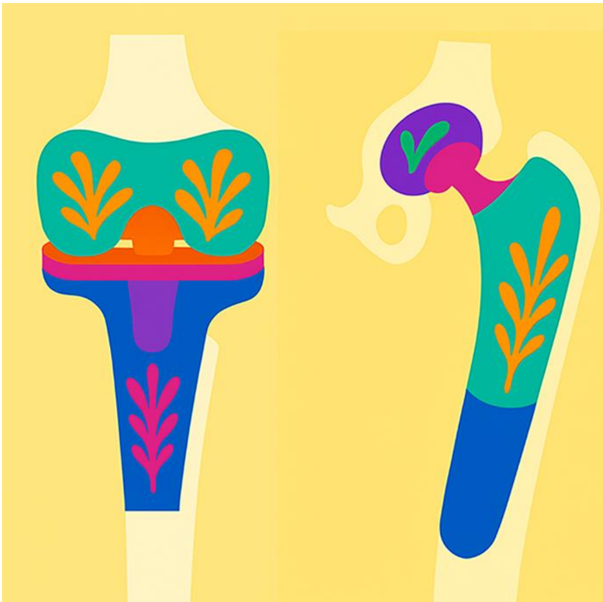
Chapter 8 – The reliability of revision rates following primary shoulder arthroplasty as a quality indicator to rank hospital performance: a national registry analysis including 13,104 shoulders and 87 hospitals



Lotje A. Hoogervorst[1,2], Peter van Schie[1,2], Jochem Nagels[1], Rob G.H.H. Nelissen[1], Perla J. Marang-van de Mheen[2]

[1] Department of Orthopaedics, Leiden University Medical Center, Leiden, the Netherlands
[2] Department of Biomedical Data Sciences & Medical Decision Making, Leiden University Medical Center, Leiden, the Netherlands

J Shoulder Elbow Surg. 2023 Jan;32(1):59-67

Abstract

Background: To assess the extent of between-hospital variation in revision following primary shoulder arthroplasty (SA), both overall and for specific revision indications to guide quality improvement initiatives, and to assess whether revision rates are suitable as quality indicators to reliably rank hospital performance.

Methods: All primary SAs performed between 2014 and 2018 were included from the Dutch Arthroplasty Register to examine 1-year revision and all primary SAs performed between 2014 and 2016 for 1 and 3-year revisions. For each hospital, the observed number (O) of revisions was compared with that expected (E) based on case-mix and depicted in funnel plots with 95% control limits to identify outlier hospitals. The rankability (i.e. the reliability of ranking hospitals) was calculated as the percentage of total hospital variation due to true between-hospital differences rather than chance and categorised as low (<50%), moderate (50%-75%), and high (>75%).

Results: A total of 13,104 primary SAs (87 hospitals) in 2014-2018 were included, of which 7,213 were performed between 2014 and 2016. Considerable between-hospital variation was found in 1-year revision in 2014-2016 (median 1.6%, interquartile range 0.0%-3.1%), identifying three outlier hospitals having overall significantly more revisions than expected (O/E range 1.9-2.3) and for specific indications (cuff pathology and infection). Results for 2014-2018 were similar. For 3-year revision, three outlier hospitals were identified (O/E range 1.7-3.3). Rankabilities for all outcomes were low.

Conclusion: Considerable between-hospital variation was observed for 1 and 3-year revision rates following primary SA, where outlier hospitals could be identified based on large differences in revision for specific indications to direct quality improvement initiatives. However, rankabilities were low, meaning that much of the other (smaller) variation in performance could not be detected, rendering

revisions unsuitable to rank hospital performances following primary SA.

Background

The number of primary shoulder arthroplasties (SAs) is expected to increase in the coming decades predominantly because of ageing of the population.[1,4,10] If complication risks remain the same, the absolute number of complications (e.g. revisions) will likely increase. This increase will have an impact on patients and have economic and societal ramifications. Therefore, an essential part of clinical practice should involve reducing complication risks by continuously improving the quality of care and thereby minimising patients' morbidity and unnecessary costs. Such continuous quality improvement can be done in two ways: i) individual hospitals can monitor their performances over time and examine if their outcomes are improving or deteriorating, and ii) hospitals performing worse than other hospitals while treating similar patients, might be studied to guide improvement initiatives based on the practices of better-performing hospitals.

Several studies have shown variation in clinical outcomes (e.g. revisions and readmissions) between hospitals performing SA, suggesting that improvement is possible.[12,16,20] Revision surgery is a major complication following primary SA and the most frequently used clinical outcome to monitor the quality of care delivered.[11,13,18] However, variation in overall revision rates do not give sufficient information to guide where to improve care as revisions may be performed for different indications (e.g. infection or surgical technique failure) each aligned with different quality improvement initiatives. The nationwide between-hospital variation in overall revision and for specific indications has not been assessed for primary SA.

