A systematic review and meta-analysis of the volume outcome relationship in the surgical treatment of breast cancer. Are breast cancer patients better off with a high volume provider?


To cite this version:

HAL Id: hal-00612390
https://hal.archives-ouvertes.fr/hal-00612390
Submitted on 29 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Title: A systematic review and meta-analysis of the volume outcome relationship in the surgical treatment of breast cancer. Are breast cancer patients better of with a high volume provider?


PII: S0748-7983(10)00200-3
DOI: 10.1016/j.ejso.2010.06.024
Reference: YEJSO 3012

To appear in: European Journal of Surgical Oncology

Received Date: 7 June 2010
Accepted Date: 9 June 2010


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.
A systematic review and meta-analysis of the volume outcome relationship in the surgical treatment of breast cancer. Are breast cancer patients better off with a high volume provider?

GA Gooiker¹, W van Gijn¹, PN Post³, CJH van de Velde¹, RAEM Tollenaar¹ and MWJM Wouters¹²

¹Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands
²Department of Surgical Oncology, Netherlands Cancer Institute – Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands
³The Dutch Institute for Healthcare Improvement, CBO, Utrecht, the Netherlands

Study performed by the ‘Quality of cancer care’ taskforce of the Signalling Committee Cancer of the Dutch Cancer Society

Corresponding author:
prof.dr. R.A.E.M. Tollenaar
Leiden University Medical Center, dept of Surgery
PO Box 9600
2300 RC Leiden
Phone: +31 71 5263610
Fax: +31 71 5266750
E-mail: r.a.e.m.tollenaar@lumc.nl
Abstract

Aims
To conduct a systematic review of the literature on the volume-outcome relationship for the surgical treatment of breast cancer with consideration of the methodological quality of the available evidence and to perform a meta-analysis on the studies of considered good quality.

Methods
A systematic search was done to identify all articles examining the effects of hospital or surgeon volume on clinical outcome of the surgical treatment of breast cancer. Reviews, opinion articles and surveys were excluded. All articles were critically appraised on methodological quality and risk of bias. After strict inclusion, meta-analysis assuming a random effects model was done to estimate the effect of higher hospital or surgeon volume on patient outcome.

Results
We found 12 studies of good methodological quality which could be included for meta-analysis. The results showed a significant association between high volume providers and an improved survival. The association is the most robust for surgeon volume (HR 0.80 (0.71-0.90) and RR 0.85 (0.80-0.90). In addition there is an effect of hospital volume on the in-hospital mortality, although the mortality was very low (0.1-0.2%). Results of meta-analysis were heterogeneous. Sensitivity analysis showed a larger effect size for studies also adjusting for comorbidity for both studies on hospital and surgeon volume. The data were not suggestive for publication bias.

Conclusions
The results show that survival after breast cancer surgery is significantly associated with high volume providers.
Introduction

Should breast cancer treatment be centralized in high volume breast units or not? Since Sainsbury and Gillis showed better survival after breast cancer surgery with specialized high volume surgeons, this question dominates the debate on the quality of breast cancer care in many western countries. 1,2

Minimal volume standards are supported by many studies, that have shown differences in mortality and survival between high and low volume providers, especially with high risk, low volume procedures. 3-8 These studies suggest that the superior outcomes are the result of more experience and better hospital based services with high volume providers. Centralization of cancer care in high volume centres is therefore expected to improve the outcome for many patients. This could not just apply to high risk, low volume procedures, but also to more common, low risk procedures, such as the surgical treatment of breast cancer.

Though the evidence on the volume outcome relationship of the treatment of cancer seems convincing, there is also solid criticism on the methodological quality of these studies. Most studies are based on administrative data instead of clinical data and do not allow for differences in case mix between low and high volume providers. 9,10 Moreover, a minimal volume standard cannot be identified. And though the association between procedural volume and clinical outcomes is subject of research of many studies over more than a decade, the strength of the evidence and the clinical importance remains unclear.

