BMJ Open

Volume-outcome relationship in transcatheter aortic valve implantations in Germany 2008-2014

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020204
Article Type:	Research
Date Submitted by the Author:	23-Oct-2017
Complete List of Authors:	Kaier, Klaus; Medical Center-University of Freiburg, Institute for Medical Biometry and Statistics Oettinger, Vera; Heart Center Freiburg University, Department of Cardiology and Angiology I Reinecke, H; University of Muenster, Department of Cardiology and Angiology, Adult Congenital and Valvular Heart Disease Center Muenster Schmoor, Claudia; Medical Centre, University of Freiburg Frankenstein, L.; University of Heidelberg, Dpt. of Cardiology, Angiology, Pulmology Vach, Werner; Institute of Medical Biometry and Medical Informatics, Hehn, Philip; Faculty of Medicine and Medical Center – University of Freiburg von zur Mühlen, Constantin; Heart Center Freiburg University, Department of Cardiology and Angiology I Bode, Christoph; Heart Center Freiburg University, Department of Cardiology and Angiology I Zehender, Manfred; Heart Center Freiburg University, Department of Cardiology and Angiology I Reinöhl, Jochen; Heart Center Freiburg University, Department of Cardiology and Angiology I
Keywords:	Valvular heart disease < CARDIOLOGY, HEALTH ECONOMICS, CARDIOLOGY, Adult cardiology < CARDIOLOGY

Volume-outcome relationship in transcatheter aortic valve implantations in Germany 2008-2014

Klaus Kaier, PhD[1,2]; Vera Oettinger[2]; Holger Reinecke, MD[3]; Claudia Schmoor, PhD[4]; Lutz Frankenstein, MD[5]; Werner Vach, PhD[1]; Philip Hehn[1]; Constantin von zur Mühlen, MD[2]; Christoph Bode, MD[2]; Manfred Zehender, MD, PhD[2]; Jochen Reinöhl, MD[2]

[1] Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center – University of Freiburg, Germany

[2] Department of Cardiology and Angiology I, Heart Center Freiburg University, Germany

[3] Division of Vascular Medicine, Department of Cardiovascular Medicine, University Hospital Muenster, Germany

[4] Clinical Trials Unit, Medical Center – University of Freiburg, Germany

[5] Department of Cardiology, Angiology, Pulmonology, University of Heidelberg, Germany

Corresponding author:

Klaus Kaier, PhD

Medical Center – University of Freiburg, Institute of Medical Biometry and Statistics

Stefan-Meier-Str. 26

79104 Freiburg, Germany

phone: + 49 (0)761 203-6807

e-mail: **kaier@imbi.uni-freiburg.de**

Word count excluding title page, abstract, references, figures and tables: 2838

ABSTRACT

Keywords: TAVI, volume-outcome; minimum volume standards; hospital quality; mortality

Objectives: We examine the volume-outcome relationship in isolated transcatheter aortic valve implantations (TAVI) in Germany between 2008 and 2014 (N=43,996). Our interest was whether the volume-outcome relationship for TAVI exists on the center level, whether it occurs equally for different outcomes, and how it develops over time.

Methods: The comprehensive German Federal Bureau of Statistics DRG database was queried for data on all isolated TAVI procedures performed in Germany between 2008 and 2014. Logistic and linear regression analyses were carried out for the endpoints in-hospital mortality, bleeding, stroke, probability of ventilation >48 hours, length of hospital stay, and reimbursement. Risk-adjustment was applied using a predefined set of patient characteristics to account for differences in the risk factor composition of the patient populations between centers and over time. Centers performing TAVI were stratified into groups performing <50, 50-99, and ≥100 procedures per year.

Results: Risk-adjusted in-hospital mortality steadily decreases over the years and is lower the higher the annual procedure volume at the respective center is. The magnitude of the latter effect declines over the observation period. Overall, our results indicate a ceiling effect in the volume-outcome relationship: The volume-outcome relationship is eminent in circumstances of relatively unfavorable outcomes. Alongside improving outcomes, however, the volume-outcome relationship decreases. In addition, a volume-outcome relationship seems to be absent in circumstances of constantly low event rates.

Conclusions: The hypothesized volume-outcome relationship for TAVI exists but diminishes and may disappear over time. This might be the case for other interventional procedures, too, which should be taken into account when considering mandatory minimum thresholds.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Study based on administrative data; coding errors are inevitable. However, about 20% of all cardiovascular diagnosis–related groups are reviewed by independent teams of physicians on behalf of the health insurers.
- Risk-adjustment included a number of parameters whose reliability cannot be fully secured, and we cannot guarantee that all parameters of relevance are included in the model. A major limitation is that the data source does not include information on the type of device used in individual TAVI procedures. In addition, information regarding the experience of surgeons at each centre would be highly relevant for the analysis but is unavailable.
- The dataset omits patients with a baseline diagnosis of pure aortic regurgitation, as well as those who underwent TAVI with any other concomitant cardiac procedure. This makes sense from a clinical perspective, but further complicates direct comparisons with other administrative datasets.
- The study provides comprehensive data on everyday TAVI practice in a large industrialized country over a multiyear period.

KEY QUESTIONS

What is already known about this subject?

After their introduction, outcomes of new interventions are subject to a learning curve effect, meaning that outcomes improve over a period of time and then level off. The volume of procedures performed at an institution can influence this process, and is thought to have some effect on patient outcomes even after learning is complete (volume-outcome-hypothesis).

What does this study add?

This study tracks patient outcomes by center procedure volume in all transcatheter aortic valve implantation (TAVI) procedures performed in Germany between the procedure's introduction in 2008 and 2014, providing empirical evidence on shape and extent of the above described effects for this procedure.

How might this impact on clinical practice?

This data is of interest to clinical practitioners, hospital administrators, and policy makers involved in the implementation of new clinical procedures.

COMPETING INTERESTS AND FUNDING

The authors declare no conflicts of interest. There was no external funding for this work.

DATA SHARING STATEMENT

No additional data available.

AUTHORSHIP STATEMENT

KK and JR developed the research question and designed the methodology. VÖ, HR, LF, CvzM, CB, MZ and JR provided the medical knowledge of German TAVI practice informing the study design. KK defined the categories, outcomes and measures and developed and implemented the formal analysis and statistical with support by WV and CS. KK and JR collected the data and evidence. KK, VÖ, and WV interpreted and contextualized the results. KK and PH wrote the initial draft of the article, with JR contributing. All authors participated in the critical revision of the article and provided final approval of the version to be published.

INTRODUCTION

Transcatheter aortic valve implantation (TAVI) is a rapidly evolving technique for therapy of aortic stenosis, with a very early and pronounced utilization in Germany [1]. Previous studies report hospital-specific learning curves with respect to in-hospital outcomes such as procedural success, mortality and clinical complications of varying lengths and magnitudes [2–6]. In general, learning curve effects within and between centers can to some degree be explained by the volume of procedures performed at the center. This relationship can be summed up as the "practice-makes-perfect hypothesis", according to which quality of care either increases with the number of patients as a result of economies of scale, with a competing explanation of "selective-referral", according to which higher-quality hospitals attract greater demand and therefore have a greater volume of patients [7,8].

There are a number of criticisms on empirical analyses on the volume-outcome relationship: Many studies lack appropriate adjustment for differences in the risk factor composition of the patient populations between centers [9,10]. Secondly, most studies focus on in-hospital mortality only [11], which is easy to measure, but it is recommended to include additional quality measurements. Finally, most studies divided patients into groups of equal size for analyzing the volume-outcome relationship, which makes it difficult to make use of such results when justifying specific volume thresholds[6,12,13].

Although the evidence regarding the existence of an inverse relationship between the number of TAVI procedures and related outcomes is limited [14,15], medical authorities in Germany and several other countries have issued guidelines calling for minimum numbers of procedures for primary operators performing TAVI [16–19]. There however remains some question whether, firstly, the volume-outcome relationship outlined above exists on the center level regarding TAVI and, secondly, whether or not it takes place in all outcomes and complications equally, and how an existing volume-outcome relationship might change over the years.

To address these questions, we calculated annual procedure volumes for all German hospitals that performed TAVI procedures between January 2008 and December 2014. In order to account for differences in the patient population between high-, medium-, and low-volume centers and over time, we carried out baseline-adjusted regression analyses for the endpoints in-hospital mortality, bleeding, stroke, probability of ventilation >48 hours, length of hospital stay, and reimbursement.

METHODS

Data

Since 2005, data on all hospitalizations in Germany have been available for scientific use via the Diagnosis Related Groups (DRG) statistics collected by the Research Data Center of the Federal Bureau of Statistics (DESTATIS). These hospitalization data, including diagnoses and procedures, are a valuable source of representative nationwide data on the in-hospital treatment of patients. This database represents a virtually complete collection of all hospitalizations in German hospitals that are reimbursed according to the DRG system. From this database [1], we have extracted data on 43,996 cases of isolated TAVI for our analysis.

Our study did not involve direct access by the investigators to data on individual patients but only access to summary results provided by the Research Data Center. Therefore, approval by an ethics committee and informed consent were determined not to be required, in accordance with German law. All summary results were anonymized by DESTATIS. In practice, this means that any information allowing the drawing of conclusions regarding a single patient or a specific hospital are censored by DESTATIS to guarantee data protection. Especially the use of the anonymous, persistent "institute indicator of hospitals" is highly restricted in order not to publish any information directly attributable to a single hospital.

As described previously [1,20], we were able to use the OPS codes (OPS codes: 5-35a.0 in 2007 and 5-35a.00, 5-35a.01 from 2008) to identify all TAVI procedures performed (and reimbursed) in Germany between 2008 and 2014. Patients with a baseline diagnosis of pure aortic regurgitation (main or secondary diagnosis other than I35.0, I35.2, I06.0, I06.2) and those with concomitant cardiac surgery or percutaneous coronary intervention were not included in this analysis. A complete list of procedure codes as well as a more detailed discussion of the validity of the data source may be found in a previous manuscript [1,20].

Measures

Regarding the in-hospital complications, bleeding was defined as requiring a transfusion of more than 5 units of red blood cells (RBC). For all other comorbidities and complications the existing anamnestic or acute distinctive codes were used (we have discussed OPS and ICD codes in greater detail previously [20]).

In order to analyze possible effects of the above discussed mandatory minimum quantities, the number of procedures per year and center was categorized (i.e. n<50, 50≤n<100, n≥100) on the basis of an anonymous, persistent "institute indicator of hospitals" provided by DESTATIS. These particular thresholds are applied because the minimum number of 50 procedures is often mentioned in official TAVI-guidelines [16–19], and these thresholds are widely applied in the literature [21–23].

The primary outcome was in-hospital mortality. Secondary outcomes include post-procedural complications such as stroke and bleeding events (transfusion of >=5 RBC), as well as reimbursement, length of hospital stay and proportion of patients with ventilation >48h.

Statistical analysis

In a first step, multivariate regression analyses were carried out for the different endpoints, with all available patient and procedural characteristics (as defined by Reinöhl et al. [1]) included as covariates (all covariates listed in Table 1). In addition, an interaction term between time (in years) and the above mentioned annual volume categories was included in the regression analyses in order to investigate the volume-outcome relationship over the years.

Table 1: Baseline characteristics (2008-2014)

N	43,996
Female	55.87%
Age in years, mean/SD	80.95/6.11
Estimated logistic EuroSCORE[1], mean/SD	22.21%/13.57%
Aortic valve stenosis	68.22%
Combined aortic valve diseases	26.56%
Heart failure	
NYHA II	8.26%
NYHA III or IV	41.66%
Hypertension	62.66%
CAD	46.88%
Previous myocardial infarction	
within 4 months	1.59%
within 1 year	0.75%
after 1 year	4.35%
Previous CABG	12.75%
Previous cardiac surgery	18.06%
Peripheral vascular disease	12.39%
Carotid disease	6.17%
COPD	15.14%
Pulmonary hypertension	22.32%
Renal disease	
GFR <15ml/min	2.95%
GFR <30ml/min	4.90%
Atrial fibrillation	45.93%
Diabetes	33.30%

[1]For calculation of the logistic EuroSCORE, we were able to populate all fields except for critical preoperative state and left ventricular function. In these we assumed an inconspicuous state (i.e. no critical preoperative state and no left ventricular dysfunction) and thus calculated a best-case scenario.

Abbreviations: NYHA – New York Heart Association Functional Classification; CAD – coronary artery disease; CABG – coronary artery bypass graft; COPD – chronic obstructive pulmonary disease; GFR – glomerular filtration rate.

Please note that in comparison to the data published by Reinöhl et al., one TA-TAVI procedure (in 2010) needed to be removed from the dataset due to incomplete information.

Logistic and linear regression analyses are applied for dichotomous and continuous endpoints, respectively. The question of how to account for patients treated in the same hospital was discussed previously [13,24,25]. As recommended in a previous study that also used data from the German DRG-statistic [13], we used cluster-robust standard errors to account for this dependency.

Risk-adjusted rates and means within each year and hospital volume category were obtained by computing the corresponding predicted probabilities or means, respectively, for an artificial subject with each confounder set to its mean value (prediction at the means, see Table 1 for mean values of all confounders).

The visualization of these risk-adjusted rates or means together with their 95% confidence intervals constitutes the main analytical approach in this paper. To assess the statistical significance of the observed volume-outcome relationship, of the time trend and a potential change of the volume-outcome relationship over time, we applied to the estimated rates or means a random effects meta regression (command metareg [26]) with time and volume as continuous covariates. A model with an interaction term was used to assess the change in the volume-outcome relationship. A model without an interaction was used to assess the main effects.

All analyses were carried out using Stata 13.1 (StataCorp, College Station, Texas, USA).

RESULTS

Between 2008 and 2014, a total of 43,996 TAVI procedures were performed in 113 different centers in Germany. The total number of TAVI procedures performed per year increased markedly over the observation period, from 1,122 in 2008 to 11,559 in 2014 (see Table 2).

Table 2: Number of procedures with regard to the performed TAVI volume of a distinct center in a given year.

TAVI Volume in Center	2008	2009	2010	2011	2012	2013	2014
<50 procedures, n (number of centers)	613 (40)	1,234 (61)	1,155 (51)	1,107 (43)	960 (36)	765 (31)	617 (30)
50-99 procedures, n (number of centers)	236 (3)	658 (10)	1,875 (26)	1,957 (27)	1,569 (20)	1,930 (25)	1,135 (16)
>=100 procedures, n (number of centers)	273 (n/a*)	707 (n/a)	1,776 (3)	3,459 (7)	5,711 (16)	6,452 (9)	9,807 (20)
Total number, n (number of centers)	**1,122 (>=44)**	**2,599 (>=72)**	**4,806 (80)**	**6,523 (77)**	**8,240 (72)**	**9,147 (65)**	**11,559 (66)**

Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year. Number of centres in parentheses.

* n/a = not available, exact number censored by DESTATIS due to data protection concerns

As reported previously [1], substantial reductions in in-hospital mortality have been achieved between 2008 and 2013, and we find this trend to continue into 2014. Regarding center-specific procedure volumes of all TAVI procedures, it appears that the differences in unadjusted in-hospital mortality between the procedure volume groups (<50, 50-99, and >=100) steadily decline over the years (see Table 3). Figure 1 A provides risk-adjusted in-hospital mortality rates allowing for comparison despite possible differences in the patient selection process and consequently the risk factor composition between hospitals in the different procedure volume groups and over time (See Table S1 for details of the process used to generate the results shown in Figure 1A). These results indicate that risk-adjusted in-hospital mortality rates (1) steadily decrease over the years (annual change: -0.58 percentage points (pp), p<0.001), are (2) lower the higher the procedure volume at the hospital is (volume effect: -0.74pp, p=0.002), but that (3) this volume effect declines over the seven year observation period (p-value of interaction term: p=0.027; annual change of volume effect: 0.2pp).

Table 3: Unadjusted in-hospital outcomes with regard to the performed TAVI volume of a distinct center in a given year.

	Mortality, %	Stroke, %	Bleeding, %	Length of stay, mean in days	Reim-bursement, mean in €	Proportion of patients with ventilation >48h, %
2008						
<50 procedures	10.11%	3.26%	14.36%	19.2		9.79%
50-99 procedures	9.32%	2.12%	11.44%	21.8		6.78%
>=100 procedures	6.59%	2.56%	7.33%	14.7		4.76%
2009						
<50 procedures	9.81%	3.57%	14.18%	21.6		9.48%
50-99 procedures	8.36%	3.34%	11.25%	18.5		7.14%
>=100 procedures	6.08%	2.12%	7.21%	18.0		7.36%
2010						
<50 procedures	9.00%	2.51%	12.12%	21.0	37,071€	8.74%
50-99 procedures	8.11%	2.56%	11.41%	19.1	36,173€	8.69%
>=100 procedures	6.14%	2.20%	6.25%	17.0	35,074€	5.01%
2011						
<50 procedures	7.68%	2.35%	9.39%	20.0	35,984€	8.04%
50-99 procedures	8.02%	2.35%	9.04%	19.3	35,424€	8.28%
>=100 procedures	5.87%	3.01%	9.31%	17.3	35,046€	7.29%
2012						
<50 procedures	6.15%	2.29%	8.44%	18.7	35,294€	7.29%
50-99 procedures	7.07%	2.42%	8.41%	18.9	34,798€	5.48%
>=100 procedures	5.03%	2.10%	6.30%	16.7	34,233€	5.39%
2013						
<50 procedures	5.49%	2.09%	9.28%	20.2	35,808€	6.93%
50-99 procedures	5.85%	2.33%	6.53%	18.2	34,650€	4.56%
>=100 procedures	5.29%	2.70%	5.98%	16.3	34,456€	5.29%
2014						
<50 procedures	5.34%	2.75%	5.99%	19.9	35,993€	6.15%
50-99 procedures	4.58%	2.20%	5.73%	18.3	34,904€	4.32%
>=100 procedures	3.70%	2.28%	4.22%	15.3	34,771€	3.92%

Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year.

Over the seven years of data we analyzed, a slight decreasing trend was visible in the risk-adjusted in-hospital stroke rate, which started out at 2-2,5% in 2008-2009 and ranged from 1,5-2% in 2013-2014 (Figure 1 B). Volume-outcome relationship was actually negative for years following 2010, with higher-volume centers having higher stroke rates.

Risk-adjusted bleeding rates (Figure 1 C), in contrast, showed a clear beneficial effect of higher center procedure volumes for all years but 2011. The magnitude of the effect was distinct from 2008-2010 and decreased in the following years in parallel with an ongoing marked decrease in the general likelihood of bleeding complications, but still was present in 2013/2014.

For risk-adjusted in-hospital ventilation rate (>48h) (Figure 1 D), a pronounced beneficial effect of higher center procedure volumes persisted throughout the observation period. In addition, risk-adjusted in-hospital ventilation rates decreased substantially over the years. As for bleeding, the magnitude of the volume effect was distinct in the first years but steadily declined over the seven year period (annual change of the volume effect: 0.30pp, p=0.041).

Risk-adjusted in-hospital length of stay shows a strong beneficial effect of center procedure volume (Figure 2 A). Unlike the situation found for the endpoints mortality and bleeding, the magnitude of the effect did not decrease much over the observed timeframe. There also is a slight reduction in average length of stay over the years.

Standardized reimbursement data (Figure 2 B) is only available starting in 2010 due to a change in the reimbursement system making previous data difficult to compare. In Germany, reimbursement is based on DRGs which are defined by the patients' diagnoses, gender and age, treatment procedures, complications or comorbidities, and further attributes. Based on this data, a predetermined reimbursement rate per case is calculated. Hospitals receive additional reimbursement for long-stay outlier cases [27]. Furthermore, additional reimbursement is possible for very complex intensive care treatments, which have to be proven by documentation of illness severity and treatment effort during ICU stay [28]. As shown in Figure 2 B, there is a drop in the overall reimbursement level from 2010-2012, but reimbursement stays roughly the same thereafter. In much the same way as found for length of hospital stay, risk-adjusted amount of reimbursement decreased only slightly over time, and showed a large volume effect which did not change over the five year period.

CONCLUSIONS

Our study shows mixed results regarding a volume-outcome relationship in TAVI procedures in German hospitals. First of all, TAVI-related in-hospital mortality decreased substantially between 2008 and 2014 and was lower the higher the procedure volume at the respective hospital is. The magnitude of this volume-outcome relationship, however, declines over the observation period. Especially in later years (2012-2014) differences in mortality between low-, medium-, and high-volume centers are small.

Regarding in-hospital mortality and secondary endpoints, a volume-outcome relationship is eminent in circumstances of relatively unfavorable outcomes (see early years of mortality, bleeding, and ventilation) and decreases as outcomes improve (later years of mortality, bleeding, and ventilation), but is not present in circumstances of constantly low event rates (see stroke). In addition, in most of the cases when we observe a distinct annual decrease, we also observe a decreasing volume effect over time. Presumably, the small centers succeed in participating at the system level learning curve to a degree which allows them to catch up to some degree to the group of high-volume. Unfortunately, our data does not allow addressing the question whether this is due to exchange of expertise or to increasing cumulative experience. The group of small centers may also benefit from there being only a reduced capacity for improvement even in large volume centers some years after the introduction of a new procedure.

Interestingly, decreases in the volume effect over time were not observed for the endpoints of in-hospital length of stay and reimbursement. Presumably, this might be due to the fact that high-volume centers are at a major advantage in streamlining clinical workflows before and after the procedure.

Two recent studies showed volume-outcome relationships for TAVI procedures performed in US hospitals in 2012 [14,15]. In both studies, patients were divided into groups of equal sample size. Disregarding the accompanying problems regarding the external validity of the results [12,13], the results shown in these studies are similar to ours: Among others, inverse volume-outcome relationships were shown for the endpoints death and bleeding [14,15]. One of the two studies also included the endpoints length of stay and hospitalization costs and identified significant differences between the observed hospital volume quartiles (TAVI/year cutoffs <=5, 6-10, 11-20 and >20) [15]. The other study also included the endpoint stroke and did not show significant differences between volume groups (TAVI cutoffs: 20 or 10 cases for different access routes) [14].

As stated before, medical authorities in several countries have issued guidelines calling for minimum numbers of procedures for primary operators performing TAVI [16–19].

In Germany, such mandatory minimums are not yet implemented, but a mandatory number of 50 TAVI procedures annually is officially recommended [19], and this number is also mentioned in guidelines from the UK, Canada and Portugal [16–18]. Our results confirm the existence of a volume-outcome relationship for TAVI procedures between 2008 and 2014, and these effects are in line with existing evidence from TAVI procedures performed in US hospitals [6,14,15]. The above discussed weakening of the volume-outcome relationship over time, however, relativizes the rationale behind mandatory minimum numbers of procedures: The volume-outcome relationship may be considerable in the years following the introduction of a new procedure when there still is a lot of room for improvement (in the two of the cited studies [14,15], i.e. 2012). After a few years, then, the association between procedure numbers and better performance may diminish (see our results regarding the year 2014 and presumably thereafter). In the worst case, the volume effect is already gone by the time mandatory minimums are finally implemented, or the implementation hinders the system to reach optimal health service without restrictions. It should be, however, noted that the average number of TAVI procedures per hospital is larger in Germany compared to most other countries, and that hence the time span until such a point is reached may be longer in other countries.

This might be especially problematic since mandatory minimum quantities on the center level are not free of further disadvantages. They are thought to lead to centralization of procedures in large hospitals, necessitating costly patient transfers and potentially worse aftercare. In addition, it is unclear how an optimal threshold could be set (and adjusted yearly) and by whom, how effects of physician volume and hospital volume should be combined, whether low-volume hospitals and their surgeons perceive the thresholds as new incentives to operate, and how new and innovative hospitals might be able to enter the market [29]. The latter question is especially relevant for TAVI since a recent study showed that between 2010 and 2015 a new center entering the TAVI market needed to perform 54 procedures to achieve clinical outcomes comparable to those reported in high-volume centers [30]. According to the authors of the study, this represents more than 2 years of continuous activity [30].

In addition, the question remains how to integrate the observed volume effects into the existing theory. The “practice-makes-perfect hypothesis” implies a contrary causal relationship than the theory of

"selective-referral" [7,8], and we cannot answer the question whether volume generates quality (practice makes perfect), quality generates volume (selective referral), or both.