Variation between hospitals is often shown in funnel plots, where the "funnel" is wider on the left side reflecting hospitals with smaller volume of patients. Revision rates from hospitals only performing a few annual procedures will be rather unreliable, making it harder to detect differences in performance. Even though we can identify some hospitals with relatively large differences in revision rate outside the upper control limit of the funnel plot, the question is also whether we can sufficiently discriminate the performance among all hospitals, particularly because revision rates are mostly low, making it harder to detect differences. The rankability has been previously used to assess the reliability of ranking hospitals and represents the percentage of total hospital variation in, for example, revision that is due to true between-hospital differences rather than chance.[5,7,22,23] Previous studies have shown that the rankability varies between 0% and 71% depending on the outcome and patient population,[22] and was 62% for 1-year revision after total hip arthroplasty (THA) and 46% after total knee arthroplasty (TKA), which means that 62% of the total variation in 1-year revisions after THA and 46% after TKA is due to true hospital differences rather than chance.[23] The rankability for revision rates after primary SA is currently unknown, but relevant to assess the overall usefulness of revisions as a reliable nationwide quality indicator. However, because of the low number of events (revisions), the rankability may be low, so that using multiple years of data or extending the duration of patients' follow-up may be a solution, but it is unknown whether this will result in significantly better rankability.

Therefore, this study aimed to assess the extent of between-hospital variation for both 1 and 3-year overall revision rates following primary SA in the Netherlands, as well as for specific indications for revision, to guide subsequent quality improvement initiatives for hospitals with significantly worse performance. Second, we evaluated whether 1 and 3-year revision rates could be used as reliable nationwide quality indicators to rank hospital performance.

Methods

Study design and setting

This observational study used nationwide data from the Dutch Arthroplasty Register (LROI). Within the Dutch health care system, there is mandatory standard health insurance that covers costs of, for example, hospital treatment and therefore also both primary and revision SA, and the optional additional insurance to cover costs not included in the standard package. All orthopedic surgeons are first trained for 1.5 years in general surgery and then 4.5 years in orthopedic surgery, before they are allowed to perform SAs.

Patients' baseline demographics and surgical details are collected in the LROI by surgeons for all patients undergoing a primary or revision SA in the Netherlands using predefined forms.[21] Data completeness is currently above 96% (primary SA) and 91% (revision SA).[9] The LROI uses barcode scanning of prosthetic components for traceability and linkage with the Dutch national insurance database on health care to determine patients' vital status (dead/alive).[3] The medical ethical committee of the Leiden University Medical Center waived the need for ethical approval under Dutch law (P.15.198).

Procedure and definitions

All primary SAs performed in the Netherlands (January 1, 2014, to December 31, 2018) were included. Available patients' baseline characteristics were age, sex, body mass index, current smoking status (yes/no), American Society of Anesthesiologists classification (I, II, III-IV), Walch classification (A1, A2, B1, B2, B3, C, not applicable), primary SA indication, categorised into: i) osteoarthritis; ii) rotator cuff disease (rotator cuff arthropathy; iii) and/or irreparable rotator cuff tears); iv) trauma (acute and/or post-traumatic fractures); v) other (osteonecrosis, rheumatoid arthritis, inflammatory arthritis, primary tumour, metastasis, and/or ''other'').

Outcome variables
Revision is defined as any replacement, removal, or addition of at least one prosthetic component. Surgeons registered at least one indication for every revision, categorised into three groups: i) infection; ii) surgical technique failure (malalignment, instability, periprosthetic fracture, loosening of the glenoid component, and/or loosening of the humeral component), and iii) cuff pathology (cuff rupture and/or cuff arthropathy).

For the 2014-2016 cohort, overall and specific indication for 1 and 3-year revision rates were calculated. The 1-year revision rates were also calculated for all shoulder surgeries (2014-2018). In this way, the impact of including more years of data on the reliability of ranking 1-year revision rates (2014-2018 *versus* 2014-2016) and the impact of extending follow-up to 3- rather than 1-year revisions (for the 2014-2016 cohort) could be assessed. To ensure fair comparison, hospitals without procedures performed in 2014-2016 were excluded from all analyses.

Statistical analysis
Descriptive statistics were used to examine hospital-level variations in patients' baseline characteristics and the crude 1 and 3-year revision rates for the 2014-2016 and 2014-2018 cohorts.

