This is a repository copy of Is there a relationship between surgical case volume and mortality in congenital heart disease services? A rapid evidence review.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/94178/

Version: Published Version

Article:

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Is there a relationship between surgical case volume and mortality in congenital heart disease services? A rapid evidence review

L Preston, J Turner, A Booth, C O'Keeffe, F Campbell, A Jesurasa, K Cooper, E Goyder

ABSTRACT

Objective: To identify and synthesise the evidence on the relationship between surgical volume and patient outcomes for adults and children with congenital heart disease.

Design: Evidence synthesis of interventional and observational studies.

Data sources: MEDLINE, EMBASE, CINAHL, Cochrane Library and Web of Science (2009–2014) and citation searching, reference lists and recommendations from stakeholders (2003–2014) were used to identify evidence.

Study selection: Quantitative observational and interventional studies with information on volume of surgical procedures and patient outcomes were included.

Results: 31 of the 34 papers identified (91.2%) included only paediatric patients. 25 (73.5%) investigated the relationship between volume and mortality, 7 (20.6%) mortality and other outcomes and 2 (5.9%) non-mortality outcomes only. 88.2% were from the US, 97% were multicentre studies and all were retrospective observational studies. 20 studies (58.8%) included all congenital heart disease conditions and 14 (41.2%) single conditions or procedures. No UK studies were identified. Most studies showed a relationship between volume and outcome but this relationship was not consistent. The relationship was stronger for single complex conditions or procedures. We found limited evidence about the impact of volume on non-mortality outcomes. A mixed picture emerged revealing a range of factors, in addition to volume, that influence outcome including condition severity, individual centre and surgeon effects and clinical advances over time.

Conclusions: The heterogeneity of findings from observational studies suggests that, while a relationship between volume and outcome exists, this is unlikely to be a simple, independent and directly causal relationship. The effect of volume on outcome relative to the effect of other, as yet undetermined, health system factors remains a complex and unresolved research question.

INTRODUCTION

An extensive evidence base supports an association between organisational factors and patient outcomes in elective surgery provision. The existence of a causal relationship between volume of activity and better patient outcomes is based on assumptions that more activity may be associated with better facilities, more experienced multidisciplinary teams and more experienced and specialist clinicians, rather than being simply attributable to increased workload. The volume and outcome association has been most extensively studied in the surgical specialities and for complex procedures where institutional and surgical experience and specialisation might be especially important in optimising outcomes. However, the underlying reasons for the observed associations between greater volumes of surgical activity and better outcomes for patients remain unclear and observed variations in outcomes, including mortality, remain unexplained.

Evidence on the relationship between volume and outcome of surgery is dominated by studies evaluating the relationship with...
mortality. However, volume may exert important effects on other patient outcomes such as morbidity and quality of life as well as service consequences, such as length of stay in hospital and costs.

Services for congenital heart disease (CHD) have been subject to scrutiny for over a decade, in the UK and internationally. In 2012 a series of recommendations was made for the reconfiguration of cardiac services for children in England. However, the process for making these recommendations was challenged and, following a judicial review, service reconfiguration was not implemented and a new service review considering the whole lifetime pathway for CHD undertaken.

The objective of this evidence synthesis was to inform the service review by examining whether there is evidence for a relationship between institutional and individual surgeon surgical volume and patient outcomes in CHD services. Evidence for other explanatory variables, including organisational features and other outcomes (such as complications) were examined in the full review.

Here we summarise the evidence for the specific relationship between surgical volume and outcome.

METHODS

We undertook a keyword-based systematic literature search using a predefined protocol, (see online supplementary file 1) enhanced by supplementary search methods. Reporting follows the PRISMA guidelines.

The review was completed within 3 months.

Search strategy

Relevant articles were identified using a database search strategy adapted from an earlier systematic review completed in 2009 (see online supplementary file 2). Search terms included population, volume, other organisational factors (eg, proximity to other services such as intensive care) and patient-related outcomes. We conducted searches (January and March 2014) of MEDLINE, EMBASE, CINAHL, Cochrane Library and Web of Science for the years 2009–2014.

These formal keyword-based literature searches were supplemented by four additional search methods designed to identify additional studies not included in the earlier systematic review for the 11-year period 2003–2009. These included citation searching using key references; responses from patient and public groups and clinical experts following a call for evidence; scrutiny of the reference lists of included papers and examination of the reference lists of published reviews, guideline documents and reports.

