BMJ

BMJ 2014;348:g3284 doi: 10.1136/bmj.g3284 (Published 23 May 2014)


RESEARCH

# Relation between surgeon volume and risk of complications after total hip arthroplasty: propensity score matched cohort study



Bheeshma Ravi *resident physician*[1], Richard Jenkinson *assistant professor*[1], Peter C Austin *senior scientist*[2 3], Ruth Croxford *epidemiologist*[2], David Wasserstein *fellow*[1], Benjamin Escott *fellow*[1], J Michael Paterson *scientist*[2 3], Hans Kreder *professor*[1 2 3], Gillian A Hawker *professor*[2 3 4]

[1]Division of Orthopaedic Surgery, Department of Surgery, University of Toronto, Toronto, Canada; [2]Institute for Clinical Evaluative Sciences, Toronto, Canada; [3]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada; [4]Division of Rheumatology, Department of Medicine, Women's College Hospital, Toronto, Canada

## Abstract

**Objectives** To identify a cut point in annual surgeon volume associated with increased risk of complications after primary elective total hip arthroplasty and to quantify any risk identified.

**Design** Propensity score matched cohort study.

**Setting** Ontario, Canada

**Participants** 37 881 people who received their first primary total hip arthroplasty during 2002-09 and were followed for at least two years after their surgery.

**Main outcome measure** The rates of various surgical complications within 90 days (venous thromboembolism, death) and within two years (infection, dislocation, periprosthetic fracture, revision) of surgery.

**Results** Multivariate splines were developed to visualize the relation between surgeon volume and the risk for various complications. A threshold of 35 cases a year was identified, under which there was an increased risk of dislocation and revision. 6716 patients whose total hip arthroplasty was carried out by surgeons who had done ≤35 such procedure in the previous year were successfully matched to patients whose surgeon had carried out more than 35 procedures. Patients in the former group had higher rates of dislocation (1.9% *v* 1.3%, P=0.006; NNH 172) and revision (1.5% *v* 1.0%, P=0.03; NNH 204).

**Conclusions** In a cohort of first time recipients of total hip arthroplasty, patients whose operation was carried by surgeons who had performed 35 or fewer such procedures in the year before the index procedure were at increased risk for dislocation and early revision. Surgeons should consider performing 35 cases or more a year to minimize the risk for complications. Furthermore, the methods used to visualize the relationship between surgeon volume and the occurrence of complications can be easily applied in any jurisdiction, to help inform and optimize local healthcare delivery.

## Introduction

Associations between volume and outcome for the occurrence of short term complications (such as mortality, deep vein thrombosis, early revision) after total hip arthroplasty have been variably reported.[1] Most studies show that the risk for these complications is roughly inversely proportional to the volume of procedures carried out by the operating surgeon.[2-5] Such findings have implications for the centralization of delivery of arthroplasty,[6] particularly as these complications are associated with considerable morbidity and increased healthcare costs.[7-10]

There is, however, a lack of consensus around what constitutes a "low" annual volume, with definitions ranging from less than six to less than 52 procedures a year.[11-13] The lack of a consistent definition is a reflection of the fact that volume thresholds have typically been created to ensure an even distribution of patients across volume categories. As such, there cannot be a reasonable expectation that these definitions are generalizable across regions nor can one expect that any resultant conclusions about the impact of volume on the risk for complications are accurate.[14-19] Thus, we do not know whether a threshold exists, or the amount of potential benefit, if any, of receiving surgery from a surgeon who carries out a higher number of procedures.

We clarified the relation between surgeon volume (defined as the number of total hip arthroplasties performed by the surgeon in the year before the index arthroplasty) and the risk for complications within 90 days (venous thromboembolism and death) or within two years (periprosthetic fracture, infection, dislocation, revision arthroplasty). Our specific objectives were to graphically describe the relation between surgeon volume and the risk for complications and to identify a cut point that predicts differential risk for complications, if one exists; and to

Correspondence to: B Ravi bheeshma.ravi@mail.utoronto.ca

quantify the increased risk of complication in cases in which the procedures were performed by surgeons with lower annual volumes.

## Methods

### Study sample

We used health administrative databases from Ontario, Canada (the country's most populous province, with a population of 13.5 million in 2012). Ontarians are insured under a single payer system, which covers all medically necessary procedures, including total hip arthroplasties. The main data sources were hospital discharge abstracts from the Canadian Institute for Health Information Discharge Abstract Database (CIHI-DAD), physician claims from the Ontario Health Insurance Plan (OHIP), and demographic information on each physician from the Ontario Physician Human Resources Data Centre (OPHRDC) and OHIP Corporate Provider Database (CPDB). Using specific procedure and diagnostic codes from the Canadian version of the 10th revision of the international statistical classification of diseases (ICD-10) and the Canadian classification of health interventions (ICD-10-CA/CCI), we defined a cohort of patients who received their first primary elective total hip arthroplasty for osteoarthritis from 1 April 2002 to 31 March 2009.

### Primary outcome: surgical complications

We identified the occurrence of venous thromboembolism and death within 90 days of the index total hip arthroplasty. Occurrence of a venous thromboembolism (deep vein thrombosis or pulmonary embolism) was identified by using diagnostic codes in the CIHI-DAD or National Ambulatory Care Reporting System (NACRS) databases. We identified death within 90 days of operation using the healthplan's registered persons database. We also identified the occurrence of infection, dislocation, periprosthetic fracture, and dislocation within two years of the index arthroplasty. To identify infections we used occurrence of an ICD-10-CA diagnostic code for intra-articular infection, with a confirmatory code for an irrigation and debridement; occurrence of an healthplan's code for a spacer insertion; and/or occurrence of a procedure code for a peripheral intravenous central catheter after the total joint arthroplasty, when the referring physician was an orthopedic surgeon. Dislocations were defined as the occurrence of a diagnostic code for dislocation or a procedure code for closed/open hip reduction. Periprosthetic fractures were defined as the occurrence of a diagnostic code for fracture after insertion of an implant. Revision procedures were identified with ICD-10-CA/CCI procedure codes accompanied by the supplementary status attribute "R."

### Covariates of interest

We measured and controlled for several patient and provider covariates that have been previously shown to affect the risk of occurrence of complications after joint replacement. Patient age and sex was obtained from the healthplan's registered persons databases (RPDB).[20-22] Comorbidities listed on hospital discharge abstracts in the three years before the index admission for arthroplasty were categorized according to an adaptation of the Charlson comorbidity index.[23] Adjusted clinical groups (ACGs), based on diagnosis codes from admission to hospital and physician visits in the two years before the index admission were used to classify recipients as "frail" (yes/no) at the time of the index procedure.[24] We identified patients with a history of pre-existing cardiovascular disease,[25] diabetes,[26] hypertension,[27] and chronic obstructive pulmonary disease[28] using validated algorithms.

