratification models in open-heart surgery. Eur Heart J 2006;27:867–74.
   An excellent review on the various risk models in cardiac surgery. Describes which variables are used in which models.
2. Parsonnet V, Dean D, Bernstein AD. A method of uniformmstratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation 1989;79:I3–12.
   One of the first articles describing the use of a score to evaluate risk in cardiac surgery.
3. Nashef SA, Roques F, Michel P, et al. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9–13.
   Original article describing the EuroSCORE.
4. Ambler G, Omar RZ, Royston P, et al. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224–31.
5. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol 2001;37:885–92.
6. Gardner SC, Grunwald GK, Rumsfeld JS, et al. Comparison of short-term mortality risk factors for valve replacement versus coronary artery bypass graft surgery. Ann Thorac Surg 2004;77:549–56.
7. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting inhospital mortality for cardiac valve surgery. Ann Thorac Surg 2007;83:921–9.
8. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005;80:471–9.
9. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77:1966–77.
10. Rankin JS, Hammill BG, Ferguson TB Jr, et al. Determinants of operative mortality in valvular heart surgery. J Thorac Cardiovasc Surg 2006;131:547–57.
11. Florath I, Rosendahl UP, Mortasawi A, et al. Current determinants of operative mortality in 1400 patients requiring aortic valve replacement. Ann Thorac Surg 2003;76:75–83.
12. Kuduvalli M, Grayson AD, Au J, et al. A multi-centre additive and logistic risk model for in-hospital mortality following aortic valve replacement. Eur J Cardiothorac Surg 2007;31:607–13.
13. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453–73.
    Article describing what is meant by validation of a risk model.
14. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007;165:710–8.
15. Akins CW, Miller DC, Turina MI, et al. Guidelines for reporting mortality and morbidity after cardiac valve interventions. J Thorac Cardiovasc Surg 2008;135:732–8.
    Guidelines on how to report the outcome after surgical valve implantation.
16. Nashef SA, Roques F, Hammill BG, et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. Eur J Cardiothorac Surg 2002;22:101–5.

17. Tjang YS, van Hees Y, Korfer R, et al. Predictors of mortality after aortic valve replacement. Eur J Cardiothorac Surg 2007;32:469–74.
18. van Gameren M, Kappetein AP, Steyerberg EW, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg 2008;85:921–30.
Study indicating the need for risk models for heart valve surgery and heart valve surgery combined with coronary surgery.
19. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.
20. Lee SJ, Lindquist K, Segal MR, et al. Development and validation of a prognostic index for 4-year mortality in older adults. JAMA 2006;295:801–8.
Description of a ''frailty'' score to predict mortality in elderly people.
21. de Craen AJ, Westendorp RG. Prognostic index for 4-year mortality in older adults. JAMA 2006;296:648; author reply 648–9.
22. Toumpoulis IK, Anagnostopoulos CE, Toumpoulis SK, et al. EuroSCORE predicts long-term mortality after heart valve surgery. Ann Thorac Surg 2005;79:1902–8.
23. Society of Thoracic Surgeons. STS Online Calculator. http://www.sts.org/sections/stsnationaldatabase/riskcalculator/ (accessed 24 December 2008).
Link to the web based STS risk calculator.

CHAPTER 3



Do We Need Separate Risk Stratification Models for Hospital Mortality After Heart Valve Surgery?

Van Gameren M
Kappetein AP
Steyerberg EW
Venema AC
Berenschot EAJ
Hannan EL
Bogers AJJC
Takkenberg JJM

Ann Thorac Surg. 2008; 85: 921-30

ABSTRACT

Background

The EuroSCORE is often used to benchmark and predict hospital mortality after cardiac surgery. Based mainly upon coronary surgery patients, EuroSCORE may not be optimal for valve surgery patients. We evaluated the New York (NY) State dedicated valve surgery models and compared their performance to the EuroSCORE model.

Methods

Required model variables were collected prospectively for all patients, followed by calculation of predictive mortality rates using the logistic and additive EuroSCORE, the logistic and additive NY State models for valve surgery without concomitant coronary surgery (isolated valve surgery) and the logistic and additive NY State models for combined valve and coronary surgery.

Results

Observed mortality was 2.8% (25/904) for isolated valve surgery and 6.8% (27/395) for valve+coronary surgery. Logistic NY State and EuroSCORE expected mortality for isolated valve surgery was respectively 3.0% and 6.1%, and for valve+coronary surgery 5.9% and 7.8%. The logistic NY State model for isolated valve surgery showed better discrimination (c-index 0.86 versus 0.76) and calibration than the logistic EuroSCORE. Discriminatory power for the logistic NY State model for valve+coronary surgery was comparable to the logistic EuroSCORE (c-index 0.74 versus 0.72), as was calibration.

Conclusions

Our results suggest that dedicated risk models for valve surgery may be useful to provide more valid estimates of hospital mortality after heart valve surgery. Further exploration is needed to demonstrate general applicability of our results and assess the possible additional value of separate models for isolated valve surgery and valve+coronary surgery, and/or aortic and mitral valve surgery.

INTRODUCTION

Risk stratification models are increasingly important in the current clinical practice for two purposes. They can serve as a hospital performance benchmark, but can also be used to provide the surgeon and the patient with a quantitative estimate of the procedural risk, or to study the impact of particular risk factors on outcome [1]. In many European institutions, the EuroSCORE model is used to estimate hospital mortality after cardiac surgery [2].

The EuroSCORE model serves as a general cardiac surgery risk stratification model, while within the cardiac surgery population, there is a wide variety of procedures with different determinants of early mortality. Based on a dataset with mainly coronary surgery patients and only 30% valve surgery patients [3], EuroSCORE therefore may not be an optimal predictive model for valve surgery patients. Recently, several new models to predict early mortality after valve surgery have been developed that show an adequate to good performance [4-9]. However, while EuroSCORE may not be ideal to model early mortality after valve operations, the use of multiple models for different types of adult cardiac surgery may be more time and resource consuming, and may perhaps not have sufficient added value compared to the use of one common risk model.

Therefore we sought to answer the question whether in our institution a valve dedicated risk stratification model would provide improved hospital mortality risk prediction for patients undergoing valve surgery. This was done by validating 2 recently developed valve dedicated risk stratification models (a model for valve surgery without concomitant coronary surgery (isolated valve surgery) and a model for valve surgery with concomitant coronary surgery (valve+coronary surgery) from New York (NY) State (USA) [4] and comparing their performance with the EuroSCORE model.

METHODS

Study design and data collection

Patient and procedural risk factors for all adult patients who undergo open heart surgery in Rotterdam are systematically collected at the time of procedure, resulting in an almost completely prospectively collected dataset. The dataset was completed by using operative reports to retrospectively identify patients who

presented an extensively calcified ascending aorta during surgery.

All collected variables were compliant with published EuroSCORE and NY State model variable definitions [2,4]. Variables that initially did not meet all model definitions were all converted. EuroSCORE and NY State variables use different binary creatinine level and age cut off points. Since we stored these as continuous variables, both definitions could be met. Myocardial infarction and surgery interval was also stored as a continuous variable, enabling a correct binary conversion for all models. Systolic left ventricular function (LVF) was measured qualitatively (good, impaired, moderate and bad), thus complying to the EuroSCORE definition but deviating from the NY State model definitions (ejection fraction percentage; EF). It was therefore assumed that good LVF equals EF≥40%, impaired LVF equals EF 30-39%, moderate LVF equals EF 20-29% and bad LVF equals EF<20%.

For this study, all patients who underwent surgery between January 2003 and January 2007 were selected. To meet all investigated model requirements, we excluded patients with concomitant ascending aortic surgery, aortic valve repair surgery, isolated tricuspid surgery or isolated pulmonary valve surgery. Whenever patients underwent multiple operations within one month or during the same hospital admission, only the first (index) procedure was included. The data was stored in a local cardiac surgery database. Two patient groups were created from this database; patients who underwent isolated valve surgery were separated from patients who underwent valve surgery with concomitant coronary surgery. For all patients, predictive mortality rates were calculated using the published logistic and additive EuroSCORE coefficients [10]. Additional predictive mortality rates were added using the published NY State logistic and additive model coefficients [4] for their corresponding target group.
The dependent variable in this study was hospital mortality.

Table 1. Legend.
[a]Equals 2.5 mg/dL;
[b]Meets EuroSCORE definition of recent myocardial infarction.
[c]Meets NY State model definitions.

AVR = aortic valve replacement; COPD = chronic obstructive pulmonary disease; LVF = left ventricular function; MI = myocardial infarction; MV = mitral valve; MVR = mitral valve replacement; PA = pulmonary artery

Table 1. Baseline patient characteristics .

	Isolated valve (N=904)	**Valve+CABG (N=395)**
Age, Mean (SD; median)	60.1 years (15; 63)	68.4 years (9; 70)
Sex, M/F-ratio	479/425	272/123
Creatinine (umol/L), Mean (SD; median)	95.7 (60; 86)	101.2 (43; 93)
Renal failure		
Creatinin > 221micromol/L[a]	1.2%	1.3%
Dialysis	0.4%	0.3%
Diabetes mellitus	5.3%	13.7%
COPD	10.1%	10.1%
Extracardiac arteriopathy	3.8%	11.4%
Severe neurologic dysfunction	1.3%	1.0%
Prior cardiac surgery	16.6%	5.3%
Active endocarditis	4.8%	1.8%
Preoperative critical state	2.8%	6.8%
Unstable angina	0.3%	4.8%
Systolic LVF		
Good	72.0%	53.2%
Impaired	16.2%	22.0%
Moderate	9.8%	19.2%
Bad	2.0%	5.6%
MI		
<90 days[b]	1.4%	11.4%
<14 days[c]	0.2%	-
1-7 days[c]	-	0.5%
6-23 hours[c]	-	1.0%
<6 hours[c]	-	0%
Systolic PA pressure >60 mmHg	6.0%	7.8%
Emergency procedure	2.1%	2.5%
Hemodynamic state	2.5%	5.8%
Unstable	0.2%	1.0%
Shock		
Type valve surgery		
AVR isolated	43.8% (N=396)	51.1% (N=202)
MVR isolated	13.4% (N=121)	8.9% (N=35)
MV repair isolated	24.2% (N=219)	28.6% (N=113)
Multiple valve	18.6% (N=168)	11.4% (N=45)
Hepatic failure	0.1%	0.5%
Cerebrovascular disease	5.4%	6.1%
Extensively calcified ascending aorta	0.8%	2.0%

Statistical analysis

Continuous data are presented as mean ±1 standard deviation, and median. Categorical data are presented as proportions. To evaluate the performance of risk models, we tested both discrimination (resolution) and calibration (reliability).

Discriminatory power was assessed using the c-index (area under the receiver operating characteristic curve) with 95% confidence limits (CI). A c-index of 1.0 would indicate perfect discrimination, whereas a c-index of 0.50 indicates total absence of discrimination. A value between these extremes is the quantitative measure of a model's ability to distinguish between survivors and non-survivors. To demonstrate significant differences between c-indices, a bootstrapping cycle of 2000 runs was performed [11]. Tests between c-indices were 2 sided with $p<0.05$ considered to be a significant difference.

Calibration was evaluated by the Hosmer-Lemeshow goodness-of-fit test (H-L) and graphically by a calibration plot. The dashed smooth curve in a calibration plot reflects the non-parametric relation between observed mortality and predicted probability of mortality. Perfect calibration is represented by the straight dotted line through the origin. Triangles are based on quintiles of patients with similar predicted probabilities. Spikes at the bottom of a calibration plot represent the distribution of predicted probabilities. Models with Hosmer-Lemeshow p-values above 0.05 were considered to be calibrated well for our population.

Descriptive statistical analyses were performed with SPSS version 13.0 (SPSS, Chicago, Illinois) and R version 2.5.1 (The R Foundation for Statistical Computing) was used for bootstrapping, calculating c-values with 95% CI, Hosmer-Lemeshow p-values and constructing ROC (receiver operating characteristic) curves and calibration plots.

RESULTS

Patient population

All 1516 isolated valve surgery procedures (valve surgery without coronary surgery) or valve surgery procedures with concomitant coronary surgery that were performed in our institution between January 2003 and January 2007 were evaluated for study inclusion. From this cohort 201 procedures involving patients who underwent concomitant ascending aortic surgery, aortic valve repair surgery, isolated tricuspid surgery or isolated pulmonary valve surgery were

Table 2. Model performance in isolated valve surgery patients.

	Discrimination		Calibration	
	AUC c-index	(95% CI)	H-L p-value	Chi Square
NY State logistic	0.86	(0.78-0.94)	0.63	7.02
NY State additive	0.86	(0.78-0.94)	0.62	7.15
EuroSCORE logistic	0.76	(0.66-0.86)	0.004	24.16
EuroSCORE additive	0.77	(0.67-0.87)	0.007	22.76

AUC = area under curve; CI = confidence interval; H-L = Hosmer-Lemeshow

Table 3. Model performance in valve+CABG surgery patients.

	Discrimination		Calibration	
	AUC c-index	(95% CI)	H-L p-value	Chi Square
NY State logistic	0.74	(0.65-0.82)	0.03	18.59
NY State additive	0.72	(0.64-0.81)	0.44	8.91
EuroSCORE logistic	0.72	(0.63-0.81)	0.76	5.78
EuroSCORE additive	0.71	(0.62-0.80)	0.78	5.57

AUC = area under curve; CI = confidence interval; H-L = Hosmer-Lemeshow

excluded. Furthermore, there were 16 patients who underwent multiple valve operations within the same month or admission; we excluded 16 procedures for this reason. This resulted in a database of 1299 procedures. Baseline patient characteristics are presented in Table 1.

Observed hospital mortality was 2.8% (25/904) for isolated valve surgery and 6.8% (27/395) for valve+coronary surgery.

Isolated valve surgery

PREDICTED MORTALITY: The NY State logistic risk model for isolated valve surgery predicted a hospital mortality of 3.0% (95% CI 2.7-3.3) versus a 2.8% observed rate for all isolated valve surgery patients. The NY State additive model predicted a mortality of 3.4% (95% CI 3.1-3.7) for this group, the logistic EuroSCORE predicted a mortality of 6.1% (95% CI 5.6-6.6) and the additive EuroSCORE a mortality of 5.3% (95% CI 5.1-5.5).

DISCRIMINATION: Based upon discriminatory performance in the isolated valve surgery group, ROC curves (Figure 1) were generated for the logistic and additive NY State and EuroSCORE models. Table 2 presents the corresponding

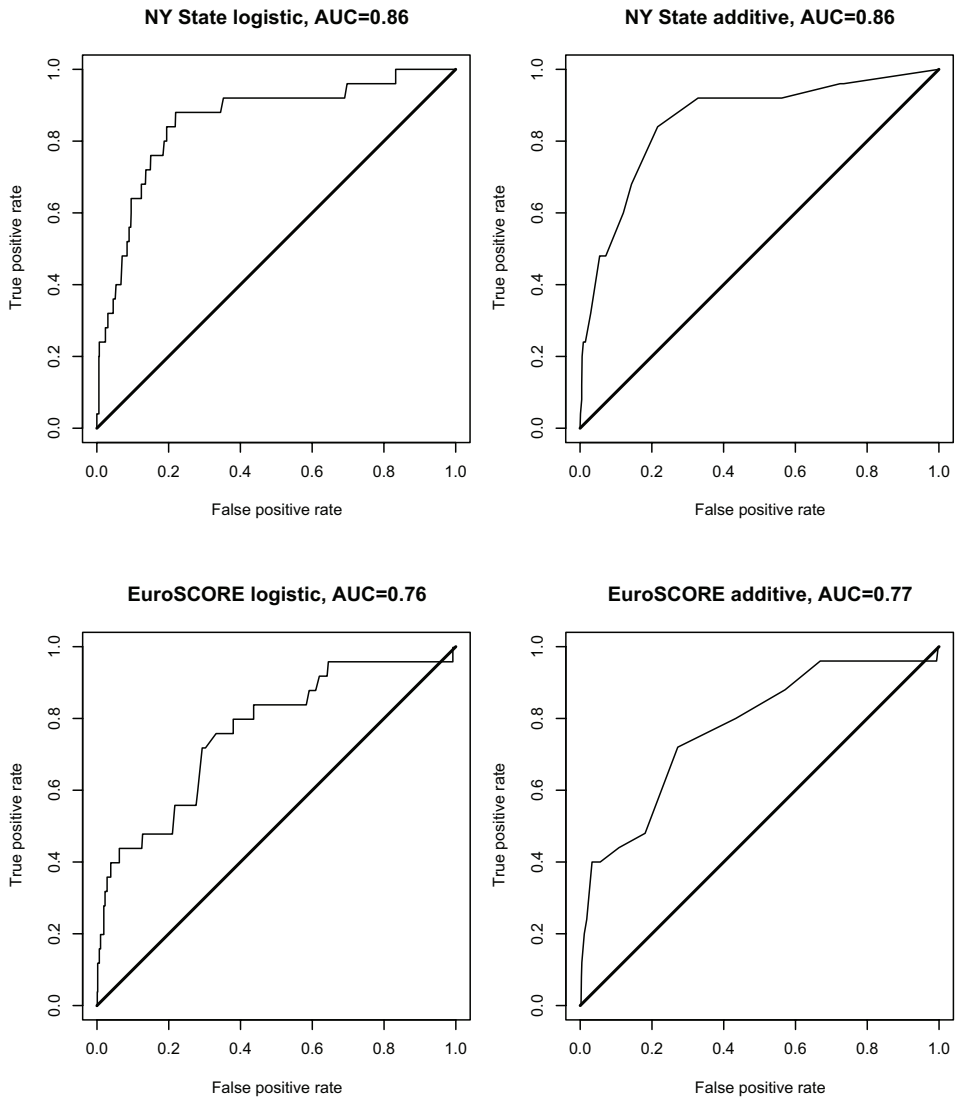


Figure 1. ROC curve graphs for the isolated valve surgery group (n=904).

c-indices. The c-indices of the NY State models are significantly higher than the EuroSCORE model c-indices (0.86 and 0.86 versus 0.74 and 0.76; all $p<0.05$), demonstrating a better discriminatory power for the NY State models in our population.

CALIBRATION: For the isolated valve surgery group, calibration of the NY State models for isolated valve surgery and the EuroSCORE models is graphically

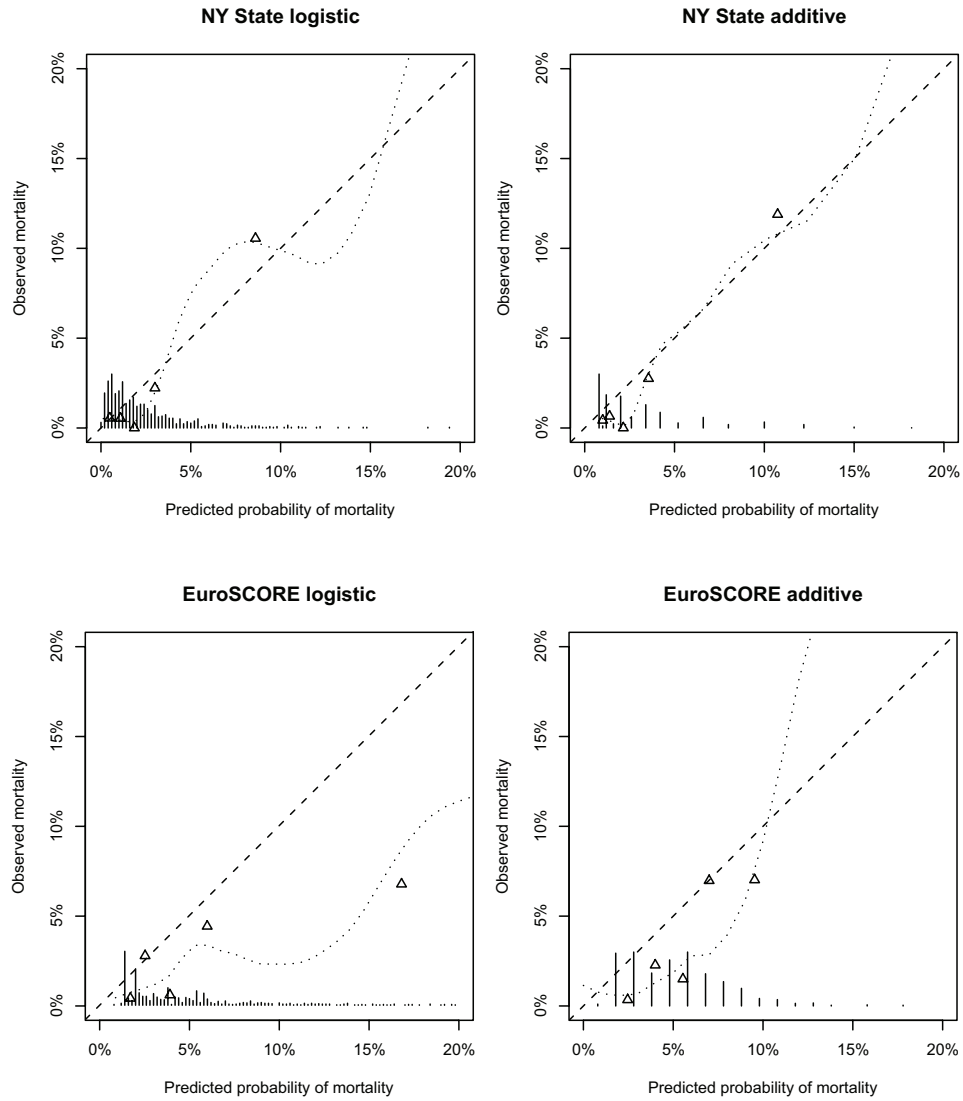


Figure 2. Calibration plots for the isolated valve surgery group (n=904).

presented in calibration plots (Figure 2). In both NY State models, the lines that represent the relation between observed frequency and predicted probability are closer to the ideal, diagonal line. While Hosmer-Lemeshow statistics for the NY State models demonstrate a good calibration, the EuroSCORE models do not fit adequately (see Table 2).

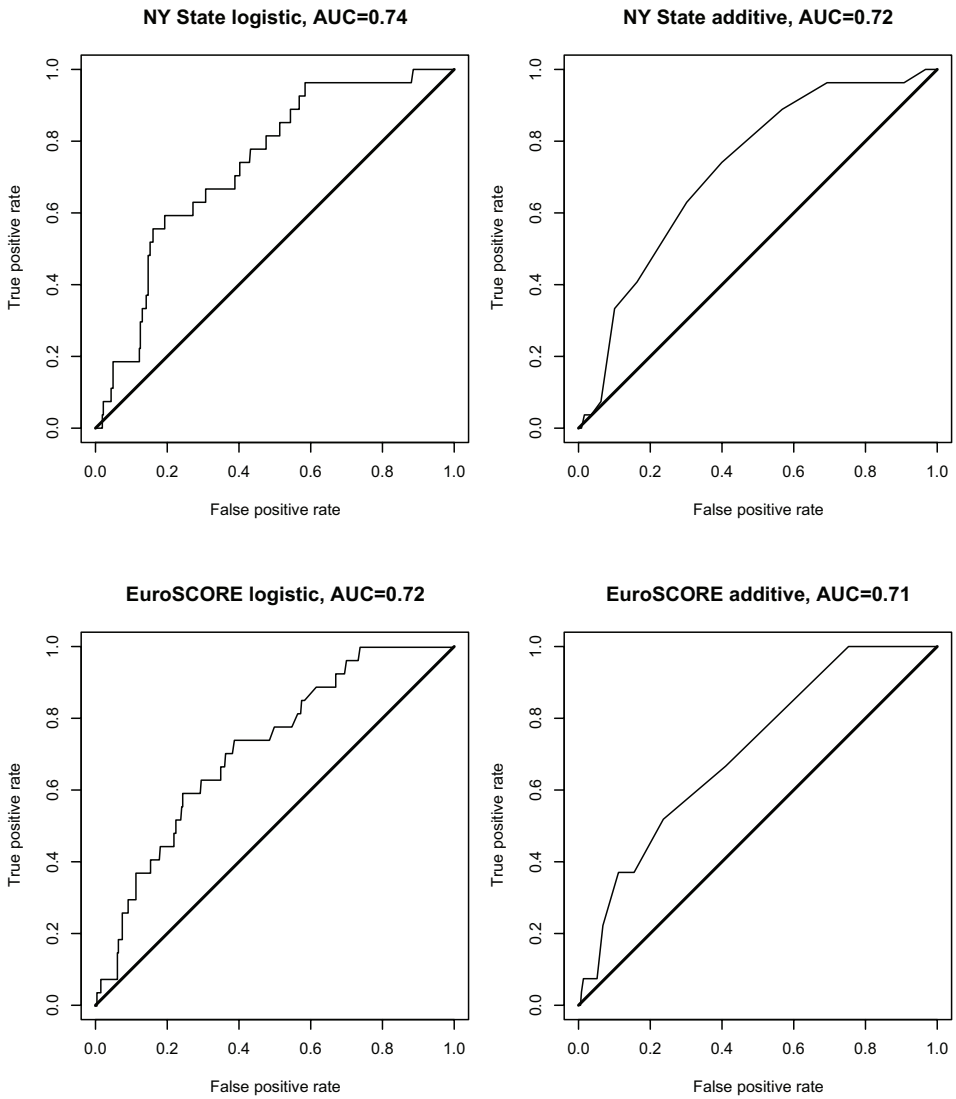


Figure 3. ROC curve graphs for the valve+coronary surgery group (n=395).

Valve surgery with concomitant coronary surgery

PREDICTED MORTALITY: The NY State logistic risk model for valve+coronary surgery predicted a hospital mortality of 5.9% (95% CI 5.4-6.4) for all valve+coronary surgery patients versus a 6.8% observed rate. The NY State additive model predicted a mortality of 6.2% (95% CI 5.6-6.7), the logistic EuroSCORE predicted a mortality of 7.8% (95% CI 7.0-8.6) and the additive EuroSCORE predicted 6.4%

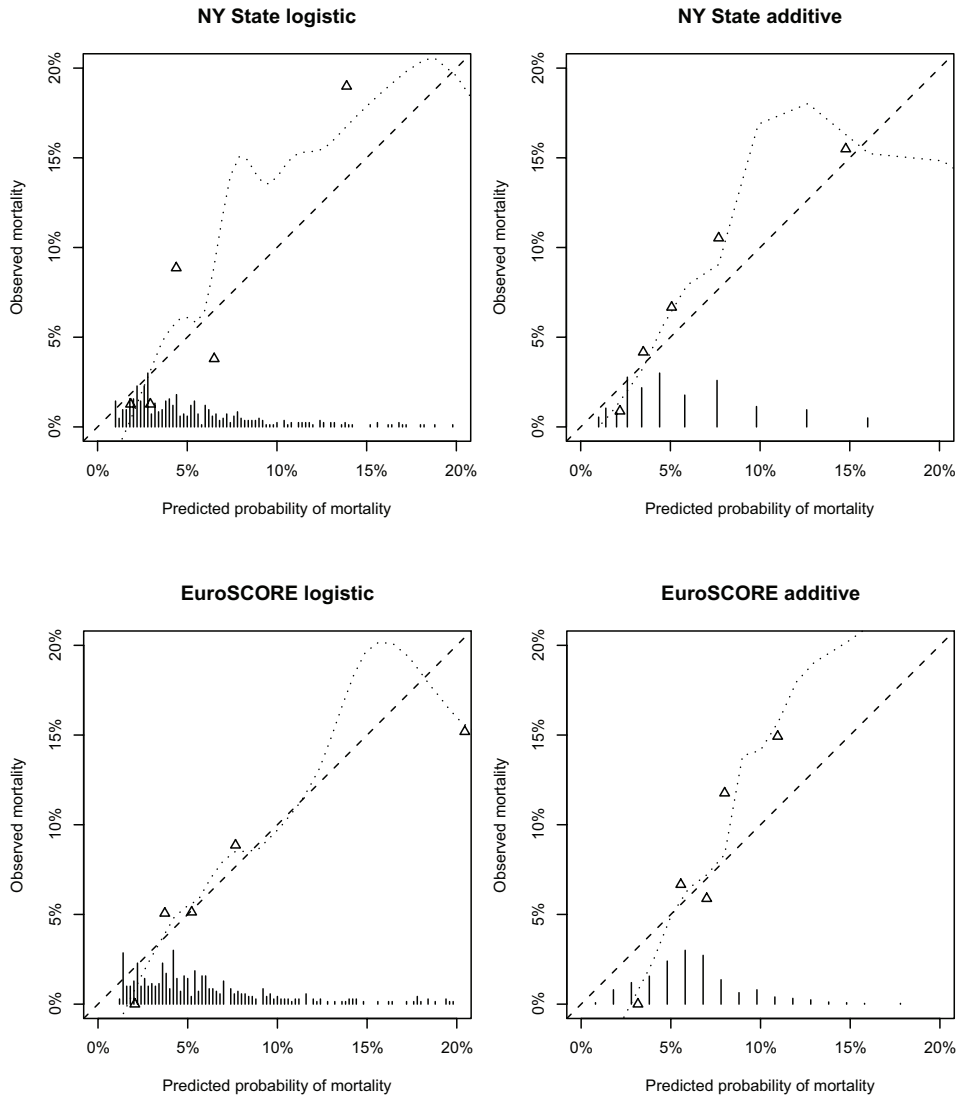


Figure 4. Calibration plots for the valve+coronary surgery group (n=395).

(95% CI 6.1-6.7) for this group.

DISCRIMINATION: When comparing the c-indices of the NY State models for valve+coronary surgery to the EuroSCORE model c-indices (see Table 3), no significant differences were found in the valve+coronary surgery group; none of the models outperformed the others regarding discrimination. A visual representation using ROC curves is given in Figure 3.

CALIBRATION: With a significant Hosmer-Lemeshow p-value (see Table 3), the NY State logistic model for valve+coronary surgery proved to be poorly calibrated for our population, while its additive counterpart and the EuroSCORE models fit our population quite well. However, the corresponding calibration plots for all models (see Figure 4) show that the curves that represent the relation between observed mortality and predicted probability of mortality deviate substantially from the ideal line in all models. This is especially true for patients with hospital mortality risks of over 10%.

COMMENT

The purpose of this study was to determine whether in our institution a valve dedicated risk stratification model would provide improved hospital mortality risk prediction for patients undergoing valve surgery. Little is known about EuroSCORE performance regarding mortality in a valve surgery population when compared with a dedicated valve surgery model. The results show that in an isolated valve surgery patient population, the logistic and additive NY State isolated valve models outperform the EuroSCORE models, both with regard to discrimination and calibration. This suggests that indeed for this particular patient population there is room for improvement of operative risk prediction. On the other hand, for patients in our institution who undergo valve surgery combined with coronary surgery, there is no added value in using the NY State valve+coronary model over the EuroSCORE models. In fact, all models demonstrated a poor performance, suggesting that for this particular patient population a better predictive model still needs to be developed, or that valve+coronary surgery populations differ too much to allow usage of externally developed models.

Rotterdam and New York State population

When comparing the Rotterdam population with the one that the NY State models were based upon, only a few differences were found. Except for a higher prevalence of diabetes mellitus (valve+coronary surgery group: 13.7% versus 29.0%), cerebrovascular disease (valve+coronary surgery group: 6.1% versus 22.6%) and chronic obstructive pulmonary disease (isolated valve surgery group: 10.1% versus 18.5%) in the NY State population, both populations appear very similar.

Observed hospital mortality rates were slightly higher for the NY State population compared to the Rotterdam population for both the isolated valve sur-

gery patient groups (4.4% versus 2.8%) and the valve with concomitant coronary surgery groups (8.9% versus 6.8%). Nevertheless, the NY State isolated valve surgery models only provide a slight overestimate of observed mortality in the Rotterdam population (3.0% and 3.4% versus 2.8% observed), while the NY State valve+coronary surgery model slightly underestimated observed mortality in the Rotterdam dataset (5.9% and 6.2% versus 6.8%).

Apart from some small differences in patient characteristics between NY State and Rotterdam and the observed mortality differences, no evident explanation from our data can be found for the fact that the NY State valve with concomitant coronary surgery model does not outperform the EuroSCORE like the NY State isolated valve model does.

Risk factors

In the past 5 years, several new models to predict early mortality after valve surgery have been developed that show an adequate to good performance [4-9]. Table 4 gives an overview of these models and their included risk factors, along with the EuroSCORE model. Many risk factors known to be relevant to predict mortality after valve surgery are included in the EuroSCORE [2,12]. However, several other important factors like diabetes mellitus and valve position however are missing in this model [4-9]. We assume a correlation between diabetes mellitus and factors that are included in the EuroSCORE model, like extracardiac arteriopathy, is responsible for the absence of diabetes as a risk factor in this model.

The weights of identical risk factors also differ, as illustrated for age by the following example. An eighty year old male patient without comorbidities has a predicted mortality of 6% after aortic valve replacement according to the logistic EuroSCORE, while the same patient only scores 2% when using the logistic NY State model.

Emergency surgery is a well-known predictor of higher mortality rates, and the EuroSCORE model does include emergency procedures as a risk factor, as do all other recently published valve dedicated models. However, in both NY State models, no direct risk factor for emergency procedures was included. This was explained by a correlation between emergency status and shock or hemodynamic instability, factors that are included in the NY State models [13].

Furthermore, all recently developed models mainly use comorbidities as risk factors to predict procedure mortality. Cardiac and valve anatomy are how-

Table 4. Recently published valve dedicated risk stratification modelsa and the EuroSCORE .

Model	**Hannan [4]**		**Rankin [5]**	**Jin [6]**	**Ambler [7]**	**Gardner [9]**		**Nowicki [8]**		**Nashef [2]**
Dataset	NY State		STS	PHS	SCTS	VA		NNE		EuroSCORE
Region	USA		USA	USA	UK	USA		USA		Europe
Year of collection	2001-2003		1994-2003	1997-2004	1995-2003	1991-2001		1991-2001		1995
Year of publication	2007		2006	2005	2005	2004		2004		1999
Model subgroups	isolated valve	valve +coronary	-	-	-	aortic	mitral	aortic	mitral	-
Sample size	10,702	8,823	409,100	4,914	32,839	7,450	1,870	5,793	3,150	13,302
c-inde•	0.79	0.75	0.74	AVR 0.79 MVRR 0.84	0.77	N/A	N/A	0.75	0.79	N/A for valve surgery
Mortality (%)	4.4	8.9	7.1	5.0	6.4	6.1	8.9	6.2	9.4	4.7
Number of factors	11	12	14	12	14[c]	14	8	11[c]	11[c]	17[d]
Factor (# of models)										
Age (10)	•	•	•	•	•	•	•	•	•	•
Renal failure (9)	•	•		•	•	•	•	•	•	•
Priority (8)			•	•	•	•	•	•	•	•
Re-operation (8)	•	•	•	•	•	•		•		•
Female sex (7)	•	•	•	•	•				•	•
Concomitant CABG (6)			•[b]	•	•	•	•	•		
Operation year (6)			•		•	•	•	•	•	
Endocarditis (5)	•	•	•			•				•
CVA/CVD (5)		•				•		•	•	•
Vascular disease (5)	•			•		•	•			•
NYHA (4)				•		•		•	•	
Diabetes (4)		•			•	•			•	

Table 4. Continued.

Model	Hannan [4]		Rankin [5]	Jin [6]	Ambler [7]	Gardner [9]		Nowicki [8]		Nashef [2]
MI (4)	•	•				•				•
LVF/EF (4)		•	•		•					•
Valve position(s) (3)	•	•	•							
Body size (3)				•	•			•		
CAD (3)			•			•			•	
CHF (3)			•					•	•	
Arrhythmia (3)				•	•			•		
Valve replacement (2)			•						•	
COPD (3)	•			•						•
Hemodynamic state (3)	•	•								•
Life dependence (2)						•	•			
Pulmonary HT (2)				•						•
ECAA (1)		•								
Hepatic failure (1)	•									
CCS (1)							•			
Mitral or Ao + mitral (1)					•					
Concomitant tricuspid (1)					•					
Extra comorbidities (1)			•							
Hypertension (1)					•					
Severe lesion (1)			•							
Surgery thoracic aorta (1)										•
Postinfarct septal rupt. (1)										•
UAP (1)										•

Table 4. Legend.
[a] Model inclusion criteria: sample size at least 1000 patients, published since 2004
[b] This model includes the factor concomitant surgery instead of concomitant CABG
[c] Logistic model includes more factors than additive model; logistic model factors are presented
[d] One variable ('Other than isolated CABG') is always scored 'yes' in valve surgery and therefore not included in this table

AVR = aortic valve replacement; CABG = coronary artery bypass grafting; CAD = coronary artery disease; CCS = Canadian Cardiovascular Society class; CHF = congestive heart failure; COPD = chronic obstructive pulmonary disease; CVA = cerebral vascular accident; CVD = cerebral vascular disease; ECAA = extensively calcified ascending aorta; EF = ejection fraction; LVF = left ventriculair function; MI = myocardial infarction; MVRR = mitral valve replacement or repair; NNE = Northern New England; NYHA = New York Heart Association functional class; PHS = Providence Health System; PVD = peripheral vascular disease; SCTS = Society of Cardiothoracic Surgeons of Great Britain and Ireland; STS = Society of Thoracic Surgeons; UAP = Unstable Angina Pectoris; VA = Veteran Affairs

ever considered to be of even more importance in predicting procedure success, but are absent from almost all current valve risk models, except the NY State model for valve+coronary surgery [14], which includes 'severely calcified ascending aorta' as a risk factor.

Early mortality

The dependent variable in both the Euroscore and NY State model is hospital mortality. One can debate whether this is an appropriate end point. It is much easier to collect patient information only during hospital stay and most other models also use hospital mortality, but on the other hand postoperative hospital stay varies widely between centers and in addition it is well-known that an increased post-cardiac surgery mortality risk does not disappear at patient discharge but continues for months after the procedure [15]. The new guidelines for reporting morbidity and mortality after cardiac valve interventions recommend that early mortality should be reported at 30, 60 and/or 90 days (submitted to AATS/EACTS/STS for final review). Therefore it is advisable that future valve models focus on a time-related endpoint independent of the location of the patient.

Even though all recently developed valve dedicated risk stratification models are designed to predict short term mortality only, they are probably able to predict one year (and possibly even longer) mortality as well. General cardiac surgery risk stratification models that were designed to predict short term

mortality have been demonstrated to possess this quality as well [16]. Future research is necessary to confirm this hypothesis. It would also be interesting to assess valve model performance in predicting other quality related outcomes. Again, risk stratification models that were built to predict short term mortality after general cardiac surgery have been successfully used to predict morbidity and intensive care unit stay [17] and total length of stay [4]. Xu et al. already published a model that successfully predicts intensive care unit stay after valve surgery (c-index 0.76; Hosmer-Lemeshow p=0.25) [18]. Taking into account these facts, models included in Table 4 are likely to possess this quality as well.

Concomitant coronary surgery

Of all recently published valve dedicated models (see Table 4), only the NY State model developers offer a separate model for isolated valve surgery and valve surgery with concomitant coronary surgery. In 2001, Edwards et al. also developed separate models for isolated valve surgery and valve+coronary surgery using the STS database [19].

The developers of all other recently published models have not employed this approach and simply included concomitant coronary surgery (or other type of surgery) as a risk factor in their valve models.

The reason for the development of a separate model for isolated valve surgery and a model for valve+coronary surgery according to Hannan et al. lies in the substantially different patient characteristics between both groups, an observation that is confirmed in the Rotterdam population.

For example, mitral valve repair is a lower risk procedure than aortic valve replacement in the isolated valve model while mitral valve repair combined with coronary surgery is a higher risk procedure than aortic valve replacement with concomitant coronary surgery using the valve+coronary surgery model, illustrating the difference between degenerated or myxomatous mitral valves in patients who need valve surgery only and the ischaemic nature of mitral valves in patients requiring concomitant coronary surgery.

Aortic and mitral surgery

Some model developers offer a separate model for aortic valve surgery and mitral surgery [8,9]. This way, more significant risk factors can be distilled from these separate, more homogeneous groups. Related to the pathophysiologic differences between these groups, are differences between early mortality rates.

As supported by the STS cohort, the models for mitral valve surgery produce higher predictive mortality rates [14]. However, models for aortic and mitral valve surgery have many similar risk factors [6].

Furthermore, in case of surgery on multiple valves, these models do not apply. A model that includes valve position(s!) as a model variable might therefore be preferred.

EuroSCORE, a vintage model?

Apart from both NY State models being valve dedicated models, their development was undertaken on a 2001-2003 dataset, while EuroSCORE was based on data collected in 1995. Since the introduction of the EuroSCORE model, increasing procedure safety and decreasing mortality rates probably contribute to the better performance of the NY State isolated valve model in our population. For this reason, an improvement could also have been expected for the NY State valve+coronary surgery model.

In this age of ubiquitous computing power, the simplicity of additive models might not justify their existence. And even though the EuroSCORE developers nowadays consider their additive EuroSCORE model to be inferior to the logistic model [10], we were not able to demonstrate a different performance between the logistic and additive EuroSCORE.

At the time of writing, the developers of the EuroSCORE just launched their plan to update the EuroSCORE in 2008 to reflect contemporary practice [20]. This new model will probably predict mortality after valve surgery better than the current EuroSCORE model due to improvement of the calibration of a new EuroSCORE model. However, it may only outperform the NY State model for isolated valve surgery when specific factors that are more predictive of hospital mortality after valve operations, are included.

Study limitations

This study is limited by the small sample size of the Rotterdam population, especially the valve+coronary surgery group. Even though the patient group sizes were sufficient to perform decent statistical analyses, a larger population will result in more reliable model performance results. Another study limitation is its single center approach. Hospital and surgeon related factors might have resulted in study results that are not applicable to other centers. However, the purpose of this study was to evaluate model performance and usability in our center.

Other models from Table 4 were not included in our comparison because many risk factors that are required for these models were not available. Retrospective collection of these extra model variables would contaminate the prospectively collected data we used for the current comparison.

Valve model implementation

Firstly, the importance of clinical judgement should never be underestimated and regarding individual preoperative mortality prediction, our search has provided us merely with a new tool to support clinical decisions regarding valve surgery. A very useful tool though, with a use not limited to preoperative risk prediction only.

The question remains whether a center needs a general model that has been based upon an enormous population and was validated throughout a particular region or a smaller model constructed by only using its own patients. Both types have advantages that cannot be combined into one model to fit all needs.

Perhaps centers should use two separate valve dedicated models; one model tailored to the center to predict actual and individual patient mortality and study the impact of different risk factors on outcome, and one model based upon a national or continental population, like the STS derived models, to provide a benchmark for hospital performance. This way, all possible purposes for valve dedicated risk stratification models are covered.

Noteworthy is that not only population selection is an important consideration when developing models for a specific purpose. For instance, the NY State risk stratification models were developed to serve as a tool to evaluate surgeon performance. By including the risk factor 'extensively calcified ascending aorta', best measured by palpation during surgery, in the valve+coronary surgery model, this model is less suitable for preoperative risk prediction purposes.

It is important to keep in mind that treatment for cardiac valve disease is still continuously improving and the danger of outdated models is considerable. In addition to periodic recalibration, future models will therefore need to be subject to improvements made by adjusting population sample size and re-evaluation of risk factor inclusion and weights. The biggest step forward in model use should consist of implementing risk stratification models into the clinical process, whereby we use an optimal model for a pre-specified patient population not only to monitor but also to improve outcome of our patients.

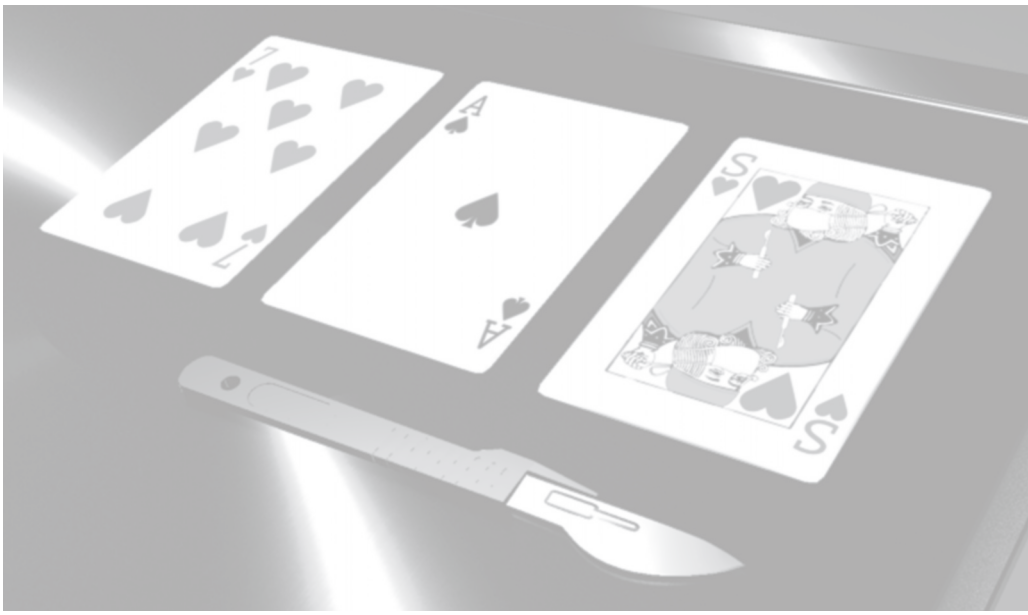
In conclusion, we demonstrated that in our clinical practice (even in a relatively small population) there is an added value using a valve dedicated model like the NY State model in the setting of preoperative risk prediction and assessment of the impact of individual risk factors. We encourage other centers to conduct comparable studies to verify our results and increase general applicability. Several valve dedicated models actually consist of two sub models [4,8,9]. Further exploration is needed to assess whether these separate models for isolated valve surgery and valve+coronary surgery, and/or aortic and mitral valve surgery are really needed or whether one common valve model for each clinical purpose will suffice. With the emergence of less invasive valve repair and replacement techniques, the implementation and further development of risk stratification models in valve surgery will be of increasing importance in optimizing patient treatment.

REFERENCES

1. Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models: a position article. Ann Thorac Surg. Nov 2004;78(5):1868-77.
2. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S and Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. Jul 1999;16(1):9-13.
3. Roques F, Nashef SA, Michel P, et al. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients. Eur J Cardiothorac Surg. Jun 1999;15(6):816-22; discussion 822-3.
4. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg. Mar 2007;83(3):921-9.
5. Rankin JS, Hammill BG, Ferguson TB, Jr., et al. Determinants of operative mortality in valvular heart surgery. J Thorac Cardiovasc Surg. Mar 2006;131(3):547-57.
6. Jin R, Grunkemeier GL and Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg. Aug 2005;80(2):471-9.
7. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE and Taylor KM. Generic, simple risk stratification model for heart valve surgery. Circulation. Jul 12 2005;112(2):224-31.
8. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg. Jun 2004;77(6):1966-77.
9. Gardner SC, Grunwald GK, Rumsfeld JS, et al. Comparison of short-term mortality risk factors for valve replacement versus coronary artery bypass graft surgery. Ann Thorac Surg. Feb 2004;77(2):549-56.
10. Roques F, Michel P, Goldstone AR and Nashef SA. The logistic EuroSCORE. Eur Heart J. May 2003;24(9):881-2.
11. Efron B and Tibshirani RJ. An introduction to the bootstrap. Boca Raton, FL: Chapman & Hall/CRC; 1993
12. Tjang YS, van Hees Y, Korfer R, Grobbee DE and van der Heijden GJ. Predictors of mortality after aortic valve replacement. Eur J Cardiothorac Surg. Sep 2007;32(3):469-474.
13. Hannan EL, Racz MJ, Jones RH, et al. Predictors of mortality for patients undergoing cardiac valve replacements in New York State. Ann Thorac Surg. Oct 2000;70(4):1212-8.
14. Nowicki ER. What is the future of mortality prediction models in heart valve surgery? Ann Thorac Surg. Aug 2005;80(2):396-8.
15. Blackstone EH and Kirklin JW. Death and other time-related events after valve replacement. Circulation. Oct 1985;72(4):753-67.
16. Nilsson J, Algotsson L, Hoglund P, Luhrs C and Brandt J. Comparison of 19 pre-operative risk stratification models in open-heart surgery. Eur Heart J. Apr 2006;27(7):867-74.
17. Nilsson J, Algotsson L, Hoglund P, Luhrs C and Brandt J. EuroSCORE predicts intensive care unit stay and costs of open heart surgery. Ann Thorac Surg. Nov 2004;78(5):1528-34.
18. Xu J, Ge Y, Hu S, Song Y, Sun H and Liu P. A simple predictive model of prolonged intensive care unit stay after surgery for acquired heart valve disease. J Heart Valve Dis. Mar 2007;16(2):109-15.
19. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve

replacement surgery. J Am Coll Cardiol. Mar 1 2001;37(3):885-92.
20. EuroSCORE 2008. Available at: http://www.euroscore.org/EuroSCORE2008.htm. Accessed September 12, 2007.

CHAPTER 4



Therapeutic decisions for patients with symptomatic severe aortic stenosis: room for improvement?

Van Geldorp MWA
Van Gameren M
Kappetein AP
Arabkhani B
De Groot-de Laat LE
Takkenberg JJM
Bogers AJJC

Eur J Cardiothorac Surg. 2009; 35: 953-7

ABSTRACT

Objective

Symptomatic severe aortic stenosis is an indication for aortic valve replacement. Some patients are denied intervention. This study provides insight into the proportion of conservatively treated patients and into the reasons why conservative treatment is chosen.

Methods

Of a patient cohort presenting with severe aortic stenosis between 2004 and 2007, medical records were retrospectively analyzed. Only symptomatic patients (n=179) were included. We studied their characteristics, treatment decisions, and survival.

Results

Mean age was 71 years, 50% were male. During follow-up (mean 17 months, 99% complete) 76 (42%) patients were scheduled for surgical treatment (63 conventional valve replacement, 10 transcatheter, 1 heart transplantation, 2 waiting list) versus 101 (56%) who received medical treatment. Reasons for medical treatment were: perceived high operative risk (34%), symptoms regarded mild (19%), stenosis perceived non-severe (14%), and patient preference (9%). In 5% the decision was pending at the time of the analysis and in 20% the reason was other/unclear. Mean age of the surgical group was 68 years versus 73 years for medically treated patients (p=0.004). Predicted mortality (EuroSCORE) was 7.8% versus 11.3% (p=0.006). During follow-up 12 patients died in the surgical group (no 30-day operative mortality), versus 28 in the medical group. Two-year survival was 90% versus 69%.

Conclusions

A large proportion (56%) of symptomatic patients does not undergo aortic valve replacement. Often operative risk is estimated (too) high or hemodynamic severity and symptomatic status are misclassified. Interdisciplinary team discussions between cardiologists and surgeons should be encouraged to optimize patient selection for surgery.

INTRODUCTION

The prevalence of aortic stenosis increases with age to up to 8% in the elderly [1]. Meanwhile the Western population increases to age during the last decades and this trend is expected to continue [2]. Therefore aortic stenosis constitutes a growing health burden.

While the treatment of asymptomatic patients with severe aortic stenosis remains debatable, both European and American guidelines on the management of valvular heart disease recommend that symptomatic patients have aortic valve replacement [3,4]. This recommendation is not only based on the survival advantage that can be expected after surgery but also on the improvement in functional class, even in elderly patients [3-5].

Recent literature suggests that a considerable proportion (33-60%) of patients with symptomatic severe aortic stenosis do not receive aortic valve replacement (AVR) [6-9]. We sought to confirm that many symptomatic patients remain unoperated and were interested in the reasons and the consequences of the decision to operate or not. The goal of our study therefore was to gain insight into decision making and survival in patients with severe symptomatic aortic stenosis.

METHODS

Study design and data collection

A retrospective search in the echocardiography database of our department revealed 115 patients with severe aortic stenosis. An additional 140 patients were recruited from the echocardiography laboratories in the outpatient cardiology clinics of 7 hospitals in the Rotterdam region. All echocardiograms were made between October 2004 and December 2007. Patients had at least one of the following inclusion criteria: aortic valve area <1.0 cm^2, maximum aortic jet velocity >4.0 m/s, peak aortic gradient >64 mmHg or mean aortic gradient >40 mmHg. To avoid missing low-output aortic stenosis, patients were also included if the ratio between the velocity time integral over the aortic valve and the left ventricular outflow tract was >4.0.

Information was gathered on medical history, cardiovascular risk factors, and symptomatic status at the time of the echocardiogram. Asymptomatic pa-

tients were excluded from the eventual analysis. For all symptomatic patients, anticipated operative risk was calculated using the logistic EuroSCORE risk model (http://www.euroscore.org).

Treatment strategies and their reasons were retrieved from notes in the patients' medical charts. Reasons for 'conservative/medical treatment' were classified in six main categories: (1) anticipated high operative risk (including advanced age or left ventricular dysfunction); (2) only mild symptoms; (3) stenosis non-severe; (4) patient preference; (5) decision not final yet; and (6) other, including 'reason unclear'.

The study protocol was approved by the institutional review board, patient informed consent was waived (MEC 06-066, MEC 08-022). The authors had full access to the data and take responsibility for its integrity. All authors have read and agree to the manuscript as written.

Study population

Of the 255 patients that were initially identified, 73 asymptomatic patients were excluded plus 3 patients, of whom symptomatic status could not be retrieved, leaving 179 symptomatic patients in the study cohort. Mean age was 71 years, 50% were male.

During follow-up (mean 17 months, median 13.6, range 0.1-40) 76 patients

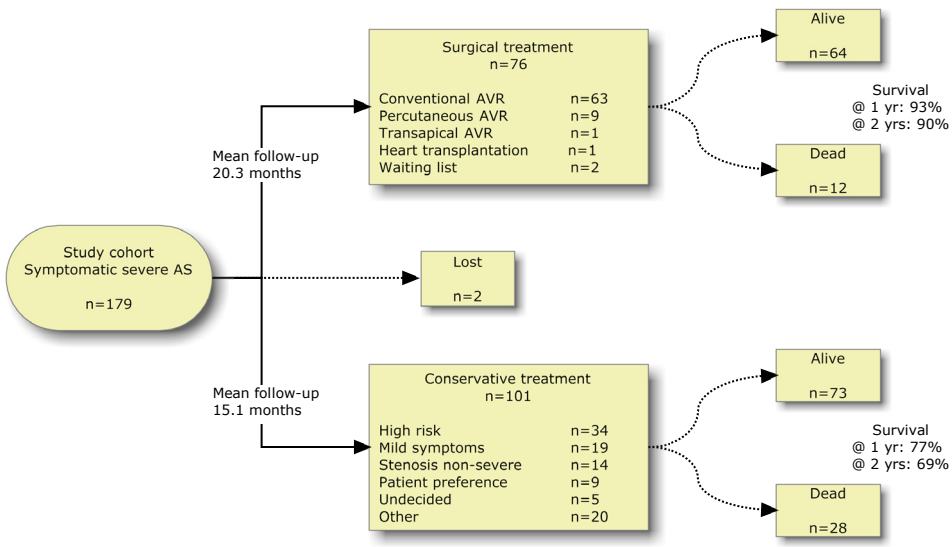


Figure 1. Flow chart of main results.

Table 1. Patient characteristics.