Furthermore, Gandjour et al. differentiated the "practice-makes-perfect hypothesis" into learning curve effects, economies of scope, and the concept of a focused factory [31]. Improved outcomes may result from economies of scale: every time doctors perform a procedure, they gain experience. Economies of scope, in contrast, would occur from the simultaneous performance of dissimilar procedures. In the TAVI context, this means that a high-volume center might see improved TAVI outcomes as a result of the performance of high numbers of other procedures. Accordingly, Epstein already raised the question whether similar procedures should also be counted towards a set volume threshold [29]. The focused factory concept, in contrast, assumes that focusing on a small number of procedures could also be favorable [31]. Unfortunately, none of the existing approaches analyzed whether the volume-outcome relationship differs in accordance to the number of other (closely related) procedures conducted in the respective center.

We conclude that the hypothesized volume-outcome relationship for TAVI exists but diminishes and may disappear over time. This might be the case for other interventional procedures, too, which should be taken into account when considering mandatory minimum thresholds.

Reference List

1. Reinöhl, J., Kaier, Klaus, Reinecke, Holger, Schmoor, Claudia, Frankenstein, Lutz, Vach, Werner, et al. Effect of Availability of Transcatheter Aortic Valve Replacement on Clinical Practice. N. Engl. J. Med. 2015;373:2438–47.

2. Gurvitch R, Tay EL, Wijesinghe N, Ye J, Nietlispach F, Wood DA, et al. Transcatheter aortic valve implantation: lessons from the learning curve of the first 270 high-risk patients. Catheter. Cardiovasc. Interv. 2011;78:977–84.

3. Cerillo AG, Murzi M, Glauber M, Berti S. Quality Control and the Learning Curve of Transcatheter Aortic Valve Implantation. JACC Cardiovasc. Interv. 2012;5:456.

4. Alli OO, Booker JD, Lennon RJ, Greason KL, Rihal CS, Holmes DR. Transcatheter Aortic Valve Implantation. JACC Cardiovasc. Interv. 2012;5:72–9.

5. Arai T, Lefèvre T, Hovasse T, Hayashida K, Watanabe Y, O'Connor SA, et al. Evaluation of the learning curve for transcatheter aortic valve implantation via the transfemoral approach. Int. J. Cardiol. 2016;203:491–7.

6. Carroll JD, Vemulapalli S, Dai D, Matsouaka R, Blackstone E, Edwards F, et al. Procedural Experience for Transcatheter Aortic Valve Replacement and Relation to Outcomes. J. Am. Coll. Cardiol. 2017;70:29–41.

7. Gaynor M, Seider H, Vogt WB. The volume-outcome effect, scale economies, and learning-by-doing. Am. Econ. Rev. 2005;95:243–7.

8. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health Serv. Res. 1987;22:157–82.

9. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann. Intern. Med. 2002;137:511–20.

10. Shahian DM, Normand S-LT. The volume-outcome relationship: from Luft to Leapfrog. Ann. Thorac. Surg. 2003;75:1048–58.

11. Barker D, Rosenthal G, Cram P. Simultaneous relationships between procedure volume and mortality: do they bias studies of mortality at specialty hospitals? Health Econ. 2011;20:505–18.

12. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ. The Leapfrog volume criteria may fall short in identifying high-quality surgical centers. Ann. Surg. 2003;238:447–57.

13. Hentschker C, Mennicken R. The Volume-Outcome Relationship and Minimum Volume Standards–Empirical Evidence for Germany. Health Econ. 2015;24:644–58.

14. Kim LK, Minutello RM, Feldman DN, Swaminathan RV, Bergman G, Singh H, et al. Association between transcatheter aortic valve implantation volume and outcomes in the United States. Am. J. Cardiol. 2015;116:1910–5.

15. Badheka AO, Patel NJ, Panaich SS, Patel SV, Jhamnani S, Singh V, et al. Effect of hospital volume on outcomes of transcatheter aortic valve implantation. Am. J. Cardiol. 2015;116:587–94.

16. Webb J, Rodés-Cabau J, Fremes S, Pibarot P, Ruel M, Ibrahim R, et al. Transcatheter Aortic Valve Implantation: A Canadian Cardiovascular Society Position Statement. Can. J. Cardiol. 2012;28:520–8.

17. British Cardiovascular Intervention Society (BCIS) and the Society of Cardiothoracic Surgeons (SCTS). A position statement of the British Cardiovascular Intervention Society (BCIS) and the Society of Cardiothoracic Surgeons (SCTS) [Internet]. 2013 [cited 2015 Nov 12]. Available from: http://www.bcis.org.uk/resources/documents/BCIS%20SCTS%20position%20statement.pdf

18. Campante Teles R, Ribeiro VG, Patrício L, Neves JP, Vouga L, Fragata J, et al. Position statement on transcatheter aortic valve implantation in Portugal. Rev. Port. Cardiol. Engl. Ed. 2013;32:801–5.

19. Kuck KH, Eggebrecht H, Figulla HR, Haude M, Katus H, Möllmann H, et al. Qualitätskriterien zur Durchführung der transvaskulären Aortenklappenimplantation. Kardiologe. 2015;9:11–26.

20. Reinöhl, J., K. Kaier, Reinecke H, Schmoor C, Frankenstein, L., Vach, W., et al. Effect of Availability of Transcatheter Aortic Valve Replacement on Clinical Practice: Supplementary Appendix. N. Engl. J. Med. 2015;373:2438–47.

21. Fedeli U, Alba N, Schievano E, Visentin C, Rosato R, Zorzi M, et al. Diffusion of good practices of care and decline of the association with case volume: the example of breast conserving surgery. BMC Health Serv. Res. 2007;7:167.

22. Nguyen NT, Paya M, Stevens CM, Mavandadi S, Zainabadi K, Wilson SE. The relationship between hospital volume and outcome in bariatric surgery at academic medical centers. Ann. Surg. 2004;240:586–94.

23. Hollenbeak CS, Rogers AM, Barrus B, Wadiwala I, Cooney RN. Surgical volume impacts bariatric surgery mortality: a case for centers of excellence. Surgery. 2008;144:736–43.

24. Farley DE, Ozminkowski RJ. Volume-outcome relationships and inhospital mortality: the effect of changes in volume over time. Med. Care. 1992;30:77–94.

25. Hamilton BH, Hamilton VH. Estimating surgical volume—outcome relationships applying survival models: accounting for frailty and hospital fixed effects. Health Econ. 1997;6:383–95.

26. Harbord RM, Higgins JP. Meta-regression in Stata. Stata J. 2008;8:493–519.

27. Quentin W, Geissler A, Scheller-Kreinsen D, Busse R. DRG-type hospital payment in Germany: The G-DRG system. Euro Obs. 2010;12:4–6.

28. Braun J-P, Bause H, Bloos F, Geldner G, Kastrup M, Kuhlen R, et al. Peer reviewing critical care: a pragmatic approach to quality management. GMS Ger. Med. Sci. 2010;8.

29. Epstein A. Volume and outcome - it is time to move ahead. N. Engl. J. Med. 2002;346:1161–4.

30. Lunardi M, Pesarini G, Zivelonghi C, Piccoli A, Geremia G, Ariotti S, et al. Clinical outcomes of transcatheter aortic valve implantation: from learning curve to proficiency. Open Heart. 2016;3:e000420.

31. Gandjour A, Lauterbach KW. The practice-makes-perfect hypothesis in the context of other production concepts in health care. Am. J. Med. Qual. 2003;18:171–5.

Table 1: Baseline characteristics (2008-2014)

N	43,996
Female	55.87%
Age in years, mean/SD	80.95/6.11
Estimated logistic EuroSCORE[1], mean/SD	22.21%/13.57%
Aortic valve stenosis	68.22%
Combined aortic valve diseases	26.56%
Heart failure	
NYHA II	8.26%
NYHA III or IV	41.66%
Hypertension	62.66%
CAD	46.88%
Previous myocardial infarction	
within 4 months	1.59%
within 1 year	0.75%
after 1 year	4.35%
Previous CABG	12.75%
Previous cardiac surgery	18.06%
Peripheral vascular disease	12.39%
Carotid disease	6.17%
COPD	15.14%
Pulmonary hypertension	22.32%
Renal disease	
GFR <15ml/min	2.95%
GFR <30ml/min	4.90%
Atrial fibrillation	45.93%
Diabetes	33.30%

[1]For calculation of the logistic EuroSCORE, we were able to populate all fields except for critical preoperative state and left ventricular function. In these we assumed an inconspicuous state (i.e. no critical preoperative state and no left ventricular dysfunction) and thus calculated a best-case scenario.

Abbreviations: NYHA – New York Heart Association Functional Classification; CAD – coronary artery disease; CABG – coronary artery bypass graft; COPD – chronic obstructive pulmonary disease; GFR – glomerular filtration rate.

Please note that in comparison to the data published by Reinöhl et al., one TA-TAVI procedure (in 2010) needed to be removed from the dataset due to incomplete information.

Table 2: Number of procedures with regard to the performed TAVI volume of a distinct center in a given year.

TAVI Volume in Center	2008	2009	2010	2011	2012	2013	2014
<50 procedures, n (number of centers)	613 (40)	1,234 (61)	1,155 (51)	1,107 (43)	960 (36)	765 (31)	617 (30)
50-99 procedures, n (number of centers)	236 (3)	658 (10)	1,875 (26)	1,957 (27)	1,569 (20)	1,930 (25)	1,135 (16)
>=100 procedures, n (number of centers)	273 (n/a*)	707 (n/a)	1,776 (3)	3,459 (7)	5,711 (16)	6,452 (9)	9,807 (20)
Total number, n (number of centers)	**1,122 (>=44)**	**2,599 (>=72)**	**4,806 (80)**	**6,523 (77)**	**8,240 (72)**	**9,147 (65)**	**11,559 (66)**

Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year. Number of centers in parentheses.

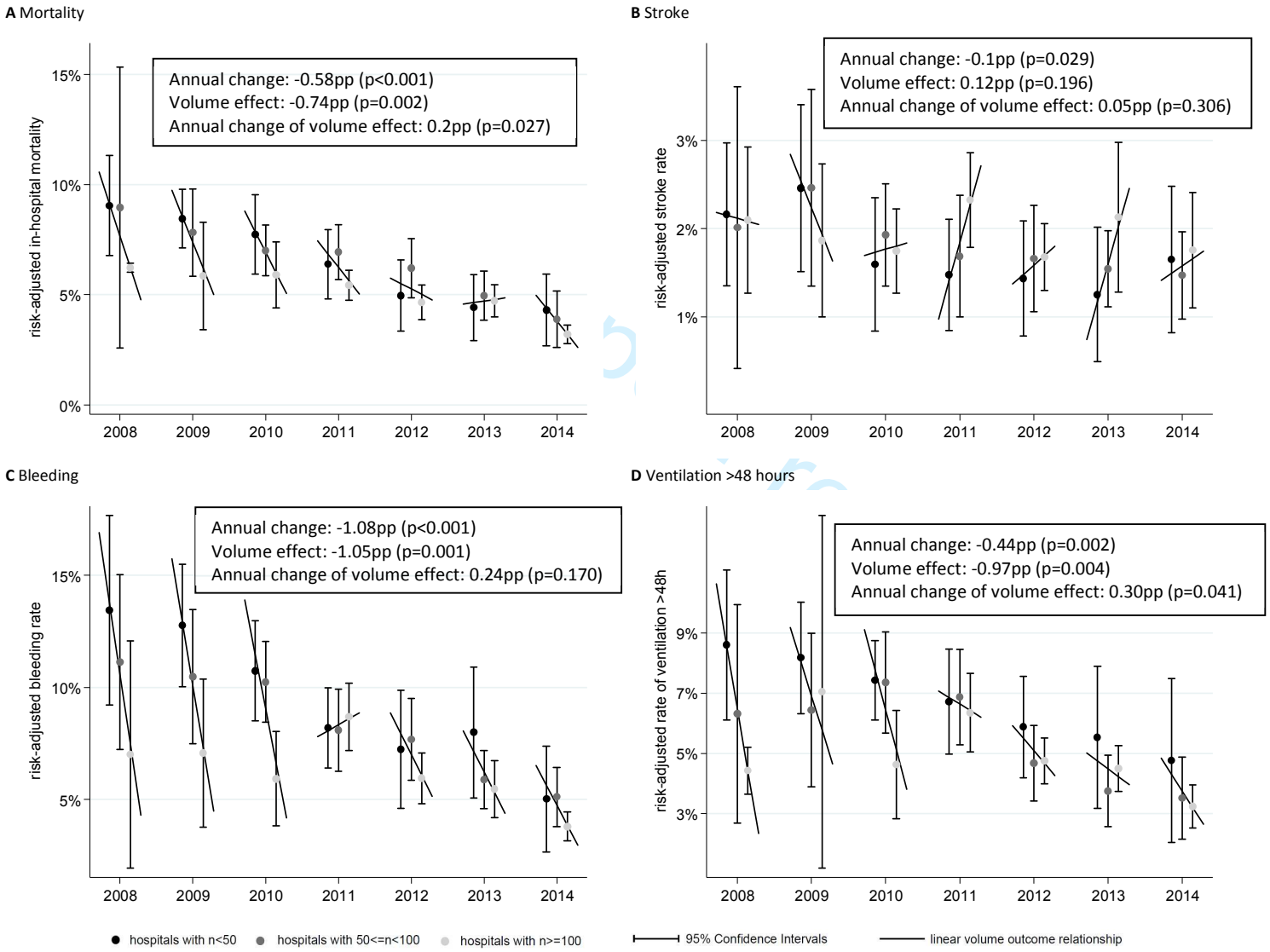
* n/a = not available, exact number censored by DESTATIS due to data protection concerns

Table 3: Unadjusted in-hospital outcomes with regard to the performed TAVI volume of a distinct center in a given year.

	Mortality, %	Stroke, %	Bleeding, %	Length of stay, mean in days	Reim-bursement, mean in €	Proportion of patients with ventilation >48h, %
2008						
<50 procedures	10.11%	3.26%	14.36%	19.2		9.79%
50-99 procedures	9.32%	2.12%	11.44%	21.8		6.78%
>=100 procedures	6.59%	2.56%	7.33%	14.7		4.76%
2009						
<50 procedures	9.81%	3.57%	14.18%	21.6		9.48%
50-99 procedures	8.36%	3.34%	11.25%	18.5		7.14%
>=100 procedures	6.08%	2.12%	7.21%	18.0		7.36%
2010						
<50 procedures	9.00%	2.51%	12.12%	21.0	37,071€	8.74%
50-99 procedures	8.11%	2.56%	11.41%	19.1	36,173€	8.69%
>=100 procedures	6.14%	2.20%	6.25%	17.0	35,074€	5.01%
2011						
<50 procedures	7.68%	2.35%	9.39%	20.0	35,984€	8.04%
50-99 procedures	8.02%	2.35%	9.04%	19.3	35,424€	8.28%
>=100 procedures	5.87%	3.01%	9.31%	17.3	35,046€	7.29%
2012						
<50 procedures	6.15%	2.29%	8.44%	18.7	35,294€	7.29%
50-99 procedures	7.07%	2.42%	8.41%	18.9	34,798€	5.48%
>=100 procedures	5.03%	2.10%	6.30%	16.7	34,233€	5.39%
2013						
<50 procedures	5.49%	2.09%	9.28%	20.2	35,808€	6.93%
50-99 procedures	5.85%	2.33%	6.53%	18.2	34,650€	4.56%
>=100 procedures	5.29%	2.70%	5.98%	16.3	34,456€	5.29%
2014						
<50 procedures	5.34%	2.75%	5.99%	19.9	35,993€	6.15%
50-99 procedures	4.58%	2.20%	5.73%	18.3	34,904€	4.32%
>=100 procedures	3.70%	2.28%	4.22%	15.3	34,771€	3.92%

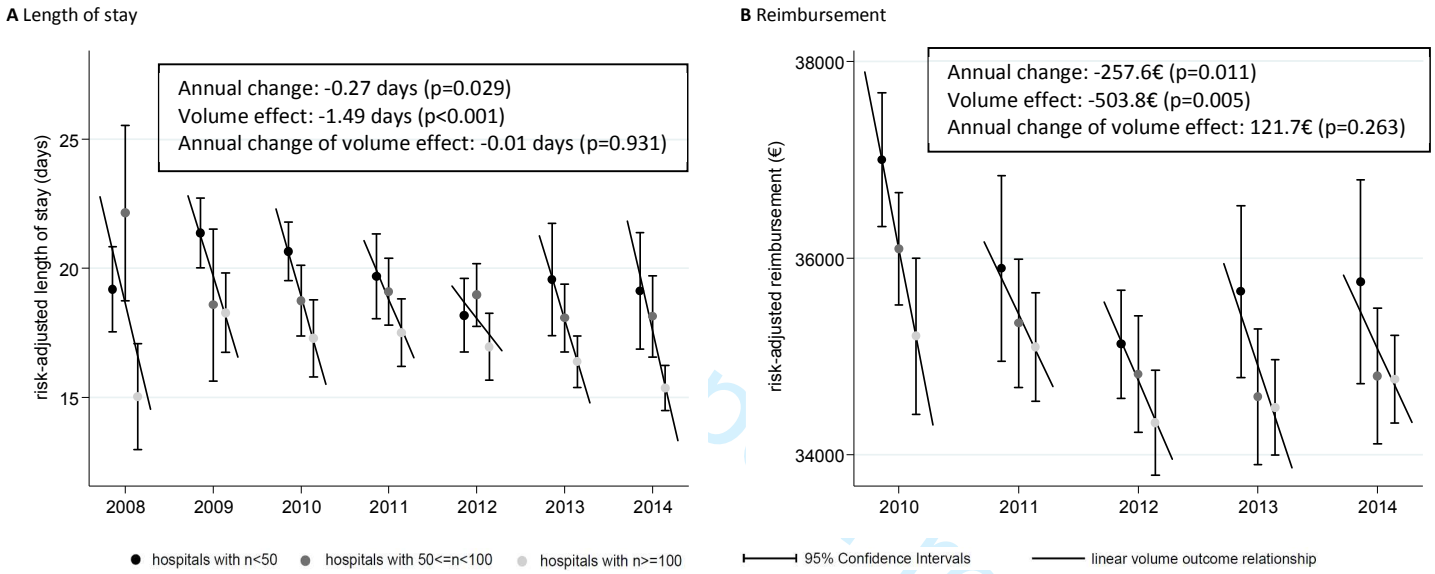
Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year.

Figure 1: Risk-adjusted in-hospital mortality, stroke, bleeding and ventilation rates and their association with center-specific procedure volumes in a given year.



Estimates are based on risk-adjusted logistic regression analysis including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated rates. A separate model with an interaction term was used to assess the change in the volume-outcome relationship. pp= percentage points

Figure 2: Risk-adjusted in-hospital length of stay and reimbursement and their association with center-specific procedure volumes in a given year.



Estimates are based on risk-adjusted linear regression analyses including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated means. A separate a model with an interaction term was used to assess the change in the volume-outcome relationship.

Table S1: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with in-hospital mortality as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	OR	p-value	95% CI
procedure volume <50	1.000		
procedure volume 50-99	0.989	0.98	0.437 - 2.241
procedure volume >=100	0.668	0.004	0.506 - 0.882
year			
2008	1.000		
2009	0.929	0.671	0.661 - 1.305
2010	0.844	0.376	0.579 - 1.229
2011	0.686	0.032	0.486 - 0.968
2012	0.525	0.004	0.340 - 0.812
2013	0.465	0.001	0.302 - 0.718
2014	0.453	0.001	0.288 - 0.712
n_ik_year_50_100#year			
procedure volume 50-99 # 2009	0.929	0.874	0.376 - 2.298
procedure volume 50-99 # 2010	0.909	0.828	0.384 - 2.151
procedure volume 50-99 # 2011	1.105	0.822	0.464 - 2.631
procedure volume 50-99 # 2012	1.280	0.597	0.512 - 3.202
procedure volume 50-99 # 2013	1.141	0.781	0.452 - 2.877
procedure volume 50-99 # 2014	0.908	0.845	0.346 - 2.385
procedure volume >=100 # 2009	1.007	0.981	0.583 - 1.740
procedure volume >=100 # 2010	1.120	0.598	0.735 - 1.705
procedure volume >=100 # 2011	1.260	0.223	0.869 - 1.827
procedure volume >=100 # 2012	1.399	0.164	0.872 - 2.244
procedure volume >=100 # 2013	1.606	0.046	1.009 - 2.558
procedure volume >=100 # 2014	1.100	0.687	0.692 - 1.749
Female	0.902	0.045	0.815 - 0.998
Age in years	1.009	0.155	0.997 - 1.022
Estimated logistic EuroSCORE	1.022	0.000	1.015 - 1.029
Aortic valve stenosis	0.636	0.000	0.504 - 0.802
Combined aortic valve diseases	0.553	0.000	0.447 - 0.685
NYHA II	0.551	0.000	0.423 - 0.717
NYHA III or IV	1.550	0.000	1.264 - 1.900
CAD	1.034	0.517	0.934 - 1.144
Hypertension	0.698	0.000	0.612 - 0.797
Previous MI (within 4 months)	0.683	0.048	0.467 - 0.997
Previous MI (within 1 year)	1.042	0.881	0.608 - 1.785
Previous MI (after 1 year)	0.979	0.816	0.821 - 1.169
Previous CABG	1.017	0.884	0.809 - 1.278
Previous cardiac surgery	0.808	0.117	0.619 - 1.055
Peripheral vascular disease	1.118	0.140	0.964 - 1.295
Carotid disease	0.896	0.165	0.768 - 1.046
COPD	0.979	0.744	0.863 - 1.111
Pulmonary hypertension	0.852	0.021	0.744 - 0.976
GFR <15%	1.770	0.000	1.443 - 2.170
GFR <30%	1.414	0.000	1.167 - 1.714
Atrial fibrillation	1.211	0.000	1.115 - 1.315
Diabetes	1.024	0.640	0.926 - 1.133

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year	Prob.	p-value	95% CI
procedure volume <50 # 2008	0.090	0.000	0.068 - 0.113
procedure volume <50 # 2009	0.085	0.000	0.071 - 0.098
procedure volume <50 # 2010	0.077	0.000	0.059 - 0.095
procedure volume <50 # 2011	0.064	0.000	0.048 - 0.08
procedure volume <50 # 2012	0.050	0.000	0.033 - 0.066
procedure volume <50 # 2013	0.044	0.000	0.029 - 0.059
procedure volume <50 # 2014	0.043	0.000	0.027 - 0.059
procedure volume 50-99 # 2008	0.090	0.006	0.026 - 0.153
procedure volume 50-99 # 2009	0.078	0.000	0.058 - 0.098
procedure volume 50-99 # 2010	0.070	0.000	0.059 - 0.082
procedure volume 50-99 # 2011	0.069	0.000	0.057 - 0.082
procedure volume 50-99 # 2012	0.062	0.000	0.049 - 0.076
procedure volume 50-99 # 2013	0.050	0.000	0.038 - 0.061
procedure volume 50-99 # 2014	0.039	0.000	0.026 - 0.052
procedure volume >=100 # 2008	0.062	0.000	0.06 - 0.064
procedure volume >=100 # 2009	0.058	0.000	0.034 - 0.083
procedure volume >=100 # 2010	0.059	0.000	0.044 - 0.074
procedure volume >=100 # 2011	0.054	0.000	0.047 - 0.061
procedure volume >=100 # 2012	0.047	0.000	0.039 - 0.054
procedure volume >=100 # 2013	0.047	0.000	0.040 - 0.055
procedure volume >=100 # 2014	0.032	0.000	0.028 - 0.036

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

Volume effect	-0.0074	0.002 -	-0.0031
Annual change	-0.0058	0.000 -	-0.0041

Fourth step: A second random effects meta regression model was applied including also an interaction term.

