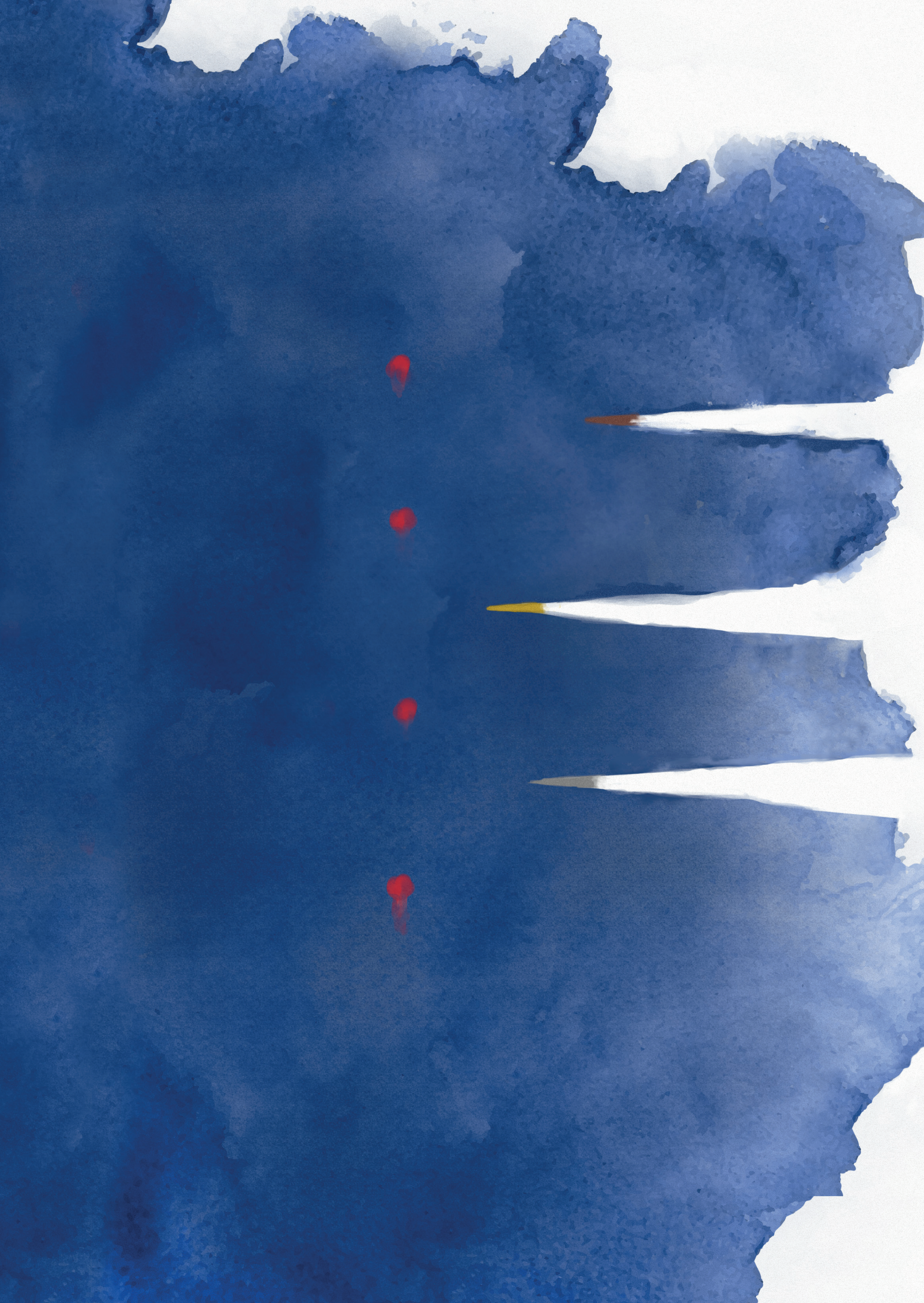
What makes the best performing hospital? the IQ Joint study

Schie, P. van



Chapter 3

Between-Hospital Variation in Revision Rates After Total Hip and Knee Arthroplasty in the Netherlands

Directing Quality-Improvement Initiatives

Peter van Schie[1,2], Liza N. van Steenbergen[3], Leti van Bodegom-Vos[2], Rob G.H.H. Nelissen[1], Perla J. Marang- van de Mheen[2]

1. Department of Orthopaedics, Leiden University Medical Centre, Leiden, Netherlands;
2. Department of Biomedical Data Sciences, Medical Decision Making, Leiden University Medical Centre, Leiden, Netherlands;
3. Dutch Arthroplasty Register (LROI), 's-Hertogenbosch, Netherlands.

J Bone Joint Surg Am. 2020 Feb 19;102(4):315-324

Abstract

Background

Variation in 1-year revision rates between Dutch hospitals after primary total hip and knee arthroplasty (THA and TKA) may direct quality-improvement initiatives if this variation accurately reflects true hospital differences. The aim of the present study was to assess the extent of variation, both overall and for specific indications, as well as the statistical reliability of ranking hospitals.

Methods

All primary THAs and TKAs that were performed between January 2014 and December 2016 were included. Observed/expected (O/E) ratios regarding 1-year revision rates were depicted in a funnel plot with 95% control limits to identify outliers based on 1 or 3 years of data, both overall and by specific indication for revision. The expected number was calculated on the basis of patient mix with use of logistic regression models. The statistical reliability of ranking hospitals (rankability) on these outcomes indicates the percentage of total variation that is explained by "true" hospital differences rather than chance. Rankability was evaluated using fixed and random effects models, for overall revisions and specific indications for revision, including 1 versus 3 years of data.