Selection criteria

Studies were eligible for inclusion if they reported an association between surgical volume (surgical unit or individual surgeon) and patient outcomes for children and/or adults undergoing treatment (surgical or interventional) for congenital heart disease. All types of patient-related outcomes (mortality, complications and quality of life) and health service outcomes (length of stay, costs) were eligible.

Studies eligible for inclusion were (1) observational studies and reports from trials. Qualitative or questionnaire-based studies were excluded. (2) Evidence from Organisation for Economic Cooperation and Development countries to ensure relative health system comparability to the UK. The review only included original research articles published in English and data from conference abstracts was excluded as these did not yield sufficient information. (3) Published in peer-reviewed journals to ensure that the evidence being synthesised had undergone methodological and expert scrutiny.

One author (LP) screened titles and abstracts using the inclusion and exclusion criteria. Second screening was undertaken for 10% of the references retrieved via the database searches by a second reviewer (KC) then five reviewers (JT, KC, AJ, CO and FC) screened the full text of any potentially relevant article. Each reviewer independently assessed the eligibility of each study, and the final list of included studies was agreed by consensus.

Data extraction

Five reviewers (JT, KC, AJ, CO and FC) independently extracted information into a standardised data extraction form, piloted on three studies and refined accordingly. The data extraction form collected information on the characteristics of each study, the results as reported and outcome.

One author (JT) appraised the evidence for the relationship between volume and outcome using a digital appraisal tool. A reliability rating was possible only for the primary outcome as differences were already evident at the level of study inclusion.

Quality assessment

As this was a rapid review we did not conduct a quality appraisal of individual included studies using a

---

Table 1: Items included in data extraction

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th>Study findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study dates</td>
<td>Volume analysed as continuous</td>
</tr>
<tr>
<td>Study aim</td>
<td>or categorical variable</td>
</tr>
<tr>
<td>Study design</td>
<td>Volume thresholds for</td>
</tr>
<tr>
<td>Data source and type</td>
<td>categorical variables</td>
</tr>
<tr>
<td>Study population</td>
<td>Covariates used in the analysis</td>
</tr>
<tr>
<td>Condition(s)</td>
<td>Crude associations of volume and outcome</td>
</tr>
<tr>
<td>Unit characteristics</td>
<td>Adjusted associations of volume and outcome</td>
</tr>
<tr>
<td>Intervention/procedure</td>
<td>Linear or non-linear relationship</td>
</tr>
<tr>
<td>Definition of volume</td>
<td>Summary of main findings</td>
</tr>
<tr>
<td>Outcomes measured</td>
<td>Summary of limitations identified by authors</td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
</tr>
<tr>
<td>Number of participants</td>
<td></td>
</tr>
<tr>
<td>Number of events</td>
<td></td>
</tr>
</tbody>
</table>

---

conventional quality assessment tool. Instead we used two complementary approaches to quality assessment to examine the collective contribution of the evidence base as a whole. First, we assessed the adequacy of the included evidence in addressing the aim of the research using a simple yes/no checklist for relevant factors including the study characteristics (eg, whether the study was single or multicentre or included more than one intervention/condition), the quality of the source data (eg, whether data collection was voluntary or mandatory) and the statistical analysis/adjustment (eg, whether the study adjusted for severity of condition and/or age). These relevance criteria for each included study are provided in online supplementary file 3. Second, we performed a study design level quality assessment to identify generic weaknesses. Similar study designs were coded with shared limitations. Judgements on quality were informed by limitations explicitly reported by study authors in the included studies. Assessment was undertaken by three authors (AB, JT and LP) and disagreements resolved by discussion.

Data synthesis
We extracted and tabulated the study information and used this to produce a narrative synthesis. A meta-analysis was not feasible given the considerable heterogeneity in the design, methods and settings of the included studies.

RESULTS
Study selection
The database search identified 2256 unique references of which 14 met inclusion criteria. An additional 20 papers were identified using additional non-database search methods giving a total of 34 included papers (figure 1).

Characteristics of the reviewed studies
The characteristics of the 34 included papers from 34 individual studies are summarised in table 2.

The majority of studies (88.3%) were conducted in the USA and most were multicentre (97%). No UK studies were identified. 31/34 included only paediatric patients. Thirty-one studies used routine data sets, including 19 voluntary clinical and/or administrative data sets, 12 using mandatory administrative data sets.5 Twenty-five of the 34 studies (73.5%) measured mortality or survival as the only outcome, eight studies measured mortality and other non-mortality outcomes including complications, length of stay in hospital, re-operation rates, length of ventilator treatment and time to extubation and costs. Two studies measured only non-mortality outcomes. Only 8 (25%) of the 32 studies reporting mortality measured this outcome postdischarge.