Volume effect	-4.5364	0.026 -	-0.5997
Annual change	-0.0113	0.000 -	-0.0061
Annual change of volume effect	0.0023	0.027 -	0.0042

BMJ Open

Volume-outcome relationship in transcatheter aortic valve implantations in Germany 2008-2014: A secondary data analysis of electronic health records

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020204.R1
Article Type:	Research
Date Submitted by the Author:	05-Mar-2018
Complete List of Authors:	Kaier, Klaus; Medical Center-University of Freiburg, Institute for Medical Biometry and Statistics Oettinger, Vera; Heart Center Freiburg University, Department of Cardiology and Angiology I Reinecke, H; University of Muenster, Department of Cardiology and Angiology, Adult Congenital and Valvular Heart Disease Center Muenster Schmoor, Claudia; Medical Centre, University of Freiburg Frankenstein, L.; University of Heidelberg, Dpt. of Cardiology, Angiology, Pulmology Vach, Werner; Institute of Medical Biometry and Medical Informatics, Hehn, Philip; Faculty of Medicine and Medical Center – University of Freiburg von zur Mühlen, Constantin; Heart Center Freiburg University, Department of Cardiology and Angiology I Bode, Christoph; Heart Center Freiburg University, Department of Cardiology and Angiology I Zehender, Manfred; Heart Center Freiburg University, Department of Cardiology and Angiology I Reinöhl, Jochen; Heart Center Freiburg University, Department of Cardiology and Angiology I
Primary Subject Heading:	Cardiovascular medicine
Secondary Subject Heading:	Health economics
Keywords:	Valvular heart disease < CARDIOLOGY, HEALTH ECONOMICS, CARDIOLOGY, Adult cardiology < CARDIOLOGY

Volume-outcome relationship in transcatheter aortic valve implantations in Germany 2008-2014: A secondary data analysis of electronic health records

Klaus Kaier, PhD[1,2]; Vera Oettinger[2]; Holger Reinecke, MD[3]; Claudia Schmoor, PhD[4]; Lutz Frankenstein, MD[5]; Werner Vach, PhD[1,6]; Philip Hehn[1]; Constantin von zur Mühlen, MD[2]; Christoph Bode, MD[2]; Manfred Zehender, MD, PhD[2]; Jochen Reinöhl, MD[2]

[1] Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center – University of Freiburg, Germany

[2] Department of Cardiology and Angiology I, Heart Center Freiburg University, Germany

[3] Division of Vascular Medicine, Department of Cardiovascular Medicine, University Hospital Muenster, Germany

[4] Clinical Trials Unit, Medical Center – University of Freiburg, Germany

[5] Department of Cardiology, Angiology, Pulmonology, University of Heidelberg, Germany

[6] Department of Orthopaedics and Traumatology, University Hospital Basel, Switzerland

Corresponding author:

Klaus Kaier, PhD

Medical Center – University of Freiburg, Institute of Medical Biometry and Statistics

Stefan-Meier-Str. 26

79104 Freiburg, Germany

phone: + 49 (0)761 203-6807

e-mail: **kaier@imbi.uni-freiburg.de**

Word count excluding title page, abstract, references, figures and tables: 2947

ABSTRACT

Keywords: TAVI, volume-outcome; minimum volume standards; hospital quality; mortality

Objectives: We examine the volume-outcome relationship in isolated transcatheter aortic valve implantations (TAVI). Our interest was whether the volume-outcome relationship for TAVI exists on the center level, whether it occurs equally for different outcomes, and how it develops over time.

Design: Secondary data analysis of electronic health records. The comprehensive German Federal Bureau of Statistics DRG database was queried for data on all isolated TAVI procedures performed in Germany between 2008 and 2014. Logistic and linear regression analyses were carried out. Risk-adjustment was applied using a predefined set of patient characteristics to account for differences in the risk factor composition of the patient populations between centers and over time. Centers performing TAVI were stratified into groups performing <50, 50-99, and ≥100 procedures per year.

Setting: Germany 2008 - 2014.

Participants: All patients undergoing isolated TAVI in the observation period.

Interventions: none.

Primary and secondary outcome measures: In-hospital mortality, bleeding, stroke, probability of ventilation >48 hours, length of hospital stay, and reimbursement.

Results: Between 2008 and 2014, a total of 43,996 TAVI procedures were performed in 113 different centers in Germany with a total of 2,532 cases of in-hospital mortality. Risk-adjusted in-hospital mortality decreases over the years and is lower the higher the annual procedure volume at the center is. The magnitude of the latter effect declines over the observation period. Our results indicate a ceiling effect in the volume-outcome relationship: The volume-outcome relationship is eminent in circumstances of relatively unfavorable outcomes. Alongside improving outcomes, however, the volume-outcome relationship decreases. Also, a volume-outcome relationship seems to be absent in circumstances of constantly low event rates.

Conclusions: The hypothesized volume-outcome relationship for TAVI exists but diminishes and may disappear over time. This might be the case for other interventional procedures, too, which should be taken into account when considering mandatory minimum thresholds.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Study based on administrative data; coding errors are inevitable. However, about 20% of all cardiovascular diagnosis–related groups are reviewed by independent teams of physicians on behalf of the health insurers.
- Risk-adjustment included a number of parameters whose reliability cannot be fully secured, and we cannot guarantee that all parameters of relevance are included in the model. A major limitation is that the data source does not include information on the type of device used in individual TAVI procedures. In addition, information regarding the experience of surgeons at each centre would be highly relevant for the analysis but is unavailable.
- Hospital volume was classified into three fixed categories (<50, 50-99, >=100), which is in line with thresholds mentioned in official guidelines and previously applied in the literature. Possible effects related to very high volumes, however, might be hidden in the analyzed group of patients treated in hospitals with >=100 cases per year.
- The dataset omits patients with a baseline diagnosis of pure aortic regurgitation, as well as those who underwent TAVI with any other concomitant cardiac procedure. This makes sense from a clinical perspective, but further complicates direct comparisons with other administrative datasets.
- The study provides comprehensive data on everyday TAVI practice in a large industrialized country over a multiyear period.

KEY QUESTIONS

What is already known about this subject?

After their introduction, outcomes of new interventions are subject to a learning curve effect, meaning that outcomes improve over a period of time and then level off. The volume of procedures performed at an institution can influence this process, and is thought to have some effect on patient outcomes even after learning is complete (volume-outcome-hypothesis).

What does this study add?

This study tracks patient outcomes by center procedure volume in all transcatheter aortic valve implantation (TAVI) procedures performed in Germany between the procedure's introduction in 2008 and 2014, providing empirical evidence on shape and extent of the above described effects for this procedure.

How might this impact on clinical practice?

This data is of interest to clinical practitioners, hospital administrators, and policy makers involved in the implementation of new clinical procedures.

COMPETING INTERESTS AND FUNDING

The authors declare no conflicts of interest. There was no external funding for this work.

DATA SHARING STATEMENT

No additional data available.

AUTHORSHIP STATEMENT

KK and JR developed the research question and designed the methodology. VÖ, HR, LF, CvzM, CB, MZ and JR provided the medical knowledge of German TAVI practice informing the study design. KK defined the categories, outcomes and measures and developed and implemented the formal analysis and statistical with support by WV and CS. KK and JR collected the data and evidence. KK, VÖ, and WV interpreted and contextualized the results. KK and PH wrote the initial draft of the article, with JR contributing. All authors participated in the critical revision of the article and provided final approval of the version to be published.

INTRODUCTION

Transcatheter aortic valve implantation (TAVI) is a rapidly evolving technique for therapy of aortic stenosis, with a very early and pronounced utilization in Germany [1]. Previous studies report hospital-specific learning curves with respect to in-hospital outcomes such as procedural success, mortality and clinical complications of varying lengths and magnitudes [2–6]. In general, learning curve effects within and between centers can to some degree be explained by the volume of procedures performed at the center. This relationship can be summed up as the "practice-makes-perfect hypothesis", according to which quality of care either increases with the number of patients as a result of economies of scale, with a competing explanation of "selective-referral", according to which higher-quality hospitals attract greater demand and therefore have a greater volume of patients [7,8].

There are a number of criticisms on empirical analyses on the volume-outcome relationship: Many studies lack appropriate adjustment for differences in the risk factor composition of the patient populations between centers [9,10]. Secondly, most studies focus on in-hospital mortality only [11], which is easy to measure, but it is recommended to include additional quality measurements. Finally, most studies divided patients into groups of equal size for analyzing the volume-outcome relationship, which makes it difficult to make use of such results when justifying specific volume thresholds [6,12–14].

Although the evidence regarding the existence of an inverse relationship between the number of TAVI procedures and related outcomes is limited [15,16], medical authorities in Germany and several other countries have issued guidelines calling for minimum numbers of procedures for primary operators performing TAVI [17–20]. There however remains some question whether, firstly, the volume-outcome relationship outlined above exists on the center level regarding TAVI and, secondly, whether or not it takes place in all outcomes and complications equally, and how an existing volume-outcome relationship might change over the years.

To address these questions, we calculated annual procedure volumes for all German hospitals that performed TAVI procedures between January 2008 and December 2014. In order to account for differences in the patient population between high-, medium-, and low-volume centers and over time, we carried out baseline-adjusted regression analyses for the endpoints in-hospital mortality, bleeding, stroke, probability of ventilation >48 hours, length of hospital stay, and reimbursement.

METHODS

Data

Since 2005, data on all hospitalizations in Germany have been available for scientific use via the Diagnosis Related Groups (DRG) statistics collected by the Research Data Center of the Federal Bureau of Statistics (DESTATIS). These hospitalization data, including diagnoses and procedures, are a valuable source of representative nationwide data on the in-hospital treatment of patients. This database represents a virtually complete collection of all hospitalizations in German hospitals that are reimbursed according to the DRG system. From this database [1], we have extracted data on 43,996 cases of isolated TAVI for our analysis.

Our study did not involve direct access by the investigators to data on individual patients but only access to summary results provided by the Research Data Center. Therefore, approval by an ethics committee and informed consent were determined not to be required, in accordance with German law. All summary results were anonymized by DESTATIS. In practice, this means that any information allowing the drawing of conclusions regarding a single patient or a specific hospital are censored by DESTATIS to guarantee data protection. Especially the use of the anonymous, persistent "institute indicator of hospitals" is highly restricted in order not to publish any information directly attributable to a single hospital.

As described previously [1,21], we were able to use the OPS codes (OPS codes: 5-35a.0 in 2007 and 5-35a.00, 5-35a.01 from 2008) to identify all TAVI procedures performed (and reimbursed) in Germany between 2008 and 2014. Patients with a baseline diagnosis of pure aortic regurgitation (main or secondary diagnosis other than I35.0, I35.2, I06.0, I06.2) and those with concomitant cardiac surgery or percutaneous coronary intervention were not included in this analysis. Although some concomitant procedures might be informative (a cardiac surgery procedure during the same hospital stay as TAVI might likely represent a complication following a TAVI procedure), these cases cannot be consistently identified in our dataset as, in many cases, concomitant procedures might have taken place in another center. A complete list of procedure codes can be found in Table S1, a more detailed discussion of the data source may be found in a previous manuscript [1,21].

Patient and Public Involvement

The development of the research question was guided by the intention to provide hospitals and policymakers with empirical evidence that enables them to structure the infrastructure in such a way as to deliver the best possible outcomes to patients. The selected outcome measures represent the most severe complications to the procedure and are of high significance to patient quality of life after the intervention. There was, however, no direct involvement of patients in the design, the recruitment and conduct of the study, nor will the results be disseminated to study participants as the study was based on anonymized administrative data.

Measures

Regarding the in-hospital complications, bleeding was defined as requiring a transfusion of more than 5 units of red blood cells (RBC). For all other comorbidities and complications the existing anamnestic or acute distinctive codes were used (we have discussed OPS and ICD codes in greater detail previously [21]).

In order to analyze possible effects of the above discussed mandatory minimum quantities, the number of procedures per year and center was categorized (i.e. $n<50$, $50\leq n<100$, $n\geq 100$) on the basis of an anonymous, persistent "institute indicator of hospitals" provided by DESTATIS. These particular thresholds are applied because the minimum number of 50 procedures is often mentioned in official TAVI-guidelines [17–20], and these thresholds are widely applied in the literature [22–24].

The primary outcome was in-hospital mortality. Secondary outcomes include post-procedural complications such as stroke and bleeding events (transfusion of >=5 RBC), as well as reimbursement, length of hospital stay and proportion of patients with ventilation >48h.

Statistical analysis

In a first step, multivariate regression analyses were carried out for the different endpoints. In a previous study, Reinöhl et al. [1] identified 21 baseline patient characteristics to describe risk profiles between procedural groups. For risk adjustment, all of these 21 baseline patient characteristics were included as covariates (all covariates listed in Table 1) in the respective regression analyses. In addition, an interaction term between time (in years) and the above mentioned annual volume categories was

included in the regression analyses in order to investigate the volume-outcome relationship over the years.

Table 1: Baseline characteristics (2008-2014)

N	43,996
Female	55.87%
Age in years, mean/SD	80.95/6.11
Estimated logistic EuroSCORE[1], mean/SD	22.21%/13.57%
Aortic valve stenosis as main diagnosis	68.22%
Combined aortic valve diseases as main diagnosis	26.56%
Heart failure	
NYHA II	8.26%
NYHA III or IV	41.66%
Hypertension	62.66%
CAD	46.88%
Previous myocardial infarction	
within 4 months	1.59%
within 1 year	0.75%
after 1 year	4.35%
Previous CABG	12.75%
Previous cardiac surgery	18.06%
Peripheral vascular disease	12.39%
Carotid disease	6.17%
COPD	15.14%
Pulmonary hypertension	22.32%
Renal disease	
GFR <15ml/min	2.95%
GFR <30ml/min	4.90%
Atrial fibrillation	45.93%
Diabetes	33.30%

[1]For calculation of the logistic EuroSCORE, we were able to populate all fields except for critical preoperative state and left ventricular function. In these we assumed an inconspicuous state (i.e. no critical preoperative state and no left ventricular dysfunction) and thus calculated a best-case scenario.

Abbreviations: NYHA – New York Heart Association Functional Classification; CAD – coronary artery disease; CABG – coronary artery bypass graft; COPD – chronic obstructive pulmonary disease; GFR – glomerular filtration rate.

Please note that in comparison to the data published by Reinöhl et al., one TA-TAVI procedure (in 2010) needed to be removed from the dataset due to incomplete information.

Logistic and linear regression analyses are applied for dichotomous and continuous endpoints, respectively. The question of how to account for patients treated in the same hospital was discussed previously [13,25,26]. As recommended in a previous study that also used data from the German DRG-statistic [13], we used cluster-robust standard errors to account for this dependency.

Risk-adjusted rates and means within each year and hospital volume category were obtained by computing the corresponding predicted probabilities or means, respectively, for an artificial subject with each confounder set to its mean value (prediction at the means, see Table 1 for mean values of all confounders). Thereby, risk-adjusted rates and means are taking two aspects into account: (1) change in the patients risk factors compositions over the years, and (2) differences in the patients risk factors compositions within different hospital volume categories. Risk-adjusted rates and means are therefore interpreted as the 'true' procedure-related outcomes independent of changes in the patient population over the years and differences between low, medium, and high-volume centers. Please note that this implies the assumption that all outcome relevant parameters are used for risk-adjustment. Unfortunately, we cannot guarantee that all parameters of relevance are included in the model. In fact, the administrative dataset lacks relevant clinical information (such as echocardiographic findings or anatomical characteristics).

The visualization of these risk-adjusted rates or means together with their 95% confidence intervals constitutes the main analytical approach in this paper. To assess the statistical significance of the observed volume-outcome relationship, of the time trend and a potential change of the volume-outcome relationship over time, we applied to the estimated rates or means a random effects meta regression (command metareg [27]) with time and volume as continuous covariates. A model with an interaction term was used to assess the change in the volume-outcome relationship. A model without an interaction was used to assess the main effects.

Standardized reimbursement data is only available starting in 2010 due to a change in the reimbursement system making previous data difficult to compare. In Germany, reimbursement is based on DRGs which are defined by the patients' diagnoses, gender and age, treatment procedures, complications or comorbidities, and further attributes. Based on this data, a predetermined reimbursement rate per case is calculated. Hospitals receive additional reimbursement for long-stay outlier cases [28]. Furthermore,

additional reimbursement is possible for very complex intensive care treatments, which have to be proven by documentation of illness severity and treatment effort during ICU stay [29].

All analyses were carried out using Stata 13.1 (StataCorp, College Station, Texas, USA).

RESULTS

Between 2008 and 2014, a total of 43,996 TAVI procedures were performed in 113 different centers in Germany. The total number of TAVI procedures performed per year increased markedly over the observation period, from 1,122 in 2008 to 11,559 in 2014 (see Table 2).

Table 2: Number of procedures with regard to the performed TAVI volume of a distinct center in a given year.

TAVI Volume in Center	2008	2009	2010	2011	2012	2013	2014
<50 procedures, n (number of centers)	613 (40)	1,234 (61)	1,155 (51)	1,107 (43)	960 (36)	765 (31)	617 (30)
50-99 procedures, n (number of centers)	236 (3)	658 (10)	1,875 (26)	1,957 (27)	1,569 (20)	1,930 (25)	1,135 (16)
>=100 procedures, n (number of centers)	273 (n/a*)	707 (n/a)	1,776 (3)	3,459 (7)	5,711 (16)	6,452 (9)	9,807 (20)
Total number, n (number of centers)	**1,122 (>=44)**	**2,599 (>=72)**	**4,806 (80)**	**6,523 (77)**	**8,240 (72)**	**9,147 (65)**	**11,559 (66)**

Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year. Number of centres in parentheses.

* n/a = not available, exact number censored by DESTATIS due to data protection concerns

As reported previously [1], substantial reductions in in-hospital mortality have been achieved between 2008 and 2013, and we find this trend to continue into 2014. Regarding center-specific procedure volumes of all TAVI procedures, it appears that the differences in unadjusted in-hospital mortality between the procedure volume groups (<50, 50-99, and >=100) steadily decline over the years (see Table 3). Figure 1 A provides risk-adjusted in-hospital mortality rates allowing for comparison despite possible differences in the patient selection process and consequently the risk factor composition between hospitals in the different procedure volume groups and over time (See Table S2 –Table S7 for details of the process used to generate the results shown in Figure 1A). These results indicate that risk-adjusted in-hospital mortality rates (1) steadily decrease over the years (annual change: -0.58 percentage points (pp), p<0.001), are (2) lower the higher the procedure volume at the hospital is (volume effect: -0.74pp, p=0.002), but that (3) this volume effect declines over the seven year observation period (p-value of interaction term: p=0.027; annual change of volume effect: 0.2pp).

Table 3: Unadjusted in-hospital outcomes with regard to the performed TAVI volume of a distinct center in a given year.

	Mortality, %	Stroke, %	Bleeding, %	Length of stay, mean in days	Reim-bursement, mean in €	Proportion of patients with ventilation >48h, %
2008						
<50 procedures	10.11%	3.26%	14.36%	19.2		9.79%
50-99 procedures	9.32%	2.12%	11.44%	21.8		6.78%
>=100 procedures	6.59%	2.56%	7.33%	14.7		4.76%
2009						
<50 procedures	9.81%	3.57%	14.18%	21.6		9.48%
50-99 procedures	8.36%	3.34%	11.25%	18.5		7.14%
>=100 procedures	6.08%	2.12%	7.21%	18.0		7.36%
2010						
<50 procedures	9.00%	2.51%	12.12%	21.0	37,071€	8.74%
50-99 procedures	8.11%	2.56%	11.41%	19.1	36,173€	8.69%
>=100 procedures	6.14%	2.20%	6.25%	17.0	35,074€	5.01%
2011						
<50 procedures	7.68%	2.35%	9.39%	20.0	35,984€	8.04%
50-99 procedures	8.02%	2.35%	9.04%	19.3	35,424€	8.28%
>=100 procedures	5.87%	3.01%	9.31%	17.3	35,046€	7.29%
2012						
<50 procedures	6.15%	2.29%	8.44%	18.7	35,294€	7.29%
50-99 procedures	7.07%	2.42%	8.41%	18.9	34,798€	5.48%
>=100 procedures	5.03%	2.10%	6.30%	16.7	34,233€	5.39%
2013						
<50 procedures	5.49%	2.09%	9.28%	20.2	35,808€	6.93%
50-99 procedures	5.85%	2.33%	6.53%	18.2	34,650€	4.56%
>=100 procedures	5.29%	2.70%	5.98%	16.3	34,456€	5.29%
2014						
<50 procedures	5.34%	2.75%	5.99%	19.9	35,993€	6.15%
50-99 procedures	4.58%	2.20%	5.73%	18.3	34,904€	4.32%
>=100 procedures	3.70%	2.28%	4.22%	15.3	34,771€	3.92%

Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year.

Over the seven years of data we analyzed, a slight decreasing trend was visible in the risk-adjusted in-hospital stroke rate, which started out at 2-2,5% in 2008-2009 and ranged from 1,5-2% in 2013-2014 (Figure 1 B). Volume-outcome relationship was actually negative for years following 2010, with higher-volume centers having higher stroke rates.

Risk-adjusted bleeding rates (Figure 1 C), in contrast, showed a clear beneficial effect of higher center procedure volumes for all years but 2011. The magnitude of the effect was distinct from 2008-2010 and decreased in the following years in parallel with an ongoing marked decrease in the general likelihood of bleeding complications, but still was present in 2013/2014.

For risk-adjusted in-hospital ventilation rate (>48h) (Figure 1 D), a pronounced beneficial effect of higher center procedure volumes persisted throughout the observation period. In addition, risk-adjusted in-hospital ventilation rates decreased substantially over the years. As for bleeding, the magnitude of the volume effect was distinct in the first years but steadily declined over the seven year period (annual change of the volume effect: 0.30pp, p=0.041).

Risk-adjusted in-hospital length of stay shows a strong beneficial effect of center procedure volume (Figure 2 A). Unlike the situation found for the endpoints mortality and bleeding, the magnitude of the effect did not decrease much over the observed timeframe. There also is a slight reduction in average length of stay over the years.

As shown in Figure 2 B, there is a drop in the overall reimbursement level from 2010-2012, but reimbursement stays roughly the same thereafter. In much the same way as found for length of hospital stay, risk-adjusted amount of reimbursement decreased only slightly over time, and showed a large volume effect which did not change over the five year period.

CONCLUSIONS

Our study shows mixed results regarding a volume-outcome relationship in TAVI procedures in German hospitals. First of all, TAVI-related in-hospital mortality decreased substantially between 2008 and 2014 and was lower the higher the procedure volume at the respective hospital is. The magnitude of this volume-outcome relationship, however, declines over the observation period. Especially in later years (2012-2014) differences in mortality between low-, medium-, and high-volume centers are small.

Regarding in-hospital mortality and secondary endpoints, a volume-outcome relationship is eminent in circumstances of relatively unfavorable outcomes (see early years of mortality, bleeding, and ventilation) and decreases as outcomes improve (later years of mortality, bleeding, and ventilation), but is not present in circumstances of constantly low event rates (see stroke). In addition, in most of the cases when we observe a distinct annual decrease, we also observe a decreasing volume effect over time. Presumably, the small centers succeed in participating at the system level learning curve to a degree which allows them to catch up to some degree to the group of high-volume. Unfortunately, our data does not allow addressing the question whether this is due to exchange of expertise or to increasing cumulative experience. The group of small centers may also benefit from there being only a reduced capacity for improvement even in large volume centers some years after the introduction of a new procedure.

Interestingly, decreases in the volume effect over time were not observed for the endpoints of in-hospital length of stay and reimbursement. Presumably, this might be due to the fact that high-volume centers are at a major advantage in streamlining clinical workflows before and after the procedure.

Two recent studies showed volume-outcome relationships for TAVI procedures performed in US hospitals in 2012 [15,16]. In both studies, patients were divided into groups of equal sample size. Disregarding the accompanying problems regarding the external validity of the results [12,13], the results shown in these studies are similar to ours: Among others, inverse volume-outcome relationships were shown for the endpoints death and bleeding [15,16]. One of the two studies also included the endpoints length of stay and hospitalization costs and identified significant differences between the observed hospital volume quartiles (TAVI/year cutoffs <=5, 6-10, 11-20 and >20) [16]. The other study also included the endpoint stroke and did not show significant differences between volume groups (TAVI cutoffs: 20 or 10 cases for different access routes) [15].