	AVR n= 76	**Conservative n= 101**
Age (mean ± SD in years)	67.9 ± 12.4	73.3 ± 12.3
Male (%)	49	51
Follow-up (mean ± SD in months)	20.3 ± 11.8	15.1 ± 11.5
Echocardiographic parameters (mean ± SD)		
Maximal transaortic velocity (m/s)	4.4 ± 0.8	4.0 ± 0.8
Peak gradient (in mmHg)	82 ± 32	66 ± 26
AV/LVOT VTI ratio	4.9 ± 1.8	5.0 ± 2.3
Aortic Valve Area (cm^2)	0.68 ± 0.24	0.71 ± 0.26
NYHA class (%)		
II	42.1	54.5
III	38.2	34.7
IV	13.2	8.9
Missing	6.6	2.0
Left Ventricular Function (%)		
Good/impaired (EF >50%)	56.6	57.4
Moderate (EF 30-50%)	38.2	30.7
Poor (EF <30%)	2.6	7.9
Missing	1.3	4.0
Logistic EuroSCORE (mean ± SD)	7.8 ± 7.9	11.3 ± 9.6

(42%) underwent AVR or were scheduled for surgery (Figure 1). There were 63 conventional aortic valve replacements, 9 percutaneous and 1 transapical valve implantations. Two patients were on a waiting list for AVR and one patient required a heart transplantation during follow-up. Medical treatment was given in 101 patients (56%). Two patients were lost to follow-up (99% completeness). Mean age of the surgical group was 68 years versus 73 years for the medically treated patients (p=0.004). Predicted operative mortality according to the logistic EuroSCORE was 7.8% versus 11.3% (p=0.009). More patient characteristics are given in Table 1.

Statistical analysis

Continuous data are presented as mean ±1 standard deviation, and median. Categorical data are presented as proportions.

Chi-square testing was used for comparison of categorical variables. Con-

tinuous variables were compared using the Student's t-test. A p value <0.05 was considered significant. Survival curves were estimated by the Kaplan-Meier method. Differences in survival were not statistically assessed.

Statistical analyses were performed with SPSS for Windows (release 15.0; SPSS Inc., Chicago, Illinois).

RESULTS

There was no 30-day mortality. During follow-up 12 patients died in the surgical group, versus 28 patients in the medical group. One- and 2-year survival was, respectively, 93% and 90% for the AVR group and for the conservative group 77% and 69% (Figure 1).

Reasons for choosing non-surgical treatment were: operative risk deemed 'too high' (34%), symptoms regarded as 'mild' (19%), stenosis regarded as 'non-severe' (14%), and patient preference (9%). In 5% the decision to operate was still under consideration by cardiologist and/or patient. In 20 patients (20%) the reason behind decision making could not be retrieved accurately. Of the latter 20 patients, 11 were in NYHA class II, 6 were in NYHA III, and 3 were in NYHA class IV.

Of the 34 patients in whom the reason not to operate was 'high risk', the mean age was 75.7 years and the mean EuroSCORE was 11.6%. Eight of them had a history of malignancy or active malignancy (six of these patients eventually died during follow-up). Eighteen patients had a EuroSCORE <10% and only 9 of the 34 patients in whom the operative risk was deemed too high had a EuroSCORE >15%.

DISCUSSION

Although treatment consensus seems to exist on symptomatic patients with severe aortic stenosis, it is not uncommon to diverge from these guidelines [6,8-10]. Advanced age and left ventricular dysfunction are known reasons to deny surgery in a symptomatic patient [6,11]. Instead of using patient characteristics to predict whether a patient gets AVR or not, our study was designed to investigate the decision making. Therefore it provides a different perspective: in our

cohort an overestimation of operative risk, underestimation of symptoms, and misclassification of hemodynamic severity are common causes why symptomatic patients are denied AVR. Furthermore, we found that survival of the conservative group is not as pessimistic as reported by others [12,13].

'Overestimation' of operative risk?

In a third of the patients who were treated conservatively, an anticipated high operative risk was the main reason not to go for AVR. This subgroup had a mean age of only 76 years, and only 9 of the 34 patients had a EuroSCORE >15%. Perhaps it is even more important that more than half (18 patients) had a relatively low operative risk with a EuroSCORE <10%.

From the literature it is known that remission of symptoms after starting medical treatment can be a reason to stay conservative and that patients who are treated conservatively are generally older and more often have impaired left ventricular function than surgically treated patients [6,7]. Yet, both remission of symptoms, advanced age, and depressed left ventricular function are debatable reasons not to operate on a symptomatic patient. Even elderly patients can be operated upon with acceptable morbidity and mortality, and can expect a considerable quality of life [5,11].

Note that 10 patients in the AVR group underwent a minimally invasive valve replacement. They were deemed not amenable for surgery. This indicates that even in a region with a tertiary center that uses new percutaneous and transapical techniques to replace the aortic valve, the majority of patients are treated conservatively.

Eight patients had either a malignancy in medical history or an active malignancy, risk factors which are not taken into account by the EuroSCORE. Another issue with risk models is that they do not score characteristics such as vitality or biological age. Furthermore there is a large variability between different risk models, and the one most commonly used (EuroSCORE) seems to overestimate the actual operative risk most [14]. Perhaps this adds to the large variance in treatment advice that exists among cardiologists which was already found by Bouma et al. [7].

Underestimation of symptoms?

Due to inactivity or gradual adjustment of daily activities to developing symptoms, patients with aortic stenosis often do not acknowledge the presence

of symptoms or attribute them to the ageing process. Exercise testing is recommended in asymptomatic patients with aortic stenosis in order to exclude symptoms with more certainty [15-17], and up to 37% of patients previously considered asymptomatic have limiting symptoms when they are tested [17]. According to the European Heart Survey exercise testing is highly underused [3,18]. This could lead to an underestimation of the proportion of symptomatic patients treated medically that was reported by others and in the current study [6,8-10].

In this study, the classical aortic stenosis symptoms such as dyspnea, syncope, or angina were documented for several patients but regarded as mild or non-debilitating. Having only mild symptoms does not exclude a patient from being an AVR candidate [3,4]. It is furthermore known that even if symptoms are recognized, the resulting functional disability is often underestimated by physicians [19]. Symptomatic patients with severe aortic stenosis from our cohort suffer from both physical and emotional impairment hampering normal daily activities (unpublished data). These are clear reasons to assess symptomatic status accurately, and to reconsider a conservative approach when symptoms are present.

Underestimation of hemodynamic severity

As much as 14% of the symptomatic patients who were denied surgery were not referred because the stenosis was classified non-severe by the treating cardiologist during the initial assessment. According to the guidelines they should however have been classified as severe [3,4]. Since only patients with a severe stenosis are recommended to have surgery, these misclassified patients are at increased risk of left ventricular deterioration and sudden death [20].

Even if the stenosis severity is only just below the severe threshold, it can be disputed that watchful waiting is the best treatment. Peak aortic gradient increases 10-15 mmHg/year and aortic valve area decreases 0.1-0.12 cm^2/year [21-23]. Given these progression rates, borderline patients will enter the severe category within a few months or at most a year later. Meanwhile left ventricular function will only get worse.

Survival in the conservative and in the surgical group

Survival in the medically treated group cannot easily be compared with the surgically treated group because the patients have quite different characteristics,

which could account for a large part of the difference in survival. It is therefore questionable if, and to what extent, the survival of the total study group would have improved if more patients had aortic valve replacement.

From the survival curve of the non-AVR group it can be seen that a decline in survival already occurs in the first year after the echocardiogram (Figure 2). Still, survival in the conservative group is not as bad as expected based on previous reports [12,13,20]. Perhaps improvement in medical treatment over the past years plays a role, but survival in the conservative group highly depends on referral strategy as well; if more high risk patients are operated upon, the patients with a really bad prognosis are left for conservative treatment, resulting in low survival in this category. Therefore the relatively good prognosis of our medically treated group could be a reflection of the conservative approach of the cardiologists in our region.

Because of its dependence on referral, natural history of aortic stenosis is very difficult to study. If one would like to gain a clear view on natural history, theoretically all eligible patients should be excluded from having AVR, or they

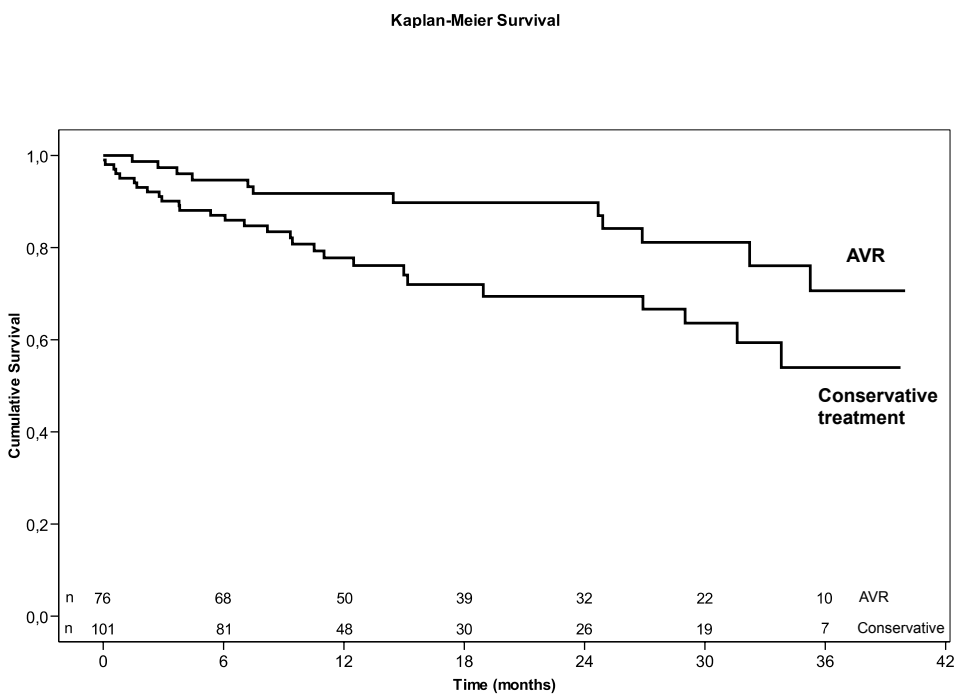


Figure 2. Kaplan-Meier survival for the conservatively treated group and the AVR group.

should be randomized to receive either surgical or conservative treatment. In practice this would be impossible and ethically incorrect.

Future prospects

Microsimulation methods can accurately estimate life expectancy for patients after AVR [24,25], but have yet to be developed for patients who are treated conservatively. Our department intends to develop these models, but this requires large datasets with extensive numbers of variables and some patient factors, such as vitality, will be difficult to grasp in a model.

CONCLUSION

A considerable proportion of patients with symptomatic severe aortic stenosis are not referred for surgery although theoretically they have an indication for aortic valve replacement. Often operative risk is estimated (too) high, and misclassification of both hemodynamic severity and symptomatic status occurs frequently.

Most patients who were treated conservatively were simply not *referred* to a surgical department. Referral to surgical departments should be encouraged in order to have more interdisciplinary team discussions between cardiologists and surgeons. Hopefully, this will result in better patient selection for surgery, possibly resulting in better survival of patients with severe symptomatic aortic stenosis.

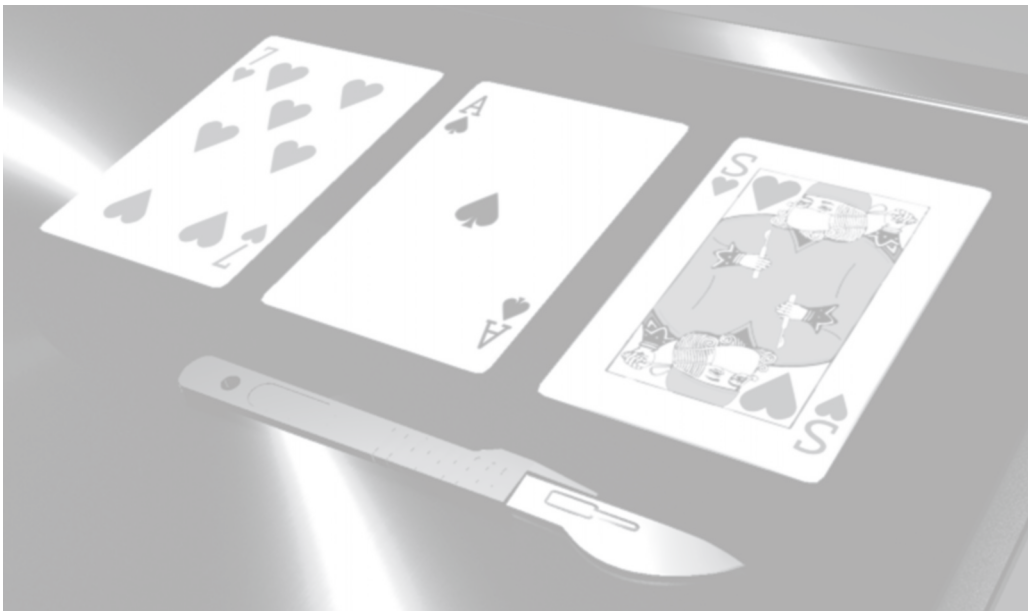
The authors would like to thank the patients, cardiologists, echo laboratory workers and secretaries of the following participating hospitals for their kind cooperation: Havenziekenhuis, Rotterdam; St. Franciscus Gasthuis, Rotterdam; IJsselland Ziekenhuis, Capelle aan den IJssel; Vlietland Ziekenhuis, Vlaardingen; Albert Schweitzer Ziekenhuis, Dordrecht; Medisch Centrum Rijnmond Zuid, Rotterdam; and Erasmus University Medical Center, Rotterdam.

REFERENCES

1. Lindroos M, Kupari M, Heikkila J, Tilvis R. Prevalence of aortic valve abnormalities in the elderly: an echocardiographic study of a random population sample. J Am Coll Cardiol 1993;21:1220-5.
2. Devision UNP. World Population Prospects: The 2006 Revision Population Database, 2006.
3. Bonow RO, Carabello BA, Kanu C, de Leon AC, Jr., Faxon DP, Freed MD, Gaasch WH, Lytle BW, Nishimura RA, O'Gara PT, O'Rourke RA, Otto CM, Shah PM, Shanewise JS, Smith SC, Jr., Jacobs AK, Adams CD, Anderson JL, Antman EM, Faxon DP, Fuster V, Halperin JL, Hiratzka LF, Hunt SA, Lytle BW, Nishimura R, Page RL, Riegel B. ACC/AHA 2006 guidelines for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (writing committee to revise the 1998 Guidelines for the Management of Patients With Valvular Heart Disease): developed in collaboration with the Society of Cardiovascular Anesthesiologists: endorsed by the Society for Cardiovascular Angiography and Interventions and the Society of Thoracic Surgeons. Circulation 2006;114:e84-231.
4. Vahanian A, Baumgartner H, Bax J, Butchart E, Dion R, Filippatos G, Flachskampf F, Hall R, Iung B, Kasprzak J, Nataf P, Tornos P, Torracca L, Wenink A. Guidelines on the management of valvular heart disease: The Task Force on the Management of Valvular Heart Disease of the European Society of Cardiology. Eur Heart J 2007;28:230-68.
5. Varadarajan P, Kapoor N, Bansal RC, Pai RG. Survival in elderly patients with severe aortic stenosis is dramatically improved by aortic valve replacement: Results from a cohort of 277 patients aged > or =80 years. Eur J Cardiothorac Surg 2006;30:722-7.
6. Iung B, Cachier A, Baron G, Messika-Zeitoun D, Delahaye F, Tornos P, Gohlke-Barwolf C, Boersma E, Ravaud P, Vahanian A. Decision-making in elderly patients with severe aortic stenosis: why are so many denied surgery? Eur Heart J 2005;26:2714-20.
7. Bouma BJ, van der Meulen JH, van den Brink RB, Arnold AE, Smidts A, Teunter LH, Lie KI, Tijssen JG. Variability in treatment advice for elderly patients with aortic stenosis: a nationwide survey in The Netherlands. Heart 2001;85:196-201.
8. Charlson E, Legedza AT, Hamel MB. Decision-making and outcomes in severe symptomatic aortic stenosis. J Heart Valve Dis 2006;15:312-21.
9. Bach DS, Cimino N, Deeb GM. Unoperated patients with severe aortic stenosis. J Am Coll Cardiol 2007;50:2018-9.
10. Bouma BJ, van Den Brink RB, van Der Meulen JH, Verheul HA, Cheriex EC, Hamer HP, Dekker E, Lie KI, Tijssen JG. To operate or not on elderly patients with aortic stenosis: the decision and its consequences. Heart 1999;82:143-8.
11. Bouma BJ, van den Brink RB, Zwinderman K, Cheriex EC, Hamer HH, Lie KI, Tijssen JG. Which elderly patients with severe aortic stenosis benefit from surgical treatment? An aid to clinical decision making. J Heart Valve Dis 2004;13:374-81.
12. Horstkotte D, Loogen F. The natural history of aortic valve stenosis. Eur Heart J 1988;9 Suppl E:57-64.
13. Pai RG, Kapoor N, Bansal RC, Varadarajan P. Malignant natural history of asymptomatic severe aortic stenosis: benefit of aortic valve replacement. Ann Thorac Surg 2006;82:2116-22.

14. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. J Thorac Cardiovasc Surg 2008;135:180-7.
15. Gibbons RJ, Balady GJ, Bricker JT, Chaitman BR, Fletcher GF, Froelicher VF, Mark DB, McCallister BD, Mooss AN, O'Reilly MG, Winters WL, Jr., Gibbons RJ, Antman EM, Alpert JS, Faxon DP, Fuster V, Gregoratos G, Hiratzka LF, Jacobs AK, Russell RO, Smith SC, Jr. ACC/AHA 2002 guideline update for exercise testing: summary article: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1997 Exercise Testing Guidelines). Circulation 2002;106:1883-92.
16. Bonow RO, Carabello B, de Leon AC, Edmunds LH, Jr., Fedderly BJ, Freed MD, Gaasch WH, McKay CR, Nishimura RA, O'Gara PT, O'Rourke RA, Rahimtoola SH, Ritchie JL, Cheitlin MD, Eagle KA, Gardner TJ, Garson A, Jr., Gibbons RJ, Russell RO, Ryan TJ, Smith SC, Jr. ACC/AHA Guidelines for the Management of Patients With Valvular Heart Disease. Executive Summary. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Management of Patients With Valvular Heart Disease). J Heart Valve Dis 1998;7:672-707.
17. Das P, Rimington H, Chambers J. Exercise testing to stratify risk in aortic stenosis. Eur Heart J 2005;26:1309-13.
18. Iung B, Baron G, Butchart EG, Delahaye F, Gohlke-Barwolf C, Levang OW, Tornos P, Vanoverschelde JL, Vermeer F, Boersma E, Ravaud P, Vahanian A. A prospective survey of patients with valvular heart disease in Europe: The Euro Heart Survey on Valvular Heart Disease. Eur Heart J 2003;24:1231-43.
19. Calkins DR, Rubenstein LV, Cleary PD, Davies AR, Jette AM, Fink A, Kosecoff J, Young RT, Brook RH, Delbanco TL. Failure of physicians to recognize functional disability in ambulatory patients. Ann Intern Med 1991;114:451-4.
20. Chizner MA, Pearle DL, deLeon AC, Jr. The natural history of aortic stenosis in adults. Am Heart J 1980;99:419-24.
21. Otto CM, Burwash IG, Legget ME, Munt BI, Fujioka M, Healy NL, Kraft CD, Miyake-Hull CY, Schwaegler RG. Prospective study of asymptomatic valvular aortic stenosis. Clinical, echocardiographic, and exercise predictors of outcome. Circulation 1997;95:2262-70.
22. Otto CM, Pearlman AS, Gardner CL. Hemodynamic progression of aortic stenosis in adults assessed by Doppler echocardiography. J Am Coll Cardiol 1989;13:545-50.
23. Faggiano P, Ghizzoni G, Sorgato A, Sabatini T, Simoncelli U, Gardini A, Rusconi C. Rate of progression of valvular aortic stenosis in adults. Am J Cardiol 1992;70:229-33.
24. Takkenberg JJ, Puvimanasinghe JP, Grunkemeier GL. Simulation models to predict outcome after aortic valve replacement. Ann Thorac Surg 2003;75:1372-6.
25. van Geldorp MW, Jamieson WR, Kappetein AP, Puvimanasinghe JP, Eijkemans MJ, Grunkemeier GL, Takkenberg JJ, Bogers AJ. Usefulness of microsimulation to translate valve performance into patient outcome: patient prognosis after aortic valve replacement with the Carpentier-Edwards supra-annular valve. J Thorac Cardiovasc Surg 2007;134:702-9.

CHAPTER 5



A Comparison of Patient Characteristics and 30-day Mortality Outcomes after Transcatheter Aortic Valve Implantation and Surgical Aortic Valve Replacement for the Treatment of Aortic Stenosis: a Two-Center Study

Piazza N, Van Gameren M, Jüni P, Wenaweser P, Carrel T, Onuma Y, Gahl B, Hellige G, Otten A, Kappetein AP, Takkenberg JJM, Van Domburg R, De Jaegere P, Serruys PW, Windecker S

EuroIntervention. 2009; 5: 580-8

ABSTRACT

Aims

It is unclear whether transcatheter aortic valve implantation (TAVI) addresses an unmet clinical need for those currently rejected for surgical aortic valve replacement (SAVR) and whether there is a subgroup of high-risk patients benefiting more from TAVI compared to SAVR. In this two-centre, prospective cohort study, we compared baseline characteristics and 30-day mortality between TAVI and SAVR in consecutive patients undergoing invasive treatment for aortic stenosis.

Methods and results

We pre-specified different adjustment methods to examine the effect of TAVI as compared with SAVR on overall 30-day mortality: crude univariable logistic regression analysis, multivariable analysis adjusted for baseline characteristics, analysis adjusted for propensity scores, propensity score matched analysis, and weighted analysis using the inverse probability of treatment (IPT) as weights. A total of 1,122 patients were included in the study: 114 undergoing TAVI and 1,008 patients undergoing SAVR. The crude mortality rate was greater in the TAVI group (9.6% vs. 2.3%) yielding an odds ratio [OR] of 4.57 (95%-CI 2.17-9.65). Compared to patients undergoing SAVR, patients with TAVI were older, more likely to be in NYHA class III and IV, and had a considerably higher logistic EuroSCORE and more comorbid conditions. Adjusted OR depended on the method used to control for confounding and ranged from 0.60 (0.11-3.36) to 7.57 (0.91-63.0). We examined the distribution of propensity scores and found scores to overlap sufficiently only in a narrow range. In patients with sufficient overlap of propensity scores, adjusted OR ranged from 0.35 (0.04-2.72) to 3.17 (0.31 to 31.9). In patients with insufficient overlap, we consistently found increased odds of death associated with TAVI compared with SAVR irrespective of the method used to control confounding, with adjusted OR ranging from 5.88 (0.67-51.8) to 25.7 (0.88-750). Approximately one third of patients undergoing TAVI were found to be potentially eligible for a randomised comparison of TAVI versus SAVR.

Conclusions

Both measured and unmeasured confounding limit the conclusions that can be drawn from observational comparisons of TAVI versus SAVR. Our study indicates that TAVI could be associated with either substantial benefits or harms. Randomised comparisons of TAVI versus SAVR are warranted.

INTRODUCTION

Transcatheter aortic valve implantation (TAVI) is currently restricted to patients with aortic stenosis in whom surgical aortic valve replacement (SAVR) would be associated with a high or prohibitive risk of morbidity and/or mortality. The aim of TAVI is to provide a minimally invasive treatment that is at least as effective and associated with less morbidity and mortality compared to conventional valve surgery in high-risk patients. It is unclear whether TAVI addresses an unmet clinical need for those currently rejected for SAVR and whether there is a subgroup of high-risk SAVR patients who can benefit more from TAVI than SAVR.

Contemporary studies indicate that the 30-day mortality rate following TAVI is 8% to 12% [1-3]. Following SAVR, the 30-day mortality rate in high-risk patient subsets was reported to be between 4.6% to 13.5% for octogenarians [4-11] and 6% to 33% for patients with left ventricular dysfunction [12,13]. To our knowledge, direct comparisons between TAVI and SAVR are not yet available. In this two-centre, prospective cohort study, we set out to compare the characteristics at baseline and 30-day mortality rates between TAVI and SAVR in consecutive patients undergoing invasive treatment for aortic stenosis.

METHODS

Patients and methods

Between January 1st, 2006 and December 31st, 2008, we prospectively enrolled consecutive patients with aortic stenosis who underwent invasive treatment for aortic stenosis at the Erasmus Medical Center, Rotterdam, The Netherlands and Bern University Hospital, Bern, Switzerland. The study complies with the Declaration of Helsinki and was approved by the local Research Ethics Committees. All patients provided written informed consent.

Patients and interventions

Patients were included if they underwent TAVI or SAVR for the treatment of aortic stenosis at one of the two institutions. Patients undergoing invasive treatment who had a primary diagnosis of aortic regurgitation, multiple valve interventions, or concomitant aortic root reconstruction were excluded. Contraindications for TAVI or SAVR typically included sepsis, bleeding diathesis or coagulopathy, any

Table 1. Baseline characteristics.

	TAVI (n=114)	SAVR (n=1008)	Difference (95% CI)*	P
Age, y (SD)	82.8 (5.5)	69.9 (11.4)	12.9 (10.8 to 15.1)	< 0.001
Female, n (%)	64 (56.1%)	408 (41.5%)	15.7% (6.1 to 25.3%)	0.001
Logistic EuroSCORE, % (SD)	20.1 (13.4)	9.1 (10.2)	11.0 (9.0 to 13.1)	< 0.001
NYHA class, n (%)				<0.001**
I	1 (0.9%)	147 (14.6%)		
II	14 (12.3%)	403 (40.0%)		
III	78 (68.4%)	356 (35.3%)		
IV	21 (18.4%)	102 (10.1%)		
Diabetes mellitus, n (%)	26 (22.8%)	243 (24.1%)	–1.3% (–9.4 to 6.8%)	0.76
Hypertension, n (%)	72 (63.1%)	631 (62.6%)	0.6% (–8.8 to 9.9%)	0.91
Coronary artery disease, n (%)	64 (56.1%)	512 (50.8%)	5.3% (–4.3 to 15.0%)	0.28
Previous coronary bypass surgery, n (%)	28 (24.6%)	42 (4.2%)	20.4% (12.4 to 28.4%)	<0.001
Left ventricular ejection fraction, n (%)				<0.001**
> 50%	67 (58.8%)	834 (82.7%)		
30-50%	40 (35.1%)	135 (13.4%)		
< 30%	7 (6.1%)	39 (3.9%)		
Atrial fibrillation, n (%)	22 (19.3%)	90 (8.9%)	10.4% (2.9 to 17.8%)	<0.001
Cereberovascular disease, n (%)	20 (17.5%)	50 (5.0)	12.6% (5.5 to 19.7%)	<0.001
Peripheral vascular disease, n (%)	21 (18.4%)	47 (4.7%)	13.8% (6.5 to 21.0%)	<0.001
COPD, n (%)	24 (21.1%)	134 (13.3%)	7.8% (0.0 to 15.5%)	0.024
Pulmonary hypertension, n (%)	34 (29.8%)	86 (8.5%)	21.3% (12.7 to 29.9%)	< 0.001
Creatinine, umol/L (SD)	114 (92.9)	98.6 (61.5)	15.7 (3.0 to 28.3)	0.016
Creatinine above 200 umol/L, n (%)	7 (6.1%)	31 (3.1%)	3.1% (–1.5 to 7.6%)	0.086
MI within 90 days of procedure, n (%)	4 (3.5%)	34 (3.4%)	0.1% (–3.4 to 3.7%)	0.94

COPD = chronic obstructive pulmonary disease; MI = myocardial infarction; NYHA = New York Heart Association; SAVR = Surgical aortic valve replacement; TAVI = Transcatheter aortic valve implantation.

* Differences relate to percentages for dichotmous variables and means for continuous vari ables.

** Test for trend.

condition considered a contraindication to extracorporeal assistance, or estimated life expectancy of less than one year.

TAVI was performed using the third generation CoreValve ReValving System (Medtronic, Minneapolis, MN, USA). For patients to undergo TAVI, an interventional cardiologist and cardiac surgeon had to agree that conventional open-heart surgery would be associated with excessive morbidity and mortality. Details of the device, and technical aspects of the procedure have been previously published [14]. The procedure was performed with the patient under general anaesthesia or with local anaesthesia and mild sedation depending on patient characteristics. Vascular access was obtained percutaneously using a "pre-closing" device (10 Fr Prostar XL) or by surgical cut-down [14]. Coronary revascularisation was performed if deemed necessary prior to or during the index procedure using percutaneous coronary intervention (PCI) [15].

The choice of implant used and the technique for SAVR was left to the discretion of the treating cardiac surgeon. The procedure was performed with the patient under general anaesthesia using cardiopulmonary bypass. Coronary revascularisation was performed if deemed necessary during the index procedure using coronary artery bypass graft surgery (CABG).

Baseline characteristics

The selection of baseline characteristics was based on expert opinion and a review of the literature. We considered patient age (years), sex, logistic EuroSCORE [16], New York Heart Association class, ejection fraction (>50%, 30-50%, or <30%), creatinine concentration (umol/L), a history of coronary artery disease, myocardial infarction in the previous 90 days, CABG, atrial fibrillation, peripheral vascular disease, cerebrovascular disease, pulmonary hypertension, and COPD.

Procedural characteristics

We recorded whether the procedure was an emergency measure, defined as a procedure carried out on referral before the beginning of the next working day [16] and whether there was a concomitant coronary revascularisation using PCI in patients undergoing TAVI and CABG in patients undergoing SAVR. Primary outcome and definitions The primary outcome was all-cause mortality within 30 days [17]. Patients were actively followed-up until 30 days after the index procedure and their vital status was confirmed by outpatient clinical visit, telephone visit, review of medical records or through civil registries.

Statistical methods

We compared baseline characteristics between patients who had undergone TAVI and SAVR using a χ^2 test for categorical variables and an unpaired t test for continuous variables, and used uni- and multivariable logistic regression models to determine the association of baseline characteristics with 30-day mortality. The propensity scores of SAVR patients were estimated using a probit model with all baseline characteristics as described above as independent variables. The propensity score is the probability that a patient would have been treated with SAVR given that patient's observed baseline characteristics. Patients with the same propensity score must have the same distribution of baseline characteristics. To determine whether this assumption of balanced baseline characteristics was satisfied, we used standard algorithms implemented in the statistical package. Then, we calculated the cstatistic (the area under the receiver operating characteristics curve for the ability of propensity scores to predict actual treatment) as an estimate of the adequacy of the model.

To ensure the quality of the matching using propensity scores, we applied the common support assumption. The common support assumption indicates that patients with the same propensity score have a non-zero probability of receiving both, TAVI or SAVR [18]. If probability densities are too low in one of the groups, there is insufficient overlap of propensity scores and the assumption of common support is unlikely to be satisfied. To address this, we estimated the probability densities of propensity scores of patients in the TAVI and SAVR group using Epanechnikov kernel probability density estimates based on propensity score increments of 0.025 [19]. We pre-specified a probability density of >0.5 to ensure sufficient overlap of propensity scores [20]. Below this pre-specified probability density (established at a propensity score <0.675), patients who had undergone TAVI were considered unlikely to be comparable with patients of the same propensity score who had undergone SAVR.

We used logistic regression models to compare 30-day mortality between groups. We pre-specified the following seven approaches for analysis: (1) crude univariable analysis; (2) multivariable analysis adjusted for all baseline characteristics described above; (3) bivariable analysis adjusted for the propensity score as a linear term [21]; (4) multivariable analysis adjusted for the propensity score and for all baseline characteristics; (5) univariable analysis after caliper matching on the propensity score in a range of ±0.05 [22]; (6) weighted univariable analysis using the inverse probability of treatment (IPT) as weights [23,24];

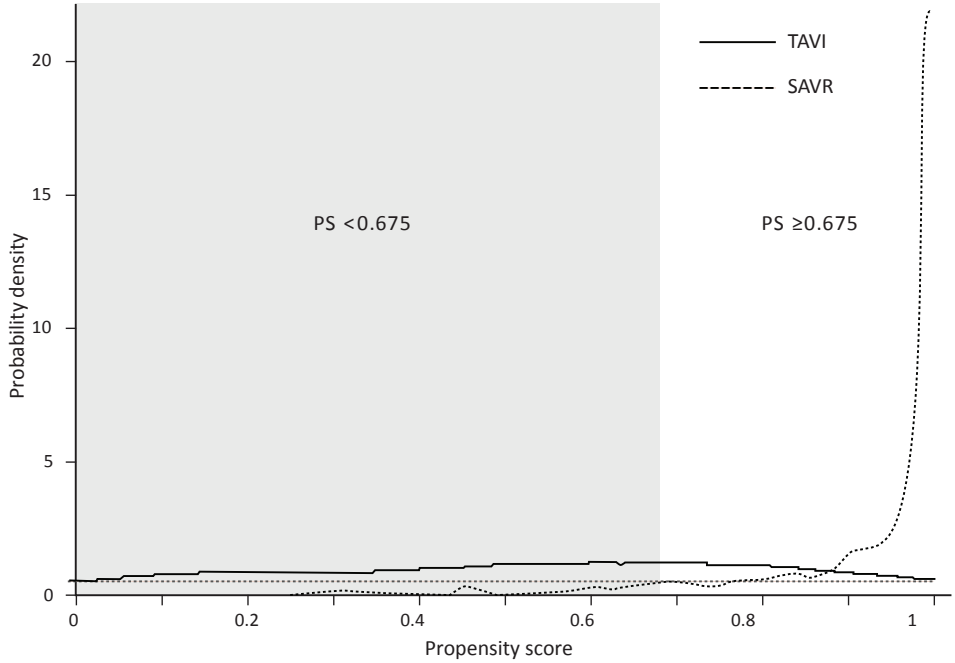


Figure 1. Probability density function of the propensity score for 114 patients undergoing TAVI (black solid line) and 1,008 patients undergoing SAVR (black dotted line). The grey dotted line represents the pre-specified probability density of >0.5; propensity score matching above this probability density ensures sufficient overlap of propensity scores. The white area indicates the range of propensity scores with sufficient overlap (propensity score ≥0.675) and identifies patients potentially eligible for a randomised comparison of TAVI and SAVR. The grey area indicates the range of propensity scores with insufficient overlap and identifies patients likely to be ineligible for a randomised comparison.

(7) IPT weighted multivariable analysis adjusted for all baseline characteristics. The IPT weighted analyses used the inverse of the propensity score as weights in SAVR patients and the inverse of 1 minus the propensity score in TAVI patients. All analyses were based on logistic regression models and were performed in the overall dataset and stratified according to presence or absence of sufficient overlap of propensity scores (propensity score <0.675 versus ≥0.675). To determine whether there was an interaction between estimated odds ratios of death and overlap of propensity scores, we performed a formal interaction test based on z-scores. Then, we performed a sensitivity analysis (not pre-specified), additionally adjusting IPT weighted analyses restricted to patients with propensity scores

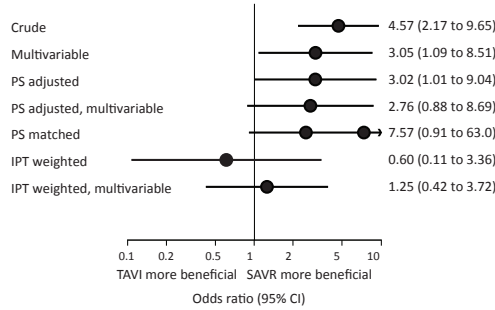


Figure 2. Impact of different approaches used to control for confounding on mortality estimates. The propensity score matched analysis was based on 162 patients (81 pairs), all other analyses were based on 1,122 patients.
PS = propensity score; IPT = inverse probability of treatment.

≥0.675 for the procedural characteristics described above. Finally, we considered patients undergoing TAVI with propensity scores ≥0.675 to be potentially eligible for a randomised comparison of TAVI and SAVR and compared pre-treatment characteristics of these patients with characteristics of patients undergoing TAVI with propensity scores <0.675 who were deemed ineligible for a randomised comparison with SAVR. All p-values are two-sided. Statistical analyses were performed using Stata Version 10 (Stata Corporation, College Station, TX, USA).

RESULTS

Between January 2006 and December 2008, 1,633 consecutive patients underwent invasive treatment for aortic stenosis using TAVI or SAVR. After exclusion of 508 patients because of a primary diagnosis of aortic regurgitation, multiple valve surgery, or concomitant aortic root surgery, 1,122 patients (1,122 procedures) were included in the study: 114 undergoing TAVI (52 Bern and 62 Rotterdam) and 1,008 patients undergoing SAVR (645 Bern and 363 Rotterdam). No patient was lost to follow-up at 30 days after the index procedure.

Baseline characteristics of patients

Baseline characteristics of the TAVI and SAVR patients are summarised in Table 1. Compared to patients undergoing SAVR, patients with TAVI were clearly older, more often female, more likely to be in NYHA class III and IV, had a considerably

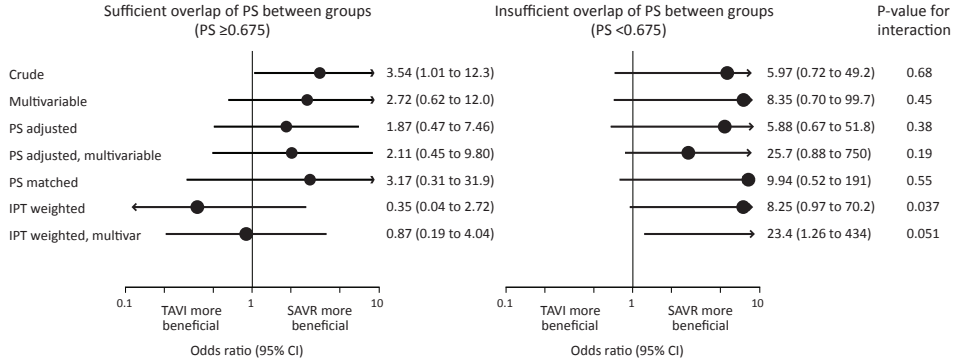


Figure 3. Analyses stratified according to overlap of propensity scores (propensity scores ≥0.675 versus <0.675). Among patients with sufficient overlap (left), the propensity score matched analysis was based on 78 patients (39 pairs), all other analyses were based on 996 patients. Among patients with insufficient overlap (right), the propensity score matched analysis was based on 84 patients (42 pairs), all other analyses were based on 126 patients.

higher logistic EuroSCORE and more comorbid conditions, including previous coronary artery bypass surgery, left ventricular dysfunction, atrial fibrillation, cerebral and peripheral vascular disease, chronic obstructive pulmonary disease, and pulmonary hypertension.

Procedural characteristics

The procedure was performed as an emergency measure in two TAVI patients (1.8%) and 43 SAVR patients (4.3%) (difference –2.4%, 95% confidence interval [CI] –5.1 to 0.3%). Concomitant revascularisation was performed in 10 TAVI patients (8.8%) and 436 SAVR patients (43.5%) (difference -34.7%, 95% CI –40.7 to –28.7%). An additional 10 TAVI patients (8.8%) had coronary revascularisation within four weeks prior to valve implantation as a staged procedure.

Association of baseline characteristics with 30-day mortality

Table 2 presents the crude and adjusted odds ratios for 30-day mortality. Univariable analysis revealed that age, logistic EuroSCORE, previous coronary artery bypass, atrial fibrillation, renal failure (creatinine >200 µmol/L) and myocardial infarction within 90 days of the procedure were associated with 30-day mortality at $p < 0.05$; a trend was noted for diabetes mellitus ($p = 0.051$). In multivariable

Table 2. Crude and adjusted odds ratios (OR) for 30-day mortality.

	Crude OR (95% CI)	P	Adjusted OR (95% CI)	P
Age (per 10 years)	1.46 (1.01 to 2.11)	0.043	1.06 (0.71 to 1.59)	0.78
Female	1.39 (0.70 to 2.75) 0.34 1.28 (0.59 to 2.75) 0.53			
Logistic EuroSCORE (per 10%)	1.61 (1.35 to 1.91) <0.001 1.65 (1.20 to 2.25) 0.002			
NYHA class		0.005		0.32
I	Reference category		Reference category	
II	1.25 (0.26 to 6.07)		1.11 (0.22 to 5.55)	
III	3.16 (0.72 to 13.78)		1.64 (0.35 to 7.67)	
IV	4.41 (0.90 to 21.61)		1.74 (0.32 to 9.58)	
Diabetes mellitus	2.01 (0.99 to 4.07)	0.052	1.97 (0.90 to 4.30)	0.089
Hypertension	1.25 (0.61 to 2.60)	0.54	0.95 (0.43 to 2.12)	0.91
Coronary artery disease	1.37 (0.68 to 2.73)	0.38	0.82 (0.36 to 1.88)	0.65
Previous coronary surgery	1.35 (0.39 to 4.64)	0.64	1.35 (0.39 to 4.64)	0.64
Left ventricular ejection fraction		0.10		0.44
> 50%	Reference category		Reference category	
30-50%	1.03 (0.39 to 2.73)		0.41 (0.14 to 1.23)	
< 30%	3.34 (1.11 to 10.02)		1.04 (0.23 to 4.72)	
Atrial fibrillation	4.03 (1.87 to 8.66)	<0.001	4.75 (2.10 to 11.26)	<0.001
Cereberovascular disease	0.94 (0.22 to 3.99)	0.93	0.43 (0.09 to 2.15)	0.30
Peripheral vascular disease	1.94 (0.41 to 5.42)	0.55	1.49 (0.41 to 5.42)	0.55
COPD	1.61 (0.69 to 3.76)	0.27	1.22 (0.49 to 3.07)	0.67
Pulmonary hypertension	1.83 (0.74 to 4.52)	0.19	0.64 (0.22 to 1.84)	0.40
Creatinine above 200 umol/L	4.13 (1.38 to 12.39)	0.011	1.87 (0.49 to 7.10)	0.36
MI within 90 days of procedure,	4.13 (1.38 to 12.39)	0.011	2.10 (0.54 to 8.16)	0.29

COPD = chronic obstructive pulmonary disease; MI = myocardial infarction; NYHA = New York Heart Association

analysis, only logistic EuroSCORE and atrial fibrillation were associated with mortality at p<0.05, and a trend remained for diabetes mellitus (p=0.089).

Propensity scores

The model used to estimate propensity scores yielded a c-statistic of 0.93. Figure 1 shows the distribution of propensity scores. The mean propensity score to receive SAVR was 0.53±0.28 for patients undergoing TAVI compared with 0.94±0.12 for patients actually undergoing SAVR. The distribution was symmetrical around 0.5 for the TAVI group, but clearly shifted towards 1 for the SAVR group. Based on the common support assumption, the area of sufficient overlap of propensity scores was small (Figure 1, white area to the right): in patients undergoing SAVR, the probability density fell definitely below the pre-specified value of 0.5 at a propensity score of 0.675. In patients undergoing TAVI, 39 patients (34.2%) were above and 75 (65.8%) were below the cut-off of 0.675. In patients undergoing SAVR, 957 patients (94.9%) were above and 51 (5.1%) were below the cut-off. Using simply the propensity scores with a calliper matching range of 0.05, 81 patients undergoing TAVI (71.1%) could each be matched to a patient in the SAVR group. Of these 81 pairs of patients, 39 were above and 42 pairs were below the cut-off of 0.675.

Impact of different approaches used for control of confounding on mortality estimates in overall dataset

Figure 2 presents the results of the stepwise procedure to control for confounding in the analysis of the overall dataset. Within 30 days of the index procedure, 11 patients (9.6%) who had undergone TAVI and 23 patients (2.3%) who had undergone SAVR died (crude odds ratio [OR] 4.57, 95% CI, 2.17 to 9.65, Figure 2). The estimated odds ratios decreased only slightly after multivariable adjustment and propensity score adjustment. This suggested an approximately threefold increase in the odds of dying with TAVI as compared with SAVR. The estimated odds ratio increased in the analysis of propensity score matched patients to 7.57 (95% CI 0.91 to 63.0). Conversely, when the IPT weights were used, we found the estimated odds ratios decreased to 0.60 in the univariable analysis (95% CI 0.11 to 3.36) and to 1.25 in the multivariable analysis (95% CI 0.42 to 3.72).

Analyses stratified according to overlap of propensity scores

Figure 3 shows the results of the analyses stratified according to overlap of propensity scores. Among patients with propensity scores ≥ 0.675 and hence sufficient overlap of propensity scores, three out of 39 patients (8.0%) who had undergone TAVI and 22 out of 957 patients (2.8%) who had undergone SAVR died

Table 3. Table 3. Comparison of patients with transcatheter aortic valve implantation potentially eligible for a randomised trial (propensity score ≥0.675) with those likely to be ineligible (propensity score <0.675).

	Eligible (n=39)	Ineligible (n=75)	Difference (95% CI)*	P
Age, y (SD)	79.3 (5.8)	84.7 (4.4)	–5.3 (–7.3 to –3.43)	<0.001
Female, n (%)	23 (60.0%)	41 (54.7%)	4.3% (–14.8 to 23.4%)	0.66
Logistic EuroSCORE, % (SD)	15.7 (10.4)	22.4 (14.25)	–6.7 (–11.9 to –1.6)	0.010
NYHA class, n (%)				0.002
I	0 (0%)	1 (1.3%)		
II	11 (28.2%)	3 (4.0%)		
III	22 (56.4%)	56 (74.7%)		
IV	6 (15.4%)	15 (20.0%)		
Diabetes mellitus, n (%)	7 (17.9%)	19 (25.3%)	–7.4% (–22.9 to 8.2%)	0.37
Hypertension, n (%)	25 (64.1%)	47 (62.7%)	1.4% (–17.2 to 20.1%)	0.88
Coronary artery disease, n (%)	19 (48.7%)	45 (60.0%)	–11.3% (–30.5 to 7.9%)	0.25
Previous coronary bypass surgery, n (%)	2 (5.1%)	26 (34.7%)	–29.5% (–42.3 to –16.7%)	<0.001
Left ventricular ejection fraction, n (%)				0.037
> 50%	29 (74.4%)	38 (50.7%)		
30-50%	8 (20.5%)	32 (42.7%)		
< 30%	2 (5.1%)	5 (6.7%)		
Atrial fibrillation, n (%)	3 (7.7%)	19 (25.3%)	–17.6% (–30.6 to –4.7%)	0.024
Cereberovascular disease, n (%)	3 (7.7%)	17 (22.7%)	–15.0% (–27.6 to –2.3%)	0.046
Peripheral vascular disease, n (%)	3 (7.7%)	18 (24.0%)	–16.3% (–29.1 to –3.5%)	0.033
COPD, n (%)	6 (15.4%)	18 (24.0%)	–8.6% (–23.5 to 6.3%)	0.28
Pulmonary hypertension, n (%)	6 (15.4%)	28 (37.3%)	–21.9% (–37.7 to –6.2%)	0.015
Creatinine, umol/L (SD)	97 (42)	123 (110)	–27.0 (–63.1 to –9.2)	0.14
Creatinine above 200 umol/L, n (%)	1 (2.6%)	6 (8.0%)	–5.4% (–13.3 to 2.5%)	0.25
MI within 90 days of procedure, n (%)	0 (0%)	4 (5.3%)	–5.3% (–10.4 to –0.2%)	0.14

COPD = chronic obstructive pulmonary disease; MI = myocardial infarction; NYHA = New York Heart Association.

* Differences relate to percentages for dichotmous variables and means for continuous vari ables.

(crude OR 3.54, 95% CI 1.01 to 12.3, Figure 2, top left). The estimated odds ratios decreased after multivariable adjustment (OR 2.72), propensity score adjustment (1.87), but not after propensity score matching (3.17). The 95% CIs, however, largely overlapped the line of null effect at 1. When using the IPT weights, we found the estimated odds ratios to decrease below 1 in both the univariable (OR 0.35, 95% CI 0.04 to 2.72) and multivariable analyses (OR 0.87, 95% CI 0.19 to 4.04). We found similar results after adjusting the IPT weighted analyses for procedural characteristics (emergency measure and concomitant revascularisation); the estimated odds ratios in patients with propensity scores ≥ 0.675 was 0.73 (95% CI 0.07 to 7.70) after the weighted univariable analysis and 1.63 (95% CI 0.22 to 12.1) after the weighted multivariable analysis. Among patients with insufficient overlap of propensity scores (propensity scores <0.675), eight out of 75 patients (10.7%) who had undergone TAVI and one out of 51 patients (2.0%) who had undergone SAVR died (crude OR 5.97, 95% CI 0.72 to 49.2, Figure 2, top right). The estimated odds ratios were between 5.88 and 25.7 depending on the approach used to control for confounding. Tests of interaction between estimated odds ratios and overlap were positive at the conventional significance level for univariable IPT weighted analyses (p=0.037) and of borderline significance for multivariable IPT weighted analyses (p=0.051, Figure 3).

Patients potentially eligible for a randomised trial

Table 3 presents the baseline characteristics of TAVI patients who had propensity scores ≥0.675 and were considered to be potentially eligible for a randomised comparison of TAVI and SAVR compared with the characteristics of TAVI patients with propensity scores <0.675 deemed to be ineligible. Patients eligible for a randomised trial were younger, had less severe symptoms as measured by the NYHA class, had lower logistic EuroSCORES, and were less likely to report a history of previous coronary artery bypass, atrial fibrillation, stroke, peripheral vascular disease, or pulmonary hypertension.

DISCUSSION

In this prospective cohort study we found that both measured and unmeasured confounding complicated the observational comparison of 30-day mortality of TAVI versus SAVR. Compared to SAVR patients, TAVI patients were clearly older,

more often female, more likely to be in NYHA class III and IV, had a considerably higher logistic EuroSCORE and had more comorbid conditions, including previous coronary artery bypass surgery, left ventricular dysfunction, atrial fibrillation, cerebral and peripheral vascular disease, chronic obstructive pulmonary disease, and pulmonary hypertension. When naively analysing the overall dataset of 1,122 consecutive patients included in our study, we found a 4.5-fold increase in the odds of death associated with TAVI as compared with SAVR. The odds ratio, however, ranged from 0.6 to 7.57 depending on the method used to control for confounding. These results indicate that TAVI may be associated with potential benefits or harm. Given the fact that TAVI was performed in either high risk or inoperable patients, analysis of the overall dataset may suffer from confounding by indication. For the inoperable patients, the relevant comparison would have been medically managed patients. In order to minimise confounding by indication, we identified patients from the TAVI and SAVR group who, in addition to having similar propensity scores, would also be appropriate for either type of treatment.

Below a propensity score cut-off of 0.675, it was deemed unlikely that two patients with the same propensity score, one who actually received TAVI, the other SAVR, had indeed identical probabilities to undergo SAVR. Two-thirds of patients undergoing TAVI, but only 5% of those undergoing SAVR had scores below this cut-off. Restricting the analysis to patients with sufficient overlap of propensity scores ≥ 0.675, we found little evidence for an excess mortality in TAVI patients. On the other hand, the estimated odds ratios greatly increased between 5.9 and 25.7 in those patients with insufficient overlap of propensity scores. These results suggest that unmeasured confounding factors complicated the observational comparisons of TAVI versus SAVR, albeit more in patients with insufficient than sufficient overlap of propensity scores. Examples of unmeasured confounding variables that could be associated with both the selection of treatment and prognosis include porcelain aorta, history of mediastinal radiation, or frailty. In addition, we were unable to record the clinical judgment of the treating physicians, which may be a relevant proxy for prognosis. In light of the high probability of residual confounding complicating any observational study in this field, randomised controlled trials comparing TAVI and SAVR are clearly warranted.

Through our examination of propensity score distributions, we characterised a subgroup of TAVI patients likely to be eligible for a randomised trial (Table

3). In our institutions, the following criteria are used to guide the selection of patients for TAVI: age ≥75 years, a logistic EuroSCORE ≥15%, and/or age ≥65 years associated with severe limiting comorbid conditions. Our study indicates that approximately one-third of patients selected for TAVI (based on these criteria) would be eligible for a randomised comparison of TAVI and SAVR. These patients were younger, had lower logistic EuroScores, and were less likely to report a history of previous coronary artery bypass, atrial fibrillation, stroke, peripheral vascular disease, or pulmonary hypertension than potentially ineligible TAVI patients. Based on the exploratory nature of our study, we cannot provide a firm basis for the identification of patients eligible for such a trial. The small number of patients undergoing TAVI and the small number of accumulated deaths was a limitation of this study. The multivariable analyses was difficult particularly when numbers were cut down through stratification of analyses according to overlap of propensity scores. The stratified multivariable analyses should therefore be interpreted with discretion. The small number of patients and events also meant that the most reliable analyses, which were restricted to patients with sufficient overlap of propensity scores, were imprecise. Whereas we have little evidence to suggest that TAVI is associated with increased mortality as compared with SAVR, the 95% CIs from these analyses are wide and compatible with either substantial benefits or harms of this new intervention. The strengths of this study include the following: To our knowledge, this is a unique endeavour to compare the characteristics and outcomes of patients with aortic stenosis undergoing TAVI and SAVR. Secondly, an interdisciplinary team of cardiologists, cardiac surgeons, statisticians and clinical epidemiologists performed a thorough examination of the potential for measured and unmeasured confounding. Thirdly, we had complete follow-up within the first 30 days of the index procedure.

The unadjusted 30-day mortality rates were 9.6% and 2.3%, for TAVI and SAVR groups, respectively. The results for the TAVI group compare favourably with contemporary outcome reports for the CoreValve ReValving System [1-3]. The 30-day mortality rate for the SAVR group was more than two times lower than the unadjusted mortality rate of 5.7% reported by Rankin et al for 216,245 patients undergoing single aortic valve replacement [25]. The TAVI group was mainly comprised of octogenarians and approximately 40% of patients had left ventricular dysfunction (EF <50%). Table 4 presents the in-hospital mortality rates of high-risk patients undergoing SAVR. The 30-day mortality rate of 9.6% observed in TAVI patients in our study is encouraging in light of the reported

inhospital mortality rates of selected octogenarians (4.6% to 13.5%) [4-11], patients with left ventricular dysfunction (6% to 33%) [12,13,26-31] and patients with high logistic EuroSCORES (7.8%) undergoing surgical aortic valve replacement [32].

While surgical heart valve replacement remains the standard of care, several studies have demonstrated that 30% to 60% of patients with symptomatic severe aortic valve stenosis are denied or not referred for surgery [33-36]. Patients who are denied or not referred for surgical aortic valve replacement are typically older, have moderate impairment of ejection fraction and more non-cardiac co-morbidities than patients undergoing valve surgery [33-36]. TAVI meets an unmet clinical need in these patients and may therefore be considered a breakthrough technology [37].

CONCLUSIONS

Both measured and unmeasured confounding limits the conclusions that can be drawn from observational comparisons of TAVI versus SAVR. Our study indicates that TAVI could be associated with either substantial benefits or harms. Randomised comparisons of TAVI and SAVR are warranted to provide evidencebased medicine and legitimise the use of TAVI in the eyes of the medical community, non-invasive cardiologist, cardiac surgeon, and health authorities. In addition, these trials will play a crucial role in reimbursement policies.

We are grateful to Jonathan Sterne and Marcel Zwahlen for helpful discussions related to the analytical strategies used in this study.
The analysis was funded by intramural grants provided by CTU Bern, Bern University Hospital, the Institute of Social and Preventive Medicine, University of Bern, and the Swiss National Science Foundation to Peter Jüni, a PROSPER fellow funded by the Swiss National Science Foundation. CTU Bern is supported by the Swiss National Science Foundation.