Results

The present study included 86,468 THAs and 73,077 TKAs from 97 and 98 hospitals, respectively. Thirteen hospitals performing THAs were identified as negative outliers (median O/E ratio, 1.9; interquartile range [IQR], 1.5-2.5), with 5 hospitals as outliers in multiple years. Eight negative outliers were identified for periprosthetic joint infection; 4, for dislocation; and 2, for prosthesis loosening. Seven hospitals performing TKAs were identified as negative outliers (median O/E-ratio, 2.3; IQR, 2.2-2.8), with 2 hospitals as outliers in multiple years. Two negative outlier hospitals were identified for periprosthetic joint infection and 1 was identified for technical failures. The rankability for overall revisions was 62% (moderate) for THA and 46% (low) for TKA.

Conclusion

There was large between-hospital variation in 1-year revision rates after primary THA and TKA. For most outlier hospitals, a specific indication for revision could be identified as contributing to worse performance, particularly for THA; these findings are starting points for quality-improvement initiatives.

Introduction

Recent studies showed variation in hospital performance after total hip arthroplasty (THA) and total knee arthroplasty (TKA) in terms of outcomes such as length of stay, readmission, complications, allogeneic blood transfusion, and mortality as well as costs, suggesting that improvement is possible for some hospitals(1-10). For instance, a fivefold-higher postoperative complication rate was found in some United States hospitals compared with others, which may be due to patient characteristics, true hospital differences, and random variation(2,9). The statistical reliability of ranking (rankability) indicates the percentage of variation that is due to "true" hospital differences as opposed to random (chance) variation(11-16). Random variation is particularly likely if the number of events is small(17). Therefore, rankability adds crucial information to calculated between-hospital variation in clinical outcomes. Conclusions about the quality of care delivered are only reliable if an outcome is ranked reliably.

Most arthroplasty registries publish annual reports with the aim of monitoring hospital performance and providing hospitals with feedback. In orthopaedics, the 1-year revision rate is an important clinical outcome that is often used to monitor the quality of care delivered. The consequences of revision—both for the patient and in terms of associated costs—are considerable. Because of low revision rates, several years of outcomes are usually combined to have more events(18-20). As a result, it may take longer before deteriorating hospital performance is noticed. However, if the rankability for single years of data analysis is acceptable, a reliable earlier signal can be given when performance deteriorates. Furthermore, variation in 1-year revision rates for specific indications (infection, dislocation, etc.) may direct quality-improvement initiatives for worse-performing hospitals. The variation in 1-year revision rates among hospitals is currently unknown, as are the variations for specific indications and the statistical reliability of ranking on revisions.

The aims of the present investigation were (1) to assess the extent of variation in 1-year revision rates between Dutch hospitals after primary THA and TKA procedures in the period 2014-2016 and (2) to estimate rankability to determine the extent to which our findings represent true hospital differences. In both cases, we evaluated both 3 years of procedures as well as single years for both overall revision and specific indications for revision.

Materials and Methods

Study design and setting

This observational study used routinely collected data from the nationwide Dutch Arthroplasty Register (LROI). The LROI was established in 2007, and by 2012 all Dutch hospitals were participating. For every arthroplasty procedure, the product number of the implanted component is registered to identify the prosthesis, as well as the encrypted social security number of each patient so that it is possible to connect a revision procedure to the hospital where the primary procedure took place. Surgeons provide information about the procedure, patient characteristics, and any possible revisions (including the type of, and indication for, revision). The vital status of patients (dead or alive) is obtained from Vektis, the Dutch insurance health-care database. The LROI uses the opt-out system for informed consent, whereby patients must actively object in order not to be included. The completeness of data on primary THA and TKA procedures is checked against hospital electronic health records and currently exceeds 98% for primary procedures and 96% for revisions, meaning that >98% of primary procedures performed are included in the register(21,22). THAs were performed at 97 hospitals, and TKAs were performed at 98.

Patients and outcomes

Anonymous data on all Dutch patients undergoing a primary THA or TKA procedure between January 2014 and December 2016 were included. The following patient characteristics that may influence the need for revision were recorded: age, sex, body mass index (BMI), smoking status (yes or no), American Society of Anaesthesiologists (ASA) classification (I, II, III-IV), Charnley score (A, B1, B2, C, not applicable), and diagnosis (osteoarthritis or non-osteoarthritis)(23). Revision within 1 year (yes or no) was the primary outcome measure (defined as exchange, removal, or addition of any component). When a revision was performed, surgeons registered 1 of the available indications for revision:

- THA: wear (cup and/or insert), dislocation, prosthesis removal, prosthesis loosening (femur and/or acetabulum), periarticular ossification, symptomatic metal-on-metal bearing;

- TKA: patellar pain, wear (modular tibial polyethylene insert), patellar dislocation, malalignment, instability, prosthesis loosening (femur, tibia, and/or patella), progressive patellofemoral osteoarthritis, arthrofibrosis;

- THA and TKA: periprosthetic joint infection, periprosthetic fracture, “other” indication.

To direct quality-improvement initiatives if a hospital performed more revisions than expected given its patient mix, the indications for revision were categorized into the following groups:

- THA: infection, loosening (acetabular and/or femur), and dislocation;

- TKA: infection, loosening (femur, tibia, and/or patella), and technical failure (malalignment, instability, and/or patellar dislocation).