In the Netherlands, the Dutch Cancer Society formed a study group to investigate whether there is a potential benefit of centralization of the treatment of breast cancer in the Dutch hospitals. As a part of the report the study group decided to review all available literature on the association of hospital and surgeon volume with postoperative mortality and survival for the surgical treatment of breast cancer with consideration of the methodological quality of the available evidence and to perform a meta-analysis on the studies of considered good quality.
Method

Systematic Search Strategy

A specialised librarian and one of the investigators (GG) performed a systematic search to identify all relevant studies describing the association between hospital or surgeon volume and clinical outcomes. We conducted a search in the electronic databases Medline, Embase and the Cochrane Library, with a combination of MESH terms and text words. Table 1 shows the search terms used for Medline and Embase. Because the volume of procedures is not well indexed in the electronic databases, we formulated the search terms as sensitive as possible to ensure no publications were missed. In addition, reference lists of relevant articles were hand-searched to identify additional articles, which could have been missed in the search. We also used the “related articles” function in Pubmed. The last search was on February 1 2010.

Table 1  Search terms used in the search in the databases Medline and Embase

<table>
<thead>
<tr>
<th>Medline (Pubmed)</th>
</tr>
</thead>
</table>
| ("hospital volume" OR "surgeon volume" OR "surgical volume” OR workload OR caseload OR "procedure volume” OR "procedural volume”)
AND
AND
("Surgical Procedures, Operative"[Mesh] OR surgery OR surgical OR surgeon)                                                                                                                                 |

<table>
<thead>
<tr>
<th>Embase</th>
</tr>
</thead>
<tbody>
<tr>
<td>hospital volume.mp. OR surgeon volume.mp. OR workload.mp. OR Workload/ OR caseload.mp. OR procedure volume.mp.</td>
</tr>
</tbody>
</table>
AND                                                                                                                                                                                                           |
| surgical complications.mp. OR postoperative complications.mp. OR exp Postoperative Complication/ OR Mortality/ OR mortality.mp. OR exp Survival Rate/ OR exp Survival/ OR survival.mp. OR exp Cancer Recurrence/ OR neoplasm recurrence.mp. OR treatment |
outcome.mp. OR exp Treatment Outcome/ OR surgical mortality.mp. OR exp Surgical Mortality/ OR exp recurrent disease/ or exp tumor recurrence/ OR exp cancer survival/ OR disease free survival.mp. or exp Disease Free Survival/ AND surgery/ OR exp cancer surgery/ OR surgeon.mp. OR exp Surgeon/ OR surgical procedures.mp. or exp Surgical Technique/

This search strategy retrieved 1319 articles in Pubmed and 616 in Embase. After combining these results and removing duplicates, 1400 studies rested for primary selection of which 63 studies had the (surgical) treatment of breast cancer as study subject.

Study selection
Two reviewers (GG and WvG) independently screened titles and abstracts of all 63 retrieved articles. Studies were selected using the following inclusion criteria.

- The subject of the study is the surgical treatment of breast cancer.
- Hospital or surgeon volume is an independent variable.
- The outcome parameter is postoperative mortality or survival.
- The study does not describe a single hospital or surgeon.
- The study uses primary data (e.g. editorials, systematic reviews, are excluded)

After the first selection 24 articles remained. These were obtained in full text and were further selected by the same two reviewers using the following exclusion criteria:

- Multiple publications are based on the same database. In this case the study with the lowest quality is excluded. When studies have similar quality, the study with the shortest period or the least recent period is excluded.
- The study period is older than 20 years. Studies with a study period starting before 1988 are excluded, since surgical and clinical care has considerably changed.
- There is no multivariate analysis done with adjustment for at least age and gender.
Volume is not defined as a distinct number or cut-off value (e.g. studies that defined volume as 'specialization' are excluded).

No language constraints were placed. We did not use abstracts or unpublished data. In the appendix is an overview of the excluded studies.

Any discrepancies regarding inclusion or exclusion of a study were solved by discussion or by a third author (PP, MW).

Assessment of study quality & Data-extraction

Two authors (GG, WvG) critically appraised each study on methodological quality and risk of bias. Data of the included studies were gathered in a data-extraction form, which was based on the STROBE criteria, to record study characteristics, methodological quality and the study results (www.strobe-statement.org). The form was recorded in Access (Microsoft Corporation, Redmond WA, USA).

From each study characteristics were collected regarding the tumour type, the type of surgery, the study period, the unit of analysis (hospital or surgeon), and the number of analysed patients, hospitals and surgeons and the country.