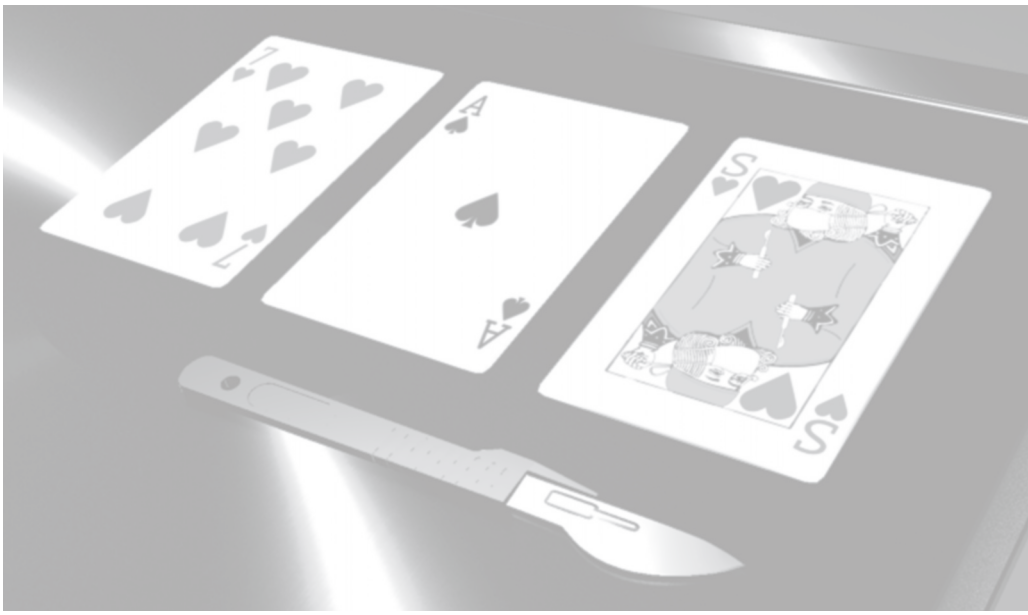
REFERENCES

1. Webb JG, Pasupati S, Humphries K, Thompson C, Altwegg L, Moss R, Sinhal A, Carere RG, Munt B, Ricci D, Ye J, Cheung A, Lichtenstein SV. Percutaneous transarterial aortic valve replacement in selected high-risk patients with aortic stenosis. Circulation 2007;116:755-763.
2. Piazza N, Grube E, Gerckens U, den Heijer P, Linke A, Luha O, Ramondo A, Ussia G, Wenaweser P, Windecker S, Laborde JC, de Jaegere P, Serruys PW. Procedural and 30-day outcomes following transcatheter aortic valve implantation using the third generation (18F) CoreValve ReValving System - Results from the Multicenter, Expanded Evaluation Registry 1-year following CE mark approval. EuroIntervention 2008;4:242-249.
3. Grube E, Buellesfeld L, Mueller R, Sauren B, Zickmann B, Nair D, Beucher H, Felderhoff T, Iversen S, Gerckens U. Progress and current status of percutaneous aortic valve replacement:Results of three device generations of the CoreValve ReValving System. Circ Cardiovasc Intervent 2008;1:167-175.
4. de Vincentiis C, Kunkl AB, Trimarchi S, Gagliardotto P, Frigiola A, Menicanti L, Di Donato M. Aortic valve replacement in octogenarians:is biologic valve the unique solution? Ann Thorac Surg 2008;85:1296-1301.
5. Gulbins H, Malkoc A, Ennker J. Combined cardiac surgical procedures in octogenarians: operative outcome. Clin Res Cardiol 2008;97:176-180.
6. Melby SJ, Zierer A, Kaiser SP, Guthrie TJ, Keune JD, Schuessler RB, Pasque MK, Lawton JS, Moazami N, Moon MR, Damiano RJ, Jr. Aortic valve replacement in octogenarians:risk factors for early and late mortality. Ann Thorac Surg 2007;83:1651-1656; discussion 1656-1657.
7. Kolh P, Kerzmann A, Honore C, Comte L, Limet R. Aortic valve surgery in octogenarians:predictive factors for operative and long-term results. Eur J Cardiothorac Surg 2007;31:600-606.
8. Langanay T, Verhoye JP, Ocampo G, Vola M, Tauran A, De La Tour B, Derieux T, Ingels A, Corbineau H, Leguerrier A. Current hospital mortality of aortic valve replacement in octogenarians. J Heart Valve Dis 2006;15(5):630-637; discussion 637.
9. Asimakopoulos G, Edwards MB, Taylor KM. Aortic valve replacement in patients 80 years of age and older:survival and cause of death based on 1100 cases:collective results from the UK Heart Valve Registry. Circulation 1997;96:3403-3408.
10. Bose AK, Aitchison JD, Dark JH. Aortic valve replacement in octogenarians. J Cardiothorac Surg 2007;2:33.
11. Filsoufi F, Rahmanian PB, Castillo JG, Chikwe J, Silvay G, Adams DH. Excellent early and late outcomes of aortic valve replacement in people aged 80 and older. J Am Geriatr Soc 2008;56:255-261.
12. Quere JP, Monin JL, Levy F, Petit H, Baleynaud S, Chauvel C, Pop C, Ohlmann P, Lelguen C, Dehant P, Gueret P, Tribouilloy C. Influence of preoperative left ventricular contractile reserve on postoperative ejection fraction in low-gradient aortic stenosis. Circulation 2006;113:1738-1744.
13. Subramanian H, Kunadian B, Dunning J. Is it ever worth contemplating an aortic valve replacement on patients with low gradient severe aortic stenosis but poor left ventricu-

lar function with no contractile reserve? Interactive cardiovascular and thoracic surgery 2008;7:301-305.
14. de Jaegere P, van Dijk L, Laborde JC, Sianos G, Ramos FJO, Lighart J, Kappetein AP, vander Ent M, Serruys PW. True percutaneous implantation of the CoreValve aortic valve prosthesis by the combined use of ultrasound guided vascular access, Prostar XL and the TandemHeart. EuroIntervention 2007;2:500-505.
15. Piazza N, Serruys PW, de Jaegere P. Feasibility of complex coronary intervention in combination with percutaneous aortic valve implantation in patients with aortic stenosis using percutaneous left ventricular assist device (TandemHeart). Catheter Cardiovasc Interv 2009;73:161-166.
16. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9-13.
17. Akins CW, Miller DC, Turina MI, Kouchoukos NT, Blackstone EH, Grunkemeier GL, Takkenberg JJ, David TE, Butchart EG, Adams DH, Shahian DM, Hagl S, Mayer JE, Lytle BW. Guidelines for reporting mortality and morbidity after cardiac valve interventions. J Thorac Cardiovasc Surg 2008;135:732-738.
18. Heckman J, Lalonde R, Smith J. The economics and econometrics of active labour market programs, in A. Ashenfelter and D. Card, eds., Handbook of Labour Economics, vol. 3, Amsterdam: Elsevier Science. 1999.
19. Silverman BW. Density Estimation for Statistics and Data Analysis. Chapman and Hall, London. 1986.
20. Smith J, Todd P. Does matching overcome LaLonde's critique of nonexperimental estimators? Journal of Econometrics 2005;125:305-353.
21. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in medicine 1998;17:2265-2281.
22. Rubin DB, Thomas N. Matching using estimated propensity scores:relating theory to practice. Biometrics 1996;52:249-264.
23. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology, Cambridge, Mass 2000;11:550-560.
24. Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, Robins JM. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. American journal of epidemiology 2006;163:262-270.
25. Rankin JS, Hammill BG, Ferguson TB, Jr., Glower DD, O'Brien SM, DeLong ER, Peterson ED, Edwards FH. Determinants of operative mortality in valvular heart surgery. J Thorac Cardiovasc Surg 2006;131:547-557.
26. Clavel MA, Fuchs C, Burwash IG, Mundigler G, Dumesnil JG, Baumgartner H, Bergler-Klein J, Beanlands RS, Mathieu P, Magne J, Pibarot P. Predictors of outcomes in low-flow, low-gradient aortic stenosis:results of the multicenter TOPAS Study. Circulation 2008;118(14 Suppl):S234-242.
27. Levy F, Laurent M, Monin JL, Maillet JM, Pasquet A, Le Tourneau T, Petit-Eisenmann H, Gori M, Jobic Y, Bauer F, Chauvel C, Leguerrier A, Tribouilloy C. Aortic valve replacement for low-flow/low-gradient aortic stenosis operative risk stratification and long-term outcome:a European multicenter study. J Am Coll Cardiol 2008;51:1466-1472.

28. Tarantini G, Buja P, Scognamiglio R, Razzolini R, Gerosa G, Isabella G, Ramondo A, Iliceto S. Aortic valve replacement in severe aortic stenosis with left ventricular dysfunction:determinants of cardiac mortality and ventricular function recovery. Eur J Cardiothorac Surg 2003;24:879-885.
29. Sharony R, Grossi EA, Saunders PC, Schwartz CF, Ciuffo GB, Baumann FG, Delianides J, Applebaum RM, Ribakove GH, Culliford AT, Galloway AC, Colvin SB. Aortic valve replacement in patients with impaired ventricular function. Ann Thorac Surg 2003;75:1808-1814.
30. Pereira JJ, Lauer MS, Bashir M, Afridi I, Blackstone EH, Stewart WJ, McCarthy PM, Thomas JD, Asher CR. Survival after aortic valve replacement for severe aortic stenosis with low transvalvular gradients and severe left ventricular dysfunction. J Am Coll Cardiol 2002;39:1356-1363.
31. Powell DE, Tunick PA, Rosenzweig BP, Freedberg RS, Katz ES, Applebaum RM, Perez JL, Kronzon I. Aortic valve replacement in patients with aortic stenosis and severe left ventricular dysfunction. Archives of internal medicine 2000;160:1337-1341.
32. Grossi EA, Schwartz CF, Yu PJ, Jorde UP, Crooke GA, Grau JB, Ribakove GH, Baumann FG, Ursumanno P, Culliford AT, Colvin SB, Galloway AC. High-risk aortic valve replacement:are the outcomes as bad as predicted? Ann Thorac Surg 2008;85:102-106; discussion 107.
33. Bouma BJ, van Den Brink RB, van Der Meulen JH, Verheul HA, Cheriex EC, Hamer HP, Dekker E, Lie KI, Tijssen JG. To operate or not on elderly patients with aortic stenosis:the decision and its consequences. Heart (British Cardiac Society) 1999;82:143-148.
34. Charlson E, Legedza AT, Hamel MB. Decision-making and outcomes in severe symptomatic aortic stenosis. The Journal of heart valve disease 2006;15:312-321.
35. Iung B, Cachier A, Baron G, Messika-Zeitoun D, Delahaye F, Tornos P, Gohlke-Barwolf C, Boersma E, Ravaud P, Vahanian A. Decision-making in elderly patients with severe aortic stenosis:why are so many denied surgery? Eur J H 2005;26:2714-2720.
36. Pellikka PA, Sarano ME, Nishimura RA, Malouf JF, Bailey KR, Scott CG, Barnes ME, Tajik AJ. Outcome of 622 adults with asymptomatic, hemodynamically significant aortic stenosis during prolonged follow-up. Circulation 2005;111:3290-3295.
37. Food and Drug Administration Website. Guidance for industry and FDA staff:Expedited review of premarket submissions for devices. 2008; http://www.fda.gov/CDRH/mdufma/guidance/108.html

CHAPTER 6



Relationship between the logistic EuroSCORE and the Society of Thoracic Surgeons Predicted Risk of Mortality score in patients implanted with the CoreValve ReValving System: a Bern-Rotterdam Study

Piazza M, Wenaweser P, Van Gameren M, Pilgrim T, Tsikas A, Otten A, Nuis R, Onuma Y, Cheng JM, Kappetein AP, Boersma H, Juni P, De Jaegere P, Windecker S, Serruys PW

Am Heart J. 2010; 159: 323-9

ABSTRACT

Background

Surgical risk scores, such as the logistic EuroSCORE (LES) and Society of Thoracic Surgeons Predicted Risk of Mortality (STS) score, are commonly used to identify high-risk or "inoperable" patients for transcatheter aortic valve implantation (TAVI). In Europe, the LES plays an important role in selecting patients for implantation with the Medtronic CoreValve System. What is less clear, however, is the role of the STS score of these patients and the relationship between the LES and STS.

Objective

The purpose of this study is to examine the correlation between LES and STS scores and their performance characteristics in high-risk surgical patients implanted with the Medtronic CoreValve System.

Methods

All consecutive patients (n = 168) in whom a CoreValve bioprosthesis was implanted between November 2005 and June 2009 at 2 centers (Bern University Hospital, Bern, Switzerland, and Erasmus Medical Center, Rotterdam, The Netherlands) were included for analysis. Patient demographics were recorded in a prospective database. Logistic EuroSCORE and STS scores were calculated on a prospective and retrospective basis, respectively.

Results

Observed mortality was 11.1%. The mean LES was 3 times higher than the mean STS score (LES 20.2% ± 13.9% vs STS 6.7% ± 5.8%). Based on the various LES and STS cutoff values used in previous and ongoing TAVI trials, 53% of patients had an LES ≥15%, 16% had an STS ≥10%, and 40% had an LES ≥20% *or* STS ≥10%. Pearson correlation coefficient revealed a reasonable (moderate) linear relationship between the LES and STS scores, r = 0.58, P b .001. Although the STS score outperformed the LES, bothmodels had suboptimal discriminatory power (c-statistic, 0.49 for LES and 0.69 for STS) and calibration.

Conclusions

Clinical judgment and the Heart Team concept should play a key role in selecting patients for TAVI, whereas currently available surgical risk score algorithms should be used to guide clinical decision making.

INTRODUCTION

Surgical risk scores, such as the logistic EuroSCORE (LES) and Society of Thoracic Surgeons Predicted Risk of Mortality (STS-PROM) score, are commonly used in clinical trials to identify high-risk surgical or "inoperable" patients for transcatheter aortic valve implantation (TAVI) [1-4]. Furthermore, these surgical risk scores are occasionally used as benchmark performance measures for TAVI procedures—a practice that can lead to complacency on the part of the treating physician, especially when the score grossly overestimates the actual mortality risk of TAVI [4].

Among other inclusion criteria, the 21F CoreValve Safety and Efficacy trial and subsequent 18F Safety and Efficacy trial required an LES ≥20% and ≥15%, respectively [5]. For inclusion into the Edwards PARTNER EU trial, patients required an LES ≥20% or STS score ≥10% [6]. These 3 trials were European-based, prospective, nonrandomized multicenter safety and efficacy trials. In contrast, the pivotal randomized PARTNER US trial is enrolling patients in whom the calculated STS score is ≥10% or the surgeon and cardiologist (Heart Team) concur that the individual patient's predicted risk of operative mortality is ≥15%.

Recently, questions have been raised regarding the reliability of risk algorithms in predicting early operative outcomes in high-risk patients undergoing surgical aortic valve replacement [7-11]. This has led some investigators to question the use of risk algorithms in selecting high-risk or "inoperable" patients for TAVI.

Despite these concerns regarding the reliability of risk scores, the LES continues to play an important role in selecting patients for CoreValve implantation in Europe. What is less clear, however, is the importance of the STS score of patients undergoing CoreValve implantation in Europe and the relationship between the LES and STS scores in these patients. In this retrospective study, we examine the correlation between the LES and STS score and their performance characteristics in high-risk surgical patients implanted with the CoreValve ReValving System at 2 European centers.

METHODS

Patients

We reviewed the records of 168 consecutive patients in whom a CoreValve ReValving System was implanted between November 2005 and June 2009 at 2 centers: Bern University Hospital (n = 76), Bern, Switzerland, and Erasmus Medical Center (n = 92), Rotterdam, The Netherlands. Patients were referred for TAVI implantation after a team of physicians (typically including interventional cardiologists and cardiac surgeons) agreed that surgical replacement would be associated with either high or prohibitive risk. The LES was prospectively calculated with an online calculator (http:// www.euroscore.org) to estimate the baseline surgical operative risk. A cutoff value of 15% was used to guide clinical decision making.

The LES comprises 17 variables divided into 3 categories (patient related, cardiac related, and operation related) [12]. The β coefficients for the LES regression model can be obtained on the Web site for the European System for Cardiac Operative Risk Evaluation (http://www.euroscore.org).

Although the STS for aortic valve procedures was calculated retrospectively using an online calculator (dataset 2.61, http://209.220.160.181/STSWebRiskCalc261/de.aspx), clinical variables required for its calculation had been collected prospectively. The STS includes 41 variables divided into 8 categories (demographics, risk factors, previous interventions, preoperative cardiac status, preoperative medications, hemodynamics and catheterization, operative, and valve surgery). β coefficients for the STS risk model recently have become available [13]. In our practice, the STS score was not used for clinical decision making.

Primary end point

Periprocedural death (mortality within 30 days of the procedure or during the index hospitalization) was confirmed through medical records, civil registries, or telephone visits.

Device and procedural details

The initial 11 patients underwent implantation using the second-generation 21F CoreValve ReValving System, and the subsequent 157 consecutive patients underwent implantation using the third-generation 18F CoreValve ReValving Sys-

Table 1. Baseline characteristics of the study population (n = 168).

	n = 168
Age, y (SD)	82.8 (5.5)
Female (%)	56.1%
LES, % (SD)	20.2 (13.9)
NYHA class	
I	0.9%
II	12.3%
III	68.4%
IV	18.4%
Diabetes mellitus (%)	22.8%
Hypertension (%)	63.1%
Coronary artery disease (%)	56.1%
Previous coronary bypass surgery (%)	24.6%
LVEF (%)	
>50%	58.8%
30%-50%	35.1%
<30%	6.1%
Atrial fibrillation (%)	19.3%
Cerebrovascular disease (%)	17.5%
Peripheral vascular disease (%)	18.4%
COPD (%)	21.1%
Pulmonary hypertension (%)	29.8%
Creatinine, µmol/L (SD)	92.9%
Creatinine above 200 µmol/L (%)	6.1%
MI within 90 d of procedure (%)	3.5%

NYHA = New York Heart Association; LVEF = left ventricular ejection fraction; COPD = chronic obstructive pulmonary disease; MI = myocardial infarction.

tem. Details of the CoreValve device and procedure have been previously published [14].

Statistical analysis

Continuous variables are presented as means (±1 SD). Categorical variables are presented as frequencies and percentages. Comparisons between continuous variables were performed using Student t test.

The performance of the LES and STS risk algorithms were evaluated in terms

of their discrimination and calibration. Discriminatory power was assessed using the c-index (area under the receiver operating characteristic curve) with 95% CI—a c-index of 0.5 indicates no predictive ability, whereas a c-index of 1.0 represents perfect discrimination [15]. Calibration, comparing the observed and predicted probabilities, was evaluated by the Hosmer-Lemeshow goodness-of-fit test and graphically represented by a calibration plot [16]. The dashed smooth curve in a calibration plot reflects the nonparametric relation between observed and predicted risk mortality. The straight dotted line through the origin of a calibration plot represents perfect calibration. Triangles are based on quintiles of patients with similar predicted probabilities. Spikes at the bottom of the calibration plot illustrate the distribution of predicted probabilities. Hosmer-Lemeshow P values above .05 indicate a well-calibrated model for the study population in question.

The relationship between LES and STS scores was graphically represented using a correlation (scatter) plot. A line of best fit was drawn to further study the correlation between the LES and STS in our data set. Linear regression analysis was used to determine the equation for the best-fit line, y = a + bx (where x = independent variable [LES score], y = dependent variable [STS score], a = y-intercept, and b = regression coefficient). Pearson correlation coefficient indicated

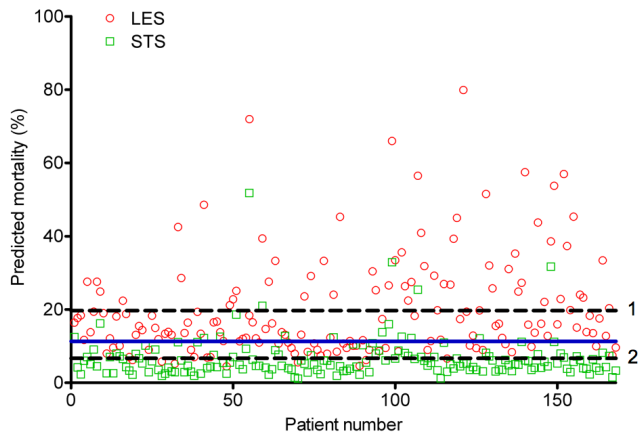


Figure 1. Scatter plot of the LES and STS scores for individual patients (n = 168). The dotted lines 1 and 2 represent the mean LES (20.2% ± 13.9%) and STS scores (6.7% ± 5.8%). The solid line represents the observed mortality rate (11.1%)..

the strength and direction of the linear relationship between the 2 variables.

Statistical analyses were performed with SPSS software version 15 (SPSS Institute, Chicago, IL). The R version software 2.5.1 (R Foundation for Statistical Computing, Vienna, Austria) was used for calculating c-values and Hosmer-Lemeshow P values and for constructing receiver operating characteristic curves and calibration plots.

RESULTS

Between November 2005 and June 2009, 168 consecutive patients were implanted with the Medtronic CoreValve aortic valve bioprosthesis and were included in the study analysis. Baseline characteristics of the study patients are summarized in Table 1.

Follow-up for the primary end point was 100% complete. There were 19 periprocedural deaths among the 168 patients (11.3%). The presumed causes of death are summarized in Table 2.

Logistic EuroSCORE and STS risk score analysis

Individual patient level data for LES and STS scores are shown in Figure 1, and the mathematical difference between these risk scores is shown in Figure 2. The mean LES was 3 times higher than the mean STS score (LES 20.2% ± 13.9% vs STS 6.7 ± 5.8%), and the mean difference between the scores was 13.0% ± 12.0%. The LES was higher than the STS score in all but 3 patients; in these 3 patients, the absolute difference between the STS and LES was 0.35%, 0.72%, and 0.53% (Figure 2). Cutoff tertiles were 12% and 21% for LES and 4% and 7% for STS score.

Based on the various LES and STS cutoff values used in previous and ongoing TAVI trials, 35% had an LES ≥20%, 53% had an LES ≥15%, 16% had an STS ≥10%, 40% had an LES ≥20% or STS ≥10%, and 11% had an LES ≥20% and STS ≥10%.

Logistic EuroSCORE and STS correlation

Pearson correlation coefficient revealed a reasonable (moderate) linear relationship between the LES and STS scores, r (166) = 0.58, P b .001 (Figure 3). The regression equation that best described the relationship between the LES and STS scores for our data set was y = 1.79 + 0.243x (where y = dependent variable [STS

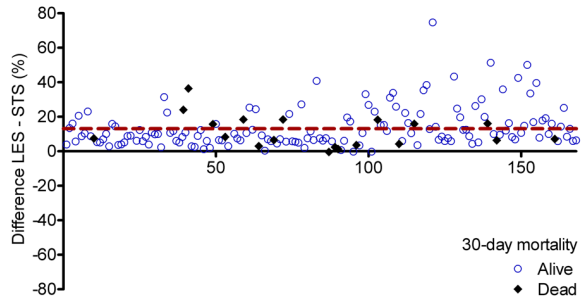


Figure 2. Scatter plot of the difference between LES and STS for individual patients (n = 168). The black diamonds represent patients who died. The dotted line represents mean difference between LES and STS scores (13 ± 12%).

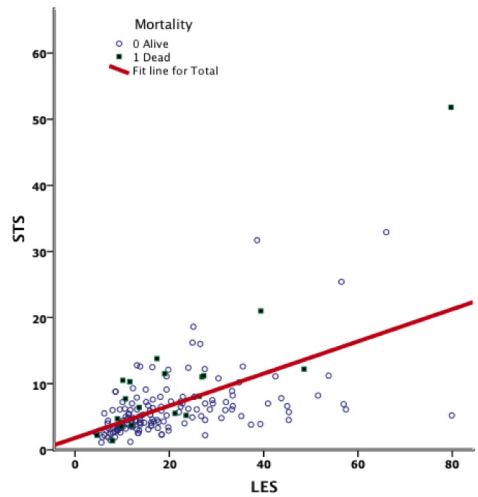


Figure 3. Correlation plot between the LES score (%) (x-axis) and STS score (%) (y-axis). The black diamonds represent patients who died. Pearson correlation coefficient, r (166) = 0.58, P < .001. The regression equation for the "best-fit" line was y = 1.79 + 0.243x (where y = dependent variable [STS score], a = constant [or y-intercept], b = regression coefficient, x = independent variable [LES score]).

score], a = constant [or y-intercept], b = regression coefficient, x = independent variable [LES score]).

Table 2. Summary of the presumed causes of death (n= 19 deaths).

	Day of death	Cause of death	LES (%)	STS (%)
1	6	Tamponade	19.0	11.5
2	29	Sepsis	48.6	12.2
3	0	Induction	21.2	5.5
4	24	Stroke	12.0	3.7
5	11	Sepsis	71.6	51.8
6	8	Asystole	39.4	21.0
7	9	Stroke	10.7	7.7
8	29	Sudden cardiac death	23.6	5.2
9	0	Tamponade	10.1	10.5
10	14	Cardiac Failure (severe pAR)	4.6	2.2
11	0	Electrical mechanical dissociation	11.6	10.3
12	30	Cardiac failure	8.0	1.4
13	0	Cardiac failure (severe AR post-BAV)	17.4	13.8
14	14	Cardiac failure	26.4	8.1
15	0	Tamponade	9.0	4.7
16	14	Sepsis	27.0	11.0
17	14	Stroke	27.3	11.2
18	25	Stroke	9.8	3.5
19	27	Stroke	13.6	6.4

AR = Aortic regurgitation; pAR = paravalvular aortic regurgitation; BAV = balloon aortic valvuloplasty.

Logistic EuroSCORE and STS performance measures

Mortality predicted by LES score was 20.2% (95% CI 18.1-22.3) and by STS was 6.7% (95% CI 5.8-7.6). The observed mortality was 11.1%. The area under the curve (c-index) was 0.49 (95% CI 0.35-0.63) for the LES and 0.69 (95% CI 0.53-0.82) for the STS score (Figure 4). Although the calibration plot favored the STS over the LES, both risk scores demonstrated significant differences between predicted and observed mortalities throughout the entire range of mortality risk (Figure 4). As confirmatory information, the Hosmer-Lemeshow test demonstrated that both models were poorly calibrated for our patient population (LES, P = .005, and STS, P = .013).

DISCUSSION

This study describes the correlation between LES and STS scores and their performance characteristics in highrisk surgical patients implanted with the Medtronic CoreValve System in 2 European centers. Whereas LES was calculated prospectively and used for clinical decision making, STS was calculated retrospectively and used for study purposes.

Several aspects were of note. First, LES consistently provided higher estimates of operative mortality than STS (mean LES score 20.2% ± 13.9% vs mean STS score 6.7% ± 5.8%). Second, our data set demonstrated a moderate correlation between the LES and STS score described by the linear equation y (STS) = 1.79 + 0.24x (LES). Third, although the STS score outperformed the LES, both risk scores had suboptimal discriminatory power (ie, ability to correctly classify dead or alive patients) and calibration (ie, ability to correctly estimate risk of mortality). The ensuing discussion attempts to relate these findings to the following clinical questions: (1) What is the role of surgical risk scores in selecting patients for TAVI trials? (2) What is the role of surgical risk scores in predicting outcomes of patients undergoing TAVI?

The large discrepancy between LES- and STS-predicted risk of mortality might invite a reappraisal as to whether some patients implanted with the CoreValve device were indeed high-risk surgical patients. One caveat, and a general shortcoming of all surgical risk algorithms, is the omission of several measurable and unmeasurable risk factors known to influence patient selection and mortality [12,13,17-20]. For example, both models fail to include porcelain aorta, mediastinal radiation, malnutrition, and frailty of the patient.

Several investigators, however, have shown that LES overestimates absolute mortality risk among high-risk patients and octogenarians undergoing surgical aortic valve replacement [7-10]. Based on a data set composed chiefly of coronary surgery patients (approximately 60%), LES may be less well adapted than dedicated valve surgery models for prediction of operative mortality in patients undergoing valve surgery [13,18-24]. Moreover, many believe that the LES, initially developed in 1999, has become increasingly uncalibrated owing to its overestimation of absolute risk [9,11]. Despite this, Osswald et al [9] suggest that LES still retains its ability to risk stratify patients.

At this time, a uniform approach to risk stratification based on absolute percentages is lacking. In fact, a number of studies examining the reliability of LES

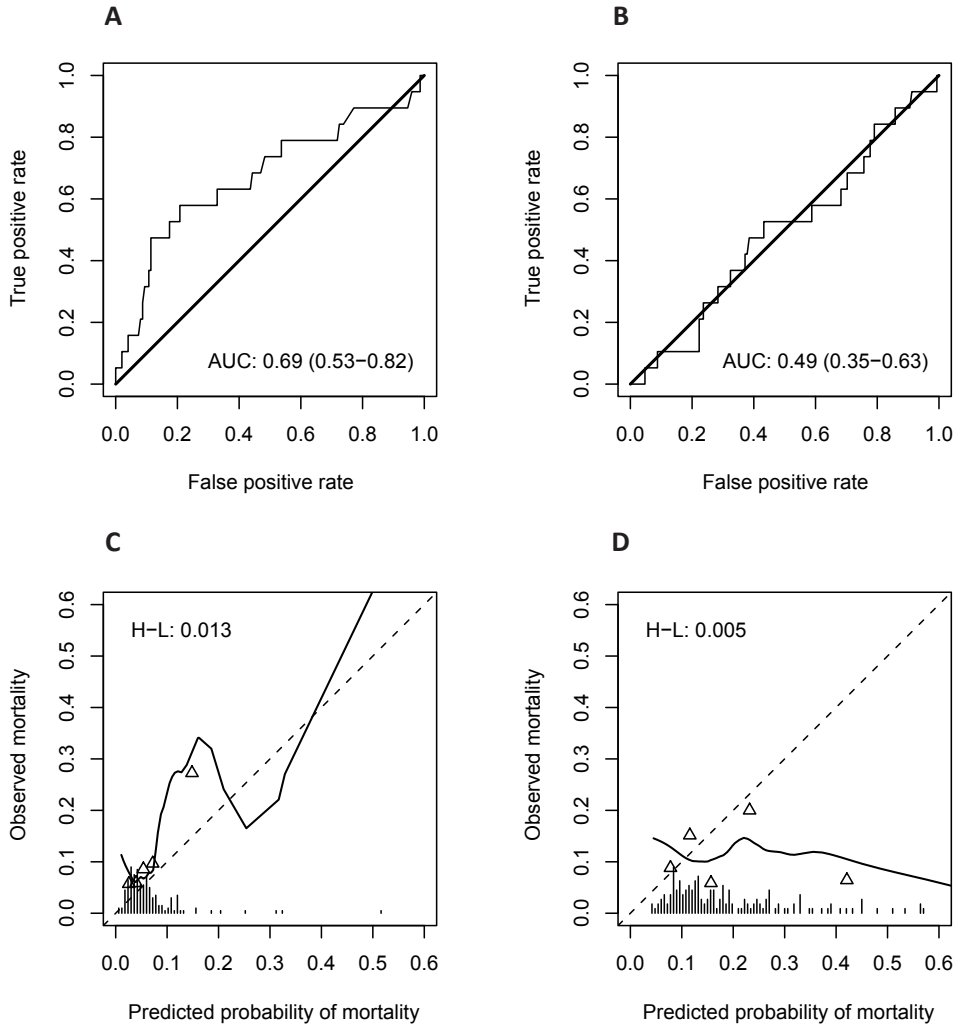


Figure 4. *A* and *B*, Receiver operating characteristic curve graphs for the STS and LES scores, respectively (n = 168 patients). *C* and *D*, Calibration plots for the STS and LES scores, respectively. The dashed smooth curve reflects the relation between observed mortality and predicted probability of mortality. Perfect calibration is represented by the straight line through origin whereby observed equals predicted probability of mortality. Spikes at the bottom represent the distribution of predicted probabilities (AUC = area under the curve).

and its implications for TAVI define "high risk" differently: LES ≥6% [9] versus LES ≥20%10 versus LES ≥90th percentile [8]. Furthermore, clinical TAVI trials use various cutoff values of risk scores (LES and/or STS) to define high risk. Based on the results of our study, 53%, 16%, and 40% had an LES ≥15% (eg, CoreValve Ex-

panded Evaluation Registry), STS ≥10% (eg, Edwards PARTNER US trial), and LES ≥20% or STS ≥10% (eg, Edwards PARTNER EU trial), respectively. These observations need to be interpreted in light of the additional criteria that may be driving patient selection (such as mediastinal radiation porcelain aorta or frailty).

Investigators should be discouraged from using 2 different risk scores as inclusion criteria for TAVI trials unless the relationship between the 2 risk scores is known—otherwise, this can lead to enrollment of 2 different patient populations. Among other inclusion criteria, the PARTNER EU trial enrolled patients with either an LES ≥20% or STS ≥10%. The findings of our study question the equivalency of these risk score cutoff values to identify high-risk patient subgroups. On the basis of a reasonable correlation between LES and STS, we attempted to establish a linear mathematical relationship between these 2 scores whereby STS = 1.79 + 0.243 LES. Imputing a 20% LES to this equation would result in an STS score of 6.7%, whereas imputing a 10% STS score would result in an LES of 33.8%. In view of the high-risk study population, this equation may not apply to “low” or “intermediate” risk subgroups. Furthermore, the equation would require validation.

Although diagnostic models are concerned with accurately classifying individuals into their true disease states, prognostic models add the element of time and chance. Evaluation of prognostic models requires an appreciation of both discrimination (ability to correctly classify a binary outcome, eg, dead/alive) and calibration (comparison between observed and predicted probabilities) [25]. Although the STS score outperformed the LES, both LES and STS had suboptimal discriminatory power and calibration. This may not come as a surprise when one considers the notion of extrapolating a surgical risk score to the outcomes of a distinct form of therapy such as TAVI. Furthermore, using surgical risk scores as benchmark performance measures can lead to complacency on the part of the treating physician especially when the surgical risk score grossly overestimates the actual mortality risk of the TAVI procedure.

In summary, 2 prognostic models are needed: (1) a surgical model to help identify high-risk patients for TAVI (developed using a reasonable proportion of high-risk patients) and (2) a transcatheter model to help make treatment decisions, advise patients, and serve as a benchmark performance measure. Risk scores for these purposes are currently lacking and should be the focus of future studies. Clinical judgment should play a key role in the patient selection process, and risk scores (keeping in mind their respective shortcomings) should be used

to guide clinical decision making.

Limitations

Although all necessary baseline variables were collected prospectively, a limitation of this study includes the retrospective calculation of the STS score. Notwithstanding the sufficient sample size to perform decent statistical analyses, a larger population could result in more reliable model performance results. Another limitation is that only 2 centers were involved in this study; the Heart Team selection process and inherent patient selection bias might have led to findings that are not applicable to other centers.

Conclusions

The study results confirm those of previous studies that found the LES to provide substantially higher surgical risk estimates than the STS risk algorithm. In contrast to previous studies that examined the comparative performance of risk scores in patients undergoing surgical aortic valve replacement, the current study uniquely examined patients undergoing TAVI [12,13,17-20]. Importantly, the purpose of this study was not to discredit "per se" the validity of the surgical risk score for purposes they were not intended, but rather to underline that patient selection and predicting outcomes after TAVI requires dedicated models.

Until a validated risk score becomes available and a better understanding of what constitutes "high risk" is elucidated, clinical judgment based on the Heart Team concept should be used to guide patient selection.

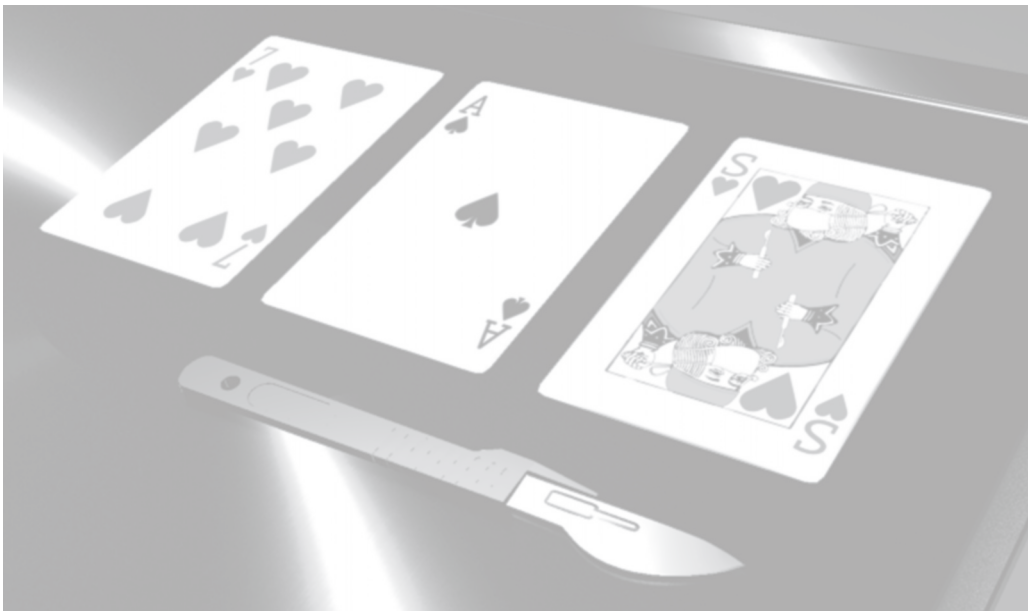
REFERENCES

1. Cribier A, Eltchaninoff H, Tron C, et al. Treatment of calcific aortic stenosis with the percutaneous heart valve: mid-term follow-up from the initial feasibility studies: the French experience. J Am Coll Cardiol 2006;47:1214-23.
2. Grube E, Schuler G, Buellesfeld L, et al. Percutaneous aortic valve replacement for severe aortic stenosis in high-risk patients using the second- and current third-generation self-expanding CoreValve prosthesis: device success and 30-day clinical outcome. J Am Coll Cardiol 2007;50:69-76.
3. Piazza N, Grube E, Gerckens U, et al. Procedural and 30-day outcomes following transcatheter aortic valve implantation using the third generation (18F) CoreValve ReValving System—results from the multicenter, expanded evaluation registry 1-year following CE mark approval. Eurointervention 2008:242-9.
4. Webb JG, Pasupati S, Humphries K, et al. Percutaneous transarterial aortic valve replacement in selected high-risk patients with aortic stenosis. Circulation 2007;116:755-63.
5. Buellesfeld L. 1 year results from the CoreValve 18F safety and efficacy study. Barcellona: EuroPCR; 2009.
6. Schächinger VobotPEI. Results from the PARTNER EU Trial: prospective multicentric European registry of transcatheter aortic valve implantation—primary endpoint analysis. Barcelona: EuroPCR; 2009.
7. Brown ML, Schaff HV, Sarano ME, et al. Is the European System for Cardiac Operative Risk Evaluation model valid for estimating the operative risk of patients considered for percutaneous aortic valve replacement? J Thorac Cardiovasc Surg 2008;136:566-71.
8. Dewey TM, Brown D, Ryan WH, et al. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. J Thorac Cardiovasc Surg 2008;135:180-7.
9. Osswald BR, Gegouskov V, Badowski-Zyla D, et al. Overestimation of aortic valve replacement risk by EuroSCORE: implications for percutaneous valve replacement. Eur Heart J 2009; 30:74-80.
10. Leontyev S, Walther T, Borger MA, et al. Aortic valve replacement in octogenarians: utility of risk stratification with EuroSCORE. Ann Thorac Surg 2009;87:1440-5.
11. van Gameren M, Kappetein AP, Steyerberg EW, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg 2008;85:921-30.
12. Nashef SA, Roques F, Michel P, et al. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9-13.
13. O'Brien SM, Shahian DM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2—isolated valve surgery. Ann Thorac Surg 2009;88:S23-42.
14. de Jaegere P, van Dijk L, Laborde JC, et al. True percutaneous implantation of the CoreValve aortic valve prosthesis by the combined use of ultrasound guided vascular access, Prostar XL and the TandemHeart. Eurointervention 2007;2:500-5. 328 Piazza et al American Heart Journal February 2010
15. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation 2007;115:654-7.
16. Hosmer D, Lemeshow S. A goodness-of-fit test for the multiple logistic regression model.

Comm Stat 1980;A10:1043-69.

17. Rankin JS, Hammill BG, Ferguson Jr TB, et al. Determinants of operative mortality in valvular heart surgery. J Thorac Cardiovasc Surg 2006;131:547-57.
18. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting inhospital mortality for cardiac valve surgery. Ann Thorac Surg 2007; 83:921-9.
19. Kuduvalli M, Grayson AD, Au J, et al. A multi-centre additive and logistic risk model for in-hospital mortality following aortic valve replacement. Eur J Cardiothorac Surg 2007;31:607-13.
20. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77: 1966-77.
21. Ambler G, Omar RZ, Royston P, et al. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224-31.
22. Florath I, Rosendahl UP, Mortasawi A, et al. Current determinants of operative mortality in 1400 patients requiring aortic valve replacement. Ann Thorac Surg 2003;76:75-83.
23. Gardner SC, Grunwald GK, Rumsfeld JS, et al. Comparison of shortterm mortality risk factors for valve replacement versus coronary artery bypass graft surgery. Ann Thorac Surg 2004;77:549-56.
24. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005; 80:471-9.
25. van Gameren M, Piazza N, Bogers AJJC, et al. How to assess the risk of valve surgery: quality, implementation and future risk models. Heart 2009 Accepted for publication.

CHAPTER 7



Early complications of stenting in patients with congenital heart disease: a multicentre study

Van Gameren M
Witsenburg M
Takkenberg JJM
Boshoff D
Mertens L
Van Oort AM
De Wolf D
Freund M
Sreeram N
Bökenkamp R
Talsma MD
Gewillig M

Eur Heart J. 2006; 27: 2709-15

ABSTRACT

Aims

Stenting has become an established interventional cardiology procedure for congenital heart disease. Although most stent procedures are completed successfully, complications may occur. This multicentre study evaluated early complications after stenting in patients with congenital heart disease, including potential risk factors.

Methods and results

In this combined Dutch-Belgian retrospective study, 309 consecutive patients had undergone 366 catheterizations and received 464 stents in 13 different anatomical positions (418 sites). Seventy-two stenting-related complications (19%) occurred, of which 24 (5.7%) were major. Seven procedure-related deaths were documented (2.3%). Stent malpositioning and embolization were most common (7.7%). The use of non-premounted stents tended to be associated with higher complication rates. Centre inexperience with stenting and stenting of native vs. post-surgical stenosis tended to be associated with increased major complication rates.

Conclusion

After stenting, complications are common for congenital heart disease. The vast diversity of stenotic sites combined with relatively small patient populations makes these procedures sensitive to complications. Combining operator experience may reduce the risks of stenting in congenital heart disease. The availability of premounted stents for greater vessel diameters will likely reduce incidences of stent migration and embolization.

INTRODUCTION

Since the first report on the use of stents in congenital heart disease in 1991 [1], stenting has gradually replaced —or postponed to a later date— a variety of surgical interventions of either a corrective or palliative nature [2,3]. At present, stenting is the preferred intervention for a large variety of primary or acquired stenoses in congenital heart disease [1,4].

Intravascular stents are mainly used for resolving stenoses that do not respond to conventional balloon dilation. These include: compliant obstructions, stenoses due to kinking or external compression and several post-operative stenoses, aortic re-coarctation being a good example of this.

The use of stents for treating a stenosis in children is challenging, as ideally the final vessel diameter after stenting should ultimately approach the adult vessel size. Currently, only non-premounted stents can be dilated up to large diameters. When stents are used at younger ages, re-dilatation should be possible to accommodate for the expected vessel growth over time, unless stenting is being used to serve as a bridge for a surgical intervention. At present, stenting is therefore, mostly performed in older children and adults.

The annual number of stent procedures for congenital heart disease is relatively small and stented anatomical sites and patient characteristics vary widely. It appears that even in experienced hands, stenting can be complicated by different problems, ranging from malpositioning to vessel rupture and even procedure-related death [5]. To understand this, we performed a retrospective study in seven centres in the Netherlands and Belgium, evaluating rates and types of short-term complications after stenting for congenital heart disease. In addition, we looked at potential determinants of these complications.

METHODS

All five centres in the Netherlands and two in Belgium that perform stenting for congenital heart disease contributed to this study. We retrospectively evaluated all initial stent implantations performed in each centre since the start of their stenting practice. Stent re-dilatation procedures were not included.

Table 1 shows relevant statistics for each centre. Most stent implantation sites were evenly represented across centres. However, stenting for pulmonary

Table 1. Population distribution across centres.

	Patients	Procedures	Sites	Stents
Centre A	141	180	209	231
Centre B	33	38	41	45
Centre C	30	31	35	37
Centre D	29	33	40	49
Centre E	29	29	37	41
Centre F	28	33	34	39
Centre G	19	22	22	22
Total	309	366	418	464

vein stenosis, major aortapulmonary collateral arteries (MAPCA), patent arterial duct and atrial septum (AS) fenestration were almost exclusively performed in centre A.

Data collection and study design

Data was obtained from patient records and registered using an MS Access 2002 database (Microsoft Corporation, Silicon Valley, USA). Complications were categorized as major or minor. Major complications included all events leading to death, life-threatening haemodynamic compromise, the need for surgical intervention, or a substantial permanent anatomical or functional lesion. Minor complications included all transient un-anticipated events resulting from the stent procedure.

All complications were grouped in the following categories: balloon rupture, vessel dissection, stent malpositioning, migration or embolization, inaccessible implantation site, stent-induced pulmonary oedema, arrhythmias, and other complications. Stent malpositioning, migration, and embolization were grouped together because these events are fairly comparable and often hard to distinguish, the more so when they are being collected retrospectively.

Vessel dissection was considered as a major complication if it required surgical intervention or an additional procedure. Limited extravasation was assigned to the minor complications. Multiple similar complications at one stent site (e.g. repeated balloon rupture) were counted as one.

Study population

One centre started stenting practice in May 1992, the others followed in the

years between 1992 and 1997. Data was obtained through June 2004. Overall, 464 stents had been implanted in 309 patients including 180 males and 129 females. Ages ranged from 1 day to 68 years, at a mean of 11 years (±10.8 SD). Mean height and weight were 123 cm (±40.1 SD) and 32 kg (±24.5 SD), respectively. All but three of the 309 patients included in this study had congenital heart disease: two had Kawasaki disease-related arterial vasculitis, and one had severe iliac vein thrombosis. In 306 patients, stenting had been performed during cardiac catheterization. Three received one or multiple stents during open heart surgery. Two-hundred-and-sixty-three patients underwent one catheterization, 38 underwent two catheterizations, five patients with three catheterizations, and three patients underwent four catheterizations. Table 2 links the various stent implantation sites with corresponding stent and patient characteristics. Some specific details are given below.

The left and right pulmonary artery groups together reflected 114 post-surgical, 90 primary, four induced, and 15 unknown stenoses. Induced stenoses were those created by other previously implanted stents or transcatheter devices. The aorta group comprised 22 native and 47 recurrent coarctations. The systemic vein group included five primary stenoses, 12 post-surgical stenoses, two thrombosed sites, and one unknown stenosis. Four of the 12 postsurgical stenoses were modified Glenn anastomoses. Two stents in the atrial septal fenestrations group had been placed within previously implanted stents. This group also includes one fenestration stent in a total cavo-pulmonary connection. The right ventricular outflow tract and pulmonary artery trunk group reflects both primary and post-surgical stenoses. Eight of all these 11 stenoses were in surgically implanted conduits (i.e. allograft etc.). All stented aorto-pulmonary shunts were surgically constructed conduits, mostly modified Blalock-Taussig shunts.

Stent types

The 464 documented stents are hepatobiliary, renal, iliac, and coronary stents of different types and sizes. The great diversity reflects the multicentre set-up and long-term frame of this study. Of the stents, 255 were non-premounted, 145 premounted, and 41 selfexpanding. Of 23 stents, the type and manufacturer could not be identified retrospectively. Non-premounted stents were mostly the Cordis Palmaz stents. In premounted stents, a large variation in types and manufacturers were found. Of the self-expanding stents, the Boston Scientific Wallstent was mostly used.

Table 2. Stented site characteristics.

Stented site	N	Stent diameter (mm) (mean±SD)	Patient age (years) (mean±SD)	Patient weight (kg) (mean±SD)
Left pulmonary artery	129	10.5 (3.5)	7.6 (6.9)	26.5 (19.1)
Right pulmonary artery	94	10.1 (3.8)	9.9 (9.7)	32.0 (24.2)
Aorta	69	14.2 (4.8)	18.6 (14.5)	51.0 (24.4)
MAPCA	25	7.2 (2.0)	13.4 (7.9)	37.1 (16.4)
Arterial duct	23	3.8 (0.4)	0.1[a]	3.5 (0.9)
Systemic vein	20	14.0 (5.0)	9.7 (8.4)	30.5 (28.6)
Atrial septal fenestration	16	8.7 (3.2)	9.4 (12.3)	20.4 (18.0)
Mustard baffle	12	20.6 (3.5)	18.8 (6.5)	56.0 (17.3)
RVOT, pulmonary artery trunk	11	12.5 (6.5)	5.1 (5.9)	20.4 (20.4)
Aorto-pulmonary shunt	9	5.3 (2.1)	10.2 (13.2)	23.9 (21.5)
Pulmonary vein	5	5.4 (2.1)	1.6 (0.9)	8.5 (6.4)
Coronary artery	3	3.3 (0.6)	11.7 (1.5)	48.0 (26.5)
Other sites	2	8.0 (5.7)	1.5 (2.1)	12.0 (0)
Total	418			

RVOT = right ventricular outflow tract.

[a] Mean age in duct patients was 20 days.

Statistical analysis

For descriptive statistical analysis, SPSS 11.0.1 for Windows software (SPSS Inc., Chicago, USA) was used. Continuous variables are presented as means+SD. Discrete variables are presented both as proportions and counts. Potential determinants of complications, major complications, and death were analysed by SAS 9.1 (SAS Institute Inc., Cary, NC, USA). A logistic regression model was constructed for all 366 procedures corrected for multiple procedures in the same patient [generalized estimating equation (GEE)]. Repeated measures were coded per procedure; a compound symmetric correlation matrix was used. Outcome measures were all complications, major complications, and death. The following potential determinants of outcome were entered into the model: centre experience (the first 20 procedures in each centre were considered to be inexperienced, and all consecutive procedures to be experienced), premounted vs. unmounted stents, age , 1 year, selfexpandable vs. balloon-expandable stents, and native vs. post-surgical stenosis. Furthermore, using the same GEE method, a subset analysis

Table 3. Complication rates per site.

Stented site	N	Total complications (%)	Major complications (%)	Minor complications (%)
Left pulmonary artery	129	25 (19)	10 (8)	15 (12)
Right pulmonary artery	94	12 (13)	4 (4)	8 (9)
Aorta	69	9 (13)	2 (3)	7 (10)
MAPCA	25	4 (16)	0 (0)	4 (16)
Arterial duct	23	8 (35)	3 (13)	5 (22)
Systemic vein	20	3 (15)	0 (0)	3 (15)
Atrial septal fenestration	16	2 (13)	0 (0)	2 (13)
Mustard baffle	12	1 (8)	1 (8)	0 (0)
RVOT, pulmonary artery trunk	11	5 (46)	2 (18)	3 (27)
Aorto-pulmonary shunt	9	2 (22)	1 (11)	1 (11)
Pulmonary vein	5	1 (20)	1 (20)	0 (0)
Coronary artery	3	0 (0)	0 (0)	0 (0)
Other sites	2	0 (0)	0 (0)	0 (0)
Total	418	72 (17)	24 (5.7)	48 (11.3)

RVOT = right ventricular outflow tract.

was done for those procedures that involved pulmonary artery stenting. This analysis is aimed at determining the influence of simultaneous stenting of the right and left pulmonary artery on outcome. Finally, a similar subset analysis was accomplished for those procedures that involved stenting of a coarctation, aimed at determining the influence of native vs. re-coarctation on outcome.

RESULTS

Complications were documented for 69/366 (19%) procedures at 72/418 (17%) stented sites. With 24 major complications, the procedure-related major complication rate was 6.5% and the anatomical site-related major complication rate 5.7%.

Complications by implantation site

Table 3 shows complication rates for the various implantation sites. Branch pul-

monary artery stenting had been most frequently performed, followed by stenting of the aorta. Thirty-three patients underwent stenting of both a left and a right side pulmonary artery stenosis in the same procedure. The number of stent implantations for all other anatomical sites did not exceed 25. Complication rates were highest (45%) in the right ventricular outflow tract/ pulmonary trunk group; this was a small group including several post-surgical patients. Relatively high complication rates were also found in the arterial duct group (35%), the aorto-pulmonary shunt group (22%), and the pulmonary vein group (20%). Table 4 shows a more detailed distribution of complications by implantation site. Table 5 shows frequencies of the different types of complications which were broken down for complication categories, and which include major complications and death. Occurring in 7.7% of all implantations, the combined stent malpositioning, migration, and/or embolization complication type was most frequent.

Mortality

The death of seven patients was directly related to the stent implantation procedure, resulting in a 2.3% procedural mortality. These patients' details are presented in Table 6. Three were younger than 1 year (4.5% fatal complication rate in infants). Most of these infants were severely ill. For patients older than 1 year, the fatal complication rate was 1.1%.

Potential determinants of outcome

Table 7 shows that the use of a premounted stent tends to be significantly associated with both a lower overall complication rate and lower major complication rate as compared with the use of a non-premounted stent. Furthermore, although centre experience with stenting does not affect overall complication and death rates, there is a tendency towards less major complications in experienced centres. Children younger than 1 year at the time of the procedure tend to be at increased risk for death, but not for complications. The use of a self-expandable stent does not affect outcome. Finally, stenting of a native lesion tends to be associated with more major complications as compared with a post-surgical stenosis.

Subset analysis of the procedures that involved pulmonary artery stenting shows no association between simultaneous stenting of the right and left pulmonary artery and outcome [overall complications: OR 1.57 (95% CI 0.65–3.79), P ¼ 0.32; major complications: OR 2.47 (95% CI 0.31–19.7), P ¼ 0.39]. Overall com-

Table 4. Complication categories (major/minor) by implantation site.

Stented sites (418)	Left PA (129)	Right PA (94)	Aorta (69)	MAP-CA (25)	Arterial duct (23)	Systemic vein (20)	AS fenestration (16)	Mustard baffle (12)	RVOT, PA trunk (11)	Aorto-pulmonary shunt (9)	Pulmonary vein (5)
Malposition, migration or embolization	5 / 7	2 / 4	- / 4	- / 3	- / 3	- / 2	- / -	- / 1	- / 1	- / -	- / -
Inaccessible site	2 / 4	- / 2	- / -	- / -	- / -	- / -	- / -	- / -	- / -	- / -	- / -
Balloon rupture	1 / 4	- / 1	- / 3	- / -	- / -	- / 1	- / -	- / -	- / 1	- / -	- / -
Dissection	1 / -	- / -	2(1)[a] / -	- / -	1 / 1	- / -	- / -	- / -	- / 1	- / -	- / -
Induced pulmonary oedema	1 / -	1(1)[a] / -	- / -	- / -	- / -	- / -	- / -	- / -	- / -	- / -	- / -
Arrhythmia	- / -	- / 1	- / -	- / -	1 / 1	- / -	- / 1	- / -	- / -	- / -	- / -
Other	- / -	1(1)[a] / -	- / -	- / 1	1(1)[a] / -	- / -	- / 1	- / -	2(1)[a] / -	1(1)[a] / -	1(1)[a] / -

AS = atrial septum; PA = pulmonary artery; RVOT = right ventricular outflow tract.

[a] Number of complications leading to death.

Table 5. Distribution of complication categories.

Complication type	Total	Major	Mortality
Malposition, migration or embolization	32	8	0
Balloon rupture	11	1	0
Inaccessible site	8	2	0
Vessel dissection	6	4	1
Arrhythmia	5	1	0
Induced pulmonary oedema	2	2	1
Other	8	6	5

Table 6. Mortality.

Age	Weight (kg)	Sex (M/F)	Diagnosis	Cause of death	Stent location
4 months	3	F	Fallot, severe right ventricular outflow tract obstruction	Severe bradycardia prior to stent dilation	RVOT-pulmonary artery trunk
7 months	6	F	Aortic re-coarctation	Retro-peritoneal bleeding	Aorta
1 month	3	M	Severe aortic valve stenosis, LV hypoplasia	Ductal closure due to insufficient ductus coverage by stent	Arterial duct
2 years	13	F	Idiopathic pulmonary vein stenosis	Thrombotic middle cerebral artery occlusion secondary to pulmonary vein stent thrombosis	Pulmonary vein
11 years	50	F	Post-operative Fallot, pulmonary branch stenosis	Coronary ischaemia due to stent induced compression, fatal LV dysfunction	Right pulmonary artery
4 years	21	F	Post-operative Fallot, LV hypoplasia	Possible vessel rupture, hypotensia and tachycardia at ICU after procedure[a]	Blalock-Taussig shunt
51 years	79	M	Native right pulmonary artery stenosis	Hyper-perfusion right lung, followed by acute pulmonary oedema	Right pulmonary artery

[a] Post-mortem examination was not performed.

plication rates in this subset are significantly lower with the use of premounted stents [OR 0.28 (95% CI 0.10–0.86), P ¼ 0.03].