For each included patient in the 2014-2016 cohort, the expected 1-year revision risk was calculated using logistic regression analysis, with the above patients' baseline characteristics as the independent variables and revision as the dependent variable. In case of missing patients' baseline characteristics (<9% for all variables, Supplementary Table 1 – online available), the mean for numeric variables or the mode for categorical variables was used to impute missing values. These expected risks per patient were summed to get the expected number of revisions for each included hospital. The observed (O) number of revisions can then be compared with the expected (E) number to obtain the observed *versus* expected (O/E) ratios for each hospital, and depicted in funnel plots with 95% control limits. Hospitals with ratios above the upper limit had significantly higher revision rates than expected (negative outliers). In contrast, hospitals with ratios below the lower limit had significantly lower revision rates than expected (positive outliers), and hospitals within the control limits performed as expected (nonoutliers). These analyses were repeated for the three revision indications groups. The same analyses were carried out for the 2014-2018 cohort and for 3-year revisions (2014-2016 cohort). Analyses were performed using SPSS, version 25 (IBM, Armonk, NY, USA).

The rankabilities (i.e. the reliabilities of ranking) were calculated for each cohort and outcome, as the percentage of between hospital variation relative to the total variation, using the following formula:[22,23]

$$\text{rankability} = \frac{\text{between} - \text{hospital variation } (t^2)}{\text{between} - \text{hospital variation } (t^2) \ + (\text{within} - \text{hospital variation } (\sigma^2)}$$

The between-hospital variation was calculated using random effects logistic regression models, with hospital as random effect and the above patients' baseline characteristics as fixed effects. The within-hospital variation was calculated using fixed effect logistic regression models, including hospital and the patients' baseline characteristics as fixed factors. The median squared standard error of the hospital estimate represented the within hospital variation. Rankabilities were categorised into three levels: i) low (<50%); ii) moderate (50%-75%), and iii) high (>75%).[22] Analyses were performed using Stata, version 14 (StataCorp, College Station, TX, USA).

Results

Ninety-one hospitals performed primary SA during 2014-2018. Of these, four hospitals had no procedures in 2014-2016 and were excluded. Overall, 13,104 primary SAs performed in 2014-2018 (in

87 hospitals) were included, of which 7,213 were performed between 2014-2016. Of all procedures, 9,796 (75%) were performed in females, and the mean age was 71 years (standard deviation 9.6). The reverse total SA was the most frequently performed type of arthroplasty (64%), followed by total SA (22%) and hemiarthroplasty (14%). Mean overall patient-level 1 and 3-year revision rates were 2.4% and 4.3%, respectively. Surgical technique failure was the most common revision indication, occurring in 1.4% and 2.3% of procedures within 1 and 3-years, respectively.

Between-hospital differences

The clinical diversity of patients' baseline characteristics is demonstrated in Table 1. There was considerable variation in outcomes as shown by the interquartile range (IQR), with the median hospital having a crude 1-year revision rate of 1.6% (IQR: 0.0%-3.1%) in the 2014-2016 cohort and 1.8% (IQR: 0.9%-3.0%) in 2014-2018. The median hospital had a crude 3-year revision rate of 3.9% (IQR: 2.1%-6.3%). After adjustment for differences in patient-mix, there was still considerable variation as shown in the funnel plot in Figure 1, where small volume hospitals appear on the left side of the *x* axis and larger volume hospitals on the right side.

Including the 2014-2016 cohort, three hospitals were identified as negative outliers regarding their performance on 1-year revision (O/E ratios: 1.9, 2.3, and 2.3; hospital one, two, and nine, respectively) (Table 2; red dots in Figure 1). Two of these hospitals were a negative outlier for one single year (data not shown). We were able to give direction for quality improvement for all negative outlier hospitals as they were outliers for specific revision indications: all hospitals had more revisions for cuff pathology and two also for infection (hospitals one and nine). Mean annual number of performed procedures for these outlier hospitals in 2014-2016 was 105 (hospital one), 78 (hospital two), and 66 (hospital nine). No positive outliers were identified.

Extending the follow-up period to 3-year revision, identified three negative outliers (O/E ratios: 1.7, 2.7, and 3.3; hospital 2, 22, and 73, respectively) (Table 2). One hospital was a negative outlier in one single year (data not shown). Only one negative outlier hospital was also an outlier for a specific revision indication (hospital 22, cuff pathology). Mean annual number of performed procedures for these outlier hospitals in 2014-2016 was 78 (hospital 2), 36 (hospital 22), and 13 (hospital 73). Figure 2 shows the performance of hospitals on 3-year revision, with the red dots indicating the three hospitals with significantly worse 1-year revision rates (from Figure 1), showing that from these three hospitals, one hospital was also an outlier on 3-year revision (hospital two). In addition, the green dots in Figure 2 indicate the two hospitals with significantly better 3-year revision rates (i.e. positive outliers).