Several validated surrogate measures for socioeconomic status and living conditions were obtained from the registered persons databases, including fifth of neighborhood income distribution, rurality index of Ontario, and the Ontario marginalization index. Neighbourhood income fifths categorize small geographic areas into five roughly equal population groups, with the lowest fifth referring to the least affluent neighborhoods.[29 30] The rurality index of Ontario uses a weighted formula, which considers three key elements: population size and density, travel time to nearest basic referral centre, and travel time to nearest advanced referral centre. Census subdivisions are then assigned a score from 0 to 100, with higher scores indicative of increasing rurality.[31 32] The Ontario marginalization index comprises four elements: ethnic concentration, residential instability, dependency, and deprivation.[33] Each element is sorted into fifths, arranged from least (lowest fifth) to most marginalized (highest fifth). The index has been shown to be stable across time periods and across different geographic areas and to be associated with health outcomes including depression,[34] smoking,[35] alcohol consumption,[36] and body mass index (BMI).[37]

For each total hip arthroplasty, we defined hospital volume as the number of hip arthroplasty procedures (both primary and revision) performed at the hospital where the surgery was performed in the 365 days before the index procedure. We defined teaching hospitals as those who were members of the Council of Academic Hospitals of Ontario (www.cahohospitals.com). The date of birth of the primary surgeon was obtained from the Ontario Health Insurance Plan and was used to determine the surgeon's age at the time of the index total hip arthroplasty.

### Main exposure variable: surgeon volume

For each total hip arthroplasty, surgeon volume was defined as the number of hip arthroplasty procedures (both primary and revision) performed by the operating surgeon in the 365 days before the index procedure, as such surgeon volume could change dynamically over time.

### Statistical analyses

We used restricted cubic splines with four knots[38] to model the relation between surgeon volume and the occurrence of each surgical complication, after adjustment for patient age, sex, rurality index, fifth of income distribution, marginalization index, Charlson score, frailty, presence of specific comorbidities (pre-existing cardiovascular disease, hypertension, diabetes, chronic obstructive pulmonary disease, and chronic kidney disease), annual hospital volume, age of the primary surgeon, and teaching hospital status. We examined the non-linear relation between surgeon volume and the risk of each complication to identify any inflection point that could be used to dichotomize annual volume into categories in a clinically meaningful way. If we observed an area of inflection, we used multivariable logistic regression to determine the area under the curve for the models relating various cut points of surgeon volume to the risk of the relevant complication. The surgeon volume with the maximum area under the curve was selected as the cut point to dichotomize surgeon volume.

Patients in the cohort were then classified according to whether the index total hip arthroplasty had been performed by a surgeon who had carried out ≤35 or more than 35 procedures in the 365 days before the index surgery. Baseline cohort characteristics were calculated with proportions and medians as appropriate and were compared between groups with Wilcoxon rank sum tests for continuous variables and $\chi^2$ tests for categorical variables. We determined a propensity score for receipt of a total hip arthroplasty from a surgeon with ≤35 procedures using a logistic regression model.[39 40] The covariates entered into the propensity score were sociodemographics (age, sex, income fifth, rurality index, Ontario marginalization index), health status (Charlson score, frailty, hypertension, chronic obstructive pulmonary disease, congestive heart failure, diabetes, chronic kidney disease), and provider characteristics (teaching hospital, annual hospital volume, experience of the primary surgeon at the time of the index procedure). Patients who received total hip arthroplasty from a surgeon with ≤35 procedures were matched to those from a surgeon with >35 procedures on the logit of the propensity score by using calipers of width equal to 0.2 of the standard deviation of the logit of the propensity score.[41] A matching ratio of 1:1 was used.[42]

We estimated standardized differences for all covariates before and after matching, with a standardized difference of 10% or more considered indicative of imbalance.[43] The occurrence of complications (venous thromboembolism and death within 90 days; infection, dislocation, periprosthetic fracture, and revision within two years) were compared between the two groups after matching, by using methods appropriate for the analysis of matched data in estimating the treatment effect and its significance. When we found a significant difference, we estimated the absolute risk increase and the number needed to treat to harm. We determined the hazard ratio for occurrence of a complication with Cox proportional hazards, after taking pair matching into account and using robust variance estimation.[44] We also performed several sensitivity analyses: we examined the effects of stratifying the matched analysis by teaching hospital status; surgeon experience (≤35 *v* >35 procedures/year); limiting inclusion to procedures performed by surgeons that graduated from a Canadian medical school; and limiting inclusion to procedures performed by surgeons with more than five years of experience. Finally, to assess the sensitivity of our findings to unmeasured confounding, we used an array approach.[45] In doing so, we determined the necessary effect of an unmeasured confounder on the risk of the adverse outcome and the necessary imbalance in the distribution of this unmeasured confounder between the two exposure groups that were needed to negate the significant volume effects that we observed. All analyses were performed at the Institute for Clinical Evaluative Sciences (www.ices.on.ca) with SAS version 9.3 for UNIX (SAS Institute, Cary, NC). The type I error probability was set to 0.05 for all analyses.

## Results

### Patient and characteristics

From 1 April 2002 to 31 March 2009, there were 37 881 eligible recipients of total hip arthroplasty (fig 1⇓, tables 1⇓ and 2⇓). The procedures that comprise our cohort were performed by 350 surgeons. The median annual surgeon volume was 55 procedures (interquartile range 35-85), and the median surgeon experience (number of years in practice at the time of the arthroplasty) was 19 years (11-27).

### Regression splines describing relation between surgeon volume and risk of complications

The restricted cubic splines for the risk of venous thromboembolism, death, infection, and periprosthetic fracture did not display an obvious relation between the risk for these complications and annual surgeon volume (fig 2⇓). The splines relating annual surgeon volume to dislocation and revision, however, had similar shapes—both were negatively sloped with inflection points at about 35 procedures a year, after which the rates of complications continued to decrease with increased surgeon volume but at a lower rate (figs 2 and 3⇓⇓). The shapes of these splines remained unchanged after we included an interaction term for surgeon volume and hospital volume. Receiver operating characteristics curves were generated relating annual surgeon volume (with cut points of 20, 25, 30, 35, 40, 45, and 50 procedures a year) to the risk for dislocation or revision within two years of total hip arthroplasty. The ideal cut point was found to be 35 a year, with an area under curve for dislocation and revision of 0.650 and 0.605, respectively (fig 4⇓).

Based on these curves, we dichotomized surgeon volume at a surgeon volume of 35 procedure (≤35 or >35 in the 365 days before the procedure). Patients whose surgeon was in the first group were slightly older (70 *v* 67; P<0.001), more likely to be women (56% *v* 53%; P<0.001), had a higher rurality index (14.8 *v* 13.0; P<0.001), and a higher median deprivation index (third fifth *v* second fifth; P<0.001) and median dependency index (fourth fifth *v* third fifth; P<0.001) (table 3⇓. They were also less likely to receive their surgery at a teaching hospital (14% *v* 41%; P<0.001) and received their surgery at hospitals with lower annual volumes (159 *v* 237 procedures/year; P<0.001) (table 4⇓).

### Matching

We successfully matched 6716 patients whose total hip arthroplasty was carried out by surgeons who had done ≤35 such procedure in the previous year with patients whose surgeons had carried out more than 35 procedures. After matching, the absolute standardized differences were less than 10% for all variables entered into the propensity score, indicating an adequate match.