Subset analysis of the procedures that involved stenting of aortic coarctations shows no differences in outcome after stenting of native vs. recurrent coarctations [OR 4.76 (95% CI 0.56–40.62), P ¼ 0.15).

DISCUSSION

Though the annual number of stent procedures for the treatment of congenital heart defects has gradually increased since its introduction over 15 years ago, it still remains minimal. Given the large variety in anatomical locations and stent types, it is not surprising that early complications should be quite common. We found a 19% overall complication rate and a 5.7% major complication rate. Other studies on early complications after stent implantation for congenital heart disease are rare. Agnoletti et al. [6] reported an even higher complication rate and a more frequent need for surgical intervention, but lack of details preclude a valid comparison of findings. Fortunately, many problems during stent implantation are transient and can be solved without ending the procedure, or may be solved by surgery with acceptable results [7].

Procedural mortality in the present study tended to be higher in infants when compared with children older than 1 year at the time of the procedure. Stenting in two cases —an infant with Tetralogy of Fallot and an infant with aortic re-coarctation— may be debatable in retrospect, especially because reliable surgical solutions have been available for these indications for years. Several centres now routinely perform ductal stenting and have provided technical details that should prevent inappropriate ductal coverage by the stent [2,8]. Pulmonary vein stenosis bears a grim prognosis and surgical solutions are often not available. It requires aggressive anticoagulation during and after pulmonary vein stenting, which will not always prevent systemic embolic complications like the one reported here. Our documented case of coronary compression induced by branch pulmonary artery stenting illustrates the indispensability of detailed morphological work-up either by MRI or angiography before stenting is performed.

The use of non-premounted stents, in our study, tended to be associated with higher complication rates, with stent malpositioning, migration and embolization occurring most frequently. The comparison with premounted stents is

biased, however, as the latter cannot be used in larger vessels. Non-premounted stents therefore are predominant in older patients. Other studies comparing premounted with non-premounted stents related to congenital heart disease were not found. Premounted stents have now become standard in coronary heart disease. Schneider et al. [9] compared the use of premounted vs. nonpremounted stents in acquired coronary artery stenosis. No significant differences in procedural success, restenosis, and complication rates were noted. Ease of use is the main reason for using premounted stents for coronary artery sclerosis. Ease of use explains the predilection for the larger premounted stents used in congenital heart disease as well.

Cheung et al. [10] advised against the use of self-expanding stents in growing children, in view of their significant neointimal ingrowth, unyielding design to over-dilation, and complications of distal migration. Accordingly, we found self-expanding stents almost exclusively implanted in patients older than 15 years. This age bias may explain that complication rates in this study for self-expanding stent and balloon-expandable stent did not significantly differ. Nearly all self-expanding stent-related complications were malpositioning, migration or embolization complications, thus supporting the findings of Cheung et al.

Subset analysis of the procedures that involved pulmonary artery stenting showed fewer overall complications with the use of premounted stents than with the use of nonpremounted stents, implying that premounted stents may be preferred for this patient group. Pulmonary artery stenting is nowadays preferred over balloon dilation alone, because of its higher immediate success rate and lower mid-term incidence of restenosis [11]. In this study, pulmonary artery stenting was the most common procedure. As with all stent implantations in growing children, the need for future stent enlargement should be anticipated. Stent re-dilatation in pulmonary arteries was found safe and effective for up to 3 years [12,13]. A large single centre study by Vitiello et al. reported 22 complications (six major) in 162 pulmonary artery stent implantations [5]. Their 14% overall complication rate compares well with our 17% outcome when combining the left and right pulmonary artery groups. The major complication rates are comparable as well; 6% in our study vs. 4% in the Vitiello study.

Centre inexperience with stenting tended to be associated with a higher major complication rate. McMahon et al. in their 12-year retrospective, single-centre study on pulmonary artery stenting did not observe any morbidity or mortality in the last 5 years of their experience [14]. The long time-frame, the

Table 7. Potential determinants of complications and death after stenting.

	Complications		**Major complications**		**Death**	
	OR (95% CI)	**P-value**	**OR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**
Experience	1.10 (0.65–1.87)	0.71	0.45 (0.19–1.06)	0.07	0.32 (0.10–2.11)	0.47
Premounted stent	0.56 (0.29–1.08)	0.08	0.34 (0.10–1.19)	0.09	0.67 (21.53 to 2.88)	0.55
Age >1 year	0.61 (0.32–1.19)	0.15	0.58 (20.40 to 1.57)	0.24	0.27 (0.06–1.23)	0.09
Self expandable stent	0.37 (20.72 to 1.45)	0.51	0.51 (0.06–4.19)	0.53	1.93 (0.22–16.96)	0.55
Post-surgical vs. native	0.77 (0.45–1.30)	0.32	0.43 (0.18–1.08)	0.07	0.25 (21.26 to 1.76)	0.75

many different operators, and the introduction of new indications, like ductal stenting in more recent years, could explain that the effect of centre inexperience in our study is not more striking. Combining experience from interventional paediatric cardiologists may contribute to lower complication rates.

Subset analysis of the procedures that involved stenting of aortic coarctations showed no differences in outcome after stenting of native vs. recurrent coarctations. This may be due to the small number of procedures. Vessel dissection was found twice, both in recurrent coarctation sites. Historically, there is more concern about vessel dissection in native stenoses [15]. This is not supported by our study. Other studies focus mainly on results and complications for combined native and recurrent coarctation. Johnston et al [16]. demonstrated a low complication rate for this combined group. With the introduction of newer, more flexible, round-edged stents over the past years, complication rates of stenting coarctations will probably further drop [17].

Stenting in pulmonary vein stenosis was documented five times only. As mentioned above, one of these implantations proved fatal as in-stent thrombosis resulted in embolization of a middle cerebral artery. This outcome illustrates the necessity of weighing the benefits and adverse effects of stent treatment in individual cases, especially in this rare and frequently progressive disease.

Stenting for congenital heart disease is disadvantageous mainly in growing children. To accommodate for vessel growth over time, re-dilatation may be necessary. Proper stent selection is therefore crucial. Several reports of stent

re-dilatation [13,18] describe this intervention to be mostly effective and safe. However, fatal vessel rupture in re-dilatation may occur [19].

In this study, 'parking' of malpositioned or embolized stents at benign positions other than the indicated stenotic sites, was considered to be a minor complication. Long-term complications in growing children, however, may be more severe, as they may need re-dilatation or possibly even surgical stent removal, even when these stents were implanted at an originally 'benign' site.

Study limitations

Retrospective studies of this kind carry the risk of underreporting of procedures and complications. Indeed, occasionally procedural information could not be retrieved. Also, complications like balloon rupture, haematoma, and arrhythmia without severe consequences are likely to have not been documented properly. Statistical analyses based upon group totals may have been biased by population subgroup differences as well. This is especially true for patient age, as certain implantation sites are fit for specific age ranges only. Sites therefore may differ in complication rates for this reason. Unfortunately, patient numbers in this study are too small to allow for multivariate analysis of potential risk factors for complications. Using the GEE method, we attempted to correct for multiple procedures in the same patient.

A final limitation of this study is its short-term approach. Apart from short-term complications, stent implantation will most likely also carry a risk of long-term complications. Restenosis, neointima proliferation, and stent fracture may occur [13,18,20]. Long-term complications were beyond the scope of our study, but definitely require further exploration.

Conclusions and recommendations

Current clinical practice should not underestimate complication rates of stenting in patients with congenital heart disease. It is difficult, however, to fully master every possible procedure seeing the vast diversity of indications combined with a relatively small population for stenting. Procedures therefore remain complication-sensitive. Our study nevertheless showed infrequent major complications and few fatal complications over a large variety of cases.

Especially, in these small patient groups with a wide range of diagnoses and anatomical substrates to be treated, adding on the experience of interventional paediatric cardiologists may limit the number of complications. In addition,

more precise visualization of the morphology gained using MRI or multislice CT will improve our understanding of the anatomical substrate to be treated and will thus enable more precise selection of stents.

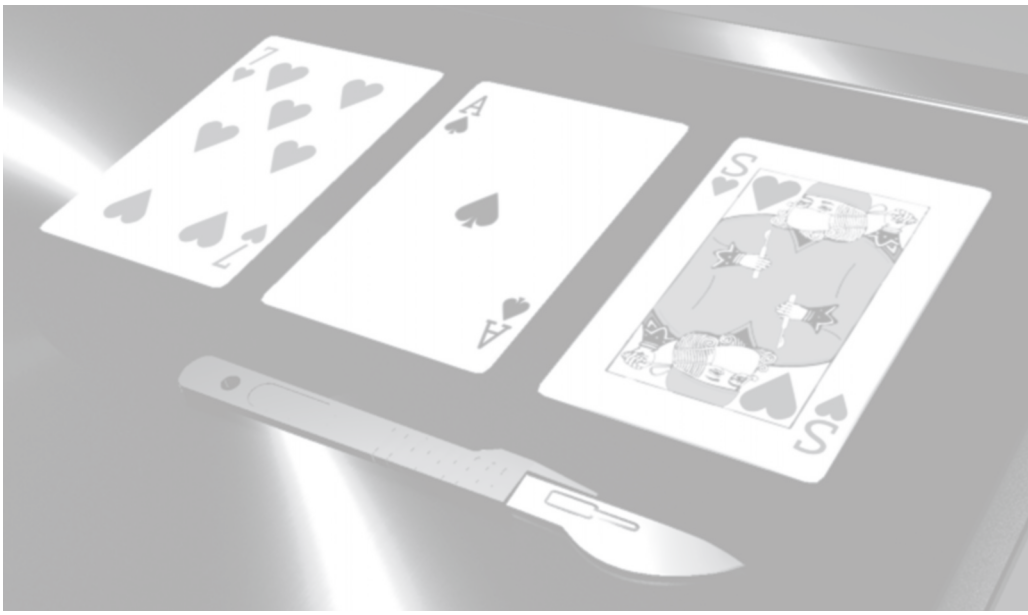
Because the use of non-premounted stents is associated with significantly higher complication rates, we recommend the use of premounted stents if an appropriate final diameter can be achieved. The industry should further expand the range of available premounted stents towards larger diameters, which could contribute to bringing down the incidences of stent migration and embolization.

REFERENCES

1. O'Laughlin MP, Perry SB, Lock JE, Mullins CE. Use of endovascular stents in congenital heart disease. Circulation 1991;83:1923–1939.
2. Gewillig M, Boshoff DE, Dens J, Mertens L, Benson LN. Stenting the neonatal arterial duct in duct dependent pulmonary circulation: new techniques, better results. J Am Coll Cardiol 2004;43:107–112.
3. Radtke WA, Waller BR, Hebra A, Bradley SM. Palliative stent implantation for aortic coarctation in premature infants weighing ,1500 g. Am J Cardiol 2002;90:1409–1412.
4. Mullins CE. Intravascular stents in congenital heart disease—general considerations. In: Cardiac catheterisation in congenital heart disease: pediatric and adult. Blackwell Futura; 2006. p. 537–660.
5. Vitiello R, McCrindle BW, David Nykanen D, Freedom RM, Benson LN. Complications associated with pediatric cardiac catheterisation. J Am Coll Cardiol 1998;32:1433–1440.
6. Agnoletti G, Boudjemline Y, Aggoun Y, Abdel-Massih T, Bonnet C, Bonnet D, Sidi D. Complications pre´coces de l'implantation de stents chez l'enfant et le jeune adulte ayant une cardiopathie conge´nitale. Arch Mal Coeur 2003;96:473–478.
7. McElhinney DB, Reddy VM, Moore P, Brook MM, Hanley FL. Surgical intervention for complications of transcatheter dilation procedures in congenital heart disease. Ann Thorac Surg 2000;69:858–864.
8. Boucek MM, Mashburn C, Kunz E, Chan KC. Ductal anatomy: a determinant of successful stenting in hypoplastic left heart syndrome. Pediatr Cardiol 2005;26:200–205.
9. Schneider TI, Ho¨pp HW, Vlaho D, Wassmer G, Fu¨ssl R, Fassbender S, Wehr G, Konz KH, Spa¨h F, Baer FM. Randomized comparison of mounted versus non premounted stents: the multicentre COMUS trial. Am Heart J 2003;145:e4.
10. Cheung Y, Sanatani S, Leung MP, Human DG, Chau AK, Culham JA. Early and intermediate-term complications of self-expanding stents limit its potential application in children with congenital heart disease. J Am Coll Cardiol 2000;354:1007–1015.
11. Allen HD, Beekman RH III, Garson A Jr, Hijazi ZM, Mullins CE, O'Laughlin MP, Taubert KA. Pediatric therapeutic cardiac catheterisation: a statement for healthcare professionals from the council on cardiovascular disease in the young. Circulation 1998;97:609–625.
12. Fogelman R, Nykanen D, Smallhorn JF, McCrindle BW, Freedom RM, Benson LN. Endovascular stents in the pulmonary circulation: clinical impact on management and medium term follow-up. Circulation 1995;92:881–885. 2714 M. van Gameren et al.
13. Duke C, Rosenthal E, Qureshi SA. The efficacy and safety of stent redilatation in congenital heart disease. Heart 2003;89:905–912.
14. McMahon CJ, ElSaid HG, Vincent JA. Grifka RG, Nihill MR, Ing FF, Fraley JK, Mullins CE. Refinements in the implantation of pulmonary arterial stents: impact on morbidity and mortality of the procedure over the last two decades. Cardiol Young 2002;12: 445–452.
15. Tynan M, Finley JP, Fontes V, Hess J, Kan J. Balloon angioplasty for the treatment of native coarctation: results of Valvuloplasty and Angioplasty of Congenital Anomalies Registry. Am J Cardiol 1990;65: 790–792.
16. Johnston TA, Grifka RG, Jones TK. Endovascular stents for treatment of coarctation of the aorta: acute results and follow-up experience. Catheter Cardiovasc Interv 2004;62:499–

505.
17. Cheatham JP. Stenting of coarctation of the aorta. Catheter Cardiovasc Interv 2001;54:112–126.
18. McMahon CJ, El-Said HG, Grifka RG, Fraley JK, Nihill MR, Mullins CE, neointimal proliferation. J Am Coll Cardiol 2001;38: 521–526.
19. Fejzic Z, Van Oort A. Fatal dissection of the descending aorta after implantation of a stent in a 19-year-old female with Turners's syndrome. Cardiol Young 2005;15:529–531.
20. Shaffer M, Mullins CE, Grifka RG, O'Laughlin MP, McMahon W, Ing FF, Nihill MR. Intravascular stents in congenital heart disease: short and long term results from a large singe-centre. J Am Coll Cardiol 1998;311:661–667.

CHAPTER 8



Seventeen years of adult congenital heart surgery: a single centre experience

Putman LM
Van Gameren M
Meijboom FJ
De Jong PL
Roos-Hesselink JW
Witsenburg M
Takkenberg JJM
Bogers AJJC

Eur J Cardiothorac Surg. 2009; 36: 96-104

ABSTRACT

Objective

With a growing number of children with congenital heart disease (CHD) reaching adulthood, an extensive experience with cardiac surgery in adults with CHD is accumulating. To increase insight in this patient category we report our 17-year single centre experience including predictors for adverse outcome and EuroSCORE performance.

Methods

Patients and operative characteristics of all consecutive adult CHD patients operated upon between January 1990 and January 2007 were collected. Categorisation was done according to the EACTS/STS congenital database. Early and late morbidity and mortality were assessed with follow-up extending up to 17 years. EuroSCORE performance was assessed.

Results

Nine hundred and sixty-three procedures were performed in 830 patients (mean age 39.3 years, 50.3% male). A total of 49% were re-do procedures, frequent procedures were for left heart lesions (37%), right heart lesions (31%) and septal defects (8%). The 51% primary procedures largely consisted of less complex procedures but also included 1.4% of tetralogy of Fallot repairs, 4.1% of aortic coarctation repairs and 2.7% of Ebstein's disease repairs. Thirty-day mortality was 1.5% (n = 14); predicted mortality by logistic EuroSCORE was 4.6%. c-index was 0.61 (95% CI 0.46-0.75). Major complications such as tamponade requiring intervention occurred in 3.2%, postoperative bleeding requiring re-exploration in 7.1% and renal insufficiency requiring dialysis in 4 (0.4%). Pulmonary hypertension was a strong predictor for short-term mortality; impaired ventricular function and cyanosis for long-term mortality. Overall 17-year survival was 71% (95% CI 61%-82%). Eighty percent of patients were in NYHA class I at last follow-up, 17% in II, 3% in III, 0% in IV.

Conclusions

Surgery in adult CHD patients can be performed with low operative mortality and good clinical outcome. EuroSCORE is not a good model for risk assessment in this group of patients.

INTRODUCTION

The innovations in cardiac surgery from the 1950s onward were crucial in the development of surgical treatment for congenital heart disease (CHD) patients towards it current status. Nowadays over 85% of children with CHD survive into adulthood, presenting themselves with various issues related to their disease. Problems such as degeneration of previously implanted prosthetic materials and residual defects occur. Management of this relatively new group of patients requires special attention, not only limited to general cardiac care, but also during other situations that increase cardiac load, for example pregnancy and anaesthesia for non-cardiac surgery [1,2]. In this setting, centres that provide care for grown-up patients with congenital heart disease are emerging.

The group of adult CHD patients is steadily growing and is expected to become, if not already, larger than the paediatric CHD population [3,4]. Based upon the vast diversity of diagnoses in CHD, it has been shown that larger patient load and surgical experience in paediatric cardiac surgery are beneficial for patient outcome [5]. This probably holds for adult congenital cardiac surgery as well. In daily practice, surgery in these patients is performed in a limited number of tertiary referral centres [1,6].

The number of reports on outcome of cardiac surgery in adults with CHD is limited, and most articles only provide a relatively short follow-up [7-12].

Although at paediatric age the Aristotle score and the RACHS-1 are being used for assessment of risk of mortality in paediatric congenital heart surgery, these models have not been validated in an adult congenital population, while the EuroSCORE has gained a widespread use for acquired heart disease in adults [13,14]. We sought to determine both shortterm and long-term predictors for adverse outcome, and assessed the performance of the EuroSCORE model in the adult congenital surgery population. In this light we report on our single centre experience, with a follow-up that extends up to almost 18 years.

METHODS

Data were collected retrospectively and consist of all consecutive adult (age 18 years and older at time of surgery) patients who had congenital cardiac surgery in a single tertiary care centre in the Netherlands between January 1, 1990 and

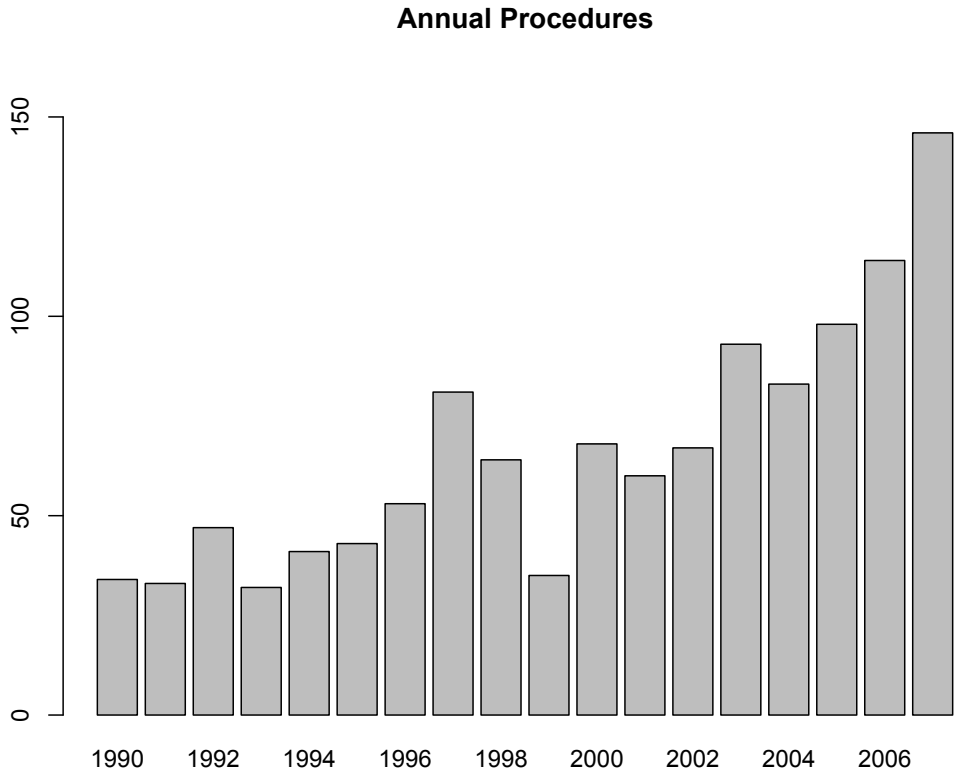


Figure 1. Number of annual procedures performed. The year 2007 was only included to show the rapidly growing number of procedures performed.

January 1, 2007. The adult congenital team meets weekly to discuss patients. During the 17-year period discussed in this article, this setup remained the same. Currently the team consists of dedicated congenital cardiothoracic surgeons treating both adult and paediatric patients and dedicated adult congenital cardiologists and paediatric cardiologists. Dedicated interventional cardiologists (collaborating paediatric and adult interventional cardiologists) are involved as well. A heart- and lung-transplant programme is available.

A database was designed in MicrosoftW Access 2000 (Microsoft Corporation, Redmond, WA) and contains data on demographics, preoperative risk factors, detailed surgical procedural information, in-hospital complications, and follow-up. Current survival status was collected from civil registries between November 7 and November 11 2007. Survival of the Dutch population was derived from the Central Bureau of Statistics (available at http://www.cbs.nl).

Table 1. Patient characteristics.

Preoperative status	Number (%) (n = 963)
Age at time of surgery	39.3 15.5 years (median 36, range 18.0-85.5)
Procedures per patient	
1	720 (86.7%)
2	91 (10.9%)
3	16 (1.9%)
4	2 (0.2%)
5	1 (0.1%)
Male sex	417 (50.3%)
Sinus rhythm	741 (79.1%)
Ventricular function	
Good	754 (78.5%)
Impaired	173 (18%)
Bad	15 (1.6%)
Pulmonary hypertension	36 (3.7%)
COPD	42 (4.4%)
Extracardiac arteriopathy	20 (2.1%)
Emergency procedure	13 (1.3%)
Cyanosis	87 (9.0%)
Smoking	98 (10.2%)
Diabetes	25 (2.6%)
Obesity	58 (6.2%)
NYHA	
I	298 (30.9%)
II	338 (35.1%)
III	174 (18.1%)
IV	32 (3.4%)
Active endocarditis	22 (2.3%)
Creatinine > 200 mmol/l	5 (0.5%)
Previous congenital cardiac surgery	476 (49.4%)
Additive EuroSCORE	5.1±2.3 (median 5, range 2—14)
Logistic EuroSCORE	5.4%±5.32 (95% CI 5.0—5.7)

A record of no-objection against the study was given by the Erasmus MC medical ethics committee.

Diagnoses, procedures and complications were classified according to the European Association of Cardio-Thoracic Surgery (EACTS) and Society of Tho-

Table 2. Original primary anatomic diagnoses, classified according to EACTS congenital database [15].

Anatomic diagnosis	**Number (%) (n = 963)**
Septal defects	332 (34.5%)
Atrial septal defect, secundum	196 (20.4%)
Ventricular septal defect	57 (5.9%)
Atrio-ventricular septal defect, partial	56 (5.8%)
Atrial septal defect, sinus venous	12 (1.2%)
Ventricular septal defect + aortic coarctation	3 (0.3%)
Atrio-ventricular septal defect, complete	3 (0.3%)
Truncus arteriosus	5 (0.5%)
Left heart lesions	242 (25.1%)
Aortic stenosis, valvar	162 (16.8%)
Aortic stenosis, subvalvar	28 (2.9%)
Aortic insufficiency	18 (1.9%)
Aortic insufficiency & stenosis	11 (1.1%)
Aortic stenosis, subvalvar + valvar	7 (0.7%)
Mitral valve stenosis	7 (0.7%)
Sinus of Valsalva aneurysm	5 (0.1%)
Mitral insufficiency	2 (0.2%)
Cardiomyopathy	2 (0.2%)
Right heart lesions	194 (20.1%)
Tetralogy of Fallot	123 (12.8%)
Ebstein's disease	31 (3.2%)
Pulmonary stenosis, valvar	18 (1.9%)
Pulmonary stenosis, valvar + subvalvar	8 (0.8%)
Pulmonary atresia, VSD including TOF/PA	8 (0.8%)
Pulmonary atresia, IVS	3 (0.3%)
Pulmonary stenosis, subvalvar	2 (0.2%)
Double-chambered right ventricle (DCRV)	1 (0.1%)
Thoracic arteries & veins	98 (10.2%)
Aortic coarctation	43 (4.5%)
Aortic coarctation + aortic valve stenosis	34 (3.5%)
Aortic aneurysm	3 (0.3%)
Anomalous origin of the left coronary artery from the pulmonary artery	3 (0.3%)
Coronary fistula	3 (0.3%)
Coronary artery anomaly, origin	1 (0.1%)
Patent ductus arteriosus	6 (0.6%)
Vascular ring	5 (0.5%)

Table 2. Continued.

Anatomic diagnosis	Number (%) (n = 963)
Transposition of the Great Arteries (TGA)	37 (3.8%)
TGA, IVS, including LVOTO	17 (1.8%)
TGA, VSD, including LVOTO	12 (1.4%)
Congenitally corrected TGA (ccTGA)	4 (0.4%)
ccTGA, VSD, including LVOTO	4 (0.4%)
Single ventricle	35 (3.6%)
Tricuspid atresia	18 (1.9%)
Double inlet left ventricle (DILV)	8 (0.8%)
Mitral atresia	4 (0.4%)
Double outlet right ventricle (DORV)	4 (0.4%)
DILV & DORV	1 (0.1%)
Electrophysiologic	16 (1.7%)
Arrhythmia, heart block, congenital	16 (1.7%)
Pulmonary venous anomalies	7 (0.7%)
Partial anomalous pulmonary venous connection	5 (0.5%)
Cor triatriatum	2 (0.2%)
Miscellaneous	2 (0.2%)
Aneurysm, atrial	2 (0.2%)

racic Surgeon's (STS) congenital heart surgery nomenclature [15]. Classification of these was based upon the data entered by the attending surgeon. Data were collected using our hospital information system (HIS) and patient charts including data sheets used during multidisciplinary preoperative patient evaluation.

Procedures or diagnoses checked to be not congenital, despite previous coding, were excluded.

Any additional congenital cardiac procedures that had been performed upon included patients, but in another institution, outside the study period or whenever the patient was under the age of 18, were only noted in terms of diagnosis requiring surgery; no additional procedural data were collected.

Collected data included: clinical status, ventricular function as ejection fraction (EF) assessed by angiography, magnetic resonance imaging or echocardiography (classified as poor (EF <30%), moderate (EF 30-50%) or good (EF >50%)) [16]. New York Heart Association (NYHA) classification was collected when available or retrospectively classified according to information in patient records. Chronic obstructive pulmonary disease (COPD) was defined as longterm use of

bronchodilators or steroids for lung disease, diabetes when diagnosis was previously made, pulmonary hypertension (PHT) as systolic pulmonary artery pressure >60 mmHg, extracardiac arteriopathy as any one or more of the following: claudication, carotid occlusion or stenosis >50%, and previous or planned intervention on the abdominal aorta, limb arteries or carotids.

Predicted additive and logistic EuroSCORE mortality rates were retrospectively calculated from available variables [16,17]. Data on neurologic dysfunction, unstable angina and recent myocardial infarction were not collected. Missing data in any of the EuroSCORE variables were furthermore regarded as not having that particular risk factor to be able to calculate a score [17].

Outcome and complications during hospital stay were collected. Complications were classified as follows: arrhythmia as any cardiac rhythm other than sinus rhythm; heart failure as a clinical presentation of right or left heart failure; pericardial effusion as requiring intervention; prolonged ventilation as postoperative ventilation >2 days; infections as positive cultures from either superficial or deep wound swabs; postoperative bleeding as requiring surgical intervention; sternum left open; pneumothorax as requiring drainage; low output as cardiac output requiring inotropic medication; mechanical circulatory support as the need for postoperative mechanical circulatory support; pulmonary hypertensive crisis as systolic pulmonary arterial pressure > systolic systemic arterial pressure.

Readmission and reason for readmission to any hospital within 30 days of the procedure was collected.

Collected follow-up data included NYHA class, arrhythmia (bundle-branch blocks were not regarded as arrhythmia), heart-transplant and pacemaker implantation. All were collected at the latest follow-up date available.

Statistical analysis

Statistical analysis was done using SPSS 15.0 (SPSS, Chicago, IL). Continuous data are presented as mean SD, and median with range. Categorical data are presented as proportions. Completeness of follow-up was calculated according to Clark et al. [18]. Logistic regression was used to identify risk factors for early mortality, both 30-day and 1- year, and Cox regression analysis was used to identify risk factors for late outcome. Potential risk factors were entered into multivariate analysis if $p > 0.10$. For multivariate analysis backward conditional stepping and 5-9 events per variable, as proposed by Vittinghoff and McCulloch were used

[19]. A p value < 0.05 was considered statistically significant.

Kaplan-Meier plots were used to demonstrate overall survival and survival of certain subgroups of diagnoses and procedures. Cumulative survival was analysed using the Kaplan-Meier method. For survival analysis of the overall cohort the survival of a patient started at the time of first operation and ended at the time of death (event) or at the last follow-up (censoring) (patient based). The survival analysis of subgroups started at the time of each procedure in that subgroup and ended with patient death (event) or last follow-up (censoring) (procedure based).

Discriminatory power of the EuroSCORE among this group of patients was assessed using the c-index (area under the receiver operating characteristic (ROC) curve) with 95% confidence intervals (95% CI).

RESULTS

Procedural and patient characteristics

Nine hundred and sixty-three surgical procedures were performed in 830 patients. Figure 1 highlights the growing number of annual procedures. Age at time of surgery was 39.3±15.5 years (median 36, range 18.0-85.5). Preoperative patient characteristics are shown in Table 1.

Four hundred and eighty-seven (50.6%) of the procedures were primary and 476 (49.4%) had been preceded by at least one congenital cardiac surgical procedure. Four hundred and seventeen male (50.3%) and 413 female (49.7%) patients were included. Complete data on anatomic diagnosis and surgical procedures are listed in Tables 2 and 3.

Intensive care unit length of stay was 47.3±96 h (median 27 h, range 0-1514 h). Hospital length of stay was 14.4±13.6 days (median 11 days, range 0-224 days).

Thirty-three (3.4%) procedures included concomitant coronary surgery, of which 31 were planned and 2 were performed to correct damaged coronary arteries during the planned procedure. Extra-corporal circulation (ECC) was used during 869 (90.2%) procedures. Procedures without ECC included: aortic coarctation repair, pacemaker procedures, vascular ring repairs and various shunt procedures.

Table 3. Actual procedures performed, classified according to EACTS congenital database [15]. Concomitant procedures are shown in italic.

Main procedures	**Number (%) (n = 963)**
Left heart lesions	366 (38%)
Aortic valve replacement, mechanical *10x LVOT, 3x VSD, 4x MVP, 3x MVR, 4x CABG, 3x ASD*	77 (8.0%)
Aortic valve replacement, homograft *5x LVOT, 1x CABG, 1x PAA*	67 (7.0%)
Ross procedure *2x PAA, 2x aorta plasty, 1x CABG, 2x LVOT*	51 (5.3%)
Aortic root replacement, mechanical (Bentall) *1x ASD, 2x LVOT, 1x MVR, 2x pulmonary mechanical root, 3x PHG, 3x MVP, 1x CABG, 2x prosthetic aortic arch*	46 (4.8%)
Mitral valvuloplasty *5x ASD, 1x AHG 5x CABG, 1x PDA, 1x MAZE*	41 (4.0%)
Mitral valve replacement *7x ASD, 6x TVP, 1x AVSD, 2x aortic valve, 1x MAZE, 2x CABG, 1x PDA*	31 (3.1%)
Subvalvular aortic repair *2x AVRM, 1x PHG, 1x VSD, 1x RVOT, 1x Shunt takedown, 1x aorta plasty, 1x CABG*	22 (2.3%)
Aortic valve replacement, Bioprosthetic *2x MVP, 3x ASD, 1x CABG*	17 (1.7%)
Aortic valvuloplasty *1x PAA, 2x LVOT, 1x MVP*	6 (0.6%)
Sinus of Valsalva, aneurysm repair	3 (0.3%)
Transplant, heart	3 (0.3%)
Aortic root replacement, valve sparing *1x MVP*	2 (0.2%)
Septal defects	255 (26.5%)
Atrial septal defect repair *7x MAZE, 15x CABG, 3x PAP, 13x MVP, 10x TVP, 1x PVP*	176 (18.3%)
Ventricular septal defect repair *7x ASD, 7x PHG, 7x RVOT, 1x LVOT, 2x DCRV, 2x AVRM, 1x CABG*	53 (5.5%)
Partial atrioventricular septal defect repair *1x LVOT, 4x TVP, 4x ASD, 1x AHG*	26 (2.7%)
Right heart lesions	180 (18.8%)
Pulmonary valve replacement *12x VSD, 30x RVOT, 7x PAP, 10x ASD, 2x AVRM, 2x TVR, 10x TVP, 4x PM*	121 (12.6%)
Ebstein's repair *2x Glenn, 11x ASD, 1x PM*	17 (1.8%)
Tricuspid valve replacement	12 (1.2%)
Tetralogy of Fallot primary repair	11 (1.1%)
Valvuloplasty tricuspid valve *4x ASD, 1x MAZE, 2x MVR, 1x MVP*	8 (0.9%)
RVOT procedure *2x ASD, 1x VSD, 1x LVOT, 1x DCRV*	6 (0.6%)
Pulmonary artery plasty	3 (0.3%)
Double chambered right ventricle (DCRV) repair	1 (0.1%)
Occlusion MAPCA(s)	1 (0.1%)

[a] Include 2 re-do for sclerosed homografts and 8 conversions (2 extracardiac and 6 lateral tunnel).

Table 3. Continued.

Main procedures	Number (%) (n = 963)
Electrophysiologic	62 (6.4%)
Pacemaker procedure	62 (6.4%)
Thoracic arteries & veins	58 (6.0%)
Coarctation repair, end-to-end	25 (2.6%)
Coarctation repair, interposition graft	17 (1.8%)
Coronary artery fistula ligation	3 (0.3%)
ALCAPA *1x coronary fistula*	3 (0.3%)
PDA closure *1x TVP*	2 (0.4%)
Vascular ring repair	3 (0.3%)
Aortic aneurysm repair	3 (0.3%)
Coronary artery bypass for anomalous coronary artery	1 (0.1%)
Single ventricle	14 (1.4%)
Fontan, revision or conversion (re-do Fontan)[a]	10 (1.0%)
Fontan, TCPC, Lateral Tunnel *1x MAZE, 1x PM*	4 (0.4%)
Palliative procedures	8 (0.8%)
Glenn procedure *1x PDA, 2x Blalock takedown, 1x PM, 1x PAP*	3 (0.3%)
Shunt, modified Blalock-Taussig	3 (0.3%)
Shunt, central	1 (0.1%)
Shunt, ligation and takedown	1 (0.1%)
Transposition of the great arteries	8 (0.8%)
Atrial baffle procedure, Mustard revision *2x RVOT, 2x PM, 1x ASD*	8 (0.8%)
Pulmonary venous anomalies	6 (0.6%)
Partial anomalous pulmonary venous connection repair	4 (0.4%)
Cor triatriatum repair	2 (0.2%)
Miscellaneous procedures	6 (0.4%)
Aneurysm, atrial, repair	2 (0.2%)
Cardiac tumor resection *2x ASD*	2 (0.2%)
Pulmonary embolectomy, acute	1 (0.1%)
Pulmonary embolectomy, chronic *1x ASD*	1 (0.1%)

ASD: atrial septal defect closure, VSD: ventricular septal defect closure, AVSD: atrio-ventricular septal defect repair LVOT: left ventricular outflow tract procedure, RVOT: right ventricular outflow tract procedure, PAA: prosthetic ascending aorta MVP: mitral valvuloplasty, MVR: mitral valve replacement, TVP: tricuspid valvuloplasty, TVR: tricuspid valve replacement, PVP: pulmonary valvuloplasty, PDA: patent ductus arteriosus closure, PHG: pulmonary homograft procedure, AHG: aortic homograft procedure, AVRM: mechanical aortic valve replacement PAP: plasty of the pulmonary artery(ies), ALCAPA: anomalous origin of the left coronary artery from the pulmonary artery, PM: pacemaker procedure. Individual combinations of concomitant procedures are not shown.

Table 4. Postoperative complications.

Complications	Number (%) (n = 963)
Death (30-day)	14 (1.5%)
Arrhythmia	175 (18.2%)
Atrial fibrillation	120 (12.5%)
Atrial flutter	5 (0.5%)
Supra-ventricular tachycardia	14 (1.5%)
Ventricular tachycardia (including non-sustained)	9 (0.9%)
Ventricular fibrillation	10 (1.0%)
AV-block (3rd degree/complete only)	7 (1.5%)
Heart failure	26 (2.7%)
Pericardial effusion/tamponade	31 (3.2%)
Prolonged ventilation (>2 days)	37 (3.8%)
Infections	39 (4.0%)
Deep	19 (2.0%)
Superficial	9 (0.9%)
Endocarditis	11 (1.1%)
Sternum left open	2 (0.2%)
Pneumothorax	23 (2.4%)
Bleeding	68 (7.1%)
Cardiac arrest	16 (1.7%)
Low output	14 (1.5%)
Mechanical circulatory support	0
Pulmonary hypertensive crisis	1 (0.1%)
Renal insufficiency	4 (0.4%)
Pleural effusion	23 (2.4%)
Pneumonia	8 (0.8%)

Early outcome

Fourteen (1.5%) patients died within 30 days of the procedure (in the first 9 years of the series the early mortality was 1.6% and in the last 8 years 1.3%). Two patients died intra-operatively, 11 patients died postoperative in-hospital and one died after discharge. Mortality occurred after three atrial septal defect (ASD) closure procedures (two isolated ASD closure procedures and one with concomitant mitral valve replacement), two pulmonary homograft procedures, two mechanical aortic valves, one aortic homograft, one autograft procedure, three Bentall procedures, one Ebstein's disease repair, and one pulmonary thrombec-

tomy with tricuspid valve replacement and PFO closure. Causes of death in these patients were: heart failure (n = 4), cardiac tamponade (n = 2), suture dehiscence (n = 2), gastro-intestinal bleeding (n = 1), brain death (n = 1), low output due to a coronary anastomosis problem (n = 1), aspiration bronchopneumonia (n = 1), persistent pulmonary hypertension (n = 1) and sepsis after gastro-intestinal ischaemia (n = 1).

One or more complications occurred after 319 (33.1%) procedures. Complication rates are presented in Table 4.

Logistic regression analysis results for 30-day (n = 14) and one-year mortality (n = 26) are shown in Table 5.

Readmission within 30 days of discharge was necessary in 72 patients (7.5%), mostly caused by late cardiac tamponade and arrhythmia (20% and 25.7%, respectively). Other reasons were: congestive heart failure, coagulation problems and respiratory problems.

Late outcome

Mean follow-up was 6.8 5.1 years (median 6.3 years, range 0-17.8 years). Total follow-up was 6585 patient-years, potential follow-up was 7361 patient-years, the completeness of follow-up index is 0.89.

Cox regression results for long-term mortality (n = 85) are shown in Table 5. Kaplan-Meier curves are shown in Figure 2. Figure 2a shows the overall group and the 18-year survival of the age-matched Dutch population, the latter being 95.4%. Figure 2b shows various main groups of procedures.

Pre- and postoperative NYHA class distribution is shown in Figure 3.

A total of 8.5% of patients had any arrhythmia at followup. Three patients had a heart transplant and one patient had a heart-lung transplantation. A pacemaker had been implanted in 9.8% of the patients at follow-up.

EuroSCORE performance

ROC curves were constructed to test EuroSCORE performance in the adult CHD patients. When used to predict 30- day mortality for the entire cohort, the c-index was 0.586 (95% CI 0.451-0.720) for the logistic EuroSCORE and 0.599 (95% CI 0.468-0.730) for the additive EuroSCORE. When the EuroSCORE was used to predict 1-year mortality, c-indices were: 0.670 (95% CI 0.562-0.778) for the logistic and 0.686 (95% CI 0.581-0.792) for the additive EuroSCORE.

DISCUSSION

This study demonstrates that a growing experience with adult congenital heart surgery is associated with a low mortality risk and acceptable risk of complications. Late survival however is markedly impaired compared to that of the general age-matched Dutch population, and varies widely between the different diagnoses. The EuroSCORE model did not adequately predict mortality.

It can be appreciated from Table 3 that a large diversity of procedures is performed in adult congenital heart surgery and a major part of these procedures is accompanied by one or more concomitant procedures. This underlines the complexity of these patients.

Notable is the number of re-do procedures within this 17-year period; over 11% of patients had two or more procedures as an adult.

Early outcome

Overall 30-day mortality was 1.5%. This is comparable to the results shown among the tertiary care adult congenital heart surgery centres study by Niwa et al. [20]. When comparing our results to the larger multicentre studies [9,11,12], it can be noted that early mortality is low, ranging from 1.7% to 3.1%. Individual centres in these multicentric studies show different results, with mortality ranging from 0% to 15.3% in Stellin et al. [11]. No comment is made on these differences, but this may be partially related to the complexity of patients. Median number of procedures per centre in their study was 45 (range 11-202), indicating that some centres may lack a sufficient number of adult CHD patients to keep up experience which might contribute to different study results. Patient load and experience can be expected to influence results in this group of patients. This could suggest performing adult CHD procedures only in centres with an acceptable number of procedures.

Only two patients died after procedures classified by the Aristotle scoring system as simple [14]. The other 12 deaths occurred after procedures that can be regarded as complex, ranging from an Aristotle score of 6.5 to 10.3 (on a scale of 1.5-15).

As seen in Table 5, predictors for mortality were preoperative pulmonary hypertension, non-sinus rhythm, impaired ventricular function, critical preoperative status, preoperative cyanosis and age at surgery. Pulmonary hyperten-

Table 5. Predictors for early and late mortality; 30-day and one-year are logistic regression, long-term is Cox regression. Odds ratios are presented with 95% CI between brackets.

	30-day		One-year		Long-term	
	Univariate OR	Multivariate OR	Univariate OR	Multivariate OR	Univariate OR	Multivariate OR
Pulmonary HT	7.82 (2.08-29.38)	7.72 (1.99-29.86)	7.20 (2.54-20.38)	7.59 (2.55-22.59)	NS	-
Arrhythmia	2.02 (1.01-4.03)	1.92 (0.90-4.07)	1.79 (1.04-3.07)	NS	2.02 (1.44-2.82)	1.97 (1.34-2.91)
Age at surgery	1.03 (1.00-1.06)	1.03 (0.99-1.06)	1.03 (1.01-1.06)	1.03 (1.00-1.05)	1.05 (1.03-1.06)	1.04 (1.02-1.06)
Impaired VEF	NS	-	3.71 (1.69-8.15)	3.61 (1.60-8.15)	4.02 (2.52-6.42)	3.71 (2.27-6.07)
Active endocarditis	NS	-	6.30 (1.74-22.80)	NS	NS	-
Preoperative ventilation	NS	-	8.04 (2.21-29.23)	6.67 (1.66-26.86)	NS	-
Additive EuroSCORE	NS	-	1.32 (1.14-1.54)	NS	1.29 (1.17-1.42)	NS
Logistic EuroSCORE	NS	-	1.10 (1.04-1.16)	NS	1.08 (1.04-1.14)	NS
NYHA III or IV	NS	-	3.28 (1.49-7.21)	NS	3.01 (1.89-4.82)	-
Cyanosis	NS	-	NS	-	1.90 (1.07-3.37)	2.37 (1.30-4.33)
Diabetes	NS	-	NS	-	NS	-
COPD	NS	-	NS	-	4.82 (2.39-9.74)	2.57 (1.19-5.52)
Creatinine	NS	-	NS	-	NS	-
Smoking	NS	-	NS	-	2.31 (1.18-4.55)	3.17 (1.58-6.36)
Extracardiac art.	NS	-	NS	-	NS	-
Male sex	NS	-	NS	-	NS	-

NS: non-significant, -: not tested multivariate, HT: hypertension, VEF: ventricular function.

sion was especially a strong predictor for early mortality, both 30-day and 1-year. Analyses however were done with limited number of events; the use of larger cohorts, and thus more events, may reveal additional predictors for mortality.

Postoperative complications occurred frequently, but largely consisted of relatively minor and (reversible) complications such as atrial fibrillation. Where atrial fibrillation may not necessarily be a surgical complication, it is the most frequent cardiologic reason for admission in CHD patients [21]. The number of major complications, such as (suspected) tamponade and persistent blood loss was limited.

Long-term outcome

Completeness of follow-up was 0.89, demonstrating adequately complete data on long-term survival.

Differences were observed regarding long-term survival among several subgroups. The excellent survival of the group of thoracic arteries and veins, comparable to that of the Dutch population, is striking. The largest groups, septal defects, right-heart lesions and left-heart lesions, show a good survival of over 80% at 15 years. The other groups were small and were therefore not plotted. The impaired survival of various subgroups may be caused by the lack of an actual cure of the disease. Several defects cause abnormal volume, flow and pressure conditions on the heart and great vessels. Lesion-specific post-treatment outcomes have been described elsewhere, with a majority of lesions, even after treatment, remaining at risk for arrhythmias, sudden death and (premature) atherosclerosis and heart failure [1].

Risk factors for long-term mortality were preoperative impaired ventricular function, arrhythmia, cyanosis, COPD and smoking. The first three are associated with the disease itself, whereas COPD and smoking are known factors for mortality.

EuroSCORE performance

The EuroSCORE system was not developed for adult congenital cardiac surgery, and included low numbers of congenital patients, ASD closure for example was performed in only 1.2% of the included patients. Jacquet et al. have previously tested EuroSCORE in predicting outcome in a group of adult CHD patients. The interquartile range (IQR) of both additive and logistic EuroSCORE in our study is comparable to the IQR in their study [22]. Their conclusion on the overestima-

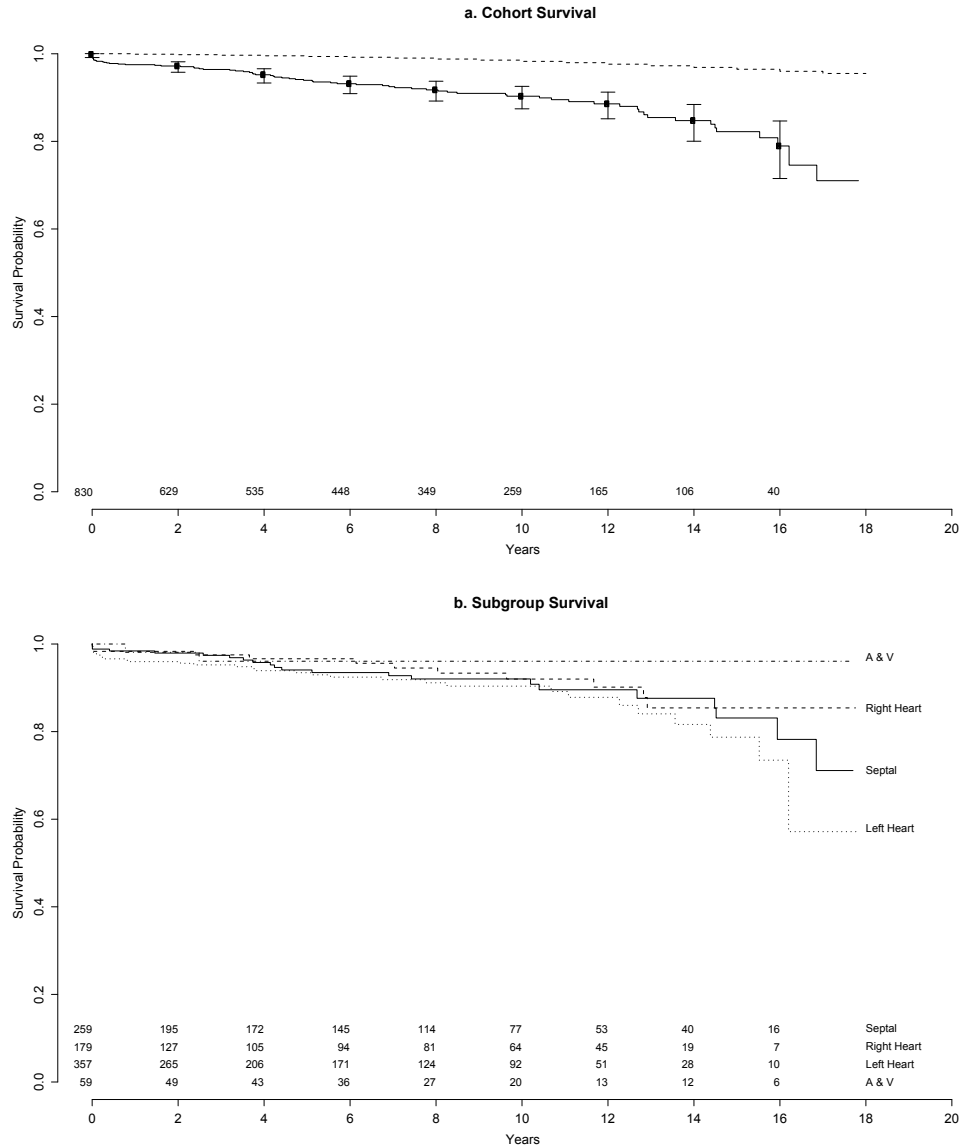


Figure 2. Kaplan-Meier curves for survival: (a) entire cohort (solid line with 95% CI bars) and age-matched Dutch population (dotted line) and (b) survival of the four largest subgroups of procedures.
A&V = thoracic arteries & veins.

tion of the EuroSCORE in this particular patient group is supported by our data. Price et al. did a study in a similar population and found no statistical difference

between predicted and observed mortality, but did find an over-estimation in low-risk patients and an underestimation in high-risk patients [23]. Both Jacquet and Price did limited analysis in testing the performance of the EuroSCORE model in these patients, lacking discriminatory testing using the c-statistic.

The use of EuroSCORE is not useful in this particular group in predicting operative mortality; discriminatory power is lacking with 95% CIs of the c-index including 0.50. A c-index of 0.75 or larger is required for an accurate test. The lack of inclusion of an operative complexity parameter and the modest number of end points (low observed mortality) might contribute to this finding.

Clinical status at last follow-up

NYHA class distribution as a very basic measure of quality of life was improved by the surgical procedure in a majority of patients. NYHA class deterioration was found in few patients, suggesting good results not only in terms of survival but also in terms of symptoms. The less than 5% missing follow-up NYHA classes mostly consisted of information lacking to reliably classify those patients. Quality of life, in a subgroup of operated adult ventricular septal defect patients in our institution, was shown to be nearly equal to that of the general population [24]. Comparable positive results are shown in a more diverse adult CHD population by Moons et al. [25]. The relatively low number of arrhythmias at follow-up, 8.5%, might be due to paroxysmal nature of some of these arrhythmias.

Limitations

Limitations of this study arise from the retrospective single tertiary centre setup. There may be to some degree an underestimation of the preoperative risk factors and complications rates, death excluded. In addition bias may have occurred due to the 89% completeness of follow-up. A further limitation of our study is the inclusion of only the surgically treated patients, thus providing only a selection of the total group of adult congenital heart disease patients.

NYHA classification was collected only at latest follow-up available and is therefore not completely comparable among the group, especially when the largest number of operations was done in the latter part of the study period, giving shorter follow-up and possibly showing better results before deterioration occurs in that group.

The lacking points of the EuroSCORE in unstable angina, recent myocardial infarction and neurological dysfunction are not expected to have changed the

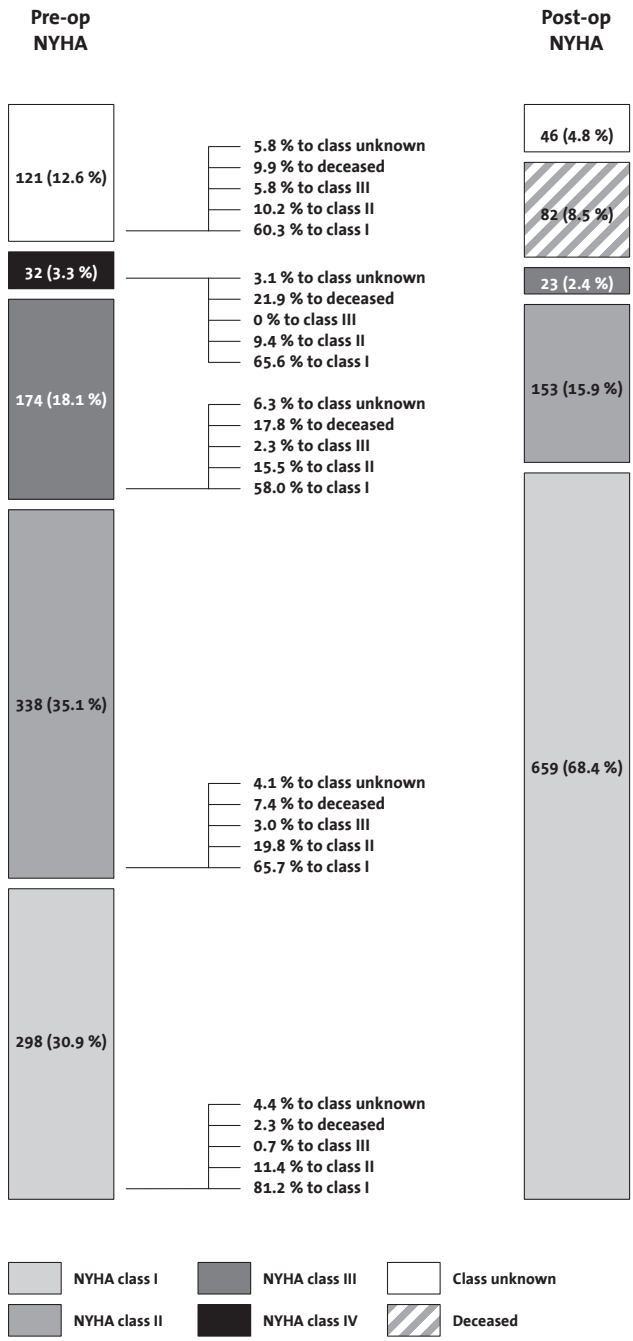


Figure 3. NYHA-class distribution. In the left bar preoperative NYHA class and in the right bar NYHA at follow-up are shown. The percentages in the middle represent the redistribution from preoperative to follow-up NYHA class.

results substantially. If they would have been included, only a small number would be relevant because these factors are uncommon in adult CHD patients. Age distribution and low rates of diabetes and obesity make coronary lesions less likely and the low number of procedures with concomitant coronary surgery also reflects the low number of coronary lesions. The preoperative neurological dysfunction may again be expected to be very low, due to the low number of surgical procedures carried out in such patients.

Conclusions

Our results regarding surgery for adult CHD show that low mortality and limited serious morbidity can be achieved in the setting of a tertiary referral centre. The EuroSCORE is not a good predictor of mortality in this group of adult CHD patients.

Further research is necessary to address issues such as decision-making in this adult group of patients, optimal timing of intervention and improved prediction models, despite the large variation of anatomic diagnoses and procedures.