Statistical analysis

First, we estimated the between-hospital variation in 1-year revision rates after primary THA and TKA procedures, with adjustment for differences in patient mix (using the same method as the LROI). For each patient, the expected revision risk was calculated with use of logistic regression analysis, with all of the patient characteristics listed above as independent variables and 1-year revision the dependent variable. Missing patient characteristic values (<10% for all variables) were imputed with the mean for numeric variables or the mode for categorical variables (so that the most frequently occurring category was imputed). All expected revision risks at the patient level were summed within a hospital to obtain the aggregated expected number (E) of revisions per hospital. The observed number (O) of revisions per hospital was then divided by the expected number to calculate an O/E ratio for each hospital.

The O/E ratios were depicted in a funnel plot with 95% control limits, including 3 years of procedures (2014-2016) and single years(24). Hospitals outside these limits either have significantly lower revision rates than expected and thus perform better (positive outlier, represented by green dots on the funnel plot) or have significantly higher revision rates and perform worse (negative outlier, represented by red dots on the funnel plot). Hospitals within these limits do not perform different than expected. Feedback based on a single year of performance data may be given sooner, but 3-year data may have better power. The analyses were repeated for specific indications for revision as these indications might point to starting points for quality improvement.

Statistical reliability of ranking (Rankability)

Rankability was introduced in previous research and refers to the statistical reliability of ranking hospitals(11-13). Rankability reflects a signal-to-noise ratio and is expressed as the percentage of the hospital variation being due to 'true' hospital differences rather than chance variation with use of the following equation(11-14):

$$\text{Rankability} = \frac{\text{Between−hospital variation}(\tau^2)}{\text{Between−hospital variation}(\tau^2) + \text{Within−hospital variation}(\sigma^2)}\ X\ 100\%$$

Between-hospital variation was estimated with use of a random effects logistic regression model to adjust for clustering of patients within hospitals, with the hospital variable as a random factor and patient characteristics as fixed factors. Within-hospital variation was estimated with use of a fixed effect logistic regression model, with all patient characteristics as fixed factors to adjust for patient mix and hospital as a categorical variable. The median squared standard error (SE) of the hospital estimate was taken to represent the within-hospital variance(15). The between and within-hospital variation from these models are included in the equation above to evaluate the extent to which hospital variation can be attributed to "true" hospital differences. If the within-hospital variation is relatively large, for example, as a result of low-frequency outcomes, it will become harder to detect any between-hospital differences. Rankability was classified as low (<50%), moderate (50% to 75%), or high (>75%), as previously suggested(15). Rankability was determined for THA and TKA separately for both 3 years and single years of data as well as by the indication for revision. If hospitals are ranked reliably over a shorter period of time (e.g., single years), deteriorating hospital performance is identified sooner and hospitals do not have to wait several years for data, potentially preventing more failures from occurring(16).

The analyses on hospital variations were performed with use of SPSS (version 23; IBM). Analyses on rankability were performed with use of STATA (version 14; StataCorp). The LUMC Medical Ethical Committee waived the need for ethical approval under Dutch law.

Results

In total, 86,468 primary THA procedures from 97 hospitals and 73,077 primary TKA procedures from 98 hospitals were included. Less than 4% of patient mix variables were missing, except for smoking, which was below 10%. The average 1-year revision rate across all patients was 1.8% for THA and 1.2% for TKA, with a median rate per hospital of 1.6% (interquartile range [IQR], 1.0%-2.3%) for THA and 1.1% (IQR, 0.7%-1.6%) for TKA (Table I).

Table I Distribution of Patient Characteristics and 1-Year Revisions in Dutch Hospitals in the Period 2014-2016.

	THA (*n*=97 hospitals)		**TKA (*n*=98 hospitals)**	
	Median (IQR)	**Range**	**Median (IQR)**	**Range**
Procedures (*n*)	759 (526-1173)	2-2502	699 (463-938)	9-1998
Mean age (years)	69.3 (67.8-70.1)	50.6-71.8	68.8 (67.4-69.7)	56.5-72.2
Gender, female (%)	66.1 (63.3-68.0)	0.0-74.1	65.2 (61.9-67.8)	8.3-100.0
Mean BMI (kg/m^2)	27.3 (27.0-27.8)	25.9-28.6	29.8 (29.3-30.4)	20.5-31.0
Smoking (%)	13.2 (10.7-15.2)	0.0-27.9	9.8 (8.4-11.8)	1.0-20.5
ASA classification (%)				
• ASA I	17.4 (14.2-21.4)	3.3-100	11.8 (9.8-16.0)	3.8-54.5
• ASA II	65.0 (59.8-70.4)	0.0-96.7	68.7 (63.7-73.6)	42.5-91.6
• ASA III-IV	15.6 (11.5-20.4)	0.0-40.1	16.6 (10.8-21.8)	0.0-50.6
Charnley score (%)				
• A	49.3 (43.7-53.9)	23.7-78.2	45.3 (35.6-52.4)	13.1-100.0
• B1	27.8 (22.9-33.4)	3.6-50.7	33.0 (27.3-40.3)	0.0- 57.8
• B2	20.1 (18.1-22.9)	4.7-28.3	19.4 (16.2-21.5)	0.0-28.0
• C	1.9 (1.0-3.3)	0.0-12.2	2.3 (1.1-4.2)	0.0-17.4
Diagnosis (%)				
• OA	87.1 (83.5-90.8)	42.2-100.0	96.6 (95.5-97.9)	58.6-100.0
• Non-OA	12.9 (9.3-16.5)	0.0-57.8	3.4 (2.1-4.5)	0.0-41.4
1-year revision (%)	1.6 (1.0-2.3)	0.0-7.0	1.1 (0.7-1.6)	0.0-16.7
• Infection*	0.3 (0.1-0.8)	0.0-4.5	0.3 (0.1-0.6)	0.0-8.3
• Loosening**	0.3 (0.1-0.4)	0.0-1.5	0.1 (0.0-0.2)	0.0-1.0
• Dislocation***	0.5 (0.2-0.8)	0.0-2.5		
• Technical failure****			0.3 (0.1-0.5)	0.0-8.3