### Outcomes after matching

Rates of dislocation (1.9% *v* 1.3%; P=0.006) and revision (1.5% *v* 1.0%; P=0.03) within two years of surgery were higher in hip replacement recipients whose surgeons had an annual volume of ≤35 procedures (table 5⇓). The numbers needed to treat to harm for dislocation and revision were 172 (95% confidence interval 164 to 182) and 204 (193 to 217), respectively. These recipients were at higher risk of both dislocation (hazard ratio 1.48, 95% confidence interval 1.21 to 1.80; P<0.001) and revision (1.44, 1.15 to 1.80; P=0.001) relative to recipients with surgeons with annual volumes of over 35 procedures.

### Sensitivity analyses

#### *Stratifying by teaching hospital status*

A total of 24 903 total hip arthroplasties were performed at a non-teaching hospital. We successfully matched 5748 (71%) patients whose arthroplasty was carried out by surgeons who had done ≤35 such procedure in the previous year with 5748 (34.4%) patients whose surgeons had carried out more than 35 procedures a year (standardized difference <10% for all matched variables). Patients in the first group had a higher risk for both dislocation (hazard ratio 1.41, 95% confidence interval 1.13 to 1.75; P=0.003) and revision (1.49, 1.15 to 1.92; P=0.002).

### *Stratifying by hospital volume*

We also repeated the propensity score match after stratifying by the median hospital volume (<211 *v* ≥211 procedures/year). Among patients who received their surgery at a hospital with fewer than 211 procedures a year, we successfully matched 4809 (68%) recipients with surgeon with ≤35 procedures a year to 4809 (40%) recipients with surgeons with more than 35 procedures a year (standardized difference <10% for all matched variables). Patients in the first group had a higher risk for dislocation (hazard ratio 1.38, 95% confidence interval 1.10 to 1.74; P=0.006) and revision (1.32, 1.02 to 1.69; P=0.03).

Among patients who received their arthroplasty at a hospital with ≥211 procedures a year, we successfully matched 1730 (71%) recipients with a surgeon with ≤35 procedures a year to 1730 (11%) recipients with surgeons with more than 35 procedures a year (standardized difference <10% for all matched variables). Patients in the first group were at a higher risk for dislocation (1.81, 1.16 to 2.84; P=0.009) but not for revision (1.57, 0.97 to 2.55; P=0.07).

### *Limiting to procedures performed by graduates of Canadian medical schools*

Canadian medical graduates performed 29 359 procedures. We successfully matched 5105 (82%) recipients with surgeons with >35 procedures a year to 5105 (22%) recipients with surgeons with >35 procedures a year (standardized difference <10% for all matched variables). Patients in the first group had a higher risk for dislocation (hazard ratio 1.76, 95% confidence interval 1.37 to 2.25; P<0.001) and revision (1.88, 1.42 to 2.49; P<0.001).

### *Limiting to procedures performed by surgeons with more than five years of experience*

Exclusion of procedures performed by surgeons with less than five years of experience left 32 867 procedures. We successfully matched 5749 recipients (22%) recipients with surgeons with >35 procedures a year to 5749 recipients (80%) recipients with surgeons with more than 35 procedures a year (standardized difference <10% for all matched variables). Patients in the first group had a higher risk for dislocation (hazard ratio 1.42, 1.15 to 1.76; P=0.001) and revision (1.28, 1.02 to 1.62; P=0.037).

### *Determining strength of unmeasured confounder*

An unmeasured confounding variable (such as obesity or smoking), if not collinear with these other covariates and present in only one of the two groups, would have to have had a prevalence of at least 65% in that one group, and a relative risk ratio of at least 0.75 (if found only among recipients with surgeons with annual volumes of >35 procedures) or 1.33 (if found only among recipients with surgeons with annual volumes of ≤35 procedures) to account for the observed effects of surgeon volume on the risk for dislocation.

## Discussion

## Principal findings

This study defines a threshold of surgeon volumes related to complication rates after total hip arthroplasty in first time recipients with osteoarthritis. We used a novel method to visually describe the relation between surgeon volume and the occurrence of a surgical complication within two years of surgery. Although we found no obvious relation between surgeon volume and either infection, periprosthetic fracture, venous thromboembolism, or death, the models indicated that as surgeon volume increased, the risks for dislocation and early revision decreased. For both these complications, we observed an inflection point at about 35 procedures a year, after which the rate of decrease in the risks for complications leveled off. In patients operated on by surgeons with annual volumes ≤35 procedures, the risks for dislocation and revision increased by about 48% and 44%, respectively.

Through the use of restricted cubic splines, we found that there was a noticeable decrease in likelihood of dislocation and revision as the surgeon's yearly volume of hip arthroplasty increased; however, the relation was not linear. Surgeons with extremely low volumes have predicted rates of dislocation of about 4%, with a drop in likelihood to about 2% at 25-50 procedures a year. While the relative improvement in complication rates with increasing surgeon volume attenuated after this point, there continued to be a downward trend in the risks for dislocation and revision, indicating that increased surgeon volume continues to have a beneficial impact, although one that is less pronounced.

The spline curves showed that surgeon volume has a differential effect and relation on the risk of specific complications. That the splines relating surgeon volume to dislocation and revision were most indicative of an association is not unexpected, as these complications (of the ones examined) are the most likely to be affected by surgeon technique.[46-48] The splines for venous thromboembolism, death, infection, and periprosthetic fracture did not show an obvious relation between surgeon volume and risk of complication. This is consistent with our current knowledge around the risk factors for these complications,[1 4 49] which are mainly to do with the patient (for example, male sex, increased comorbidity, and frailty)[49-51] and would not be influenced by surgeon volume. While both patient and provider factors can influence risk for early periprosthetic fracture,[52 53] the rate of this complication was low in our cohort (0.4%), minimizing our statistical power to find a relation if one exists.

A surgeon with an increased rate of complications can opt to do fewer joint replacements in response; therefore, surgeon volume could in fact be a reflection of a surgeon's skill. The latter is a complex entity that includes innate skill, training, and experience. We attempted to control for surgeon skill by performing secondary analyses in which we limited inclusion to procedures performed by surgeons who graduated from a Canadian medical school and to procedures performed by a surgeon with at least five years of experience. In both analyses, our findings remained consistent—surgeon volume of ≤35 procedures a year was associated with an increased risk of dislocation and revision, indicating that this finding is not the result of variable training or lack of experience.

The relation between surgeon volume and outcomes has been examined for several surgical procedures, ranging from radical prostatectomy to repair of thoracoabdominal aortic aneurysms,[54-60] with the general finding that increased volumes contribute to lower rates of complications. In several jurisdictions in North America and the United Kingdom, these findings have contributed to the regionalization of care into specialized centers.[61-63] In the current study, we found that patients who had their operation performed by surgeon with annual volumes ≤35 procedures a year were more likely to be frail, live in rural areas, and receive their hip replacements from lower volume hospitals. This patient profile is consistent with previous descriptions of patients who utilize lower volume centers.[64] These patients might be adversely affected by a policy of selective referral for arthroplasty to high volume centres, particularly if there are longer waiting times in the latter.[64 65] While our findings indicate that greater surgeon volume is associated with reduced risk of complications, the identified threshold of surgeon volume of 35 procedures a year is not onerous and does not necessarily require centralization to achieve, particularly when one considers that the median volume of surgeons in non-teaching hospitals (that is, community hospitals) was 46 procedures a year.