REFERENCES

1 Deanfield J, ThaulowE,Warnes C,Webb G, Kolbel F, Hoffman A, Sorenson K, Kaemmer H, Thilen U, Bink-Boelkens M, Iserin L, Daliento L, Silove E, Redington A, Vouhe P, Priori S, Alonso MA, Blanc JJ, Budaj A, Cowie M, Deckers J, Fernandez Burgos E, Lekakis J, Lindahl B, Mazzotta G, Morais J, Oto A, Smiseth O, Trappe HJ, KleinW, Blomstrom-Lundqvist C, de Backer G, HradecJ,MazzottaG,ParkhomenkoA,PresbiteroP,Torbicki A.Management of grown up congenital heart disease. Eur Heart J 2003;24(11):1035-84.

2. Warnes CA, Williams RG, Bashore TM, Child JS, Connolly HM, Dearani JA, del Nido P, Fasules JW, Graham Jr TP, Hijazi ZM, Hunt SA, King ME, Landzberg MJ, Miner PD, Radford MJ, Walsh EP, Webb GD, Smith Jr SC, Jacobs AK, Adams CD, Anderson JL, Antman EM, Buller CE, Creager MA, Ettinger SM, Halperin JL, Hunt SA, Krumholz HM, Kushner FG, Lytle BW, Nishimura RA, Page RL, Riegel B, Tarkington LG, Yancy CW. ACC/AHA 2008 guidelines for the management of adults with congenital heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing committee to develop guidelines on the management of adults with congenital heart disease). Developed in collaboration with the American Society of Echocardiography, Heart Rhythm Society, International Society for Adult Congenital Heart Disease, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol 2008;52(23):e1-21.

3. Daebritz SH. Update in adult congenital cardiac surgery. Pediatr Cardiol 2007;28(2):96-104.

4. Monro J. The changing state of surgery for adult congenital heart disease. Heart 2005;91(2):139-40.

5. Hannan EL, Racz M, Kavey RE, Quaegebeur JM, Williams R. Pediatric cardiac surgery: the effect of hospital and surgeon volume on in-hospital mortality. Pediatrics 1998;101(6):963-9.

6. Daenen W, Lacour-Gayet F, Aberg T, Comas JV, Daebritz SH, Di Donato R, Hamilton JR, Lindberg H, Maruszewski B, Monro J. Optimal structure of a congenital heart surgery department in Europe. Eur J Cardiothorac Surg 2003;24(3):343-51.

7. Chatzis AC, Giannopoulos NM, Tsoutsinos AI, Chryssostomidis G, Panagiotou C, Zavaropoulos P, Kirvasilis G, Sarris GE. Successful surgical correction of congenital heart disease in adults: seven years' experience. Hellenic J Cardiol 2005;46(2):128-34.

8. Dore A, Glancy DL, Stone S, Menashe VD, Somerville J. Cardiac surgery for grown-up congenital heart patients: survey of 307 consecutive operations from 1991 to 1994. Am J Cardiol 1997;80(7):906-13.

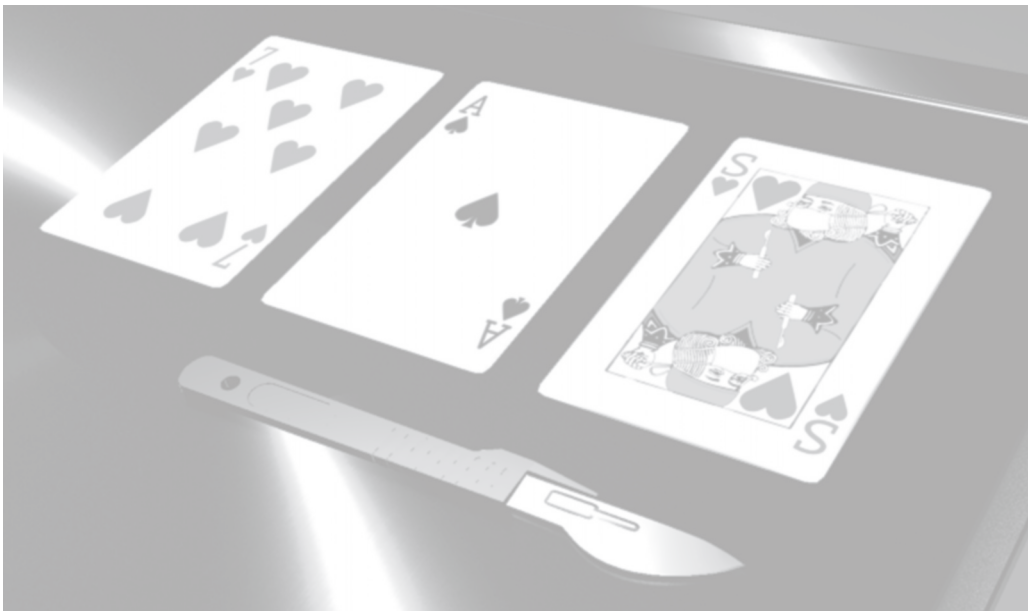
9. Padalino MA, Speggiorin S, Rizzoli G, Crupi G, Vida VL, BernabeiM, Gargiulo G, GiambertiA, Santoro F, Vosa C, PacileoG,Calabro R, Daliento L, Stellin G. Midterm results of surgical intervention for congenital heart disease in adults: an Italian multicenter study. J Thorac Cardiovasc Surg 2007;134(1): 106-13. 13 e1-9.

10. Srinathan SK, Bonser RS, Sethia B, Thorne SA, Brawn WJ, Barron DJ. Changing practice of cardiac surgery in adult patients with congenital heart disease. Heart 2005;91(2):207-12.

11. Stellin G, Vida VL, Padalino MA, Rizzoli G. Surgical outcome for congenital heart malformations in the adult age: a multicentric European study. Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu 2004;7:95-101.

12. Vida VL, BerggrenH, BrawnWJ,DaenenW, Di Carlo D, Di Donato R, Lindberg HL, Corno AF, Fragata J, Elliott MJ, Hraska V, Kiraly L, Lacour-Gayet F, Maruszewski B, Rubay J, Sairanen H, Sarris G, Urban A, Van Doorn C, Ziemer G, Stellin G. Risk of surgery for congenital heart disease in the adult: a multicentered European study. Ann Thorac Surg 2007;83(1):161-8.
13. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. J Thorac Cardiovasc Surg 2002;123(1):110-8.
14. Lacour-Gayet F, Clarke D, Jacobs J, Comas J, Daebritz S, Daenen W, Gaynor W, Hamilton L, Jacobs M, Maruszsewski B, Pozzi M, Spray T, Stellin G, Tchervenkov C, Mavroudis C. The Aristotle score: a complexityadjusted method to evaluate surgical results. Eur J Cardiothorac Surg 2004;25(6):911-24.
15. Mavroudis C, Jacobs JP. Congenital Heart Surgery Nomenclature and Database Project: overview and minimum dataset. Ann Thorac Surg 2000;69(4 Suppl.):S2-17.
16. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16(1):9-13.
17. Roques F, Michel P, Goldstone AR, Nashef SA. The logistic EuroSCORE. Eur Heart J 2003;24(9):881-2.
18. Clark TG, Altman DG, De Stavola BL. Quantification of the completeness of follow-up. Lancet 2002;359(9314):1309-10.
19. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007;165(6):710-8.
20. Niwa K, Perloff JK, Webb GD, Murphy D, Liberthson R, Warnes CA, Gatzoulis MA. Survey of specialized tertiary care facilities for adults with congenital heart disease. Int J Cardiol 2004;96(2):211-6.
21. Somerville J. Management of adults with congenital heart disease: an increasing problem. Annu Rev Med 1997;48:283-93.
22. Jacquet L, Vancaenegem O, Rubay J, Laarbaui F, Goffinet C, Lovat R, Noirhomme P, El Khoury G. Intensive care outcome of adult patients operated on for congenital heart disease. Intensive Care Med 2007;33(3): 524-8.
23. Price S, Jaggar SI, Jordan S, Trenfield S, Khan M, Sethia B, Shore D, Evans TW. Adult congenital heart disease: intensive care management and outcome prediction. Intensive Care Med 2007;33(4):652-9.
24. Bol Raap G, Meijboom FJ, Kappetein AP, Galema TW, Yap SC, Bogers AJ. Long-term follow-up and quality of life after closure of ventricular septal defect in adults. Eur J Cardiothorac Surg 2007;32(2):215-9.
25. Moons P, Van Deyk K, De Bleser L, Marquet K, Raes E, De Geest S, Budts W. Quality of life and health status in adults with congenital heart disease: a direct comparison with healthy counterparts. Eur J Cardiovasc Prev Rehabil 2006;13(3):407-13.

CHAPTER 9



Risk stratification for adult congenital heart surgery

Van Gameren M
Putman LM
Takkenberg JJM
Bogers AJJC

Eur J Cardiothorac Surg. Published online ahead of print on September 14th 2010

ABSTRACT

Objectives

At this moment, no risk stratification models are available for adult congenital cardiac surgery. This study aims to identify a suitable stratification tool for the adult congenital heart surgery population. Pediatric congenital cardiac surgery score models were therefore tested in an adult congenital population. In addition, an age component was added to these models and performance was compared with the original score systems.

Methods

The RACHS-1, Basic Aristotle Score, STS-EACTS Score and Comprehensive Aristotle Score were calculated for all adult patients who underwent congenital cardiac surgery between January 1990 and January 2007 in a single center (N=963). In addition, an age component was added to these models. Discrimination was then tested for all models with and without the age component.

Results

Application of the original pediatric risk scores resulted in c statistics for 30 day mortality of 0.60, 0.60, 0.60, and 0.66 respectively. Combining these models with the age component resulted in significantly higher c statistics of 0.69, 0.70, 0.69, and 0.76 respectively. Age as a sole predictor already resulted in a c statistic of 0.67. Comparable results were found for one year mortality.

Conclusions

The discriminatory power of the pediatric risk scores was suboptimal, but increased when adding age as a score component. The best performance was achieved by the combination of age and the Comprehensive Aristotle score, both for 30 day and one year mortality.

INTRODUCTION

Risk stratification models play an important role in outcome assessment. With increasing numbers of adult congenital surgery patients, a proper risk stratification tool should therefore be available for this heterogeneous and often complex patient group. At this moment however, no model is available for this population.

For pediatric congenital heart surgery, two important risk stratification scores are available.

In 2002, the Risk Adjustment in Congenital Heart Surgery (RACHS-1) system was published [1]. This score provides 6 risk categories based upon procedure complexity. A comparable approach was published in 2004: the Aristotle Score [2]. This score assigns points to each procedure type based upon mortality, morbidity and technical difficulty. Points for this Basic Aristotle Score range from 1.5 to 15. In addition, the Comprehensive Aristotle Score was presented. Ten extra points can be scored based upon several procedure dependent and procedure independent variables.

Both the RACHS-1 system and the Aristotle Score rely heavily upon expert opinion but have previously been validated with good results in pediatric congenital populations [3-7].

In 2009, a new score based upon 77,294 procedures from the Society of Thoracic Surgeons (STS) and European Association for Cardiothoracic Surgery (EACTS) congenital databases was presented as an alternative to the RACHS-1 system and the Aristotle Score [8]. The main advantage of this STS-EACTS Congenital Heart Surgery Mortality Score (ranging from 0.1 to 5.0 points) is that it was developed using data instead of expert opinion.

For adult cardiac surgery, a variety of risk stratification models are available [9-12]. Although based upon a majority of coronary surgery patients, the EuroSCORE is still a widely used risk stratification model for all types of cardiac surgery, often including adult congenital procedures [9]. However, the EuroSCORE was already identified as a less suitable stratification tool for mortality after adult congenital surgery [13].

While the EuroSCORE does contain age as a model variable, this variable assigns additional risk to patients over 60 years only [14]. The Basic Aristotle, RACHS-1 and STS-EACTS score systems do not contain any age related variables suitable for an adult population [1, 2, 8]. The Comprehensive Aristotle Score for

Table 1. Age distribution and corresponding points.

Age in years	Additional points	Incidence (%)
18 – 29	1	327 (34.0%)
30 – 39	2	226 (23.4%)
40 – 49	3	180 (18.7%)
50 – 59	4	106 (11.0%)
60 and above	5	124 (12.9%)

example contains prematurity as an age related risk factor [2]. A logistic regression model provided by the RACHS-1 developers, includes age below 1 month as a model variable [1].

As age is generally considered an important risk factor in cardiac surgery, adding an age component to these congenital scoring systems might improve their performance in an adult population.

This study aims to identify a suitable stratification tool for the adult congenital heart surgery population. The discriminatory performance of pediatric congenital score models was therefore tested in a single center adult congenital surgery population. In addition, an age component was added to these models and performance was compared with the original score systems.

MATERIALS AND METHODS

Data

All adult (age 18 years and older at time of surgery) patients who underwent congenital cardiac surgery in the Erasmus University Medical Center between January 1990 and January 2007 were included in this study (N=963).

Mean patient age is 39.3 ± 15.5 years, with a median of 36 years and a range between 18.0 and 85.5 years. Detailed patient and procedural characteristics are previously described by Putman and colleagues [13].

The dependent variables in this study are 30 day mortality and one year mortality, regardless of patient location. Thirty day mortality is 1.5% (N=14), and one year mortality is 2.7% (N=26).

Score application

The RACHS-1, Basic Aristotle Score, STS-EACTS Score and Comprehensive Aristo-

tle Score were calculated for each patient. In addition, an age component was added to these scores. Table 1 provides the distribution of age and corresponding points (range 1 – 5). To allow for an equal contribution to the combined scores, the age component was multiplied by a factor 0.6 for the RACHS-1 System, 1.5 for the Basic Aristotle Score, 0.5 for the STS-EACTS Score, and 2.5 for the Comprehensive Aristotle Score. As a result, the original score - age component ratio is 2:1 for all combined scores.

The EuroSCORE was calculated in a previous study based upon the same population [13].

Statistical analysis

Continuous data are presented as mean ± 1 standard deviation, and median with range. Categorical data are presented as proportions.

To evaluate the performance of the risk scores, performance was tested by assessment of discrimination only. Calibration was not assessed in this population.

Discriminatory power was assessed using the c-index (area under the receiver operating characteristic curve) with 95% confidence limits (CI). A c-index of 1.0 would indicate perfect discrimination, whereas a c-index of 0.50 indicates total absence of discrimination. A value between these extremes is the quantitative measure of a model's ability to distinguish between survivors and non-survivors. To demonstrate significant differences between c-indices, a bootstrapping cycle of 2000 runs was performed [15]. Tests between c-indices were 2 sided with $p<0.05$ considered to be a significant difference.

R version 2.9.0 (R Foundation for Statistical Computing, Vienna, Austria) was used for bootstrapping, all other statistical analyses were performed with SPSS version 16.0 (SPSS, Chicago, Illinois).

RESULTS

RACHS-1 distribution

The numbers of patients corresponding with each RACHS-1 risk category are presented in Table 2. Please note that category 3 is by far the largest category, while no patients were assigned to category 5 or 6. Mean of the RACHS-1 is 2.1 ± 0.9, median 2 (range 1 – 4). In combination with age, mean is 3.6 ± 1.2, median

3.6 (range 1.6 – 6).

Basic Aristotle Score distribution

Mean of the Basic Aristotle Score is 6.4 ± 2.3, median 6.5 (range 1.5 – 10.5). In combination with age, mean is 10.0 ± 2.8, median 10.3 (range 4.5 – 16.3). A more detailed distribution is presented in Table 3. The complexity distribution is comparable to the RACHS-1 distribution; no patients are found in the most complex regions, while the majority of patients reside in the medium complexity regions.

STS-EACTS Score distribution

Mean of the STS-EACTS Score is 0.5 ± 0.3, median 0.4 (range 0.1 – 2.1). In combination with age, mean is 1.7 ± 3.4, median 1.5 (range 0.6 – 4.2). A more detailed distribution is presented in Table 4. As with the RACHS-1 and Basic Aristotle Score distributions, no patients are found in the most complex regions.

Comprehensive Aristotle Score distribution

Table 5 provides the incidence of the procedure independent factors used by the Comprehensive Aristotle Score. Factors are scored according to the published Aristotle Score definitions, unless stated otherwise. Mean of the Comprehensive Aristotle Score is 7.6 ± 2.9, median 8.5 (range 3.0 – 20.5). In combination with age, mean is 13.8 ± 4.0, median 13.5 (range 5.5 – 30.0).

Discriminatory performance

The discriminatory power of all score systems is presented in Table 6. For the sake of clarity, results of the EuroSCORE model are included in this table as well.

Age as a single predictor for both endpoints already outperforms the logistic EuroSCORE, the RACHS-1 system, the Basic Aristotle Score and the STS-EACTS Score (all $p<0.001$).

Adding the age component to the Basic Aristotle Score or the Comprehensive Aristotle Score results in a c-index increase of 0.10, for both 30 day as one year mortality ($p<0.001$). The combinations age and the Basic Aristotle Score, RACHS-1 system, or STS-EACTS Score however show a comparable discriminatory power as age alone.

The combination of age and the comprehensive Aristotle Score had the highest discriminatory ability of all approaches ($p<0.001$ when compared to any

Table 2. Distribution of RACHS-1 risk categories.

RACHS-1 risk category	Number of patients (%)
1	286 (29.7%)
2	259 (26.9%)
3	412 (42.8%)
4	6 (0.6%)
5	0 (0%)
6	0 (0%)

RACHS = Risk Adjustment in Congenital Heart Surgery

Table 3. Distribution of Basic Aristotle points.

Basic Aristotle Score	Number of patients (%)
1.5 – 3	240 (24.9%)
> 3 – 4	30 (3.1%)
> 4 – 5	6 (0.6%)
> 5 – 6	56 (5.8%)
> 6 – 7	262 (27.2%)
> 7 – 8	156 (16.2%)
> 8 – 9	156 (16.2%)
> 9 – 10	6 (0.6%)
> 10 – 11	51 (5.3%)
> 11 – 12	0 (0%)
> 12 – 13	0 (0%)
> 13 – 14	0 (0%)
> 14 – 15	0 (0%)

other approach), in both 30 day as one year mortality.

DISCUSSION

This study has tested the RACHS-1, Basic Aristotle Score, STS-EACTS Score and Comprehensive Aristotle Score in an adult congenital population with suboptimal results in terms of discrimination.

Results were suboptimal for all score systems. When adding an age component

Table 4. Distribution of STS-EACTS Score points.

STS-EACTS Score points	Number of patients (%)
0.1 – 0.5	762 (79.1%)
> 0.5 – 1.0	152 (15.8%)
> 1.0 – 1.5	27 (2.8%)
> 1.5 – 2.0	21 (2.2%)
> 2.0 – 2.5	1 (0.1%)
> 2.5 – 5	0 (0%)

STS = Society of Thoracic Surgeons, EACTS = European Association for Cardiothoracic Surgery

Table 5. Incidence of risk factors of the Comprehensive Aristotle Score.

Additional factors	Number (%)
Mechanical cardio-pulmonary support	18 (1.9%)
Myocardial dysfunction	15 (1.6%)
Mechanical ventilation	18 (1.9%)
Pulmonary hypertension*	35 (3.6%)
Endocarditis	22 (2.3%)
Renal dysfunction	6 (0.6%)
Diabetes mellitus**	25 (2.6%)
Redo sternotomy	476 (49.4%)

* Replaces 'Elevated lung resistances'

** Both insulin and non-insulin dependent types score 1 point

to these scores, a significant increase was found in discriminatory power for all three models. C-indices for the Basic Aristotle Score and Comprehensive Aristotle Score increased by 0.10 when assessing 30 day mortality.

Combining the Comprehensive Aristotle Score and age resulted in the best discriminatory performance of all approaches ($p<0.001$).

Additional variables

Adding several variables to the RACHS-1 system, the Basic Aristotle Score or the STS-EACTS Score has previously been studied in pediatric populations and results in a better discriminatory performance for this patient group [8]. A logistic regression model provided by the RACHS-1 developers that includes the com-

Table 6. Discrimination for all risk scores.

Score	30 day mortality c-index	95% CI	1 year mortality c-index	95% CI
Logistic EuroSCORE [13]	0.59	0.45 – 0.72	0.67	0.56 – 0.78
Age	0.67	0.54 – 0.80	0.66	0.56 – 0.76
RACHS-1	0.60	0.45 – 0.75	0.59	0.47 – 0.70
RACHS-1 + Age	0.69	0.54 – 0.84	0.66	0.56 – 0.77
Basic Aristotle	0.60	0.44 – 0.76	0.56	0.44 – 0.68
Basic Aristotle + Age	0.70	0.53 – 0.86	0.64	0.53 – 0.76
STS-EACTS	0.60	0.44 – 0.76	0.57	0.46 – 0.68
STS-EACTS + Age	0.69	0.54 – 0.84	0.67	0.57 – 0.77
Comprehensive Aristotle	0.66	0.54 – 0.79	0.65	0.54 – 0.75
Comprehensive Aristotle + Age	0.76	0.62 – 0.90	0.74	0.65 – 0.84

RACHS = Risk Adjustment in Congenital Heart Surgery, STS = Society of Thoracic Surgeons, EACTS = European Association for Cardiothoracic Surgery

plexity score and a few other risk factors also showed an increased performance in their study population [1]. This model was not assessed in this study because the majority of variables in this model are present in pediatric patients only.

Including additional risk factors to the RACHS-1 model will likely improve results in an adult congenital population as well, comparable to the Comprehensive Aristotle Score.

Complexity of adult congenital surgery

As one can appreciate from Table 2, RACHS-1 categories 4, 5 and 6 are not well represented. Although this effect can be seen in the higher Aristotle Score and STS-EACTS Score points as well, they offer a larger variety of risk categories. Procedures assigned to these higher RACHS-1 categories are obviously generally performed in childhood. This results in a sub-optimal classification for adult congenital patients, possibly contributing to the limited performance of this score. A revision of these complexity based categories might improve results. This also goes for the STS-EACTS Score.

Calibration

Calibration was not assessed in this study. For all pediatric scores included in

this study, the risk of mortality corresponding with the resulting scores and categories has been reported, but these risks differ greatly from our adult population. These risks clearly depend on the selection of the reference data set. For example, it has been reported that in a specific pediatric data set, RACHS-1 categories 2 and 3 have a mortality of respectively 3.8 and 9.5%. As thirty day mortality in our study is 1.5% while most patients reside in categories 2 and 3, testing calibration is futile.

As this is a single center study, no efforts have been made to construct score-mortality conversions and validate these conversions. Due to lower mortality rates it is likely that all complexity based scores developed for pediatric patients require recalibration when used in an adult congenital population [13, 16].

Limitations

Limitations of this study are the single center approach and the modest cohort size. Differences in patient selection and surgical skills may cause results to vary from those in other centers. For this reason, analyses are limited to validation of an existing model in combination with age; no new model was derived from this study cohort. For the sake of parsimony, the age component in this study consists of a linear relation between age and points, while in real life a more exponential relation might be found.

Conclusions and recommendations

Although general applicability of results from this study remains undetermined, the encouraging results of the inclusion of age should incent the EACTS and STS database committees to explore these findings in a larger adult congenital population or – if necessary– to start collecting additional variables to enable such investigations. Furthermore, just like unification of the Aristotle Score and the RACHS-1 for pediatric purposes was suggested by those that developed them [17], a combined effort for the adult congenital population might be considered as well.

A combination of procedure complexity and known adult surgery risk factors, including age and for example smoking history and preoperative renal dysfunction, will probably result in a well performing risk stratification model for adult congenital heart surgery. Ideally, this model will truly be based upon an adult congenital heart surgery patient cohort. This way, risk factors that might be specific to adults with congenital heart disease, such as the number of repeat

sternotomy, can be examined further [18].

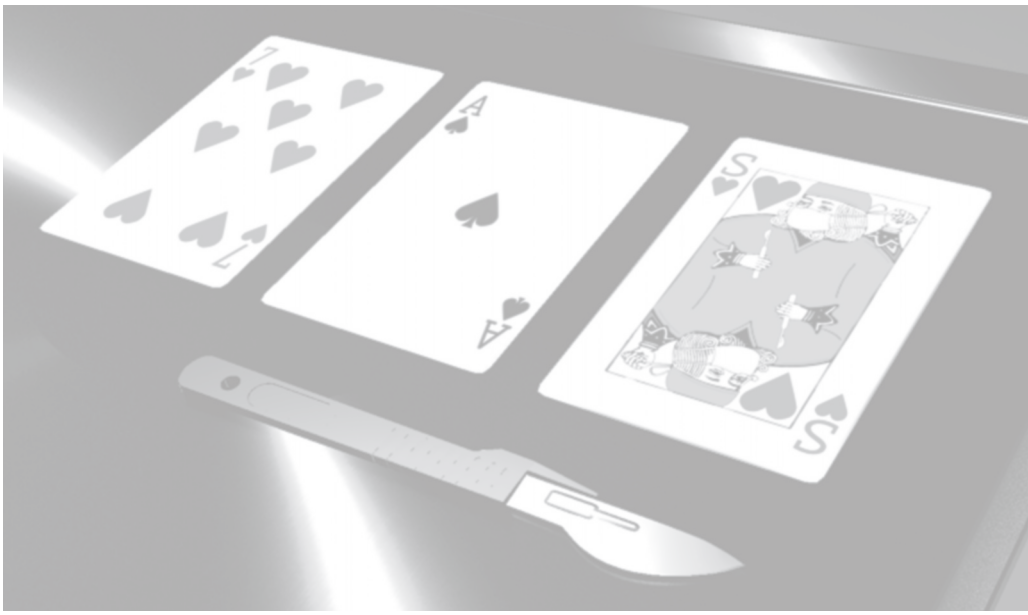
Until dedicated prediction models for adult congenital surgery are available, the simple combination of the Comprehensive Aristotle Score and age appears a promising risk stratification alternative in this population.

REFERENCES

1. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH and Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. J Thorac Cardiovasc Surg. 2002;123(1):110-8.
2. Lacour-Gayet F, Clarke D, Jacobs J, Comas J, Daebritz S, Daenen W, Gaynor W, Hamilton L, Jacobs M, Maruszsewski B, Pozzi M, Spray T, Stellin G, Tchervenkov C and Mavroudis And C. The Aristotle score: a complexity-adjusted method to evaluate surgical results. Eur J Cardiothorac Surg. 2004;25(6):911-24.
3. Allen SW, Gauvreau K, Bloom BT and Jenkins KJ. Evidence-based referral results in significantly reduced mortality after congenital heart surgery. Pediatrics. 2003;112(1 Pt 1):24-8.
4. Boethig D, Jenkins KJ, Hecker H, Thies WR and Breymann T. The RACHS-1 risk categories reflect mortality and length of hospital stay in a large German pediatric cardiac surgery population. Eur J Cardiothorac Surg. 2004;26(1):12-7.
5. Jacobs JP, Lacour-Gayet FG, Jacobs ML, Clarke DR, Tchervenkov CI, Gaynor JW, Spray TL, Maruszewski B, Stellin G, Gould J, Dokholyan RS, Peterson ED, Elliott MJ and Mavroudis C. Initial application in the STS congenital database of complexity adjustment to evaluate surgical case mix and results. Ann Thorac Surg. 2005;79(5):1635-49; discussion 1635-49.
6. Larsen SH, Pedersen J, Jacobsen J, Johnsen SP, Hansen OK and Hjortdal V. The RACHS-1 risk categories reflect mortality and length of stay in a Danish population of children operated for congenital heart disease. Eur J Cardiothorac Surg. 2005;28(6):877-81.
7. O'Brien SM, Jacobs JP, Clarke DR, Maruszewski B, Jacobs ML, Walters HL, 3rd, Tchervenkov CI, Welke KF, Tobota Z, Stellin G, Mavroudis C, Hamilton JR, Gaynor JW, Pozzi M and Lacour-Gayet FG. Accuracy of the aristotle basic complexity score for classifying the mortality and morbidity potential of congenital heart surgery operations. Ann Thorac Surg. 2007;84(6):2027-37; discussion 2027-37.
8. O'Brien SM, Clarke DR, Jacobs JP, Jacobs ML, Lacour-Gayet FG, Pizarro C, Welke KF, Maruszewski B, Tobota Z, Miller WJ, Hamilton L, Peterson ED, Mavroudis C and Edwards FH. An empirically based tool for analyzing mortality associated with congenital heart surgery. J Thorac Cardiovasc Surg. 2009;138(5):1139-53.
9. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S and Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. 1999;16(1):9-13.
10. O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2--isolated valve surgery. Ann Thorac Surg. 2009;88(1 Suppl):S23-42.
11. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg. 2009;88(1 Suppl):S43-62.
12. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1--coronary artery bypass grafting

surgery. Ann Thorac Surg. 2009;88(1 Suppl):S2-22.
13. Putman LM, van Gameren M, Meijboom FJ, de Jong PL, Roos-Hesselink JW, Witsenburg M, Takkenberg JJ and Bogers AJ. Seventeen years of adult congenital heart surgery: a single centre experience. Eur J Cardiothorac Surg. 2009;36(1):96-104; discussion 104.
14. Roques F, Michel P, Goldstone AR and Nashef SA. The logistic EuroSCORE. Eur Heart J. 2003;24(9):881-2.
15. Efron B and Tibshirani R. An Introduction to the bootstrap. Chapman & Hall, New York, 1993.
16. Vida VL, Berggren H, Brawn WJ, Daenen W, Di Carlo D, Di Donato R, Lindberg HL, Corno AF, Fragata J, Elliott MJ, Hraska V, Kiraly L, Lacour-Gayet F, Maruszewski B, Rubay J, Sairanen H, Sarris G, Urban A, Van Doorn C, Ziemer G and Stellin G. Risk of surgery for congenital heart disease in the adult: a multicentered European study. Ann Thorac Surg. 2007;83(1):161-8.
17. Jacobs ML, Jacobs JP, Jenkins KJ, Gauvreau K, Clarke DR and Lacour-Gayet F. Stratification of complexity: the Risk Adjustment for Congenital Heart Surgery-1 method and the Aristotle Complexity Score--past, present, and future. Cardiol Young. 2008;18 Suppl 2:163-8.
18. Dearani JA, Connolly HM, Martinez R, Fontanet H and Webb GD. Caring for adults with congenital cardiac disease: successes and challenges for 2007 and beyond. Cardiol Young. 2007;17 Suppl 2:87-96.

CHAPTER 10



Complexity of Coronary Vasculature Predicts Outcome of Surgery for Left Main Disease

Birim Ö
Van Gameren M
Bogers AJJC
Serruys PW
Mohr FW
Kappetein AP

Ann Thorac Surg. 2009; 87: 1097-104

ABSTRACT

Background

The SYNTAX score, a comprehensive angiographic scoring system, was recently developed as a tool for risk stratification during the SYNTAX trial (randomized trial comparing coronary artery bypass grafting with percutaneous coronary intervention). We applied the SYNTAX score in patients with left main coronary artery disease who underwent coronary artery bypass grafting to examine its role in predicting incidences of major adverse cardiac and cerebrovascular events (MACCE) within 30 days and 1 year.

Methods

One hundred forty-eight patients were studied. Their angiograms were scored according to the SYNTAX score. The MACCE-free survival curves were estimated by the Kaplan–Meier method. Univariate and multivariate analyses determined risk factors for MACCE. Performance of the SYNTAX score was studied with respect to discrimination by receiver-operating characteristic curves with their area under the curve (c-index). Classification and regression tree analysis was performed to identify the best outcome predictors and develop a risk stratification model.

Results

Overall SYNTAX score ranged from 11 to 53 (mean, 24±9). At 30 days and 1 year, 15 (10%) and 19 (13%) patients experienced MACCE. Patients with a higher SYNTAX score had a significantly ($p < 0.0001$) poorer MACCE-free survival. In multivariate analysis, SYNTAX score, female sex, and incomplete revascularization were associated with a higher rate of MACCE in 30 days. The SYNTAX score was the single predictor for MACCE in 1 year. The c-index of the SYNTAX score was 0.88 for 30 days and 0.90 for 1 year, respectively. The SYNTAX score was the best single discriminator between patients with and those without MACCE, with a discrimination level of 36.5.

Conclusions

The SYNTAX score is the first coronary vasculature complexity score predictive for postoperative outcome in patients with left main coronary artery disease undergoing coronary artery bypass grafting. The outcomes of the ongoing SYNTAX trial will definitively define the role of the SYNTAX score in predicting short-term and long-term incidence of MACCE.

INTRODUCTION

Coronary artery bypass graft surgery (CABG) is considered the standard of care for the treatment of left main coronary artery disease [1–4]. However, continued technical evolution of percutaneous coronary intervention (PCI), including the introduction of drug-eluting stents, has renewed the interest for the percutaneous treatment of left main coronary artery disease [5–7]. In addition, whereas PCI has improved, CABG has also progressed with better perioperative management, more frequent use of arterial grafting, and improved techniques with minimally invasive and off-pump surgery as options [8, 9]. As a result of continually improving the treatment strategy in patients with coronary artery disease, the SYNTAX (Synergy Between Percutaneous Intervention With TAXUS Drug-Eluting Stent and Cardiac Surgery) trial has recently been initiated [10, 11]. In this prospective randomized trial, consecutive patients with de novo three-vessel disease or left main coronary artery disease (isolated, or in combination with one-, two-, or three-vessel disease) are randomly assigned to either PCI or CABG, if both can achieve comparable revascularization. The SYNTAX score is a comprehensive, angiographic scoring system, aiming to assist in patient selection and risk stratification. It merges several previously validated angiographic classifications based on morphology and location of coronary artery lesions within the coronary tree. Recently, the predictive value of the SYNTAX score was assessed in patients who underwent PCI [12]. A validation of this angiographic classification tool is lacking for patients undergoing CABG. We applied the SYNTAX score in patients with left main coronary artery disease (isolated, or in combination with one-, two-, or three-vessel disease) who underwent primary CABG to examine its role in predicting short-term and long-term incidences of major adverse cardiac and cerebrovascular events (MACCE).

MATERIAL AND METHODS

Study Design and Patient Population

Retrospectively, the medical records of 148 patients who underwent primary CABG for left main coronary artery disease at the Department of Cardio-Thoracic Surgery of the Erasmus MC Rotterdam between January 1, 2001, and March 31, 2004, have been reviewed. The Ethics Committee of the Erasmus MC Rotterdam

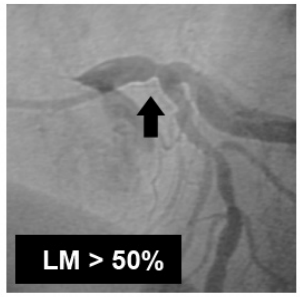
Table 1. Baseline characteristics of the study population .

Characteristic	SYNTAX score			p value
	≤ 19 (n= 49)	> 19 – 25 (n= 48)	> 25 (n= 51)	
Age (years)	62 ± 9	63 ± 9	67 ± 10	0.028
Female sex	4 (8%)	10 (21%)	9 (18%)	0.199
Body mass index (kg/m^2)	27 ± 4	27 ± 4	27 ± 4	0.703
Diabetes	7 (14%)	10 (21%)	15 (29%)	0.183
Insulin dependent	2 (4%)	1 (2%)	4 (8%)	0.389
Hypertension	13 (27%)	23 (48%)	25 (49%)	0.038
Hypercholesterolemia	35 (71%)	29 (60%)	36 (71%)	0.435
Smoking	14 (29%)	8 (17%)	11 (22%)	0.367
Left ventricular ejection fraction (%)	57 ± 8	56 ± 11	52 ± 10	0.042
Chronic obstructive pulmonary disease	4 (8%)	3 (6%)	5 (10%)	0.811
Peripheral vascular disease	8 (16%)	3 (6%)	13 (26%)	0.034
Prior myocardial infarction	13 (27%)	15 (31%)	27 (53%)	0.014
Unstable angina	21 (43%)	18 (38%)	22 (43%)	0.816
Intra-aortic balloon pumping support	6 (12%)	9 (19%)	8 (16%)	0.676
Urgent surgery	3 (6%)	7 (15%)	3 (6%)	0.225
Prior percutaneous coronary intervention	10 (21%)	7 (15%)	8 (16%)	0.717
Prior cerebrovascular accident	3 (6%)	1 (2%)	3 (6%)	0.575
Chronic renal failure	5 (10%)	5 (10%)	4 (8%)	0.888
EuroSCORE	3.1 ± 3.1	3.8 ± 3.2	4.8 ± 3.4	0.035
Parsonnet Score	5.0 ± 6.9	7.1 ± 6.8	9.3 ± 7.5	0.016

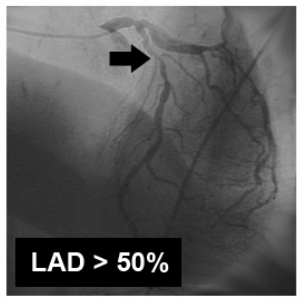
has approved the study. Individual consent for the study was waived. Patients were followed up with regular visits to the outpatient clinic. Surgical risk profile of all patients was scored according to the EuroSCORE [13] and the Parsonnet score [14]. The patient's angiograms (coronary vasculature complexity) were scored according to the SYNTAX score [10, 11].

SYNTAX Score and Angiographic Analysis

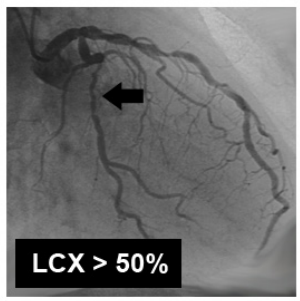
Each coronary lesion producing 50% or greater luminal obstruction in vessels 1.5 mm or greater was separately scored and added to provide the overall SYNTAX score. The SYNTAX score was calculated using dedicated software that integrates the number of lesions with their specific weighting factors based on the amount of myocardium distal to the lesion according to the score of Leaman and col-



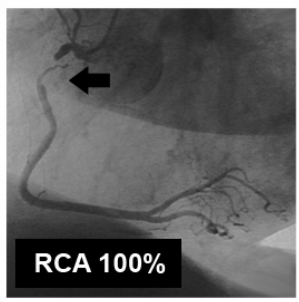
Lesion 1	
Segment 5	10
Bifurcation Type A	1
Lesion 1 score	**11**



Lesion 2	
Segment 6	7
Length >20 mm	1
Lesion 2 score	**8**



Lesion 3	
Segment 13	1
Length >20 mm	1
Heavy calcification	2
Lesion 3 score	**4**



Lesion 4	
Segment 1	5
Age T.O. is yes	1
Blunt stump	1
Bridging	1
First segment visualised: 2	0
Lesion 4 score	**8**

SYNTAX SCORE	**31**

Figure 1. SYNTAX score of a patient with left main coronary disease in combination with three-vessel disease.

LAD = left anterior descending coronary artery; LCX = left circumflex coronary artery; LM = left main coronary artery; RCA = right coronary artery.

Table 2. Angiographic and procedural characteristics of the study population.

Characteristic	SYNTAX score			p value
	≤ 19 (n= 49)	> 19 – 25 (n= 48)	> 25 (n= 51)	
Left main stenosis (%)	75 ± 13	74 ± 14	73 ± 11	0.775
Isolated left main disease	14 (29%)	0	0	< 0.001
Left main + one-vessel disease	24 (49%)	20 (42%)	2 (4%)	< 0.001
Left main + two-vessel disease	11 (22%)	16 (33%)	13 (26%)	0.461
Left main + three-vessel disease	0	12 (25%)	36 (71%)	< 0.001
Type of graft				
Total arterial	9 (18%)	9 (19%)	2 (4%)	0.047
Arterial and venous	37 (76%)	37 (77%)	44 (86%)	0.350
Total venous	3 (6%)	2 (4%)	5 (10%)	0.524
Complete revascularization	44 (90%)	44 (92%)	40 (78%)	0.111

leagues [15] and the morphologic features of each single lesion, as previously reported [11]. An example of the SYNTAX score calculation in 1 patient is shown in Figure 1. The patient's SYNTAX score was stratified into SYNTAX score tertiles [12]. All diagnostic angiograms were scored by one experienced investigator (Ö.B.) who was blinded as to procedural data and clinical outcome.

Study Objectives and End Points

The primary objective of this study was to analyze the value of the SYNTAX score in predicting short-term and long-term outcomes in patients with left main coronary artery disease who underwent primary CABG. The primary end point was the incidence of MACCE, defined as a composite of all-cause death, cerebrovascular events, myocardial infarction, and repeat revascularization by either PCI or CABG. We report the incidence of the primary end point at 30 days and 1 year.

End Point Definitions

Death from all causes was reported. A cerebrovascular accident is any acute event related to the impairment of the cerebral circulation that lasts more than 24 hours and results in irreversible brain damage or permanent body impairment. Myocardial infarction was considered if there was documentation of new abnormal Q waves and a ratio of serum creatinine kinase-MB isoenzyme to total cardiac enzyme that was greater than 0.1 or a creatinine kinase to creatinine

Table 3. Thirty-day and one-year outcome.

Characteristic	SYNTAX score			p value
	≤ 19 (n= 49)	> 19 – 25 (n= 48)	> 25 (n= 51)	
30 days				
MACCE	0	3 (6%)	12 (24%)	0.0001
Death	0	1 (2%)	6 (12%)	0.014
Myocardial infarction	0	3 (6%)	6 (12%)	0.038
Cerebrovascular accident	0	0	2 (4%)	0.329
Revascularization	0	1 (2%)	0	0.324
1 year				
MACCE	0	3 (6%)	16 (31%)	< 0.0001
Death	0	1 (2%)	8 (16%)	0.002
Myocardial infarction	0	3 (6%)	6 (12%)	0.038
Cerebrovascular accident	0	0	3 (6%)	0.106
Revascularization	0	1 (2%)	1 (2%)	0.770

kinase-MB value that was five times the upper limit of normal [10]. Serum creatinine kinase and creatinine kinase-MB isoenzyme concentration were measured 6, 12, and 18 hours after operation. All repeat revascularization procedures by either PCI or CABG were recorded. Events were counted from the time of operation.

Statistical Analysis

Discrete variables are displayed as proportions, continuous variables as mean ± standard deviation unless specified otherwise. The X^2 (whenever n > 5 in all groups) or Fisher's exact test was used to analyze the categorical data. Differences between continuous variables were analyzed using one-way analysis of variance. When comparing three groups, a probability value of less than 0.0167 was considered significant (alpha correction according to Bonferroni). The MACCE-free survival curves were estimated by the Kaplan–Meier method. Differences in survival were compared using the log-rank test. Univariate and multivariate logistic regression analysis determined risk factors for MACCE within 30 days. Univariate and multivariate Cox proportional hazard analysis determined risk factors for MACCE within 1 year. A probability value of less than 0.05 was considered significant. The multivariate analyses were performed with a stepwise

backward regression model in which each variable with a probability value of less than 0.20 in the univariate analysis was entered in the model. Relative risks are reported with 95% confidence intervals. Performance of the SYNTAX score was studied with respect to discrimination (resolution). Discrimination refers to the ability to distinguish patients with MACCE from those without. It was assessed by receiver operating characteristic curves with their areas under the curve (c-index) with 95% confidence limits. A c-index of 1.0 would indicate perfect discrimination, whereas a c-index of 0.5 indicates total absence of discriminative power.

All variables associated with the incidence of MACCE at 1 year at a probability value of 0.10 in the Cox proportional hazard analysis were subjected to classification and regression tree analysis to identify the best outcome predictors and develop the risk stratification model [16]. This method is based on recursive partitioning analysis and involves the segregation of different values of classification variables through a decision tree composed of progressive binary splits. This approach has the advantage of uncovering possible interactions among predictors.

Descriptive statistical analyses were performed with SPSS 15.0 for Windows (SPSS, Chicago, IL), and R version 2.5.1 (R Foundation for Statistical Computing, Vienna, Austria) was used for calculating c-indices with 95% confidence limits, constructing receiver-operating characteristic curves, plotting Kaplan–Meier survival curves, and performing classification and regression tree analysis.

RESULTS

Patient Characteristics

Of the 148 patients included in this analysis, 125 (85%) were men and 23 (15%) women. The mean age at time of surgery was 64±9 years (range, 32 to 83 years). Eightyseven patients (59%) presented with stable angina and 61 patients (41%) with unstable angina. The patient's preoperative characteristics stratified according to SYNTAX score tertiles are outlined in Table 1. The rate of a prior myocardial infarction and the Parsonnet score were significantly higher in the patients within the third tertile, whereas no difference was seen in the rate of all other preoperative characteristics.

SYNTAX Score and Procedural Characteristics

The overall SYNTAX score ranged from 11 to 53, with a mean of 24±9. The overall EuroSCORE ranged from 0 to 15 (mean, 4±3) and the overall Parsonnet score ranged from 0 to 35 (mean, 7±7). Fourteen (10%) patients had isolated left main disease, 46 (31%) patients had left main disease with one-vessel disease, 40 (27%) patients had left main disease with two-vessel disease, and 48 (32%) patients had left main disease with three-vessel disease. Elective surgery accounted for 82% of the procedures, with 18% being urgent. In 23 patients (16%) preoperative intraaortic balloon pumping support was applied. Revascularization with arterial grafts only was performed in 20 patients (14%) while 10 patients (7%) were treated with venous grafts only. The majority (118 patients, 80%) was treated with a combination of arterial and venous grafts.

Complete revascularization was achieved in 128 patients (86%). Baseline angiographic and procedural characteristics of the study population stratified according to SYNTAX score tertiles are presented in Table 2. The patients within the first tertile mainly had isolated left main disease or left main with one-vessel disease, whereas the patients within the third tertile mainly had left main with two- or three-vessel disease.

Thirty-Day Outcome

At 30 days, 15 (10%) of the patients experienced MACCE. Thirty-day outcome of the study population stratified across SYNTAX score tertiles is illustrated in Table 3. Overall hospital mortality was 5% (7 of 148 patients). None of the patients within the first SYNTAX score tertile experienced MACCE. The patients within the third tertile showed the highest rate of MACCE. This difference of MACCE was mainly driven by a higher rate of postoperative deaths within the third tertile. When evaluating risk factors in the univariate analysis, the SYNTAX score, female sex, number of diseased vessels, and incomplete revascularization significantly predicted the rate of MACCE (Table 4). In multivariate analysis, SYNTAX score (relative risk, 1.2; 95% confidence interval, 1.1 to 1.3), female sex (relative risk, 6.6; 95% confidence interval, 1.5 to 29.7), and incomplete revascularization (relative risk, 4.7; 95% confidence interval, 1.1 to 20.8), were associated with a higher rate of MACCE (Table 4).

On the basis of the discriminatory performance of the SYNTAX score a receiver-operating characteristic curve was generated (Figure 2A). The corresponding c-index of the SYNTAX score was 0.88 (95% confidence interval, 0.79 to 0.97),

Table 4. Univariate and multivariate analysis of MACCE at 30 days and 1 year.

Characteristic	30 days			1 year		
	p value	RR	95% CI	p value	RR	95% CI
Univariate model						
SYNTAX score	< 0.0001	1.2	1.1 - 1.3	< 0.0001	1.2	1.1 - 1.2
Age	0.22	1.0	2.3 - 26.9	0.65	1.0	0.9 - 1.0
Female gender	0.01	4.6	1.4 - 14.4	0.05	2.7	1.0 - 7.0
Diabetes	0.87	0.9	0.2 - 3.4	0.92	0.9	0.3 - 2.8
Insulin dependent diabetes	0.12	3.9	0.7 - 22.4	0.23	2.4	0.6 - 10.6
Insulin independent diabetes	0.29	0.3	0.1 - 2.6	0.44	0.6	0.1 - 2.4
Hypertension	0.52	0.8	0.2 - 2.1	0.97	1.0	0.4 - 2.5
Hypercholesterolemia	0.94	1.0	0.3 - 3.0	0.69	0.8	0.3 - 2.1
Smoking	0.39	0.5	0.1 - 2.4	0.87	0.9	0.3 - 2.7
Body Mass Index	0.70	1.0	0.9 - 1.2	0.62	1.0	0.9 - 1.1
Prior myocardial infarction	0.18	2.2	0.7 - 6.1	0.32	1.6	0.6 - 3.9
Prior cerebrovascular accident	0.71	1.5	0.2 - 13.5	0.23	2.5	0.6 - 10.6
Prior percutaneous coronary intervention	0.74	1.3	0.3 - 4.8	0.58	1.4	0.5 - 4.1
Peripheral vascular disease	0.75	0.8	0.2 - 3.7	0.95	1.0	0.3 - 3.3
Chronic renal failure	0.70	0.7	0.1 - 5.4	0.89	1.1	0.3 - 4.8
Chronic obstructive pulmonary disease	0.83	0.8	0.1 - 6.6	0.72	1.3	0.3 - 5.7
Left ventricular ejection fraction	0.24	1.0	0.9 - 1.0	0.16	1.0	0.9 - 1.0
Number of diseased vessels	0.002	4.7	1.8 - 12.4	0.001	2.9	1.5 - 5.6
Left main stenosis (%)	0.63	1.0	0.9 - 1.1	0.36	1.0	0.9 - 1.1

Table 4. Continued.

Characteristic	30 days			1 year		
	p value	RR	95% CI	p value	RR	95% CI
Univariate model (continued)						
Left main stenosis (%)	0.63	1.0	0.9 - 1.1	0.36	1.0	0.9 - 1.1
Unstable angina	0.52	0.7	0.2 - 2.1	0.67	0.8	0.3 - 2.1
Intra-aortic balloon pumping support	0.80	0.8	0.2 - 3.9	0.99	1.0	0.3 - 3.4
Type of graft						
Arterial and venous	0.23	0.2	0.1 - 2.7	0.23	0.2	0.1 - 2.6
Total venous	0.35	0.3	0.1 - 2.4	0.56	0.6	0.1 - 2.8
Incomplete revascularization	0.004	5.7	1.8 - 18.3	0.02	3.3	1.3 - 8.7
Urgent operation	0.76	0.7	0.1 - 6.0	0.58	0.6	0.1 - 4.2
EuroSCORE	0.30	1.1	0.9 - 1.2	0.12	1.1	1.0 - 1.2
Parsonnet Score	0.61	1.0	0.7 - 1.6	0.50	0.1	1.0 - 1.1
Multivariate model						
SYNTAX score	< 0.0001	1.2	1.1 - 1.3	< 0.0001	1.2	1.1 - 1.2
Female gender	0.013	6.7	1.5 - 29.7	0.06	2.6	1.0 - 6.9
Incomplete revascularization	0.04	4.7	1.1 - 20.8	NS		
Prior myocardial infarction	NS					
Number of diseased vessels	NS			NS		
Insulin dependent diabetes	NS					
EuroSCORE				NS		
Left ventricular ejection fraction				NS		

CI = confidence interval; NS = not significant; RR = relative risk.

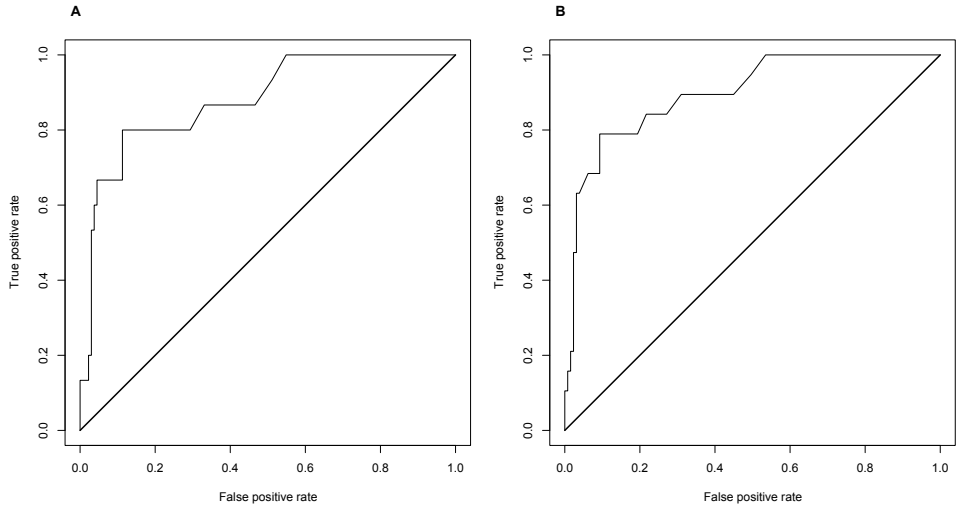


Figure 2. Receiver-operating characteristic curves for discriminatory performance of the SYNTAX score at 30 days (A) and 1 year (B).

demonstrating an excellent discriminatory performance of the SYNTAX score.

One-Year Outcome

At 1 year, 19 (13%) of the patients experienced MACCE. The distribution of MACCE according to SYNTAX score tertiles is illustrated in Table 3. All 4 MACCE (2 deaths, 1 revascularization, and 1 cerebrovascular accident) that occurred after 30 days occurred in the patients within the third tertile. The MACCE-free survival curves according to SYNTAX score tertiles are illustrated in Figure 3. Survival of the patients within the third tertile was significantly poorer than patients within the first tertile ($p < 0.0001$) and patients within the second tertile ($p < 0.002$). No significant difference ($p = 0.08$) was seen between patients within the first and those within the second tertile. In univariate analysis, the SYNTAX score, female sex, number of diseased vessels, and incomplete revascularization significantly predicted the rate of MACCE (Table 4). In multivariate analysis, only the SYNTAX score (relative risk, 1.2; 95% confidence interval, 1.1 to 1.2) was associated with a higher rate of MACCE (Table 4).

The c-index of the SYNTAX score was 0.90 (95% confidence interval, 0.82 to 0.97), demonstrating an outstanding discriminatory performance of the SYNTAX score (Figure 2B).

By evaluating all variables related to 1-year MACCE at a p value of 0.10 or

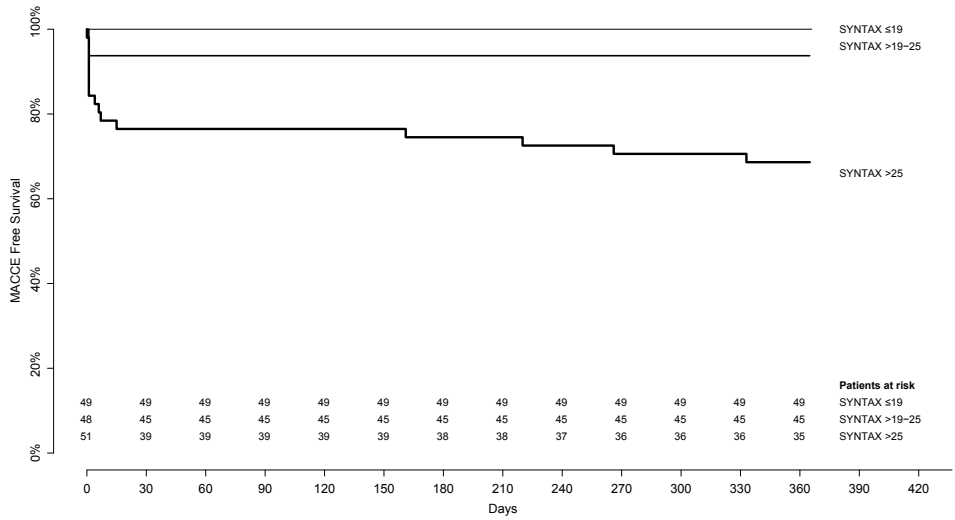


Figure 3. Major adverse cardiac and cerebrovascular event (MACCE)-free survival according to SYNTAX score tertiles.