We have classified included studies into two groups—those where the primary objective was to explore the relationship between volume of service and outcomes for a range of CHD conditions (20/34) and those where the focus was on the relationship between volume and outcome for specific single conditions or procedures (14/34). For studies involving specific conditions or procedures these were mainly complex conditions such as hypoplastic left heart syndrome, transposition of great arteries and pulmonary atresia or procedures including Norwood Procedure, arterial switch operation and Blalock Taussig Shunt Procedure.8–41 Online supplementary table S3 provides a summary of the individual study characteristics for the two groups of included studies.

Findings as reported by the study authors
Results for the included studies are summarised in online supplementary table S4. In hospital mortality refers to death during the admission for the procedure. ORs signify the risk of death when different volumes are compared with 95% CIs where reported. Detailed analysis of the results of the 34 included studies is available in the full report.5

Studies on single conditions or procedures were more likely to identify an effect of volume on mortality but these focused on high risk conditions and procedures. Even within these highly selected groups there was considerable variation in effect depending on procedure type and individual centre or centre performance. The effect of surgeon volume illustrates this variability. Of four studies that included an examination of the effects of surgeon volume as well as centre volume, two found an association of decreasing mortality with increasing surgeon volume,35 41 one found increasing surgeon volume decreased mortality for only one of four complex conditions36 and one study found no association between surgeon volume and outcome.31

The findings from studies that included broader CHD populations were more equivocal. In some studies where an effect was identified, the effect was weak or only demonstrable for specific subgroups of patients. There was no clear indication that the evidence for the volume and mortality relationship was substantially stronger than the evidence for a no effect relationship in this broader group of patients Two large, comprehensively adjusted studies showed that, while a volume relationship exists, effects are small in comparison to factors such as condition severity and associated surgical risk, and surgical era.23 26

Overall, the evidence does demonstrate a relationship between volume and outcome in the majority of studies, although this relationship is not consistent. While volume is an important factor to consider the evidence highlights the complex relationship between volume, outcome and other factors which may also have an effect.

DISCUSSION
This review found a substantial body of evidence reporting a positive relationship between volume and
outcome, particularly for highly complex cases. However, interpretation is not straightforward. The 34 included studies revealed considerable variability highlighting the complexity of this relationship, as well as identifying variation in methods and findings across individual studies, and the methodological limitations imposed by the research approaches taken. Interpreting the evidence is particularly challenging due to a lack of information on clinical and service-related processes in the literature making it difficult to disentangle the volume/outcome relationship from other clinical and service processes and outcomes.

We have identified five key findings relevant to the organisation and delivery of CHD specialist services.

First, a range of factors influence mortality in CHD of which centre volume is only one. Our data extraction identified 67 different variables used to adjust for risk in the included studies, the most influential being condition severity.

Second, the included studies show that clinical advances, increasing expertise and changes in service provision have also influenced and improved outcomes for CHD over time. Five studies that analysed data over periods spanning up to 10 years found that, irrespective of other factors including volume and despite increasing complexity, mortality decreased over the study period. Therefore, the relevance of findings from historical data to contemporary services needs to be carefully considered.

Third, many studies used aggregated data from a large number of centres. Although this approach may show a difference in mortality rates between high and

---

Figure 1  Modified PRISMA flow chart of study identification, listing reason for exclusion during review process.
low volume centres, it may mask between-centre variation. Several included studies identified this variation with some low or medium volume centres performing as well as those with high volume indicating that individual centre effects are also likely to influence outcomes.

Fourth, the available evidence on the relationship between volume and mortality is dominated by retrospective studies conducted within the USA. It includes centres with very small volumes of cases limiting generalisability. The organisation of services in the USA is very different to the UK and other countries where centralisation of CHD services has led to a consequent increase in volume as more cases become concentrated in fewer centres. It remains unclear whether the impact of volume on outcome is largely a consequence of higher volume units organising and providing a complex service with all the ‘right’ components, is an independent factor directly related to the advantages of dealing with a larger number of cases or a combination of both. The lack of any UK studies to contribute to the review indicates a serious gap in evidence relevant to NHS service provision.

Finally, few studies are able to suggest an optimal size of a CHD centre in terms of volume. Less than half of the included studies analysed volume as a continuous variable which would provide the most robust evidence from which to consider volume thresholds.

The effects of some factors, such as condition severity, are well established but the effect of processes, systems and individual clinician effects on outcome remain unknown.