As stated before, medical authorities in several countries have issued guidelines calling for minimum numbers of procedures for primary operators performing TAVI [17–20]. In Germany, such mandatory minimums are not yet implemented, but a mandatory number of 50 TAVI procedures annually is officially recommended [20], and this number is also mentioned in guidelines from the UK, Canada and Portugal [17–19]. Our results confirm the existence of a volume-outcome relationship for TAVI procedures between 2008 and 2014, and these effects are in line with existing evidence from TAVI procedures performed in US hospitals [6,15,16]. The above discussed weakening of the volume-outcome relationship

over time, however, relativizes the rationale behind mandatory minimum numbers of procedures: The volume-outcome relationship may be considerable in the years following the introduction of a new procedure when there still is a lot of room for improvement (in the two of the cited studies [15,16], i.e. 2012). After a few years, then, the association between procedure numbers and better performance may diminish (see our results regarding the year 2014 and presumably thereafter). In the worst case, the volume effect is already gone by the time mandatory minimums are finally implemented, or the implementation hinders the system to reach optimal health service without restrictions. It should be, however, noted that the average number of TAVI procedures per hospital is larger in Germany compared to most other countries, and that hence the time span until such a point is reached may be longer in other countries.

This might be especially problematic since mandatory minimum quantities on the center level are not free of further disadvantages. They are thought to lead to centralization of procedures in large hospitals, necessitating costly patient transfers and potentially worse aftercare. In addition, it is unclear how an optimal threshold could be set (and adjusted yearly) and by whom, how effects of physician volume and hospital volume should be combined, whether low-volume hospitals and their surgeons perceive the thresholds as new incentives to operate, and how new and innovative hospitals might be able to enter the market [30]. The latter question is especially relevant for TAVI since a recent study showed that between 2010 and 2015 a new center entering the TAVI market needed to perform 54 procedures to achieve clinical outcomes comparable to those reported in high-volume centers [31]. According to the authors of the study, this represents more than 2 years of continuous activity [31].

In addition, the question remains how to integrate the observed volume effects into the existing theory. The “practice-makes-perfect hypothesis” implies a contrary causal relationship than the theory of “selective-referral” [7,8], and we cannot answer the question whether volume generates quality (practice makes perfect), quality generates volume (selective referral), or both.

Furthermore, Gandjour et al. differentiated the “practice-makes-perfect hypothesis” into learning curve effects, economies of scope, and the concept of a focused factory [32]. Improved outcomes may result from economies of scale: every time doctors perform a procedure, they gain experience. Economies of scope, in contrast, would occur from the simultaneous performance of dissimilar procedures. In the TAVI context, this means that a high-volume center might see improved TAVI outcomes as a result of the

performance of high numbers of other procedures. Accordingly, Epstein already raised the question whether similar procedures should also be counted towards a set volume threshold [30]. The focused factory concept, in contrast, assumes that focusing on a small number of procedures could also be favorable [32]. Unfortunately, none of the existing approaches analyzed whether the volume-outcome relationship differs in accordance to the number of other (closely related) procedures conducted in the respective center.

We conclude that the hypothesized volume-outcome relationship for TAVI exists but diminishes and may disappear over time. This might be the case for other interventional procedures, too, which should be taken into account when considering mandatory minimum thresholds.

Figure Legends

Figure 1: Risk-adjusted in-hospital mortality, stroke, bleeding and ventilation rates and their association with center-specific procedure volumes in a given year.

Estimates are based on risk-adjusted logistic regression analysis including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated rates. A separate model with an interaction term was used to assess the change in the volume-outcome relationship. pp= percentage points.

Figure 2: Risk-adjusted in-hospital length of stay and reimbursement and their association with center-specific procedure volumes in a given year.

Estimates are based on risk-adjusted linear regression analyses including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated means. A separate a model with an interaction term was used to assess the change in the volume-outcome relationship.

Reference List

1. Reinöhl, J., Kaier, Klaus, Reinecke, Holger, Schmoor, Claudia, Frankenstein, Lutz, Vach, Werner, et al. Effect of Availability of Transcatheter Aortic Valve Replacement on Clinical Practice. N Engl J Med. 2015;373:2438–47.

2. Gurvitch R, Tay EL, Wijesinghe N, Ye J, Nietlispach F, Wood DA, et al. Transcatheter aortic valve implantation: lessons from the learning curve of the first 270 high-risk patients. Catheter Cardiovasc Interv. 2011;78:977–84.

3. Cerillo AG, Murzi M, Glauber M, Berti S. Quality Control and the Learning Curve of Transcatheter Aortic Valve Implantation. JACC Cardiovasc Interv. 2012;5:456.

4. Alli OO, Booker JD, Lennon RJ, Greason KL, Rihal CS, Holmes DR. Transcatheter Aortic Valve Implantation. JACC Cardiovasc Interv. 2012;5:72–9.

5. Arai T, Lefèvre T, Hovasse T, Hayashida K, Watanabe Y, O'Connor SA, et al. Evaluation of the learning curve for transcatheter aortic valve implantation via the transfemoral approach. Int J Cardiol. 2016;203:491–7.

6. Carroll JD, Vemulapalli S, Dai D, Matsouaka R, Blackstone E, Edwards F, et al. Procedural Experience for Transcatheter Aortic Valve Replacement and Relation to Outcomes. J Am Coll Cardiol. 2017;70:29–41.

7. Gaynor M, Seider H, Vogt WB. The volume-outcome effect, scale economies, and learning-by-doing. Am Econ Rev. 2005;95:243–7.

8. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health Serv Res. 1987;22:157–82.

9. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med. 2002;137:511–20.

10. Shahian DM, Normand S-LT. The volume-outcome relationship: from Luft to Leapfrog. Ann Thorac Surg. 2003;75:1048–58.

11. Barker D, Rosenthal G, Cram P. Simultaneous relationships between procedure volume and mortality: do they bias studies of mortality at specialty hospitals? Health Econ. 2011;20:505–18.

12. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ. The Leapfrog volume criteria may fall short in identifying high-quality surgical centers. Ann Surg. 2003;238:447–57.

13. Hentschker C, Mennicken R. The Volume-Outcome Relationship and Minimum Volume Standards–Empirical Evidence for Germany. Health Econ. 2015;24:644–58.

14. Reames BN, Ghaferi AA, Birkmeyer JD, Dimick JB. Hospital volume and operative mortality in the modern era. Ann Surg. 2014;260:244.

15. Kim LK, Minutello RM, Feldman DN, Swaminathan RV, Bergman G, Singh H, et al. Association between transcatheter aortic valve implantation volume and outcomes in the United States. Am J Cardiol. 2015;116:1910–5.

16. Badheka AO, Patel NJ, Panaich SS, Patel SV, Jhamnani S, Singh V, et al. Effect of hospital volume on outcomes of transcatheter aortic valve implantation. Am J Cardiol. 2015;116:587–94.

17. Webb J, Rodés-Cabau J, Fremes S, Pibarot P, Ruel M, Ibrahim R, et al. Transcatheter Aortic Valve Implantation: A Canadian Cardiovascular Society Position Statement. Can J Cardiol. 2012;28:520–8.

18. British Cardiovascular Intervention Society (BCIS) and the Society of Cardiothoracic Surgeons (SCTS). A position statement of the British Cardiovascular Intervention Society (BCIS) and the Society of Cardiothoracic Surgeons (SCTS) [Internet]. 2013 [cited 2015 Nov 12]. Available from: http://www.bcis.org.uk/resources/documents/BCIS%20SCTS%20position%20statement.pdf

19. Campante Teles R, Ribeiro VG, Patrício L, Neves JP, Vouga L, Fragata J, et al. Position statement on transcatheter aortic valve implantation in Portugal. Rev Port Cardiol Engl Ed. 2013;32:801–5.

20. Kuck KH, Eggebrecht H, Figulla HR, Haude M, Katus H, Möllmann H, et al. Qualitätskriterien zur Durchführung der transvaskulären Aortenklappenimplantation. Kardiologe. 2015;9:11–26.

21. Reinöhl, J., K. Kaier, Reinecke H, Schmoor C, Frankenstein, L., Vach, W., et al. Effect of Availability of Transcatheter Aortic Valve Replacement on Clinical Practice: Supplementary Appendix. N Engl J Med. 2015;373:2438–47.

22. Fedeli U, Alba N, Schievano E, Visentin C, Rosato R, Zorzi M, et al. Diffusion of good practices of care and decline of the association with case volume: the example of breast conserving surgery. BMC Health Serv Res. 2007;7:167.

23. Nguyen NT, Paya M, Stevens CM, Mavandadi S, Zainabadi K, Wilson SE. The relationship between hospital volume and outcome in bariatric surgery at academic medical centers. Ann Surg. 2004;240:586–94.

24. Hollenbeak CS, Rogers AM, Barrus B, Wadiwala I, Cooney RN. Surgical volume impacts bariatric surgery mortality: a case for centers of excellence. Surgery. 2008;144:736–43.

25. Farley DE, Ozminkowski RJ. Volume-outcome relationships and inhospital mortality: the effect of changes in volume over time. Med Care. 1992;30:77–94.

26. Hamilton BH, Hamilton VH. Estimating surgical volume—outcome relationships applying survival models: accounting for frailty and hospital fixed effects. Health Econ. 1997;6:383–95.

27. Harbord RM, Higgins JP. Meta-regression in Stata. Stata J. 2008;8:493–519.

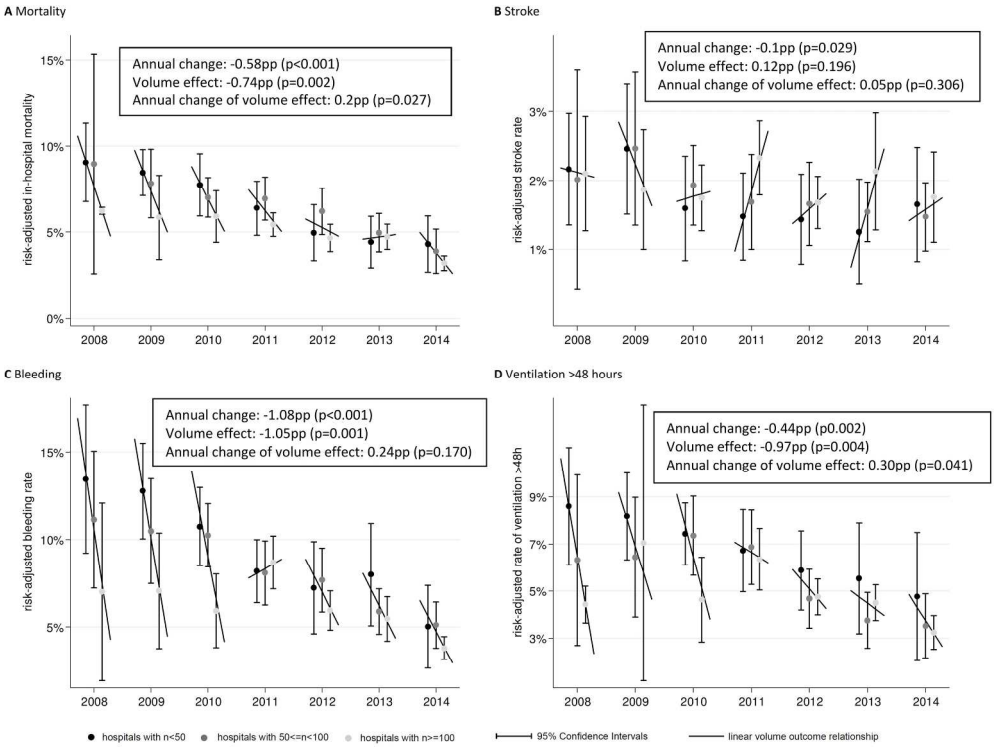
28. Quentin W, Geissler A, Scheller-Kreinsen D, Busse R. DRG-type hospital payment in Germany: The G-DRG system. Euro Obs. 2010;12:4–6.

29. Braun J-P, Bause H, Bloos F, Geldner G, Kastrup M, Kuhlen R, et al. Peer reviewing critical care: a pragmatic approach to quality management. GMS Ger Med Sci. 2010;8.

30. Epstein A. Volume and outcome - it is time to move ahead. N Engl J Med. 2002;346:1161–4.

31. Lunardi M, Pesarini G, Zivelonghi C, Piccoli A, Geremia G, Ariotti S, et al. Clinical outcomes of transcatheter aortic valve implantation: from learning curve to proficiency. Open Heart. 2016;3:e000420.

32. Gandjour A, Lauterbach KW. The practice-makes-perfect hypothesis in the context of other production concepts in health care. Am J Med Qual. 2003;18:171–5.



Caption : Figure 1: Risk-adjusted in-hospital mortality, stroke, bleeding and ventilation rates and their association with center-specific procedure volumes in a given year.‼ ┼ Estimates are based on risk-adjusted logistic regression analysis including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated rates. A separate model with an interaction term was used to assess the change in the volume-outcome relationship. pp= percentage points.

111x83mm (600 x 600 DPI)

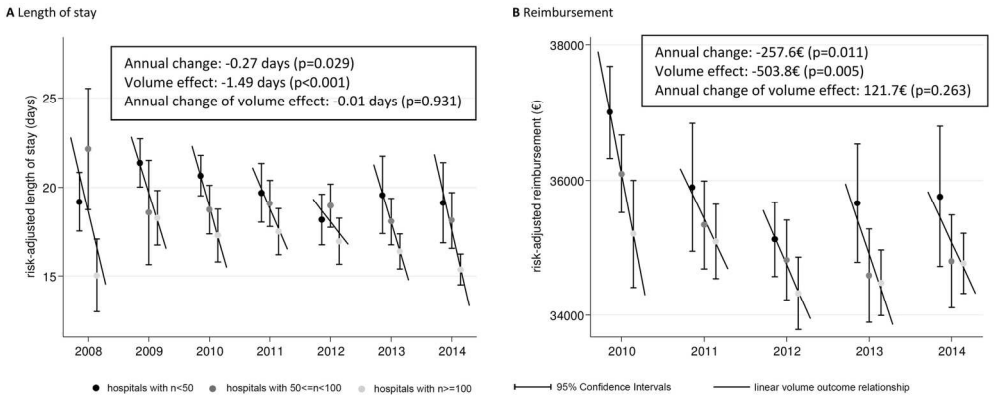


Figure 2: Risk-adjusted in-hospital length of stay and reimbursement and their association with center-specific procedure volumes in a given year.
Estimates are based on risk-adjusted linear regression analyses including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated means. A separate a model with an interaction term was used to assess the change in the volume-outcome relationship.

81x32mm (600 x 600 DPI)

Table S1: Diagnosis and procedure codes used for this analysis

OPS codes	
5-351.0*	Surgical aortic valve replacement
5-35a.0*	Transcatheter aortic valve replacement
5-361.*, 5-362.*, 5-363.*,	Coronary artery bypass graft
5-351.1*, 5-351.2*, 5-353.1, 5-353.2	Surgical mitral valve replacement/reconstruction
5-351.4*	Surgical tricuspid valve replacement
5-377.0 et seqq.	Permanent pacemaker implantation
8-800.7* since 2010: 8-800.c*	Transfusion of RBC

Diagnosis	
I35.0, I06.0	Aortic valve stenosis (degenerative/rheumatic)
I35.2, I06.2	Combined aortic valve diseases (degenerative/rheumatic)
I50.1*	Left ventricular congestive heart failure *(according to NYHA classes)*
I10*	Arterial Hypertension
I25.11, I25.12, I25.13	Coronary artery disease
I25.20, I25.21, I25.22	Previous myocardial infarction *(within 4 months/1 year/after 1 year)*
Z95.1	Previous coronary artery bypass graft
Z95.1 – Z95.4	Previous cardiac surgery
I70.20-I70.25, I70.8, I70.9, I73.9	Peripheral vascular disease
I65.2	Carotid disease
I21*	Acute myocardial infarction *(within the last 28 days)*
J44*	Chronic obstructive pulmonary disease
I27*	Pulmonary hypertension
N18*	Renal disease
N17*	Acute kidney injury
I48.1*	Atrial fibrillation
E10* - E14*	Diabetes
I63*, I64	Stroke or cerebral infarction incl. occlusion and stenosis of cerebral and precerebral arteries, resulting in cerebral infarction

Table S2: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with in-hospital mortality as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	OR	p-value	95% CI
procedure volume <50	1		
procedure volume 50-99	0.989	0.98	0.437 - 2.241
procedure volume >=100	0.668	0.004	0.506 - 0.882
year			
2008	1		
2009	0.929	0.671	0.661 - 1.305
2010	0.844	0.376	0.579 - 1.229
2011	0.686	0.032	0.486 - 0.968
2012	0.525	0.004	0.340 - 0.812
2013	0.465	0.001	0.302 - 0.718
2014	0.453	0.001	0.288 - 0.712
n_ik_year_50_100#year			
procedure volume 50-99 # 2009	0.929	0.874	0.376 - 2.298
procedure volume 50-99 # 2010	0.909	0.828	0.384 - 2.151
procedure volume 50-99 # 2011	1.105	0.822	0.464 - 2.631
procedure volume 50-99 # 2012	1.280	0.597	0.512 - 3.202
procedure volume 50-99 # 2013	1.141	0.781	0.452 - 2.877
procedure volume 50-99 # 2014	0.908	0.845	0.346 - 2.385
procedure volume >=100 # 2009	1.007	0.981	0.583 - 1.740
procedure volume >=100 # 2010	1.120	0.598	0.735 - 1.705
procedure volume >=100 # 2011	1.260	0.223	0.869 - 1.827
procedure volume >=100 # 2012	1.399	0.164	0.872 - 2.244
procedure volume >=100 # 2013	1.606	0.046	1.009 - 2.558
procedure volume >=100 # 2014	1.100	0.687	0.692 - 1.749
Female	0.902	0.045	0.815 - 0.998
Age in years	1.009	0.155	0.997 - 1.022
Estimated logistic EuroSCORE	1.022	0.000	1.015 - 1.029
Aortic valve stenosis	0.636	0.000	0.504 - 0.802
Combined aortic valve diseases	0.553	0.000	0.447 - 0.685
NYHA II	0.551	0.000	0.423 - 0.717
NYHA III or IV	1.550	0.000	1.264 - 1.900
CAD	1.034	0.517	0.934 - 1.144
Hypertension	0.698	0.000	0.612 - 0.797
Previous MI (within 4 months)	0.683	0.048	0.467 - 0.997
Previous MI (within 1 year)	1.042	0.881	0.608 - 1.785
Previous MI (after 1 year)	0.979	0.816	0.821 - 1.169
Previous CABG	1.017	0.884	0.809 - 1.278
Previous cardiac surgery	0.808	0.117	0.619 - 1.055
Peripheral vascular disease	1.118	0.140	0.964 - 1.295
Carotid disease	0.896	0.165	0.768 - 1.046
COPD	0.979	0.744	0.863 - 1.111
Pulmonary hypertension	0.852	0.021	0.744 - 0.976
GFR <15%	1.770	0.000	1.443 - 2.170
GFR <30%	1.414	0.000	1.167 - 1.714
Atrial fibrillation	1.211	0.000	1.115 - 1.315
Diabetes	1.024	0.640	0.926 - 1.133

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year	Prob.	p-value	95% CI
procedure volume <50 # 2008	0.090	0.000	0.068 - 0.113
procedure volume <50 # 2009	0.085	0.000	0.071 - 0.098
procedure volume <50 # 2010	0.077	0.000	0.059 - 0.095
procedure volume <50 # 2011	0.064	0.000	0.048 - 0.08
procedure volume <50 # 2012	0.050	0.000	0.033 - 0.066
procedure volume <50 # 2013	0.044	0.000	0.029 - 0.059
procedure volume <50 # 2014	0.043	0.000	0.027 - 0.059
procedure volume 50-99 # 2008	0.090	0.006	0.026 - 0.153
procedure volume 50-99 # 2009	0.078	0.000	0.058 - 0.098
procedure volume 50-99 # 2010	0.070	0.000	0.059 - 0.082
procedure volume 50-99 # 2011	0.069	0.000	0.057 - 0.082
procedure volume 50-99 # 2012	0.062	0.000	0.049 - 0.076
procedure volume 50-99 # 2013	0.050	0.000	0.038 - 0.061
procedure volume 50-99 # 2014	0.039	0.000	0.026 - 0.052
procedure volume >=100 # 2008	0.062	0.000	0.060 - 0.064
procedure volume >=100 # 2009	0.058	0.000	0.034 - 0.083
procedure volume >=100 # 2010	0.059	0.000	0.044 - 0.074
procedure volume >=100 # 2011	0.054	0.000	0.047 - 0.061
procedure volume >=100 # 2012	0.047	0.000	0.039 - 0.054
procedure volume >=100 # 2013	0.047	0.000	0.040 - 0.055
procedure volume >=100 # 2014	0.032	0.000	0.028 - 0.036

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	-0.007	0.002	-0.012 - -0.003
Annual change	-0.006	0.000	-0.008 - -0.004

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-4.536	0.026	-8.473 - -0.600
Annual change	-0.011	0.000	-0.016 - -0.006
Annual change of volume effect	0.002	0.027	0.000 - 0.004

Table S3: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with stroke as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100		OR	p-value	95% CI
procedure volume <50		1		
procedure volume 50-99		0.929	0.886	0.339 - 2.5473
procedure volume >=100		0.969	0.945	0.398 - 2.3631
year				
	2008	1		
	2009	1.140	0.643	0.654 - 1.989
	2010	0.732	0.308	0.401 - 1.335
	2011	0.678	0.211	0.369 - 1.246
	2012	0.659	0.197	0.350 - 1.241
	2013	0.574	0.114	0.289 - 1.142
	2014	0.759	0.426	0.385 - 1.496
n_ik_year_50_100#year				
procedure volume 50-99 #	2009	1.079	0.896	0.344 - 3.384
procedure volume 50-99 #	2010	1.310	0.637	0.428 - 4.009
procedure volume 50-99 #	2011	1.233	0.715	0.400 - 3.796
procedure volume 50-99 #	2012	1.247	0.706	0.396 - 3.921
procedure volume 50-99 #	2013	1.332	0.631	0.414 - 4.288
procedure volume 50-99 #	2014	0.957	0.942	0.290 - 3.162
procedure volume >=100 #	2009	0.779	0.649	0.265 - 2.287
procedure volume >=100 #	2010	1.134	0.810	0.407 - 3.159
procedure volume >=100 #	2011	1.640	0.331	0.606 - 4.440
procedure volume >=100 #	2012	1.209	0.712	0.441 - 3.312
procedure volume >=100 #	2013	1.769	0.282	0.626 - 4.995
procedure volume >=100 #	2014	1.099	0.857	0.393 - 3.076
Female		0.648	0.000	0.565 - 0.744
Age in years		0.900	0.000	0.890 - 0.910
Estimated logistic EuroSCORE		68461	0.000	33804 - 138650
Aortic valve stenosis		1.278	0.067	0.983 - 1.663
Combined aortic valve diseases		1.329	0.045	1.006 - 1.755
NYHA II		0.949	0.675	0.744 - 1.211
NYHA III or IV		1.096	0.166	0.962 - 1.249
CAD		1.011	0.875	0.885 - 1.154
Hypertension		0.910	0.148	0.802 - 1.034
Previous MI (within 4 months)		0.278	0.000	0.167 - 0.464
Previous MI (within 1 year)		1.246	0.491	0.667 - 2.328
Previous MI (after 1 year)		1.005	0.975	0.742 - 1.360
Previous CABG		0.857	0.350	0.620 - 1.184
Previous cardiac surgery		0.120	0.000	0.089 - 0.161
Peripheral vascular disease		0.395	0.000	0.325 - 0.480
Carotid disease		0.461	0.000	0.362 - 0.587
COPD		0.355	0.000	0.295 - 0.429
Pulmonary hypertension		0.198	0.000	0.164 - 0.239
GFR <15%		0.266	0.000	0.183 - 0.387
GFR <30%		0.297	0.000	0.226 - 0.391
Atrial fibrillation		1.093	0.165	0.964 - 1.239
Diabetes		1.079	0.260	0.945 - 1.231