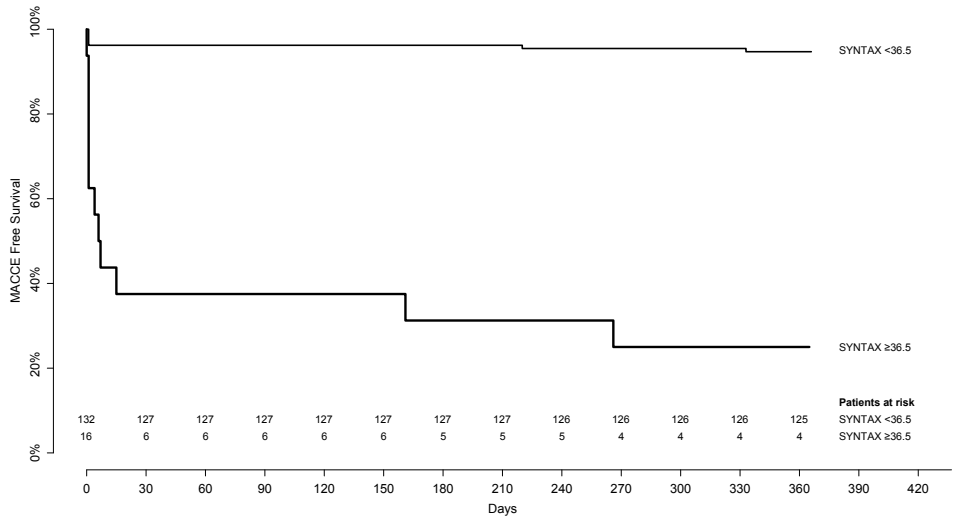


Figure 4. Based on classification and regression three analysis, the SYNTAX score emerged as the best single discriminator between patients with and those without major adverse cardiac and cerebrovascular events (MACCE) in 1 year, with a discrimination level of 36.5.

less at univariate Cox proportional hazard analysis (Table 4), the classification and regression tree method confirmed the SYNTAX score as the best single dis-

criminator between patients with and those without MACCE, with a discrimination level of 36.5 (Figure 4). When stratified into the discrimination level suggested by classification and regression tree analysis, the adjusted relative risk for MACCE was 21.5 (95% confidence interval 8.3 to 55.9; $p < 0.0001$) for patients with a high versus low SYNTAX score.

COMMENT

Despite a continuous effort to detect new and progressively more predictive markers of prognosis in patients with coronary artery disease, implementation of unconventional and expensive risk stratification algorithms in the clinical setting remains problematic. So far, prognostic scoring systems, which are used in clinical practice, such as the EuroSCORE and the Parsonnet score, consist of patient-related and operation-related risk factors. The SYNTAX score, which is a comprehensive, angiographic scoring system, was recently developed in an attempt to assist in patient selection and risk stratification of patients with extensive coronary artery disease undergoing revascularization of the left main coronary artery or the three main coronary arteries [10, 11]. Higher SYNTAX scores, indicative of more complex coronary artery disease, are assumed to represent a bigger therapeutic challenge and to have worse prognosis. A recently published study evaluating the predictive value of the SYNTAX score in patients who underwent PCI showed that the SYNTAX score had the greatest discriminatory ability for incidence of MACCE [12]. To obtain some insights into the performance of the SYNTAX score in patients who underwent CABG, 148 patients who underwent primary CABG for left main coronary artery disease were studied. In our study, the SYNTAX score performed significantly better than the EuroSCORE and Parsonnet score in terms of prognostic accuracy. The discriminatory performance of the SYNTAX score was outstanding with a c-index of 0.88 for incidence of MACCE in 30 days, and a c-index of 0.90 for incidence of MACCE in 1 year. The classification and regression tree analysis confirmed the SYNTAX score as the best single discriminator between patients with and those without MACCE, with a discrimination level of 36.5. Our findings demonstrate that the SYNTAX score may be a suitable tool to stratify risk in early and late outcomes in this subset of patients.

Most prognostic models proposed thus far have been derived from an origi-

nal dataset from a large-scale registry or a randomized controlled trial [17]. In this context, a vital aspect of prediction is to consider whether such a model is applicable to similar patients in another setting. A model that is found to pass such a test is said to have been validated [18]. The SYNTAX score was created by an international group of expert interventional cardiologists and cardiac surgeons by merging together and tailoring several previously proposed coronary artery disease scoring systems based on personal expertise [11]. The ultimate goal is to create an angiographic tool grading the complexity of coronary artery disease and obtain evidence-based guidelines for selecting the optimal technique of revascularization (CABG or PCI). The present report is the first evaluation of the predictive value of this recently developed angiographic scoring system in patients undergoing CABG. However, the SYNTAX score cannot be considered fully validated because this is the only data set of patients undergoing CABG in which the model has been tested. As such, it remains unclear whether and to what extent our present findings can be reproduced in a different group of patients with left main coronary artery disease. The outcomes of the SYNTAX trial, which is ongoing at the present time, in which the SYNTAX score will be used to predict clinical outcomes at 1 month, and 1, 3, and 5 year after the procedure will most likely define the role of the SYNTAX score in predicting clinical outcomes after CABG or PCI [10, 11].

Interpretation of diagnostic angiograms, as any clinical tool, is subject to some intraobserver and interobserver variability. A limitation of the present study is that one experienced investigator scored the angiograms, whereas in a prospective randomized controlled trial the angiograms are scored by a local heart team (composed of both a cardiothoracic surgeon and an interventional cardiologist), and as a consequence probably will decrease the interobserver variability. The present study is performed in a relatively small number of patients. Therefore, some covariates, such as age and urgent surgery, which have been shown to affect outcome in earlier studies, might have been excluded from the multivariate model.

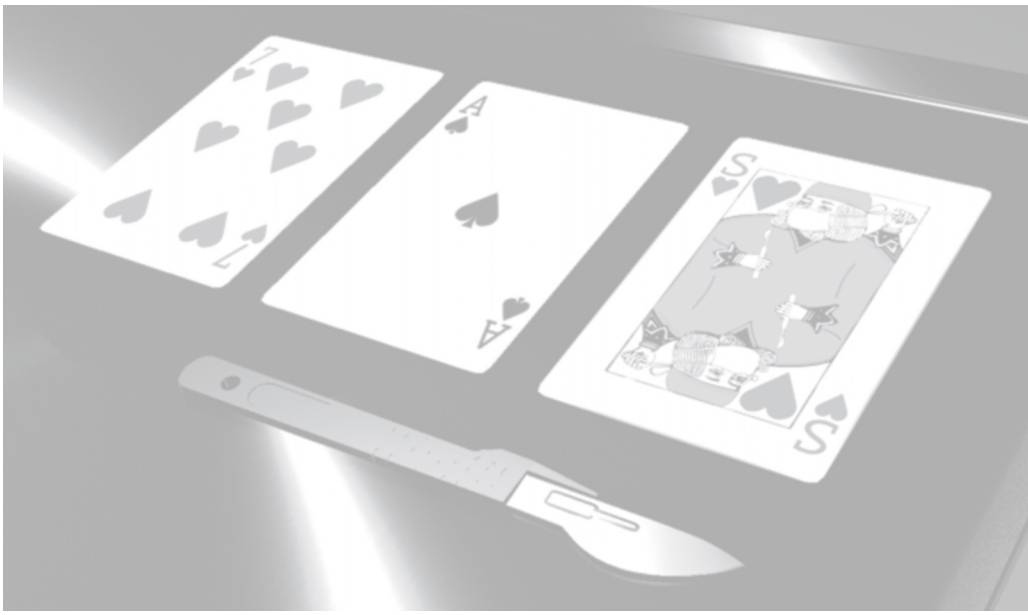
In conclusion, the SYNTAX score is the first coronary vasculature complexity score predictive for postoperative outcome in patients with left main coronary artery disease (isolated, or in combination with one-, two-, or three-vessel disease) undergoing CABG. The outcomes of the ongoing SYNTAX trial [19] will definitively define the role of the SYNTAX score in predicting short-term and long-term incidence of MACCE.

REFERENCES

1. Taggart DP, Kaul S, Boden WE, et al. Revascularization for unprotected left main stem coronary artery stenosis stenting or surgery. J Am Coll Cardiol 2008;51:885–92.
2. Smith SC Jr, Dove JT, Jacobs AK, et al. ACC/AHA guidelines of percutaneous coronary interventions (revision of the 1993 PTCA guidelines)—executive summary. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (committee to revise the 1993 guidelines for percutaneous transluminal coronary angioplasty). J Am Coll Cardiol 2001;37:2215–39.
3. Hoffman SN, TenBrook JA, Wolf MP, Pauker SG, Salem DN, Wong JB. A meta-analysis of randomized controlled trials comparing coronary artery bypass graft with percutaneous transluminal coronary angioplasty: one- to eight-year outcomes. J Am Coll Cardiol 2003;41:1293–304.
4. Serruys PW, Ong AT, van Herwerden LA, et al. Five-year outcomes after coronary stenting versus bypass surgery for the treatment of multivessel disease: the final analysis of the Arterial Revascularization Therapies Study (ARTS) randomized trial. J Am Coll Cardiol 2005;46:575– 81.
5. Stone GW, Ellis SG, Cox DA, et al. One-year clinical results with the slow-release, polymer-based, paclitaxel-eluting TAXUS stent: the TAXUS-IV trial. Circulation 2004;109: 1942–7.
6. Morice MC, Serruys PW, Sousa JE, et al. A randomized comparison of a sirolimus-eluting stent with a standard stent for coronary revascularization. N Engl J Med 2002;346:1773–80.
7. Grube E, Silber S, Hauptmann KE, et al. TAXUS I. Six- and twelve-month results from a randomized, double-blind trial on a slow-release paclitaxel-eluting stent for de novo coronary lesions. Circulation 2003;107:38–42.
8. Ferguson TB Jr, Hammill BG, Peterson ED, DeLong ER, Grover FL. A decade of change—risk profiles and outcomes for isolated coronary artery bypass grafting procedures, 1990–1999: a report from the STS National Database Committee and the Duke Clinical Research Institute. Society of Thoracic Surgeons. Ann Thorac Surg 2002;73:480 –90.
9. Nishida H, Tomizawa Y, Endo M, Kurosawa H. Survival benefit of exclusive use of in situ arterial conduits over combined use of arterial and vein grafts for multiple coronary artery bypass grafting. Circulation 2005;112(9 Suppl):I- 299–303.
10. Ong AT, Serruys PW, Mohr FW, et al. The SYNergy between percutaneous coronary intervention with TAXus and cardiac surgery (SYNTAX) study: design, rationale, and run-in phase. Am Heart J 2006;151:1194 –204.
11. Sianos G, Morel M, Kappetein A, et al. The Syntax score: an angiographic tool grading the complexity of coronary artery disease. Eurointervention 2005;1:219 –27.
12. Valgimigli M, Serruys PW, Tsuchida K, et al. Cyphering the complexity of coronary artery disease using the syntax score to predict clinical outcome in patients with three-vessel lumen obstruction undergoing percutaneous coronary intervention. Am J Cardiol 2007;99:1072–81.
13. Roques F, Nashef SA, Michel P, et al. Risk factors and outcome in European cardiac surgery: analysis of the Euro- SCORE multinational database of 19030 patients. Eur J Cardiothorac Surg 1999;15:816 –23.

14. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation 1989;79(6 Pt 2):I-3– 12.
15. Leaman DM, Brower RW, Meester GT, Serruys P, van den Brand M. Coronary artery atherosclerosis: severity of the disease, severity of angina pectoris and compromised left ventricular function. Circulation 1981;63:285–99.
16. Fonarow GC, Adams KF Jr, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. JAMA 2005;293:572– 80.
17. Antman EM, Cohen M, Bernink PJ, et al. The TIMI risk score for unstable angina/non-ST elevation MI. A method for prognostication and therapeutic decision making. JAMA 2000;284:835– 42.
18. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453–73.
19. Kappetein AP, Dawkins KD, Mohr FW, et al. Current percutaneous coronary intervention and coronary artery bypass grafting practices for three-vessel and left main coronary artery disease. Insights from the SYNTAX run-in phase. Eur J Cardiothorac Surg 2006;29:486 –91.

CHAPTER 11



Impact of the definition of renal dysfunction on Euro-SCORE performance

Van Gameren M
Klieverik LMA
Struijs A
Venema AC
Kappetein AP
Bogers AJJC
Takkenberg JJM

J Cardiovasc Surg (Torino). 2009; 50: 703-9

ABSTRACT

Background

Renal dysfunction is an important variable in the EuroSCORE (European System for Cardiac Operative Risk Evaluation) model and is currently defined as 'creatinine >200 µmol/l'. The aim of this study was to examine whether using other definitions of renal dysfunction could improve the predictive ability of the EuroSCORE.

Methods

Between January 2004 and January 2006, 1205 patients underwent cardiac surgery. Their preoperative glomerular filtration rate and EuroSCORE were calculated. Four recalibrated EuroSCORE models were constructed using (1) creatinine as a binary variable, (2) creatinine as a continuous variable, (3) glomerular filtration rate as a categorical variable, or (4) glomerular filtration rate as a continuous variable. The predictive ability of these models was assessed using receiver operating characteristic curve analysis.

Results

Hospital mortality was 4% (n=47). Receiver operating characteristic curve values were: 0.78 for the original EuroSCORE, 0.80 for the recalibrated binary creatinine model, 0.83 for the continuous creatinine model, 0.83 for the categorical glomerular filtration rate model, and 0.82 for the continuous glomerular filtration rate model.

Conclusion

The use of creatinine as a continuous variable or glomerular filtration rate as a categorical or continuous variable improves the predictive accuracy of the EuroSCORE model for hospital mortality. Given the increasing incidence of preoperative renal dysfunction and its impact on hospital mortality, future risk stratification models should include continuous creatinine or glomerular filtration rate rather than creatinine as a binary variable.

BACKGROUND

Inclusion of binary variables and thereby ease of usage was one of the main objectives during the development of the EuroSCORE (European System for Cardiac Operative Risk Evaluation) model [1]. In 1999, this resulted in a simple additive model based upon 17 independent predictors. Shortly after, studies however showed that the additive EuroSCORE model overestimates mortality in low-risk patients, while it underestimates mortality in high risk patients [2, 3]. For this reason, the logistic EuroSCORE model was published and recommended to be used instead of the additive model [4]. Although this model indeed demonstrated an improved performance, it nowadays proves to generally overestimate mortality [5]. Recently, plans were released to develop a new EuroSCORE model that reflects contemporary practice (http://www.euroscore.org/EuroSCORE2008.htm).

To produce a more accurate EuroSCORE model, existing binary variables could be replaced by continuous ones. One of these variables is renal dysfunction, defined in the original EuroSCORE model as serum creatinine exceeding 200 µmol/l. As renal dysfunction is an important predictor of mortality and morbidity after cardiac surgery [6-9], replacing this binary variable may improve EuroSCORE performance in particular.

Several studies claim that glomerular filtration rate (GFR) is a better predictor than creatinine as an independent risk factor for adverse outcome after cardiac surgery [10, 11]. This could imply that EuroSCORE performance improves further when GFR replaces the binary creatinine variable.

We sought to determine whether the replacement of 'creatinine >200 µmol/l' by continuous creatinine, categorical GFR, or continuous GFR improves the performance of the EuroSCORE model to predict hospital mortality after cardiac surgery.

MATERIALS AND METHODS

Study design and data collection

Data were collected prospectively on all patients undergoing adult cardiac surgery in our hospital between January 2004 and January 2006. All collected variables were compliant with published EuroSCORE model variable definitions [1].

Whenever patients underwent multiple operations within one month or during the same hospital admission, only the first (index) procedure was included. No restrictions on type of surgery were applied. The dependent variable (outcome of interest) was in-hospital mortality, defined as all cause mortality during hospital admission.

Analysis plan

Four models based upon the EuroSCORE model were created. The first was constructed by recalibrating the logistic EuroSCORE to fit our population, still using the original 'creatinine >200 µmol/l' variable. Three additional recalibrated EuroSCORE models were constructed; in one model creatinine as a binary variable was substituted by creatinine as a continuous variable, in another model by GFR as a categorical variable and in the last model by GFR as a continuous variable. Recalibration was performed by re-estimating all model coefficients (including intercept) using logistic regression analysis.

Finally, performance of the original and the recalibrated EuroSCORE models was compared.

GFR calculation

Preoperative GFR was calculated according to Cockcroft and Gault [12]. The following equation was used for men:

*GFR = ((140 – age in years) * bodyweight in kilograms) / (0.81 * serum creatinine in µmol/l)*

As the female body generally has a lower percentage of muscle mass, the calculated GFR was multiplied by 0.85 in female patients, as determined by Cockroft and Gault [12].

In one of the produced models, GFR was inserted as a categorical variable; values were categorized according to the five stages of chronic kidney disease according to the Kidney Disease Outcome Quality Initiative [13]. These stages are shown in Table 1.

Statistical analysis

Continuous data are presented as mean ± 1 standard deviation, and median. Categorical data are presented as proportions.

Table 1. National Kidney Foundation Kidney Disease Outcomes Quality Initiative Classification, Prevalence, and Action Plan for Stages of Chronic Kidney Disease.

Stage	Description	GFR, mL/min per 1.73 m^2	Action
1	Kidney damage with normal or increased GFR	≥90	Diagnosis and treatment; treatment of comorbid conditions; slowing progression; CVD risk reduction
2	Kidney damage with mild decreased GFR	60–89	Estimating progression
3	Moderately decreased GFR	30–59	Evaluating and treating complications
4	Severely decreased GFR	15–29	Preparation for kidney replacement therapy
5	Kidney failure	<15 (or dialysis)	Kidney replacement (if uremia present)

CVD = Cardiovascular Disease; GFR = Glomerular Filtration Rate.

The study population was compared with the EuroSCORE population for prevalence of risk factors. Chi-square testing was used for comparison of categorical variables. Continuous variables were assessed using the Student's t-test. P-values under 0.05 were considered significant.

Univariable logistic regression was used to identify potential risk factors for hospital mortality. Predictors with p-values below 0.10 were entered into a multivariable regression model using backward conditional stepping and limited by 5-9 events per variable, as proposed by Vittinghoff and McCulloch [14]. Predictors with a p-value below 0.10 were considered significant. To assess risk-factor weightings, odds ratio (OR) values were calculated for all significant predictors.
The fit of the logistic regression model was measured in terms of its discrimination and calibration. Discrimination, which is measured using the c-index (area under the receiver operating characteristic curve) with 95% confidence limits (CI), captures the model's ability to distinguish between patients who die in the hospital and patients who are discharged alive. It is defined as the proportion of the time that a patient who survives is assigned a higher probability of survival than a patient who dies in the hospital. A value of 1.0 is perfect, and a value of 0.5 denotes only random ability to distinguish between deaths and survivors.
To demonstrate significant differences between c-indices, a bootstrapping cycle

of 2000 runs was performed [15]. Tests between c-indices were 2 sided with $p<0.05$ considered to be a significant difference.

Calibration was evaluated by the Hosmer-Lemeshow goodness-of-fit test (H-L) and graphically by a calibration plot. The smooth curve in a calibration plot reflects the non-parametric relation between observed mortality and predicted probability of mortality. Perfect calibration is represented by the straight dotted line through the origin. Triangles are based on quintiles of patients with similar predicted probabilities. Spikes at the bottom of a calibration plot represent the distribution of predicted probabilities. Models with Hosmer-Lemeshow p-values above 0.05 were considered to be well calibrated.

Basic statistical analyses were performed with SPSS version 15.0 (SPSS, Chicago, Illinois). R version 2.5.1 (The R Foundation for Statistical Computing) was used for bootstrapping, calculating c-values with 95% CI, Hosmer-Lemeshow p-values and constructing receiver operating characteristic (ROC) curves and calibration plots.

RESULTS

Study population

A total of 1205 patients were included in the study and their overall in-hospital mortality was 4% (N=47). Mean patient age was 62 years (range 18-87 years). Patient characteristics are presented in Table 2.

Isolated coronary surgery was performed in only 36% (n=440) of the procedures, while this was 64.% for the EuroSCORE population. Combined coronary and valve surgery was performed in 14% (n=172), isolated valve surgery in 40% (n=488), and 9% (n=105) were other procedures.

Table 2 reveals several other evident differences between the study population and the EuroSCORE population. Chronic pulmonary disease and pulmonary hypertension were more prevalent in our population and more of our patients underwent surgery on the thoracic aorta or had previously undergone cardiac surgery.

Independent risk factors for hospital mortality

Four separate logistic regression analyses were undertaken. Each analysis included all collected EuroSCORE variables, but only one of the investigated renal

Table 2. Prevalence of risk factors in our population and the EuroSCORE population.

	Rotterdam (%)	EuroSCORE (%)	p-value
N=	1205	19030 [21]	
Age	62.0 (SD 14.1, 18-87)	62.5 (SD 10.7, 17-94)	
Female	31.1	27.8	0.003
Chronic pulmonary disease	8.5	3.9	0.000
Extracardiac arteriopathy	8.3	11.3	0.001
Neurological dysfunction disease	0.7	1.4	0.058
Previous cardiac surgery	12.7	7.3	0.000
Renal function			
Serum creatinin >200 µmol/l	1.9	1.8	0.788
Serum creatinin mean	98.6 (SD 52.0, 32-896)		-
Calculated GFR mean	81.6 (SD 31.0, 7.5-208.1)		-
Calculated GFR			
≥90	34.1		-
60–89	39.7		-
30–59	24.3		-
15–29	1.5		-
<15	0.4		-
Active endocarditis	2.2	1.1	0.001
Critical preoperative state	6.7	4.1	0.000
Unstable angina	6.5	8.0	0.057
Poor LV function (EF <30%)	4.4	5.8	0.042
Recent myocardial infarct	9.5	9.7	0.785
Pulmonary hypertension	3.7	2.0	0.000
Emergency	6.1	4.9	0.072
Other than isolated CABG	63.5	36.4	0.000
Surgery on thoracic aorta	6.9	2.4	0.000
Postinfarct septal rupture	0.2	0.2	0.798

CABG = Coronary Artery Bypass Grafting; EF = Ejection Fraction; GFR = Glomerular Filtration Rate; LV = Left Ventricular; SD = Standard Deviation.

dysfunction variables at the time. Continuous GFR and both binary and continuous creatinine were significant independent predictors of hospital mortality (OR 1.02 (95%CI 1.01-1.03), OR 3.78 (95%CI 1.08-13.19) and OR 1.005 (95%CI 1.002-1.007) respectively). Categorical GFR was also a significant independent predictor of hospital mortality: compared to the reference group with a normal GFR

(≥90), patients had an increasing higher risk of hospital mortality with decreasing GFR: GFR 60-89 OR 1.53 (0.63-3.67), GFR 30-59 OR 3.89 (1.70-8.91), GFR 15-29 OR 10.00 (2.41-41.52), and GFR<15 OR 12.50 (1.25-124.7).

Model performance: discrimination

C-statistics were: 0.78 for the original logistic EuroSCORE, 0.80 for the recalibrated EuroSCORE, 0.83 for both the recalibrated EuroSCORE with continuous creatinine as with categorical GFR, and 0.82 for the recalibrated EuroSCORE with continuous GFR. These results are listed with 95% CI's in Table 3. A graphic representation in the form of ROC curves is presented in Figure 1.

The recalibrated EuroSCORE models with continuous creatinine, categorical GFR or continuous GFR all showed a better discriminatory performance than the original EuroSCORE (p=0.001, p=0.040 and p=0.006 respectively). The recalibrated EuroSCORE (creatinine still as a binary variable) c-statistic was not significantly different from the original EuroSCORE.

While the recalibrated EuroSCORE with creatinine as a continuous variable had a significantly higher c statistic than the recalibrated EuroSCORE with creatinine as a binary variable (p=0.01), the discriminatory advantage of the recalibrated EuroSCORE models with continuous or categorical GFR over the recalibrated EuroSCORE with creatinine as a binary variable is not significant.

Model performance: calibration

All recalibrated models show a better calibration than the original EuroSCORE

Table 3. model performance.

Logistic EuroSCORE model	Predicted mortality	Discrimination		Calibration	
		AUC	95% CI	Hosmer-Lemeshow	Chi-Square
Original EuroSCORE	6.4%	0.77	0.72 - 0.84	0.034	18.09
EuroSCORE binary creatinin	4.2%	0.80	0.74 - 0.86	0.35	10.05
EuroSCORE continuous creatinin	4.2%	0.83	0.78 - 0.89	0.66	6.75
EuroSCORE categorical GFR	4.1%	0.83	0.78 - 0.88	0.58	7.52
EuroSCORE continuous GFR	4.0%	0.82	0.77 - 0.87	0.73	6.11

AUC = Area Under Curve; CI = Confidence Interval; EuroSCORE = European System for Cardiac Operative Risk Evaluation; H-L = Hosmer-Lemeshow.

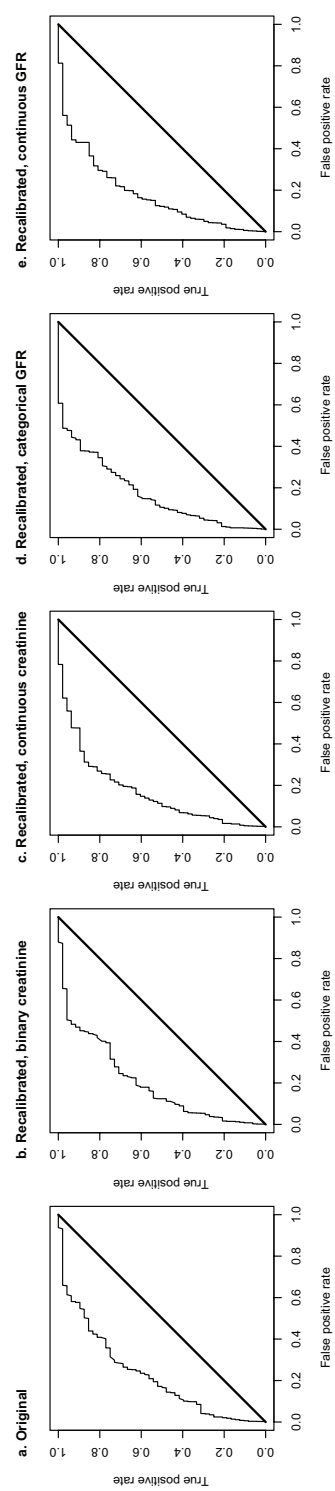


Figure 1. Discrimination

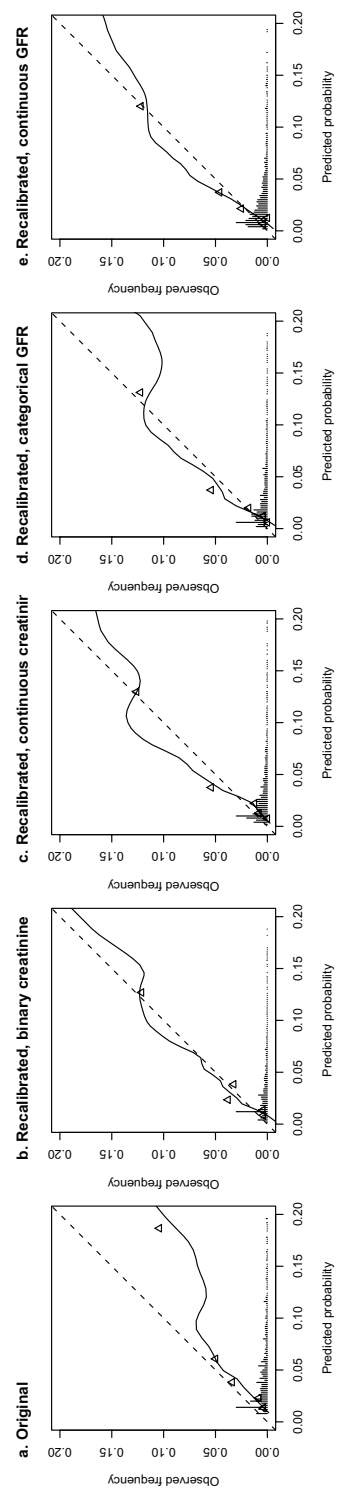


Figure 2. Calibration.

(see Table 3). When comparing the four recalibrated models mutually, the models with creatinine as a continuous variable, GFR as a categorical value or GFR as a continuous variable show higher Hosmer-Lemeshow p-values than the model with creatinine as a binary variable. The calibration plots (see Figure 2) confirm these findings.

DISCUSSION

This study has shown that the ability of the EuroSCORE model to predict hospital mortality increases when its binary creatinine variable is replaced by creatinine as a continuous variable, by GFR as a categorical variable or by GFR as a continuous variable.

Recalibration

All recalibrated EuroSCORE models performed better than the original EuroSCORE model. No relevant conclusions can be drawn from this comparison as they have been adjusted for our population, and thereby also corrected for a decade of improved care (the EuroSCORE is based upon a cohort from 1995). A comparison of discrimination and calibration of the four recalibrated EuroSCORE models with each other has to be made to evaluate which variable improves the EuroSCORE most.

More valve surgery patients were present in our cohort than in the EuroSCORE population, perhaps also affecting results. A relative decrease of coronary surgery of over 10% throughout Europe since 1995 might partly explain this difference [16].

Mild renal dysfunction

When using the original EuroSCORE model to predict hospital mortality, increased risk related to renal dysfunction is assigned to patients with high serum creatinine levels only. Previous studies demonstrated that mild renal dysfunction (with creatinine levels as low as 130 μmol/l) is a risk factor for adverse outcome after coronary surgery as well [17, 18]. A significant increase in the need for mechanical renal support, total postoperative stay, duration of special care and short term mortality was found in these patients [18]. With continuous creatinine levels incorporated in the EuroSCORE, these higher risk patients should be

better identified.

Categorical glomerular filtration rate approach

The use of continuous variables in risk stratification models usually outperforms the use of binary or categorical variables. Calculated GFR values based upon creatinine (like the Cockroft and Gault equation) are accurate enough to successfully assign patients to the correct stage of kidney disease. They may however not be precise and accurate enough (especially in the normal range of GFR) to directly use as a continuous variable [11]. For this reason, a model with a categorical GFR variable was examined as well.

In this study, patients without kidney disease are assigned to the stage I category. When looking at the five stages of kidney disease however, even stage I is defined as disease being present. Results are not influenced by assigning these patients to stage I, as only patients with a GFR of at least 60 (stage II and onwards) are considered at increased risk [13].

Glomerular filtration rate versus creatinine as a predictor

According to the STS database, preoperative renal dysfunction (assessed using GFR) is a common finding and a powerful predictor of mortality and morbidity [9]. This study supports this finding, both when GFR was used as an independent predictor as when used as part of a EuroSCORE model. The predictors themselves were not compared in this study, only the risk models they were incorporated in.

Wang et al. have previously compared creatinine level and GFR as predictors for adverse events after cardiac surgery. They found that GFR was a better predictor than creatinine level, particularly in patients with normal creatinine levels [10].

Renal dysfunction incorporated in the EuroSCORE model

The effect of replacing creatinine by binary or continuous GFR in the EuroSCORE model was investigated in a previous study [19]. In that study, GFR-models performed better than the binary creatinine model when looking at discrimination alone. Significant differences were not tested. More recently, Jin et al. also recommended the use of (estimated) GFR instead of creatinine as a risk factor in risk models [20].

In this study, the EuroSCORE models with GFR (categorical or continuous)

were not superior to the EuroSCORE model with continuous creatinine, despite better results in previous studies. This could be explained by combining the EuroSCORE model and the Cockroft-Gault equation. When comparing GFR and creatinine when incorporated in the EuroSCORE model, overlapping variables should be taken into account. The possible advantage of GFR over creatinine as an independent risk factor is partially eliminated, since 3 out of 4 Cockroft-Gault variables are incorporated in the EuroSCORE as well. Age, female sex and (continuous) creatinine form the base of the calculated GFR, resulting in bodyweight and the interdependent relation of these variables as the only new aspects of GFR incorporation in the EuroSCORE.

Limitations

When interpreting the results from this study, one should keep in mind that data was used from a single center. Hospital and surgeon related factors might therefore have resulted in study results that are not applicable to other European centers or a future model based upon data from these centers.

Implementation

Optimal implementation of risk models in clinical practice, serving the clinic rather than increasing its work load, is a key factor for success and will result in continuous quality improvement. As optimal implementation also comprises a periodical update of the model, the new EuroSCORE model is eagerly awaited.
Whether or not neural networking, submodel usage (for coronary surgery, valve surgery and combined coronary and valve surgery), additional variables, different endpoints or other improvements will be applied, the use of continuous variables will most likely improve model performance on their own.
Our modest cohort already reveals positive results, thus contributing to this hypothesis.

Conclusion

Given the high incidence of preoperative renal dysfunction and its impact on hospital mortality, future risk stratification models like a new EuroSCORE should include GFR or continuous creatinine rather than creatinine as a binary variable.

Apart from the inclusion of additional procedural variables and possibly the use of submodels, implementation of a minor modification as demonstrated in

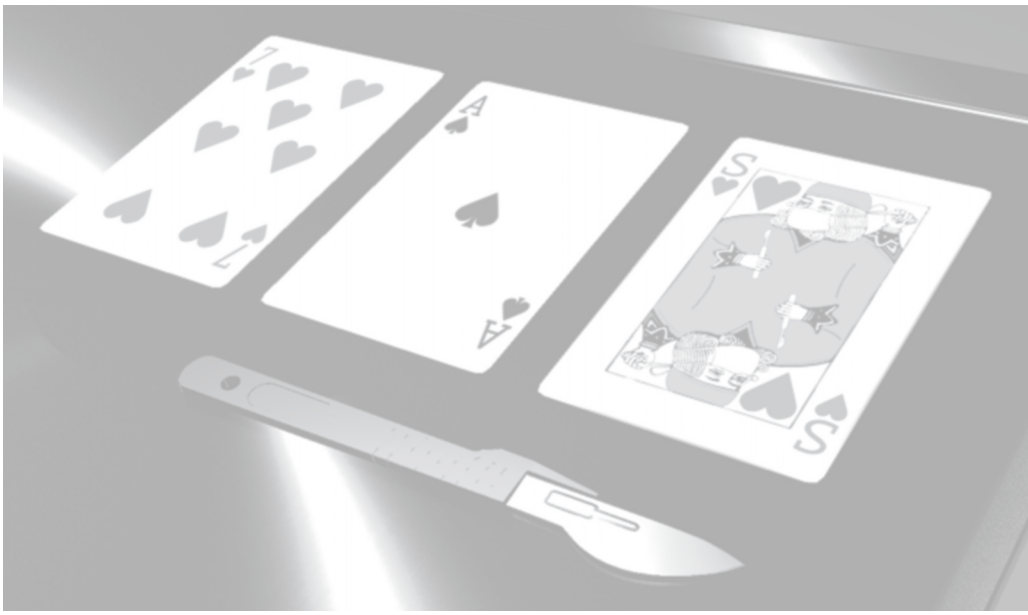
this study will provide European surgeons with a much improved tool for benchmarking and prediction of mortality after cardiac surgery.

REFERENCES

1. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S and Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. Jul 1999;16(1):9-13.
2. Gogbashian A, Sedrakyan A and Treasure T. EuroSCORE: a systematic review of international performance. Eur J Cardiothorac Surg. May 2004;25(5):695-700.
3. Karthik S, Srinivasan AK, Grayson AD, Jackson M, Sharpe DA, Keenan DJ, et al. Limitations of additive EuroSCORE for measuring risk stratified mortality in combined coronary and valve surgery. Eur J Cardiothorac Surg. Aug 2004;26(2):318-22.
4. Roques F, Michel P, Goldstone AR and Nashef SA. The logistic EuroSCORE. Eur Heart J. May 2003;24(9):881-2.
5. Bhatti F, Grayson AD, Grotte G, Fabri BM, Au J, Jones M, et al. The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk? Heart. Dec 2006;92(12):1817-20.
6. Hillis GS, Croal BL, Buchan KG, El-Shafei H, Gibson G, Jeffrey RR, et al. Renal function and outcome from coronary artery bypass grafting: impact on mortality after a 2.3-year follow-up. Circulation. Feb 28 2006;113(8):1056-62.
7. Devbhandari MP, Duncan AJ, Grayson AD, Fabri BM, Keenan DJ, Bridgewater B, et al. Effect of risk-adjusted, non-dialysis-dependent renal dysfunction on mortality and morbidity following coronary artery bypass surgery: a multi-centre study. Eur J Cardiothorac Surg. Jun 2006;29(6):964-70.
8. Zimmet AD, Almeida A, Goldstein J, Shardey GC, Pick AW, Lowe CE, et al. The outcome of cardiac surgery in dialysis-dependent patients. Heart Lung Circ. Sep 2005;14(3):187-90.
9. Cooper WA, O'Brien SM, Thourani VH, Guyton RA, Bridges CR, Szczech LA, et al. Impact of renal dysfunction on outcomes of coronary artery bypass surgery: results from the Society of Thoracic Surgeons National Adult Cardiac Database. Circulation. Feb 28 2006;113(8):1063-70.
10. Wang F, Dupuis JY, Nathan H and Williams K. An analysis of the association between preoperative renal dysfunction and outcome in cardiac surgery: estimated creatinine clearance or plasma creatinine level as measures of renal function. Chest. Nov 2003;124(5):1852-62.
11. Coresh J and Stevens LA. Kidney function estimating equations: where do we stand? Curr Opin Nephrol Hypertens. May 2006;15(3):276-84.
12. Cockcroft DW and Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron. 1976;16(1):31-41.
13. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis. Feb 2002;39(2 Suppl 1):S1-266.
14. Vittinghoff E and McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. Mar 15 2007;165(6):710-8.
15. Efron B and Tibshirani R. An Introduction to the bootstrap. Chapman & Hall, New York, 1993.
16. Keogh BE. Third European Adult Cardiac Surgical Database Report. Dendrite Clinical Systems, Oxfordshire, 2007.
17. Zakeri R, Freemantle N, Barnett V, Lipkin GW, Bonser RS, Graham TR, et al. Relation between mild renal dysfunction and outcomes after coronary artery bypass grafting. Circulation. Aug

30 2005;112(9 Suppl):I270-5.
18. Weerasinghe A, Hornick P, Smith P, Taylor K and Ratnatunga C. Coronary artery bypass grafting in non-dialysis-dependent mild-to-moderate renal dysfunction. J Thorac Cardiovasc Surg. Jun 2001;121(6):1083-9.
19. Walter J, Mortasawi A, Arnrich B, Albert A, Frerichs I, Rosendahl U, et al. Creatinine clearance versus serum creatinine as a risk factor in cardiac surgery. BMC Surg. 2003;3:4.
20. Jin R, Grunkemeier GL, Brown JR and Furnary AP. Estimated glomerular filtration rate and renal function. Ann Thorac Surg. Jul 2008;86(1):1-3.
21. Nashef SA, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. Eur J Cardiothorac Surg. Jul 2002;22(1):101-5.

CHAPTER 12



Parsimony and Performance

Van Gameren M
Kappetein AP
Bogers AJJC
Takkenberg JJM

Circ Cardiovasc Qual Outcomes. 2009; 2: 583-90: online reaction

PARSIMONY AND PERFORMANCE

Dr Mehta and colleagues [1] provided an informative overview of incidence, risk factors, time trends, and outcomes regarding reoperations for bleeding after coronary surgery.

In addition, they performed a logistic regression analysis to provide a model that predicts the risk of reoperation for bleeding. The resulting model was derived from an impressive database and consists of 19 variables. With a c statistic of only 0.60, this model showed a very modest discrimination. A c statistic of 1 indicates perfect discrimination while 0.5 equals flipping a coin. Generally, prediction models with a c statistic of at least 0.75 are considered to discriminate well.

A more parsimonious bedside tool was also constructed based upon the derived model. The number of variables was reduced to 12, but discriminatory performance was not reported.

As the study performed by Dr. Mehta and colleagues already contains most known important preoperative variables that influence bleeding, perhaps preoperative variables only play a limited role in predicting the risk of reoperation for bleeding. When keeping the tradeoff between parsimony and performance in mind, either discriminatory performance should be better or the level of parsimony of both model and bedside tool should be higher to justify their implementation in clinical practice.

One of the variables used is serum creatinin level. As this is a continuous variable, readily available, and a well known predictor for many adverse outcomes, we chose to construct a model that predicts the risk of reoperation for bleeding based upon this variable alone. Between January 2003 and January 2008, 1873 patients underwent coronary surgery in our center, of which 108 (5.8%) required a reoperation for bleeding. We used 80% of the cohort for the development of the model and the remaining 20% of the cohort for model validation. The derived model already obtained a c statistic of 0.63 when applied to the validation cohort.

Although these results are from a cohort of different size and characteristics, comparable results are likely to be obtained when applying this approach to the dataset used by Dr Mehta and colleagues.

In other words: the simple clinical rule that increased serum creatinin levels result in a higher risk of reoperation for bleeding could well have a discrimina-

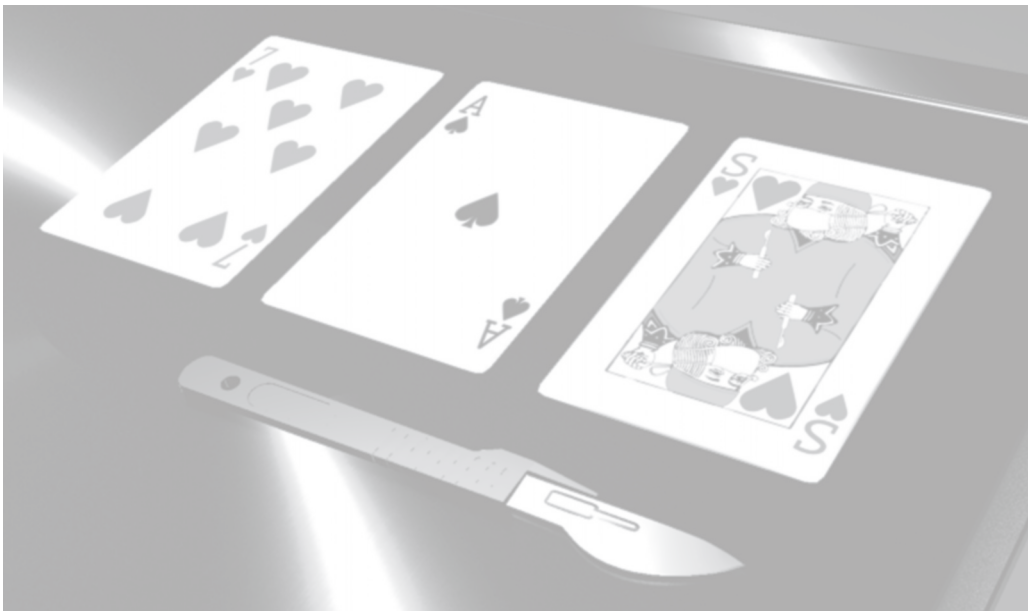
tory capability comparable to the model described by Mehta et al.
What are in Dr Mehta's study the discriminatory results for the bedside tool and what c statistic is obtained when (for example) only serum creatinin is used in the model?

In any case one can argue that with a c statistic of only 0.60 common clinical sense will likely outperform preoperative models that predict the risk of reoperation for bleeding after coronary surgery in terms of parsimony and discrimination.

REFERENCES

1. Mehta. Reoperation for Bleeding in Patients Undergoing Coronary Artery Bypass Surgery. Circ outcomes. 2009.

CHAPTER 13



Optimizing intensive care unit capacity using preoperative prediction models for length of stay after cardiac surgery: tree modeling versus linear regression

Van Gameren M
Van der Pijl R
Wu Y
Grunkemeier GL
Kappetein AP
Bogers AJJC
Takkenberg JJM

In preparation

ABSTRACT

Background

Prediction models are often used to benchmark center and surgeon performance and to support clinicians in treatment selection and patient counseling. They can also be used to predict length of stay (LOS) in the intensive care unit after cardiac surgery. This study explores the use of models to predict LOS in the intensive care unit for adult patients requiring CABG and/or valve replacement during the preoperative planning process.

Methods

A regression tree model and a linear regression model were constructed using a cohort of 2540 patients to identify potential predictors. Performance of these models was assessed by calculating the total reduction of under- and overestimated days in a validation cohort (N=948). Results in terms of under- and overestimation were then compared for each model and with results from an approach without a model.

Results

Procedure type, critical preoperative state, serum creatinin level and patient age were identified as the most important predictors of increased LOS in the study cohort. All variables present in the regression tree are also found in the linear regression model. A decrease in underestimation and overestimation of length of stay was achieved by both the tree model (6 and 225 days respectively) and the linear regression model (20 and 194 days respectively).

Conclusions

Predictions models using individual patient data can be used to predict LOS in the intensive care unit after cardiac surgery. This study demonstrates that a decrease in underestimation and overestimation of LOS can be achieved by using them. When taking its simplicity into account, constructing a regression tree seems an ideal first step when implementing a model to predict LOS.

INTRODUCTION

Available intensive care unit (ICU) capacity is among the hospital resources that greatly influence patient planning for cardiac surgery [1]. Apart from the number of beds and nursing staff availability, ICU capacity also depends upon the length of stay (LOS) of the postoperative patients. As ICU capacity is limited, optimal patient planning is required to maximize its use.

By preoperatively predicting individual patient LOS, the operating schedule can be adjusted accordingly to support an effective throughput of patients in the ICU. For instance, patients who are likely to occupy an ICU bed for a longer period, can be spread over time (and should preferably not be operated upon when few ICU beds are available). Planning is further improved when an anticipated date at which an ICU bed will once again become available is known. In addition, preoperative information on expected LOS facilitates scheduling of individual surgical procedures on specific dates.

Nowadays, the use of clinical prediction models in cardiac surgery is common practice. These models are most often used to benchmark center and surgeon performance and to support clinicians in treatment selection and patient counseling. With the importance of optimizing ICU capacity in mind, it seems natural to also integrate results from prediction models based upon patient and procedural factors in preoperative patient planning.

This study aims to explore the possibility of using prediction models to optimize patient planning and maximize ICU capacity. An easy to interpret and implement regression tree model was therefore constructed to predict LOS. In addition, a linear regression model was constructed for the same purpose. The potential gain of these models was then tested in a validation cohort. Results in terms of under- and overestimation of LOS were compared mutually and with results from an approach in which no model was used.

METHODS

Data

All adult patients who underwent elective coronary artery bypass grafting, valve surgery, or a combination of these procedures in the Erasmus University Medical Center between January 2003 and January 2007 were included in the study

Table 1. Patient and procedural characteristics.

	Development Cohort (N=2540)		Validation Cohort (N=948)		p-value
Age, mean (SD)	63.9 years	(12.0)	65.5 years	(11.7)	< 0.001
Female sex	29.6%	753	31.6%	300	0.270
Diabetes	17.9%	455	23.6%	224	< 0.001
Previous cardiac surgery	8.7%	221	8.2%	78	0.707
Serum creatinin, mean (SD)	95.7 umol/l	(48.0)	98.6 umol/l	(71.7)	0.214
Cerebrovascular accident	4.0%	101	3.5%	33	0.563
Infectious endocarditis	1.1%	27	1.4%	13	0.599
COPD	8.3%	212	9.0%	85	0.606
Extracardiac arteriopathy	9.1%	230	9.5%	90	0.739
Neurological dysfunction	0.5%	13	0.5%	5	0.835
Procedure type					
Coronary surgery	55.6%	1412	56.8%	538	0.565
Valve surgery	30.9%	784	27.7%	263	0.080
Valve + coronary	13.5%	344	15.5%	147	0.153
Recent MI (<90 days)	10.2%	258	15.7%	149	< 0.001
Unstable AP	6.2%	158	6.2%	59	0.940
Pulmonary hypertension	3.1%	79	1.9%	18	0.096
Systolic left ventricular function					
Good	84.4%	2143	84.3%	799	0.991
Moderate	12.6%	321	11.9%	113	0.607
Poor	2.9%	73	3.8%	36	0.199
Critical preoperative state	3.2%	81	3.5%	33	0.746

AP = angina pectoris; COPD = chronic obstructive pulmonary disease; MI = myocardial infarction; SD = standard deviation.

development cohort (N=2540).

Patients undergoing concomitant procedures were excluded. Patients undergoing emergency procedures were also excluded; a reservation policy for patients undergoing emergency procedures is assumed. Whenever patients underwent multiple operations within one month or during the same hospital admission, only the first (index) procedure was included. Patient and procedural characteristics are presented in Table 1. Thirty-nine (1.5%) patients died within

Table 2. Linear regression analysis results for LOS in the ICU.

	Univariable		Multivariable	
	Coefficient (95% CI)	**p-value**	**Coefficient (95% CI)**	**p-value**
Critical preoperative state	0.900 (0.749 – 1.050)	< 0.0001	0.765 (0.622 – 0.908)	< 0.0001
Procedure type: valve + coronary surgery	0.504 (0.428 – 0.581)	< 0.0001	0.438 (0.364 – 0.511)	< 0.0001
Infectious endocarditis	0.653 (0.389 – 0.916)	< 0.0001	0.389 (0.142 – 0.636)	< 0.0001
Previous cardiac surgery	0.266 (0.171 – 0.362)	< 0.0001	0.329 (0.238 – 0.420)	< 0.0001
Pulmonary hypertension (>60 mmHg)	0.480 (0.325 – 0.635)	< 0.0001	0.309 (0.165 – 0.453)	< 0.0001
Recent MI	0.133 (0.043 – 0.222)	0.004	0.105 (0.022 – 0.188)	0.013
Female sex	0.060 (0.001 – 0.119)	0.047	0.083 (0.028 – 0.139)	0.003
Age (continuous)	0.008 (0.006 – 0.011)	< 0.0001	0.007 (0.005 – 0.009)	< 0.0001
Serum creatinin level (continuous)	0.003 (0.002 – 0.003)	< 0.0001	0.002 (0.002 – 0.003)	< 0.0001
Poor systolic left ventricular function	0.354 (0.192 – 0.516)	< 0.0001	-	
Cerebrovascular accident	0.123 (-0.016 – 0.262)	0.082	-	
Extracardiac arteriopathy	0.120 (0.025 – 0.214)	0.013	-	
COPD	0.100 (0.002 – 0.198)	0.045	-	
Unstable AP	0.110 (-0.002 – 0.222)	0.055	-	
Model intercept	-		-0.420	

AP = angina pectoris; COPD = chronic obstructive pulmonary disease; CI = confidence interval; MI = myocardial infarction.

30-days of the index procedure.

The department of cardiac surgery has 10 ICU beds and 4 operating rooms. The outcome variable of the present study was LOS, defined as the time in days between admission and discharge from the ICU. For patients discharged to the ward and readmitted to the ICU, the intervening stay on the ward was excluded in the LOS.

All variables were collected prospectively according to the EuroSCORE definitions [2], with age and serum creatinin level as unmodified continuous vari-

ables.

Continuous data are presented as mean ± SD, the median is also given for non-normal distributions. Categorical data are presented as proportions.

Data preparation: LOS distribution, variance stabilizing

Given the inherently skewed distribution of LOS, a natural log transformation (ln[LOS]) was applied prior to model construction [3]. The skewed distribution of LOS in the study population was confirmed by the Kolmogorov-Smirnov test ($p < 0.001$).

For a regression tree analysis, there is no a need to consider transforms of the predictor variables as the classification and regression tree (CART) algorithm examines only the rank order of the predictors and not their values; no assumptions are made regarding the underlying distribution of values of the predictor variables. This however does not hold true for the dependent variable as splits are assessed based on their ability to reduce within-node variance in the dependent variable. Stabilizing variance using the natural log transformation was therefore used for the endpoint in the regression tree analysis as well.

After construction of the models, a smearing factor to correct the natural logarithmic 'back transformation' bias was needed to obtain the estimated LOS [4].

Tree model construction

Collected variables in the development cohort were subjected to classification and regression tree (CART) analysis to identify the most important outcome predictors and to develop a regression tree. This method is based on recursive partitioning analysis and involves the segregation of different values of variables through a decision tree composed of progressive binary splits [ref Statpage Wu] + [5]. The binary splits in this tree resulted in the formation of several groups with a different mean LOS (ln [LOS]).

The number of cross validations used to create the CART model was set to 10. The full tree, that might be overfit, was then pruned according to the optimal complexity parameter [6].

The primary split of the tree indicates the variable that is the most important predictor for increased LOS.

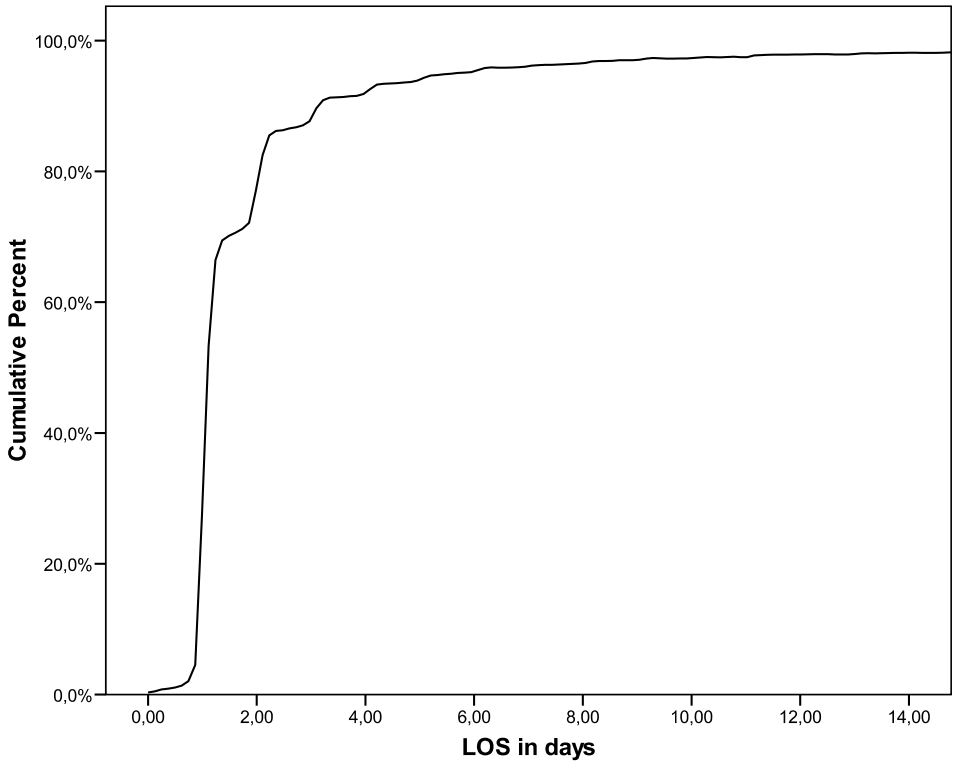


Figure 1. Cumulative distribution of LOS in days.

Linear regression model construction

To place the tree model results into perspective, a conventional univariable linear regression analysis using identical data was undertaken to test which of the variables contributed to LOS with $p < 0.10$. Significant variables in the univariable analyses were entered as potentially prognostic variables into a backward, stepwise selection procedure to construct a multivariable linear model that provides a natural logarithm transformed prediction of LOS (ln [LOS]). A p value below 0.05 was used to retain variables in the model. The predictive power of the model was expressed as the percentage explained variation (R2). This is a measure of the global fit of the model and represents the proportion of variability in LOS that may be attributed to the combination of the model variables.

Tree modeling was performed with R version 2.9.0 using the RPART library (R Foundation for Statistical Computing, Vienna, Austria). Calibration plotting was also performed with R version 2.9.0. All other analyses were performed with

SPSS for Windows (release 16.0; SPSS Inc, Chicago, Illinois).

Model application

After internal validation of the models, the potential gain in terms of usage in routine clinical practice was assessed in a second cohort. This cohort consisted of patients that underwent elective coronary artery bypass grafting, valve surgery, or a combination of these procedures in the Erasmus University Medical Center between January 2007 and July 2008. Exclusion criteria were identical to those used for the development cohort. This validation cohort consisted of 948 patients. Patient and procedural characteristics are presented in Table 1.

Differences between cohort characteristics were assessed using the Pearson Chi-square test for categorical variables and the Mann-Whitney test for continuous variables (skewed distributions confirmed by the Kolmogorov-Smirnov test). Patients in the validation cohort were slightly older and more suffered from diabetes and recent myocardial infarctions. No other differences were found between the two cohorts.