The methods we used in this study take advantage of several existing statistical techniques, foremost of which was the creation of multivariate restricted cubic splines. To our knowledge, this is the first study to utilize splines to visualize the relation between volume and the risk of complications after adjustment for multiple potential confounders, and to attempt to identify a variable cut point associated with a differential risk for complications. While the identified cut point of 35 procedures a year might not be generalizable to other settings,[66] the technique used to define it can be applied in different geographic areas and for various procedures, thus providing local administrators and policy makers with more relevant information on the interplay of provider volume and surgical complications in their specific setting. In turn, this will inform their strategies around delivery of these procedures. The use of this technique, however, requires the availability of population based data, the ability to accurately determine the volume of each operating surgeon in the year before the surgery, and specific patient level data including comorbidity and sociodemographic variables.

## Strengths and limitations

Strengths of our study include the use of population based health administrative data to assemble a large sample of first time recipients of total hip arthroplasty and consideration of patient, hospital, and surgeon predictors of complications after the procedure. Our use of restricted cubic splines allowed us to visualize the relation between surgeon volume and occurrence of complications, enabling selection of a cut point for surgeon volume with more confidence than in previous studies. Our use of a propensity score matched analysis allowed us to balance several characteristics of patients (such as age, sex, comorbidity, various indices of socioeconomic status) and providers (such as surgeon experience, teaching hospital status, hospital volume) between groups. We also found that that our results remained robust after stratifying our analysis by teaching hospital status and surgeon experience.

There were some limitations in addition to those already noted. First, we did not have any information on outcomes reported by patients and thus could not identify a specific threshold for surgeon volume as it relates to these outcomes. As the occurrence of surgical complications has been linked with worse outcomes reported by patients,[67-70] it is possible that volumes greater than 35 procedures a year contribute to improved outcomes. Second, there were other potential confounders that we were unable to capture and thus control for, such as BMI and smoking status.[71 72] Both these factors, however, are strongly associated with other factors that were measured and balanced between matched groups, including diabetes,[73 74] hypertension,[75 76] congestive heart failure,[77 78] chronic obstructive pulmonary disease,[79 80] chronic kidney disease,[81 82] frailty,[83 84] and various socioeconomic indices.[35 37 85] We were also unable to control for technical aspects of the procedure—such as surgical approach,[86] implant type,[87] and use of bone cement[88 89]—all of which have been linked with complication rates after total joint arthroplasty. Surgeons with higher volumes might systematically differ from those with lower volumes with respect to surgical techniques; if so, this could account, at least in part, for the volume effect observed in the current study. Further research, potentially with data sources that capture this information, is recommended to confirm or refute these hypotheses.

## Conclusions and policy implications

In summary, among first time primary elective recipients of total hip arthroplasty, patients who were operated on by surgeons who performed fewer than 35 procedures in the year before the surgery were at higher risk for dislocation or revision within two years. We found no relation between surgeon volume and the occurrence of venous thromboembolism, death, infection, or periprosthetic fracture. Our findings indicate that restricted cubic splines allow for a visualization of the relation between surgeon volume and the occurrence of complications. This technique will allow for more relevant cut points for surgeon volume and will allow for more informed decision making around standards for volume.

Contributors: All authors contributed to the development, analysis and final manuscript. BR is guarantor.

Funding: This study was supported by a grant from the Canadian Institutes of Health Research and by the Institute for Clinical Evaluative Sciences, a non-profit research institute funded by the Ontario Ministry of Health and Long-Term Care. The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by the Institute for Clinical Evaluative Sciences or the Ontario Ministry of Health and Long-Term Care is intended or should be inferred. GAH is supported in part by the FM Hill Chair in Academic Women's Medicine. ICES received support from the Ministry of Health and Long-Term Care. CIHR Grant number: MOP-15468.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Transparency: The lead author (the manuscript's guarantor) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Data sharing: Please contact the corresponding author regarding data sharing.

1 Paterson JM, Williams JI, Kreder HJ, Mahomed NN, Gunraj N, Wang X, et al. Provider volumes and early outcomes of primary total joint replacement in Ontario. *Can J Surg* 2010;53:175-83.
2 Lavernia CJ, Guzman JF. Relationship of surgical volume to short-term mortality, morbidity, and hospital charges in arthroplasty. *J Arthroplasty* 1995;10:133-40.
3 Katz JN, Phillips CB, Baron JA, Fossel AH, Mahomed NN, Barrett J, et al. Association of hospital and surgeon volume of total hip replacement with functional status and satisfaction three years following surgery. *Arthritis Rheum* 2003;48:560-8.
4 Losina E, Barrett J, Mahomed NN, Baron JA, Katz JN. Early failures of total hip replacement: effect of surgeon volume. *Arthritis Rheum* 2004;50:1338-43.
5 Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, et al. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States medicare population. *J Bone Joint Surg Am* 2001;83-A:1622-9.
6 Morris AH. The association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States Medicare population: health policy implications. *J Bone Joint Surg Am* 2001;83-A:1754-5.
7 Kluess D, Martin H, Mittelmeier W, Schmitz K-P, Bader R. Influence of femoral head size on impingement, dislocation and stress distribution in total hip replacement. *Med Eng Phys* 2007;29:465-71.
8 De Palma L, Procaccini R, Soccetti A, Marinelli M. Hospital cost of treating early dislocation following hip arthroplasty. *Hip Int* 2012;22:62-7.
9 Clyburn TA, Cui Q. Antibiotic laden cement: current state of the art. *J Am Acad Orthop Surg Now* 2013. www.aaos.org/news/bulletin/may07/clinical7.asp.
10 Burns AWR, Bourne RB, Chesworth BM, MacDonald SJ, Rorabeck CH. Cost effectiveness of revision total knee arthroplasty. *Clin Orthop* 2006;446:29-33.

### What is already known on this topic

Though increased surgeon volume is associated with reduced risk for surgical complications, it is not clear if there is a specific threshold that is associated with a reduced risk

It is not known if this beneficial effect of increased volume persists after adjustment for relevant confounders, including hospital volume and surgeon age

### What this study adds

After primary total hip arthroplasty, the risks for dislocation and early revision in patients whose surgeons had carried out ≤35 procedures in the previous year were about 48% and 44% higher, respectively, than in patients whose surgeons had carried out more than 35 procedures

This study used a novel method to visually describe the relation between surgeon volume and specific surgical complications, which could be applied to other procedures after primary total hip arthroplasty