Increasing the sample size by including more patients (2014-2018) identified three negative outliers (O/E ratios: 1.6, 2.0, and 2.5; hospital 1, 4, and 13, respectively) and three positive outliers for 1-year revision (Table 2 and Figure 3). All hospitals were also an outlier in one single year (data not shown). All negative outlier hospitals were outliers for specific revision indications, making it possible to direct quality improvement initiatives: all hospitals had more revisions for infection and one also for both cuff pathology and surgical technique failure (hospital four). Mean annual number of performed procedures for these outlier hospitals in 2014-2018 was 133 (hospital 1), 68 (hospital 4), and 45 (hospital 13). Figure 3 shows the performance of hospitals in 2014-2018 on 1-year revision, with the red dots indicating the three hospitals with significantly worse 1-year revision rates in 2014-2016 (Figure 1), showing that from these three hospitals, one hospital remained a negative outlier for 2014-2018 (hospital one). The green dots in Figure 3 indicate the two hospitals with significantly better 3-year revision rates (from Figure 2), showing that one of these two hospitals also had significantly better 1-year revision rates during

2014-2018.

Rankability

Rankability for 1-year revision, including patients from 2014-2016, was low (26%). This means that even though several outlier hospitals were identified, the majority of variation depicted in Figure 1 is due to random variation, likely because of the large number of low-volume hospitals on the left side of the funnel plot. Increasing the sample size by combining more years (2014-2018) resulted in the expected lower within-hospital variation (the hospital revision rates can be estimated more precisely), but the between-hospital variation was also lower, which resulted in a slightly lower rankability (23%). This was mainly caused by the low between-hospital variation in 2017 (Figure 4). Extending the follow-up period by examining 3-year revision rates, the rankability increased to 32% but is still classified as low. Rankabilities for single years were all classified as low, for both 1 and 3-year revision rates (median of 10% *versus* 25%, respectively; data not shown). Because revision rates for low-volume hospitals will be fairly uncertain, this may contribute to the relatively large random variation and thereby low reliability to rank hospitals (rankability). Therefore, as a sensitivity analysis, we excluded hospitals based on their mean annual volume of procedures (Supplementary Table 2 – online available). In general, rankability increased but with fewer hospitals included and was still classified as low in most cases. Moderate rankabilities could only be achieved when including hospitals with at least 50 procedures annually (n=13 hospitals) (Supplementary Table 3 – online available).

Discussion

To our knowledge, this is the first nationwide registry-based study reporting the between-hospital variation and reliability of ranking hospitals on revision rates following primary SA. Considerable between-hospital variation for both 1 and 3-year revision rates following primary SA was observed, with seven outlier hospitals identified with relatively large differences in revision compared with the other hospitals and/or sufficiently high annual volume of procedures. For all negative outliers on 1-year revision, we were able to direct quality improvement initiatives, as they were also identified as outlier for one or more specific revision indications. However, overall rankabilities were low regardless of increasing the number of years or extending the follow-up period (range 23-32%), meaning that at best only 32% of the observed variation reflects ''true'' differences in hospital performance. Only when including the larger volume hospitals (>50 procedures annually), moderate rankability could be achieved for 1-year revision, but this would exclude the majority of hospitals. Thus, 1 and 3-year revision rates following primary SA do not seem useful as reliable nationwide quality indicators to rank hospital performance as much of the smaller differences in performance will remain undetected. However, both outcomes are still relevant for individual hospitals to monitor their own performances over time and to implement quality improvement initiatives if performance deteriorates.

Revision procedures are considered one of the most important clinical outcomes following SA. In addition, it is a widely accepted endpoint for treatment failure and relatively easy to monitor based on the date of revision surgery. Hence, almost every SA registry reports revision surgery as their primary outcome.[13] However, surgeons across different hospitals could have different thresholds for performing specific revision procedures, influencing these revision rates. Furthermore, despite the advantages of using revision procedures as the primary outcome, our study demonstrated that 1 and 3-year revision rates are not well-suited nationwide quality indicators to reliably discriminate hospitals in their performance because of the low rankabilities. In other words, the estimated revision rate for a particularly low-volume hospital is surrounded by a rather large confidence interval, making it harder to detect a statistically significant difference with another hospital. This is illustrated by the fact that five of the seven negative outlier hospitals identified performed at least 50 procedures annually. The other two outlier hospitals were identified on their 3-year revision rates, indicating that for these lower volume hospitals it takes more time for deteriorating to be detected and quality improvement can be initiated. This also raises the question whether centralisation of SA procedures is warranted, to enable more reliable assessment of hospital performance. In addition, higher-volume hospitals have been shown to be associated with a decreased occurrence of intraoperative blood loss, procedure time, and in-hospital long length of stay following SA.[19]