The prediction models were assessed by comparing the total overestimation and underestimation of the required ICU days if the mean LOS was used (no model) with the total overestimation and underestimation of the required ICU days if the prediction models were used (tree model, the primary split of the tree model only, and the linear regression model). The overestimation and underestimation for the no model approach were calculated by subtracting the observed LOS from the mean LOS. The overestimation and underestimation for the model approaches were calculated by subtracting the observed LOS from the predicted LOS. The mean LOS was used rather than the median LOS, because use of the median would favour the prediction model because the LOS is skewed. Therefore, use of the mean LOS will result in a more conservative gain in comparison with the median LOS.

Each approach has three possible outcomes: negative, indicating that the ICU bed was reserved for too long and that the number of ICU days was overestimated; zero, indicating perfect prediction; or positive, indicating that the ICU bed was reserved for an insufficient period and that the number of ICU days was underestimated.

Table 3. Model application results: underestimated and overestimated LOS in the validation cohort (N=948) using either the primary tree split model, the full tree model or the linear model.

	No model	Primary tree split	Tree model	Linear model
Patients with underestimation (n)	131	206	180	181
Total underestimated days	830	876	824	810
Patients with overestimation (n)	817	742	768	767
Total overestimated days	894	610	669	700
Reduction of underestimated days	-	-46	6	20
Reduction of overestimated days	-	284	225	194

RESULTS

Study population

Cumulative distribution of LOS is presented in Figure 1. Mean LOS was 2.3 (5.6) days, median LOS was 1.1 days. Of all 2540 patients, 780 (30.7%) patients were discharged from the ICU within 1 day, 1996 (78.6%) patients were discharged within 2 days, and 2236 (88.0%) were discharged within 3 days.

Regression tree

The derived regression tree is presented in Figure 2. The variables procedure type (valve surgery, coronary surgery or both), serum creatinin level (binary), critical preoperative state, previous cardiac surgery and age (binary) are included in the final tree. The R2 for the entire tree model was 10%. Mean predicted LOS was 2.1 (0.8; median 1.6) days.

The primary split of the regression tree was procedure type. Mean predicted LOS of the primary split according to the primary split only was 1.9 (0.4; median 1.8) days. The R2 for the primary split was 8%.

Linear regression model

Results for the univariable and multivariable linear regression analyses are presented in Table 2.

Based upon their coefficients, critical preoperative state, combined valve and coronary surgery, active endocarditis, previous cardiac surgery and pulmonary hypertension were the variables that contributed most to the increase in LOS for the linear regression model.

All variables that were used by the tree also proved to be significant in the linear regression analysis.

The R^2 for the linear regression model was 13%. Mean predicted LOS was 2.1 (1.1) days, median predicted LOS was 1.8.

Model application

Table 3 presents the assessment results (under- and overestimation of LOS) of all models in the validation cohort of 948 patients. Observed mean LOS in this cohort was 2.2 (5.6) days, median LOS was 1.1 days. For example, when using no model (prediction based upon mean LOS of the development cohort), 131 out of 948 patients had an observed LOS longer than predicted (underestimation of 830 ICU days), and the remaining 817 had an observed LOS shorter than predicted (overestimation of 894 ICU days).

As can be deducted from Table 3, the total underestimation of ICU days decreased by 6 when using the tree model and by 20 when using the linear model. The total underestimation of ICU days increased by 46 when using the primary split only. The total overestimation of ICU days decreased by 284 when using the primary split only, by 225 when using the tree model and by 194 when using the linear model.

DISCUSSION

This study aimed to explore the possibility of preoperatively using (patient data driven) prediction models to optimize ICU capacity. The developed predictive models both identified procedure type, critical preoperative state, serum creatinin level and age as the most important predictors of increased LOS in the study cohort. All models enhanced the accuracy of estimated LOS after cardiac surgery compared to the no model approach; a decrease in underestimation and overestimation of LOS was achieved by both the tree model and the linear regression model over an 18-month period.

Only a few other prediction models to determine risks for prolonged LOS after cardiac surgery have previously been developed or tested [7-10]. In these studies the investigators developed or tested a logistic regression model, with risk for prolonged LOS as the main outcome. This outcome measure is claimed to improve planning and therefore cost-effectiveness of hospitals. However, the

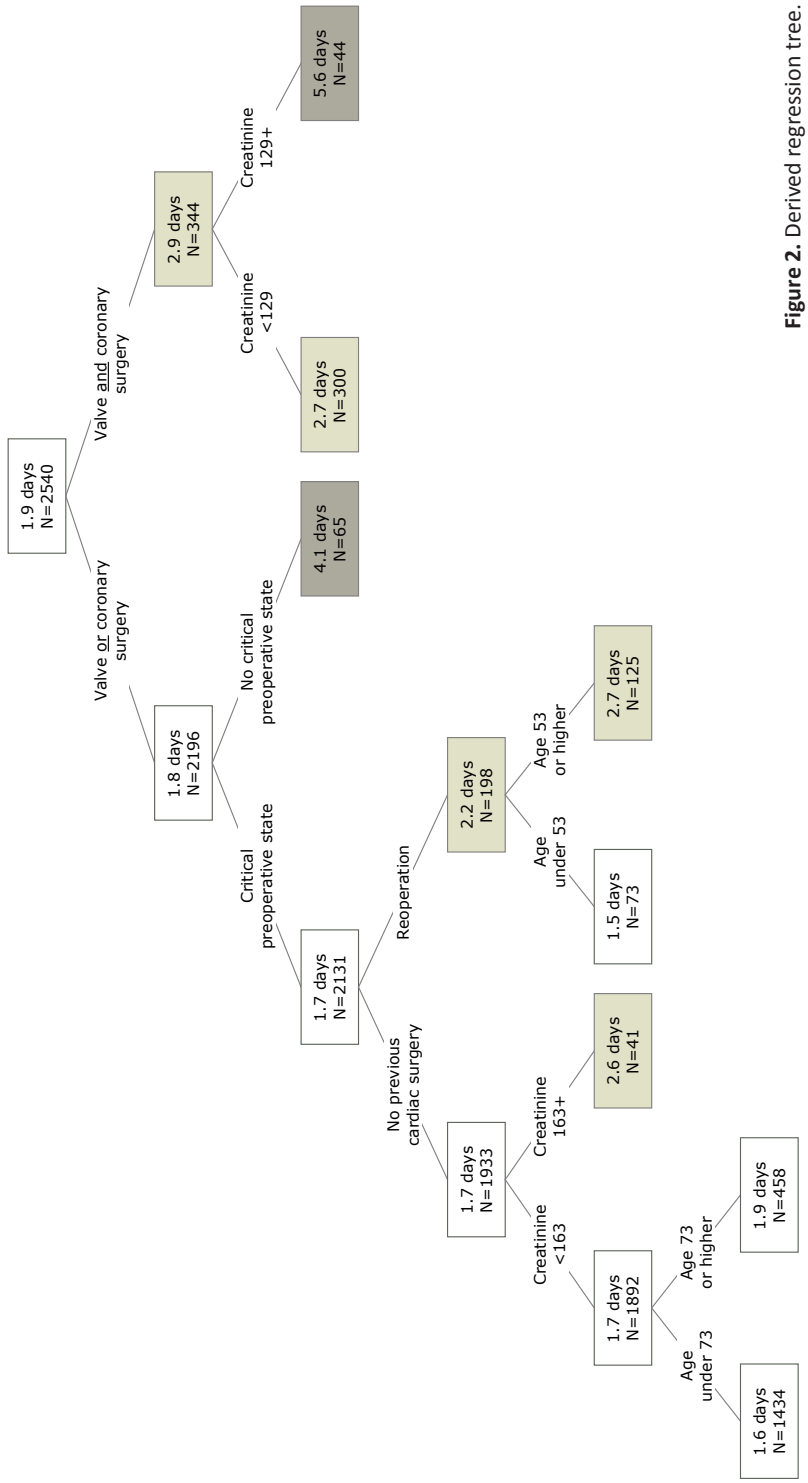


Figure 2. Derived regression tree.

results from these studies are less suitable for scheduling of individual patients in the ICU as they only provide the risk for prolonged LOS given a certain cut-off point. In contrast, the modeling approach proposed here permits individual patient scheduling in the ICU in daily clinical practice by providing an estimated number of days a patient will reside in the ICU.

Regression tree or linear regression?

The focus of this study lies on the construction and performance of simple models that predict LOS and provide insight in important factors influencing LOS. Because regression trees are easy to construct, interpret and implement, using them seems an ideal approach. As long as performance is adequate, they can serve as an alternative for linear regression models for planning purposes and the exploration of challenges in the planning process.

In the present study, procedure type was identified as the most important factor (primary split) increasing LOS, followed by critical preoperative state and serum creatinin level.

Although the linear regression model (Table 2) provides insight in factors influencing LOS as well, they are presented in a less intuitive way.

For example, procedure type is clearly presented as the primary (hence most important factor) split in the regression tree. According to the linear regression model, critical preoperative state has the highest coefficient (strong predictor), but has a low incidence (3.2%) and therefore a limited influence in the entire population. A more difficult procedure type however affects more patients (13.5%), therefore taking a more prominent place in the tree. In other words, factor weights and incidences are presented in a combined fashion in a tree model while for linear regression models, incidences need to be deducted from the confidence intervals of the model coefficients.

Preoperative approach

The R^2 values reveal a considerable variation for all approaches and confirm that preoperative LOS predictions are very difficult and can only partly be made preoperatively. Although it is possible that not all preoperative variables that influence LOS were collected, further predictive improvements are most likely to arise from the inclusion of peri- and post-operative variables. Examples are ICU nursing staff availability fluctuations during the day or the occurrence of adverse events that are not directly related to a patient's preoperative status.

A similar study as the present study involving oesophagectomy patients demonstrated that a model using intra-ICU variables as well performed better than a model based solely upon preoperative variables [11].

Implementation

Since this study shows that the use of simple predictive models can substantially increase ICU capacity, the next logical step seems to be the development of a simple planning tool that incorporates such a model, and thus allows for optimal planning of elective adult cardiac surgery patients.

A recent study successfully modeled logistics planning/optimization in our cardiac surgery department up to a certain level [1]. No patient specific information however was entered into that model apart from procedure type. By combining that model with the approach used in this present study (along with peri- and postoperative model variables), resource optimization might be improved beyond LOS in the ICU alone.

Usually, the patient flow in a cardiac surgery department consists of scheduled patients (elective patients) and non-scheduled patients (emergency patients requiring immediate surgery) [1].

Preoperative prediction models like those in the present study can only be used to predict ICU bed usage by elective patients. Determining a reservation policy for emergency patients is therefore an essential part of the implementation of these preoperative models. Although emergency patients usually only account for a limited percentage of all cardiac surgery patients, they can influence the entire planning process.

In clinical practice, patients in a 'critical preoperative state' often require surgery on short notice. Even if they are not labeled as emergency patients, they cannot always be scheduled in advance. This might also influence model implementation.

Cardiac surgery planning involves different resources such as operating theatre time, ward beds, ICU beds and nursing staff. In the daily practice the planning usually focuses on optimal use of operating theatre time, though the performance of a department as a whole is often more limited by these other resources. Apart from individually optimizing usage of these resources as well, one should furthermore take into account dependencies and interactions between these resources. For instance, patients are usually not transferred from the ICU to the ward during the night, as the nursing staff is not fully available at that mo-

ment. As a result, patients might reside in the ICU for some extra hours.

Final implementation depends upon the hospital, its performance and its resource capabilities. All models constructed in the present study consist of variables that are well known to predict adverse events and prolonged length of stay in the ICU in other populations as well [9, 12, 13]. However, general applicability of our models is limited because of the single center setup of this study. For instance, the percentage of valve procedures, the primary split of the tree in this study, varies greatly from center to center. We therefore encourage other centers to construct their own trees for planning purposes. Instructions, advantages and disadvantages of tree modeling are discussed further in the StatPage [ref Statpage Wu].

Limitations and other considerations

Although calculating the over- and underestimated ICU days does not simulate reality accurately, this simplified application gives an indication of performance of the proposed planning strategies. Ideally, one should use a more accurate model of reality, including day/night shift patterns, weekends, and other resource dependencies as described above.

Conclusion

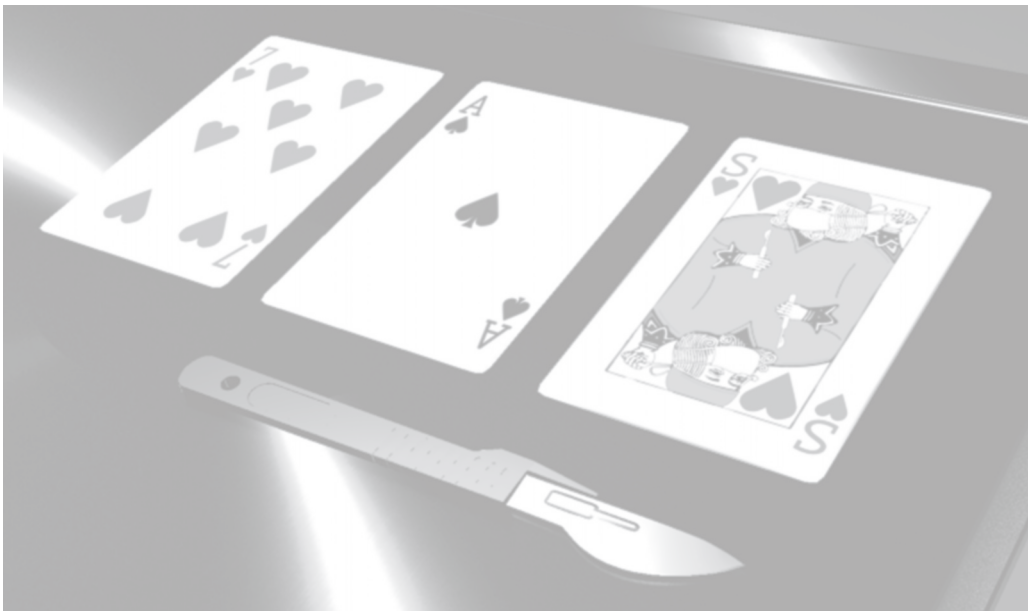
Predictions models using individual patient data can be used to further optimize ICU capacity by predicting LOS after cardiac surgery. Although the predictive ability of these preoperative models is limited, this study demonstrates that by using them, a decrease in underestimation and overestimation of LOS can be achieved. When taking the simplicity of the regression tree into account, it performed well enough to not only use it for exploration challenges in optimization of procedure planning but to justify implementation in the planning process itself as well. For purposes that require more accurate predictions, such as surgeon and center benchmarking, conventional regression analyses will however provide more accurate predictions.

Optimal implementation requires full integration of these models along with other resource related models and models containing peri- and post- operative information. Until then, this study demonstrates how to easily construct a model that will provide planners with a tool to predict the LOS in the ICU after cardiac surgery, thereby (contributing to) optimizing ICU capacity.

REFERENCES

1. Adan I, Bekkers J, Dellaert N, Vissers J and Yu X. Patient mix optimisation and stochastic resource requirements: a case study in cardiothoracic surgery planning. Health Care Manag Sci. Jun 2009;12(2):129-41.
2. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S and Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. Jul 1999;16(1):9-13.
3. Strum DP, May JH and Vargas LG. Modeling the uncertainty of surgical procedure times: comparison of log-normal and normal models. Anesthesiology. Apr 2000;92(4):1160-7.
4. Duan N. Smearing estimate: a nonparametric retransformation method. J Am Stat Assoc. 1986(383):605-610.
5. Breiman L. Classification and regression trees. Wadsworth, Belmont, Calif., 1984.
6. R: A Language and Environment for Statistical Computing. Available at: http://cran.r-project.org/doc/manuals/fullrefman.pdf. Accessed April 23th, 2009.
7. Janssen DP, Noyez L, Wouters C and Brouwer RM. Preoperative prediction of prolonged stay in the intensive care unit for coronary bypass surgery. Eur J Cardiothorac Surg. Feb 2004;25(2):203-7.
8. Kurki TS, Jarvinen O, Kataja MJ, Laurikka J and Tarkka M. Performance of three preoperative risk indices; CABDEAL, EuroSCORE and Cleveland models in a prospective coronary bypass database. Eur J Cardiothorac Surg. Mar 2002;21(3):406-10.
9. Nilsson J, Algotsson L, Hoglund P, Luhrs C and Brandt J. EuroSCORE predicts intensive care unit stay and costs of open heart surgery. Ann Thorac Surg. Nov 2004;78(5):1528-34.
10. Toumpoulis IK, Anagnostopoulos CE, Swistel DG and DeRose JJ, Jr. Does EuroSCORE predict length of stay and specific postoperative complications after cardiac surgery? Eur J Cardiothorac Surg. Jan 2005;27(1):128-33.
11. Van Houdenhoven M, Nguyen DT, Eijkemans MJ, Steyerberg EW, Tilanus HW, Gommers D, et al. Optimizing intensive care capacity using individual length-of-stay prediction models. Crit Care. 2007;11(2):R42.
12. Xu J, Ge Y, Pan S, Liu F and Shi Y. A preoperative and intraoperative predictive model of prolonged intensive care unit stay for valvular surgery. J Heart Valve Dis. Mar 2006;15(2):219-24.
13. Ghotkar SV, Grayson AD, Fabri BM, Dihmis WC and Pullan DM. Preoperative calculation of risk for prolonged intensive care unit stay following coronary artery bypass grafting. J Cardiothorac Surg. 2006;1:14.

CHAPTER 14



Effects of the introduction of an e-learning supported multimodality protocol to reduce blood loss

Van Gameren M
Kappetein AP
Hofland J
Van Thiel RJ
Boks R
Takkenberg JJM
Bogers AJJC

Submitted

ABSTRACT

Background

Excessive blood loss after cardiac surgery is a health burden for the patient and a logistic nuisance and costly complication for the institution. The goal of this study is to measure the effectiveness of a newly introduced, guideline based, multimodality protocol to reduce blood loss after adult cardiac surgery in comparison to previous separate interventions.

Methods

Blood loss (within 12 hours after surgery) and related adverse outcomes (blood product administration, need for re-exploration) during a period after the introduction of the new protocol (December 2008– June 2009) were compared with data from a previous period (January 2008–July 2008). Protocol implementation included deployment of e-learning modules and a clinical decision support system. In addition, the effect of the level of adherence to the new protocol was investigated.

Results

Data of 948 patients were analyzed. The introduction of the new protocol was associated with less blood loss (668 versus 759mL;p=0.038) and less red blood cell (RBC) unit (2.3 versus 2.8;p=0.01) administration. Less patients required RBC transfusion (66.1% versus 76.8%;p<0.001).
More protocol violations (≥2) were associated with more blood loss (893 versus 605mL;p<0.001), more RBC and platelet unit administration (2.8 versus 2.1 and 0.9 versus 0.7 respectively;p=0.045 and p=0.021), and more re-explorations (19.1% versus 7.6%;p<0.001).

Conclusions

Blood loss after adult cardiac surgery can be decreased by the introduction of a guideline based, multimodality protocol focused at reducing blood loss. In addition, the degree of protocol adherence is associated with the amount of blood loss. To successfully improve the quality of care and reduce the costs of care and resource utilization, an important part of the implementation of a protocol consists of ensuring an adequate level of adherence.

INTRODUCTION

Excessive blood loss after cardiac surgery is a health burden for the individual patient and a logistic nuisance and costly complication for the institution, in particular when a surgical re-exploration is required. Approximately 3% of patients undergoing isolated coronary surgery require re-exploration due to excessive blood loss [1, 2]. Excessive blood loss during and after cardiac surgery is an established marker of increased hospital mortality and morbidity, and longer Intensive Care Unit (ICU) and hospital length of stay [2-6].

Causes of excessive blood loss in adults undergoing cardiac surgery are multifactorial, and can be categorized into patient-related variables, procedure-related variables and process-related variables [7]. Current guidelines recommend a systematic approach to reduce postoperative blood loss that focuses on adequate preoperative assessment of bleeding risk, optimal use of blood conservation techniques (drugs, devices, interventions), consensus of participating disciplines with common goals for limiting postoperative blood loss, and a multimodality approach to blood conservation combining all of the above in a continuous quality assurance approach [7, 8].

Several medical interventions that prove to be effective in reducing postoperative blood loss can be employed in patients undergoing cardiac surgery [7, 9-13]. Recently, in our center these interventions were incorporated in a multimodality protocol, involving a multidisciplinary team consisting of clinicians involved in the care for patients who undergo cardiac surgery, and experts in the field of clinical epidemiology, clinical decision making, cost-effectiveness and medical informatics. To facilitate implementation of this protocol, e-learning modules and a clinical decision support system (CDSS) were made available.

The goal of this study is to measure the effectiveness of a new integrated protocol in comparison to previous separate interventions to reduce blood loss in adult cardiac surgery. This was done by (1) comparing blood loss amounts since the introduction of a new multimodality protocol with previous results and (2) investigating whether the level of adherence to the new protocol influences blood loss.

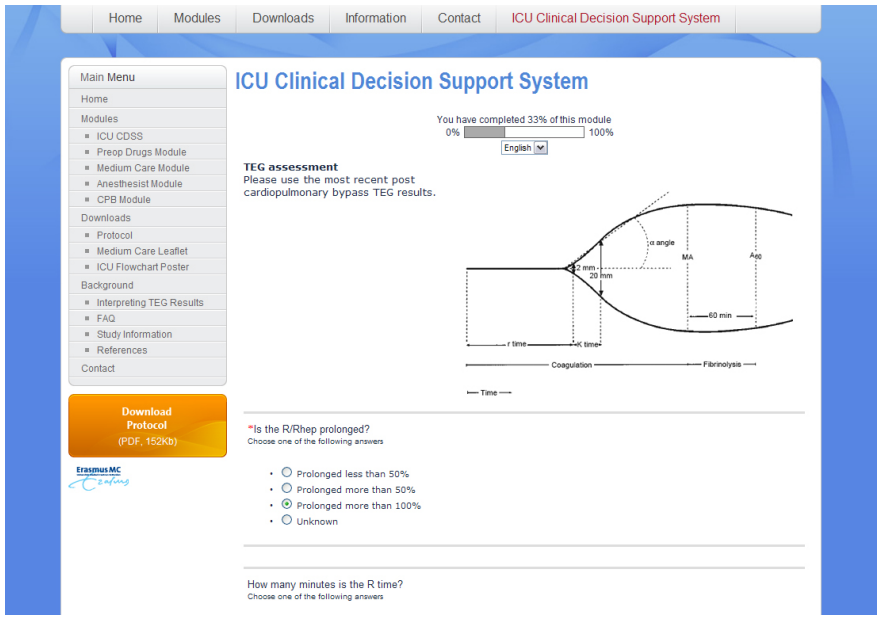


Figure 1. In this clinical decision support system screenshot, thromboelastography (TEG) results can be entered for a given patient. An explanation of TEG results is avalaible on request.

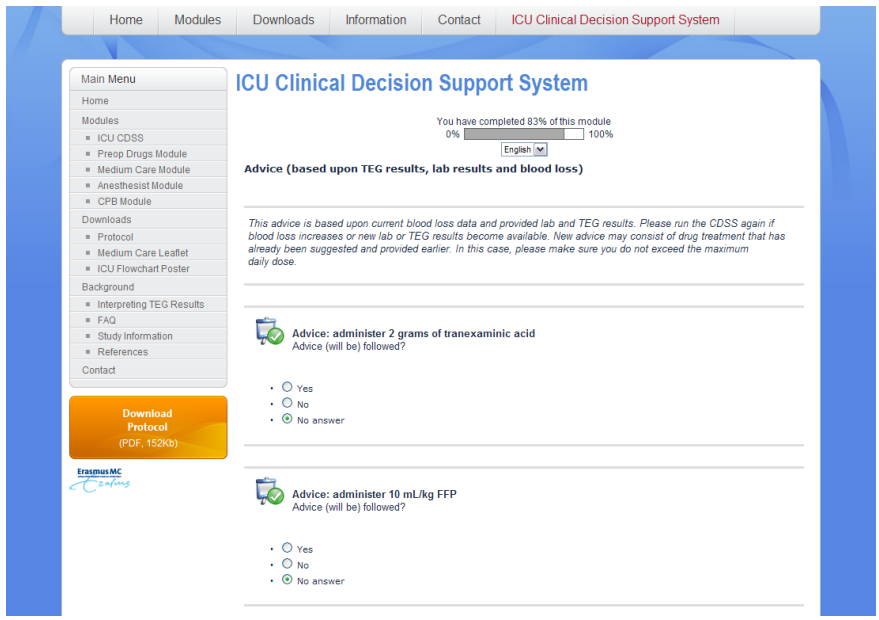


Figure 2. Based upon provided data, the clinical decision support system now provides an advice for this patient, thereby guiding the physician through the transfusion algorithm. Meanwhile, data is collected for study analyses.

Table 1. Incidence of violation types.

Violation	Incidence (N=534)
Clopidogrel not discontinued in time before surgery when indicated	16 (3.0 %)
Aspirin not discontinued in time before surgery when indicated	41 (7.7 %)
Coumarin derivatives not discontinued in time before surgery	0 (0 %)
Not enough tranexamic acid used during surgery	65 (12.2 %)
Baseline TEG not made	33 (6.2 %)
Autologous blood was not taken when indicated	6 (1.1 %)
Ht dropped below 0.20 during hyperdynamic hemodilution	0 (0 %)
Follow up TEG was not made after procedure	73 (13.7 %)
Patient was not warmed up properly at the end of the procedure	159 (29.8 %)
Hemofiltration was not performed when indicated	0 (0 %)
Perfusion stopped at poor Hb or Ht levels	0 (0 %)
Lab results were not available in the ICU	0 (0 %)
Protamine was not provided (or inadequate dose) in ICU when indicated	32 (6.0 %)
Tranexamic acid was not provided (or inadequate dose) in ICU when indicated	25 (4.7 %)
Any of the above violations	304 (56.9 %)

TEG = Thromboelastography, Ht = hematocrit, Hb = hemoglobin, ICU = Intensive Care Unit.

METHODS

Protocol

The introduced protocol was based upon recent literature and guidelines [7, 14-24] and contains only those measures that were supported by the multidisciplinary team. The protocol can be found in Appendix I.

The new protocol was presented to all medical professionals –nurses, perfusionists, anesthesists, intensivists, residents and surgeons– in our center responsible for cardiac surgery patients. In addition, referring centers were notified of changes in preoperative workup.

E-learning modules and CDSS

E-learning modules were deployed to help in correctly following the new protocol guidelines. To further facilitate implementation of the new protocol, a rule based CDSS was developed and made available to guide clinicians through the

Table 2. Key characteristics of the control cohort and the study cohort.

	Control cohort (N=414)		Study cohort (N=534)		p
Age	65	(64-66)	64	(63-65)	0.315
Female sex	31.2%	(129)	28.3%	(151)	0.200
Serum creatinin level	96	(90-101)	96	(91-101)	0.252
Type of surgery*					
Coronary +/- concomitant	51.4%	(213)	48.5%	(259)	0.368
Valve +/- concomitant	32.4%	(134)	32.4%	(173)	0.992
Valve + coronary +/- conc	11.4%	(47)	13.7%	(73)	0.287
Other**	4.8%	(20)	5.4%	(29)	0.679
Logistic EuroSCORE	6.8	(6.0-7.7)	6.8	(6.0-7.6)	0.674
Diabetes	19.8%	(82)	21.8%	(109)	0.461
Previous CVA	3.6%	(15)	3.6%	(18)	0.985
COPD	10.4%	(43)	13.6%	(68)	0.139
Previous cardiac surgery	9.9%	(41)	6.7%	(32)	0.077
Critical preoperative state	6.5%	(27)	6.8%	(34)	0.867
Emergency procedure	5.1%	(21)	5.0%	(25)	0.400

N or 95% CI in brackets

Mann-Whitney or Chi-square tested

CVA = CardioVascular Accident, COPD = Chronic Obstructive Pulmonary Disease

* Concomitant: Maze surgery, Atrial Septal Defect (ASD) closure, Ventricular Septal Defect (VSD) closure, aortic surgery

** Myxomectomy, isolated Maze surgery, dissection, ASD closure, VSD closure

postoperative transfusion algorithm. Both the e-learning modules and the CDSS are web-based and available on each computer attached to the hospital intranet.

Figures 1 and 2 are screenshots of the CDSS. Protocol violations were simultaneously registered by the support modules. At the same time, the clinician could explain why the protocol violation was deemed necessary.

Study cohorts and data collection

Data were collected on (1) patient, procedure and process characteristics, including the multidisciplinary protocol components throughout the entire chain, starting at the preoperative outpatient clinic until discharge from the hospital, and (2) patient blood loss, need for reintervention due to blood loss, mortal-

ity, use of blood products. Data obtained from two patient cohorts were compared.

The first patient cohort contained retrospective data of all consecutive adult patients who underwent cardiac surgery between January 2008 and July 2008. This was the control group.

The second patient cohort consisted of all consecutive adult patients who underwent cardiac surgery between December 2008 and June 2009. All patients in the second cohort were treated according to the multimodality protocol as described in Appendix I. This protocol was introduced in September 2008.

Patients who underwent cardiac transplantation, ventricular assist device implantation or percutaneous valve implantation were excluded from this study.

Study approval was obtained from our institutional review board. The need for informed consent was waived.

Endpoints

The primary endpoint of this study is postoperative blood loss (drain production in mL) during the first 12 postoperative hours. Secondary endpoints are: use of platelet, plasma and RBC transfusion within 24 hours prior to surgery and 48 hours after surgery, need for reintervention due to blood loss (including cardiac tamponade), and mortality at thirty days.

A power calculation was performed indicating that 143 patients in each group would be required to detect a decrease in blood loss of at least 50 mL (alpha 0.05; power 0.80).

The number of and reasons for protocol violations in the study group were studied to determine whether adherence to the protocol influences study endpoints.

Protocol adherence

Of the possible protocol violations listed in Appendix I, the violations presented in Table 1 were considered important enough for additional evaluation. The sum of these protocol violations was calculated for each patient.

Statistical methods

Continuous data are presented as mean (95% confidence interval; median) and were compared using the unpaired T-test. When data were not normally distrib-

Table 3. Control versus control cohort results.

	Control cohort (N=414)	Study cohort (N=534)	p
Blood loss (mL; 12 hours)	759 (699-818; 525)	668 (626-708; 525)	0.038
RBC units (275 mL)	2.8 (2.5-3.1)	2.3 (2.0-2.5)	0.010
Plasma units (325 mL)	3.4 (3.0-3.8)	3.4 (3.2-3.7)	0.117
Platelet units (250 mL)	0.7 (0.6-0.8)	0.7 (0.6-0.8)	0.091
Re-explorations for blood loss	8.9% (37)	10.1% (54)	0.542
30 day mortality	2.2% (9)	1.3% (7)	0.306

N or 95% CI + median in brackets
Mann-Whitney or Chi-square tested
RBC = Red Blood Cell

Table 4. Key characteristics of patients with 0-1 protocol violations versus patients with ≥2 protocol violations .

	0-1 violations (N=419)		≥2 violations (N=115)		p
Age	64	(63-66)	64	(63-65)	0.217
Female sex	31.6%	(125)	24.8%	(26)	0.172
Serum creatinin level	94	(89-99)	103	(89-118)	0.416
Type of surgery*					
Coronary +/- concomitant	48.9%	(205)	47.0%	(54)	0.708
Valve +/- concomitant	32.9%	(138)	30.4%	(35)	0.612
Valve + coronary +/- conc	12.9%	(54)	16.5%	(19)	0.315
Other**	5.3%	(22)	6.1%	(7)	0.726
Logistic EuroSCORE	6.9	(6.1-7.8)	6.4	(4.4-8.4)	0.275
Diabetes	21.3%	(84)	23.8%	(25)	0.575
Previous CVA	3.8%	(15)	2.9%	(3)	0.646
COPD	14.4%	(57)	10.5%	(11)	0.293
Previous cardiac surgery	7.6%	(30)	1.9%	(2)	0.034
Critical preoperative state	6.8%	(27)	6.7%	(7)	0.951
Emergency procedure	5.1%	(20)	4.8%	(5)	0.900

N or 95% CI in brackets
Mann-Whitney or Chi-square tested
CVA = CardioVascular Accident, COPD = Chronic Obstructive Pulmonary Disease
* Concomitant: Maze surgery, Atrial Septal Defect (ASD) closure, Ventricular Septal Defect (VSD) closure, aortic surgery
** Myxomectomy, isolated Maze surgery, dissection, ASD closure, VSD closure

Table 5. Results for patients with less than two violations versus patients with ≥2 violations.

	0-1 violations (N=419)	≥2 violations (N=115)	p
Blood loss (mL; 12 hours)	605 (566-645; 475)	893 (773-1014; 675)	< 0.0001
RBC units (275 mL)	2.1 (1.9-2.3)	2.8 (2.3-2.4)	0.045
Plasma units (325 mL)	3.3 (3.0-3.6)	4.0 (3.2-3.7)	0.141
Platelet units (250 mL)	0.7 (0.6-0.7)	0.9 (0.7-1.1)	0.021
Re-explorations for blood loss	7.6% (32)	19.1% (22)	< 0.0001
30 day mortality	1.2% (5)	1.7% (2)	0.649

N or 95% CI + median in brackets
Mann-Whitney or Chi-square tested
RBC = Red Blood Cell

uted (Kolmogorov-Smirnov test) the Mann-Whitney U-test was used. Categorical data are presented as frequencies and were compared using the Chi-Square test or the Fisher Exact test where appropriate. All tests were 2-sided, with an alpha level of 0.05.

Spearman's correlation tests were used to determine if there was a significant correlation between blood loss and blood product transfusion.

A classification and regression tree (CART) analysis was performed to determine at which number of protocol violations a split results in the largest difference in postoperative blood loss [25]. This method is based on recursive partitioning analysis and involves the segregation of different values of classification variables through a decision tree composed of progressive binary splits.

Linear regression was used to identify risk factors for blood loss. Potential risk factors were entered into multivariate analysis if $p < 0.10$. For multivariate analysis backward conditional stepping and 5-9 events per variable, as proposed by Vittinghoff et al., were used [26]. A p-value < 0.05 was considered statistically significant.

RESULTS

Protocol effect: study cohort versus control cohort

Data of a total number of 948 patients were analyzed. Patient and procedure characteristics are presented in Table 2. No significant demographic differences

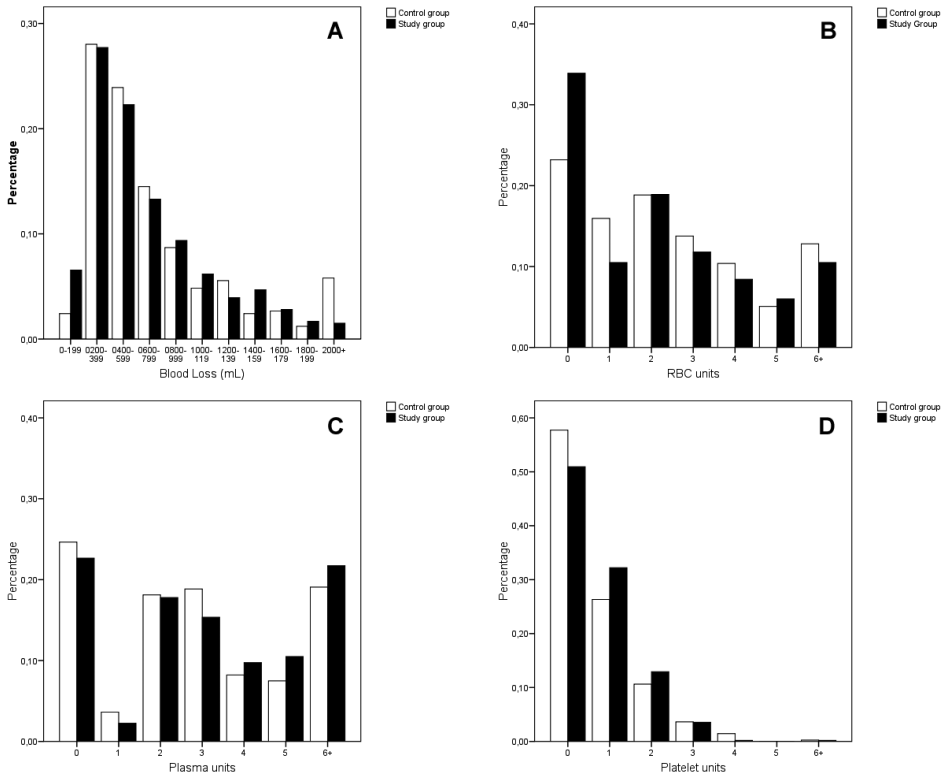


Figure 3. Distribution of blood loss and used blood products before and after protocol introduction.

were found between the study group and the control group.

Results regarding blood loss, blood products used, re-exploration for excessive blood loss and mortality are presented in Table 3. A significant decrease in blood loss and RBC unit administration after introduction of the new protocol was found. In addition, significantly fewer patients required RBC transfusion; 66.1% (353) versus 76.8% (318) ($p<0.001$). No difference was found between plasma and platelet transfusion, mortality and re-explorations for blood loss. A positive correlation was found between blood loss and RBC units, plasma units, platelet units, and re-explorations (all $p<0.001$). Distributions of blood loss and used blood products are shown in Figure 3.

Influence of protocol violations

Table 1 presents how often the protocol was violated. No protocol violations

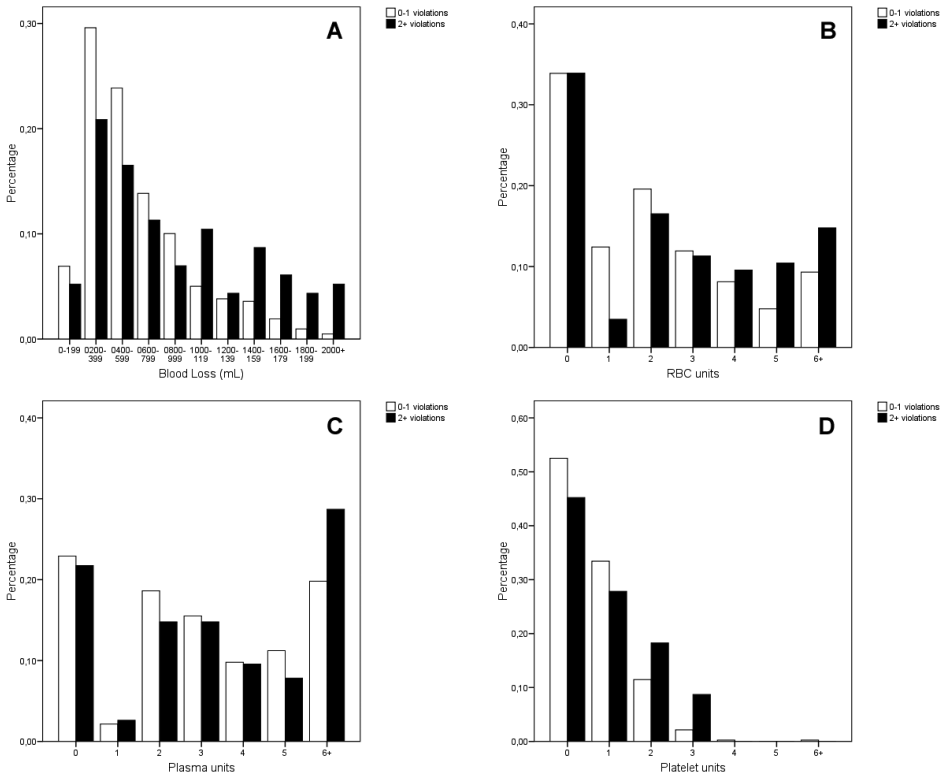


Figure 4. Distribution of blood loss and used blood products in the study group for 0-1 versus 2+ violations.

were found in 230 patients (43.1 %). One violation took place in 189 patients (35.4 %), two in 88 patients (16.5 %), three in 24 patients (4.5 %), four in 2 patients (0.4 %), and five in 1 patient (0.2 %).

The CART analysis revealed that the greatest difference in postoperative blood loss was between patients who had no or one protocol violations and patients who had two or more protocol violations.

Table 4 provides the characteristics for these two patient groups. Apart from the fact that the incidence of previous cardiac surgery was lower in the ≥2 violations group there were no differences between the two groups.

Results regarding blood loss and the other study endpoints for these two groups are presented in Table 5. In the group of patients who had less than two protocol violations there was a significantly lower amount of blood loss and RBC and platelet unit transfusion compared to the group of patients who had two

or more protocol violations. In addition, more re-explorations were performed in the group of patients with two or more protocol violations. Distributions of blood loss and used blood products are shown in Figure 4.

Reasons for protocol violations

For most protocol violations, no reason was given. However, reasons for deviating from the protocol by not discontinuing clopidogrel or aspirin preoperatively could successfully be collected (Figure 5). Among the reasons for not discontinuing these drugs were rescheduled procedures not related to urgency and poor instructions by referring hospitals.

Protocol violations (≥2) as an independent predictor for blood loss

Univariate linear regression analysis for '≥2 protocol violations' as a predictor for blood loss reveals a coefficient of 288 (95% CI 191-385; $p < 0.001$). A multivariate linear regression analysis including all variables from Table 1 and all individual EuroSCORE variables –significant in univariate linear regression analyses– revealed '≥2 protocol violations' as an independent predictor for blood loss. Detailed analysis results are presented in Table 6. Given the skewed distribution of blood loss, the regression analyses were repeated with natural logarithmic transformed values to confirm the results from Table 6. The included model variables were identical and coefficients were comparable after natural logarithmic 'back transformation'.

DISCUSSION

This study demonstrates that blood loss after adult cardiac surgery in our center can be reduced by the introduction of a multimodality protocol. After protocol introduction, less RBC units were administrated and fewer patients received RBC units.

Furthermore, the level of adherence to this new, multimodality protocol influenced blood loss as well. Patients for whom the protocol was violated twice or more had more blood loss, received more RBC and platelet units, and had to undergo a re-exploration for blood loss more frequently.

Reasons for protocol violations

Although the establishment of a protocol is an essential first step, results maybe disappointing if adherence is lacking. Research has shown that clinical protocols are often violated, thereby limiting their full potential [27, 28]. Known reasons for decreased guideline adherence are: out-of-date guidelines, lack of awareness, agreement, or self-efficacy, lack of outcome expectancy, the inertia of previous practice, and external barriers [29]. These reasons undoubtedly also apply to guideline-based protocol adherence.

Apart from preoperative protocol violations regarding discontinuing anticoagulant drugs, the reasons for most protocol violations were not documented well. As can be deducted from Figure 5, a number of protocol violations are preventable. When rescheduling patients for example, patients who preoperatively use aspirin or clopidogrel should be given enough time to discontinue.

Individual protocol violation analysis

In general, the inclusion of protocol items should mainly depend upon guideline recommendations with a high level of evidence. Once protocol items are selected, knowing the effect of each individual protocol violation on blood loss could substantially contribute to future protocol development and allow for a more fair comparison of protocol violations as well. For instance, failing to cease coumarine derivatives in time is likely to be a more severe protocol violation than not providing a preoperative thromboelastography (TEG) measurement. This study was however not designed and insufficiently powered to calculate significance or weights for each individual protocol violation. Under the assumption that each protocol violation is important, the combination of protocol violations (two or more violations) however was tested and proved to be an independent predictor for increased blood loss.

Table 6. Independent predictors of postoperative blood loss.

Model variable	Coefficient	95% CI	p
≥2 protocol violations	271	173-368	< 0.001
Urgent procedure	203	21-385	< 0.001
Male sex	231	144-317	0.029
Model constant	664	611-716	-

CI = Confidence Interval

Protocol quality

Besides protocol adherence, results obtained from this study depend upon the quality of the tested protocol. The protocol that was investigated in this study is largely based upon published guideline recommendations. Some recommendations were however modified or not incorporated in the protocol because no consensus could be formed in the multidisciplinary team or because of technical limitations. For this reason the effect of the used protocol might be limited.
Periodic protocol updates based upon new evidence are warranted to allow for inclusion of new insights in blood reduction and management.

Protocol implementation

We chose not to conduct a randomized trial at the time of introduction, since the introduction of a multimodality approach in clinical practice usually has an important effect (Hawthorne effect) on standard care: the exposure of health care workers to the multidisciplinary approach will affect their standard behavior [30]. An organizational change is hence best investigated in a before-after design.

Further measures to monitor protocol adherence and effectiveness will be scheduled, assuming that periodic feedback to all involved personnel reinforces protocol adherence. Ideally, a full continuous quality improvement (CQI) program should be implemented.

Surgical aspects

Of all precautions and instructions incorporated in the introduced protocol, surgical aspects are almost absent. When creating the protocol, the surgeons were considered to limit blood loss during surgery in all possible ways. It is self explanatory that no consensus on each possible surgical technique can be formed and included in the implemented blood reduction protocol. To a certain extent, the surgeon is regarded as a 'black box' during surgery. As the surgical team remained unaltered during the entire study time frame, the influence of the surgeons on the study results is limited.

No reason was found to explain the fact that the incidence of previous cardiac surgery was lower in the ≥2 violations group. It could be speculated that more caution was taken with patients who underwent a reoperation.

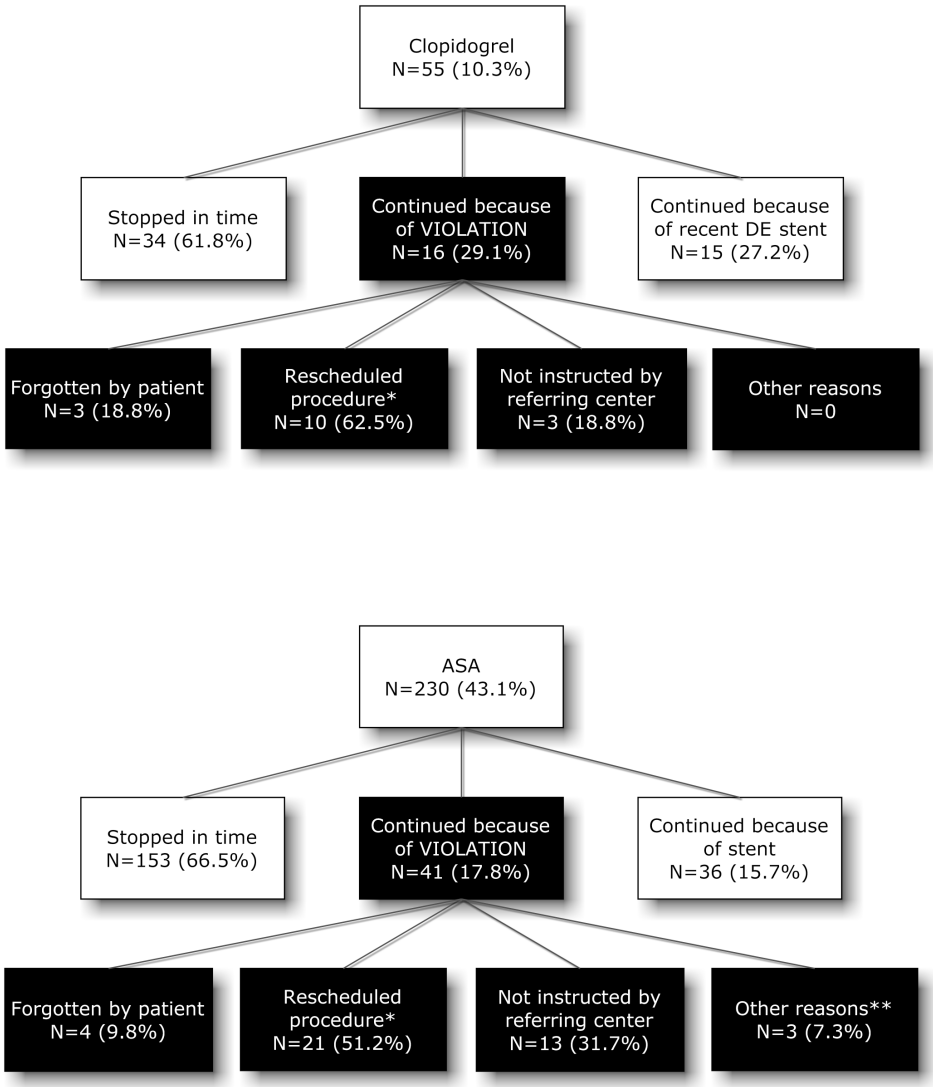


Figure 5. Reasons for protocol violations regarding clopidogrel and aspirin .

DE = Drug Eluting

* Contains no rescheduled procedures based on urgency

** Main stem stenosis (N=1); severe three vessel disease (N=1); recent transient ischemic at tacks (N=1)

Costs

A simple cost reduction example can be calculated based upon decreased blood product usage since protocol introduction exclusively. According to our hospital management system, one RBC unit costs at least € 300.00. A decrease of 0.5 RBC unit per operated patient therefore on average results in an annual cost reduction of € 150,000.00 based upon 1,000 operated patients (300*0.5*1000).

As suggested by the lower incidence of re-explorations for excessive blood loss in the group of patients with less than two protocol violations, an additional reduction of costs could be achieved by increasing the level of protocol adherence. These costs and potential cost reduction were not calculated. One should however keep in mind that the costs of a re-explorative operation include not only the re-intervention and prolonged and more complicated hospital stay (estimated to be 1 extra ICU day and 1.5 days extra in the general ward [4]), but also include costs associated with the surgical treatment delay of other patients.

Timing of re-exploration for blood loss

The decision to re-operate a patient for excessive blood loss is influenced by many factors. The introduced protocol does not provide instructions for the moment of reoperation. Consensus during protocol development could not be formed as team members tend to accept different amounts of blood loss and base the timing of re-explorations on other circumstances, including logistics, as well.

Certain items regarding the timing of re-explorations could however be included in a blood loss reduction protocol. For instance, literature shows that when all TEG values are normal, a surgical focus is present in 90% of the patients [21].

In conclusion, blood loss after cardiac surgery can be decreased by the introduction of a multimodality protocol aimed at blood loss reduction. Protocols to reduce postoperative blood loss should ideally be implemented in every department of cardiac surgery to achieve less postoperative morbidity and mortality, shorter length of ICU and hospital stay, a decrease in the utilization of platelet, plasma and erythrocytes transfusion, and a concomitant reduction in societal costs. Ensuring a high level of protocol adherence appears to be a vital aspect of implementation. Ideally, a blood loss reduction protocol should be part of daily practice through implementation in the patient information system and the in-

stitution of a continuous quality assurance system for excessive blood loss.

REFERENCES

1. Mehta RH, Sheng S, O'Brien SM, Grover FL, Gammie JS, Ferguson TB and Peterson ED. Re-operation for bleeding in patients undergoing coronary artery bypass surgery: incidence, risk factors, time trends, and outcomes. Circ Cardiovasc Qual Outcomes. 2009;2(6):583-90.
2. Hall TS, Brevetti GR, Skoultchi AJ, Sines JC, Gregory P and Spotnitz AJ. Re-exploration for hemorrhage following open heart surgery differentiation on the causes of bleeding and the impact on patient outcomes. Ann Thorac Cardiovasc Surg. 2001;7(6):352-7.
3. Dacey LJ, Munoz JJ, Baribeau YR, Johnson ER, Lahey SJ, Leavitt BJ, Quinn RD, Nugent WC, Birkmeyer JD and O'Connor GT. Reexploration for hemorrhage following coronary artery bypass grafting: incidence and risk factors. Northern New England Cardiovascular Disease Study Group. Arch Surg. 1998;133(4):442-7.
4. Karthik S, Grayson AD, McCarron EE, Pullan DM and Desmond MJ. Reexploration for bleeding after coronary artery bypass surgery: risk factors, outcomes, and the effect of time delay. Ann Thorac Surg. 2004;78(2):527-34; discussion 534.
5. Murphy GJ, Reeves BC, Rogers CA, Rizvi SI, Culliford L and Angelini GD. Increased mortality, postoperative morbidity, and cost after red blood cell transfusion in patients having cardiac surgery. Circulation. 2007;116(22):2544-52.
6. Surgenor SD, DeFoe GR, Fillinger MP, Likosky DS, Groom RC, Clark C, Helm RE, Kramer RS, Leavitt BJ, Klemperer JD, Krumholz CF, Westbrook BM, Galatis DJ, Frumiento C, Ross CS, Olmstead EM and O'Connor GT. Intraoperative red blood cell transfusion during coronary artery bypass graft surgery increases the risk of postoperative low-output heart failure. Circulation. 2006;114(1 Suppl):I43-8.
7. Ferraris VA, Ferraris SP, Saha SP, Hessel EA, 2nd, Haan CK, Royston BD, Bridges CR, Higgins RS, Despotis G, Brown JR, Spiess BD, Shore-Lesserson L, Stafford-Smith M, Mazer CD, Bennett-Guerrero E, Hill SE and Body S. Perioperative blood transfusion and blood conservation in cardiac surgery: the Society of Thoracic Surgeons and The Society of Cardiovascular Anesthesiologists clinical practice guideline. Ann Thorac Surg. 2007;83(5 Suppl):S27-86.
8. Wolfe R, Bolsin S, Colson M and Stow P. Monitoring the rate of re-exploration for excessive bleeding after cardiac surgery in adults. Qual Saf Health Care. 2007;16(3):192-6.
9. Baric D, Biocina B, Unic D, Sutlic Z, Rudez I, Vrca VB, Brkic K and Ivkovic M. Topical use of antifibrinolytic agents reduces postoperative bleeding: a double-blind, prospective, randomized study. Eur J Cardiothorac Surg. 2007;31(3):366-71; discussion 371.
10. Boodhwani M, Williams K, Babaev A, Gill G, Saleem N and Rubens FD. Ultrafiltration reduces blood transfusions following cardiac surgery: A meta-analysis. Eur J Cardiothorac Surg. 2006;30(6):892-7.
11. Davies L, Brown TJ, Haynes S, Payne K, Elliott RA and McCollum C. Cost-effectiveness of cell salvage and alternative methods of minimising perioperative allogeneic blood transfusion: a systematic review and economic model. Health Technol Assess. 2006;10(44):iii-iv, ix-x, 1-210.
12. Dietrich W, Thuermel K, Heyde S, Busley R and Berger K. Autologous blood donation in cardiac surgery: reduction of allogeneic blood transfusion and cost-effectiveness. J Cardiothorac Vasc Anesth. 2005;19(5):589-96.

13. Spalding GJ, Hartrumpf M, Sierig T, Oesberg N, Kirschke CG and Albes JM. Cost reduction of perioperative coagulation management in cardiac surgery: value of “bedside” thrombelastography (ROTEM). Eur J Cardiothorac Surg. 2007;31(6):1052-7.
14. Dunning J, Versteegh M, Fabbri A, Pavie A, Kolh P, Lockowandt U and Nashef SA. Guideline on antiplatelet and anticoagulation management in cardiac surgery. Eur J Cardiothorac Surg. 2008;34(1):73-92.
15. Ferraris VA, Ferraris SP, Moliterno DJ, Camp P, Walenga JM, Messmore HL, Jeske WP, Edwards FH, Royston D, Shahian DM, Peterson E, Bridges CR and Despotis G. The Society of Thoracic Surgeons practice guideline series: aspirin and other antiplatelet agents during operative coronary revascularization (executive summary). Ann Thorac Surg. 2005;79(4):1454-61.
16. Klein HG, Spahn DR and Carson JL. Red blood cell transfusion in clinical practice. Lancet. 2007;370(9585):415-26.
17. Mannucci PM and Levi M. Prevention and treatment of major blood loss. N Engl J Med. 2007;356(22):2301-11.
18. Moulton MJ, Creswell LL, Mackey ME, Cox JL and Rosenbloom M. Reexploration for bleeding is a risk factor for adverse outcomes after cardiac operations. J Thorac Cardiovasc Surg. 1996;111(5):1037-46.
19. Ralley FE. Programmatic blood conservation in cardiac surgery. Semin Cardiothorac Vasc Anesth. 2007;11(4):242-6.
20. Shore-Lesserson L, Manspeizer HE, DePerio M, Francis S, Vela-Cantos F and Ergin MA. Thromboelastography-guided transfusion algorithm reduces transfusions in complex cardiac surgery. Anesth Analg. 1999;88(2):312-9.
21. Spiess BD. Blood transfusion for cardiopulmonary bypass: the need to answer a basic question. J Cardiothorac Vasc Anesth. 2002;16(5):535-8.
22. Spiess BD, Gillies BS, Chandler W and Verrier E. Changes in transfusion therapy and reexploration rate after institution of a blood management program in cardiac surgical patients. J Cardiothorac Vasc Anesth. 1995;9(2):168-73.
23. Stroncek DF and Rebulla P. Platelet transfusions. Lancet. 2007;370(9585):427-38.
24. Sun JC, Crowther MA, Warkentin TE, Lamy A and Teoh KH. Should aspirin be discontinued before coronary artery bypass surgery? Circulation. 2005;112(7):e85-90.
25. Breiman L. Classification and regression trees. Wadsworth, Belmont, Calif., 1984.
26. Vittinghoff E and McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165(6):710-8.
27. Hoeks SE, Scholte op Reimer WJ, Lenzen MJ, van Urk H, Jorning PJ, Boersma E, Simoons ML, Bax JJ and Poldermans D. Guidelines for cardiac management in noncardiac surgery are poorly implemented in clinical practice: results from a peripheral vascular survey in the Netherlands. Anesthesiology. 2007;107(4):537-44.
28. Khan IA, Mehta NJ, Gowda RM, Sacchi TJ and Vasavada BC. Reinforcement as a means for quality improvement in management of coronary syndromes: adherence to evidence-based medicine. Int J Cardiol. 2004;95(2-3):281-3.
29. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA and Rubin HR. Why don’t physicians follow clinical practice guidelines? A framework for improvement. Jama. 1999;282(15):1458-65.