In addition we appraised the quality of the data source (administrative or clinical data), the study design (prospective or retrospective) and the degree of risk adjustment. We noted the casemix factors for which statistical adjustment was done. Casemix factors were categorized as age and gender; comorbidities, tumour characteristics; (neo-)adjuvant treatment or urgency of the operation. We also checked the inclusion criteria of the study to verify if there was a probability of selection bias.

Cut-off values for high and low volume used by the studies were noted per volume group, along with how these cut-off points were determined. It was not possible to categorize the volume groups of all studies in equal volume categories, since there was a high variation in cut-off values.

The study results were recorded separately for each unit (surgeon or hospital) and outcome parameter (30-day mortality, in-hospital mortality, postoperative mortality and 2 year or 5 year survival), the crude outcomes for each volume group were noted (when reported).

Subsequently, we noted for each volume group and outcome parameter the estimated effect
size after adjustment, expressed as odds ratio’s (OR), hazard ratio’s (HR) or relative risks (RR) with confidence intervals (CI) and measures of significance.

Synthesis of the data
The reference category varied between studies. Therefore, we had to convert the effect sizes so that the lowest group was the reference in all studies. As a result, the OR of mortality or the HR of survival reflected the odds of mortality in the highest volume group compared to the odds of mortality in the lowest volume group. For each study, we only included one comparison. We chose to compare the results for the highest volume groups with the results of the lowest volume groups.
To determine a pooled estimated effect, we used the random effect model for the meta-analysis. The random effects model accounts for expected heterogeneity, which is more appropriate with pooling of observational studies. Data of the included studies were represented in a forest plot for each unit and outcome parameter, with the pooled effect size represented at the bottom.
Heterogeneity was quantified by the $I^2$ test. An $I^2 > 50$ was considered as notable heterogeneity.\textsuperscript{11} We conducted a sensitivity analysis of meta-analyses including more than three studies to further explore heterogeneity and assess the impact of subgroups. A subgroup analysis was done for data source (administrative vs clinical), case mix adjustment (adjustment for comorbidity or severity) and study country.
Publication bias was assessed for meta-analyses with more than three studies with an Egger’s regression intercept and shown in a funnelplot.\textsuperscript{12}
The meta-analysis was conducted with Comprehensive Meta Analysis, professional version 2.2 (©2006, Biostat inc. Englewood, USA).
Results

Study characteristics

Our initial search retrieved 63 potentially relevant studies concerning the volume outcome relationship of the surgical treatment of breast cancer (figure 1). After the first screening we excluded 39 studies: 15 studies examining other outcome parameters; 10 studies in which hospital or surgeon volume is not the independent variable (but specialization, teaching status, or urban vs. rural hospital, as independent variable); and 14 studies contained no primary data (editorials, letters). After the first selection, 24 studies remained and were fully obtained for more detailed evaluation. A total of 12 studies were further excluded, because the study period was to old (5), because identical databases were used (4), because no risk adjustment was done (2) and volume was not an independent variable in multivariate analysis (1).

Table 2 shows the characteristics of the remaining 12 studies. We included five studies with hospital volume as the independent factor, three studies with surgeon volume as the independent factor, and four studies with both hospital and surgeon volume as independent factors. Five studies were from the United States, one from Australia, two from Canada, three from the United Kingdom and one from Taiwan.

There was considerable variation between the included studies in the used cut-off values of the volume groups. For hospital volume, cut-off values of the highest volume strata varied between 40-195 procedures each year. The cut-off values of the lowest hospital volume strata varied between 10-86 procedures each year. As a result, it occurred that high volume in one study could be considered as low volume in another study and vice versa.

Methodological quality of the studies

All included studies had an observational design and 7/12 studies were based on clinical data, collected in cancer registries. The other five studies were based on administrative data. The results of all included studies were adjusted for age and 7/12 were adjusted for comorbidity as well. The majority was adjusted for tumour stage and adjuvant treatment (11/12).

Study results
Two studies were included concerning in-hospital mortality after breast cancer surgery. The mortality was low (0.1-0.2%), but was significantly lower in high volume hospitals.

Eight studies on hospital volume and survival were included of which seven studies showed a beneficial effect of a higher hospital volume on the survival. All seven included studies concerning surgeon volume and survival, showed a better survival with high volume surgeons.

