Surgical outcomes for low volume versus high volume surgeons in Gynecology surgery: a systematic review and meta-analysis

Dr Alex Mowat, MBBS FRANZCOG, A/Professor Christopher Maher, FRANZCOG CU, Dr Emma Ballard, B.App.Sc (Hon) PhD

PII: S0002-9378(16)00390-2
DOI: 10.1016/j.ajog.2016.02.048
Reference: YMOB 10972


Received Date: 6 February 2016
Accepted Date: 24 February 2016


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Title: Surgical outcomes for low volume versus high volume surgeons in Gynecology surgery: a systematic review and meta-analysis.

Trial Registration No. CRD42015026154 (PROSPERO register)

Authors:
Dr Alex MOWAT. Royal Brisbane and Women’s Hospital, University of Queensland. MBBS FRANZCOG.
A/Professor Christopher MAHER. Royal Brisbane and Women’s Hospital, University of Queensland. FRANZCOG CU.
Dr Emma BALLARD. Queensland Institute of Medical Research. B.App.Sc (Hon) PhD

Corresponding author:
Dr Alex MOWAT. Butterfield St., Herston 4006 Qld, Australia. ph(m) +61 404 303 334. Email zanhmowat@gmail.com

The authors report no conflicts of interest or disclosures.

Word Count: Abstract 370, Main text 4132
Mowat, A

25 **Short Title:** Low volume versus high volume surgery in Gynecology
26 **Condensation:** Gynecologists performing surgery approximately once a month or
27 less have greater morbidity than their high volume colleagues.
Abstract

Objective: The aim of this study was to determine the impact of gynecologic surgeon volumes on patient outcomes.

Data Sources: Eligible studies were selected through an electronic literature search from database inception up until September 2015 and references in published studies. Search terms included “surgical volume,” “surgeon volume,” “low-volume OR high-volume,” “gynecology OR hysterectomy OR sling OR pelvic floor repair OR continence procedure”.

Study Eligibility: The literature search was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We defined a low volume surgeon (LVS) as one performing the procedure once a month or less and studies were excluded if their definition of LVS was > +/- 33% of our definition. Primary outcomes were total complications, intraoperative complications and postoperative complications.

Study appraisal and synthesis methods

All outcome data for individual studies were entered into Review Manager 5.3 systematic review software. When two or more studies evaluated a designated outcome meta-analysis of the entered data was undertaken as per the Cochrane database methodology. Data analysis was entered into GRADEpro, software, which generated a Summary of Findings (SOF) table that included structured and qualified grading (very low to high) of the quality for the evidence of the individual outcomes and provided a measure of effect.

Results: Fourteen peer-reviewed studies with 741 760 patients were included in the systematic review. For gynecology the low volume surgeon (LVS) group had an increased rate of total complications; OR 1.3 95% CI 1.2 to 1.5, intraoperative
complications; OR 1.6 95% CI 1.2 to 2.1 and postoperative complications; OR 1.4 95% CI 1.3 to 1.4. In gynecologic oncology the LVS group had higher mortality; OR 1.9 95% CI 1.3 to 2.6. In the urogynecology group a single study reported that the LVS group had a higher rate of any complication; RR 1.4 95% CI 1.2-1.6. Another single study found that LVS had higher rates of reoperation for mesh complications after mid-urethral sling procedures; RR 1.4 95% CI 1.2 to 1.5. The evidence is of moderate to very-low quality.

Conclusions: Gynecologists performing procedures approximately once a month or less were found to have higher rates of adverse outcomes in gynecology, gynecologic oncology and urogynecology, with higher mortality in gynecologic oncology.

Key words: gynecology, outcomes, surgeon volume
Introduction

In his popular book *Outliers*\(^1\), Malcolm Gladwell famously argued that the mastering of any skill requires 10 000 hours, or twenty hours a week for ten years, of deliberate practice. It seems a feasible theory that to obtain a skill with a repetitive technique, be it surgery or landing a plane, recurrent practice is required. In Australia, airline pilots for example, are required to land at least three times every 90 days to maintain their proficiency certificates\(^2\).

In many surgical fields the relationship between surgical volume and outcome is well established. An American study of greater than 470 000 Medicare patients undergoing either cardiovascular procedures or cancer resections found that the operative mortality rate was strongly and inversely related to surgeon volume for each procedure.\(^3\)

Several papers have examined this relationship in the field of gynecology, specifically looking at mid-urethral slings\(^4,5\), pelvic reconstructive procedures\(^6\), hysterectomies for benign indications\(^7-11\), myomectomies\(^12,13\) and gynecologic-oncological procedures\(^14-21\) and have reported conflicting results. A 2013 review article, without meta-analysis, of surgeon volumes and outcomes for benign hysterectomy concluded that morbidity was higher for low volume surgeons and high volume surgeons were more efficient\(^22\).

The lifetime risk of undergoing major gynecologic surgery in many developed countries is in the order of 15-40\%.\(^23,24,25\) Estimating the association between adverse outcomes and risk factors that can potentially be addressed through practice or policy changes, such as surgeon volume, is an important public health concern.
We performed a systematic review and meta-analysis to determine if gynecologic surgeon volumes impacted upon patient outcomes. Our null hypothesis is that surgeon volume had no impact on surgical outcomes or surgical efficiency.

**Methods**

**Eligibility**

Eligible studies were selected through an electronic literature search from inception up until September 2015 using Pubmed, the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, Medline and clinicaltrials.gov. Search terms included the following keywords “surgical volume,” “surgeon volume,” “low-volume OR high-volume,” “gynecology OR hysterectomy OR sling OR pelvic floor repair OR continence procedure”. There were no exclusion criteria for language or geographic location. Additional records were identified from references of articles identified through database searching.

**Study Selection**

The literature search was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and included prospective and retrospective studies that compared surgical complications or markers of surgical efficiency between high volume surgeons (HVS) and low volume surgeons (LVS) for any major gynecological procedure. All the studies involved women over 18 years of age undergoing major gynecological surgery. We defined a low volume surgeon as one performing the procedure once a month or less (twelve or less procedures a year) and studies were excluded if their definition of a LVS was > +/- 33% of our definition (range 8 to 16). Studies which divided surgeons into low, medium and high volume surgeons were included if two of the groups could be merged to fit our inclusion criteria.
Primary outcomes were total complications, intraoperative complications and/or postoperative complications. Secondary outcomes were mortality, medical complications, cystotomy, ureteric injury, bowel injury, vascular injury, transfusion rates, operating time, length of stay, estimated blood loss, readmission rates and reoperation rates. For the primary outcome of total complications, studies were included if they provided data for an outcome of "total complications", "total morbidity" or "any complication". Intraoperative complications included ureteric, bladder, bowel, vascular and other intraoperative injuries. Postoperative complications included wound complications (including vault hematoma), hemorrhage, ileus, bowel obstruction and venous thromboembolism (VTE). Medical complications included cardiopulmonary arrest, stroke, respiratory failure, pneumonia, renal failure, gastrointestinal complications, sepsis, fever and urinary tract infections. Papers that did not include any of these outcomes were excluded. Gynecology has a wide range of surgical interventions and analysis was divided into gynecology, gynecologic oncology and urogynecology. Before data extraction the review was registered with PROSPERO International Prospective Register of Systematic Reviews (registration no. CRD42015026154).

**Data extraction and analysis**

Data extraction was undertaken independently by two reviewers to ensure accuracy. If disagreement occurred a decision was made by mutual agreement. Outcome data for individual studies were entered into Review Manager 5.3 systematic review software. When two or more studies evaluated a designated outcome meta-analysis was performed as per the Cochrane methodology. For analysis of the categorical variables we calculated the odds ratio (OR; odds of women with a certain outcome in relation to the odds of women without the outcome in the group). For continuous
variables we used means and standard deviations to derive a mean difference (MD).

Unless otherwise stated the outcomes in this meta-analysis were calculated from the raw data reported in the papers and presented without adjustment for confounders.

Where possible, adjusted odds ratios and risk ratios were combined to give outcomes adjusted for possible confounders, including patient age, body mass index (BMI) and co-morbidities. If there was significant heterogeneity in the outcomes recorded in different studies as defined by the $I^2$ calculation being greater than 75%, a random effects model was used, otherwise a fixed-effect model was used for the calculation of summary estimates and their 95% confidence intervals (CI). Data analysis was entered into GRADEpro software, which generated a Summary of Findings (SOF) table that included structured and qualified grading (very low to high) of the quality for the evidence of the individual outcomes and provided a measure of effect.

**Results**

**Study selection and characteristics**

A total of 2151 abstracts met the initial search criteria. Of those 2123 were excluded by reviewers because they did not meet the pre-defined criteria. Twenty-eight full articles were assessed for eligibility and 14 were excluded for not meeting the defined inclusion criteria as outlined in the PRISMA flow study. (Fig 1). Fourteen peer-reviewed studies from three countries (The Netherlands, The United States of America and Canada) with a total of 741 760 patients were included in the systematic review. The two urogynecology studies were unable to be combined and so twelve studies were combined in meta-analysis. Two studies (Wallenstein, Roga-Gupta) potentially included the same patients from the Premier (Perspective) database between 2004 and 2007 for the three outcomes intraoperative
Mowat, A

complications, postoperative complications and medical complications. To minimize the risk of duplication, data from Roga-Gupta et al was excluded from the analysis for these three outcomes.

In the gynecology group five studies evaluated hysterectomy\textsuperscript{7,8,9,10,11} and two evaluated myomectomy\textsuperscript{12,13}. In the gynecologic oncology group three studies reported on endometrial cancer\textsuperscript{14,15,16} and two on ovarian cancer\textsuperscript{16,17}. In the urogynecology group one study evaluated pelvic reconstructive surgeries\textsuperscript{6} and another evaluated reoperation rates after mid-urethral sling surgery\textsuperscript{5}.

Patient characteristics including age and co-morbidities were reported in nine of the 14 articles and in six studies\textsuperscript{5,7,9,11,13,15} the HVS group had older women and/or women with more co-morbidities, in one study the LVS group had older women with more co-morbidities\textsuperscript{8} and in two studies the pre-intervention groups were similar\textsuperscript{12,14}.

Table 1.

Synthesis of results

1. Low volume surgeon versus high volume surgeon and outcomes in gynecology.

\textit{Total in-hospital complications}

Low volume surgeons had a higher rate of total in-hospital complications than high volume surgeons as reported in four studies\textsuperscript{7,8,10,11}; OR 1.3 95\% CI 1.2 to 1.5, SOF Table 1. This means that if in-hospital complications occur in 97 per 1000 patients in the HVS group, between 114-137 per 1000 patients in the LVS group would develop in-hospital complications. Two studies\textsuperscript{10,11} provided data for total in-hospital complications adjusted for age and comorbidities and the increased risk of any in-hospital complication in the LVS group was slightly greater; OR 1.4 95\% CI 1.3 to 1.5, SOF Table 1. On a number needed to treat analysis this translated to one in-
hospital complication being avoided for every 30 operations that were performed by a high volume surgeon rather than a low volume surgeon. Hanstede\textsuperscript{11} re-analysed their data \textit{excluding gynecologic oncologists} and reported a greater difference between the LVS and HVS groups OR of 2.8 95% CI 2.1 to 3.6, adjusted for age and comorbidities. On a number needed to treat analysis this translated to one in-

\textit{Intraoperative complications}

Three studies\textsuperscript{8,10,11} reported on this outcome and the LVS group had a higher rate of intraoperative complications compared with the HVS group; OR 1.6 95% CI 1.2 to 2.1, SOF Table 1. This means that if intraoperative complications occur in 22 per 1000 patients in the HVS group, between 26-45 per 1000 patients in the LVS group would develop intraoperative complications. Two studies\textsuperscript{9,11} calculated an OR adjusted for age and comorbidities and a further increase in the risk of intraoperative complications was seen in the LVS group compared to the HVS group; OR 1.8 95% CI 1.1 to 3.2, SOF Table 1. On a number needed to treat analysis this translates to one intraoperative complication being avoided for every 38 operations that are performed by a high volume surgeon rather than a low volume surgeon.

Hanstede\textsuperscript{11} re-analysed their data, \textit{without gynecologic oncologists}, and the greater risk of intraoperative complications in the LVS group was more evident; OR 3.4 95% CI 2.0 to 5.9, SOF Table 1. On a number needed to treat analysis this translates to one intraoperative complication being avoided for every 20 operations that are performed by a high volume surgeon rather than a low volume surgeon.

\textit{Postoperative complications}
Four studies\textsuperscript{7,8,10,11} reported on this outcome and the LVS group had a higher rate of postoperative complications compared with the HVS group; OR 1.4 95% CI 1.3 to 1.4, SOF Table 1. This means that if postoperative complications happen in 39 per 1000 patients in the HVS group, between 50-54 per 1000 patients in the LVS group would develop postoperative complications. Two studies\textsuperscript{9,11} calculated an OR adjusted for age and comorbidities and a further increase in the risk of postoperative complications was seen when comparing the LVS and HVS groups; OR 1.5 95% CI 1.2 to 1.7, SOF Table 1. On a number needed to treat analysis this translates to one postoperative complication being avoided for every 41 operations that are performed by a high volume surgeon rather than a low volume surgeon.

Hanstede\textsuperscript{11} re-analysed, \textit{without gynecologic oncologists}, and the difference between the two groups was more pronounced; OR 2.4 95% CI 1.8 to 3.2, SOF Table 1. On a number needed to treat analysis this translates to one postoperative complication being avoided for every 15 operations that are performed by a high volume surgeon rather than a low volume surgeon.

\textbf{Mortality}

There was no difference in mortality between the two groups.

\textbf{Medical complications}

Three studies\textsuperscript{8,9,10,11} reported on medical complications and found that medical complications were more common in the LVS group compared with the HVS group; OR 1.6 95% CI 1.5 to 1.6, SOF Table 2. This means that if medical complications happen in 79 per 1000 patients in the HVS group, between 115-122 per 1000 women in the LVS group would develop medical complications.

\textbf{Visceral and Vascular Injuries}
Ureteric injury was reported in two studies\textsuperscript{8,10} and was more likely in the LVS group; OR 1.7 95% CI 1.4 to 2.1, SOF Table 2. Bowel injury was reported in two studies\textsuperscript{8,10} and the risk was higher in the LVS group; OR 1.1 95% CI 1.1 to 1.2, SOF Table 2. Urinary tract injury (ureteric and bladder injury combined) was reported in three studies\textsuperscript{8,10,11} and was more likely in the LVS group; OR 1.4 95% CI 1.1 to 1.9, SOF Table 2.

There was no difference in risk of cystotomy or vascular injury, SOF Table 2.

\textbf{Operating time}

Four studies\textsuperscript{7,11,12,13} reported on this outcome and found that operating time was longer in the LVS group; MD 17.7 minutes 95% CI 10.4 to 25.0, SOF Table 2.

\textbf{Estimated blood loss and transfusion rates}

Two studies\textsuperscript{7,13} reported that estimated blood loss was greater in the LVS group; MD 59.3 ml 95% CI 32.0 to 86.6mls, SOF Table 2. There was no difference in transfusion rates.

\textbf{Length of stay}

There was no difference in mean length of stay between the HVS group and the LVS group. Wallenstein\textsuperscript{8} found that women in the LVS group were more likely to stay in hospital for over two days; OR 1.4 95% CI 1.3 to 1.4, SOF Table 2. Boyd\textsuperscript{10} reported that women in the HVS group had a shorter LOS by 0.4 days 95% CI 0.4 to 0.5, adjusted for mode of hysterectomy, comorbidities and post-operative complications.

\textbf{Readmission rates}

Readmission rates were reported in three studies\textsuperscript{7,9,11} and were lower for the LVS group; OR 0.8 95%CI 0.7 to 0.9, SOF Table 2. This means that if 20 in 1000 patients are readmitted in the HVS group, 14-18 in 1000 patients would be readmitted in the LVS group.
Reoperation rates

There was no difference in reoperation rates.

2. Low volume surgeon versus high volume surgeon and outcomes in gynecologic oncology.

Mortality

Mortality was reported in four studies\textsuperscript{14,15,16,18} and was higher in the LVS group compared with the HVS group; OR 1.9 95% CI 1.3 to 2.6, SOF Table 3. This means that if the mortality rate is 7 per 1000 patients in the HVS group, the rate would be between 9-18 per 1000 in the LVS group. Three of these studies\textsuperscript{14,16,18} adjusted the outcomes for age and comorbidities and the difference between the two groups became more significant; OR 2.5 95% CI 1.7 to 3.8, SOF Table 3. On a number needed to treat analysis this translates to one perioperative death being avoided for every 97 operations that are performed by a high volume surgeon rather than a low volume surgeon.

Complications

Two studies\textsuperscript{14,15} reported on complications. There was no difference in total in-hospital complications. Intraoperative complications were higher in the LVS group than in the HVS group; OR 1.2 95% CI 1.1 to 1.5. This means that if intraoperative complications occur in 62 per 1000 patients in the HVS group, between 67-87 per 1000 patients in the LVS group would develop intraoperative complications. In-hospital postoperative complications occurred more often in the LVS group than in the HVS group; OR 1.2 95% CI 1.1 to 1.4. This means that if in-hospital postoperative complications occur in 110 per 1000 patients in the HVS group, between 120-144 per 1000 patients in the LVS group would develop in-hospital postoperative complications, SOF Table 3.
Length of Stay

A single study by Wright\textsuperscript{15} looked at this outcome and found that patients in the LVS group were more likely to stay over two days when compared with patients in the HVS group; OR 1.3 95% CI 1.1 to 1.6, SOF Table 3.

Transfusion rates

Two studies\textsuperscript{14,15} reported on this outcome and found that blood transfusions were required more often in the HVS group than in the LVS group; OR 0.7 95% CI 0.6 to 0.8. This means that if transfusions are required in 71 per 1000 patients in the HVS group, they would be required in between 43-60 per 1000 patients in the LVS group, SOF Table 3.

5-year survival

Vernooij\textsuperscript{17} looked at 5-year survival in ovarian cancer patients. Results were presented as hazard ratios adjusted for age and stage of cancer and it was reported that surgery by a HVS reduced mortality by 29% (HR 0.7 95% CI 0.5-1.0).

3. Low volume surgeon versus high volume surgeon and outcomes in urogynecology.

Due to the heterogeneity of both datum format and outcomes, we were unable to combine the two studies in this group. One study by Sung\textsuperscript{6} evaluated the impact of low versus high volume surgeons on complications in pelvic reconstructive surgery. Raw data was not presented in this paper and was not able to be provided by author when contacted by email. Thus we were unable to combine the findings of this study with those of any other. In this study from the United States of America 310 759 pelvic reconstructive surgeries were evaluated comparing surgeons who performed less than eight operations a year (low volume surgeons), 8-18 operations a year (medium volume surgeons, MVS) and those who performed over 18 operations a year...
year (high volume surgeons). Women in the LVS and MVS groups had a higher rate of any complication (RR 1.4 95% CI 1.2 to 1.6, RR 1.2 95% CI 1.1 to 1.4 respectively) compared with the HVS group when adjusted for age and comorbidities. Women in the LVS and MVS groups also had a higher risk of a non-routine discharge defined as patients transferred to a skilled nursing facility or other short-term facility (RR 2.0 95% CI 1.4 to 2.5, RR 1.6 95% CI 1.1 to 2.1) when adjusted for age and comorbidities.

A large Canadian population-based study (n=59,887) by Welk evaluated the impact of surgeon volume on the rate of reoperation for mesh complications after mid-urethral sling procedures over a 10 year period. The specific nature of the outcome meant that this study was unable to be combined with other studies. A low volume surgeon was defined as one who performed 16 or fewer mid-urethral sling procedures per year. The LVS group had a higher rate of reoperation than the HVS group; RR 1.4 95% CI 1.2 to 1.5. This means that if reoperation is performed in 20 in 1000 patients in the HVS group, reoperation would be performed in 24-30 in 1000 patients in the LVS group. The result was unchanged when multivariable analysis was performed to account for possible confounders.

Comment

Main Findings

We demonstrated a 30% increase in the risk of experiencing any in-hospital complication, a 60% increase in the risk of incurring an intraoperative complication, a 40% increase in the risk of incurring an in-hospital postoperative complication in gynecology, and a 90% increase in the mortality rate for gynecologic oncology in the LVS group as compared with the HVS group. After adjusting for possible confounders the magnitude of the effect between high and low volume surgeons was
increased such that there was a 40% increase in the risk of experiencing any in-hospital complication, an 80% increase in the risk of incurring an intraoperative complication, a 50% increase in the risk of incurring an in-hospital postoperative complication in gynecology and a 250% increase in the mortality rate for gynecologic oncology. This implies that HVS operate on patients with greater morbidities. Interestingly, in Wallenstein\textsuperscript{8} it was the low volume surgeon group which had the older and sicker patients. In that study a multivariate analysis was undertaken adjusting for patient characteristics and when they adjusted for these differences, there was little change in the relative risk ratios of the outcomes total in-hospital complications (1.3 95% CI 1.3 to 1.4; 1.3 95% 1.2 to 1.5), intraoperative complications (1.3 95% CI 1.2 to 1.4; 1.2 95% CI 1.0 to 1.4) and in-hospital postoperative complications (1.3 95% CI 1.2 to 1.4; 1.4 95% CI 1.2 to 1.7). This indicated that the pre-intervention difference in patient characteristics did not account for the higher morbidity in the low volume surgeon group as compared with the high volume surgeon group.

In the urogynecology group we demonstrated a 37% increase in the risk of reoperation for mesh complications after mid-urethral sling procedure in the LVS group compared with the HVS group.

While these findings are clinically relevant to the patient, some of our findings, although statistically significant, may not be clinically relevant. For example, in the gynecology group, a higher blood loss in the LVS group of 60mls is not of clinical relevance.

**Strengths and Limitations**

The strengths of this study are the large numbers and that data was collected from government databases which increased the strength and reliability of the data. The
Mowat, A

data from such databases are samples and are frequently without patient-specific
identifiers, meaning that the complication and the operation cannot be reliably linked.
However these concerns are nullified by the original trial methodologies ensuring
that any impact of this equally affects both the control and the intervention group.
A weakness of the paper is that the data only captures complications that are
managed in hospital and so is likely to underestimate the true incidence of
complications related to gynecologic surgery. However you would expect that the
more serious complications, that by their very nature require readmission, would be
captured. Five of the 14 studies failed to adjust for patient comorbidities as possible
confounders. When possible, adjusted data was combined in meta-analysis,
however the evidence would be strengthened if all studies had adjusted for possible
confounders. Only two studies in this group evaluated urogynecology and, although
the patient number is very large, this may limit the generalisability of our findings.
The authors accept that by allowing a range of values for the high and low volume
surgeon cut-off there may be an under-estimation of the volume-outcome effect at
the upper limit and an over-estimation at the lower limit. Large population based
studies retrieving data retrospectively appear to be the only reasonable avenue for
sourcing data on complication outcomes with a relatively low incidence. GRADEpro
rates all retrospective data as a priori “low grade” and downgrades it to “very low”
based on risk of bias, inconsistency, indirectness, imprecision, and publication bias,
and upgrades it to “moderate” based on large effect, plausible confounders and dose
response gradient. Randomised controlled trials would provide a higher grade of
evidence however, based on the present data, these may be unethical and
impractical. While this evidence is graded from very low to moderate resulting in
some uncertainty regarding the findings for the reasons outlined above, it is unlikely that a higher grade of evidence will become available.

Comparison with existing literature

The findings of this meta-analysis reflect those of studies in other fields of surgery which have concluded that surgical outcomes improve with surgeon volume. Studies have shown an inverse relationship between surgeon volume and mortality in the areas of cardiovascular procedures, colectomy, gastrectomy, esophagectomy, pancreatic resection, nephrectomy, cystectomy, lung lobectomy, and pneumonectomy\textsuperscript{3,26}. Further studies show a similar relationship between surgeon volume and reduced morbidity in the areas of colorectal procedures, esophagectomy, gastrectomy, pancreatectomy, thyroidectomy, coronary artery bypass graft surgery, and carotid endarterectomy\textsuperscript{27}.

Defining a cut-off for the definition to distinguish a low volume surgeon from a high volume surgeon was one of the challenges of this systematic review. We decided on a cut-off of one operation a month to reflect the definitions in the majority of the potential studies for inclusion and this definition also reflects a readily achievable number of procedures. To facilitate maximum inclusion of trials in the meta-analysis and in recognition of varying definitions of high and low volume surgeons by the contributing authors, we allowed for a range of plus or minus 33\% from 12 procedures/year (8-16). Despite varying complexities of the surgical procedures in the included studies we found this to be a consistent cut-off, suggesting that familiarity with surgical technique is important across a range of procedures in gynecology. There is no consensus in the surgical literature on appropriate volume cut-offs, but what is clear is that there is a dose-response relationship between surgeon volume and outcomes\textsuperscript{3,26,27} and a recent publication on surgeon volume and
outcomes for rectal cancer surgery defined a high volume surgeon as one
performing the procedure at least ten times a year, a definition which is compatible
with our own.\textsuperscript{28}

\textbf{Conclusion and implications}

There is significant morbidity in patients following gynecology surgery, with 10% of
all patients reporting in-hospital complications in this meta-analysis. Surgeons
performing procedures approximately once a month or less (range 8-16 procedures
a year) had higher rates of adverse outcomes in gynecology, gynecologic oncology
and urogynecology and a higher mortality rate and lower 5-year survival outcomes in
gynecologic oncology. While the grade of evidence for the majority of primary
outcomes is moderate, it is very low to low for most of the secondary outcomes, and
this should be reflected if this data were to be utilised for the formulation of
guidelines relating to the performance of gynecologic surgery.

This systematic review and meta-analysis finds that surgeon volume is an important
factor in surgical outcomes in gynecology. In the United States of America the
mounting evidence linking surgeon and hospital volume to patient outcomes,
including mortality, has led to initiatives which encourage patients to choose high
volume surgeons and hospitals for elective procedures.\textsuperscript{29} An example of this is the
Leapfrog Group for Patient Safety, a national consortium of private and public
purchasers of health insurance, which produces a website encouraging patients to
choose a surgeon and a hospital with procedure-specific experience.\textsuperscript{30} There is
literature indicating that high volume surgeons are increasingly performing a larger
proportion of elective surgery\textsuperscript{27}.

Further research is required to ensure that patient co-morbidities are fully controlled
for in the assessment of surgeon volume and outcomes, and to improve the
generalisability and quality of the evidence. Additionally, research is required to determine whether individual surgeon characteristics such as inherent surgical ability, training and total experience impact upon the number of surgical repetitions needed to minimise patient morbidity.

References
Mowat, A


Mowat, A


Mowat, A  


## Summary of Findings Table 1: Low volume compared to high volume surgeons in gynecology: Primary Outcomes

**Patient or population:** women undergoing major gynecology surgery  
**Intervention:** low volume surgeons  
**Control:** high volume surgeons

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Risk in high volume surgeon group</th>
<th>Risk in low volume surgeon group</th>
<th>Relative effect (95% CI)</th>
<th>No. of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total complications</td>
<td>97 per 1000</td>
<td>125 per 1000 (114 to 137)</td>
<td>OR 1.3 (1.2 to 1.5)</td>
<td>283 119 (4 studies)</td>
<td>⊕⊕⊕⊕ very low¹</td>
</tr>
<tr>
<td>Total complications adjusted OR</td>
<td></td>
<td></td>
<td>OR 1.4 (1.3 to 1.5)</td>
<td>153 660 (2 studies)</td>
<td>⊕⊕⊕ moderate²</td>
</tr>
<tr>
<td>Total complications adjusted OR excluding gynecologic oncology</td>
<td>68 per 1000</td>
<td>167 per 1000 (133 to 208)</td>
<td>OR 2.8 (2.1 to 3.6)</td>
<td>3427 (1 study)</td>
<td>⊕⊕ moderate³</td>
</tr>
<tr>
<td>Intraoperative complications</td>
<td>22 per 1000</td>
<td>35 per 1000 (26 to 45)</td>
<td>OR 1.6 (1.2 to 2.1)</td>
<td>358 296 (3 studies)</td>
<td>⊕⊕⊕⊕ very low⁴</td>
</tr>
<tr>
<td>Intraoperative complications adjusted OR</td>
<td></td>
<td></td>
<td>OR 1.8 (1.1 to 3.2)</td>
<td>84 275 (2 studies)</td>
<td>⊕⊕⊕ moderate²</td>
</tr>
<tr>
<td>Intraoperative complications adjusted OR excluding gynecologic oncology</td>
<td>15 per 1000</td>
<td>50 per 1000 (30 to 83)</td>
<td>OR 3.4 (2.0 to 5.9)</td>
<td>3427 (1 study)</td>
<td>⊕⊕⊕ moderate³</td>
</tr>
<tr>
<td>Postoperative complications</td>
<td>39 per 1000</td>
<td>52 per 1000 (50 to 54)</td>
<td>OR 1.4 (1.3 to 1.4)</td>
<td>359 528 (4 studies)</td>
<td>⊕⊕⊕ moderate²</td>
</tr>
<tr>
<td>Postoperative complications adjusted OR</td>
<td>OR 1.5 (1.2 to 1.7)</td>
<td>84,275 (2 studies)</td>
<td>⊕⊕⊕ moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
<td>---------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postoperative adjusted OR excluding gynecologic oncology</td>
<td>53 per 1000</td>
<td>117 per 1000 (89 to 153)</td>
<td>OR 2.4 (1.8 to 3.2)</td>
<td>3427 (1 study)</td>
<td>⊕⊕⊕ moderate</td>
</tr>
</tbody>
</table>

**CI:** Confidence interval; **OR:** Odds ratio

1. Imprecision: Significant heterogeneity with an I²=85%. Downgraded the quality rating by 1 level
2. Plausible Confounder: Inclusion of gynecologic-oncologists diminishes the volume. Upgraded the quality rating by 1 level
3. Magnitude effect: The magnitude of effect was large (OR ≥ 2). Upgraded the quality rating by 1 level
4. Imprecision: Significant heterogeneity with an I²=96%. Downgraded the quality rating by 1 level

**GRADE Working Group grades of evidence**

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.
Summary of Findings Table 2: Low volume compared to high volume surgeons in gynecology: Secondary Outcomes

Patient or population: women undergoing major gynecology surgery
Intervention: low volume surgeons
Control: high volume surgeons

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Risk in high volume surgeon group</th>
<th>Risk in low volume surgeon group</th>
<th>Relative effect (95% CI)</th>
<th>No. of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>0 per 1000 (0 to 2)</td>
<td>1 per 1000 (0 to 2)</td>
<td>OR 1.3 (0.4 to 4.7)</td>
<td>351 148 (3 studies)</td>
<td>⊕⊕⊕⊕, very low¹</td>
</tr>
<tr>
<td>Medical complications</td>
<td>79 per 1000 (115-122)</td>
<td>118 per 1000 (115-122)</td>
<td>OR 1.6 (1.5 to 1.6)</td>
<td>358 296 (3 studies)</td>
<td>⊕⊕⊕, low</td>
</tr>
<tr>
<td>Operating time (mins)</td>
<td>mean operating time (mins)</td>
<td>17.7 higher (10.4 to 25.0)</td>
<td></td>
<td>9335 (4 studies)</td>
<td>⊕⊕⊕, very low²</td>
</tr>
<tr>
<td>Transfusion</td>
<td>55 per 1000 (37 to 84)</td>
<td>56 per 1000 (37 to 84)</td>
<td>OR 1.0 (0.7 to 1.6)</td>
<td>234 203 (4 studies)</td>
<td>⊕⊕⊕⊕, very low³</td>
</tr>
<tr>
<td>Estimated blood loss (mls)</td>
<td>mean estimated blood loss (mls)</td>
<td>59.3 higher (32.0 to 86.6)</td>
<td></td>
<td>1754 (2 studies)</td>
<td>⊕⊕⊕, low</td>
</tr>
<tr>
<td>Cystotomy</td>
<td>8 per 1000 (8 to 12)</td>
<td>9 per 1000 (8 to 12)</td>
<td>OR 1.1 (0.9 to 1.4)</td>
<td>273 949 (2 studies)</td>
<td>⊕⊕⊕⊕, very low⁴</td>
</tr>
<tr>
<td>Ureteric Injury</td>
<td>1 per 1000 (1 to 2)</td>
<td>2 per 1000 (1 to 2)</td>
<td>OR 1.7 (1.4 to 2.1)</td>
<td>274 039 (2 studies)</td>
<td>⊕⊕⊕, low</td>
</tr>
<tr>
<td>Cystotomy or ureteric injury</td>
<td>9 per 1000 (10-17)</td>
<td>13 per 1000 (10-17)</td>
<td>OR 1.4 (1.1 to 1.9)</td>
<td>281 205 (3 studies)</td>
<td>⊕⊕⊕⊕</td>
</tr>
<tr>
<td>Condition</td>
<td>Incidence</td>
<td>Incidence</td>
<td>OR</td>
<td>N</td>
<td>Quality Rating</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------</td>
<td>-------</td>
<td>----------------</td>
</tr>
<tr>
<td>Bowel injury</td>
<td>12 per 1000</td>
<td>14 per 1000</td>
<td>OR 1.1</td>
<td>274,039</td>
<td>low (1.1 to 1.2)</td>
</tr>
<tr>
<td>Vascular injury</td>
<td>0 per 1000</td>
<td>1 per 1000</td>
<td>OR 2.2</td>
<td>131,781</td>
<td>low (0.4 to 11.6)</td>
</tr>
<tr>
<td>Readmission</td>
<td>20 per 1000</td>
<td>16 per 1000</td>
<td>OR 0.8</td>
<td>85,489</td>
<td>very low (0.7 to 0.9)</td>
</tr>
<tr>
<td>Reoperation</td>
<td>3 per 1000</td>
<td>2 per 1000</td>
<td>OR 0.9</td>
<td>124,615</td>
<td>very low (0.7 to 1.2)</td>
</tr>
<tr>
<td>LOS&gt;2 days</td>
<td>63 per 1000</td>
<td>84 per 1000</td>
<td>OR 1.4</td>
<td>124,615</td>
<td>low (1.3 to 1.4)</td>
</tr>
</tbody>
</table>

CI: Confidence interval; OR: Odds ratio

1 Imprecision: Significant heterogeneity with an \( \text{i}^2 = 82\% \). Downgraded the quality rating by 1 level
2 Imprecision: Significant heterogeneity with an \( \text{i}^2 = 83\% \). Downgraded the quality rating by 1 level
3 Imprecision: Significant heterogeneity with an \( \text{i}^2 = 98\% \). Downgraded the quality rating by 1 level
4 Imprecision: Significant heterogeneity with an \( \text{i}^2 = 85\% \). Downgraded the quality rating by 1 level
5 Imprecision: Significant heterogeneity with an \( \text{i}^2 = 88\% \). Downgraded the quality rating by 1 level
6 Imprecision: Significant heterogeneity with an \( \text{i}^2 = 74\% \). Downgraded the quality rating by 1 level

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.
Summary of Findings Table 3: Low volume compared to high volume surgeons in gynecologic oncology

**Patient or population:** women undergoing major gynecologic oncology surgery  
**Intervention:** low volume surgeons  
**Control:** high volume surgeons

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Risk in high volume surgeon group</th>
<th>Risk in low volume surgeon group</th>
<th>Relative effect (95% CI)</th>
<th>No. of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>7 per 1000 (9-18)</td>
<td>13 per 1000 (9-18)</td>
<td>OR 1.9 (1.3 to 2.6)</td>
<td>18 045 (4 studies)</td>
<td>⊙dì ⊙dì ⊙dì low</td>
</tr>
<tr>
<td>Mortality adjusted OR</td>
<td></td>
<td></td>
<td>OR 2.5 (1.7 to 3.8)</td>
<td>13 908 (3 studies)</td>
<td>⊙dì ⊙dì ⊙dì moderate³</td>
</tr>
<tr>
<td>Total complications</td>
<td>396 per 1000 (305 to 673)</td>
<td>488 per 1000 (305 to 673)</td>
<td>OR 1.5 (0.7 to 3.1)</td>
<td>10 152 (2 studies)</td>
<td>⊙dì ⊙dì ⊙dì very low²</td>
</tr>
<tr>
<td>Intraoperative complications</td>
<td>62 per 1000 (67 to 87)</td>
<td>75 per 1000 (67 to 87)</td>
<td>OR 1.2 (1.1 to 1.5)</td>
<td>10 152 (2 studies)</td>
<td>⊙dì ⊙dì ⊙dì low</td>
</tr>
<tr>
<td>Postoperative complications</td>
<td>110 per 1000 (120 to 144)</td>
<td>128 per 1000 (120 to 144)</td>
<td>OR 1.2 (1.1 to 1.4)</td>
<td>10 152 (2 studies)</td>
<td>⊙dì ⊙dì ⊙dì low</td>
</tr>
<tr>
<td>Medical complications</td>
<td>232 per 1000 (149 to 446)</td>
<td>274 per 1000 (149 to 446)</td>
<td>OR 1.3 (0.6 to 2.7)</td>
<td>10 152 (2 studies)</td>
<td>⊙dì ⊙dì ⊙dì very low³</td>
</tr>
<tr>
<td>Transfusion</td>
<td>71 per 1000 (43 to 60)</td>
<td>51 per 1000 (43 to 60)</td>
<td>OR 0.7 (0.6 to 0.8)</td>
<td>10 152 (2 studies)</td>
<td>⊙dì ⊙dì ⊙dì very low⁴</td>
</tr>
<tr>
<td>Cystotomy</td>
<td>8 per 1000 (7-17)</td>
<td>11 per 1000 (7-17)</td>
<td>OR 1.3 (0.9 to 2.0)</td>
<td>10 152 (2 studies)</td>
<td>⊙dì ⊙dì ⊙dì low</td>
</tr>
<tr>
<td>Ureteric Injury</td>
<td>17 per 1000 (11 to 22)</td>
<td>15 per 1000 (11 to 22)</td>
<td>OR 0.9 (0.6 to 1.3)</td>
<td>10 152 (2 studies)</td>
<td>⊙dì ⊙dì ⊙dì low</td>
</tr>
<tr>
<td>Cystotomy or ureteric injury</td>
<td>13 per 1000 (10 to 30)</td>
<td>18 per 1000 (10 to 30)</td>
<td>OR 1.4 (0.8 to 2.4)</td>
<td>4137 (1 study)</td>
<td>⊙dì ⊙dì ⊙dì low</td>
</tr>
<tr>
<td>Bowel injury</td>
<td>16 per 1000</td>
<td>21 per 1000</td>
<td>OR 1.3</td>
<td>10 152</td>
<td>⊙dì ⊙dì ⊙dì low</td>
</tr>
<tr>
<td>Event</td>
<td>Incidence</td>
<td>CI</td>
<td>OR</td>
<td>Studies</td>
<td>Quality Rating</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td>Vascular injury</td>
<td>1 per 1000</td>
<td>(15 to 29)</td>
<td>0.6</td>
<td>2 studies</td>
<td>low</td>
</tr>
<tr>
<td></td>
<td>1 per 1000</td>
<td>(0 to 2)</td>
<td>(0.1 to 2.8)</td>
<td>(2 studies)</td>
<td></td>
</tr>
<tr>
<td>LOS &gt; 2 days</td>
<td>122 per 1000</td>
<td>(129 to 179)</td>
<td>1.3</td>
<td>1 study</td>
<td>low</td>
</tr>
<tr>
<td></td>
<td>152 per 1000</td>
<td>(129 to 179)</td>
<td>(1.1 to 1.6)</td>
<td>(1 study)</td>
<td></td>
</tr>
<tr>
<td>Reoperation</td>
<td>9 per 1000</td>
<td>(4 to 12)</td>
<td>0.8</td>
<td>2 studies</td>
<td>low</td>
</tr>
<tr>
<td></td>
<td>7 per 1000</td>
<td>(4 to 12)</td>
<td>(0.5 to 1.3)</td>
<td>(2 studies)</td>
<td></td>
</tr>
</tbody>
</table>

CI: Confidence interval; OR: Odds ratio

1. Magnitude effect: The magnitude of effect was large (OR ≥ 2). Upgraded the quality rating by 1 level
2. Imprecision: Significant heterogeneity with an I² = 98%. Downgraded the quality rating by 1 level
3. Imprecision: Significant heterogeneity with an I² = 96%. Downgraded the quality rating by 1 level
4. Significant heterogeneity with an I² = 85%. Downgraded the quality rating by 1 level

GRADE Working Group grades of evidence
- **High quality**: Further research is very unlikely to change our confidence in the estimate of effect.
- **Moderate quality**: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- **Low quality**: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- **Very low quality**: We are very uncertain about the estimate.
2142 records identified through database searching

10 additional records identified from references of articles identified through database searching

2151 records after duplicates removed

2151 records screened

2123 records excluded as did not meet pre-defined criteria (i.e. not on topic)

28 full-text articles assessed for eligibility

14 excluded

- 2 hospital volume rather than surgeon volume
- 4 due to volume being total experience rather than annual volume
- 1 review
- 1 inappropriate subject matter (comparing gynecologists and urologists)
- 3 due to not meeting volume definition eligibility criteria
- 2 due to not meeting outcome definition eligibility criteria
- 1 due stating the results in the text but not providing any raw data - did not respond to email request

14 studies included in qualitative synthesis (systematic review)
Figure 1: PRISMA flow chart

12 studies included in quantitative synthesis (meta-analysis)