The values under ''Median (IQR)" indicate the mean or the percentage of the median hospital. The value under ''Range" indicate the highest or lowest mean or percentage of the hospitals. ASA = American Society of Anaesthesiologists, BMI = body mass index, IQR = Interquartile range, OA = osteoarthritis, THA = total hip arthroplasty and TKA = total knee arthroplasty.

*Revision within 1 year because of infection. **Revision within 1 year because of prosthesis loosening; following THA (acetabulum and/or femur or TKA (femur, tibia and/or patella). ***Revision within 1 year because of dislocation following THA only. ****Revision within 1 year because of technical failure (malalignment, instability and/or patella dislocation) following TKA only.

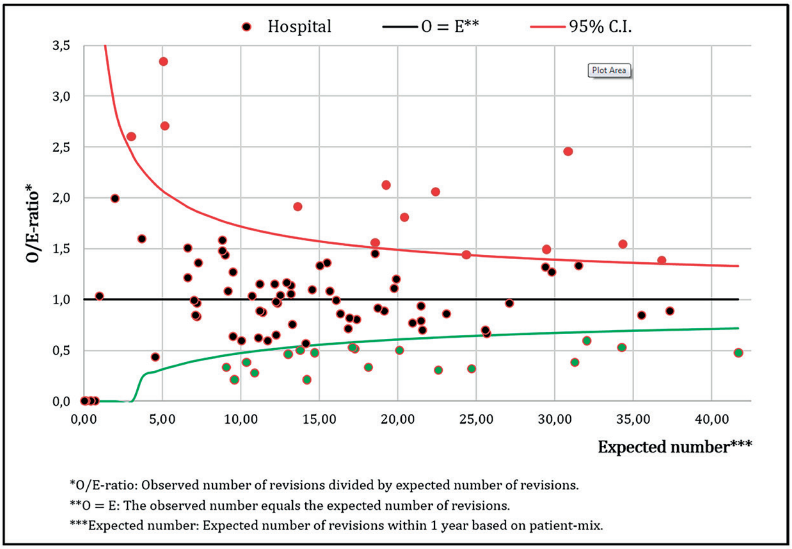


Figure 1 Funnel-plot of Hospital Variation in 1-year Revisions after THA during 2014-2016.

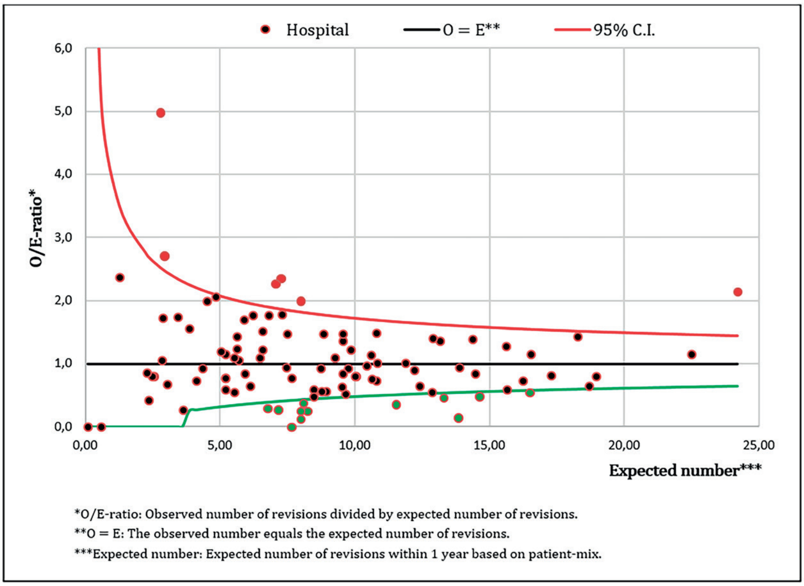


Figure 2 Funnel-plot of Hospital Variation in 1-year Revisions after TKA during 2014-2016.

Hospital variation

Thirteen hospitals performing THA were negative outliers (median O/E ratio, 1.9; IQR, 1.5-2.5), and 18 hospitals were positive outliers (median O/E ratio, 0.4; IQR, 0.3-0.5) (Fig. 1). Seven hospitals performing TKA were negative outliers (median O/E ratio, 2.3; IQR, 2.2-2.8), and 14 hospitals were positive outliers (median O/E ratio, 0.3; IQR, 0.2-0.5) (Fig. 2). Of the 13 negative outliers for THA, 1 hospital was an outlier in all 3 years, 4 were outliers in 2 years, 6 were outliers in 1 year, and 2 were not an outlier in any year (Table II). Of the 7 negative outliers for TKA, 2 hospitals were an outlier in 2 years, 4 hospitals were an outlier in 1 year, and 1 hospital was not an outlier in any single year (Table III). Some outliers in the 3-year period were not an outlier in any single year because only slightly more revisions were performed than expected, causing the difference to become significant only when a lager sample size was available (e.g., hospital 4) (Table II) or because the low volume of procedures at that hospital resulted in wider funnel-plot control limits reflecting the smaller sample size (e.g., hospital 90) (Table II).