A previous study from our group by van Schie et al,[23] investigating variation in 1-year revision rates following primary THA and TKA in the Netherlands, found higher rankabilities during 2014-2016 than the present study (62% for THA and 46% for TKA *versus* 26% for SA). Given that the median hospital-level 1-year revision rates were similar or lower than in the present study (1.6% for THA and 1.1% for TKA *versus* 1.6% for SA), the higher rankability is likely due to the higher number of THA and TKA performed in hospitals compared with primary SA, so that the hospital revision rate can be estimated more precisely and thereby between-hospital differences more easily detected. This seems supported by the fact that annual rankabilities for THA and TKA, when volumes are lower, were also all low. The present study adds to this literature that extending the follow-up period by using 3-year instead of 1-year revision rates may slightly increase the rankability (32% *versus* 26%, respectively), although in this case it was still classified as low.

Besides revision procedures, many other clinical outcomes measure aspects of the quality of care delivered (e.g. readmissions or patient-reported outcome measurements (PROMs)). PROMs are increasingly used in orthopedics to guide clinical decision making and will likely become an integral part of health care quality improvement.[15,24] In addition, the majority of SA registries are already capturing PROMs.[13] The advantage of using improvement in

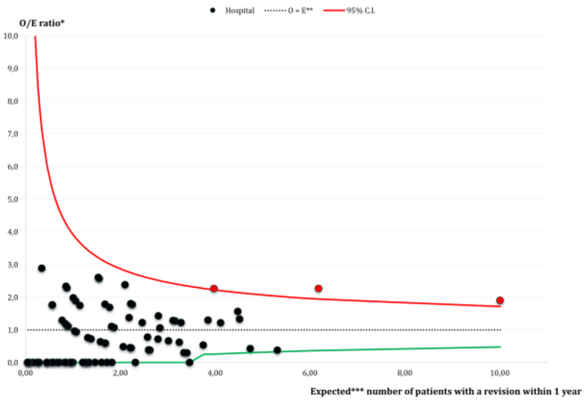
PROMs between the preoperative and postoperative measurement as a quality indicator is that all SA patients will potentially have a difference in PROMs score between these two measurements on a continuous scale, rather than a small group experiencing unintended consequences like revision. The power to detect differences is likely higher for a continuous rather than a dichotomous outcome, so that improvement in PROMs could be a more reliable quality indicator than the revision rate. However, it is challenging to achieve sufficiently high response rates in patients filling in the PROMs questionnaires so that the improvement in PROMs can be calculated for all patients, which might impede the generalizability of the results.[6,9,14,17] For example, in 2019, the LROI achieved only a 51% response rate for preoperative, 37% for 3-month postoperative, and 46% for 12-month postoperative measurements, making calculations of improvements based on these questionnaires unrepresentative for the total population who underwent SA.[9] Furthermore, it is debated whether improvement in PROMs reflects the quality of care delivered as it is also influenced by many other variables, particularly when considering the 12-months postoperative measurements. This means that other indicators such as readmissions or complications may be more suitable to monitor the quality of care delivered following primary SA.

A strength of this study is that it is the first nationwide registry-based study with a large and representative cohort of patients undergoing primary SA in the Netherlands. However, some limitations remain. First, there may be residual confounding as we were limited in the patient characteristics registered in the LROI, whereas more factors (e.g. diabetes mellitus) may affect the likelihood of SA revisions.[8] Second, revisions may have been underreported if procedures were performed outside the Netherlands and thus not included in the LROI; the LROI captures revisions performed in other Dutch hospitals through barcode scanning. Third, hospitals with longer waiting times for revision procedures may appear to perform better (i.e. lower revision rates) than hospitals with shorter waiting times. Fourth, surgeons may be operating in more than one centre, which may influence the hospital's performance if a surgeon performs much better or worse than other surgeons in that hospital. Analysing outcomes (e.g. revision rates) on both hospital and surgeon-level, which is already reported in some registries (e.g. the National Joint Registry), may provide a more in-depth overview of this. However, numbers will be even lower at surgeon-level and has been shown to result in low power to detect significantly higher readmission rates after THA and TKA, whereas power to detect differences at hospital level was excellent.[2]