**Pooled estimated effect size**

Figure 2a shows the forest plot of the included studies regarding hospital volume and in-hospital mortality. The pooled estimated effect size was significant in favour of high volume providers in the analysis of postoperative mortality (OR 0.40 (CI 0.22-0.74). The results showed little heterogeneity ($I^2=12$)

Figure 2b shows the forest plots of the included studies on hospital volume and survival. The meta-analysis showed mixed results. The analysis of studies with Hazard Ratio’s as effect size showed a significant pooled estimated effect size in favour of high volume hospitals (HR 0.84 (CI 0.76-0.93)). Though this result was heterogeneous ($I^2=76$). The pooled estimated effect size of two studies with relative risks as effect size measure, was not significant (RR 0.90 (0.66-1.22)) and large heterogeneity was observed ($I^2=96$).

Figure 2c shows the forest plots of the included studies regarding the effect of surgeon volume on survival. The meta-analysis showed a significant effect in favour of high volume. The pooled estimated effect size of studies with hazard ratio’s was HR 0.82 (0.72-0.93) and with risk ratio’s was RR 0.85 (0.80-0.91). The first analysis was statistically heterogenous ($I^2=59$) and the second analysis was statistically homogeneous ($I^2=0$).

**Sensitivity analysis**

Table 3 shows the results of the sensitivity analysis of the 6 included studies on hospital volume and survival with hazard ratio’s as effect size. Subgroup analysis of these studies showed a larger effect size for studies adjusting for differences in comorbidity(P=0.003). The pooled estimated effect of studies which were adjusted for comorbidity was HR 0.77 (CI 0.69-0.87, $I^2=46$)
Table 3 shows also the results of the sensitivity analysis of the 4 included studies on surgeon volume and survival with hazard ratio’s as effect size. Subgroup analysis of these studies showed a larger effect size for studies adjusting for differences in comorbidity (P=0.003).

Risk of publication bias

Figure 3 shows the qualitative analysis of publication bias of all studies regarding hospital volume and survival using HR’s. Quantitative analysis with the Egger’s regression intercept showed an intercept of -3.7 with a two-sided P value of 0.03.

Figure 3 shows also the qualitative analysis of publication bias of all studies regarding surgeon volume and survival using HR’s. Quantitative analysis with the Egger's regression intercept showed an intercept of -1.6 with a two-sided P value of 0.41.
Discussion

This is the first systematic review and meta-analysis examining the effect of surgeon and hospital volume on the outcomes of the surgical treatment of breast cancer, providing an overview of the current evidence. The present study shows that there is a strong significant association between high volume providers and improved survival in breast cancer treatment. The association is most robust for surgeon volume [HR 0.80 (0.71-0.90) and RR 0.85 (0.80-0.91)]. In addition, there is an effect of hospital volume on in-hospital mortality, though one has to consider that postoperative mortality in breast cancer surgery is very low: 0.1-0.2% in the included studies.

The fact that no randomized studies on the volume outcome relationship for breast cancer treatment are available, hampers the interpretation of evidence provided in literature. All the studies included in our meta-analysis had an observational design and it was challenging to deal with the heterogeneity of the eligible studies. To reduce heterogeneity strict inclusion criteria were applied and only studies, which were at least risk-adjusted for age and gender, were included. All included studies using survival as an endpoint were adjusted for severity (tumour grade, stage and adjuvant treatment) as well. In addition, all studies with study periods older than 20 years were excluded, since breast cancer treatment has considerably changed in the last two decades. Studies defining volume as ‘specialization’ were also excluded since specialization was usually vaguely defined. Nevertheless, considerable heterogeneity remained.

The methodological quality of studies may have influenced the reported results. Therefore we performed a sensitivity analysis to examine the influence of study characteristics on heterogeneity. Sensitivity analysis showed a larger effect size with studies adjusted for comorbidity, with studies both on hospital volume and surgeon volume. The results were not sensitive to differences in any of the other variables.

Another issue was the considerable differences in cut-off values for volume categories between the included studies. High volume in one study was considered low volume in another and vice versa. This can have influenced our results. Therefore it was not feasible to identify a minimal required volume from the included studies.
Acknowledging these limitations, the results of our meta-analysis provide evidence that quality of breast cancer surgery is superior with high volume providers. An explanatory theory for the volume-outcome relationship was already introduced by Luft et al in 1979.\textsuperscript{25,26} He described two theories to explain the association: the practice makes perfect theory and the selective referral theory. Both theories can be relevant in our findings.