11 Lau R, Perruccio A, Gandhi R, Mahomed N. The role of surgeon volume on patient outcome in total knee arthroplasty: a systematic review of the literature. *BMC Musculoskeletal Disorders* 2012;13:250.
12 Baker P, Dowen D, McMurtry I. The effect of surgeon volume on the need for transfusion following primary unilateral hip and knee arthroplasty. *Surgeon* 2011;9:13-7.
13 Katz JN, Mahomed NN, Baron JA, Barrett JA, Fossel AH, Creel AH, et al. Association of hospital and surgeon procedure volume with patient-centered outcomes of total knee replacement in a population-based cohort of patients age 65 years and older. *Arthritis Rheum* 2007;56:568-74.
14 Moser BK, Coombs LP. Odds ratios for a continuous outcome variable without dichotomizing. *Stat Med* 2004 30;23:1843-60.
15 Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. *Epidemiology* 1992;3:434-40.
16 Bakhshi E, McArdle B, Mohammad K, Seifi B, Biglarian A. Let continuous outcome variables remain continuous. *Comput Math Methods Med* 2012;2012:639124.
17 Dawson NV, Weiss R. Dichotomizing continuous variables in statistical analysis: a practice to avoid. *Med Decis Making* 2012;32:225-6.
18 Gustafson P, Le Nhu D. Comparing the effects of continuous and discrete covariate mismeasurement, with emphasis on the dichotomization of mismeasured predictors. *Biometrics* 2002;58:878-87.
19 Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006;25:127-41.
20 Agabiti N, Picciotto S, Cesaroni G, Bisanti L, Forastiere F, Onorati R, et al. The influence of socioeconomic status on utilization and outcomes of elective total hip replacement: a multicity population-based longitudinal study. *Int J Qual Health Care* 2007;19:37-44.
21 Santaguida PL, Hawker GA, Hudak PL, Glazier R, Mahomed NN, Kreder HJ, et al. Patient characteristics affecting the prognosis of total hip and knee joint arthroplasty: a systematic review. *Can J Surg* 2008;51:428-36.
22 Kralj B. Measuring “rurality” for purposes of health care planning: an empirical measure for Ontario. Ontario Medical Association, 2005.
23 Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613-9.
24 Weiner JP, Abrams C. The Johns Hopkins ACG system: technical reference guide version 10.0. John Hopkins Bloomberg School of Public Health, 2011.
25 Ko DT, Mamdani M, Alter DA. Lipid-lowering therapy with statins in high-risk elderly patients: the treatment-risk paradox. *JAMA* 2004;291:1864-70.
26 Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care* 2002;25:512-6.
27 Tu K, Campbell NR, Chen ZL, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. *Open Med* 2007;1:e18-26.
28 Gershon A, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physcian diagnosed COPD in health administrative databases. *COPD* 2009;6:388-94.
29 Information CIfH. Reducing gaps in health: a focus on socio-economic status in urban Canada. CIHI, 2008.
30 Glazier R, Badley E, Gilbert J, Rothman L. The nature of increased hospital use in poor neighbourhoods: findings from a Canadian inner city. *Can J Public Health* 2000;91:268-73.
31 Kralj B. Measuring “rurality” for purposes of health-care planning: an empirical measure for Ontario. *Ont Med Rev* 2000;Oct:37-40.
32 Hastings DE, Parker SM. Protrusio acetabuli in rheumatoid arthritis. *Clin Orthop Relat Res* 1975;108:76-83.
33 Matheson F, Dunn J, Smith K, Moineddin R, Glazier R. Ontario marginalization index (ON-Marg): user guide. Centre for Research in Inner City Health, St Michael’s Hospital (Toronto), 2011.
34 Matheson F, Moineddin R, Dunn J, Creatore M, Gozdyra P, Glazier R. Urban neighborhoods, chronic stress, gender and depression. *Soc Sci Med* 2006;63:2604-16.
35 White H, Matheson F, Moineddin R, Dunn J, Glazier R. Neighbourhood deprivation and regional inequalities in self-reported health among Canadians: are we equally at risk? *Health Place* 2011;17:361-9.
36 Matheson F, White H, Moineddin R, Dunn J, Glazier R. Drinking in context: the influence of gender and neighbourhood deprivation on alcohol consumption. *J Epidemiol Community Health* 2012;66:e4.
37 Matheson F, Moineddin R, Glazier R. The weight of place: a multilevel analysis of gender, neighborhood material deprivation, and body mass index among Canadian adults. *Soc Sci Med* 2008;66:675-90.
38 Marrie RA, Dawson NV, Garland A. Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. *J Clin Epidemiol* 2009;62:511-7.e1.
39 Austin P. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399-424.
40 Austin P. A tutorial and case study in propensity score analysis: an application to estimating the effect of in-hospital smoking cessation counseling on mortality. *Multivariate Behav Res* 2011;46:119-51.
41 Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm* 2011;10:150-61.
42 Austin PC. Comparing paired vs non-paired statistical methods of analyses when making inferences about absolute risk reductions in propensity-score matched samples. *Stat Med* 2011;30:1292-301.
43 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;28:3083-107.
44 Barlow WE. Robust variance estimation for the case-cohort design. *Biometrics* 1994;50:1064-72.
45 Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf* 2006;15:291-303.
46 Bini S, Khatod M, Cafri G, Chen Y, Paxton E. Surgeon, implant, and patient variables may explain variability in early revision rates reported for unicompartmental arthroplasty. *J Bone Joint Surg Am* 95:2195-202.
47 Dobzyniak M, Fehring TK, Odum S. Early failure in total hip arthroplasty. *Clin Orthop* 2006;447:76-8.
48 Alberton GM, High WA, Morrey BF. Dislocation after revision total hip arthroplasty : an analysis of risk factors and treatment options. *J Bone Joint Surg Am* 2002;84-A:1788-92.
49 Ravi B, Croxford R, Hollands S, Bogoch E, Paterson M, Kreder H, et al. Patients with rheumatoid arthritis are at increased risk for complications following total joint arthroplasty. *Arthritis Rheum* 2014;66:254-63.
50 Saleh K, Olson M, Resig S, Bershadsky B, Kuskowski M, Gioe T, et al. Predictors of wound infection in hip and knee joint replacement: results from a 20 year surveillance program. *J Orthop Res* 2002;20:506-15.
51 Urquhart DM, Hanna FS, Brennan SL, Wluka AE, Leder K, Cameron PA, et al. Incidence and risk factors for deep surgical site infection after primary total hip arthroplasty: a systematic review. *J Arthroplasty* 2010;25:1216-22.
52 Cook RE, Jenkins PJ, Walmsley PJ, Patton JT, Robinson CM. Risk factors for periprosthetic fractures of the hip: a survivorship analysis. *Clin Orthop* 2008;466:1652-6.
53 Franklin J, Malchau H. Risk factors for periprosthetic femoral fracture. *Injury* 2007;38:655-60.
54 Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med* 2003 27;349:2117-27.
55 Courcoulas A, Schuchert M, Gatti G, Luketich J. The relationship of surgeon and hospital volume to outcome after gastric bypass surgery in Pennsylvania: a 3-year summary. *Surgery* 2003;134:613-3.
56 Cowan JA Jr, Dimick JB, Henke PK, Huber TS, Stanley JC, Upchurch GR, Jr. Surgical treatment of intact thoracoabdominal aortic aneurysms in the United States: hospital and surgeon volume-related outcomes. *J Vasc Surg* 2003;37:1169-74.
57 Hu JC, Gold KF, Pashos CL, Mehta SS, Litwin MS. Role of surgeon volume in radical prostatectomy outcomes. *J Clin Oncol* 2003;21:401-5.
58 Stavrakis AI, Ituarte PHG, Ko CY, Yeh MW. Surgeon volume as a predictor of outcomes in inpatient and outpatient endocrine surgery. *Surgery* 2007;142:887-99.
59 Lieberman MD, Kilburn H, Lindsey M, Brennan MF. Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. *Ann Surg* 1995;222:638-45.
60 Simunovic M, To T, Theriault M, Langer B. Relation between hospital surgical volume and outcome for pancreatic resection for neoplasm in a publicly funded health care system. *CMAJ* 1999;160:643-8.
61 Collier R. Is regionalization working? *CMAJ* 2010;182:331-2.
62 Lewis S, Kouri D. Regionalization: making sense of the Canadian experience. *Healthc Pap* 2004;5:12-31.
63 Block EFJ, Rudloff B, Noon C, Behn B. Regionalization of surgical services in central Florida: the next step in acute care surgery. *J Trauma* 2010;69:640-4.
64 Losina E, Wright EA, Kessler CL, Barrett JA, Fossel AH, Creel AH, et al. Neighborhoods matter: use of hospitals with worse outcomes following total knee replacement by patients from vulnerable populations. *Arch Intern Med* 2007;167:182-7.
65 FitzGerald JD, Soohoo NF, Losina E, Katz JN. Potential impact on patient residence to hospital travel distance and access to care under a policy of preferential referral to high-volume knee replacement hospitals. *Arthritis Care Res (Hoboken)* 2012;64:890-7.
66 Perkins NJ, Schisterman EF. The inconsistency of ”optimal” cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol* 2006;163:670-5.
67 Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clin Orthop* 2010;468:57-63.
68 Lingard EA, Katz JN, Wright EA, Sledge CB, Kinemax Outcomes G. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am* 2004;86-A:2179-86.
69 Nilsdotter AK, Petersson IF, Roos EM, Lohmander LS. Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. *Ann Rheum Dis* 2003;62:923-30.
70 Papakostidou I, Dailana ZH, Papapolychroniou T, Liaropoulos L, Zintzaras E, Karachalios TS, et al. Factors affecting the quality of life after total knee arthroplasties: a prospective study. *BMC Musculoskelet Disord* 2012;13:116.
71 Kapadia B, Issa K, Pivec R, Bonutti P, Mont M. Tobacco use may be associated with increased revision and complication rates following total hip arthroplasty. *J Arthroplasty* 2013;29:777-80.