Conclusion

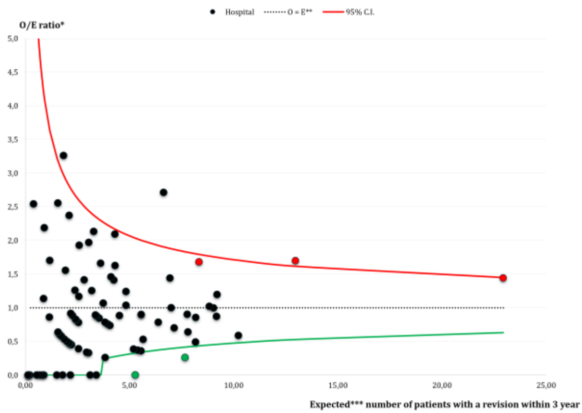
Considerable between-hospital variation in both 1 and 3-year revision rates following primary SA in the Netherlands was observed, where outlier hospitals could be identified based on relatively large differences in revision compared with other hospitals and/or a sufficiently high annual volume of procedures. Low-volume hospitals were only identified as outliers for 3-year revision, so that it takes long before quality improvement can be initiated. Given the low rankabilities for all outcomes, 1 and 3-year revision rates following primary SA do not seem reliable nationwide quality indicators to rank hospital performances, but can still be used for individual hospitals to monitor their own performances over time.

Figures and Tables



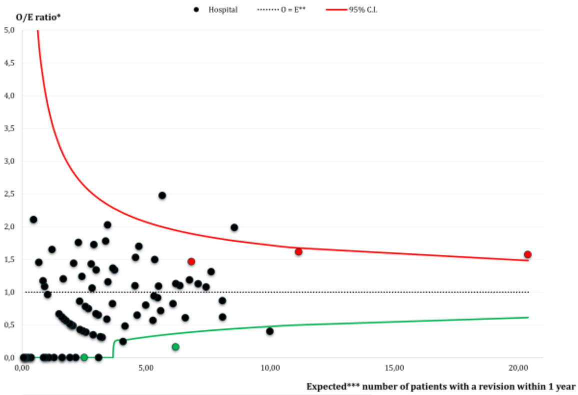
***O/E ratio:** Observed number of revisions divided by expected number of revisions
****O = E:** The observed number equals the expected number of revisions
******Expected:** Based on prediction for revision per person

Figure 1. *Funnel plot of hospital variation in 1-year revision following primary shoulder arthroplasty (2014-2016)*



***O/E ratio:** Observed number of revisions divided by expected number of revisions
****O = E:** The observed number equals the expected number of revisions
******Expected:** Based on prediction for revision per person

Figure 2. *Funnel plot of hospital variation in 3-year revision following primary shoulder arthroplasty (2014-2016)*



***O/E ratio:** Observed number of revisions divided by expected number of revisions
****O = E:** The observed number equals the expected number of revisions
*******Expected:** Based on prediction for revision per person

Figure 3. *Funnel plot of hospital variation in 1-year revision following primary shoulder arthroplasty (2014-2018)*

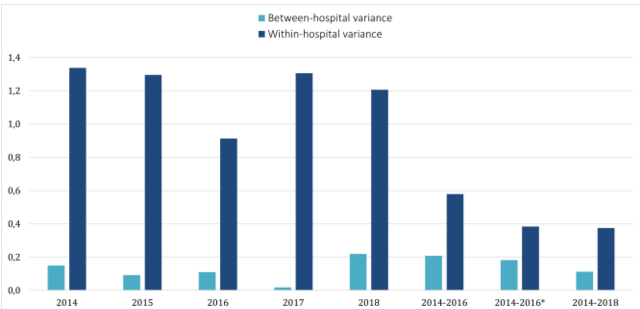


Figure 4. *Between- and within-hospital variances (2014-2018). * = 3-year revision*