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year		Prob.	p-value	95% CI
procedure volume <50 #	2008	0.022	0.000	0.012 - 0.031
procedure volume <50 #	2009	0.025	0.000	0.017 - 0.032
procedure volume <50 #	2010	0.016	0.000	0.010 - 0.022
procedure volume <50 #	2011	0.015	0.000	0.009 - 0.021
procedure volume <50 #	2012	0.014	0.000	0.008 - 0.021
procedure volume <50 #	2013	0.013	0.000	0.006 - 0.019
procedure volume <50 #	2014	0.016	0.000	0.008 - 0.025
procedure volume 50-99 #	2008	0.020	0.026	0.002 - 0.038
procedure volume 50-99 #	2009	0.025	0.000	0.014 - 0.035
procedure volume 50-99 #	2010	0.019	0.000	0.014 - 0.025
procedure volume 50-99 #	2011	0.017	0.000	0.012 - 0.022
procedure volume 50-99 #	2012	0.017	0.000	0.011 - 0.022
procedure volume 50-99 #	2013	0.015	0.000	0.011 - 0.020
procedure volume 50-99 #	2014	0.015	0.000	0.009 - 0.021
procedure volume >=100 #	2008	0.021	0.009	0.005 - 0.037
procedure volume >=100 #	2009	0.019	0.000	0.009 - 0.028
procedure volume >=100 #	2010	0.017	0.000	0.012 - 0.023
procedure volume >=100 #	2011	0.023	0.000	0.019 - 0.028
procedure volume >=100 #	2012	0.017	0.000	0.014 - 0.020
procedure volume >=100 #	2013	0.021	0.000	0.018 - 0.025
procedure volume >=100 #	2014	0.018	0.000	0.015 - 0.020

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	0.001	0.196	-0.001 - 0.003
Annual change	-0.001	0.029	-0.002 - 0.000

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-1.103	0.307	-3.311 - 1.106
Annual change	-0.002	0.086	-0.005 - 0.000
Annual change of volume effect	0.001	0.306	-0.001 - 0.002

Table S4: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with bleeding as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	OR	p-value	95% CI
procedure volume <50	1		
procedure volume 50-99	0.806	0.366	0.506 - 1.286
procedure volume >=100	0.485	0.006	0.291 - 0.811
year			
2008	1		
2009	0.942	0.677	0.711 - 1.248
2010	0.775	0.086	0.579 - 1.037
2011	0.574	0.000	0.422 - 0.781
2012	0.502	0.000	0.363 - 0.696
2013	0.559	0.001	0.399 - 0.785
2014	0.340	0.000	0.226 - 0.511
n_ik_year_50_100#year			
procedure volume 50-99 # 2009	0.992	0.978	0.571 - 1.724
procedure volume 50-99 # 2010	1.176	0.541	0.699 - 1.980
procedure volume 50-99 # 2011	1.224	0.458	0.718 - 2.087
procedure volume 50-99 # 2012	1.321	0.321	0.762 - 2.293
procedure volume 50-99 # 2013	0.894	0.693	0.511 - 1.563
procedure volume 50-99 # 2014	1.265	0.463	0.675 - 2.374
procedure volume >=100 # 2009	1.070	0.828	0.581 - 1.970
procedure volume >=100 # 2010	1.078	0.799	0.605 - 1.920
procedure volume >=100 # 2011	2.198	0.006	1.249 - 3.866
procedure volume >=100 # 2012	1.669	0.080	0.941 - 2.960
procedure volume >=100 # 2013	1.371	0.286	0.768 - 2.447
procedure volume >=100 # 2014	1.540	0.174	0.827 - 2.868
Female	1.096	0.032	1.008 - 1.193
Age in years	0.982	0.000	0.974 - 0.990
Estimated logistic EuroSCORE	5.813	0.000	3.386 - 9.979
Aortic valve stenosis	0.738	0.000	0.639 - 0.852
Combined aortic valve diseases	0.677	0.000	0.580 - 0.790
NYHA II	0.665	0.000	0.562 - 0.786
NYHA III or IV	1.313	0.000	1.216 - 1.418
CAD	1.062	0.137	0.981 - 1.149
Hypertension	0.798	0.000	0.741 - 0.861
Previous MI (within 4 months)	0.866	0.324	0.650 - 1.153
Previous MI (within 1 year)	1.071	0.742	0.711 - 1.614
Previous MI (after 1 year)	0.869	0.137	0.721 - 1.046
Previous CABG	0.530	0.000	0.447 - 0.629
Previous cardiac surgery	1.275	0.005	1.077 - 1.509
Peripheral vascular disease	1.255	0.000	1.118 - 1.409
Carotid disease	1.182	0.022	1.024 - 1.364
COPD	0.998	0.969	0.896 - 1.111
Pulmonary hypertension	0.833	0.002	0.741 - 0.935
GFR <15%	2.045	0.000	1.725 - 2.423
GFR <30%	1.446	0.000	1.240 - 1.685
Atrial fibrillation	1.418	0.000	1.316 - 1.528
Diabetes	0.968	0.418	0.894 - 1.048

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year	Prob.	p-value	95% CI
procedure volume <50 # 2008	0.134	0.000	0.108 - 0.161
procedure volume <50 # 2009	0.128	0.000	0.109 - 0.146
procedure volume <50 # 2010	0.107	0.000	0.090 - 0.125
procedure volume <50 # 2011	0.082	0.000	0.066 - 0.097
procedure volume <50 # 2012	0.072	0.000	0.057 - 0.088
procedure volume <50 # 2013	0.080	0.000	0.062 - 0.098
procedure volume <50 # 2014	0.050	0.000	0.034 - 0.066
procedure volume 50-99 # 2008	0.111	0.000	0.071 - 0.151
procedure volume 50-99 # 2009	0.105	0.000	0.082 - 0.128
procedure volume 50-99 # 2010	0.102	0.000	0.089 - 0.116
procedure volume 50-99 # 2011	0.081	0.000	0.069 - 0.093
procedure volume 50-99 # 2012	0.077	0.000	0.064 - 0.090
procedure volume 50-99 # 2013	0.059	0.000	0.049 - 0.069
procedure volume 50-99 # 2014	0.051	0.000	0.039 - 0.063
procedure volume >=100 # 2008	0.070	0.000	0.040 - 0.100
procedure volume >=100 # 2009	0.071	0.000	0.052 - 0.090
procedure volume >=100 # 2010	0.059	0.000	0.048 - 0.070
procedure volume >=100 # 2011	0.087	0.000	0.078 - 0.096
procedure volume >=100 # 2012	0.059	0.000	0.053 - 0.066
procedure volume >=100 # 2013	0.055	0.000	0.049 - 0.060
procedure volume >=100 # 2014	0.038	0.000	0.034 - 0.042

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	-0.011	0.001	-0.016 - -0.0049
Annual change	-0.011	0.000	-0.013 - -0.0083

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-4.84852	0.169	-11.97 - 2.27482
Annual change	-0.01589	0.001	-0.024 - -0.0076
Annual change of volume effect	0.0024	0.170	-0.001 - 0.00595

Table S5: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Linear regression model on 43,996 TAVI cases with Length of hospital stay as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100		Coeff	p-value	95% CI
procedure volume <50		0		
procedure volume 50-99		2.959	0.001	1.133 - 4.786
procedure volume >=100		-4.148	0.000	-5.574 - -2.721
year				
	2008	0		
	2009	2.179	0.001	0.950 - 3.409
	2010	1.472	0.019	0.247 - 2.697
	2011	0.507	0.422	-0.731 - 1.745
	2012	-1.002	0.100	-2.197 - 0.193
	2013	0.382	0.580	-0.970 - 1.734
	2014	-0.060	0.931	-1.433 - 1.312
n_ik_year_50_100#year				
procedure volume 50-99 #	2009	-5.745	0.000	-7.942 - -3.547
procedure volume 50-99 #	2010	-4.872	0.000	-6.926 - -2.819
procedure volume 50-99 #	2011	-3.555	0.001	-5.622 - -1.488
procedure volume 50-99 #	2012	-2.174	0.037	-4.219 - -0.130
procedure volume 50-99 #	2013	-4.452	0.000	-6.576 - -2.327
procedure volume 50-99 #	2014	-3.947	0.000	-6.110 - -1.783
procedure volume >=100 #	2009	1.066	0.242	-0.721 - 2.853
procedure volume >=100 #	2010	0.788	0.365	-0.915 - 2.490
procedure volume >=100 #	2011	1.972	0.021	0.292 - 3.652
procedure volume >=100 #	2012	2.929	0.000	1.302 - 4.557
procedure volume >=100 #	2013	0.970	0.275	-0.773 - 2.713
procedure volume >=100 #	2014	0.390	0.662	-1.358 - 2.139
Female		-0.483	0.000	-0.725 - -0.241
Age in years		-0.247	0.000	-0.276 - -0.218
Estimated logistic EuroSCORE		25.503	0.000	23.345 - 27.661
Aortic valve stenosis		-6.255	0.000	-6.912 - -5.598
Combined aortic valve diseases		-6.592	0.000	-7.263 - -5.921
NYHA II		0.154	0.365	-0.179 - 0.487
NYHA III or IV		2.597	0.000	2.374 - 2.821
CAD		-0.037	0.740	-0.259 - 0.184
Hypertension		-0.888	0.000	-1.111 - -0.666
Previous MI (within 4 months)		-3.355	0.000	-4.175 - -2.534
Previous MI (within 1 year)		0.015	0.980	-1.172 - 1.203
Previous MI (after 1 year)		-0.303	0.250	-0.819 - 0.213
Previous CABG		-2.938	0.000	-3.596 - -2.280
Previous cardiac surgery		-1.710	0.000	-2.405 - -1.015
Peripheral vascular disease		-0.917	0.000	-1.345 - -0.489
Carotid disease		-1.110	0.000	-1.614 - -0.606
COPD		-0.618	0.001	-0.973 - -0.263
Pulmonary hypertension		-2.226	0.000	-2.624 - -1.827
GFR <15%		1.941	0.000	0.958 - 2.923
GFR <30%		0.725	0.034	0.054 - 1.396
Atrial fibrillation		2.575	0.000	2.365 - 2.785
Diabetes		0.932	0.000	0.702 - 1.161

Second step: Marginal means are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year		Coeff	p-value	95% CI
procedure volume <50 #	2008	19.184	0.000	18.241 - 20.13
procedure volume <50 #	2009	21.364	0.000	20.577 - 22.15
procedure volume <50 #	2010	20.656	0.000	19.874 - 21.44
procedure volume <50 #	2011	19.691	0.000	18.890 - 20.49
procedure volume <50 #	2012	18.182	0.000	17.448 - 18.92
procedure volume <50 #	2013	19.566	0.000	18.599 - 20.53
procedure volume <50 #	2014	19.124	0.000	18.128 - 20.12
procedure volume 50-99 #	2008	22.144	0.000	20.579 - 23.71
procedure volume 50-99 #	2009	18.578	0.000	17.650 - 19.51
procedure volume 50-99 #	2010	18.743	0.000	18.221 - 19.27
procedure volume 50-99 #	2011	19.095	0.000	18.554 - 19.64
procedure volume 50-99 #	2012	18.967	0.000	18.413 - 19.52
procedure volume 50-99 #	2013	18.074	0.000	17.576 - 18.57
procedure volume 50-99 #	2014	18.137	0.000	17.545 - 18.73
procedure volume >=100 #	2008	15.036	0.000	13.964 - 16.11
procedure volume >=100 #	2009	18.281	0.000	17.543 - 19.02
procedure volume >=100 #	2010	17.296	0.000	16.791 - 17.8
procedure volume >=100 #	2011	17.515	0.000	17.137 - 17.89
procedure volume >=100 #	2012	16.964	0.000	16.696 - 17.23
procedure volume >=100 #	2013	16.389	0.000	16.136 - 16.64
procedure volume >=100 #	2014	15.366	0.000	15.179 - 15.55

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated means.

	Coeff	p-value	95% CI
Volume effect	-1.488	0.000	-2.021 - -0.9555
Annual change	-0.269	0.029	-0.507 - -0.0307

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	22.5075	0.935	-554.1 - 599.146
Annual change	-0.23991	0.452	-0.898 - 0.41837
Annual change of volume effect	-0.01193	0.931	-0.299 - 0.2748

Table S6: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Linear regression model on 43,996 TAVI cases with reimbursement as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	Coeff	p-value	95% CI
procedure volume <50	0		
procedure volume 50-99	-905.1	0.024	-1689.5 - -120.7
procedure volume >=100	-1792.9	0.000	-2614.4 - -971.4
year			
2010	0		
2011	-1104.3	0.026	-2078.5 - -130.1
2012	-1872.9	0.000	-2810.5 - -935.4
2013	-1339.7	0.007	-2316.5 - -362.9
2014	-1240.5	0.015	-2235.4 - -245.5
n_ik_year_50_100#year			
procedure volume 50-99 # 2011	349.5	0.545	-783.6 - 1482.5
procedure volume 50-99 # 2012	598.9	0.269	-464.0 - 1661.8
procedure volume 50-99 # 2013	-162.6	0.775	-1276.2 - 950.9
procedure volume 50-99 # 2014	-51.3	0.928	-1159.3 - 1056.7
procedure volume >=100 # 2011	994.4	0.082	-127.8 - 2116.7
procedure volume >=100 # 2012	994.3	0.063	-55.5 - 2044.2
procedure volume >=100 # 2013	614.5	0.267	-471.3 - 1700.4
procedure volume >=100 # 2014	803.0	0.150	-289.3 - 1895.4
Female	-816.4	0.000	-1022.5 - -610.4
Age in years	-134.0	0.000	-159.5 - -108.6
Estimated logistic EuroSCORE	9498.4	0.000	7514.7 - 11482.1
Aortic valve stenosis	-1480.8	0.000	-2097.4 - -864.3
Combined aortic valve diseases	-1671.1	0.000	-2296.6 - -1045.6
NYHA II	-420.2	0.001	-665.2 - -175.3
NYHA III or IV	686.8	0.000	501.3 - 872.2
CAD	133.6	0.153	-49.7 - 316.8
Hypertension	-427.8	0.000	-614.4 - -241.3
Previous MI (within 4 months)	-1677.0	0.000	-2134.6 - -1219.4
Previous MI (within 1 year)	295.6	0.574	-733.9 - 1325.2
Previous MI (after 1 year)	-483.1	0.009	-843.2 - -123.0
Previous CABG	-1118.8	0.000	-1687.9 - -549.8
Previous cardiac surgery	-574.4	0.061	-1174.8 - 25.9
Peripheral vascular disease	86.9	0.649	-287.8 - 461.7
Carotid disease	-365.0	0.106	-807.3 - 77.3
COPD	-0.6	0.997	-318.6 - 317.4
Pulmonary hypertension	-951.6	0.000	-1302.7 - -600.4
GFR <15%	1849.2	0.000	921.0 - 2777.4
GFR <30%	322.0	0.258	-235.7 - 879.6
Atrial fibrillation	913.0	0.000	741.8 - 1084.1
Diabetes	223.8	0.020	34.7 - 412.9

Second step: Marginal means are calculated by setting each confounder to its mean value (prediction at the means) using Stata's `margins` command with application of the `atmeans` option.

n_ik_year_50_100#year	Coeff	p-value	95% CI
procedure volume <50 # 2010	36999.8	0.000	36302.5 - 37697.1
procedure volume <50 # 2011	35895.5	0.000	35214.3 - 36576.7
procedure volume <50 # 2012	35126.9	0.000	34499.5 - 35754.2
procedure volume <50 # 2013	35660.1	0.000	34979.4 - 36340.7
procedure volume <50 # 2014	35759.3	0.000	35049.7 - 36469.0
procedure volume 50-99 # 2010	36094.7	0.000	35738.1 - 36451.3
procedure volume 50-99 # 2011	35339.9	0.000	34882.3 - 35797.5
procedure volume 50-99 # 2012	34820.7	0.000	34470.7 - 35170.6
procedure volume 50-99 # 2013	34592.3	0.000	34198.2 - 34986.4
procedure volume 50-99 # 2014	34803.0	0.000	34459.0 - 35147.0
procedure volume >=100 # 2010	35206.9	0.000	34775.2 - 35638.7
procedure volume >=100 # 2011	35097.1	0.000	34744.3 - 35449.8
procedure volume >=100 # 2012	34328.3	0.000	34128.7 - 34527.9
procedure volume >=100 # 2013	34481.7	0.000	34285.1 - 34678.4
procedure volume >=100 # 2014	34769.5	0.000	34623.4 - 34915.6

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated means.

	Coeff	p-value	95% CI
Volume effect	-503.8	0.005	-826.4 - -181.2
Annual change	-257.6	0.011	-444.3 - -70.9

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-245427.3	0.262	-702076.0 - 211221.4
Annual change	-518.0	0.050	-1036.3 - 0.2
Annual change of volume effect	121.7	0.263	-105.2 - 348.7

Table S7: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with ventilation as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	OR	p-value	95% CI
procedure volume <50	1		
procedure volume 50-99	0.716	0.262	0.400 - 1.283
procedure volume >=100	0.492	0.026	0.263 - 0.918
year			
2008	1		
2009	0.946	0.745	0.677 - 1.322
2010	0.853	0.362	0.605 - 1.201
2011	0.765	0.135	0.539 - 1.086
2012	0.664	0.029	0.459 - 0.959
2013	0.623	0.018	0.420 - 0.923
2014	0.532	0.004	0.345 - 0.819
n_ik_year_50_100#year			
procedure volume 50-99 # 2009	1.079	0.827	0.544 - 2.141
procedure volume 50-99 # 2010	1.382	0.322	0.729 - 2.622
procedure volume 50-99 # 2011	1.43	0.277	0.751 - 2.723
procedure volume 50-99 # 2012	1.096	0.789	0.561 - 2.142
procedure volume 50-99 # 2013	0.929	0.833	0.469 - 1.841
procedure volume 50-99 # 2014	1.016	0.967	0.489 - 2.111
procedure volume >=100 # 2009	1.732	0.131	0.848 - 3.537
procedure volume >=100 # 2010	1.232	0.555	0.616 - 2.462
procedure volume >=100 # 2011	1.914	0.059	0.975 - 3.758
procedure volume >=100 # 2012	1.624	0.163	0.822 - 3.210
procedure volume >=100 # 2013	1.636	0.165	0.817 - 3.277
procedure volume >=100 # 2014	1.358	0.402	0.664 - 2.776
Female	0.713	0.000	0.651 - 0.781
Age in years	0.959	0.000	0.951 - 0.968
Estimated logistic EuroSCORE	13.81	0.000	7.797 - 24.464
Aortic valve stenosis	0.722	0.000	0.618 - 0.843
Combined aortic valve diseases	0.663	0.000	0.561 - 0.783
NYHA II	0.498	0.000	0.404 - 0.614
NYHA III or IV	1.485	0.000	1.364 - 1.617
CAD	1.094	0.044	1.002 - 1.193
Hypertension	0.697	0.000	0.642 - 0.757
Previous MI (within 4 months)	0.804	0.163	0.591 - 1.093
Previous MI (within 1 year)	0.796	0.358	0.490 - 1.294
Previous MI (after 1 year)	0.897	0.268	0.740 - 1.087
Previous CABG	0.675	0.000	0.557 - 0.816
Previous cardiac surgery	0.848	0.091	0.701 - 1.026
Peripheral vascular disease	1.198	0.004	1.060 - 1.353
Carotid disease	0.855	0.061	0.725 - 1.007
COPD	1.211	0.001	1.085 - 1.351
Pulmonary hypertension	0.758	0.000	0.669 - 0.858
GFR <15%	1.364	0.001	1.129 - 1.647
GFR <30%	1.252	0.008	1.059 - 1.479
Atrial fibrillation	1.553	0.000	1.430 - 1.687
Diabetes	1.138	0.003	1.045 - 1.239

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year	Prob.	p-value	95% CI
procedure volume <50 # 2008	0.086	0.000	0.065 - 0.107
procedure volume <50 # 2009	0.082	0.000	0.067 - 0.096
procedure volume <50 # 2010	0.074	0.000	0.060 - 0.089
procedure volume <50 # 2011	0.067	0.000	0.053 - 0.081
procedure volume <50 # 2012	0.059	0.000	0.045 - 0.073
procedure volume <50 # 2013	0.055	0.000	0.040 - 0.070
procedure volume <50 # 2014	0.048	0.000	0.032 - 0.063
procedure volume 50-99 # 2008	0.063	0.000	0.033 - 0.094
procedure volume 50-99 # 2009	0.064	0.000	0.046 - 0.083
procedure volume 50-99 # 2010	0.074	0.000	0.062 - 0.085
procedure volume 50-99 # 2011	0.069	0.000	0.058 - 0.079
procedure volume 50-99 # 2012	0.047	0.000	0.037 - 0.057
procedure volume 50-99 # 2013	0.038	0.000	0.030 - 0.045
procedure volume 50-99 # 2014	0.035	0.000	0.025 - 0.045
procedure volume >=100 # 2008	0.044	0.000	0.020 - 0.068
procedure volume >=100 # 2009	0.070	0.000	0.052 0.089
procedure volume >=100 # 2010	0.046	0.000	0.037 0.056
procedure volume >=100 # 2011	0.063	0.000	0.056 0.071
procedure volume >=100 # 2012	0.047	0.000	0.042 0.053
procedure volume >=100 # 2013	0.045	0.000	0.040 0.050
procedure volume >=100 # 2014	0.032	0.000	0.029 0.036

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	-0.010	0.004	-0.016 - -0.004
Annual change	-0.004	0.002	-0.007 - -0.002

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-6.084	0.040	-11.870 - -0.299
Annual change	-0.011	0.004	-0.018 - -0.004
Annual change of volume effect	0.003	0.041	0.000 - 0.006

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	p1-2	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	R(ecord)1.1: p1 R1.2: p1-2 R1.3: n/a
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	p5		
Objectives	3	State specific objectives, including any prespecified hypotheses	p5		
Methods					
Study Design	4	Present key elements of study design early in the paper	p6-9		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	p6		
Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the	p6	RECORD 6.1: The methods of study population selection (such as codes or	R6.1: p6 R6.2, 6.3: n/a

		sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case		algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	p6-7	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	p6-7
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	p6-7, Supplementary tables		
Bias	9	Describe any efforts to address potential sources of bias	p6-7, 9		
Study size	10	Explain how the study size was	n/a (national cohort)		

		arrived at	(p6-7)		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	p7-10		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	p7-10		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	p6
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-	n/a

				level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	n/a (national cohort, administrative data, no follow-up)	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	n/a (national cohort, administrative data, no follow-up)
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	p8		
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures	p13		
Main results	16	(a) Give unadjusted estimates	p10-14		

		and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a		
Discussion					
Key results	18	Summarise key results with reference to study objectives	p14-15		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	p3, 9-10	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	p3, 9-10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	p 14-17		
Generalisability	21	Discuss the generalisability (external validity) of the study results	p3, 14-17		

Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	p4		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Access to public dataset: p6 Further data: Supplemental tables

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

BMJ Open

Volume-outcome relationship in transcatheter aortic valve implantations in Germany 2008-2014: A secondary data analysis of electronic health records

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020204.R2
Article Type:	Research
Date Submitted by the Author:	04-Jun-2018
Complete List of Authors:	Kaier, Klaus; Medical Center-University of Freiburg, Institute for Medical Biometry and Statistics Oettinger, Vera; Heart Center Freiburg University, Department of Cardiology and Angiology I Reinecke, H; University of Muenster, Department of Cardiology and Angiology, Adult Congenital and Valvular Heart Disease Center Muenster Schmoor, Claudia; Medical Centre, University of Freiburg Frankenstein, L.; University of Heidelberg, Dpt. of Cardiology, Angiology, Pulmology Vach, Werner; Institute of Medical Biometry and Medical Informatics, Hehn, Philip; Faculty of Medicine and Medical Center – University of Freiburg von zur Mühlen, Constantin; Heart Center Freiburg University, Department of Cardiology and Angiology I Bode, Christoph; Heart Center Freiburg University, Department of Cardiology and Angiology I Zehender, Manfred; Heart Center Freiburg University, Department of Cardiology and Angiology I Reinöhl, Jochen; Heart Center Freiburg University, Department of Cardiology and Angiology I
Primary Subject Heading:	Cardiovascular medicine
Secondary Subject Heading:	Health economics
Keywords:	Valvular heart disease < CARDIOLOGY, HEALTH ECONOMICS, CARDIOLOGY, Adult cardiology < CARDIOLOGY

Volume-outcome relationship in transcatheter aortic valve implantations in Germany 2008-2014: A secondary data analysis of electronic health records

Klaus Kaier, PhD[1,2]; Vera Oettinger[2]; Holger Reinecke, MD[3]; Claudia Schmoor, PhD[4]; Lutz Frankenstein, MD[5]; Werner Vach, PhD[1,6]; Philip Hehn[1]; Constantin von zur Mühlen, MD[2]; Christoph Bode, MD[2]; Manfred Zehender, MD, PhD[2]; Jochen Reinöhl, MD[2]

[1] Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center – University of Freiburg, Germany

[2] Department of Cardiology and Angiology I, Heart Center Freiburg University, Germany

[3] Division of Vascular Medicine, Department of Cardiovascular Medicine, University Hospital Muenster, Germany

[4] Clinical Trials Unit, Medical Center – University of Freiburg, Germany

[5] Department of Cardiology, Angiology, Pulmonology, University of Heidelberg, Germany

[6] Department of Orthopaedics and Traumatology, University Hospital Basel, Switzerland

Corresponding author:

Klaus Kaier, PhD

Medical Center – University of Freiburg, Institute of Medical Biometry and Statistics

Stefan-Meier-Str. 26

79104 Freiburg, Germany

phone: + 49 (0)761 203-6807

e-mail: **kaier@imbi.uni-freiburg.de**

Word count excluding title page, abstract, references, figures and tables: 2947

ABSTRACT

Keywords: TAVI, volume-outcome; minimum volume standards; hospital quality; mortality

Objectives: We examine the volume-outcome relationship in isolated transcatheter aortic valve implantations (TAVI). Our interest was whether the volume-outcome relationship for TAVI exists on the center level, whether it occurs equally for different outcomes, and how it develops over time.