The full review also included evidence from three studies on adult CHD. One included heart transplant patients for a range of conditions in addition to CHD and so was of limited value. Two studies explored the effect of surgeon type in relation to outcome. Both studies found adult patients with CHD had better outcomes when operated on by paediatric surgeons in specialist children’s centres.

### Strengths and weaknesses of this study

This review was commissioned to inform an ongoing service review and was completed within 3 months. Rapid reviews have evolved primarily to inform emergent decision-making in healthcare settings. The short time frame and streamlined methodology that they utilise require a compromise between the need for efficiency against exhaustive evidence identification and synthesis. An examination of recent rapid reviews found considerable variation in the methodologies adopted and acknowledges that there is not a ‘one size fits all’ approach. Methods used should therefore be clear and transparent.

The key strengths of our approach are clear and reproducible methods for evidence searches; inclusion and exclusion criteria to identify relevant evidence and structured data extraction.

Time constraints meant we did not search exhaustively but aimed to identify all key evidence of relevance. It is possible that we may have missed relevant evidence. However, we did conduct citation searches on all included studies to minimise the likelihood of omitting eligible studies. Data extraction focused on identifying critical information for evidence synthesis rather than exhaustively extracting and critiquing all available information within individual papers. We were only able to conduct limited checking for screening and data extraction. A second reviewer screened 10% of the references identified from the searches. Data extraction was undertaken by five reviewers with but double data extraction was undertaken for a sample of included papers to refine the data extraction form and queries about data extraction or inclusions were resolved by discussion within the review team.

A meta-analysis of the evidence on volume and outcome was judged to be of limited value given the identified heterogeneity of context and populations. Further review of the broader fields of cardiac surgery outside CHD could contribute to identifying clinical and

---

### Table 2: Characteristics of included studies

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number full papers included</td>
<td>34 (100)</td>
</tr>
<tr>
<td>Volume and outcome relationship all</td>
<td>20 (58.8)</td>
</tr>
<tr>
<td>conditions</td>
<td></td>
</tr>
<tr>
<td>Volume and outcome relationship</td>
<td>14 (41.2)</td>
</tr>
<tr>
<td>specific conditions/procedures</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td></td>
</tr>
<tr>
<td>USA/Canada</td>
<td>30 (88.2)</td>
</tr>
<tr>
<td>Japan</td>
<td>2 (5.9)</td>
</tr>
<tr>
<td>Germany</td>
<td>1 (2.9)</td>
</tr>
<tr>
<td>Sweden</td>
<td>1 (2.9)</td>
</tr>
<tr>
<td>Centre type</td>
<td></td>
</tr>
<tr>
<td>Multicentre</td>
<td>33 (97)</td>
</tr>
<tr>
<td>Single centre</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Data sources</td>
<td></td>
</tr>
<tr>
<td>Voluntary (STS-CHD, HCUP-KIDS, PCCC, UHC)</td>
<td>19 (55.9)</td>
</tr>
<tr>
<td>Involuntary/registry (PHIS, NIS, OSHPD, UNOS, Texas birth defects registry)</td>
<td>12 (35.3)</td>
</tr>
<tr>
<td>Study specific</td>
<td>3 (8.8)</td>
</tr>
<tr>
<td>Patient population</td>
<td></td>
</tr>
<tr>
<td>All children (0–20)</td>
<td>19 (55.9)</td>
</tr>
<tr>
<td>Newborns and infants only</td>
<td>12 (35.3)</td>
</tr>
<tr>
<td>Adults</td>
<td>3 (8.8)</td>
</tr>
<tr>
<td>Outcomes measured</td>
<td></td>
</tr>
<tr>
<td>Survival/mortality only</td>
<td>25 (73.5)</td>
</tr>
<tr>
<td>Survival/mortality and other outcomes</td>
<td>7 (20.6)</td>
</tr>
<tr>
<td>Other outcomes</td>
<td>2 (5.9)</td>
</tr>
<tr>
<td>Design</td>
<td></td>
</tr>
<tr>
<td>Retrospective cohort</td>
<td>28 (82.4)</td>
</tr>
<tr>
<td>Retrospective observational</td>
<td>1 (2.9)</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>3 (8.8)</td>
</tr>
<tr>
<td>Longitudinal</td>
<td>1 (2.9)</td>
</tr>
<tr>
<td>RCT (data source)</td>
<td>1 (2.9)</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial.
service-related processes and outcomes that may be relevant and provide a framework for future data collection.

Instead of conducting a detailed quality appraisal of individual studies, we examined study methodology and generic study design limitations, including self-reported generic limitations, to construct a collective assessment of study quality.