72 Haverkamp D, Klinkenbijl MN, Somford MP, Albers GHR, van der Vis HM. Obesity in total hip arthroplasty—does it really matter? A meta-analysis. *Acta Orthopaedica* 2011;82:417-22.
73 Hu FB. Sedentary lifestyle and risk of obesity and type 2 diabetes. *Lipids* 2003;38:103-8.
74 Will JC, Galuska DA, Ford ES, Mokdad A, Calle EE. Cigarette smoking and diabetes mellitus: evidence of a positive association from a large prospective cohort study. *Int J Epidemiol* 2001;30:540-6.
75 Grossman E, Messerli FH. Hypertension and diabetes. *Adv Cardiol* 2008;45:82-106.
76 Colosia AD PR, Khan S. Prevalence of hypertension and obesity in patients with type 2 diabetes mellitus in observational studies: a systematic literature review. *Diabetes Metab Syndr Obes* 2013;6:327-38.
77 Shah PK, Helfant RH. Smoking and coronary artery disease. *Chest* 1988;94:449-52.
78 Kenchaiah S, Evans JC, Levy D, Wilson PWF, Benjamin EJ, Larson MG, et al. Obesity and the risk of heart failure. *N Engl J Med* 2002;347:305-13.
79 Tashkin DP, Murray RP. Smoking cessation in chronic obstructive pulmonary disease. *Respir Med* 2009;103:963-74.
80 O'Donnell DE O'Donnell CD, Webb KA, Guenette JA. Respiratory consequences of mild-to-moderate obesity: impact on exercise performance in health and in chronic obstructive pulmonary disease. *Pulm Med* 2012;2012:818925.
81 Rutkowski P, Klassen A, Sebekova K, Bahner U, Heidland A. Renal disease in obesity: the need for greater attention. *J Ren Nutr* 2006;16:216-23.
82 Halimi JM, Giraudeau B, Vol S, Caces E, Nivet H, Lebranchu Y, et al. Effects of current smoking and smoking discontinuation on renal function and proteinuria in the general population. *Kidney Int* 2000;58:1285-92.
83 Hubbard RE, Lang IA, Llewellyn DJ, Rockwood K. Frailty, body mass index, and abdominal obesity in older people. *J Gerontol A Biol Sci Med Sci* 2010;65:377-81.
84 Hubbard RE, Searle SD, Mitnitski A, Rockwood K. Effect of smoking on the accumulation of deficits, frailty and survival in older adults: a secondary analysis from the Canadian Study of Health and Aging. *J Nutr Health Aging* 2009;13:468-72.
85 Information CIfH. An assessment of their health status and health determinants. How healthy are rural Canadians? 2006, CIfH.
86 Masonis JL, Bourne RB. Surgical approach, abductor function, and total hip arthroplasty dislocation. *Clin Orthop* 2002;405:46-53.
87 McMinn D, Daniel J. History and modern concepts in surface replacement. *Proc Inst Mech Eng H* 2006;220:239-51.
88 Kirk PG, Rorabeck CH, Bourne RB, Burkart B. Total hip arthroplasty in rheumatoid arthritis: comparison of cemented and uncemented implants. *Can J Surg* 1993;36:229-32.
89 Morrey BF, Kavanagh BF. Complications with revision of the femoral component of total hip arthroplasty. Comparison between cemented and uncemented techniques. *J Arthroplasty* 1992;7:71-9.

**Accepted:** 29 April 2014

Cite this as: *BMJ* 2014;348:g3284

# Tables

**Table 1| Demographic characteristics of 37 881 eligible recipients of total hip arthroplasty**

| | Data |
|---|---|
| Median (IQR) age (years) | 68 (58-75) |
| No (%) of women | 20 372 (53.9) |
| No (%) of men | 17 423 (46.1) |
| Mean (SD) rurality index | 13.45 (18.50) |
| No (%) by fifth of income: | |
| Lowest | 6114 (16.2) |
| 2 | 7257 (19.2) |
| 3 | 7375 (19.5) |
| 4 | 7909 (20.9) |
| Highest | 9105 (24.1) |
| Median (IQR) income fifth | 3 (2-5) |
| No (%) by fifth of ethnic concentration: | |
| Lowest | 6576 (19.3) |
| 2 | 7330 (21.5) |
| 3 | 7288 (21.4) |
| 4 | 6729 (19.8) |
| Highest | 6095 (17.9) |
| Median (IQR) ethnic concentration | 3 (2-4) |
| No (%) by fifth of instability index: | |
| Lowest | 7073 (20.8) |
| 2 | 7495 (22.0) |
| 3 | 6642 (19.5) |
| 4 | 6319 (18.6) |
| Highest | 6489 (19.1) |
| Median (IQR) instability index | 3 (2-4) |
| No (%) by fifth of dependency index: | |
| Lowest | 4224 (12.4) |
| 2 | 6012 (17.7) |
| 3 | 6779 (19.9) |
| 4 | 7127 (21.0) |
| Highest | 9876 (29.0) |
| Median (IQR) dependency index | 3 (2-5) |
| No (%) by fifth of deprivation index: | |
| Lowest | 8972 (26.4) |
| 2 | 8193 (24.1) |
| 3 | 7273 (21.4) |
| 4 | 5639 (16.6) |
| Highest | 3941 (11.6) |
| Median (IQR) deprivation index | 2 (1-4) |

IQR=interquartile range.