Design: Secondary data analysis of electronic health records. The comprehensive German Federal Bureau of Statistics DRG database was queried for data on all isolated TAVI procedures performed in Germany between 2008 and 2014. Logistic and linear regression analyses were carried out. Risk-adjustment was applied using a predefined set of patient characteristics to account for differences in the risk factor composition of the patient populations between centers and over time. Centers performing TAVI were stratified into groups performing <50, 50-99, and ≥100 procedures per year.

Setting: Germany 2008 - 2014.

Participants: All patients undergoing isolated TAVI in the observation period.

Interventions: none.

Primary and secondary outcome measures: In-hospital mortality, bleeding, stroke, probability of ventilation >48 hours, length of hospital stay, and reimbursement.

Results: Between 2008 and 2014, a total of 43,996 TAVI procedures were performed in 113 different centers in Germany with a total of 2,532 cases of in-hospital mortality. Risk-adjusted in-hospital mortality decreases over the years and is lower the higher the annual procedure volume at the center is. The magnitude of the latter effect declines over the observation period. Our results indicate a ceiling effect in the volume-outcome relationship: The volume-outcome relationship is eminent in circumstances of relatively unfavorable outcomes. Alongside improving outcomes, however, the volume-outcome relationship decreases. Also, a volume-outcome relationship seems to be absent in circumstances of constantly low event rates.

Conclusions: The hypothesized volume-outcome relationship for TAVI exists but diminishes and may disappear over time. This should be taken into account when considering mandatory minimum thresholds.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Study based on administrative data; coding errors are inevitable, however cardiovascular diagnosis–related groups are reviewed by independent physicians on behalf of health insurers.
- Risk-adjustment included a number of parameters whose reliability cannot be fully secured, and we cannot guarantee that all parameters of relevance are included in the model.
- Hospital volume was classified into three fixed categories, which is in line with thresholds from official guidelines and previous literature, but might hide possible effects related to very high volumes.
- The dataset omits baseline diagnoses of pure aortic regurgitation, as well as patients who underwent a concomitant cardiac procedure, which makes sense from a clinical perspective, but complicates comparisons and might cause bias.
- The study provides comprehensive data on everyday TAVI practice in a large industrialized country over a multiyear period.

KEY QUESTIONS

What is already known about this subject?

After their introduction, outcomes of new interventions are subject to a learning curve effect, meaning that outcomes improve over a period of time and then level off. The volume of procedures performed at an institution can influence this process, and is thought to have some effect on patient outcomes even after learning is complete (volume-outcome-hypothesis).

What does this study add?

This study tracks patient outcomes by center procedure volume in all transcatheter aortic valve implantation (TAVI) procedures performed in Germany between the procedure's introduction in 2008 and 2014, providing empirical evidence on shape and extent of the above described effects for this procedure.

How might this impact on clinical practice?

This data is of interest to clinical practitioners, hospital administrators, and policy makers involved in the implementation of new clinical procedures.

COMPETING INTERESTS AND FUNDING

The authors declare no conflicts of interest. There was no external funding for this work.

DATA SHARING STATEMENT

No additional data available.

AUTHORSHIP STATEMENT

KK and JR developed the research question and designed the methodology. VO, HR, LF, CvzM, CB, MZ and JR provided the medical knowledge of German TAVI practice informing the study design. KK defined the categories, outcomes and measures and developed and implemented the formal analysis and statistical with support by WV and CS. KK and JR collected the data and evidence. KK, VO, and WV interpreted and contextualized the results. KK and PH wrote the initial draft of the article, with JR contributing. All authors participated in the critical revision of the article and provided final approval of the version to be published.

INTRODUCTION

Transcatheter aortic valve implantation (TAVI) is a rapidly evolving technique for therapy of aortic stenosis, with a very early and pronounced utilization in Germany [1]. Previous studies report hospital-specific learning curves with respect to in-hospital outcomes such as procedural success, mortality and clinical complications of varying lengths and magnitudes [2–6]. In general, learning curve effects within and between centers can to some degree be explained by the volume of procedures performed at the center. This relationship can be summed up as the "practice-makes-perfect hypothesis", according to which quality of care either increases with the number of patients as a result of economies of scale, with a competing explanation of "selective-referral", according to which higher-quality hospitals attract greater demand and therefore have a greater volume of patients [7,8].

There are a number of criticisms on empirical analyses on the volume-outcome relationship: Many studies lack appropriate adjustment for differences in the risk factor composition of the patient populations between centers [9,10]. Secondly, most studies focus on in-hospital mortality only [11], which is easy to measure, but it is recommended to include additional quality measurements. Finally, most studies divided patients into groups of equal size for analyzing the volume-outcome relationship, which makes it difficult to make use of such results when justifying specific volume thresholds [6,12–14].

Although the evidence regarding the existence of an inverse relationship between the number of TAVI procedures and related outcomes is limited [15,16], medical authorities in Germany and several other countries have issued guidelines calling for minimum numbers of procedures for primary operators performing TAVI [17–20]. There however remains some question whether, firstly, the volume-outcome relationship outlined above exists on the center level regarding TAVI and, secondly, whether or not it takes place in all outcomes and complications equally, and how an existing volume-outcome relationship might change over the years.

To address these questions, we calculated annual procedure volumes for all German hospitals that performed TAVI procedures between January 2008 and December 2014. In order to account for differences in the patient population between high-, medium-, and low-volume centers and over time, we carried out baseline-adjusted regression analyses for the endpoints in-hospital mortality, bleeding, stroke, probability of ventilation >48 hours, length of hospital stay, and reimbursement.

METHODS

Data

Since 2005, data on all hospitalizations in Germany have been available for scientific use via the Diagnosis Related Groups (DRG) statistics collected by the Research Data Center of the Federal Bureau of Statistics (DESTATIS). These hospitalization data, including diagnoses and procedures, are a valuable source of representative nationwide data on the in-hospital treatment of patients. This database represents a virtually complete collection of all hospitalizations in German hospitals that are reimbursed according to the DRG system. From this database [1], we have extracted data on 43,996 cases of isolated TAVI for our analysis.

Our study did not involve direct access by the investigators to data on individual patients but only access to summary results provided by the Research Data Center. Therefore, approval by an ethics committee and informed consent were determined not to be required, in accordance with German law. All summary results were anonymized by DESTATIS. In practice, this means that any information allowing the drawing of conclusions regarding a single patient or a specific hospital are censored by DESTATIS to guarantee data protection. Especially the use of the anonymous, persistent "institute indicator of hospitals" is highly restricted in order not to publish any information directly attributable to a single hospital.

As described previously [1,21], we were able to use the OPS codes (OPS codes: 5-35a.0 in 2007 and 5-35a.00, 5-35a.01 and 5-35a.02 from 2008) to identify all TAVI procedures performed (and reimbursed) in Germany between 2008 and 2014. Patients with a baseline diagnosis of pure aortic regurgitation (main or secondary diagnosis other than I35.0, I35.2, I06.0, I06.2) and those with concomitant cardiac surgery or percutaneous coronary intervention were not included in this analysis. Although some concomitant procedures might be informative (a cardiac surgery procedure during the same hospital stay as TAVI might likely represent a complication following a TAVI procedure), these cases cannot be consistently identified in our dataset as, in many cases, concomitant procedures might have taken place in another center. A complete list of procedure codes can be found in Table S1, a more detailed discussion of the data source may be found in a previous manuscript [1,21].

Patient and Public Involvement

The development of the research question was guided by the intention to provide hospitals and policymakers with empirical evidence that enables them to structure the infrastructure in such a way as to deliver the best possible outcomes to patients. The selected outcome measures represent the most severe complications to the procedure and are of high significance to patient quality of life after the intervention. There was, however, no direct involvement of patients in the design, the recruitment and conduct of the study, nor will the results be disseminated to study participants as the study was based on anonymized administrative data.

Measures

Regarding the in-hospital complications, bleeding was defined as requiring a transfusion of more than 5 units of red blood cells (RBC). For all other comorbidities and complications the existing anamnestic or acute distinctive codes were used (we have discussed OPS and ICD codes in greater detail previously [21]).

In order to analyze possible effects of the above discussed mandatory minimum quantities, the number of procedures per year and center was categorized (i.e. $n<50$, $50\leq n<100$, $n\geq 100$) on the basis of an anonymous, persistent "institute indicator of hospitals" provided by DESTATIS. These particular thresholds are applied because the minimum number of 50 procedures is often mentioned in official TAVI-guidelines [17–20], and these thresholds are widely applied in the literature [22–24].

The primary outcome was in-hospital mortality. Secondary outcomes include post-procedural complications such as stroke and bleeding events (transfusion of >=5 RBC), as well as reimbursement, length of hospital stay and proportion of patients with ventilation >48h.

Statistical analysis

In a first step, multivariate regression analyses were carried out for the different endpoints. In a previous study, Reinöhl et al. [1] identified 21 baseline patient characteristics to describe risk profiles between procedural groups. For risk adjustment, all of these 21 baseline patient characteristics were included as covariates (all covariates listed in Table 1) in the respective regression analyses. In addition, an interaction term between time (in years) and the above mentioned annual volume categories was included in the regression analyses in order to investigate the volume-outcome relationship over the years.

Table 1: Baseline characteristics (2008-2014)

N	43,996
Female	55.87%
Age in years, mean/SD	80.95/6.11
Estimated logistic EuroSCORE[1], mean/SD	22.21%/13.57%
Aortic valve stenosis as main diagnosis	68.22%
Combined aortic valve diseases as main diagnosis	26.56%
Heart failure	
NYHA II	8.26%
NYHA III or IV	41.66%
Hypertension	62.66%
CAD	46.88%
Previous myocardial infarction	
within 4 months	1.59%
within 1 year	0.75%
after 1 year	4.35%
Previous CABG	12.75%
Previous cardiac surgery	18.06%
Peripheral vascular disease	12.39%
Carotid disease	6.17%
COPD	15.14%
Pulmonary hypertension	22.32%
Renal disease	
GFR <15ml/min	2.95%
GFR <30ml/min	4.90%
Atrial fibrillation	45.93%
Diabetes	33.30%

[1]For calculation of the logistic EuroSCORE, we were able to populate all fields except for critical preoperative state and left ventricular function. In these we assumed an inconspicuous state (i.e. no critical preoperative state and no left ventricular dysfunction) and thus calculated a best-case scenario.

Abbreviations: NYHA – New York Heart Association Functional Classification; CAD – coronary artery disease; CABG – coronary artery bypass graft; COPD – chronic obstructive pulmonary disease; GFR – glomerular filtration rate.

Please note that in comparison to the data published by Reinöhl et al., one TA-TAVI procedure (in 2010) needed to be removed from the dataset due to incomplete information.

Logistic and linear regression analyses are applied for dichotomous and continuous endpoints, respectively. The question of how to account for patients treated in the same hospital was discussed previously [13,25,26]. As recommended in a previous study that also used data from the German DRG-statistic [13], we used cluster-robust standard errors to account for this dependency.

Risk-adjusted rates and means within each year and hospital volume category were obtained by computing the corresponding predicted probabilities or means, respectively, for an artificial subject with each confounder set to its mean value (prediction at the means, see Table 1 for mean values of all confounders). Thereby, risk-adjusted rates and means are taking two aspects into account: (1) change in the patients risk factors compositions over the years, and (2) differences in the patients risk factors compositions within different hospital volume categories. Risk-adjusted rates and means are therefore interpreted as the 'true' procedure-related outcomes independent of changes in the patient population over the years and differences between low, medium, and high-volume centers. Please note that this implies the assumption that all outcome relevant parameters are used for risk-adjustment. Unfortunately, we cannot guarantee that all parameters of relevance are included in the model. In fact, the administrative dataset lacks relevant clinical information (such as echocardiographic findings or anatomical characteristics).

The visualization of these risk-adjusted rates or means together with their 95% confidence intervals constitutes the main analytical approach in this paper. To assess the statistical significance of the observed volume-outcome relationship, of the time trend and a potential change of the volume-outcome relationship over time, we applied to the estimated rates or means a random effects meta regression (command metareg [27]) with time and volume as continuous covariates. A model with an interaction term was used to assess the change in the volume-outcome relationship. A model without an interaction was used to assess the main effects.

Standardized reimbursement data is only available starting in 2010 due to a change in the reimbursement system making previous data difficult to compare. In Germany, reimbursement is based on DRGs which are defined by the patients' diagnoses, gender and age, treatment procedures, complications or comorbidities, and further attributes. Based on this data, a predetermined reimbursement rate per case is calculated. Hospitals receive additional reimbursement for long-stay outlier cases [28]. Furthermore,

additional reimbursement is possible for very complex intensive care treatments, which have to be proven by documentation of illness severity and treatment effort during ICU stay [29].

All analyses were carried out using Stata 13.1 (StataCorp, College Station, Texas, USA).

RESULTS

Between 2008 and 2014, a total of 43,996 TAVI procedures were performed in 113 different centers in Germany. The total number of TAVI procedures performed per year increased markedly over the observation period, from 1,122 in 2008 to 11,559 in 2014 (see Table 2).

Table 2: Number of procedures with regard to the performed TAVI volume of a distinct center in a given year.

TAVI Volume in Center	2008	2009	2010	2011	2012	2013	2014
<50 procedures, n (number of centers)	613 (40)	1,234 (61)	1,155 (51)	1,107 (43)	960 (36)	765 (31)	617 (30)
50-99 procedures, n (number of centers)	236 (3)	658 (10)	1,875 (26)	1,957 (27)	1,569 (20)	1,930 (25)	1,135 (16)
>=100 procedures, n (number of centers)	273 (n/a*)	707 (n/a)	1,776 (3)	3,459 (7)	5,711 (16)	6,452 (9)	9,807 (20)
Total number, n (number of centers)	**1,122 (>=44)**	**2,599 (>=72)**	**4,806 (80)**	**6,523 (77)**	**8,240 (72)**	**9,147 (65)**	**11,559 (66)**

Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year. Number of centres in parentheses.

* n/a = not available, exact number censored by DESTATIS due to data protection concerns

As reported previously [1], substantial reductions in in-hospital mortality have been achieved between 2008 and 2013, and we find this trend to continue into 2014. Regarding center-specific procedure volumes of all TAVI procedures, it appears that the differences in unadjusted in-hospital mortality between the procedure volume groups (<50, 50-99, and >=100) steadily decline over the years (see Table 3). Figure 1 A provides risk-adjusted in-hospital mortality rates allowing for comparison despite possible differences in the patient selection process and consequently the risk factor composition between hospitals in the different procedure volume groups and over time (See Table S2 –Table S7 for details of the process used to generate the results shown in Figure 1 A). These results indicate that risk-adjusted in-hospital mortality rates (1) steadily decrease over the years (annual change: -0.58 percentage points (pp), p<0.001), are (2) lower the higher the procedure volume at the hospital is (volume effect: -0.74pp, p=0.002), but that (3) this volume effect declines over the seven year observation period (p-value of interaction term: p=0.027; annual change of volume effect: 0.2pp).

Table 3: Unadjusted in-hospital outcomes with regard to the performed TAVI volume of a distinct center in a given year.

	Mortality, %	Stroke, %	Bleeding, %	Length of stay, mean in days	Reim-bursement, mean in €	Proportion of patients with ventilation >48h, %
2008						
<50 procedures	10.11%	3.26%	14.36%	19.2		9.79%
50-99 procedures	9.32%	2.12%	11.44%	21.8		6.78%
>=100 procedures	6.59%	2.56%	7.33%	14.7		4.76%
2009						
<50 procedures	9.81%	3.57%	14.18%	21.6		9.48%
50-99 procedures	8.36%	3.34%	11.25%	18.5		7.14%
>=100 procedures	6.08%	2.12%	7.21%	18.0		7.36%
2010						
<50 procedures	9.00%	2.51%	12.12%	21.0	37,071€	8.74%
50-99 procedures	8.11%	2.56%	11.41%	19.1	36,173€	8.69%
>=100 procedures	6.14%	2.20%	6.25%	17.0	35,074€	5.01%
2011						
<50 procedures	7.68%	2.35%	9.39%	20.0	35,984€	8.04%
50-99 procedures	8.02%	2.35%	9.04%	19.3	35,424€	8.28%
>=100 procedures	5.87%	3.01%	9.31%	17.3	35,046€	7.29%
2012						
<50 procedures	6.15%	2.29%	8.44%	18.7	35,294€	7.29%
50-99 procedures	7.07%	2.42%	8.41%	18.9	34,798€	5.48%
>=100 procedures	5.03%	2.10%	6.30%	16.7	34,233€	5.39%
2013						
<50 procedures	5.49%	2.09%	9.28%	20.2	35,808€	6.93%
50-99 procedures	5.85%	2.33%	6.53%	18.2	34,650€	4.56%
>=100 procedures	5.29%	2.70%	5.98%	16.3	34,456€	5.29%
2014						
<50 procedures	5.34%	2.75%	5.99%	19.9	35,993€	6.15%
50-99 procedures	4.58%	2.20%	5.73%	18.3	34,904€	4.32%
>=100 procedures	3.70%	2.28%	4.22%	15.3	34,771€	3.92%

Please note that the numbers of procedures performed per year at a given center were not constant over the observation period, so that it is possible for a center to fall into a different volume group in a different year.

Over the seven years of data we analyzed, a slight decreasing trend was visible in the risk-adjusted in-hospital stroke rate, which started out at 2-2,5% in 2008-2009 and ranged from 1,5-2% in 2013-2014 (Figure 1 B). Volume-outcome relationship was actually negative for years following 2010, with higher-volume centers having higher stroke rates.

Risk-adjusted bleeding rates (Figure 1 C), in contrast, showed a clear beneficial effect of higher center procedure volumes for all years but 2011. The magnitude of the effect was distinct from 2008-2010 and decreased in the following years in parallel with an ongoing marked decrease in the general likelihood of bleeding complications, but still was present in 2013/2014.

For risk-adjusted in-hospital ventilation rate (>48h) (Figure 1 D), a pronounced beneficial effect of higher center procedure volumes persisted throughout the observation period. In addition, risk-adjusted in-hospital ventilation rates decreased substantially over the years. As for bleeding, the magnitude of the volume effect was distinct in the first years but steadily declined over the seven year period (annual change of the volume effect: 0.30pp, p=0.041).

Risk-adjusted in-hospital length of stay shows a strong beneficial effect of center procedure volume (Figure 2 A). Unlike the situation found for the endpoints mortality and bleeding, the magnitude of the effect did not decrease much over the observed timeframe. There also is a slight reduction in average length of stay over the years.

As shown in Figure 2 B, there is a drop in the overall reimbursement level from 2010-2012, but reimbursement stays roughly the same thereafter. In much the same way as found for length of hospital stay, risk-adjusted amount of reimbursement decreased only slightly over time, and showed a large volume effect which did not change over the five year period.

CONCLUSIONS

Our study shows mixed results regarding a volume-outcome relationship in TAVI procedures in German hospitals. First of all, TAVI-related in-hospital mortality decreased substantially between 2008 and 2014 and was lower the higher the procedure volume at the respective hospital is. The magnitude of this volume-outcome relationship, however, declines over the observation period. Especially in later years (2012-2014) differences in mortality between low-, medium-, and high-volume centers are small.

Regarding in-hospital mortality and secondary endpoints, a volume-outcome relationship is eminent in circumstances of relatively unfavorable outcomes (see early years of mortality, bleeding, and ventilation) and decreases as outcomes improve (later years of mortality, bleeding, and ventilation), but is not present in circumstances of constantly low event rates (see stroke). In addition, in most of the cases when we observe a distinct annual decrease, we also observe a decreasing volume effect over time. Presumably, the small centers succeed in participating at the system level learning curve to a degree which allows them to catch up to some degree to the group of high-volume. Unfortunately, our data does not allow addressing the question whether this is due to exchange of expertise or to increasing cumulative experience. The group of small centers may also benefit from there being only a reduced capacity for improvement even in large volume centers some years after the introduction of a new procedure.

Interestingly, decreases in the volume effect over time were not observed for the endpoints of in-hospital length of stay and reimbursement. Presumably, this might be due to the fact that high-volume centers are at a major advantage in streamlining clinical workflows before and after the procedure.

Two recent studies showed volume-outcome relationships for TAVI procedures performed in US hospitals in 2012 [15,16]. In both studies, patients were divided into groups of equal sample size. Disregarding the accompanying problems regarding the external validity of the results [12,13], the results shown in these studies are similar to ours: Among others, inverse volume-outcome relationships were shown for the endpoints death and bleeding [15,16]. One of the two studies also included the endpoints length of stay and hospitalization costs and identified significant differences between the observed hospital volume quartiles (TAVI/year cutoffs <=5, 6-10, 11-20 and >20) [16]. The other study also included the endpoint stroke and did not show significant differences between volume groups (TAVI cutoffs: 20 or 10 cases for different access routes) [15].

As stated before, medical authorities in several countries have issued guidelines calling for minimum numbers of procedures for primary operators performing TAVI [17–20]. In Germany, such mandatory minimums are not yet implemented, but a mandatory number of 50 TAVI procedures annually is officially recommended [20], and this number is also mentioned in guidelines from the UK, Canada and Portugal [17–19]. Our results confirm the existence of a volume-outcome relationship for TAVI procedures between 2008 and 2014, and these effects are in line with existing evidence from TAVI procedures performed in US hospitals [6,15,16]. The above discussed weakening of the volume-outcome relationship

over time, however, relativizes the rationale behind mandatory minimum numbers of procedures: The volume-outcome relationship may be considerable in the years following the introduction of a new procedure when there still is a lot of room for improvement (in the two of the cited studies [15,16], i.e. 2012). After a few years, then, the association between procedure numbers and better performance may diminish (see our results regarding the year 2014 and presumably thereafter). In the worst case, the volume effect is already gone by the time mandatory minimums are finally implemented, or the implementation hinders the system to reach optimal health service without restrictions. It should be, however, noted that the average number of TAVI procedures per hospital is larger in Germany compared to most other countries, and that hence the time span until such a point is reached may be longer in other countries.