The practice makes perfect theory assumes that the more experienced a surgeon or a hospital is, the better the clinical outcomes. This seems a plausible explanation. A more experienced surgeon can for example have a higher percentage of radical excisions of the primary tumour, which improves recurrence-free survival. However, especially breast cancer treatment is a multidisciplinary process. Improved outcome can be also be obtained by more accurate imaging, adequate tumour biopsies, better decision making, (neo)adjuvant treatments and early detection of local recurrence and/or metastases. In short, the outcomes are the result of an adequate infrastructure in which adequate decisions are made and complex technical procedures are performed, all reflecting the high quality effort of an entire multidisciplinary team. This is further illustrated by Skinner et al, who showed that both hospital volume and surgeon volume contributed to a better survival.\textsuperscript{23}

The selective referral theory assumes that hospitals and surgeons with superior results attract more patients. This could also explain our findings. With increasing transparency of the quality of care provided by different hospitals, physicians can refer their patients to hospitals and surgeons with demonstrable good results. Patients are better informed and increasingly able to choose where they might receive optimal treatment.

However, despite the numerous publications regarding the volume-outcome relationship, the mechanisms behind better clinical outcomes are not revealed yet. Several studies examined the association of provider volume with the use of different treatments or different interventions (breast conservative surgery, vs. breast ablative surgery).\textsuperscript{27-31} These studies showed a higher positive biopsy rate, more breast conservative surgery and a higher utilization of adjuvant therapy with high volume providers. This suggests that more accurate diagnosis, better resections and adequate use of radio-, chemo- and hormonal therapy may be the underlying reasons for improved outcomes in high volume centres. Unfortunately, studies investigating other outcome measures such as tumour recurrence, positive surgical margins or postoperative complications are lacking.
The results of the present meta-analysis provide an overview of the current evidence for the volume outcome relationship in breast cancer surgery and shows that there is a strong significant relationship between high volume providers and improved survival. This can be valuable information for professionals as well as policymakers. The findings support the assumption that centralization of breast cancer has the potential of improving the quality of care. In Europe there are already several developments on this part. Specialist breast units are a key component of the European guidelines on the organisation of breast cancer care. The European Society of Breast Cancer Specialists, EUSOMA, already introduced a concept of specialised breast units in 2000 and recently established a minimal volume standard of 150 new cases per year for hospitals and 50 procedures per year for surgeons. The French National Institute of Cancer requires a minimum of 30 procedures per year for surgeons (www.e-cancer.fr/en/les-soins/recommandations). And in the United Kingdom, concentration of breast cancer care in hands of specialist doctors was already recommended in the Calman-Hine report in 1995. In Ireland, recently breast cancer treatment is centralized in 13 specialized breast units (www.europadonnaireland.ie).

However, the question remains if minimal volume standards for the surgical treatment of breast cancer actually improve the quality of care. Volume is a proxy for better quality, but remains a poor predictor of the quality of care in individual institutions. With the introduction of minimal volume standards there is a risk of selecting hospitals with inferior outcomes, lacking the facilities and expertise (future) referral centres need. Therefore a minimal volume standard should not be the only criterion in the accreditation of breast cancer units.

In addition, several other disadvantages of minimal volume standards are mentioned in the debate. The increasing elderly population with often complex comorbid disease and decreased mobility, demands for care in the region. Care is preferably provided in a hospital where the patient has been or is treated for his/her comorbid diseases. Minimal volume standards do not take geographical spread into account and can cause logistic problems. This is undesirable since breast cancer is one of the most common diseases in the developed world. Nevertheless, we can assume that to invest in infrastructure for optimal care and to maintain a certain level of expertise within the whole medical team, a certain amount of patients is needed.

Minimal volume standards alone thus seem an inadequate basis for centralization. Additional quality criteria should be formulated to direct centralization initiatives. For this purpose more
research is needed to identify essential structural or organisational characteristics and high leverage care processes that lead to the better outcomes. This kind of research can have direct implications for quality improvement programs.

**Conclusion**

This meta-analysis shows that survival after breast cancer surgery is significantly better with high volume providers. This suggests that concentration of breast cancer treatment in a limited number of centres might be beneficial. However, solid evidence for a specific minimal volume standard can not be identified in literature. And since volume is a simplistic proxy for quality of care, additional criteria have to be identified to direct centralization initiatives.

**Bibliography**

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Unit</th>
<th>Low</th>
<th>High</th>
<th>Period</th>
<th>Patients</th>
<th>Country</th>
<th>Data source</th>
<th>Data type</th>
<th>Casemix</th>
<th>Mortality %</th>
<th>Survival % (yr)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allgood 2006</td>
<td>Surgeon</td>
<td>10</td>
<td>50</td>
<td>1989-1997</td>
<td>2705</td>
<td>UK</td>
<td>Breast Test Wales</td>
<td>Combined</td>
<td>S</td>
<td>92.6 (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hospital</td>
<td>10</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bailie 2006</td>
<td>Surgeon</td>
<td>9</td>
<td>100</td>
<td>1996</td>
<td>807</td>
<td>North-Ireland</td>
<td>North-Ireland cancer registry</td>
<td>Clinical</td>
<td>C,S</td>
<td>70 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hospital</td>
<td>86</td>
<td>195</td>
<td>1997-1999</td>
<td>13360</td>
<td>Taiwan</td>
<td>Taiwan National Health Insurance Research dB</td>
<td>Administrative</td>
<td>C,S</td>
<td>74.6 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgeon</td>
<td>15</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen 2008</td>
<td>Hospital</td>
<td>86</td>
<td>195</td>
<td>1997-1999</td>
<td>13360</td>
<td>Taiwan</td>
<td>Taiwan National Health Insurance Research dB</td>
<td>Administrative</td>
<td>C,S</td>
<td>74.6 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgeon</td>
<td>15</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayforth 2007</td>
<td>Surgeon</td>
<td>19</td>
<td>20</td>
<td>1989-1999</td>
<td>1729</td>
<td>Australia</td>
<td>Western Australian CR</td>
<td>Clinical</td>
<td>S</td>
<td>83.6 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilligan 2007</td>
<td>Hospital</td>
<td>19</td>
<td>40</td>
<td>1994-1996</td>
<td>11225</td>
<td>USA</td>
<td>SEER-Medicare</td>
<td>Administrative</td>
<td>C,S</td>
<td>74.2 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guller 2005</td>
<td>Hospital</td>
<td>29</td>
<td>70</td>
<td>1988-2000</td>
<td>233247</td>
<td>USA</td>
<td>Nationwide Inpatient Sample</td>
<td>Administrative</td>
<td>C</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gutierrez 2008</td>
<td>Hospital</td>
<td>*</td>
<td>**</td>
<td>1994-2000</td>
<td>23745</td>
<td>USA</td>
<td>Florida Cancer Registry</td>
<td>Clinical</td>
<td>S</td>
<td>79 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hebert croteau 2005</td>
<td>Hospital</td>
<td>24</td>
<td>100</td>
<td>1988-1994</td>
<td>1727</td>
<td>Canada</td>
<td>Quebec Tumour Registry</td>
<td>Clinical</td>
<td>C,S</td>
<td>82 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simunovic 2006</td>
<td>Hospital</td>
<td>34</td>
<td>88</td>
<td>1991-1993</td>
<td>14346</td>
<td>Canada</td>
<td>Ontario Cancer Registry</td>
<td>Clinical</td>
<td>C,S</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stefoski Mikeljevic 2003</td>
<td>Hospital</td>
<td>49</td>
<td>100</td>
<td>1989-1994</td>
<td>11329</td>
<td>UK</td>
<td>Northern and Yorkshire Cancer Registry</td>
<td>Clinical</td>
<td>S</td>
<td>66 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgeon</td>
<td>9</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skinner 2003</td>
<td>Hospital</td>
<td>34</td>
<td>126</td>
<td>1990-1998</td>
<td>29666</td>
<td>USA</td>
<td>LA Cancer Surveillance Program</td>
<td>Administrative</td>
<td>S</td>
<td>80 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgeon</td>
<td>5</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* < 50th percentile, ** > 50th percentile; dB= database, CR= cancer registry; C= Comorbidity, S=Severity (tumour stage and/or adjuvant treatment)

Table 2 Characteristics of all studies included in meta-analysis
Table 1  Search terms used in the search in the databases Medline and Embase