Table II Outlier Hospitals with Significantly More Revisions Than Expected Within 1 Year After THA During 2014-2016 According to Year.

	THA outliers (*n*=13 hospitals)			
Hospital	**2014-2016 O/E**	**2014 (O/E)**	**2015 (O/E)**	**2016 (O/E)**
4	1.4			
6	1.5		2.1	1.8
9	2.5	2.2	2.6	2.6
13	1.5	1.8		
14	1.4		2.3	
21	2.1	3.1	1.9	
28	1.8	2.3		1.8
33	2.1		2.1	2.6
37	1.6		2.3	
52	1.9	2.5		
87	2.7			4.9
88	3.3		5.2	
90	2.6			
Median (IQR) for negative outliers	1.9 (1.5-2.5)	2.3 (2.2-2.5)	2.3 (2.1-2.5)	2.6 (1.8-2.6)
Median (IQR) for non-negative outliers	0.9 (0.5-1.1)	0.8 (0.5-1.2)	0.8 (0.5-1.2)	0.8 (0.4-1.4)

*An O/E-ratio is only provided for negative outliers during the 3-year and 1-year periods. IQR = Interquartile range, O/E = observed number of revisions within 1 year divided by the expected number of revisions within 1 year based on patient-mix, and THA = total hip arthroplasty.

Table III Outlier Hospitals with Significantly More Revisions Than Expected Within 1 Year After TKA During 2014-2016 According to Year.

	TKA outliers (*n*=7 hospitals)			
Hospital	**2014-2016 O/E**	**2014 (O/E)**	**2015 (O/E)**	**2016 (O/E)**
9	2.2		2.4	2.4
35	2.3	3.3		
39	2.0			
41	2.3			4.0
87	2.8		8.9	3.8
89	2.7			4.7
95	13.3	43.9		
Median (IQR) for negative outliers	2.3 (2.2-2.8)	23.6 (13.5-33.8)	5.7 (4.0-7.3)	3.9 (3.5-4.2)
Median (IQR) for non-negative outliers	0.8 (0.6-1.2)	0.9(0.4-1.6)	0.9 (0.3-1.3)	0.9 (0.5-1.2)

An O/E-ratio is only provided for negative outliers during the 3-year and 1-year periods. IQR = Interquartile range, O/E = observed number of revisions within 1 year divided by the expected number of revisions within 1 year based on patient-mix, and TKA = total knee arthroplasty.

Of the 13 negative outliers for THA, 8 hospitals had more revisions for infection (with the cup and/or stem being replaced in 13% of cases), 4 had more revisions for dislocation, and 2 had more revisions for prosthesis loosening. Two hospitals had more revisions for both infection and dislocation. For 1 negative outlier, no specific indication was found. Four hospitals had more revisions for infection in multiple years, and 1 hospital had more revisions for dislocation in multiple years (Table IV). Of the 7 negative outliers for TKA, 2 hospitals had more revisions for infection (with the tibial and/or femoral component being replaced in 13% of cases), with 1 hospital having more infections in all single years. Furthermore, 1 hospital had more technical failures (Table V).

Statistical reliability of ranking (Rankability)

Rankability for overall revision during 2014 to 2016 was 62% (moderate) for THAs and 46% (low) for TKAs, indicating that 62% of the observed variation for THAs and 46% for TKAs reflect "true" hospital differences rather than random variation. For THA, the 3-year indication-specific rankabilities were 61% (moderate) for infection, 39% (low) for dislocation, and 32% (low) for loosening. Rankabilities in single years were low (Table VI). For TKA, these values were 43% (low) for infection, 14% (low) for technical failures, and 11% (low) for loosening. Rankabilities for single years were low (Table VI).

Table IV Outlier Hospitals with Significantly More Revisions Than Expected Within 1 Year After THA During 2014-2016 According to Reason for Revision.

	THA outliers (n=13 hospitals)												
Hospital	**All revisions 2014-2016 O/E**	**Revisions for infection (O/E)**				**Revisions for loosening (O/E)**				**Revisions for dislocation (O/E)**			
		2014-2016	**2014**	**2015**	**2016**	**2014-2016**	**2014**	**2015**	**2016**	**2014-2016**	**2014**	**2015**	**2016**
4	1.4					2.4	3.1						
6	1.5	2.5		3.2	3.1								
9	2.5	3.0		3.6	3.7					2.5	3.6		
13	1.5									2.0		2.7	
14	1.4	2.2		3.7									
21	2.1	3.2	4.6	3.3	2.3					2.3	4.6		
28	1.8												
33	2.1	4.5	4.3	3.9	5.1								
37	1.6					4.0		6.7					
52	1.9									3.5	3.9	3.9	
87	2.7	1.6			5.3								
88	3.3	7.3		14.2									
90	2.6	5.1											
Median (IQR) for negative outlier	1.9 (1.5-2.5)	3.1 (2.4-4.7)	4.5 (4.4-4.5)	3.7 (3.4-3.9)	3.7 (3.1-5.1)	3.2 (2.8-3.6)	N/A	N/A	N/A	2.4 (2.2-2.8)	3.9 (3.8-4.3)	3.3 (3.0-3.6)	N/A
Median (IQR) for non-negative outlier	0.9 (0.5-1.1)	0.5 (0.2-1.1)	0.0 (0.0-1.0	0.5 (0.0-1.1)	0.6 (0.0-1.4)	0.8 (0.3-1.2)	0.7 (0.0-1.6)	0.7 (0.0-1.4)	0.5 (0.0-1.6)	0.8 (0.4-1.3)	0.6 (0.0-1.2)	0.8 (0.0-1.3)	0.8 (0.0-1.5)