30. Braunholtz DA, Edwards SJ and Lilford RJ. Are randomized clinical trials good for us (in the short term)? Evidence for a “trial effect”. J Clin Epidemiol. 2001;54(3):217-24.

APPENDIX 1

Preoperative section

The surgeon assesses hemoglobin (Hb), hematocrit (Ht) and platelet count when accepting the patient and treats underlying disease.

Prior to admission, the secretary or referring hospital instructs elective patients to discontinue:

- Clopidogrel 4 days prior to surgery;
  - unless a drug eluting coronary stent was implanted recently (within 6 months);
- Aspirin 3 days prior to surgery;
  - unless a coronary stent was implanted (bare metal and drug eluting);
- Coumarin derivatives 2 days prior to surgery.

During admission, the ward nurse checks if these drugs are discontinued properly. Heparin is started (20,000 IE/24h) for patients that normally use coumarin derivatives if:

- international normalized ratio (INR) < 3.0 and mitral valve replacement in history;
- INR < 2.0 and atrial fibrillation in history;
- INR < 2.5 for all other indications.

Perioperative section

The tranexamic acid dose used during the entire procedure is 50-100 mg/kg.
A baseline thromboelastogram (TEG) is made prior to initiating cardiopulmonary bypass.
Autologous blood is taken if the patient has a good left ventricular function and Hb is above 8.5 mmol/L.
Red cell saving devices are used during each procedure.
A follow up TEG is made after ending cardiopulmonary bypass.
If a hyperdynamic hemodilution approach is used, Ht levels are kept 0.20 L/L or higher.
The patient is re-warmed to at least 35.5 degrees Celcius (rectal or bladder temperature) at the end of the procedure.
Hemofiltration is performed if Ht is below 0.20 L/L and an adequate volume is

available.
The patient is weaned from cardiopulmonary bypass only if Hb is above 4.5 mmol/L and Ht is above 0.20 L/L.

Postoperative section
Ringer's lactate can be administered to a maximum of 2 L per 24 hours.
The patient is actively rewarmed further with an air warming blanket.

Transfusion algorithm

Blood loss > 200 mL/h without known TEG or lab results and no clot formation in the drains:
- 2,500 IE protamine is administered.

No or moderate blood loss (< 100 mL/h), but with abnormal lab results:
- if fibrinogen < 1.5 g/L => 5 mL/kg fresh frozen plasma (FFP);
- if platelet count < 50,000 => 1 unit platelet transfusion;
- if activated partial thromboplastin time (aPTT) > 50s, but < 65s => 2,500 IE protamine;
- if aPTT > 65s => 5,000 IE protamine.

Blood loss > 100 mL/h:
- if R/Rheparinase in TEG > 50% prolonged => 2,500 IE protamine;
- if R/Rheparinase in TEG > 100% prolonged => 5,000 IE protamine;
- if aPTT > 40s, but < 50s => 1,000-2,500 IE protamine;
- if aPTT > 50s => 5000 IE protamine;
- if prothrombin time (PT-INR) > 18s (50% prolonged) or R in TEG> 10 min => 5 mL/kg FFP;
- if PT-INR > 22s (80% prolonged) or R in TEG > 14 min => 10 mL/kg FFP;
- if mean amplitude (MA) in TEG < 45 or platelet count < 100,000 => 1 unit platelet transfusion;
- if MA in TEG < 45 and platelet count > 100,000 => 0.4 mcg/kg desmopressin acetate (DDAVP);
- if platelet count < 100.000 and recent aspirin or clopidogrel => 1 unit plate let transfusion followed by 0.4 mcg/kg DDAVP;

- if fibrinogen < 1.5 g/L => 10 mL/kg FFP;
- if Ht < 0.25 L/L => erythrocyte transfusion until Ht is 0.26-0.28 L/L;
- if LY30 in TEG >7.5% => 2 g tranexamic acid.

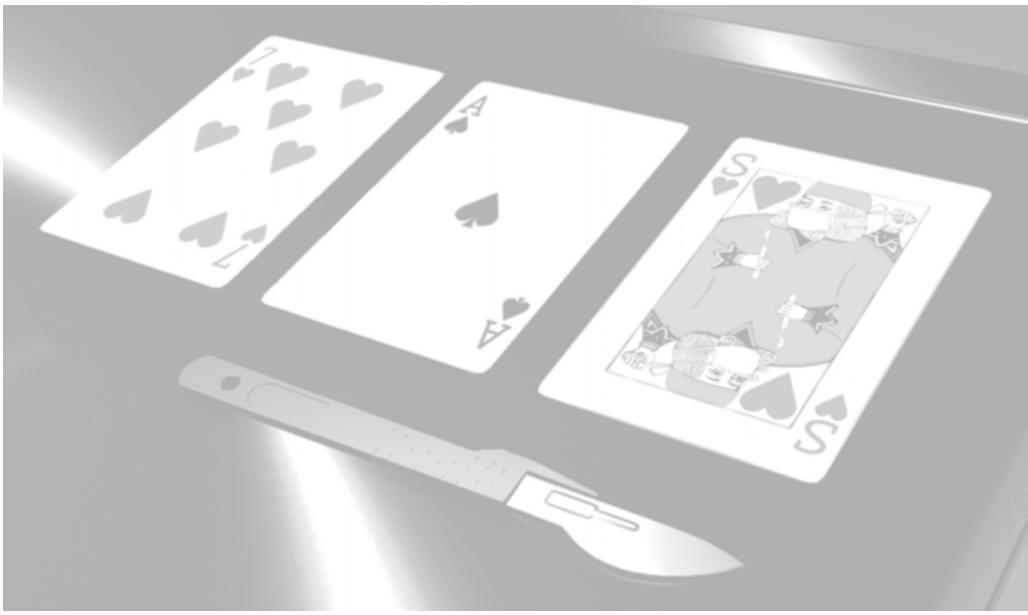
Blood loss > 200 mL/h:
As with 'blood loss > 100 mL/h' plus:

- use blood products only instead of plasma expanders if filling is required based upon hemodynamics;
- alternating transfusion of 2 platelet units and 1 FFP unit;
- after every second FFP unit (=every 6th unit) => 0.5 g CaCl2 slow IV;

Blood loss > 300 mL/h or persisting blood loss:

- if blood loss is more than 200 mL/h for 2 hours or more the surgeon is noti fied;
- if blood loss is more than 400 mL/h, more than 300 mL/h for 2 hours, or more than 200 mL/h for 3 hours and lab and TEG values are normal and no clots formation in the drains:
- 2 g tranexamic acid;
    - treat hypertension if existent;
        - the surgeon decides if a re-exploration is indicated.

CHAPTER 15



General discussion

Systematic reporting of early complications after cardiovascular interventions, systematic assessment of potential determinants of adverse outcomes (early morbidity and mortality), and the use of risk stratification systems provide us with powerful tools to measure, weigh and value the burden of complications after cardiac interventions, identify potentially modifiable risk factors and reduce the occurrence and costs of complications.

This thesis shows that there is room for improvement regarding outcome assessment and risk stratification approaches for conventional and emerging cardiovascular interventions. It also illustrates that the application of risk stratification models and outcome improvement programs offers opportunities to increase the quality of care associated with these procedures.

In this chapter, results from the studies in this thesis are discussed in the light of the research questions posed in Chapter 1.

OPTIMIZATION OF RISK STRATIFICATION MODELS FOR CARDIAC SURGERY AND TRANSCATHETER INTERVENTIONS

Risk factors differ from patient group to patient group depending for example on patient selection, case mix and background population. As set out in Chapters 1 and 2, the predictive validity of a model in a population that differs from the intended population can therefore be limited. For example, although a number of important variables, such as age, are risk factors for both coronary surgery and valve surgery (and many other procedures), other risk factors differ [1, 2]. For obvious reasons, risk stratification models for valve surgery often include variables regarding valve position or replacement versus repair, while these are absent in models developed to stratify coronary surgery.

It is therefore not surprising that a model dedicated to valve surgery has a higher predictive ability for valve surgery outcomes than the EuroSCORE, a general cardiac surgery model that was built using a database in which 65% of the patients underwent coronary surgery exclusively. This finding, presented in Chapter 3, was recently included in a meta-analysis, confirming the low discriminatory ability of the EuroSCORE for valve surgery [3].

Given these observations, the question remains to what extent models should be developed specific to the intervention that will be carried out. According to the Society of Thoracic Surgeons (STS) database committee, separate

models for different valve procedures are not more accurate than one model for all surgical valve procedures despite differences in pathophysiology and outcomes [2]. This seems like the best approach for benchmarking purposes, but for research or clinical purposes more detailed sub models are desirable. Separate models for isolated valve surgery and valve with concomitant coronary surgery were however derived from both the STS database and the New York State Cardiac Surgery Reporting system [1, 4]. The characteristics of these patient groups differ substantially and this separate model approach is therefore still favored by the STS database committee [1, 2, 5]. Currently, several valve dedicated risk stratification models are available (see Table 1 of Chapter 2). Risk stratification of patients undergoing valve surgery should ideally be performed using one of these models.

In clinical practice, classification as 'high risk' according to surgical risk stratification models is taken into account when patients are considered for transcatheter valve implantation, a novel percutaneous approach to treat severe aortic stenosis. However, existing surgical risk stratification models overestimate mortality in high risk patients undergoing aortic valve replacement [6-8]. This "outlier" population of patients that until recently were not even considered for aortic valve replacement differs too much from the population that has been used to develop the surgical risk models to reliably estimate early outcome. Results from Chapter 5 demonstrate that this also applies to the population that is currently offered transcatheter valve implantation. It is therefore likely that existing surgical risk models will not perform well when applied to the current transcatheter valve intervention population.

Chapter 6 describes that both the logistic EuroSCORE and the STS score are inadequate to (1) guide patient selection and (2) to predict outcome after transcatheter valve implantation. For these two objectives, no adequate models are yet available but are urgently needed for this fragile population. These models could respectively be derived from a group of high risk patients that underwent isolated surgical aortic valve replacement and a large population of patients that underwent transcatheter valve interventions. Although dedicated models for isolated aortic valve surgery –sub models of valve surgery models– are not favored by the STS database committee, they will likely be necessary for transcatheter procedures.

Results from surgical risk stratification models applied to the transcatheter

population should in the mean time be interpreted with great caution. Clinical judgment and patient-tailored decision-making by a team of surgeons and cardiologists plays a key role in selecting patients for transcatheter valve implantation.

As can be appreciated from Chapter 4, a considerable number of patients are denied referral for surgery based upon incorrect risk assessment. Interdisciplinary team discussions between cardiologists and surgeons should be encouraged to optimize patient selection for both transcatheter valve implantation and surgical valve replacement.

Improved awareness of the opportunities that transcatheter valve treatment offers for severe aortic valve stenosis may also lead to a better understanding of risks of and indications for surgical valve replacement, ultimately resulting in an increased referral for invasive treatment of aortic valve stenosis in patients who are currently managed conservatively.

CONGENITAL CARDIAC SURGERY AND TRANSCATHETER INTERVENTIONS

Although potential complications after cardiovascular procedures are well known for most conventional procedures, knowledge of complications and their determinants after emerging or infrequently performed procedures is incomplete. In this light, complications and their determinants were investigated for stent deployment in patients with congenital heart disease and for adult congenital cardiac surgery.

The large diversity of interventions in patients with congenital heart disease (CHD) and the widely varying severity of CHD complicates the study of associated risks. This contributes to the current situation, where knowledge on outcomes and risk stratification is not in line with emerging techniques such as stenting and surgery in the growing adult CHD population.
Knowledge on outcomes is the first step in risk assessment and subsequently allows for future risk model development efforts.

In this thesis it was shown that complications are common after stenting for CHD in both children and adults. The diversity of the underlying pathology combined with limited experience due to a relatively small patient population makes these procedures sensitive to complications. Fortunately, little major complications and few fatal complications were found. Increased operator experience

combined with technical improvements such as premounted stents for greater vessel diameters may reduce the risks of stenting in congenital heart disease. In addition, a European registry should be set up to further increase knowledge on determinants of complications of these procedures.

In recent years, more and more children with CHD are reaching adulthood, frequently requiring cardiosurgical (re-)interventions. Cardiac surgery in adult CHD patients can be performed with low operative mortality and good clinical outcome (Chapter 8). For this population, a concentration of knowledge and experience is also indicated.

The EuroSCORE is often used to stratify adult CHD surgical procedures but performs inadequate in this group of patients. The Aristotle Score and the RACHS-1 score [9, 10], both developed for pediatric congenital cardiac surgery, were therefore applied to this population. These models also had a poor discriminatory ability (Chapter 9). After studying all model variables, it became apparent that patient age was not properly represented for this population. In the EuroSCORE, age below 65 years does not affect predicted outcome, while in the pediatric models, increased age is not a risk factor. This led to the explorative approach of simply adding points for increased age to the standard pediatric congenital risk scores. Consequently, the discriminatory ability of these models increased greatly. This 'proof of concept' hopefully stimulates large database committees to further investigate risk stratification for this growing patient group.

As goes for outcome assessment of all interventions, CHD patients that were denied an intervention are not included in the analyses. These patients –who usually have extensive co-morbidity or are in a clinically critical condition– might have influenced outcome in a negative way had they been treated.

When interventional outcome is good for a particular institution, this might therefore be biased by a conservative approach. As selection and execution of treatment for congenital patients is often complex, this aspect of outcome assessment should not be overseen.

With increasing experience and improvement in techniques, updates of guidelines and risk stratification models for congenital surgery are needed on a regular basis. Professional organizations like the European Association for Cardio-Thoracic Surgery (EACTS) database may play a pivotal role in this respect.

ASPECTS OF THE ROLE OF PREDICTIVE FACTORS IN RISK STRATIFICATION MODELS

Coronary anatomy: adds value in addition to patient and procedural predictive factors

Most risk stratification models combine patient and procedure related risk factors, with little attention for variables based upon anatomical characteristics of the lesion(s) to be treated. In contrast, the SYNTAX Score describes the anatomy and complexity of coronary lesions exclusively [11]. Ultimately, the SYNTAX Score can be used to support treatment selection for patients with coronary artery disease who require an intervention by providing a complexity score that predicts adverse outcome.

The higher tertiles of the SYNTAX Score proved to be associated with an increase of major adverse cardiac and cerebrovascular events (MACCE) after PCI in the SYNTAX trial [12]. According to the results presented in Chapter 10, the SYNTAX Score can also accurately predict an increased risk of serious morbidity or death in a high risk surgical population.

Combining anatomical variables with important patient and procedure related variables might further increase accuracy of present risk stratification models. The SYNTAX Score could for example be combined with a number of other variables such as age and renal function (as for example in the ACEF score [13]) or fractional flow reserve (FFR; the ratio of maximal blood flow in a stenotic artery to normal maximal flow) values [14].

Renal dysfunction: choosing the most appropriate definition

Selection and definition of variables is important when evaluating the variables in risk stratification models. Chapter 11 focuses on renal dysfunction and how to optimally incorporate this risk factor in a risk stratification model. Model performance improved when renal dysfunction as a binary variable is replaced by a categorical or continuous variable. In addition, the measure to represent renal dysfunction also influences prediction results. According to previous studies, the estimated glomerular filtration rate is a more adequate measure to reflect renal dysfunction than serum creatinine level [15-17]. Results from Chapter 11 support these findings, although the improvement of discriminatory ability was not significant. These results illustrate the importance of careful consideration of the approach to convert risk factors to variables in risk stratification models.

When developing models for use in the current era, the use of a parsimonious approach to facilitate implementation rather than increasing workload should be pursued as long as model performance is not compromised in an unacceptable way. Fortunately, the widespread availability of computers has decreased the need to deprive models from additional information on important risk factors.

The case for parsimony: keep it simple!

In Chapter 12, the use of a risk stratification model using re-exploration for persistent blood loss as outcome measure is questioned. Mehta et al. chose to publish this model, despite its poor discriminatory ability [18]. If one chooses to accept implementation of a model with such poor performance, a more parsimonious approach might prove to work equally well; using renal function as a sole variable appears to provide a comparable discriminatory performance as the model proposed by Mehta et al based upon the STS database [18]. It may be speculated that a combination of a few core variables from their database predicts comparably mediocre as does this more complex model.

Risk stratification models, even a simple model as suggested in Chapter 12, might produce an acceptable predictive validity for groups of patients. Current models however are not sophisticated enough to direct treatment for individual patients. Unless frailty or an extensive list of risk factors, including rare comorbidities, are incorporated, clinical prediction models are to be used as tools only when used for patient education. Patients should be informed that percentages given by a model are based upon a grouped data of patients that are comparable with the patient in question and that a certain degree of uncertainty is involved. One should be aware of medical illiteracy and limited numeracy in many patients; adequate time and tools should be employed to ascertain that patients are well-informed [19].

Until now, no model has adequately incorporated patient frailty. This would require a model to contain all comorbidities, traumas and life events, as well as a functional performance evaluation. With this in mind, clinical judgment is still the best way to assess frailty.

Continuous quality improvement of risk stratification models

In the ever-changing field of cardiac surgery and interventions, continuous refinement and updating of risk stratification models is necessary to provide valid

estimates of outcome. Potential predictors should be evaluated on a regular basis and the weighing and methods of inclusion should be reassessed alike. Models that are not updated on a regular base are at risk to, for example, overestimate outcome as they have not been corrected for improved quality of care. When a model that generally overestimates mortality is used to assess observed mortality, results appear fine even when they are below par. This can lead to complacency and acceptance of current results when there is definite room for improvement. A frequently used risk stratification model that is known to overestimate outcome, is the EuroSCORE [8, 20]. This was among the reasons for recalibrating this model prior to substitution and investigation of renal dysfunction variables (Chapter 11). Data inclusion for the first EuroSCORE update started early 2010 [21], while data used for the original model dates back to 1995 [22]: a gap of 15 years. Had this model been updated on a more regular basis, performance might have even been acceptable in a valve surgery population as well (Chapter 3).

COST REDUCTION AND QUALITY IMPROVEMENT

As stated previously, the implementation of parsimonious prediction models to minimize an increased clinical workload is fine as long as performance of these models remains acceptable.

In the study presented in Chapter 13, variables are assessed that predict length of intensive care unit (ICU) stay. Although the models presented here performed statistically moderately, a decrease in over- and underestimated length of stay could be established in practice. This subsequently can improve quality as a decrease in rescheduled procedures can be realized. A more efficient use of operating theatres and ICU units will also decrease costs. More advanced models containing information on multiple hospital processes and resources should be developed to decrease costs and improve quality further. Inclusion of patient characteristics in these models will likely improve results and should therefore be examined further.

The simple tree approach in Chapter 13 to stratify patients allows for quick implementation where current (digital) structures limit the implementation of more sophisticated models for optimization of resources. Regardless of model technique, predictions made in Chapter 13 are based upon preoperative vari-

ables exclusively. Resource prediction models should be updated during the entire chain of care; such dynamic models allow for the inclusion of peri- and postoperative variables and improve accuracy of predictions for the remainder of the chain of care [23].

Although the emphasis in this thesis lies on the application of risk stratification models, another approach to cost reduction and quality improvement is also investigated. The introduction of a guideline based, multimodality approach in combination with continuous feedback has successfully decreased costs and increased quality (less RBC transfusions equals less complications) and could serve as the basis for a quality improvement program. This approach and the combined feedback and data collection methods can be implemented for other aspects of care for cardiovascular patients.

Compliance to a quality improvement program and systematic periodical review of performance and potential areas of improvement are essential to achieve progress. Results from Chapter 15 might therefore have been even better had compliance been higher.

CONCLUSION, PROSPECTS AND RECOMMENDATIONS

From the studies in this thesis it has become clear that several improvements can be made to risk stratification and outcome assessment approaches for conventional and emerging cardiac interventions in the current era. In line with the conclusions drawn in this thesis, several recommendations were formed.

Education of clinicians

Knowledge on outcome assessment and risk stratification is vital for clinicians performing interventions in the current era. Results from risk stratification models should be implemented with understanding of the fundamental aspects of model quality. Education of clinicians on the methodology of risk stratification modeling is essential for successful and appropriate application in clinical practice. Interdisciplinary team discussions between cardiologists and surgeons should be encouraged to optimize treatment selection for valve interventions. Special attention should also be given to the education of general practitioners regarding the interpretation of risks involved with available cardiovascular interventions to enable improved referral behavior.

Continuous quality improvement and proper implementation of models
Continuous refinement and updating of risk stratification models is necessary to provide valid estimates of outcome now and in the future. Database structures should allow for easy inclusion of additional variables.

Furthermore, risk stratification models are best suitable for the purpose or patient group they are used for and should be implemented accordingly. This includes recalibration at center level when used for patient counseling or clinical decision making [24].

Investments in models for emerging procedures
Risk models dedicated to new procedures or insights are not instantly available. To facilitate fast development of these models, data should be collected on a large scale as soon as new techniques emerge. There is a need for new large database initiatives for emerging procedures such as adult congenital surgery, transcatheter valve implantation and stent implantation. The importance of the modeling approach is secondary to this first step: data collection.

International collaborative efforts
Collaboration between research groups and database committees should be pursued to (1) share knowledge regarding outcome assessment and modeling approaches and to (2) increase database sizes. This allows for the development of reliable prediction models for less frequently performed procedures, such as aortic valve replacement in high risk patients or heart transplants.

The degree of model specificity should be assessed carefully for the desired model purpose. Variables and a variety of endpoints identified in the STS database should be included in new database initiatives to reflect the current standards of quality of care assessment [2, 5, 25]. This applies to both surgical procedures as catheter based procedures.

The eagerly awaited EuroSCORE update will be performed by collecting data on consecutive surgical procedures during several weeks in 2010 in centers that have volunteered to participate. Although this update will undoubtedly improve the EuroSCORE, continuous data collection on a larger scale will likely provide even more accurate models. A European version of the STS database should therefore seriously be considered. Governments or professional organizations should enforce participation in audited national data registries. These datasets could then be merged into one large European database, supervised by the

EACTS. The European Society of Cardiology (ESC) might play a similar role in the development of European interventional databases and models. When possible and appropriate, forces of surgeons and cardiologists should be joined as well.

Future research should furthermore focus on combining anatomical and physiological variables with important patient and procedure related variables.

Inclusion of patients that do not undergo invasive treatment

Until now, risk stratification databases and models contain data of patients that have been selected for and underwent treatment only. A major improvement would be to expand the current databases and include all symptomatic (or even non-symptomatic) patients that visit the cardiologist and their outcome. Patients that are currently denied surgery or interventional treatment should be included as well. Question is whether this tremendous effort is feasible. A challenging first step would be to overcome the reluctance to collect data for all cardiac patients. However, if implemented, this approach will greatly improve risk stratification for cardiac patients and ultimately provide patient-tailored state-of-the-art quality of care.

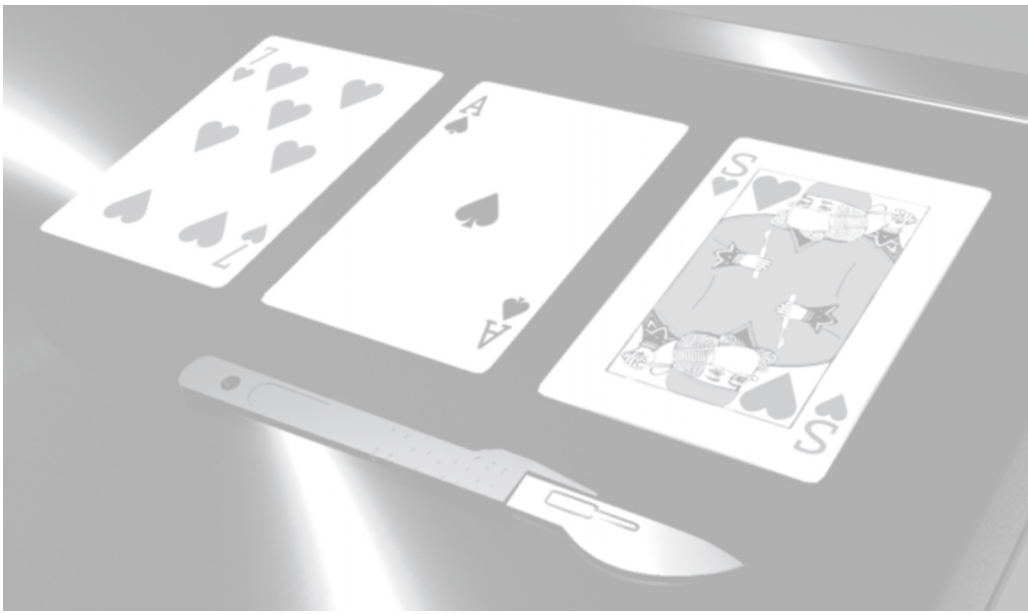
REFERENCES

1. Hannan EL, Wu C, Bennett EV, Carlson RE, Culliford AT, Gold JP, Higgins RS, Smith CR and Jones RH. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg. 2007;83(3):921-9.
2. O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2--isolated valve surgery. Ann Thorac Surg. 2009;88(1 Suppl):S23-42.
3. Parolari A, Pesce LL, Trezzi M, Cavallotti L, Kassem S, Loardi C, Pacini D, Tremoli E and Alamanni F. EuroSCORE performance in valve surgery: a meta-analysis. Ann Thorac Surg. 2010;89(3):787-93, 793 e1-2.
4. Edwards FH, Peterson ED, Coombs LP, DeLong ER, Jamieson WR, Shroyer ALW and Grover FL. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol. 2001;37(3):885-92.
5. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg. 2009;88(1 Suppl):S43-62.
6. Brown ML, Schaff HV, Sarano ME, Li Z, Sundt TM, Dearani JA, Mullany CJ and Orszulak TA. Is the European System for Cardiac Operative Risk Evaluation model valid for estimating the operative risk of patients considered for percutaneous aortic valve replacement? J Thorac Cardiovasc Surg. 2008;136(3):566-71.
7. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL and Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. J Thorac Cardiovasc Surg. 2008;135(1):180-7.
8. Osswald BR, Gegouskov V, Badowski-Zyla D, Tochtermann U, Thomas G, Hagl S and Blackstone EH. Overestimation of aortic valve replacement risk by EuroSCORE: implications for percutaneous valve replacement. Eur Heart J. 2009;30(1):74-80.
9. Lacour-Gayet F, Clarke D, Jacobs J, Comas J, Daebritz S, Daenen W, Gaynor W, Hamilton L, Jacobs M, Maruszewski B, Pozzi M, Spray T, Stellin G, Tchervenkov C and Mavroudis And C. The Aristotle score: a complexity-adjusted method to evaluate surgical results. Eur J Cardiothorac Surg. 2004;25(6):911-24.
10. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH and Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. J Thorac Cardiovasc Surg. 2002;123(1):110-8.
11. Sianos G, Morel MA, Kappetein AP, Morice MC, Colombo A, Dawkins K, van den Brand M, Van Dyck N, Russell ME, Mohr FW and Serruys PW. The SYNTAX Score: an angiographic tool grading the complexity of coronary artery disease. EuroIntervention. 2005;1(2):219-27.
12. Serruys PW, Onuma Y, Garg S, Sarno G, van den Brand M, Kappetein AP, Van Dyck N, Mack M, Holmes D, Feldman T, Morice MC, Colombo A, Bass E, Leadley K, Dawkins KD, van Es GA, Morel MA and Mohr FW. Assessment of the SYNTAX score in the Syntax study. EuroIntervention. 2009;5(1):50-6.
13. Ranucci M, Castelvecchio S, Menicanti L, Frigiola A and Pelissero G. Risk of assessing mor-

tality risk in elective cardiac operations: age, creatinine, ejection fraction, and the law of parsimony. Circulation. 2009;119(24):3053-61.

14. Tonino PA, De Bruyne B, Pijls NH, Siebert U, Ikeno F, van' t Veer M, Klauss V, Manoharan G, Engstrom T, Oldroyd KG, Ver Lee PN, MacCarthy PA and Fearon WF. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. N Engl J Med. 2009;360(3):213-24.
15. Wang F, Dupuis JY, Nathan H and Williams K. An analysis of the association between preoperative renal dysfunction and outcome in cardiac surgery: estimated creatinine clearance or plasma creatinine level as measures of renal function. Chest. 2003;124(5):1852-62.
16. Walter J, Mortasawi A, Arnrich B, Albert A, Frerichs I, Rosendahl U and Ennker J. Creatinine clearance versus serum creatinine as a risk factor in cardiac surgery. BMC Surg. 2003;3:4.
17. Jin R, Grunkemeier GL, Brown JR and Furnary AP. Estimated glomerular filtration rate and renal function. Ann Thorac Surg. 2008;86(1):1-3.
18. Mehta RH, Sheng S, O'Brien SM, Grover FL, Gammie JS, Ferguson TB and Peterson ED. Reoperation for Bleeding in Patients Undergoing Coronary Artery Bypass Surgery: Incidence, Risk Factors, Time Trends, and Outcomes. Circ Cardiovasc Qual Outcomes. 2009;2(6):583-590.
19. Fagerlin A, Ubel PA, Smith DM and Zikmund-Fisher BJ. Making numbers matter: present and future research in risk communication. Am J Health Behav. 2007;31 Suppl 1:S47-56.
20. Grossi EA, Schwartz CF, Yu PJ, Jorde UP, Crooke GA, Grau JB, Ribakove GH, Baumann FG, Ursumanno P, Culliford AT, Colvin SB and Galloway AC. High-risk aortic valve replacement: are the outcomes as bad as predicted? Ann Thorac Surg. 2008;85(1):102-6; discussion 107.
21. EuroSCORE 2010. Available at: http://www.euroscore2010.org. Accessed May 2010.
22. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S and Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. 1999;16(1):9-13.
23. Van Houdenhoven M, Nguyen DT, Eijkemans MJ, Steyerberg EW, Tilanus HW, Gommers D, Wullink G, Bakker J and Kazemier G. Optimizing intensive care capacity using individual length-of-stay prediction models. Crit Care. 2007;11(2):R42.
24. Jin R, Furnary AP, Fine SC, Blackstone EH and Grunkemeier GL. Using Society of Thoracic Surgeons risk models for risk-adjusting cardiac surgery results. Ann Thorac Surg.89(3):677-82.
25. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH and Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1--coronary artery bypass grafting surgery. Ann Thorac Surg. 2009;88(1 Suppl):S2-22.

CHAPTER 16



Summary

Samenvatting

Dankwoord

Curriculum Vitae

List of publications

PhD portfolio

SUMMARY

Chapter 1 comprises the introduction of this thesis. The purposes of outcome assessment and risk stratification are set out in the light of quality of care. This chapter describes the current situation and poses several questions to introduce the goal of this thesis, which is to study aspects of morbidity and mortality after cardiovascular procedures in current clinical practice, with special emphasis on the role of risk stratification models to improve outcome assessment.

Chapter 2 provides an overview of currently available risk stratification models for valve surgery, describing potential applications and quality of these models. In addition, future directions are given for the development of heart valve surgery risk models in the context of emerging transcatheter valve therapies. The question whether valve surgery requires dedicated risk stratification models instead of general cardiac surgery models is addressed in chapter 3. Results from this study suggest that models developed specifically for valve surgery perform better in this patient group than general cardiac surgery risk models such as the EuroSCORE.

Guidelines have been developed to assist in treatment selection for patients with valve disease. In Chapter 4, the reasons why a considerable number of patients are nonetheless denied aortic valve surgery and how proper risk assessment might influence these decisions is investigated

In chapter 5, differences between surgical valve replacement and transcatheter valve intervention patient cohorts are examined. Characteristics of these two patient cohorts differ too much to enable comparison or allow for the introduction of a randomized trial. Chapter 6 then focuses on the aspects of risk stratification and treatment selection for transcatheter valve intervention patients. Both the EuroSCORE and the STS Score perform suboptimal in this patient group.

As little is known about the outcomes and risk factors of stent deployment in patients with congenital heart disease, complications and their determinants for this population are investigated in chapter 7. Stenting in this patient group is associated with a considerable amount of complications, although only a small proportion is permanent.

Outcomes of adult congenital heart surgery are investigated in chapter 8, along with the use of the EuroSCORE risk stratification model in this population. The EuroSCORE is not suitable for this patient group. Chapter 9 further examines

risk stratification in this patient group. The use of risk stratification systems for adult congenital heart surgery is explored in this chapter. A combination of current pediatric congenital models and age seems to be a promising combination for this patient group.

For coronary surgery, many clinical prediction models are available. None of these models however include detailed anatomical variables. Recently, a consensus based risk stratification score using anatomical variables exclusively was developed for transcatheter coronary interventions and coronary surgery (the SYNTAX Score). Performance of this score is evaluated for a surgical population in chapter 10 with good results.

An important variable in many risk stratification models for cardiac surgery is renal dysfunction. Serum creatinine was incorporated as a binary variable in the EuroSCORE to reflect renal dysfunction. In chapter 11, the performance of the EuroSCORE is assessed when different definitions of renal dysfunction are incorporated. Results from this study suggest that the current way to reflect renal dysfunction in the EuroSCORE can be improved.

In chapter 12, a risk stratification using reoperation for bleeding as outcome measure is briefly discussed. What level of performance justifies clinical implementation of a risk stratification model? The level of parsimony is also addressed in this chapter.

Clinical prediction models can be used to predict length of stay, allowing for improvements in resource management. Two approaches to preoperatively predict intensive care unit length of stay after cardiac surgery are evaluated in chapter 13; a conventional approach and a tree model. They both appear to predict length of stay better than a prediction simply based on mean length of stay.

In chapter 14, the results of a multidisciplinary blood loss reduction program are evaluated. This program is implemented using a clinical decision support system and e-learning modules. The amount of lost blood and the use of blood products after introduction of this program are compared with a preceding time period. It appears that blood loss after cardiac surgery can be decreased by the introduction of a guideline based, multimodality protocol and that a higher degree of protocol adherence is associated with less blood loss.

Finally, in the general discussion in chapter 15 the results of the studies in this thesis are evaluated and conclusions are drawn. An outline for future studies is proposed.

SAMENVATTING

Hoofdstuk 1 is de introductie van dit proefschrift. In dit hoofdstuk wordt het nut van outcome assessment en risicostratificatie voor cardiovasculaire interventies beschreven, onder andere in het kader van de kwaliteit van de zorg. De huidige situatie wordt beschreven en het doel van dit proefschrift wordt uiteengezet:de bestudering van morbiditeit en mortaliteit bij cardiovasculaire ingrepen in de hedendaagse praktijk, met speciale aandacht voor de rol van risicostratificatie-modellen om outcome assessment te verbeteren.

Hoofdstuk 2 geeft een overzicht van de beschikbare risicomodellen voor hartklepchirurgie en de mogelijke toepassingen en kwaliteit van deze modellen wordt besproken. Daarnaast worden aanwijzingen gegeven voor de ontwikkeling van risicomodellen voor het plaatsen van hartkleppen middels catheters, een procedure die momenteel in opkomst is. Hoofdstuk 3 behandelt de vraag of er specifieke risicomodellen nodig zijn voor hartklepchirurgie in plaats van algemene hartchirurgische risicomodellen. De resultaten uit dit hoofdstuk suggereren dat een model als de EuroSCORE (een algemeen hartchirurgisch risicomodel) in een populatie patiënten die hartklepchirurgie ondergaat minder goed voorspelt in vergelijking met een model dat speciaal voor de hartklepchirurgie-populatie is ontwikkeld.

Voor de behandeling van patiënten met hartkleplijden zijn richtlijnen beschikbaar. In hoofdstuk 4 worden de oorzaken onderzocht waarom een aanzienlijk patiënten die volgens deze richtlijnen in aanmerking komen voor aortaklepchirurgie toch niet geopereerd worden. Hierbij wordt ook de rol van risicostratificatie bestudeerd..

In hoofdstuk 5 worden de verschillen onderzocht tussen een patientenpopulatie die chirurgische aortaklepvervanging heeft ondergaan en een patientenpopulatie die een implantatie van een aortaklep middels een catheter heeft ondergaan. Deze verschillen blijken erg groot te zijn en de groepen zijn dan ook niet goed vergelijkbaar. Slechts een klein deel van deze populaties zou eventueel geschikt zijn voor inclusie in een gerandomiseerde trial. Hoofdstuk 6 richt zich vervolgens op de aspecten van risicostratificatie bij hoog risicopatiënten met aortakleplijden. Momenteel zijn juist dit de patiënten die in aanmerking komen voor het plaatsen van een hartklep middels een catheter. Het lijkt erop dat zowel de EuroSCORE als de STS Score niet goed presteren in deze populatie.

Aangezien er weinig bekend is over de resultaten van en risicofactoren bij

het plaatsen van stents bij patiënten met een aangeboren hartafwijking, zijn complicaties en determinanten daarvan onderzocht in hoofdstuk 7.Hoewel er een aanzienlijke hoeveelheid complicaties bij deze procedures ontstaat, kan het merendeel van deze complicaties goed verholpen worden zonder blijvende schade.

De resultaten van hartchirurgie bij volwassenen met een aangeboren hartafwijking is onderzocht in hoofdstuk 8. De EuroSCORE lijkt niet geschikt te zijn voor deze patiëntengroep. Risicostratificatie in deze populatie wordt verder onderzocht in hoofdstuk 9. Het gebruik van risicomodellen wordt verder bekeken en een combinatie van huidige risicomodellen voor hartchirurgie bij kinderen met een aangeboren hartafwijking waaraan de factor leeftijd is toegevoegd geeft een betere voorspelling van vroege sterfte na hartchirurgie bij volwassenenen met een aangeboren hartafwijking.

Voor coronairchirurgie zijn vele risicomodellen beschikbaar. Deze modellen bevatten echter geen gedetailleerde anatomische variabelen. Onlangs is een risicomodel ontwikkeld dat exclusief anatomische variabelen bevat voor patiënten met coronairlijden. In hoofdstuk 10 wordt dit model onderzocht in een chirurgische populatie en met goede resultaten.

Een belangrijke variabele in veel risicomodellen is verminderde nierfunctie. Ook in de EuroSCORE is creatinine (serum) opgenomen als binaire variabele om een verminderde nierfunctie te reflecteren. In hoofdstuk 11 wordt de prestatie van dit model onderzocht wanneer een verminderde nierfunctie op andere manieren wordt geïncludeerd in dit model. Op basis van de resultaten van dit onderzoek blijkt dat de huidige aanpak voor verbetering vatbaar is.
In hoofdstuk 12 wordt een risicomodel besproken dat de kans op een reoperatie als gevolg van overmatig bloedverlies kan voorspellen. Hierbij is zowel aandacht voor het gebruikersgemak als de prestaties van dit model.

Voorspellende modellen kunnen ook gebruikt worden om de opnameduur in een ziekenhuis na een bepaalde procedure in te schatten. Hiermee kan bijvoorbeeld beter gepland worden om zo het gebruik van resources te optimaliseren. In hoofdstuk 13 worden twee manieren bestudeerd om preoperatief de verwachte opnameduur na hartchirurgie te voorspellen; een conventionele aanpak en een beslisboom. Beide methoden lijken de daadwerkelijke opnameduur beter te kunnen voorspellen dan voorspellingen op basis van een gemiddelde ligduur maar presteren gelijkwaardig ten opzichte van elkaar.

In hoofdstuk 14 worden de resultaten van de introductie van een multidis-

ciplinair protocol ter vermindering van postoperatief bloedverlies na volwassen open hartchirurgie geëvalueerd. Dit protocol is geïmplementeerd met behulp van e-learning modules en een clinical decision support system. De invoering van dit protocol heeft geleid tot een afname van de hoeveelheid bloedverlies. Daarnaast blijkt met name de mate van protocoltrouw van invloed te zijn op bloedverlies; het nauwgezet volgen van het protocol is geassocieerd met minder bloedverlies.

Ten slotte worden in hoofdstuk 15 de resultaten van alle studies in dit proefschrift besproken en worden conclusies getrokken. Daarnaast worden suggesties gegeven voor verder onderzoek.

DANKWOORD

Toen ik in 2007 de mogelijkheid kreeg om na mijn co-schappen promotieonderzoek te gaan doen, moest ik even goed nadenken. Na jaren studeren mocht ik dan eindelijk echt als dokter aan het werk en het verlaten van de kliniek leek dan ook een grote stap. Gelukkig bleek al snel dat ik omringd was door mensen die het promoveren tot een zeer aangename bezigheid maken. Dankzij hen kan ik nu terugkijken op een mooie periode en heb ik er geen seconde spijt van gehad dat ik deze weg ben ingeslagen. Een aantal mensen wil ik dan ook in het bijzonder bedanken voor hun bijdrage aan dit proefschrift.

Ten eerste mijn promotor, professor Bogers. Hartelijk dank dat u mij de mogelijkheid tot het voltooien van mijn proefschrift binnen –en nu en dan ook buiten– de afdeling Thoraxchirurgie heeft geboden. Het is bewonderenswaardig hoe u naast uw werkzaamheden in de kliniek en als afdelingshoofd nog zoveel tijd kunt vrijmaken voor de wetenschap. Met hoge snelheid komt elk manuscript weer retour afzender en zijn niet alleen alle puntjes op de i gezet, maar ook inhoudelijk zeer relevante opmerkingen geplaatst. De kalmte en rust waarmee dit alles gepaard lijkt te gaan, heeft een erg aangename werksfeer gecreëerd.

Mijn copromotoren, Pieter Kappetein en Hanneke Takkenberg. Beste AP, als Secretary General lijkt het vaak nodig te zijn om van OK naar vliegtuig te rennen en weer terug. En precies op die momenten moest ik natuurlijk mijn slag zien te slaan om het één of ander te bespreken, wat gelukkig altijd wel even mogelijk bleek. En meestal was er dan op het vliegveld wel tijd om plannen voor een grondige revisie door te mailen zodat ik voorlopig weer van de straat was. Hartelijk dank voor de vele goede ideeën en kritische blik op mijn brouwsels.
Beste Hanneke, zelfs vanuit Amerika was samenwerken met jou altijd een enorm plezier. Hoewel de colleges statistiek tijdens mijn studietijd compleet aan mij voorbij zijn gegaan, heb ik me door jouw enthousiasme met veel plezier door alle vraagstukken heen kunnen werken. Al snel bleken vele statistische vraagstukken van een probleem naar een uitdaging en uiteindelijk naar een eitje te gaan. Ik wil je niet alleen bedanken voor jouw kundige begeleiding, maar ook voor de vrolijke manier waarop je dat altijd hebt gedaan. Nu mijn proefschrift is afgerond, kan ik eindelijk de noodzaak van de oprichting van PENA rustig herevalueren. Hopelijk zijn we nog op tijd…

Professor Boersma, professor Klautz, en professor Steyerberg, de leden van mijn leescommissie. Mijn dank dat u allen bereid bent geweest om zitting te nemen in mijn promotiecommissie en de wetenschappelijke waarde van het proefschrift heeft willen beoordelen.
Beste professor Steyerberg, dank voor de nimmer aflatende interesse en bereidheid tot hulp tijdens de totstandkoming van dit proefschrift. Er liggen nog vele interessante ideeën op de plank dus ik verwacht dat ik binnenkort weer voor de deur sta om de nieuwste R-trucjes af te kijken.
Professor Roos-Hesselink en professor Klein: dank dat u bereid was in de grote promotiecommissie plaats te nemen.

Maarten Witsenburg, mijn begeleider bij mijn afstudeerscriptie, destijds nog bij de kindercardiologie. Hier is mijn interesse in onderzoek eigenlijk onstaan. Dankzij de zeer prettige begeleiding heb ik toen mijn eerste publicatie weten te voltooien, waarvoor mijn dank. Wellicht kunnen we daar in de nabije toekomst eindelijk dat goede glas whisky nog op drinken. Dan kunnen we er meteen nog één achteraan doen op dit proefschrift.

Els Berenschot, Liesbeth Duininck, Ria Hussaarts, Marijke Rozema, Liz van der Velden, Angeliek Venema, dames van de onderzoeksgroep. Dank voor al jullie hulp bij het verzamelen van kostbare data en immer snelle aanlevering hiervan.

Martijn van Geldorp. Beste Martijn, na je opleidingsplaats volgt ongetwijfeld snel ook jouw boekje. Ik hoop dat we ons daarna weer met belangrijke zaken kunnen bezighouden, zoals het restaureren van oldtimers onder het genot van een lekker biertje.

Helena Heuvelman, Mostafa Mokhles en Stuart Head, collegae promovendi. Dank voor de prettige samenwerking en veel succes met het completeren van jullie proefschrift. Ik ben benieuwd naar jullie resultaat.

Nicolo Piazza. Dear Nico, thank you for the excellent projects we did together. Your drive (and the Maggotbrain CD) really was an inspiration and working with you has always been a great pleasure. I wish you all the best, the completion of your thesis is just a matter of time. Have fun in Germany!

Alle medewerkers van het secretariaat Thoraxchirurgie, hartelijk dank voor de ondersteuning en prettige samenwerking. Connie Wemelsfelder, helemaal verstopt in de cardiologische gang, ook jij hartelijk dank voor het samenrapen en opschonen van de benodigde JANUS data.

Beste Jos Bekkers en Lex Maat, gedurende mijn tijd als jullie kamergenoot heb ik genoten van de levendige anekdotes uit de kliniek. Dank voor deze prettige afwisseling tijdens de keiharde analyses. Ik hoop dat jullie niet teveel hebben geleden onder mijn muziekinstallatie.

Robert van Thiel en Ard Struijs, hartelijk dank voor jullie intensieve hulp bij de DEBRIEF studie. Jan Hofland en alle andere anaesthesiologen, dames en heren van de perfusie, ook jullie hartelijk dank voor de hulp hierbij. Sietske Focking, Marja van der Jagt en alle andere verpleegkundigen. Dank voor het enthousiasme waarmee ook jullie data hebben verzameld en daarbij bijgedragen hebben aan de totstandkoming van dit proefschrift. Binnenkort zal ik weer eens een fijne bak met drop en chocolade afleveren.

Chirurgen, collega assistenten en nurse practitioners, bedankt voor de prettige werksfeer, gezellige momenten en alles wat ik van jullie leer.

Arjan van Cappellen en Gerard Draisma, mijn paranimfen. Dank dat jullie mij terzijde willen staan tijdens de verdediging van mijn proefschrift.

Onno, klein broertje, als vaste leverancier van arbeidsvitaminen heb je voor de nodige ontspanning gezorgd gedurende mijn gehele onderzoeksperiode. Daarnaast zorg je al zolang ik me kan herinneren voor de nodige ontspanning door je briljante humor en sta je altijd klaar, waarvoor dank.

Pap en mam, dank voor alle steun die jullie mij mijn hele leven al gegeven hebben. Nu ik *iet*wat ouder begin te worden, begin ik steeds meer te zien hoeveel jullie altijd voor ons gedaan hebben. Mijn dankbaarheid is dan ook groter dan ik hier met woorden kan weergeven.

Lieve Nina, hoewel de eerste onderzoeksjaren heel veel vrije tijd boden, is dit de laatste maanden wel anders geweest. Zoals altijd wegen de laatste loodjes het zwaarst en dat bleek ook bij de voltooiing van dit boekje (in combinatie met klinische werkzaamheden) het geval te zijn geweest. Hierdoor heb jij dikwijls niet de aandacht gekregen die je verdient. Dank je voor je begrip en steun gedurende deze periode. Vanaf nu weer méér tijd voor mooie dingen.

CURRICULUM VITAE

Menno van Gameren was born August 11, 1981 in Rotterdam, The Netherlands. After graduating from Gymnasium Erasmianum in Rotterdam in 1999, he studied medicine at the University of Antwerp, Belgium. After one year, he got the opportunity to study medicine at the Erasmus University Rotterdam and moved back to Rotterdam. In April 2007, he obtained his medical degree and in May 2007, he started his PhD project at the Department of Cardio-Thoracic Surgery of the Erasmus University Medical Center. In January 2010 he started as a resident at the Department of Cardio-Thoracic Surgery.

LIST OF PUBLICATIONS

Papers

1. van Gameren M, Putman LM, Takkenberg JJ and Bogers AJ. Risk stratification for adult congenital heart surgery. Eur J Cardiothorac Surg. Published online ahead of print on September 14th 2010.

2. van der Pijl LL, Birim O, van Gameren M, Kappetein AP, Maat AP, Steyerberg EW and Bogers AJ. Validation of a prognostic model to predict survival after non-small-cell lung cancer surgery. Eur J Cardiothorac Surg. Published online ahead of print on April 30th 2010.

3. van Gameren M, Piazza N, Bogers AJ, Takkenberg JJ and Kappetein AP. How to assess risks of valve surgery: quality, implementation and future of risk models. Heart. 2009;95(23):1958-63.

4. Birim O, van Gameren M, Bogers AJ, Serruys PW, Mohr FW and Kappetein AP. Complexity of coronary vasculature predicts outcome of surgery for left main disease. Ann Thorac Surg. 2009;87(4):1097-104.

5. Birim O, van Gameren M, de Jong PL, Witsenburg M, van Osch-Gevers L and Bogers AJ. Outcome after reoperation for atrioventricular septal defect repair. Interact Cardiovasc Thorac Surg. 2009;9(1):83-7.

6. van Geldorp MW, van Gameren M, Kappetein AP, Arabkhani B, de Groot-de Laat LE, Takkenberg JJ and Bogers AJ. Therapeutic decisions for patients with symptomatic severe aortic stenosis: room for improvement? Eur J Cardiothorac Surg. 2009;35(6):953-7.

7. Otten AM, van Domburg RT, van Gameren M, Kappetein AP, Takkenberg JJ, Bogers AJ, Serruys PW and de Jaegere PP. Population characteristics, treatment assignment and survival of patients with aortic stenosis referred for percutaneous valve replacement. EuroIntervention. 2008;4(2):250-5.

8. Piazza N, Wenaweser P, van Gameren M, Pilgrim T, Tzikas A, Otten A, Nuis R, Onuma Y, Cheng JM, Kappetein AP, Boersma E, Juni P, de Jaegere P, Windecker S and Serruys PW. Relationship between the logistic EuroSCORE and the Society of Thoracic Surgeons Predicted Risk of Mortality score in patients implanted with the CoreValve ReValving system--a Bern-Rotterdam Study. Am Heart J.159(2):323-9.

9. Piazza N, van Gameren M, Juni P, Wenaweser P, Carrel T, Onuma Y, Gahl B, Hellige G, Otten A, Kappetein AP, Takkenberg JJ, van Domburg R, de Jaegere P, Serruys PW and Windecker S. A comparison of patient characteristics and 30-day mortality outcomes after transcatheter aortic valve implantation and surgical aortic valve replacement for the treatment of aortic stenosis: a two-centre study. EuroIntervention. 2009;5(5):580-8.

10. Putman LM, van Gameren M, Meijboom FJ, de Jong PL, Roos-Hesselink JW, Witsenburg M, Takkenberg JJ and Bogers AJ. Seventeen years of adult congenital heart surgery: a single centre experience. Eur J Cardiothorac Surg. 2009;36(1):96-104.

11. van Gameren M, Kappetein AP, Steyerberg EW, Venema AC, Berenschot EA, Hannan EL, Bogers AJ and Takkenberg JJ. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg. 2008;85(3):921-30.

12. van Gameren M, Witsenburg M, Takkenberg JJ, Boshoff D, Mertens L, van Oort AM, de Wolf D, Freund M, Sreeram N, Bokenkamp R, Talsma MD and Gewillig M. Early complications of stenting in patients with congenital heart disease: a multicentre study. Eur Heart J. 2006;27(22):2709-15.

13. van Gameren M, Kappetein AP, Hofland J, van Thiel RJ, Boks R, Takkenberg JJM, Bogers. AJJC Effects of the introduction of an e-learning supported multimodality protocol to reduce blood loss. Submitted.

Letters

1. Van Gameren M, Kappetein AP, Bogers AJJC, Takkenberg JJM. Parsimony and Performance. Online reaction to Circ Cardiovasc Qual Outcomes. 2009; 2: 583-90.

Book chapters

1. A. Pieter Kappetein, Menno van Gameren, Ad J.J.C. Bogers, and Nicolo Piazza (2010). Risk Stratification for Transcatheter Aortic Valve Implantation. In: Patrick W. Serruys (Editor), Alain Cribier Md (Editor), John Webb (Editor), Jean-claude Laborde (Editor), Nicolo Piazza (Editor), Peter De Jaegere. Transcatheter Aortic Valve Implantation: Tips and Tricks to Avoid Failure. New York: Informa Healthcare. p56-63.

PHD PORTFOLIO

Name PhD student:	M. van Gameren
Erasmus MC Department:	Cardio-thoracic Surgery
Research School:	COEUR
PhD period:	2007-2010
Promotor:	Prof.dr. A.J.J.C. Bogers
Supervisors:	Dr. A.P. Kappetein
	Dr. J.J.M. Takkenberg

PhD Training	Year	Workload (ECTS)
In depth courses		
• Pathophysiology of ischemic heart disease	2007	1.5
• Cardiovascular imaging and diagnostics	2008	1.5
• Clinical cardiovascular epidemiology	2008	1.5
• NHS course: Thrombosis and Haemostatis	2009	2.0
• Congenital heart disease	2008	1.5
• FCCS course: Intensive care	2010	1.5
Presentations		
• Invited presentation on e-learning in cardiac surgery during the Ross Summit (Atlanta)	2008	0.6
• Minimizing blood loss during open heart surgery COEUR Seminar	2010	0.6
• Invited presentation on blood management during the European Association for Cardio Thoracic Surgery meeting (Geneva)	2010	0.6
Seminars and workshops		
• Pulmonary circulation	2007	0.4
• CV imaging	2007	0.4
• Tetralogy Fallot	2008	0.4
• Clinical Decision Making in Cardiovascular Interventions	2010	0.4

PhD Training	Year	Workload (ECTS)
International conferences		
• At the heart of evolution (Aortic Valve Symposium)	2007	0.3
• European Association for Cardio Thoracic Surgery annual meeting in Lisbon	2008	1.5
• Ross Summit in Atlanta	2008	0.6
• Society for Heart Valve Disease annual meeting in Berlin	2009	1.2
• European Association for Cardio Thoracic Surgery annual meeting in Geneva	2010	1.5
Others		
• Meetings of the Dutch Association for Thoracic Surgery	2007-2009	1.5
Teaching activities	**Year**	**Workload (ECTS)**
Supervising students	2008-2009	2

Additional financial support for the publication of this thesis by the following companies is gratefully acknowledged: