Meta-analysis and systematic review of the relationship between surgeon annual caseload and mortality for elective open abdominal aortic aneurysm repairs

Emily L. Young, MA, MBBS,a Peter J.E. Holt, MRCS,a Jan D. Poloniecki, PhD,b Ian M. Loftus, MD, FRCS,a and Matt M. Thompson, MD, FRCS,a London, United Kingdom

Objective: To assess the relationship between the annual caseload of elective open abdominal aortic aneurysm (AAA) repairs performed by individual surgeons and mortality.

Methods: PubMed, EMBASE, and the Cochrane library were searched for articles on the volume-outcome relationship in AAA surgery. The review conformed to the QUOROM statement. The data were meta-analyzed to compare the mortality rates of higher- and lower-volume surgeons. A critical volume threshold was calculated for better practice.

Results: Fourteen relevant articles were retrieved from the searches. A systematic review was performed, and six were meta-analyzed. A total of 115,273 elective open AAA repairs were considered, with a mean mortality rate of 5.56%. Significant relationships between higher surgeon caseload and lower mortality were demonstrated in 12 of 14 studies. From the meta-analysis, the pooled effect estimate was an odds ratio of 0.56 (95% confidence interval, 0.54-0.57) in favor of higher-volume surgeons. A critical volume threshold was identified as 13 cases per annum for individual surgeons.

Conclusions: As surgeons performed higher annual volumes of elective open AAA repairs, significantly lower mortality rates were demonstrated. Surgeons wishing to perform elective AAA repairs should achieve a minimum case volume of 13 repairs per annum. (J Vasc Surg 2007;46:1287-94.)

Elective open abdominal aortic aneurysm (AAA) repair is a common procedure, with 5000 operations per annum in England. The procedural mortality for elective repairs has been established as approximately 5% in trial hospitals, but a recent study of outcomes from this surgery using current data showed that the population mortality rate was higher than this estimate, at 7.4% in England.

Patients with aortic aneurysmal disease often present with significant medical comorbidity, which may be amenable to modification in the long term to slow disease progression. These patients are often in higher-risk categories for other nonmodifiable operative risk factors, such as age, sex, race, and socioeconomic status.

However, there is significant evidence that factors beyond patient-level factors have an effect on the mortality rate of elective AAA surgery. Both hospital- and surgeon-level data may provide a number of modifiable risk factors for surgery, in addition to the issues of case-mix. Notable among these is the proven relationship between the annual volume of surgery performed by hospitals and outcome, with increasing annual volumes being associated with improved outcomes.

It is possible that surgeon volume and subspecialization may affect mortality, in addition to hospital infrastructure and hospital volume. Therefore, the aim of this study was to review the evidence for a relationship between the annual volume of surgery undertaken by individual surgeons and outcome, in terms of mortality, for elective open AAA repairs.

METHODS

After submission to the ethics committee, it was confirmed that no ethics approval was required for this study.

Literature search. PubMed, EMBASE, and the Cochrane library medical databases were searched for all articles relating to AAA, surgeon volume, and outcome. The search terms used were AAA or abdominal aortic aneurysm and surgeon volume or surgeon and outcome or mortality. A final term of “vascular surgery and mortality” was included to increase the sensitivity, but not specificity, of the searches. Further articles were picked up through the scrutiny of the selected articles’ references.

Selection. These searches identified 2466 potentially relevant abstracts to which exclusion criteria were applied.
Exclusion criteria were investigation of hospital and not surgeon volume and outcome and, where the analysis included ruptured aneurysms, thoracic or thoracoabdominal aortic aneurysm repairs. Abstracts were assessed to narrow the search results to relevant articles.

**Validity assessment.** Having met the inclusion criteria, articles were assessed for quality and their findings. Mortality was the primary outcome measure. The degree and method of case-mix adjustment was recorded, and risk-adjusted data were used where available. Otherwise, articles were considered to have adjusted for the severity of illness if they reported separately for ruptured aneurysms or stated that they did not present ruptured aneurysms.9

**Data abstraction.** Data abstraction was performed independently by either P.I.E.H. or E.L.Y. Discrepancies were resolved by discussion. For each article, mortality rates, odds ratios, and 95% confidence intervals were taken from the article for higher- and lower-volume surgeons. The threshold value for the number of AAA repairs performed per annum between higher- and lower-volume surgeons was recorded. Datasets were dichotomized to higher- and lower-volume categories where articles published a series of volume groupings (eg, higher-, moderate-, and lower-volume surgeons).

**Study characteristics.** Analysis of the data was by meta-analysis and systematic review conforming to the QUOROM statement10 (Fig 1). Fourteen articles were included as a systematic review, and their conclusions were used to test our hypothesis, but the data in eight of these articles were not presented in a format consistent with the meta-analysis. Common problems encountered were as follows:

1. Mortality rates, but not patient numbers, for higher- and lower-volume surgeons were published.
2. No threshold value was stated between higher- and lower-volume surgeons.
3. There was nonstandard presentation of results—eg, the relative risk of mortality with doubling of surgeon volume.

**Quantitative data synthesis.** Weighted averages were calculated for meta-analysis by the Mantel-Haenszel method.11 A forest plot12 was used to illustrate the information from the individual studies used in the meta-analysis, in terms of odds ratios and 95% confidence intervals. The plot demonstrated the variation between studies and a pooled point estimate of the overall result, with 95% confidence intervals, that represented the volume-outcome relationship. A P value was generated by using the z statistic. The weighting for each study in the meta-analysis was applied to the threshold value at which the data from each study were dichotomized. This allowed the calculation of a weighted threshold value for the pooled effect estimate.

The presence, or absence, of statistical heterogeneity was assessed by using the Q statistic and tested against a χ² contingency table. The degrees of freedom were one fewer than the number of studies in the meta-analysis, and the level of heterogeneity was assessed at P = .1, which is standard due to the relative insensitivity of the method. Because the sample size was small in the meta-analysis, heterogeneity was quantified by using the I² test13; a value of less than 50% was taken as insignificant heterogeneity.

A sensitivity analysis was performed by exclusion of the largest trial from the meta-analysis and recalculation of the pooled effect estimate and heterogeneity scores. The results were checked against a DerSimonian and Laird random-effects meta-analysis by way of a further test of heterogeneity and as a second sensitivity analysis.

**RESULTS**

Twenty-five potentially relevant articles were identified by the searches. Eleven articles were excluded: four combined elective and ruptured AAA repairs, or ruptured aneurysms only14-17; six did not in fact investigate the surgeon volume–outcome relationship18-23; and one study summarized prior research, presenting no new data relevant to this review.24 After exclusions, 14 articles provided information on the annual surgeon volume of elective AAA repair and mortality.25-38 A total of 115,273 elective AAA repairs (range, 207-39,794 cases) were considered. The mean mortality rate was 5.56%, and the observed mortality rate decreased as a surgeon’s annual operative volume increased (Fig 2; Table).
was no significant heterogeneity (Q-stat, 6.22; 4 df; I² = 0.57 (95% confidence interval, 0.53-0.62). There were no physiologic data, and are prone to coding errors. How-ever, their use has been shown to be a valid surrogate for the cohort mean mortality rate of 5.56%. The trend line is a line of best fit, and no statistical inferences have been made from this.

From the meta-analysis of 51,453 cases (Fig 3), the weighted odds ratio for mortality rates of high- as compared with low-volume surgeons was 0.56 (95% confidence interval, 0.54-0.57; P < .00001) at a weighted mean threshold between higher- and lower-volume surgeons of 13 AAAs per annum. All six studies included in the meta-analysis found a statistically significant reduction in mortality with increased surgeon operating volumes.25-28,35,37 No significant statistical heterogeneity was demonstrated for the meta-analysis (Q-stat, 6.56; 5 df; I² = 23.7%).

A sensitivity analysis was performed through exclusion of the largest trial in the meta-analysis,35 and gave an odds ratio of 0.57 (95% confidence interval, 0.53-0.62). There was no significant heterogeneity (Q-stat, 6.22; 4 df; I² = 35.7%). The random-effects meta-analysis (of the complete dataset) for dichotomous outcomes gave an odds ratio of 0.58.

A further eight studies were presented in the systematic review (Table), of which six presented strong evidence for improved results with high-volume surgeons29,33,38 and two found no statistically significant association between volume and outcome.34,36

**DISCUSSION**

This systematic review and meta-analysis provided significant evidence that elective AAA repairs performed by higher-volume surgeons confer a survival advantage. Discussion of the problems encountered in this type of work, including publication bias, and the use of mortality as the principal end point in AAA surgery has been published previously.9,39,40

The use of administrative datasets for outcomes analyses has been debated because they are retrospective, contain no physiologic data, and are prone to coding errors. However, their use has been shown to be a valid surrogate for prospectively collected data.19,41

Considerable research has demonstrated a reduced mortality rate for elective open AAA surgery at high-volume hospitals,1,5,8,42 an independent predictor of increased survival and one that is additive to that conferred by surgery by a high-volume surgeon.26 That is to say, the best results were achieved by high-volume surgeons operating at high-volume institutions.5,35

It has been variously suggested that hospital factors are of more importance than surgeon factors when outcomes are examined,27,34,43,44 and vice versa,5,35,37 but that both have independent roles. The mechanisms behind the relationship between hospital volume and outcome have been discussed fully elsewhere,8,19,20 but in summary, outcomes were reliant on hospital infrastructure, the provision of expert intensive care units, the effective working of large multidisciplinary teams, and the maximum potential throughput of the hospital. These factors meant that a low-volume surgeon operating in a high-volume hospital could adopt a mortality rate close to that of the high-volume hospital for some procedures, although this has not been proven for AAA repair.44 As a confounding factor, it is likely that large effective units might also have had a greater ability to attract high-volume specialist vascular surgeons,44 thus magnifying the beneficial effects of higher hospital annual volumes. Even within high-volume institutions, high-volume surgeons had better results than low-volume surgeons.35

With there being a lack of consensus in the results from different studies, it was not possible to definitively state, or quantify, whether hospital or surgeon volume played the greater part in the volume-outcome relationships; the relative importance of each was uncertain. It was certain that both had a significant role in reducing mortality from elective AAA repair and that improved chances of survival could be found by selecting a high-volume surgeon in a high-volume hospital.35 It has been suggested that, for elective AAA repair, surgeon volume might account for up to 57% of the benefit seen at high-volume hospitals.35

**Surgeon specialty.** Three studies investigated the relationship between surgeon specialty and outcome,26,32,33 and all found that operations performed by vascular subspecialists were associated with significant reductions in mortality when compared with operations by general surgeons. This advantage was, again, in addition to the benefits gained through the provision of higher-volume services at the level of both the hospital and the surgeon.26 Therefore, high-volume specialist vascular surgeons operating in high-volume hospitals achieved the lowest mortality for elective AAA repairs.

It is worth noting that technically, high-volume surgery alone does not confer specialist status, a term that implies subspecialty or fellowship training, certification, and membership of a specialist society. Dimick et al26 defined vascular surgeons as having performed more than 75% vascular procedures in the 1997 Nationwide Inpatient Sample, having determined the effect of specialization at three thresholds (>25%, >50%, and >75% vascular procedures). The relevance of these statements is that without the abovementioned attributes, even if surgeons adhered to the minimum volume criteria enumerated here, they could not
be considered specialist vascular surgeons, who are associated with better outcomes.

Assessment of volume. The threshold values between high- and low-volume surgeons varied between studies and in some cases were not specified. One study,25 examining outcomes over a 6-year period, defined “very low volume” as a single case in 6 years and “very high volume” as more than 100 cases in 6 years, with 3 intermediate-volume categories. In their comparison, a threshold of 10 cases had an odds ratio of 3.26 (range, 1.32–8.03; \( P = .01 \)) in favor of the higher-volume surgeons. The threshold of 10 cases per annum was used in 3 other studies,26,27,30 all of which found in favor of higher-volume surgeons. Other studies used different threshold values of 4 cases,28,37 24 cases,29 26 cases,38 and 17.5 cases.35 One study simply defined the threshold as high- and low-volume,36 and two studies modeled volume as a continuous variable and calculated a cutoff value.27,31 All found significant survival benefits in the higher-volume groups. Where multiple volume groupings were published, in all but one case,25 it was found that every increase in case volume led to a further improvement in outcome.

### Table. Systematic review of the relationship between annual surgeon operative volume and mortality

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Elective cases</th>
<th>Data source</th>
<th>Open or EVAR</th>
<th>Thresholds</th>
<th>Mortality rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dardik25</td>
<td>1999</td>
<td>2335</td>
<td>Maryland Health Services</td>
<td>Open</td>
<td>1–9</td>
<td>9.9</td>
</tr>
<tr>
<td>Dimick26</td>
<td>2003</td>
<td>3912</td>
<td>Nationwide Inpatient Sample 1997</td>
<td>Open</td>
<td>&lt;10</td>
<td>2.5</td>
</tr>
<tr>
<td>Ducck29</td>
<td>2004</td>
<td>13,701</td>
<td>Ontario Health Insurance Plan 1993-1999</td>
<td>Open and EVAR</td>
<td>&lt;24</td>
<td>Not given</td>
</tr>
<tr>
<td>Hannan27</td>
<td>1992</td>
<td>3570</td>
<td>New York State Hospital Discharge Data 1986</td>
<td>Open</td>
<td>0–2</td>
<td>9.4</td>
</tr>
<tr>
<td>Kantonen31</td>
<td>1997</td>
<td>929</td>
<td>Finnvasc Registry</td>
<td>Open</td>
<td>Not used</td>
<td>Not given</td>
</tr>
<tr>
<td>Kelly34</td>
<td>1986</td>
<td>999</td>
<td>Hospital Cost and Utilization Project 1977 and 1983</td>
<td>Open</td>
<td>Not given</td>
<td>Not given</td>
</tr>
<tr>
<td>Pilcher28</td>
<td>1980</td>
<td>207</td>
<td>Vermont hospitals by request 1970-1977</td>
<td>Open</td>
<td>0.1–1.7</td>
<td>17</td>
</tr>
<tr>
<td>Tu33</td>
<td>2001</td>
<td>5878</td>
<td>Ontario Health Insurance Plan, Canadian Institute for Health Information, Ontario Registered Persons database 1998-1999</td>
<td>Open and EVAR</td>
<td>&lt;5</td>
<td>7.1</td>
</tr>
<tr>
<td>Veith38</td>
<td>1991</td>
<td>3570</td>
<td>Same sample as Hannan27</td>
<td>Open</td>
<td>1–5</td>
<td>4</td>
</tr>
<tr>
<td>AbuRahma36</td>
<td>1991</td>
<td>332</td>
<td>Southwest Virginia</td>
<td>Open</td>
<td>Low</td>
<td>4</td>
</tr>
</tbody>
</table>

*Charlson, Charlson comorbidity score; Romano-Charlson; Romano modification of Charlson comorbidity score; EVAR, endovascular aneurysm repair; AAA, abdominal aortic aneurysm.*
Three studies\textsuperscript{31,32,34} did not publish the specific volume criteria they used to test the data. Two of these found highly significant results that outcome depended on a surgeon’s experience in the last 4 years\textsuperscript{31} and that the relative risk of mortality was 0.898 with each doubling of surgeon volume when volume was modeled as a continuous variable.\textsuperscript{32} Only Kelly and Hellinger\textsuperscript{34} and AbuRahma et al\textsuperscript{36} found nonsignificant results, but even then, there were trends toward improved outcomes with increasing volume.

**Why do high-volume surgeons have better results?** It has been suggested that a “practice makes perfect” phenomenon occurs in surgery,\textsuperscript{19} such that the increased experience of high-volume surgeons accounts for their reduced mortality rate. However, as an aside, a longitudinal study showed that AAA surgeons rarely change from lower to higher volume, with new surgeons achieving higher volume quickly, without a protracted period as lower-volume surgeons.\textsuperscript{27}

Alternatively, there may be an effect of selective referral; ie, doctors preferentially refer to hospitals or surgeons with pre-existing low mortality rates, thus increasing their vol-

---

**Table. Continued**

<table>
<thead>
<tr>
<th>Mean mortality rate (%)</th>
<th>Higher-volume/ lower mortality</th>
<th>Significant</th>
<th>Risk-adjustment and modeling</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Yes</td>
<td>Yes</td>
<td>Patient complexity scores; multivariate analysis</td>
<td>Compared $&lt;10$ vs $10-40$; in meta-analysis</td>
</tr>
<tr>
<td>4.2</td>
<td>Yes</td>
<td>Yes</td>
<td>Romano-Charlson univariate and multivariate analysis</td>
<td>In meta-analysis</td>
</tr>
<tr>
<td>4.5</td>
<td>Yes</td>
<td>Yes</td>
<td>Charlson multivariate model</td>
<td>Volume modeled as a continuous distribution</td>
</tr>
<tr>
<td>7.7</td>
<td>Yes</td>
<td>Yes</td>
<td>Age, sex, severity of illness, mode of admission, comorbidities</td>
<td>In meta-analysis</td>
</tr>
<tr>
<td>5.5</td>
<td>Yes</td>
<td>Yes</td>
<td>Adjusted for race, sex, and age but not comorbidities; no modeling</td>
<td>Number of cases per group not given so not in meta-analysis</td>
</tr>
<tr>
<td>5.1</td>
<td>Yes</td>
<td>Yes</td>
<td>Multiple logistic regression, but no risk adjustment</td>
<td>Mortality rates depended on surgeon’s AAA experience in previous 4 y</td>
</tr>
<tr>
<td>10.2</td>
<td>Marginal</td>
<td>Nonsignificant trend toward improved outcome with increasing volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.7</td>
<td>Yes</td>
<td>Yes</td>
<td>Multiple logistic regression, but no risk adjustment</td>
<td>In meta-analysis</td>
</tr>
<tr>
<td>11.1</td>
<td>Yes</td>
<td>Yes</td>
<td>Charlson univariate and multivariate modeling</td>
<td>Number of cases per group not given so not in meta-analysis</td>
</tr>
<tr>
<td>4.1</td>
<td>Yes</td>
<td>Yes</td>
<td>Charlson demographic nested modeling adjusted multiple regressions</td>
<td>In meta-analysis; physician volume 57% of hospital effect</td>
</tr>
<tr>
<td>4.88</td>
<td>Yes</td>
<td>Yes</td>
<td>Charlson demographic nested modeling adjusted multiple regressions</td>
<td>Same sample as Hannan\textsuperscript{27}</td>
</tr>
<tr>
<td>Not given</td>
<td>Yes</td>
<td>Yes</td>
<td>Age, sex, severity of illness, mode of admission, comorbidities</td>
<td>Same sample as Hannan\textsuperscript{27}</td>
</tr>
<tr>
<td>3.6</td>
<td>Yes</td>
<td>No</td>
<td>Multivariate analysis</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Yes</td>
<td>Yes</td>
<td>Multivariate analysis; age, sex, severity of illness, mode of admission, comorbidities</td>
<td>Physician volume more important than hospital volume; in meta-analysis</td>
</tr>
</tbody>
</table>

---
Both selective referral and “practice makes perfect” may be applicable to AAA repair.

Countering the “practice makes perfect” argument is the “walking on water” phenomenon, anecdotal among surgeons. This suggests that increasing annual case volumes up to approximately 35 open AAA repairs per annum is associated with continued improvements in outcome (decreased mortality rates). At annual volumes above this, improvements are smaller, and professional complacency may play a part such that outcomes are at best static. The existence of this phenomenon might be supported by the findings of Dardik et al, who found that despite improvements in outcome at lower volumes, mortality rates were increased. Dueck et al also found that the benefits of increasing volume were minimal above 30 cases per annum.

Alternatively, those surgeons achieving a very high volume of surgery may be doing so on the basis of excellent pre-existing results. This in turn may lead to the referral of more complex cases to these surgeons. A more difficult case-mix may explain the plateau in the improvement in mortality observed for these surgeons.

Combining these theories, it would seem reasonable that an absolute minimum case volume of 13 cases per annum, as established through the meta-analysis, should be achieved by surgeons performing elective AAA repairs to avoid preventable deaths from elective surgery. This value was consistent with the finding of a Michaels et al study for Health Technology Assessment in the United Kingdom.

The effect of endovascular aneurysm repair. The studies in this report investigated the relationship between open aneurysm repair and outcome. However, because of the difficulties of coding new procedures, until 2007 there was no code for endovascular aneurysm repair. Most of the studies presented here, because they presented retrospective data, were concerned only with open aneurysm repair, because the availability of endovascular aneurysm repair was limited or nonexistent at the time that the operations were performed. There was, however, less certainty in the coding of the more recent studies. It is possible that high-volume surgeons have a greater availability of endovascular aneurysm repair, thereby improving their in-hospital mortality rates. Future studies must be alert to this possible confounding effect, although we believe that it did not skew these results.

The importance of case-mix. To minimize bias, account must be taken of any factors (beyond volume) likely to affect patient outcome, most notably case-mix. Although some older studies used no case-mix adjustment, more recent articles commonly used the Romano modification of the Charlson comorbidity score. Studies varied in the way in which they dealt with the problems of case-mix. One study presented medical complexity scores and mean age for each volume grouping but did not then risk-adjust the mortality data. Other studies presented standardized mortality ratios, and still others performed multivariate analyses. All of these are valid techniques, although a consensus statement might be useful for future studies so that different studies can be accurately compared.
Clinicians should be wary about publishing and interpreting data with no adjustment for case-mix, because this could have a significant effect on mortality figures.

**Ethical considerations.** Although this large body of data exists and conclusions have been drawn as to the volume of surgery that surgeons should perform to improve outcome, it might be that further assessment of outcomes at the physician level are premature. While issues surrounding coding remain and case-mix adjustment is insufficient, to make accusations regarding an individual surgeon’s outcomes would be open to litigation. Certainly, ranking surgeons on the basis of their results and naming outliers is a risky business where these imperfections remain.

**CONCLUSION**

As surgeons performed higher annual volumes of elective AAA repairs, significantly lower mortality rates were demonstrated. Surgeons wishing to perform AAA repairs should achieve a minimum case volume of 13 elective repairs per annum and attain vascular subspecialization.

**REFERENCES**


Submitted Apr 30, 2007; accepted Jun 13, 2007.

CME Credit Now Available to JVS Readers

Readers can now obtain CME credits by reading selected articles and correctly answering multiple choice questions on the Journal website (www.jvascsurg.org). Four articles are identified in the Table of Contents of each issue and 2 questions for each are posted on the website. After correctly answering the 8 questions, readers will be awarded 2 hours of Category I CME credit.