An O/E-ratios is provided if a hospital was a negative outlier for 1 of the reasons for revision in the period 2014-2016. IQR = Interquartile range, O/E = Observed number of revisions within 1 year divided by the expected number of revisions within 1 year based on patient-mix, THA = total hip arthroplasty, and N/A = not applicable.

Chapter 3

Table V Outlier Hospitals with Significantly More Revisions Than Expected Within 1 Year After TKA During 2014-2016 According to Reason for Revision.

TKA outliers (*n*=7 hospitals)													
Hospital	**All revisions 2014-2016 O/E**	**Revisions for infection (O/E)**				**Revisions for loosening (O/E)**				**Revisions for technical failure (O/E)**			
		2014-2016	**2014**	**2015**	**2016**	**2014-2016**	**2014**	**2015**	**2016**	**2014-2016**	**2014**	**2015**	**2016**
9	2.2	3.2	3.5	2.9	3.4								
35	2.3												
39	2.0												
41	2.3	4.4			7.8								
87	2.8									8.5		17.0	
89	2.7												
95	13.3												
Median (IQR) for negative outliers	2.3 (2.3-2.8)	3.8 (3.5-4.1)	N/A	N/A	5.6 (4.5-6.7)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Median (IQR) for non-negative outliers	0.8 (0.6-1.2)	0.8 (0.3-1.4)	0.0 (0.0-1.6)	0.7 (0.0-1.6)	0.6 (0.0-1.3)	0.7 (0.0-1.5)	0.0 (0.0-1.4)	0.0 (0.0-0.0)	0.0 (0.0-1.6)	0.9 (0.4-1.4)	0.0 (0.0-1.6)	0.6 (0.0-1.5)	0.8 (0.0-1.6)

An O/E-ratios is provided if a hospital was a negative outlier for 1 of the reasons for revision in the period 2014-2016. IQR = Interquartile range, O/E = Observed number of revisions within 1 year divided by the expected number of revisions within 1 year based on patient-mix, TKA = total knee arthroplasty, and N/A = not applicable.

Table VI Rankabilities of 1-Year Revision Rates per Reason for Revision during 2014-2016 According to Year.

	1-year reason for revisions	2014-2016	2014	2015	2016
THA	Infection	61%	34%	36%	31%
	Prosthesis loosening	32%	23%	28%	23%
	Dislocation	39%	29%	18%	15%
TKA	Infection	43%	28%	19%	39%
	Prosthesis loosening	11%	15%	25%	8%
	Technical failure	14%	6%	10%	13%

THA = total hip arthroplasty, TKA = total knee arthroplasty.

Discussion

The present study demonstrated large variation in overall 1-year revision rates as well as indication-specific 1-year revision rates between Dutch hospitals after primary THA and TKA. For THA, 13 hospitals performed significantly worse than expected and 18 hospitals performed significantly better than expected during 2014-2016. For TKA, these values were 7 and 14 hospitals, respectively. Eleven of the 13 negative outliers for THA and 6 of the 7 outliers for TKA were an outlier for 1 or more single years. When 3 years of data on specific indications were included, worse performance was identified for 12 of the 13 outlier hospitals for THA and 3 of the 7 outlier hospitals for TKA, with the specific indications consisting mainly of infection (after both THA and TKA) as well as dislocation (after THA). These findings are starting points for further in-depth hospital-based investigations to improve quality of care. When 3 years of data were used, rankability was moderate for THA and low for TKA. For single years of data, rankabilities were all low. When 3 years of data on specific indications were used, rankabilities were low for both THA and TKA, with the exception of infection after THA, for which rankability was moderate. Therefore, we would recommend using 3 years of data to reliably rank hospitals on their revision rates and identify areas for improvement.

Comparison with the literature

Neither hospital variation nor the statistical reliability of ranking hospitals on 1-year revision rates after primary THA and TKA have been described previously, to our knowledge. However, hospital variation for other outcomes has been described. U.S. studies have shown large variation in terms of the rates of complications (range, 1.8% to 9.0%) and blood transfusions (median, 15.9% [IQR, 5.4% to 26.2%] for THA and 11.0% [IQR, 3.5% to 18.5%] for TKA)(2,9). A Dutch study found large variation in allogeneic blood transfusions (O/E-ratio range, 0.0-4.4) but also showed that the reliability of ranking was only 34% for THA and 21% for TKA(6).

Hofstede et al. showed large between-hospital variation in the number of patients with a prolonged length of hospital stay and the number of acute readmissions after THA and TKA, with high and low rankability for these outcomes, respectively (16). For other outcomes and diseases, rankability was mostly moderate at best. Lingsma et al. found a rankability of 55% for a “poor” outcome after acute brain ischemia, and Henneman et al. found a rankability of 35% for mortality after colorectal cancer surgery(11,14). Only for surgical site infection after colectomy was a high rankability previously found(16). The present study demonstrated moderate rankability for revision after THA, which seems in line with previous findings.

Strengths and limitations

A strength of the present study is that a large population-based registry with completeness of >98% was used(21,22). However, low 1-year revision rates make detection of hospital differences difficult. In addition, low-volume hospitals are more difficult to monitor compared with high-volume hospitals because they can escape the outlier status in the funnel plot as a result of the higher uncertainty embedded in their revision estimates. This is reflected in a large range between upper and lower control limits for low-volume hospitals (on the left side of the funnel plot), which means that a higher O/E-ratio is needed to become an outlier compared with high-volume hospitals (on the right side of funnel plot). However, in this cohort, only 6 hospitals performed <25 THAs per year and 3 hospitals performed <25 TKAs per year.

The limitations of the present study are that limited patient variables are collected in registries, causality cannot be proven due to its observational nature, and there may be underreporting of revisions. However, given the high completeness of the LROI database for primary and revision procedures (currently exceeding 98% and 96%, respectively), underreporting is unlikely to affect the results(21,22). Only if a revision were performed abroad would it be missed in the registry. Another limitation may be underreporting of periprosthetic joint infections as incision and drainage is not reported when no prosthesis component is exchanged or removed(25,26). In the Netherlands, the exchange of loose prosthetic components during acute periprosthetic joint infection became standard care in 2015, which likely resulted in some underreporting before 2015(27). Limited data on patient characteristics were available, thereby limiting the possibility of case-mix correction. This factor might have resulted in overestimation of the differences now attributed to quality of care. For example, diabetes mellitus or the use of immunosuppressants influence the revision risk, but data on these factors were not available(28).

Implementation and further research

Determining hospital variation and detecting outliers seem to be a simple and efficient way to get insight into hospital performance, provided that this variation reflects "true" hospital differences with minimal random variation(12,15,29,30). A categorization for rankability is still arbitrary but is suggested to be good when >75%(13,14,31,32). Our results showed moderate rankability for THA for both the overall revision rate (62%) and infection (61%) on the basis of 3 years of data. For single years, the rankabilities were all low, making these outcomes less useful for reliably assessing performance in practice.

In 2017, the Dutch Orthopaedic Association started a procedure to detect possible negative outlier hospitals and to discuss activities to improve care, resulting in a customized plan of improvement(33). The present study used the same outlier procedure, while adding indications for revision as a method to direct quality-improvement initiatives. Most outliers for overall revision were also outliers for a specific indication for revision, so adding these analyses seems a useful addition to direct improvement activities. Although part of the variation may be surgeon-related (e.g., surgical approach), the current approach is likely to provide a broader overview to improve the quality of care given (e.g., individual surgeons are dependent on their team to achieve the best outcomes). Furthermore, having more revisions for infection is likely not surgeon-related but rather is due to antibiotic prophylactic protocol and implementation, wound-care policy, preoperative preparation of the patient with chlorhexidine, and so on.

In the future, other outcome measures associated with quality of care after THA and TKA should be used, such as length of hospital stay, readmissions, and patient-reported outcome measures, provided that these outcomes have sufficient rankability. The use of moving periods of 3-year data over time may be useful for follow-up research because this method reliably examines the outcome over several years. Furthermore, another option might be to combine outcomes in an ordinal composite measure, which would improve rankability and thereby the ability to detect "true" hospital differences while also providing a broader quality-of-care perspective than 1 single indicator(16). In addition, other methods such as statistical process control techniques may enable detection of worsening performance sooner and thereby also improve the quality of care.

Conclusion

Large variation in 1-year revision rates after THA and TKA between Dutch hospitals was identified. The majority of outliers could be linked to a specific indication for revision, which gives clear starting points for quality-improvement initiatives. Earlier detection of worse performance by using a single year of data had low reliability.

References

1. Siciliani L, Sivey P, Street A. Differences in length of stay for hip replacement between public hospitals, specialised treatment centres and private providers: selection or efficiency? Health Econ. 2013;22(2):234-42.
2. Menendez ME, Lu N, Huybrechts KF, Ring D, Barnes CL, Ladha K, et al. Variation in Use of Blood Transfusion in Primary Total Hip and Knee Arthroplasties. J Arthroplasty. 2016;31(12):2757-63.e2.
3. Solomon DH, Losina E, Baron JA, Fossel AH, Guadagnoli E, Lingard EA, et al. Contribution of hospital characteristics to the volume-outcome relationship: dislocation and infection following total hip replacement surgery. Arthritis Rheum. 2002;46(9):2436-44.
4. So-Osman C, Nelissen R, Te Slaa R, Coene L, Brand R, Brand A. A randomized comparison of transfusion triggers in elective orthopaedic surgery using leucocyte-depleted red blood cells. Vox Sang. 2010;98(1):56-64.
5. Voorn VM, Marang-van de Mheen PJ, Wentink MM, So-Osman C, Vliet Vlieland TP, Koopman-van Gemert AW, et al. Frequent use of blood-saving measures in elective orthopaedic surgery: a 2012 Dutch blood management survey.
6. Voorn VMA, Marang-van de Mheen PJ, van der Hout A, So-Osman C, van den Akker-van Marle ME, Koopman-van Gemert A, et al. Hospital variation in allogeneic transfusion and extended length of stay in primary elective hip and knee arthroplasty: a cross-sectional study. BMJ Open. 2017;7(7):e014143.
7. Fry DE, Pine M, Nedza SM, Locke DG, Reband AM, Pine G. Risk-Adjusted Hospital Outcomes in Medicare Total Joint Replacement Surgical Procedures. J Bone Joint Surg Am. 2017;99(1):10-8.
8. Weeks WB, Kotzbauer GR, Weinstein JN. Using Publicly Available Data to Construct a Transparent Measure of Health Care Value: A Method and Initial Results. Milbank Q. 2016;94(2):314-33.
9. Bozic KJ, Grosso LM, Lin Z, Parzynski CS, Suter LG, Krumholz HM, et al. Variation in hospital-level risk-standardized complication rates following elective primary total hip and knee arthroplasty. J Bone Joint Surg Am. 2014;96(8):640-7.
10. Cram P, Vaughan-Sarrazin MS, Wolf B, Katz JN, Rosenthal GE. A comparison of total hip and knee replacement in specialty and general hospitals. J Bone Joint Surg Am. 2007;89(8):1675-84.
11. Henneman D, van Bommel AC, Snijders A, Snijders HS, Tollenaar RA, Wouters MW, et al. Ranking and rankability of hospital postoperative mortality rates in colorectal cancer surgery. Ann Surg. 2014;259(5):844-9.
12. van Dishoeck AM, Koek MB, Steyerberg EW, van Benthem BH, Vos MC, Lingsma HF. Use of surgical-site infection rates to rank hospital performance across several types of surgery. Br J Surg. 2013;100(5):628-36; discussion 37.
13. Lingsma HF, Eijkemans MJ, Steyerberg EW. Incorporating natural variation into IVF clinic league tables: The Expected Rank. BMC Med Res Methodol. 2009;9:53.
14. Lingsma HF, Steyerberg EW, Eijkemans MJ, Dippel DW, Scholte Op Reimer WJ, Van Houwelingen HC. Comparing and ranking hospitals based on outcome: results from The Netherlands Stroke Survey. QJM. 2010;103(2):99-108.
15. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. BMJ quality & safety. 2011;20(10):869-74.
16. Hofstede SN, Ceyisakar IE, Lingsma HF, Kringos DS, Marang-van de Mheen PJ. Ranking hospitals: do we gain reliability by using composite rather than individual indicators? BMJ quality & safety. 2018.

17. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA. 2004;292(7):847-51.
18. Dutch Arthroplasty Register (LROI). Annual report 2018. www.lroi-report.nl.
19. Swedisch Hip Arthroplasty Register. Annual Report 2017. https://registercentrum.blob.core.windows.net/shpr/r/Eng_Arsrapport_2017_Hoftprotes_final-Syx2fJPhMN.pdf.
20. Swedisch Knee Arthroplasty Register. Annual Report 2018. http://www.myknee.se/en/.
21. van Steenbergen LN, Denissen GA, Spooren A, van Rooden SM, van Oosterhout FJ, Morrenhof JW, et al. More than 95% completeness of reported procedures in the population-based Dutch Arthroplasty Register. Acta Orthop. 2015;86(4):498-505.
22. *LROI website. Completeness of Registering Hospitals and Completeness of Registered Arthroplasties in the LROI Based on the Hospital Information System in 2016,* http://www.lroi-rapportage.nl/data-quality-coverage-and-completeness, accessed febr 2019.
23. Jasper LL, Jones, C. A., Mollins, J., Pohar, S. L., Beaupre, L. A.. Risk factors for revision of total knee arthroplasty: a scoping review. BMC Musculoskelet Disord. 2016;17:182.
24. Rothman KJ, Boice JD. Epidemologic analysis with a programmable calculator. National Institutes of Health. 1979 31-2.
25. Gundtoft PH, Overgaard S, Schonheyder HC, Moller JK, Kjaersgaard-Andersen P, Pedersen AB. The "true" incidence of surgically treated deep prosthetic joint infection after 32,896 primary total hip arthroplasties: a prospective cohort study. Acta Orthop. 2015;86(3):326-34.
26. Veltman ES, Moojen DJF, Nelissen RG, Poolman RW. Antibiotic Prophylaxis and DAIR Treatment in Primary Total Hip and Knee Arthroplasty, A National Survey in The Netherlands. J Bone Jt Infect. 2018;3(1):5-9.
27. Dutch Orthopaedic Association. Protocol: Aanbeveling werkwijze behandeling Prothese Infecties Orthopedie, 2015. www.orthopeden.org.
28. Marchant MH, Jr., Viens NA, Cook C, Vail TP, Bolognesi MP. The impact of glycemic control and diabetes mellitus on perioperative outcomes after total joint arthroplasty. J Bone Joint Surg Am. 2009;91(7):1621-9.
29. Adab P, Rouse AM, Mohammed MA, Marshall T. Performance league tables: the NHS deserves better. BMJ. 2002;324(7329):95-8.
30. Cram P, Cai X, Lu X, Vaughan-Sarrazin MS, Miller BJ. Total knee arthroplasty outcomes in top-ranked and non-top-ranked orthopedic hospitals: an analysis of Medicare administrative data. Mayo Clin Proc. 2012;87(4):341-8.
31. Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods. 2006;11(2):193-206.
32. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.
33. CommisionQuality. Commision Quality (Collaboration between Dutch Orthopaedic Association and Dutch Arthroplasty Registry). Protocol: Quality procedure, 2017. www.orthopeden.org.