<table>
<thead>
<tr>
<th><strong>Medline (Pubmed)</strong></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Embase</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>hospital volume.mp. OR surgeon volume.mp. OR workload.mp. OR Workload/ OR caseload.mp. OR procedure volume.mp. AND surgical complications.mp. OR postoperative complications.mp. OR exp Postoperative Complication/ OR Mortality/ OR mortality.mp. OR exp Survival Rate/ OR exp Survival/ OR survival.mp. OR exp Cancer Recurrence/ OR neoplasm recurrence.mp. OR treatment outcome.mp. OR exp Treatment Outcome/ OR surgical mortality.mp. OR exp Surgical Mortality/ OR exp recurrent disease/ OR exp tumor recurrence/ OR exp cancer survival/ OR disease free survival.mp. OR exp Disease Free Survival/ AND surgery/ OR exp cancer surgery/ OR surgeon.mp. OR exp Surgeon/ OR surgical procedures.mp. OR exp Surgical Technique/</td>
</tr>
</tbody>
</table>
Table 3 Sensitivity analysis of the heterogeneity of 6 included studies of hospital volume and surgeon volume and survival with HR as effect size

<table>
<thead>
<tr>
<th>Factor</th>
<th>Subgroup</th>
<th>Nr</th>
<th>HR</th>
<th>CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Datasource</td>
<td>Administrative</td>
<td>3</td>
<td>0.78</td>
<td>0.69-0.89</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Clinical</td>
<td>3</td>
<td>0.91</td>
<td>0.83-1.00</td>
<td>0.064</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>Not adjusted</td>
<td>2</td>
<td>0.94</td>
<td>0.88-0.99</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Adjusted</td>
<td>4</td>
<td>0.77</td>
<td>0.69-0.87</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.003</td>
</tr>
<tr>
<td>Surgeon volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Datasource</td>
<td>Administrative</td>
<td>1</td>
<td>0.77</td>
<td>0.69-0.85</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Clinical</td>
<td>3</td>
<td>0.85</td>
<td>0.74-0.99</td>
<td>0.033</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>Not adjusted</td>
<td>2</td>
<td>0.90</td>
<td>0.84-0.96</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Adjusted</td>
<td>2</td>
<td>0.76</td>
<td>0.69-0.84</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.007</td>
</tr>
</tbody>
</table>
Figure legends

Figure 1 Overview of the study selection process

Figure 2a Forest plot of the included studies on hospital volume and post-operative mortality

Figure 2b Forest plots of the included studies on hospital volume and survival for two effect sizes: Hazard Ratio’s and Relative Risk.

Figure 2c Funnel plots of the included studies in meta-analysis on surgeon volume and survival with two effect sizes: Hazard Ratio’s and relative risks.

Figure 3 Qualitative analysis of risk of publication bias: funnel plot of studies included in meta-analysis on hospital volume and survival using hazard ratio’s.

Figure 4 Qualitative analysis of risk of publication bias: funnel plot of studies included in meta-analysis on surgeon volume and survival using hazard ratio’s.
Potentially relevant studies identified and screened N=63

Studies retrieved for more detailed evaluation N=24

Studies included in meta-analysis N=12

39 studies excluded:
• 15 studies examined other outcome parameters
• 14 studies contained no primary data
• 10 studies volume is not an independent variable

13 Studies excluded:
• 2 studies: no risk adjustment done
• 4 studies: use of identical databases
• 5 studies: too old
• 1 study: volume is not an independent variable in multi variate analysis
### Hospital Mortality

<table>
<thead>
<tr>
<th>Model</th>
<th>Study name</th>
<th>Year</th>
<th>Low</th>
<th>High</th>
<th>Odds ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Guller [bct]</td>
<td>2005</td>
<td>29</td>
<td>70</td>
<td>0.329</td>
<td>0.121</td>
<td>0.892</td>
<td>-2.184</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>Guller [bct]</td>
<td>2005</td>
<td>29</td>
<td>70</td>
<td>0.526</td>
<td>0.269</td>
<td>1.028</td>
<td>-1.879</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>Simunovic</td>
<td>2006</td>
<td>34</td>
<td>88</td>
<td>0.100</td>
<td>0.011</td>
<td>0.914</td>
<td>-2.040</td>
<td>0.041</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.401</td>
<td>0.218</td>
<td>0.738</td>
<td>-2.937</td>
<td>0.003</td>
</tr>
</tbody>
</table>