This might be especially problematic since mandatory minimum quantities on the center level are not free of further disadvantages. They are thought to lead to centralization of procedures in large hospitals, necessitating costly patient transfers and potentially worse aftercare. In addition, it is unclear how an optimal threshold could be set (and adjusted yearly) and by whom, how effects of physician volume and hospital volume should be combined, whether low-volume hospitals and their surgeons perceive the thresholds as new incentives to operate, and how new and innovative hospitals might be able to enter the market [30]. The latter question is especially relevant for TAVI since a recent study showed that between 2010 and 2015 a new center entering the TAVI market needed to perform 54 procedures to achieve clinical outcomes comparable to those reported in high-volume centers [31]. According to the authors of the study, this represents more than 2 years of continuous activity [31].

In addition, the question remains how to integrate the observed volume effects into the existing theory. The “practice-makes-perfect hypothesis” implies a contrary causal relationship than the theory of “selective-referral” [7,8], and we cannot answer the question whether volume generates quality (practice makes perfect), quality generates volume (selective referral), or both.

Furthermore, Gandjour et al. differentiated the “practice-makes-perfect hypothesis” into learning curve effects, economies of scope, and the concept of a focused factory [32]. Improved outcomes may result from economies of scale: every time doctors perform a procedure, they gain experience. Economies of scope, in contrast, would occur from the simultaneous performance of dissimilar procedures. In the TAVI context, this means that a high-volume center might see improved TAVI outcomes as a result of the

performance of high numbers of other procedures. Accordingly, Epstein already raised the question whether similar procedures should also be counted towards a set volume threshold [30]. The focused factory concept, in contrast, assumes that focusing on a small number of procedures could also be favorable [32]. Unfortunately, none of the existing approaches analyzed whether the volume-outcome relationship differs in accordance to the number of other (closely related) procedures conducted in the respective center.

Our study has several strengths and limitations: First of all, it is based on administrative data, and coding errors are inevitable. However, about 20% of all cardiovascular diagnosis–related groups are reviewed by independent teams of physicians on behalf of the health insurers, which should ensure a generally good reliability of the data.

Second, our risk-adjustment included a number of parameters whose reliability cannot be fully secured, and we cannot guarantee that all parameters of relevance are included in the model. A major limitation is that the data source does not include information on the type of device used in individual TAVI procedures. Therefore, information regarding the type of device and access route was not used for risk adjustment. In addition, information regarding the experience of surgeons at each centre would be highly relevant for the analysis but is also unavailable.

Thirdly, in terms of the categories used, hospital volume was classified into three fixed categories (<50, 50-99, >=100), which is in line with thresholds mentioned in official guidelines and previously applied in the literature, but might result in possible effects related to very high volumes being hidden in the analyzed group of patients treated in hospitals with >=100 cases per year.

Lastly, the dataset omits patients with a baseline diagnosis of pure aortic regurgitation, as well as those who underwent TAVI with any other concomitant cardiac procedure. This makes sense from a clinical perspective, but further complicates direct comparisons with other administrative datasets and possibly caused bias in the measurement of hospital volume and outcome.

A major strength of the study is that it provides comprehensive data on everyday TAVI practice in a large industrialized country over a multiyear period.

We conclude that the hypothesized volume-outcome relationship for TAVI exists but diminishes and may disappear over time. This should be taken into account when considering mandatory minimum thresholds.

Figure Legends

Figure 1: Risk-adjusted in-hospital mortality, stroke, bleeding and ventilation rates and their association with center-specific procedure volumes in a given year.

Estimates are based on risk-adjusted logistic regression analysis including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated rates. A separate model with an interaction term was used to assess the change in the volume-outcome relationship. pp= percentage points.

Figure 2: Risk-adjusted in-hospital length of stay and reimbursement and their association with center-specific procedure volumes in a given year.

Estimates are based on risk-adjusted linear regression analyses including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated means. A separate a model with an interaction term was used to assess the change in the volume-outcome relationship.

Reference List

1. Reinöhl, J., Kaier, Klaus, Reinecke, Holger, Schmoor, Claudia, Frankenstein, Lutz, Vach, Werner, et al. Effect of Availability of Transcatheter Aortic Valve Replacement on Clinical Practice. N Engl J Med. 2015;373:2438–47.

2. Gurvitch R, Tay EL, Wijesinghe N, Ye J, Nietlispach F, Wood DA, et al. Transcatheter aortic valve implantation: lessons from the learning curve of the first 270 high-risk patients. Catheter Cardiovasc Interv. 2011;78:977–84.

3. Cerillo AG, Murzi M, Glauber M, Berti S. Quality Control and the Learning Curve of Transcatheter Aortic Valve Implantation. JACC Cardiovasc Interv. 2012;5:456.

4. Alli OO, Booker JD, Lennon RJ, Greason KL, Rihal CS, Holmes DR. Transcatheter Aortic Valve Implantation. JACC Cardiovasc Interv. 2012;5:72–9.

5. Arai T, Lefèvre T, Hovasse T, Hayashida K, Watanabe Y, O'Connor SA, et al. Evaluation of the learning curve for transcatheter aortic valve implantation via the transfemoral approach. Int J Cardiol. 2016;203:491–7.

6. Carroll JD, Vemulapalli S, Dai D, Matsouaka R, Blackstone E, Edwards F, et al. Procedural Experience for Transcatheter Aortic Valve Replacement and Relation to Outcomes. J Am Coll Cardiol. 2017;70:29–41.

7. Gaynor M, Seider H, Vogt WB. The volume-outcome effect, scale economies, and learning-by-doing. Am Econ Rev. 2005;95:243–7.

8. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health Serv Res. 1987;22:157–82.

9. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med. 2002;137:511–20.

10. Shahian DM, Normand S-LT. The volume-outcome relationship: from Luft to Leapfrog. Ann Thorac Surg. 2003;75:1048–58.

11. Barker D, Rosenthal G, Cram P. Simultaneous relationships between procedure volume and mortality: do they bias studies of mortality at specialty hospitals? Health Econ. 2011;20:505–18.

12. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ. The Leapfrog volume criteria may fall short in identifying high-quality surgical centers. Ann Surg. 2003;238:447–57.

13. Hentschker C, Mennicken R. The Volume-Outcome Relationship and Minimum Volume Standards–Empirical Evidence for Germany. Health Econ. 2015;24:644–58.

14. Reames BN, Ghaferi AA, Birkmeyer JD, Dimick JB. Hospital volume and operative mortality in the modern era. Ann Surg. 2014;260:244.

15. Kim LK, Minutello RM, Feldman DN, Swaminathan RV, Bergman G, Singh H, et al. Association between transcatheter aortic valve implantation volume and outcomes in the United States. Am J Cardiol. 2015;116:1910–5.

16. Badheka AO, Patel NJ, Panaich SS, Patel SV, Jhamnani S, Singh V, et al. Effect of hospital volume on outcomes of transcatheter aortic valve implantation. Am J Cardiol. 2015;116:587–94.

17. Webb J, Rodés-Cabau J, Fremes S, Pibarot P, Ruel M, Ibrahim R, et al. Transcatheter Aortic Valve Implantation: A Canadian Cardiovascular Society Position Statement. Can J Cardiol. 2012;28:520–8.

18. British Cardiovascular Intervention Society (BCIS) and the Society of Cardiothoracic Surgeons (SCTS). A position statement of the British Cardiovascular Intervention Society (BCIS) and the Society of Cardiothoracic Surgeons (SCTS) [Internet]. 2013 [cited 2015 Nov 12]. Available from: http://www.bcis.org.uk/resources/documents/BCIS%20SCTS%20position%20statement.pdf

19. Campante Teles R, Ribeiro VG, Patrício L, Neves JP, Vouga L, Fragata J, et al. Position statement on transcatheter aortic valve implantation in Portugal. Rev Port Cardiol Engl Ed. 2013;32:801–5.

20. Kuck KH, Eggebrecht H, Figulla HR, Haude M, Katus H, Möllmann H, et al. Qualitätskriterien zur Durchführung der transvaskulären Aortenklappenimplantation. Kardiologe. 2015;9:11–26.

21. Reinöhl, J., K. Kaier, Reinecke H, Schmoor C, Frankenstein, L., Vach, W., et al. Effect of Availability of Transcatheter Aortic Valve Replacement on Clinical Practice: Supplementary Appendix. N Engl J Med. 2015;373:2438–47.

22. Fedeli U, Alba N, Schievano E, Visentin C, Rosato R, Zorzi M, et al. Diffusion of good practices of care and decline of the association with case volume: the example of breast conserving surgery. BMC Health Serv Res. 2007;7:167.

23. Nguyen NT, Paya M, Stevens CM, Mavandadi S, Zainabadi K, Wilson SE. The relationship between hospital volume and outcome in bariatric surgery at academic medical centers. Ann Surg. 2004;240:586–94.

24. Hollenbeak CS, Rogers AM, Barrus B, Wadiwala I, Cooney RN. Surgical volume impacts bariatric surgery mortality: a case for centers of excellence. Surgery. 2008;144:736–43.

25. Farley DE, Ozminkowski RJ. Volume-outcome relationships and inhospital mortality: the effect of changes in volume over time. Med Care. 1992;30:77–94.

26. Hamilton BH, Hamilton VH. Estimating surgical volume—outcome relationships applying survival models: accounting for frailty and hospital fixed effects. Health Econ. 1997;6:383–95.

27. Harbord RM, Higgins JP. Meta-regression in Stata. Stata J. 2008;8:493–519.

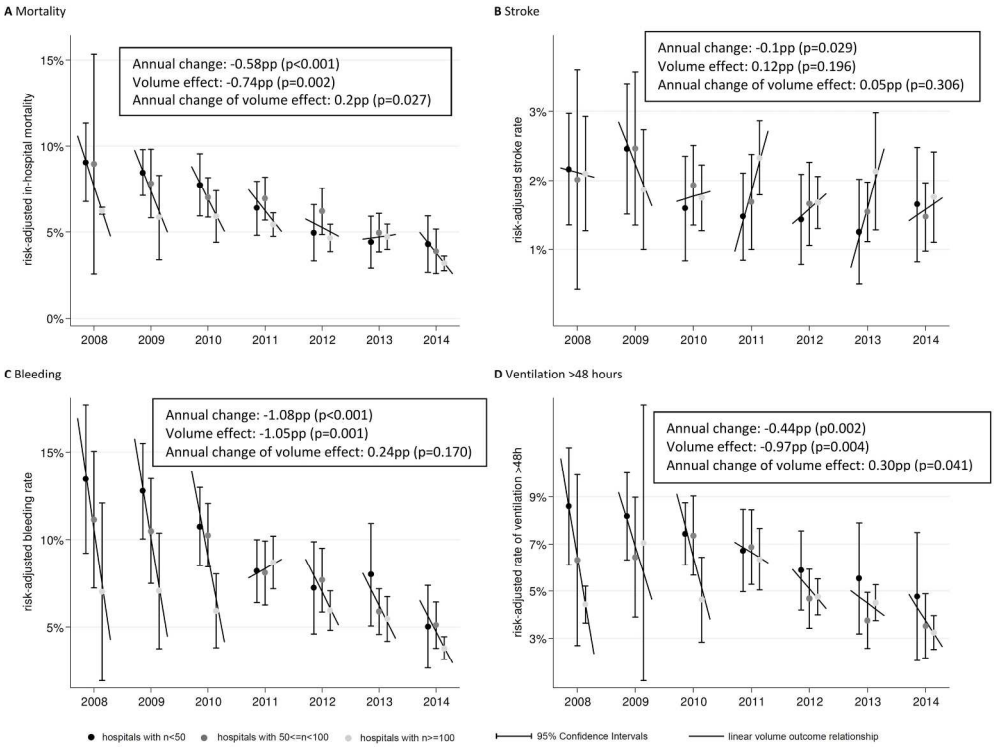
28. Quentin W, Geissler A, Scheller-Kreinsen D, Busse R. DRG-type hospital payment in Germany: The G-DRG system. Euro Obs. 2010;12:4–6.

29. Braun J-P, Bause H, Bloos F, Geldner G, Kastrup M, Kuhlen R, et al. Peer reviewing critical care: a pragmatic approach to quality management. GMS Ger Med Sci. 2010;8.

30. Epstein A. Volume and outcome - it is time to move ahead. N Engl J Med. 2002;346:1161–4.

31. Lunardi M, Pesarini G, Zivelonghi C, Piccoli A, Geremia G, Ariotti S, et al. Clinical outcomes of transcatheter aortic valve implantation: from learning curve to proficiency. Open Heart. 2016;3:e000420.

32. Gandjour A, Lauterbach KW. The practice-makes-perfect hypothesis in the context of other production concepts in health care. Am J Med Qual. 2003;18:171–5.



Caption : Figure 1: Risk-adjusted in-hospital mortality, stroke, bleeding and ventilation rates and their association with center-specific procedure volumes in a given year.‼ ┼ Estimates are based on risk-adjusted logistic regression analysis including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated rates. A separate model with an interaction term was used to assess the change in the volume-outcome relationship. pp= percentage points.

111x83mm (600 x 600 DPI)

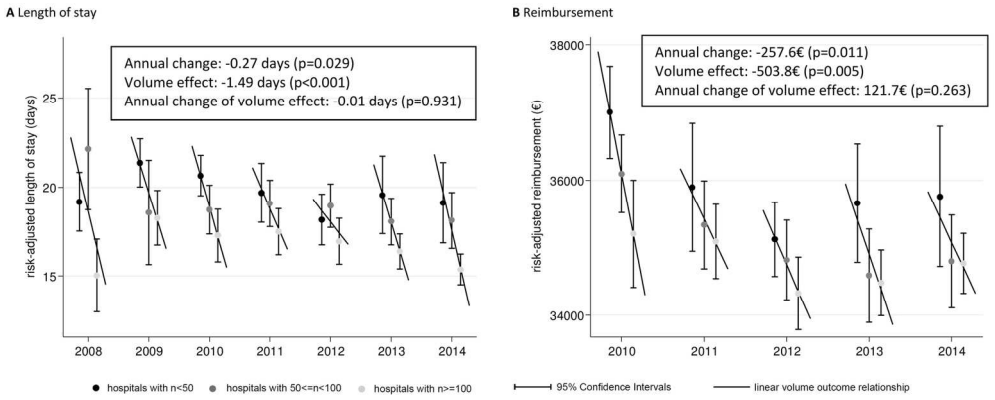


Figure 2: Risk-adjusted in-hospital length of stay and reimbursement and their association with center-specific procedure volumes in a given year.
Estimates are based on risk-adjusted linear regression analyses including all available patient characteristics as confounders (see Table 1). Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means, see Table 1 for means). Annual and volume effects were calculated using random effects meta regression based on the estimated means. A separate a model with an interaction term was used to assess the change in the volume-outcome relationship.

81x32mm (600 x 600 DPI)

Table S1: Diagnosis and procedure codes used for this analysis

OPS codes	
5-351.0*	Surgical aortic valve replacement
5-35a.0*	Transcatheter aortic valve replacement
5-361.*, 5-362.*, 5-363.*,	Coronary artery bypass graft
5-351.1*, 5-351.2*, 5-353.1, 5-353.2	Surgical mitral valve replacement/reconstruction
5-351.4*	Surgical tricuspid valve replacement
5-377.0 et seqq.	Permanent pacemaker implantation
8-800.7* since 2010: 8-800.c*	Transfusion of RBC

Diagnosis	
I35.0, I06.0	Aortic valve stenosis (degenerative/rheumatic)
I35.2, I06.2	Combined aortic valve diseases (degenerative/rheumatic)
I50.1*	Left ventricular congestive heart failure *(according to NYHA classes)*
I10*	Arterial Hypertension
I25.11, I25.12, I25.13	Coronary artery disease
I25.20, I25.21, I25.22	Previous myocardial infarction *(within 4 months/1 year/after 1 year)*
Z95.1	Previous coronary artery bypass graft
Z95.1 – Z95.4	Previous cardiac surgery
I70.20-I70.25, I70.8, I70.9, I73.9	Peripheral vascular disease
I65.2	Carotid disease
I21*	Acute myocardial infarction *(within the last 28 days)*
J44*	Chronic obstructive pulmonary disease
I27*	Pulmonary hypertension
N18*	Renal disease
N17*	Acute kidney injury
I48.1*	Atrial fibrillation
E10* - E14*	Diabetes
I63*, I64	Stroke or cerebral infarction incl. occlusion and stenosis of cerebral and precerebral arteries, resulting in cerebral infarction

Table S2: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with in-hospital mortality as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	OR	p-value	95% CI
procedure volume <50	1		
procedure volume 50-99	0.989	0.98	0.437 - 2.241
procedure volume >=100	0.668	0.004	0.506 - 0.882
year			
2008	1		
2009	0.929	0.671	0.661 - 1.305
2010	0.844	0.376	0.579 - 1.229
2011	0.686	0.032	0.486 - 0.968
2012	0.525	0.004	0.340 - 0.812
2013	0.465	0.001	0.302 - 0.718
2014	0.453	0.001	0.288 - 0.712
n_ik_year_50_100#year			
procedure volume 50-99 # 2009	0.929	0.874	0.376 - 2.298
procedure volume 50-99 # 2010	0.909	0.828	0.384 - 2.151
procedure volume 50-99 # 2011	1.105	0.822	0.464 - 2.631
procedure volume 50-99 # 2012	1.280	0.597	0.512 - 3.202
procedure volume 50-99 # 2013	1.141	0.781	0.452 - 2.877
procedure volume 50-99 # 2014	0.908	0.845	0.346 - 2.385
procedure volume >=100 # 2009	1.007	0.981	0.583 - 1.740
procedure volume >=100 # 2010	1.120	0.598	0.735 - 1.705
procedure volume >=100 # 2011	1.260	0.223	0.869 - 1.827
procedure volume >=100 # 2012	1.399	0.164	0.872 - 2.244
procedure volume >=100 # 2013	1.606	0.046	1.009 - 2.558
procedure volume >=100 # 2014	1.100	0.687	0.692 - 1.749
Female	0.902	0.045	0.815 - 0.998
Age in years	1.009	0.155	0.997 - 1.022
Estimated logistic EuroSCORE	1.022	0.000	1.015 - 1.029
Aortic valve stenosis	0.636	0.000	0.504 - 0.802
Combined aortic valve diseases	0.553	0.000	0.447 - 0.685
NYHA II	0.551	0.000	0.423 - 0.717
NYHA III or IV	1.550	0.000	1.264 - 1.900
CAD	1.034	0.517	0.934 - 1.144
Hypertension	0.698	0.000	0.612 - 0.797
Previous MI (within 4 months)	0.683	0.048	0.467 - 0.997
Previous MI (within 1 year)	1.042	0.881	0.608 - 1.785
Previous MI (after 1 year)	0.979	0.816	0.821 - 1.169
Previous CABG	1.017	0.884	0.809 - 1.278
Previous cardiac surgery	0.808	0.117	0.619 - 1.055
Peripheral vascular disease	1.118	0.140	0.964 - 1.295
Carotid disease	0.896	0.165	0.768 - 1.046
COPD	0.979	0.744	0.863 - 1.111
Pulmonary hypertension	0.852	0.021	0.744 - 0.976
GFR <15%	1.770	0.000	1.443 - 2.170
GFR <30%	1.414	0.000	1.167 - 1.714
Atrial fibrillation	1.211	0.000	1.115 - 1.315
Diabetes	1.024	0.640	0.926 - 1.133

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year	Prob.	p-value	95% CI
procedure volume <50 # 2008	0.090	0.000	0.068 - 0.113
procedure volume <50 # 2009	0.085	0.000	0.071 - 0.098
procedure volume <50 # 2010	0.077	0.000	0.059 - 0.095
procedure volume <50 # 2011	0.064	0.000	0.048 - 0.08
procedure volume <50 # 2012	0.050	0.000	0.033 - 0.066
procedure volume <50 # 2013	0.044	0.000	0.029 - 0.059
procedure volume <50 # 2014	0.043	0.000	0.027 - 0.059
procedure volume 50-99 # 2008	0.090	0.006	0.026 - 0.153
procedure volume 50-99 # 2009	0.078	0.000	0.058 - 0.098
procedure volume 50-99 # 2010	0.070	0.000	0.059 - 0.082
procedure volume 50-99 # 2011	0.069	0.000	0.057 - 0.082
procedure volume 50-99 # 2012	0.062	0.000	0.049 - 0.076
procedure volume 50-99 # 2013	0.050	0.000	0.038 - 0.061
procedure volume 50-99 # 2014	0.039	0.000	0.026 - 0.052
procedure volume >=100 # 2008	0.062	0.000	0.060 - 0.064
procedure volume >=100 # 2009	0.058	0.000	0.034 - 0.083
procedure volume >=100 # 2010	0.059	0.000	0.044 - 0.074
procedure volume >=100 # 2011	0.054	0.000	0.047 - 0.061
procedure volume >=100 # 2012	0.047	0.000	0.039 - 0.054
procedure volume >=100 # 2013	0.047	0.000	0.040 - 0.055
procedure volume >=100 # 2014	0.032	0.000	0.028 - 0.036

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	-0.007	0.002	-0.012 - -0.003
Annual change	-0.006	0.000	-0.008 - -0.004

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-4.536	0.026	-8.473 - -0.600
Annual change	-0.011	0.000	-0.016 - -0.006
Annual change of volume effect	0.002	0.027	0.000 - 0.004

Table S3: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with stroke as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100		OR	p-value	95% CI
procedure volume <50		1		
procedure volume 50-99		0.929	0.886	0.339 - 2.5473
procedure volume >=100		0.969	0.945	0.398 - 2.3631
year				
	2008	1		
	2009	1.140	0.643	0.654 - 1.989
	2010	0.732	0.308	0.401 - 1.335
	2011	0.678	0.211	0.369 - 1.246
	2012	0.659	0.197	0.350 - 1.241
	2013	0.574	0.114	0.289 - 1.142
	2014	0.759	0.426	0.385 - 1.496
n_ik_year_50_100#year				
procedure volume 50-99 #	2009	1.079	0.896	0.344 - 3.384
procedure volume 50-99 #	2010	1.310	0.637	0.428 - 4.009
procedure volume 50-99 #	2011	1.233	0.715	0.400 - 3.796
procedure volume 50-99 #	2012	1.247	0.706	0.396 - 3.921
procedure volume 50-99 #	2013	1.332	0.631	0.414 - 4.288
procedure volume 50-99 #	2014	0.957	0.942	0.290 - 3.162
procedure volume >=100 #	2009	0.779	0.649	0.265 - 2.287
procedure volume >=100 #	2010	1.134	0.810	0.407 - 3.159
procedure volume >=100 #	2011	1.640	0.331	0.606 - 4.440
procedure volume >=100 #	2012	1.209	0.712	0.441 - 3.312
procedure volume >=100 #	2013	1.769	0.282	0.626 - 4.995
procedure volume >=100 #	2014	1.099	0.857	0.393 - 3.076
Female		0.648	0.000	0.565 - 0.744
Age in years		0.900	0.000	0.890 - 0.910
Estimated logistic EuroSCORE		68461	0.000	33804 - 138650
Aortic valve stenosis		1.278	0.067	0.983 - 1.663
Combined aortic valve diseases		1.329	0.045	1.006 - 1.755
NYHA II		0.949	0.675	0.744 - 1.211
NYHA III or IV		1.096	0.166	0.962 - 1.249
CAD		1.011	0.875	0.885 - 1.154
Hypertension		0.910	0.148	0.802 - 1.034
Previous MI (within 4 months)		0.278	0.000	0.167 - 0.464
Previous MI (within 1 year)		1.246	0.491	0.667 - 2.328
Previous MI (after 1 year)		1.005	0.975	0.742 - 1.360
Previous CABG		0.857	0.350	0.620 - 1.184
Previous cardiac surgery		0.120	0.000	0.089 - 0.161
Peripheral vascular disease		0.395	0.000	0.325 - 0.480
Carotid disease		0.461	0.000	0.362 - 0.587
COPD		0.355	0.000	0.295 - 0.429
Pulmonary hypertension		0.198	0.000	0.164 - 0.239
GFR <15%		0.266	0.000	0.183 - 0.387
GFR <30%		0.297	0.000	0.226 - 0.391
Atrial fibrillation		1.093	0.165	0.964 - 1.239
Diabetes		1.079	0.260	0.945 - 1.231