**Table 2| Medical and admission characteristics of 37 881 eligible recipients of total hip arthroplasty before surgery and subsequent complications**

| | Data |
|---|---|
| **Medical characteristics** | |
| No (%) by comorbidities: | |
| Frail | 1953 (5.2) |
| Congestive heart failure | 1933 (5.1) |
| Chronic kidney disease | 912 (2.4) |
| Chronic obstructive pulmonary disease | 6117 (16.1) |
| Diabetes | 5668 (15.0) |
| Hypertension | 22 865 (60.4) |
| No (%) by Charlson score: | |
| 0 | 34 500 (91.1) |
| 1 | 1700 (4.5) |
| ≥2 | 1681 (4.4) |
| Median (IQR) Charlson score | 0 (0-0) |
| **Admission characteristics** | |
| No (%) of teaching hospitals | 12 978 (34.3) |
| Median (IQR) hospital volume* | 211 (150-341) |
| Median (IQR) surgeon experience (years) | 19 (11-27) |
| No (%) by medical school of surgeon: | |
| Canadian | 29 359 (77.5) |
| International | 6331 (16.7) |
| Unknown | 2191 (5.9) |
| **Complications** | |
| No (%) by complication within 2 years: | |
| Infection | 407 (1.1) |
| Dislocation | 458 (1.2) |
| Periprosthetic fracture | 138 (0.4) |
| Revision | 429 (1.1) |
| No (%) by complications within 90 days: | |
| Venous thromboembolism | 533 (1.4) |
| Death | 178 (0.5) |

IQR=interquartile range.

*No of total hip arthroplasties in 365 days before surgery.

**Table 3| Comparison of demographic characteristic of recipients of total hip arthroplasty before and after matching by surgeon volume**

| | Before matching | | | After matching | | |
|---|---|---|---|---|---|---|
| | ≤35 procedures (9494 patients) | >35 procedures (28 387 patients) | Standardized difference | ≤35 procedures (6716 patients) | >35 procedures (6716 patients) | Standardized difference |
| Median (IQR) age (years) | 70 (61-76) | 67 (58-75) | 0.20 | 69 (61-76) | 70 (61-76) | 0.03 |
| No (%) of women | 5290 (55.9) | 15 082 (53.2) | 0.05 | 3730 (55.7) | 3716 (55.4) | 0.00 |
| No (%) of men | 4170 (44.1) | 13 253 (46.8) | | 2965 (44.3) | 2989 (44.6) | |
| Mean (SD) rurality index | 14.77 (18.79) | 13.01 (18.38) | 0.10 | 12.28 (17.22) | 12.08 (16.81) | 0.01 |
| No (%) by fifth of income: | | | | | | |
| Lowest | 1674 (17.7) | 4440 (15.7) | 0.10 | 1106 (16.5) | 1122 (16.7) | 0.01 |
| 2 | 1966 (20.8) | 5291 (18.7) | | 1385 (20.6) | 1394 (20.8) | |
| 3 | 1898 (20.0) | 5477 (19.4) | | 1341 (20.0) | 1346 (20.0) | |
| 4 | 1924 (20.3) | 5985 (21.2) | | 1413 (21.0) | 1416 (21.1) | |
| Highest | 2011 (21.2) | 7094 (25.1) | | 1471 (21.9) | 1438 (21.4) | |
| Median (IQR) income fifth | 3 (2-4) | 3(2-5) | 0.10 | 3 (2-4) | 3 (2-4) | 0.01 |
| No (%) by fifth of ethnic concentration: | | | | | | |
| Lowest | 1595 (19.8) | 4981 (19.2) | 0.05 | 1269 (18.9) | 1221 (18.2) | 0.04 |
| 2 | 1613 (20.1) | 5717 (22.0) | | 1356 (20.2) | 1279 (19.0) | |
| 3 | 1669 (20.8) | 5619 (21.6) | | 1412 (21.0) | 1408 (21.0) | |
| 4 | 1481 (18.4) | 5248 (20.2) | | 1278 (19.0) | 1290 (19.2) | |
| Highest | 1682 (20.9) | 4413 (17.0) | | 1401 (20.9) | 1518 (22.6) | |
| Median (IQR) ethnic concentration | 3 (2-4) | 3 (2-4) | 0.05 | 3 (2-4) | 3 (2-4) | 0.04 |
| No (%) by fifth of instability index: | | | | | | |
| Lowest | 1570 (19.5) | 5503 (21.2) | 0.01 | 1338 (19.9) | 1365 (20.3) | 0.01 |
| 2 | 1813 (22.5) | 5682 (21.9) | | 1514 (22.5) | 1508 (22.5) | |
| 3 | 1612 (20.0) | 5030 (19.4) | | 1313 (19.6) | 1314 (19.6) | |
| 4 | 1581 (19.7) | 4738 (18.2) | | 1303 (19.4) | 1313 (19.6) | |
| Highest | 1464 (18.2) | 5025 (19.3) | | 1248 (18.6) | 1216 (18.1) | |
| Median (IQR) instability index | 3 (2-4) | 3 (2-4) | 0.01 | 3 (2-4) | 3 (2-4) | 0.01 |
| No (%) by fifth of dependency index: | | | | | | |
| Lowest | 893 (11.1) | 3331 (12.8) | 0.07 | 773 (11.5) | 772 (11.5) | 0.01 |
| 2 | 1381 (17.2) | 4631 (17.8) | | 1163 (17.3) | 1193 (17.8) | |
| 3 | 1560 (19.4) | 5219 (20.1) | | 1302 (19.4) | 1330 (19.8) | |
| 4 | 1697 (21.1) | 5430 (20.9) | | 1390 (20.7) | 1353 (20.1) | |
| Highest | 2509 (31.2) | 7367 (28.4) | | 2088 (31.1) | 2068 (30.8) | |
| Median (IQR) dependency index | 4 (2-5) | 3 (2-5) | 0.07 | 4 (2-5) | 4 (2-5) | 0.01 |
| No (%) by fifth of deprivation index: | | | | | | |
| Lowest | 1747 (21.7) | 7225 (27.8) | 0.15 | 1499 (22.3) | 1468 (21.9) | 0.02 |
| 2 | 1904 (23.7) | 6289 (24.2) | | 1604 (23.9) | 1557 (23.2) | |
| 3 | 1847 (23.0) | 5426 (20.9) | | 1537 (22.9) | 1548 (23.0) | |
| 4 | 1434 (17.8) | 4205 (16.2) | | 1192 (17.7) | 1228 (18.3) | |
| Highest | 1108 (13.8) | 2833 (10.9) | | 884 (13.2) | 915 (13.6) | |
| Median (IQR) deprivation index | 3 (2-4) | 2 (1-4) | 0.15 | 3 (2-4) | 3 (2-4) | 0.02 |