### Hospital Survival

<table>
<thead>
<tr>
<th>Model</th>
<th>Study name</th>
<th>Year</th>
<th>Low</th>
<th>High</th>
<th>Hazard ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Allgood</td>
<td>2006</td>
<td>10</td>
<td>50</td>
<td>0.962</td>
<td>0.699</td>
<td>1.028</td>
<td>-1.151</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>Chen</td>
<td>2006</td>
<td>86</td>
<td>195</td>
<td>0.647</td>
<td>0.572</td>
<td>0.795</td>
<td>-0.795</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Gilligan</td>
<td>2007</td>
<td>19</td>
<td>40</td>
<td>0.830</td>
<td>0.749</td>
<td>0.919</td>
<td>-3.575</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Gutierrez</td>
<td>2006</td>
<td>8*</td>
<td>65*</td>
<td>0.903</td>
<td>0.828</td>
<td>0.964</td>
<td>-2.317</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>Hebert croteau</td>
<td>2005</td>
<td>24</td>
<td>100</td>
<td>0.686</td>
<td>0.496</td>
<td>0.946</td>
<td>-2.301</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>Simunovic</td>
<td>2006</td>
<td>34</td>
<td>88</td>
<td>0.833</td>
<td>0.704</td>
<td>0.986</td>
<td>-2.124</td>
<td>0.034</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.835</td>
<td>0.755</td>
<td>0.925</td>
<td>-3.474</td>
<td>0.001</td>
</tr>
</tbody>
</table>

### Hospital Survival

<table>
<thead>
<tr>
<th>Model</th>
<th>Study name</th>
<th>Year</th>
<th>Low</th>
<th>High</th>
<th>Risk ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skinner</td>
<td>2003</td>
<td>34</td>
<td>126</td>
<td>0.770</td>
<td>0.703</td>
<td>0.843</td>
<td>-5.619</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Stefoski Mikeljevic</td>
<td>2003</td>
<td>49</td>
<td>100</td>
<td>1.053</td>
<td>0.967</td>
<td>1.145</td>
<td>1.191</td>
<td>0.234</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.901</td>
<td>0.663</td>
<td>1.224</td>
<td>-0.669</td>
<td>0.504</td>
</tr>
</tbody>
</table>

### Surgeon Survival

<table>
<thead>
<tr>
<th>Model</th>
<th>Study name</th>
<th>Year</th>
<th>Low</th>
<th>High</th>
<th>Hazard ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Allgood</td>
<td>2006</td>
<td>10</td>
<td>50</td>
<td>0.900</td>
<td>0.838</td>
<td>0.967</td>
<td>-2.870</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Baillie</td>
<td>2007</td>
<td>9</td>
<td>100</td>
<td>0.613</td>
<td>0.391</td>
<td>0.963</td>
<td>-2.126</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>Chen</td>
<td>2008</td>
<td>15</td>
<td>67</td>
<td>0.766</td>
<td>0.693</td>
<td>0.847</td>
<td>-5.190</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Clayforth</td>
<td>2007</td>
<td>19</td>
<td>20</td>
<td>0.830</td>
<td>0.639</td>
<td>1.078</td>
<td>-1.396</td>
<td>0.163</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.816</td>
<td>0.718</td>
<td>0.927</td>
<td>-3.114</td>
<td>0.002</td>
</tr>
</tbody>
</table>

### Surgeon Survival

<table>
<thead>
<tr>
<th>Model</th>
<th>Study name</th>
<th>Year</th>
<th>Low</th>
<th>High</th>
<th>Risk ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nattinger</td>
<td>2007</td>
<td>4</td>
<td>10</td>
<td>0.860</td>
<td>0.752</td>
<td>0.983</td>
<td>-2.210</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>Skinner</td>
<td>2003</td>
<td>5</td>
<td>16</td>
<td>0.840</td>
<td>0.768</td>
<td>0.918</td>
<td>-3.840</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Stefoski Mikeljevic</td>
<td>2003</td>
<td>9</td>
<td>50</td>
<td>0.870</td>
<td>0.780</td>
<td>0.969</td>
<td>-2.521</td>
<td>0.012</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.853</td>
<td>0.803</td>
<td>0.907</td>
<td>-5.073</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Figure(s)
Figure(s)

Graph 1: Standard Error vs. Log Hazard Ratio

Graph 2: Standard Error vs. Log Hazard Ratio