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year		Prob.	p-value	95% CI
procedure volume <50 #	2008	0.022	0.000	0.012 - 0.031
procedure volume <50 #	2009	0.025	0.000	0.017 - 0.032
procedure volume <50 #	2010	0.016	0.000	0.010 - 0.022
procedure volume <50 #	2011	0.015	0.000	0.009 - 0.021
procedure volume <50 #	2012	0.014	0.000	0.008 - 0.021
procedure volume <50 #	2013	0.013	0.000	0.006 - 0.019
procedure volume <50 #	2014	0.016	0.000	0.008 - 0.025
procedure volume 50-99 #	2008	0.020	0.026	0.002 - 0.038
procedure volume 50-99 #	2009	0.025	0.000	0.014 - 0.035
procedure volume 50-99 #	2010	0.019	0.000	0.014 - 0.025
procedure volume 50-99 #	2011	0.017	0.000	0.012 - 0.022
procedure volume 50-99 #	2012	0.017	0.000	0.011 - 0.022
procedure volume 50-99 #	2013	0.015	0.000	0.011 - 0.020
procedure volume 50-99 #	2014	0.015	0.000	0.009 - 0.021
procedure volume >=100 #	2008	0.021	0.009	0.005 - 0.037
procedure volume >=100 #	2009	0.019	0.000	0.009 - 0.028
procedure volume >=100 #	2010	0.017	0.000	0.012 - 0.023
procedure volume >=100 #	2011	0.023	0.000	0.019 - 0.028
procedure volume >=100 #	2012	0.017	0.000	0.014 - 0.020
procedure volume >=100 #	2013	0.021	0.000	0.018 - 0.025
procedure volume >=100 #	2014	0.018	0.000	0.015 - 0.020

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	0.001	0.196	-0.001 - 0.003
Annual change	-0.001	0.029	-0.002 - 0.000

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-1.103	0.307	-3.311 - 1.106
Annual change	-0.002	0.086	-0.005 - 0.000
Annual change of volume effect	0.001	0.306	-0.001 - 0.002

Table S4: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with bleeding as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100		OR	p-value	95% CI
procedure volume <50		1		
procedure volume 50-99		0.806	0.366	0.506 - 1.286
procedure volume >=100		0.485	0.006	0.291 - 0.811
year				
	2008	1		
	2009	0.942	0.677	0.711 - 1.248
	2010	0.775	0.086	0.579 - 1.037
	2011	0.574	0.000	0.422 - 0.781
	2012	0.502	0.000	0.363 - 0.696
	2013	0.559	0.001	0.399 - 0.785
	2014	0.340	0.000	0.226 - 0.511
n_ik_year_50_100#year				
procedure volume 50-99 #	2009	0.992	0.978	0.571 - 1.724
procedure volume 50-99 #	2010	1.176	0.541	0.699 - 1.980
procedure volume 50-99 #	2011	1.224	0.458	0.718 - 2.087
procedure volume 50-99 #	2012	1.321	0.321	0.762 - 2.293
procedure volume 50-99 #	2013	0.894	0.693	0.511 - 1.563
procedure volume 50-99 #	2014	1.265	0.463	0.675 - 2.374
procedure volume >=100 #	2009	1.070	0.828	0.581 - 1.970
procedure volume >=100 #	2010	1.078	0.799	0.605 - 1.920
procedure volume >=100 #	2011	2.198	0.006	1.249 - 3.866
procedure volume >=100 #	2012	1.669	0.080	0.941 - 2.960
procedure volume >=100 #	2013	1.371	0.286	0.768 - 2.447
procedure volume >=100 #	2014	1.540	0.174	0.827 - 2.868
Female		1.096	0.032	1.008 - 1.193
Age in years		0.982	0.000	0.974 - 0.990
Estimated logistic EuroSCORE		5.813	0.000	3.386 - 9.979
Aortic valve stenosis		0.738	0.000	0.639 - 0.852
Combined aortic valve diseases		0.677	0.000	0.580 - 0.790
NYHA II		0.665	0.000	0.562 - 0.786
NYHA III or IV		1.313	0.000	1.216 - 1.418
CAD		1.062	0.137	0.981 - 1.149
Hypertension		0.798	0.000	0.741 - 0.861
Previous MI (within 4 months)		0.866	0.324	0.650 - 1.153
Previous MI (within 1 year)		1.071	0.742	0.711 - 1.614
Previous MI (after 1 year)		0.869	0.137	0.721 - 1.046
Previous CABG		0.530	0.000	0.447 - 0.629
Previous cardiac surgery		1.275	0.005	1.077 - 1.509
Peripheral vascular disease		1.255	0.000	1.118 - 1.409
Carotid disease		1.182	0.022	1.024 - 1.364
COPD		0.998	0.969	0.896 - 1.111
Pulmonary hypertension		0.833	0.002	0.741 - 0.935
GFR <15%		2.045	0.000	1.725 - 2.423
GFR <30%		1.446	0.000	1.240 - 1.685
Atrial fibrillation		1.418	0.000	1.316 - 1.528
Diabetes		0.968	0.418	0.894 - 1.048

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year		Prob.	p-value	95% CI
procedure volume <50 #	2008	0.134	0.000	0.108 - 0.161
procedure volume <50 #	2009	0.128	0.000	0.109 - 0.146
procedure volume <50 #	2010	0.107	0.000	0.090 - 0.125
procedure volume <50 #	2011	0.082	0.000	0.066 - 0.097
procedure volume <50 #	2012	0.072	0.000	0.057 - 0.088
procedure volume <50 #	2013	0.080	0.000	0.062 - 0.098
procedure volume <50 #	2014	0.050	0.000	0.034 - 0.066
procedure volume 50-99 #	2008	0.111	0.000	0.071 - 0.151
procedure volume 50-99 #	2009	0.105	0.000	0.082 - 0.128
procedure volume 50-99 #	2010	0.102	0.000	0.089 - 0.116
procedure volume 50-99 #	2011	0.081	0.000	0.069 - 0.093
procedure volume 50-99 #	2012	0.077	0.000	0.064 - 0.090
procedure volume 50-99 #	2013	0.059	0.000	0.049 - 0.069
procedure volume 50-99 #	2014	0.051	0.000	0.039 - 0.063
procedure volume >=100 #	2008	0.070	0.000	0.040 - 0.100
procedure volume >=100 #	2009	0.071	0.000	0.052 - 0.090
procedure volume >=100 #	2010	0.059	0.000	0.048 - 0.070
procedure volume >=100 #	2011	0.087	0.000	0.078 - 0.096
procedure volume >=100 #	2012	0.059	0.000	0.053 - 0.066
procedure volume >=100 #	2013	0.055	0.000	0.049 - 0.060
procedure volume >=100 #	2014	0.038	0.000	0.034 - 0.042

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	-0.011	0.001	-0.016 - -0.0049
Annual change	-0.011	0.000	-0.013 - -0.0083

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-4.84852	0.169	-11.97 - 2.27482
Annual change	-0.01589	0.001	-0.024 - -0.0076
Annual change of volume effect	0.0024	0.170	-0.001 - 0.00595

Table S5: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Linear regression model on 43,996 TAVI cases with Length of hospital stay as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100		Coeff	p-value	95% CI
procedure volume <50		0		
procedure volume 50-99		2.959	0.001	1.133 - 4.786
procedure volume >=100		-4.148	0.000	-5.574 - -2.721
year				
	2008	0		
	2009	2.179	0.001	0.950 - 3.409
	2010	1.472	0.019	0.247 - 2.697
	2011	0.507	0.422	-0.731 - 1.745
	2012	-1.002	0.100	-2.197 - 0.193
	2013	0.382	0.580	-0.970 - 1.734
	2014	-0.060	0.931	-1.433 - 1.312
n_ik_year_50_100#year				
procedure volume 50-99 #	2009	-5.745	0.000	-7.942 - -3.547
procedure volume 50-99 #	2010	-4.872	0.000	-6.926 - -2.819
procedure volume 50-99 #	2011	-3.555	0.001	-5.622 - -1.488
procedure volume 50-99 #	2012	-2.174	0.037	-4.219 - -0.130
procedure volume 50-99 #	2013	-4.452	0.000	-6.576 - -2.327
procedure volume 50-99 #	2014	-3.947	0.000	-6.110 - -1.783
procedure volume >=100 #	2009	1.066	0.242	-0.721 - 2.853
procedure volume >=100 #	2010	0.788	0.365	-0.915 - 2.490
procedure volume >=100 #	2011	1.972	0.021	0.292 - 3.652
procedure volume >=100 #	2012	2.929	0.000	1.302 - 4.557
procedure volume >=100 #	2013	0.970	0.275	-0.773 - 2.713
procedure volume >=100 #	2014	0.390	0.662	-1.358 - 2.139
Female		-0.483	0.000	-0.725 - -0.241
Age in years		-0.247	0.000	-0.276 - -0.218
Estimated logistic EuroSCORE		25.503	0.000	23.345 - 27.661
Aortic valve stenosis		-6.255	0.000	-6.912 - -5.598
Combined aortic valve diseases		-6.592	0.000	-7.263 - -5.921
NYHA II		0.154	0.365	-0.179 - 0.487
NYHA III or IV		2.597	0.000	2.374 - 2.821
CAD		-0.037	0.740	-0.259 - 0.184
Hypertension		-0.888	0.000	-1.111 - -0.666
Previous MI (within 4 months)		-3.355	0.000	-4.175 - -2.534
Previous MI (within 1 year)		0.015	0.980	-1.172 - 1.203
Previous MI (after 1 year)		-0.303	0.250	-0.819 - 0.213
Previous CABG		-2.938	0.000	-3.596 - -2.280
Previous cardiac surgery		-1.710	0.000	-2.405 - -1.015
Peripheral vascular disease		-0.917	0.000	-1.345 - -0.489
Carotid disease		-1.110	0.000	-1.614 - -0.606
COPD		-0.618	0.001	-0.973 - -0.263
Pulmonary hypertension		-2.226	0.000	-2.624 - -1.827
GFR <15%		1.941	0.000	0.958 - 2.923
GFR <30%		0.725	0.034	0.054 - 1.396
Atrial fibrillation		2.575	0.000	2.365 - 2.785
Diabetes		0.932	0.000	0.702 - 1.161

Second step: Marginal means are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year		Coeff	p-value	95% CI
procedure volume <50 #	2008	19.184	0.000	18.241 - 20.13
procedure volume <50 #	2009	21.364	0.000	20.577 - 22.15
procedure volume <50 #	2010	20.656	0.000	19.874 - 21.44
procedure volume <50 #	2011	19.691	0.000	18.890 - 20.49
procedure volume <50 #	2012	18.182	0.000	17.448 - 18.92
procedure volume <50 #	2013	19.566	0.000	18.599 - 20.53
procedure volume <50 #	2014	19.124	0.000	18.128 - 20.12
procedure volume 50-99 #	2008	22.144	0.000	20.579 - 23.71
procedure volume 50-99 #	2009	18.578	0.000	17.650 - 19.51
procedure volume 50-99 #	2010	18.743	0.000	18.221 - 19.27
procedure volume 50-99 #	2011	19.095	0.000	18.554 - 19.64
procedure volume 50-99 #	2012	18.967	0.000	18.413 - 19.52
procedure volume 50-99 #	2013	18.074	0.000	17.576 - 18.57
procedure volume 50-99 #	2014	18.137	0.000	17.545 - 18.73
procedure volume >=100 #	2008	15.036	0.000	13.964 - 16.11
procedure volume >=100 #	2009	18.281	0.000	17.543 - 19.02
procedure volume >=100 #	2010	17.296	0.000	16.791 - 17.8
procedure volume >=100 #	2011	17.515	0.000	17.137 - 17.89
procedure volume >=100 #	2012	16.964	0.000	16.696 - 17.23
procedure volume >=100 #	2013	16.389	0.000	16.136 - 16.64
procedure volume >=100 #	2014	15.366	0.000	15.179 - 15.55

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated means.

	Coeff	p-value	95% CI
Volume effect	-1.488	0.000	-2.021 - -0.9555
Annual change	-0.269	0.029	-0.507 - -0.0307

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	22.5075	0.935	-554.1 - 599.146
Annual change	-0.23991	0.452	-0.898 - 0.41837
Annual change of volume effect	-0.01193	0.931	-0.299 - 0.2748

Table S6: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Linear regression model on 43,996 TAVI cases with reimbursement as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	Coeff	p-value	95% CI
procedure volume <50	0		
procedure volume 50-99	-905.1	0.024	-1689.5 - -120.7
procedure volume >=100	-1792.9	0.000	-2614.4 - -971.4
year			
2010	0		
2011	-1104.3	0.026	-2078.5 - -130.1
2012	-1872.9	0.000	-2810.5 - -935.4
2013	-1339.7	0.007	-2316.5 - -362.9
2014	-1240.5	0.015	-2235.4 - -245.5
n_ik_year_50_100#year			
procedure volume 50-99 # 2011	349.5	0.545	-783.6 - 1482.5
procedure volume 50-99 # 2012	598.9	0.269	-464.0 - 1661.8
procedure volume 50-99 # 2013	-162.6	0.775	-1276.2 - 950.9
procedure volume 50-99 # 2014	-51.3	0.928	-1159.3 - 1056.7
procedure volume >=100 # 2011	994.4	0.082	-127.8 - 2116.7
procedure volume >=100 # 2012	994.3	0.063	-55.5 - 2044.2
procedure volume >=100 # 2013	614.5	0.267	-471.3 - 1700.4
procedure volume >=100 # 2014	803.0	0.150	-289.3 - 1895.4
Female	-816.4	0.000	-1022.5 - -610.4
Age in years	-134.0	0.000	-159.5 - -108.6
Estimated logistic EuroSCORE	9498.4	0.000	7514.7 - 11482.1
Aortic valve stenosis	-1480.8	0.000	-2097.4 - -864.3
Combined aortic valve diseases	-1671.1	0.000	-2296.6 - -1045.6
NYHA II	-420.2	0.001	-665.2 - -175.3
NYHA III or IV	686.8	0.000	501.3 - 872.2
CAD	133.6	0.153	-49.7 - 316.8
Hypertension	-427.8	0.000	-614.4 - -241.3
Previous MI (within 4 months)	-1677.0	0.000	-2134.6 - -1219.4
Previous MI (within 1 year)	295.6	0.574	-733.9 - 1325.2
Previous MI (after 1 year)	-483.1	0.009	-843.2 - -123.0
Previous CABG	-1118.8	0.000	-1687.9 - -549.8
Previous cardiac surgery	-574.4	0.061	-1174.8 - 25.9
Peripheral vascular disease	86.9	0.649	-287.8 - 461.7
Carotid disease	-365.0	0.106	-807.3 - 77.3
COPD	-0.6	0.997	-318.6 - 317.4
Pulmonary hypertension	-951.6	0.000	-1302.7 - -600.4
GFR <15%	1849.2	0.000	921.0 - 2777.4
GFR <30%	322.0	0.258	-235.7 - 879.6
Atrial fibrillation	913.0	0.000	741.8 - 1084.1
Diabetes	223.8	0.020	34.7 - 412.9

Second step: Marginal means are calculated by setting each confounder to its mean value (prediction at the means) using Stata's `margins` command with application of the `atmeans` option.

n_ik_year_50_100#year	Coeff	p-value	95% CI
procedure volume <50 # 2010	36999.8	0.000	36302.5 - 37697.1
procedure volume <50 # 2011	35895.5	0.000	35214.3 - 36576.7
procedure volume <50 # 2012	35126.9	0.000	34499.5 - 35754.2
procedure volume <50 # 2013	35660.1	0.000	34979.4 - 36340.7
procedure volume <50 # 2014	35759.3	0.000	35049.7 - 36469.0
procedure volume 50-99 # 2010	36094.7	0.000	35738.1 - 36451.3
procedure volume 50-99 # 2011	35339.9	0.000	34882.3 - 35797.5
procedure volume 50-99 # 2012	34820.7	0.000	34470.7 - 35170.6
procedure volume 50-99 # 2013	34592.3	0.000	34198.2 - 34986.4
procedure volume 50-99 # 2014	34803.0	0.000	34459.0 - 35147.0
procedure volume >=100 # 2010	35206.9	0.000	34775.2 - 35638.7
procedure volume >=100 # 2011	35097.1	0.000	34744.3 - 35449.8
procedure volume >=100 # 2012	34328.3	0.000	34128.7 - 34527.9
procedure volume >=100 # 2013	34481.7	0.000	34285.1 - 34678.4
procedure volume >=100 # 2014	34769.5	0.000	34623.4 - 34915.6

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated means.

	Coeff	p-value	95% CI
Volume effect	-503.8	0.005	-826.4 - -181.2
Annual change	-257.6	0.011	-444.3 - -70.9

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-245427.3	0.262	-702076.0 - 211221.4
Annual change	-518.0	0.050	-1036.3 - 0.2
Annual change of volume effect	121.7	0.263	-105.2 - 348.7

Table S7: Results of the logistic regression model, transformation into discrete event probabilities and subsequent meta regressions

First step: Logistic regression model on 43,996 TAVI cases with ventilation as dependent variable, an interaction term (n_ik_year_50_100#year) between categorical time (in years) and volume categories and 22 predefined patient and procedural characteristics as potential confounder.

n_ik_year_50_100	OR	p-value	95% CI
procedure volume <50	1		
procedure volume 50-99	0.716	0.262	0.400 - 1.283
procedure volume >=100	0.492	0.026	0.263 - 0.918
year			
2008	1		
2009	0.946	0.745	0.677 - 1.322
2010	0.853	0.362	0.605 - 1.201
2011	0.765	0.135	0.539 - 1.086
2012	0.664	0.029	0.459 - 0.959
2013	0.623	0.018	0.420 - 0.923
2014	0.532	0.004	0.345 - 0.819
n_ik_year_50_100#year			
procedure volume 50-99 # 2009	1.079	0.827	0.544 - 2.141
procedure volume 50-99 # 2010	1.382	0.322	0.729 - 2.622
procedure volume 50-99 # 2011	1.43	0.277	0.751 - 2.723
procedure volume 50-99 # 2012	1.096	0.789	0.561 - 2.142
procedure volume 50-99 # 2013	0.929	0.833	0.469 - 1.841
procedure volume 50-99 # 2014	1.016	0.967	0.489 - 2.111
procedure volume >=100 # 2009	1.732	0.131	0.848 - 3.537
procedure volume >=100 # 2010	1.232	0.555	0.616 - 2.462
procedure volume >=100 # 2011	1.914	0.059	0.975 - 3.758
procedure volume >=100 # 2012	1.624	0.163	0.822 - 3.210
procedure volume >=100 # 2013	1.636	0.165	0.817 - 3.277
procedure volume >=100 # 2014	1.358	0.402	0.664 - 2.776
Female	0.713	0.000	0.651 - 0.781
Age in years	0.959	0.000	0.951 - 0.968
Estimated logistic EuroSCORE	13.81	0.000	7.797 - 24.464
Aortic valve stenosis	0.722	0.000	0.618 - 0.843
Combined aortic valve diseases	0.663	0.000	0.561 - 0.783
NYHA II	0.498	0.000	0.404 - 0.614
NYHA III or IV	1.485	0.000	1.364 - 1.617
CAD	1.094	0.044	1.002 - 1.193
Hypertension	0.697	0.000	0.642 - 0.757
Previous MI (within 4 months)	0.804	0.163	0.591 - 1.093
Previous MI (within 1 year)	0.796	0.358	0.490 - 1.294
Previous MI (after 1 year)	0.897	0.268	0.740 - 1.087
Previous CABG	0.675	0.000	0.557 - 0.816
Previous cardiac surgery	0.848	0.091	0.701 - 1.026
Peripheral vascular disease	1.198	0.004	1.060 - 1.353
Carotid disease	0.855	0.061	0.725 - 1.007
COPD	1.211	0.001	1.085 - 1.351
Pulmonary hypertension	0.758	0.000	0.669 - 0.858
GFR <15%	1.364	0.001	1.129 - 1.647
GFR <30%	1.252	0.008	1.059 - 1.479
Atrial fibrillation	1.553	0.000	1.430 - 1.687
Diabetes	1.138	0.003	1.045 - 1.239

Second step: Predicted probabilities are calculated by setting each confounder to its mean value (prediction at the means) using Stata's margins command with application of the atmeans option.

n_ik_year_50_100#year	Prob.	p-value	95% CI
procedure volume <50 # 2008	0.086	0.000	0.065 - 0.107
procedure volume <50 # 2009	0.082	0.000	0.067 - 0.096
procedure volume <50 # 2010	0.074	0.000	0.060 - 0.089
procedure volume <50 # 2011	0.067	0.000	0.053 - 0.081
procedure volume <50 # 2012	0.059	0.000	0.045 - 0.073
procedure volume <50 # 2013	0.055	0.000	0.040 - 0.070
procedure volume <50 # 2014	0.048	0.000	0.032 - 0.063
procedure volume 50-99 # 2008	0.063	0.000	0.033 - 0.094
procedure volume 50-99 # 2009	0.064	0.000	0.046 - 0.083
procedure volume 50-99 # 2010	0.074	0.000	0.062 - 0.085
procedure volume 50-99 # 2011	0.069	0.000	0.058 - 0.079
procedure volume 50-99 # 2012	0.047	0.000	0.037 - 0.057
procedure volume 50-99 # 2013	0.038	0.000	0.030 - 0.045
procedure volume 50-99 # 2014	0.035	0.000	0.025 - 0.045
procedure volume >=100 # 2008	0.044	0.000	0.020 - 0.068
procedure volume >=100 # 2009	0.070	0.000	0.052 0.089
procedure volume >=100 # 2010	0.046	0.000	0.037 0.056
procedure volume >=100 # 2011	0.063	0.000	0.056 0.071
procedure volume >=100 # 2012	0.047	0.000	0.042 0.053
procedure volume >=100 # 2013	0.045	0.000	0.040 0.050
procedure volume >=100 # 2014	0.032	0.000	0.029 0.036

Third step: A random effects meta regression (using Stata's command metareg) with time and volume as continuous covariates was applied to the estimated rates.

	Coeff	p-value	95% CI
Volume effect	-0.010	0.004	-0.016 - -0.004
Annual change	-0.004	0.002	-0.007 - -0.002

Fourth step: A second random effects meta regression model was applied including also an interaction term.

	Coeff	p-value	95% CI
Volume effect	-6.084	0.040	-11.870 - -0.299
Annual change	-0.011	0.004	-0.018 - -0.004
Annual change of volume effect	0.003	0.041	0.000 - 0.006

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	p1-2	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	R(ecord)1.1: p1 R1.2: p1-2 R1.3: n/a
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	p5		
Objectives	3	State specific objectives, including any prespecified hypotheses	p5		
Methods					
Study Design	4	Present key elements of study design early in the paper	p6-9		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	p6		
Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the	p6	RECORD 6.1: The methods of study population selection (such as codes or	R6.1: p6 R6.2, 6.3: n/a

		sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case		algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	p6-7	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	p6-7
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	p6-7, Supplementary tables		
Bias	9	Describe any efforts to address potential sources of bias	p6-7, 9		
Study size	10	Explain how the study size was	n/a (national cohort)		

		arrived at	(p6-7)		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	p7-10		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	p7-10		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	p6
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-	n/a

				level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	n/a (national cohort, administrative data, no follow-up)	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	n/a (national cohort, administrative data, no follow-up)
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	p8		
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures	p13		
Main results	16	(a) Give unadjusted estimates	p10-14		

		and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	n/a		
Discussion					
Key results	18	Summarise key results with reference to study objectives	p14-15		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	p3, 9-10	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	p3, 9-10
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	p 14-17		
Generalisability	21	Discuss the generalisability (external validity) of the study results	p3, 14-17		

Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	p4		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Access to public dataset: p6 Further data: Supplemental tables

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.