Table 4| **Comparison of comorbidities and admission characteristics in THA recipients, before and after matching**

| | Before matching | | | After matching | | |
|---|---|---|---|---|---|---|
| | ≤35 procedures | >35 procedures | Standardized difference | ≤35 procedures | >35 procedures | Standardized difference |
| **Comorbidities (No (%))** | | | | | | |
| Frail | 522 (5.5) | 1431 (5.0) | 0.02 | 354 (5.3) | 337 (5.0) | 0.01 |
| Congestive heart failure | 545 (5.7) | 1388 (4.9) | 0.04 | 367 (5.5) | 368 (5.5) | 0.00 |
| Chronic kidney disease | 236 (2.5) | 676 (2.4) | 0.01 | 180 (2.7) | 168 (2.5) | 0.01 |
| COPD | 1736 (18.3) | 4381 (15.4) | 0.08 | 1191 (17.7) | 1238 (18.4) | 0.02 |
| Diabetes | 1531 (16.1) | 4137 (14.6) | 0.04 | 1063 (15.8) | 1075 (16.0) | 0.00 |
| Hypertension | 6115 (64.4) | 16 750 (59.0) | 0.11 | 4315 (64.2) | 4383 (65.3) | 0.02 |
| No (%) by Charlson score: | | | | | | |
| 0 | 8572 (90.3) | 25 928 (91.3) | 0.03 | 6110 (91.0) | 6119 (91.1) | 0.01 |
| 1 | 475 (5.0) | 1225 (4.3) | | 302 (4.5) | 300 (4.5) | |
| ≥2 | 447 (4.7) | 1234 (4.3) | | 304 (4.5) | 297 (4.4) | |
| Median (IQR) Charlson score | 0 (0-0) | 0 (0-0) | 0.02 | 0 (0-0) | 0 (0-0) | 0.00 |
| **Admission characteristics** | | | | | | |
| (No (%) of teaching hospitals | 1303 (13.7) | 11 675 (41.1) | 0.60 | 829 (12.3) | 859 (12.8) | 0.01 |
| Median (IQR) hospital volume | 159 (98-213) | 237 (170-395) | 0.69 | 165 (103-214) | 167 (119-211) | 0.04 |
| Median (IQR) surgeon experience (years) | 18 (8-26) | 20 (12-27) | 0.16 | 18 (9-27) | 19 (10-25) | 0.01 |

COPD=chronic obstructive pulmonary disease.

Table 5| **Proportion of arthoplasty recipients with specific complications. Figures are numbers (percentage) of patients**

| | Surgeon volume/year | | |
|---|---|---|---|
| | ≤35 procedures | >35 procedures | P value* |
| Venous thromboembolism within 90 days | 113 (1.7) | 95 (1.4) | 0.21 |
| Death within 90 days | 44 (0.7) | 30 (0.5) | 0.10 |
| Revision within 2 years | 98 (1.5) | 69 (1.0) | 0.03 |
| Dislocation within 2 years | 126 (1.9) | 86 (1.3) | 0.006 |
| Infection within 2 years | 70 (1.0) | 72 (1.1) | 0.87 |
| Periprosthetic fracture within 2 years | 21 (0.3) | 20 (0.3) | 0.88 |

*McNemar's test.

# Figures

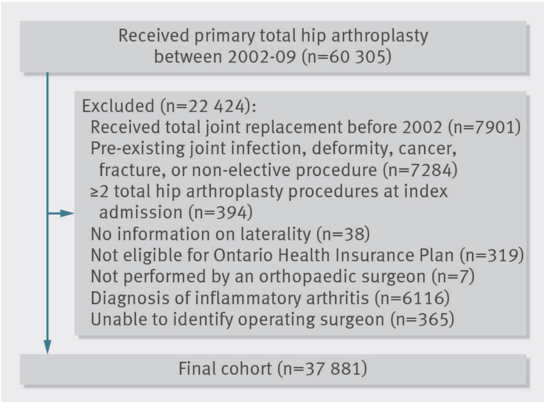


**Fig 1** Selection of patients for inclusion in study of effect surgeon volume on risk of complications after total hip arthroplasty

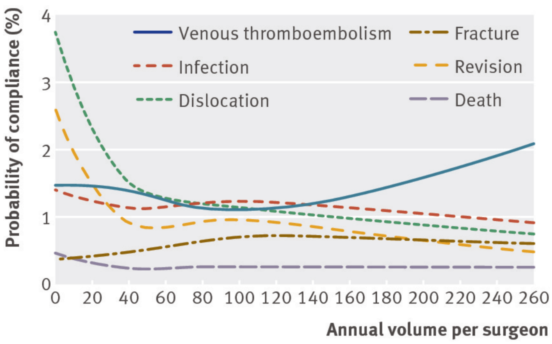


**Fig 2** Probability of specific complications after total hip arthroplasty according to surgeon volume

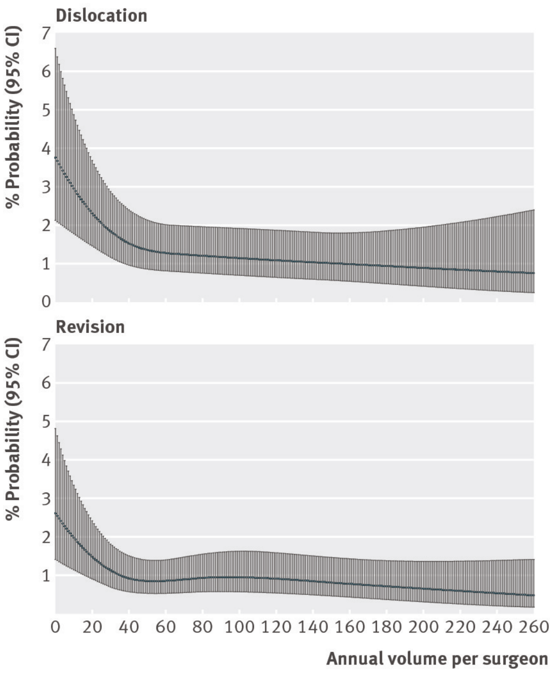


**Fig 3** Probability of dislocation and revision after total hip arthroplasty according to surgeon volume

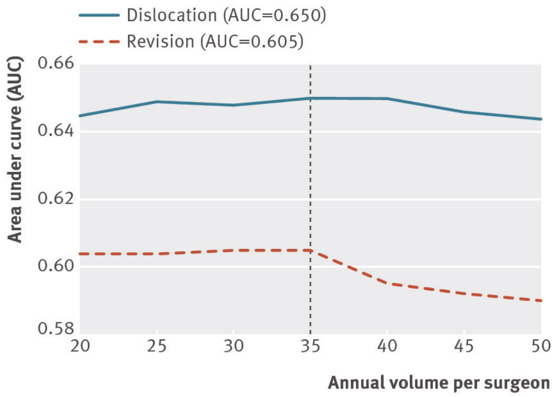


**Fig 4** Area under curve in multivariate models for various cut points of surgeon volume