	2014-2016 (87 hospitals; 7,213 patients)		2014-2018 (87 hospitals; 13,104 patients)	
	Median (IQR)	**Range**	**Median (IQR)**	**Range**
Number of procedures	63 (37.5-111.5)	1-404	120 (74-202.5)	1-754
Mean age (years)	71.6 (69.5-73.0)	60.0-78.7	71.3 (70.3-72.7)	60.0-80.9
Male sex (%)	24.8 (19.4-29.7)	0.0-100	24.8 (20.6-28.8)	12.5-100
Mean BMI	28.2 (27.6-28.8)	23.6-30.2	28.2 (27.8-28.8)	23.6-30.4
Smoking (%)	12.9 (8.9-16.3)	0.0-50.0	11.7 (9.7-14.5)	0.0-40.0
ASA classification (%)				
I	7.1 (4.0-10.1)	0.0-66.7	7.0 (4.5-10.3)	0.0-42.9
II	63.1 (56.9-70.3)	33.3-100	61.5 (55.5-68.5)	34.7-100
III-IV	29.6 (18.8-36.6)	0.0-55.3	31.4 (22.6-38.3)	0.0-60.4
Walch classification (%)				
A1	62.4 (45.8-72.8)	12.9-100	55.6 (43.8-64.3)	17.4-100
A2	19.9 (13.1-25.6)	0.0-54.8	22.1 (14.5-27.8)	0.0-3.3
B1	10.5 (5.4-18.6)	0.0-41.7	9.9 (6.3-15.0)	0.0-32.4
B2	3.4 (0.2-5.7)	0.0-25.0	3.8 (2.0-6.5)	0.0-20.0
B3	0.0 (0.0-1.6)	0.0-11.8	0.9 (0.0-2.0)	0.0-8.9
C	0.0 (0.0-0.0)	0.0-6.5	0.0 (0.0-1.1)	0.0-7.6
N/A, no osteoarthritis	0.0 (0.0-1.0)	0.0-25.0	4.0 (0.9-8.6)	0.0-60.9
Indication for arthroplasty (%)				
Osteoarthritis	42.7 (35.4-58.3)	0.0-100	42.9 (36.3-56.3)	0.0-100
Trauma*	23.0 (16.5-31.8)	0.0-100	24.2 (15.9-30.8)	0.0-100
Rotator cuff disease†	19.6 (10.8-28.0)	0.0-51.6	22.3 (11.1-28.4)	0.0-48.3
Other‡	7.3 (3.6-11.8)	0.0-43.8	6.9 (4.0-10.4)	0.0-51.7

Procedure type (%)				
Reverse TSA	57.1 (43.7-70.9)	0.0-100	61.0 (50.4-73.2)	0.0-92.4
TSA	20.3 (10.4-31.1)	0.0-66.7	22.0 (10.4-29.7)	0.0-66.7
Hemiarthroplasty	16.3 (8.0-30.5)	0.0-100	12.5 (6.3-24.1)	0.0-100
1-year revision (%)	1.6 (0.0-3.1)	0.0-7.1	1.8 (0.9-3.0)	0.0-6.2
Surgical technical failure§	0.0 (0.0-2.1)	0.0-7.1	1.0 (0.0-2.0)	0.0-5.3
Infection¥	0.0 (0.0-0.0)	0.0-3.7	0.0 (0.0-0.5)	0.0-2.8
Cuff parthology‖	0.0 (0.0-0.0)	0.0-3.4	0.0 (0.0-0.4)	0.0-1.9
3-year revision (%)	3.9 (2.1-6.3)	0.0-17.0		
Surgical technical failure§	2.0 (0.0-3.5)	0.0-12.5		
Infection¥	0.0 (0.0-1.0)	0.0-4.8		
Cuff parthology‖	0.0 (0.0-1.4)	0.0-12.5		

*BMI = body mass index; ASA = American Society of Anesthesiologists; TSA = total shoulder arthroplasty; IQR = interquartile range. The median (IQR) values indicate the mean or the percentage of the median hospital. The range values indicate the highest or lowest mean or percentage of the hospitals. * = indication for primary shoulder arthroplasty was trauma (acute fracture or post-traumatic); † = indication for primary shoulder arthroplasty was rotator cuff disease (rotator cuff arthroplasty or irreparable rotator cuff tears); ‡ = indication for primary shoulder arthroplasty was "other" (osteonecrosis, rheumatoid arthritis, inflammatory arthritis, primary tumor or metastasis); § = revision because of surgical technique failure (malalignment, instability, periprosthetic fracture, loosening glenoid component, and/or loosening humeral component); ¥ = revision because of infection; ‖ = revision because of cuff pathology (cuff rupture and/or cuff arthropathy)*

Table 1. *Hospital-level variation in patient characteristics and 1 and 3-year revision rates following primary shoulder arthroplasty in Dutch hospital*

Hospital	O/E ratio		
	2014-2016		2014-2018
	1-year revision	3-year revision	1-year revision
1	1.9[¥‖]		1.6[¥]
2	2.3[‖]	1.7	
4			2.0[¥‖§]
9	2.3[¥‖]		
13			2.5[¥]
22		2.7[‖]	
73		3.3	
Median (IQR) for non-negative outliers	0.6 (0.0-1.3)	0.8 (0.4-1.2)	0.8 (0.4-1.2)

IQR = interquartile range. An O/E ratio is provided for negative outliers during 2014-2016 and 2014-2018. O/E ratios indicate the observed number of revisions within the period divided by the expected number of revisions within 1 or 3-years based on patient mix. [¥] = negative outlier for revision because of infection; [‖] = negative outlier for revision because of cuff pathology (cuff rupture and/or cuff arthropathy); [§] = negative outlier for revision because of surgical technique failure (malalignment, instability, periprosthetic fracture, loosening glenoid component, and/or loosening humeral component)