**Strengths and weaknesses of included studies**

Information bias might result from missing data, miscoding or misinterpretation of information provided in routine databases. Several studies included in this review cited incomplete data as an issue, for example, missing surgeon identifiers, limited exploration of the surgeon volume and outcome relationship. Some data sources relied on voluntary completion, which introduces potential selection bias through coverage, membership or criteria for case submission. Inconsistency in coding, particularly over time, can lead to errors and routine databases may not include information on important contextual details about individual institutions such as team composition, training and experience, type of facility and access to specialist facilities, services and care pathways. Critical details such as non-intervention, transfers between institutions and pre-operative mortality are frequently not recorded. This lack of information means the ability to assess the impact of other aspects of care will remain constrained.

Data relating to a single institution is unlikely to be generalisable. Analysing data from a single year overcome some of the confounding effects related to structural or process changes over time and the associated danger that results measured at different time points may be misinterpreted. Study reports of a single surgical procedure can produce valuable insights for a discrete area of surgical practice but these usually involve rare and complex conditions and small numbers. This combined with the decreasing mortality reducing power, particularly as surgical procedures improve, limits the value of the reported results.

Included studies illustrate significant advances in methodology and analytic approaches over the time period covered by this review. Increasingly sophisticated tools to score for condition complexity and associated risk of mortality are being developed and methods for handling data as continuous, rather than a categorical, variables is now considered essential. The predominant method of using a step-wise volume category approach to establish a threshold for change in outcome used in many of the included studies is frequently criticised for being unsophisticated and misleading.

**Implications for future research**

Our review reveals a clear evidence gap in understanding the relationships between organisational factors in CHD services, how these can potentially predict a range of outcomes relevant to patients and their families, and the causal pathways between organisational factors and outcomes. Better understanding of these relationships is key to the development of methods for assessing and monitoring surgical performance that are not based solely on volume and mortality rates. While existing databases have value in helping understand some relationships and can help inform policy decisions there is scope to develop more comprehensive, high quality clinical and administrative databases to collect information on a range of organisational factors and outcomes related to quality of care. In the UK there is scope to expand the existing National Institute for Cardiovascular Outcomes Research (NICOR) database to capture more of this information. A more sophisticated information resource could then be used to conduct high quality studies of the relationship between organisational factors, volume and outcomes of direct relevance to the NHS and to improve the evidence base to support decisions about the organisation and delivery of CHD services.

**CONCLUSION**

This attempt to locate intervention or observational studies on the relationship between volume and other related organisational features and patient outcomes for adults and children with CHD identified a substantial volume of studies. Observational studies reported the relationship between volume and outcome in congenital heart services, particularly for paediatric surgery. This extensive body of evidence reveals a range of factors, in addition to volume, that influence outcome. These include condition severity, individual centre and surgeon effects and clinical advances over time. The heterogeneity of findings from observational studies suggests that, while a relationship between volume and outcome exists, this is unlikely to be a simple, independent and directly causal relationship. The effect of volume on outcome relative to the effect of other as yet undetermined health system factors remains a complex and unresolved research question.

**Contributors**

LP, JT, AB and EG conceived and designed the study. LP and AB designed the search strategy and undertook the searches. JT, FC, KC, AJ and COK contributed content to further drafts and outcomes. Better understanding of these relationships is key to the development of methods for assessing and monitoring surgical performance that are not based solely on volume and mortality rates. While existing databases have value in helping understand some relationships and can help inform policy decisions there is scope to develop more comprehensive, high quality clinical and administrative databases to collect information on a range of organisational factors and outcomes related to quality of care. In the UK there is scope to expand the existing National Institute for Cardiovascular Outcomes Research (NICOR) database to capture more of this information. A more sophisticated information resource could then be used to conduct high quality studies of the relationship between organisational factors, volume and outcomes of direct relevance to the NHS and to improve the evidence base to support decisions about the organisation and delivery of CHD services.

LP, JT, AB, EG and COK contributed to the analysis and interpretation of data. JT, FC, KC, AJ and COK critically revised the manuscript. LP, JT, AB, EG, FC, KC, AJ and COK contributed to the drafting of the manuscript. JT is the guarantor of the study. All authors had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis, and revised and approved the final version of the article.

**Funding**

The study was funded as an independent research project by The National Institute for Health Research HS&DR Programme grant number 13/05/12. The funders contributed to the development of the research questions to be addressed. The funders had no role in conducting the study, writing the paper or the decision to submit the paper for publication.

**Competing interests**

None declared.

**Provenance and peer review**

Not commissioned; externally peer reviewed.

**Data sharing statement**

No additional data are available.
REFERENCES