Table 2. *Negative outlier hospitals with significantly more revisions than expected*

References

1. Best MJ, Aziz KT, Wilckens JH, et al. Increasing incidence of primary reverse and anatomic total shoulder arthroplasty in the United States. J Shoulder Elbow Surg 2021;30: 1159-66
2. Bottle A, Loeffler MD, Aylin P, et al. Comparison of 3 types of readmission rates for measuring hospital and surgeon performance after primary total hip and knee arthroplasty. J Arthroplasty 2018;33: 2014-9.e2
3. Denissen GAW, van Steenbergen LN, Lollinga WT, et al. Generic implant classification enables comparison across implant designs: the Dutch Arthroplasty Register implant library. EFORT Open Rev 2019;4:344-50
4. Dillon MT, Chan PH, Inacio MCS, et al. Yearly trends in elective shoulder arthroplasty, 2005-2013. Arthritis Care Res (Hoboken) 2017;69:1574-81
5. Gonzalez AA, Girotti ME, Shih T, et al. Reliability of hospital readmission rates in vascular surgery. J Vasc Surg 2014;59:1638-43
6. Ho A, Purdie C, Tirosh O, et al. Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. Patient Relat Outcome Meas 2019;10:217-26
7. Hofstede SN, Ceyisakar IE, Lingsma HF, et al. Ranking hospitals: do we gain reliability by using composite rather than individual indicators? BMJ Qual Saf 2019;28:94-102
8. Leong NL, Sumner S, Gowd A, et al. Risk factors and complications for revision shoulder arthroplasty. HSS J 2020;16:9-14
9. Dutch Arthroplasty Register (LROI). Online LROI annual report 2020 - PDF. Available: https://www.lroireport.nl/app/uploads/2021/03/PDF-LROI-annual-report-2020.pdf (last accessed 28 October, 2021)
10. Lübbeke A, Rees JL, Barea C, et al. International variation in shoulder arthroplasty. Acta Orthop 2017;88: 592-9
11. Lübbeke A, Silman AJ, Barea C, et al. Mapping existing hip and knee replacement registries in Europe. Health Policy 2018;122:548-57
12. Mahoney A, Bosco JA 3rd, Zuckerman JD. Readmission after shoulder arthroplasty. J Shoulder Elbow Surg 2014;23:377-81
13. Rasmussen JV, Olsen BS, Fevang BT, et al. A review of national shoulder and elbow joint replacement registries. J Shoulder Elbow Surg 2012;21:1328-35
14. National Joint Registry (NJR). NJR 17th annual report 2020. Available: https://reports.njrcentre.org.uk/Portals/0/PDFdownloads/NJR%2017th%20Annual%20Report%202020.pdf
15. Rolfson O, Chenok KE, Bohm E, Lübbeke A, et al. Patient-reported outcome measures in arthroplasty registries. Acta Orthop 2016;87(Suppl 1):3-8
16. Schairer WW, Zhang AL, Feeley BT. Hospital readmissions after primary shoulder arthroplasty. J Shoulder Elbow Surg 2014;23:1349-55
17. Schamber EM, Takemoto SK, Chenok KE, et al. Barriers to completion of patient reported outcome measures. J Arthroplasty 2013;28:1449-53
18. Serra-Sutton V, Allepuz A, Espallargues M, et al. Arthroplasty registers: a review of international experiences. Int J Technol Assess Health Care 2009;25:63-72
19. Singh A, Yian EH, Dillon MT, et al. The effect of surgeon and hospital volume on shoulder arthroplasty perioperative quality metrics. J Shoulder Elbow Surg

2014;23:1187-94
20. Singh JA, Ramachandran R. Does hospital volume predict outcomes and complications after total shoulder arthroplasty in the US? Arthritis Care Res (Hoboken) 2015;67:885-90
21. Theelen LMA, Mory B, Venkatesan S, et al. Stem retention and survival in revision of anatomical convertible shoulder arthroplasty to reverse arthroplasty: a Dutch registry study. BMC Musculoskelet Disord 2021; 22:396
22. van Dishoeck AM, Lingsma HF, Mackenbach JP, et al. Random variation and rankability of hospitals using outcome indicators. BMJ Qual Saf 2011;20:869-74
23. van Schie P, van Steenbergen LN, van Bodegom-Vos L, et al. Between-hospital variation in revision rates after total hip and knee arthroplasty in the Netherlands: directing quality-improvement initiatives. J Bone Joint Surg Am 2020;102:315-24
24. Wilson I, Bohm E, Lübbeke A, et al. Orthopaedic registries with patient-reported outcome measures. EFORT Open Rev 2019;4:357-67