The relationship between procedural volume and patient outcomes for percutaneous coronary interventions: a systematic review and meta-analysis [version 1; peer review: 2 approved]

Kieran A. Walsh\textsuperscript{1}, Thomas Plunkett\textsuperscript{1}, Kirsty K. O'Brien\textsuperscript{1}, Conor Teljeur\textsuperscript{1}, Susan M. Smith\textsuperscript{2}, Patricia Harrington\textsuperscript{1+}, Máirín Ryan\textsuperscript{1,3*}

\textsuperscript{1}Health Technology Assessment (HTA) Directorate, Health Information and Quality Authority, Dublin 7, Ireland
\textsuperscript{2}Health Research Board Centre for Primary Care Research, Department of General Practice, Royal College of Surgeons in Ireland, Dublin 2, Ireland
\textsuperscript{3}Department of Pharmacology & Therapeutics, Trinity College Dublin, Dublin 8, Ireland

\textsuperscript{*} Equal contributors

\textbf{Open Peer Review}

\textbf{Invited Reviewers}

\textbf{1. Alexander Fanaroff, Hospital of the University of Pennsylvania, Philadelphia, USA}

\textbf{2. JJ Coughlan, Technische Universität München, Munich, Germany}

Any reports and responses or comments on the article can be found at the end of the article.
volume hospitals for primary PCI procedures (OR: 0.77, 95% CI: 0.62-0.94, $I^2 = 78\%$). Overall, GRADE certainty of evidence was 'very low'. There were mixed findings for secondary outcomes.

Conclusions: A volume-outcome relationship may exist in certain situations, although this relationship appears to be attenuating with time, and there is ‘very low’ certainty of evidence. While volume might be important, it should not be the only standard used to define an acceptable PCI service and a broader evaluation of quality metrics should be considered that encompass patient experience and clinical outcomes.

Systematic review registration: PROSPERO, CRD42019125288

Keywords
Systematic review, meta-analysis, volume-outcome, PCI, health services research, STEMI, heart attack, myocardial infarction.
Introduction
The volume-outcome relationship refers to the association between volumes of procedures and outcomes\(^1\). The hypothesis underpinning this relationship is that ‘practice makes perfect’, i.e. hospitals or operators that perform a larger number of procedures will achieve better outcomes than those that perform relatively fewer procedures\(^2\). Previous systematic reviews have examined the relationship between percutaneous coronary interventions (PCI) volume and postoperative mortality\(^4,6,6\). A systematic review and meta-analysis published in 2010 investigated the volume-outcome relationship for all PCI procedures, exclusively at the hospital-level\(^4\). The authors calculated a pooled effect estimate with an odds ratio (OR) of 0.87 (95% confidence interval [CI]: 0.83-0.91, \(F = 38\%\)) in favour of high-volume hospitals. A more recent systematic review and meta-analysis published in 2016, also evaluated this relationship at the hospital-level\(^8\), again finding in favour of high-volume hospitals (OR: 0.79, 95% CI: 0.72-0.86, \(F = 38\%\)). A systematic review and meta-analysis published in 2014, investigated the volume-outcome relationship at the operator-level\(^9\). The pooled effect estimate showed no significant relationship between operator volume and mortality (OR: 0.96, 95% CI: 0.86-1.08, \(F = 61\%\)) though a significant inverse relationship was found between operator volume and major adverse cardiac events [MACE] (OR: 0.62, 95% CI: 0.40-0.97, \(F = 97\%\)).

Evidence from studies in these reviews has informed the development of minimum volume criteria\(^3,9\), which in turn have informed healthcare policy and service provision standards\(^6-11\). However, many have argued that minimum volume criteria should no longer be prioritised as a key metric for PCI service delivery given advances in PCI techniques and postoperative medical management, and regionalisation of care which may have attenuated the volume-outcome relationship over time\(^12\). Furthermore, given the inclusion of older, often poorer quality studies in these previous systematic reviews\(^4,6,8\), an updated review of the volume-outcome relationship at both the hospital- and operator-level is warranted. The aim of this study is to re-examine the relationship between PCI procedural volume and patient outcomes, in light of advances in interventional cardiology and emerging evidence.

Methods
A systematic review and meta-analysis was undertaken to examine the relationship between procedural volumes (both at the hospital- and operator-level) and patient outcomes, in adults requiring PCI. The primary outcome was mortality. The secondary outcomes were procedural complications, healthcare utilisation and process outcomes. This study is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Reporting guidelines\(^7\))\(^9\). The protocol was registered on PROSPERO (CRD42019125288).

Search strategy
Electronic searches were conducted in PubMed, Embase, CINAHL Plus and the Cochrane Library for the period 1 January 2008 to 28 May 2019. Due to significant advances in PCI practices and perioperative management, only studies examining contemporary evidence, published since 2008 were included\(^12\). Grey literature sources were searched (Extended data: S2\(^9\)), along with the first five pages of Google and Google Scholar. The search strategy used search terms (Extended data: S3\(^9\)) adapted from an earlier systematic review\(^6\). Additional search methods included forward citation searching of eligible studies, hand searching relevant journals and systematic reviews and searching reference lists of included studies.

Study selection criteria, data extraction and quality appraisal
Published observational studies examining the relationship between total PCI (PCI for any acute or elective indication) or primary PCI (PPCI) (emergent angioplasty without the previous administration of fibrinolytic therapy for ST-elevation Myocardial Infarction [STEMI]) procedural volume and patient outcomes were included according to the inclusion and exclusion criteria in Table 1. Screening, data extraction and quality appraisal were all conducted independently by two reviewers from the research team with any disagreements being resolved by discussion, and where necessary, a third reviewer. Covidence (www.covidence.org) was used for data management and extraction purposes. The data extraction tool was piloted on two studies initially.

The following data were extracted from each included study:
- year of publication
- country
- clinical trial registration
- database
- data type (for example administrative or clinical)
- study period
- study design
- number of patients/procedures
- number of hospitals
- number of operators
- hospital volume classification (in terms of cases per year)
- definition of high-volume hospital
- definition of high-volume operator
- how cut-points are selected (for example data-driven, guideline-based)
- volume grouping (for example quartiles, median)
- risk adjustment covariates
  - process measures (for example distance to hospital, time to treatment, out of hospital cardiac arrest, radial artery access, use of drug-eluting stents)
  - demographics of patient population (for example sex, indication, stent)
  - patient comorbidities (for example heart failure)
Table 1. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Relationship between hospital or operator volume and PCI outcomes is investigated</td>
<td>- Multiple publications based on the same database</td>
</tr>
<tr>
<td>PCI defined as follows:</td>
<td>- No definition of procedural volume as a distinct number or cut-off value</td>
</tr>
<tr>
<td>- Total PCI - PCI for any acute or elective indication.</td>
<td>- Conference papers and abstracts where the full paper was unobtainable</td>
</tr>
<tr>
<td>- Primary PCI - emergent angioplasty without the previous administration of</td>
<td>- Paper published prior to 2008</td>
</tr>
<tr>
<td>fibrinolytic therapy to open the infarct-related artery during a STEMI.</td>
<td>- Paediatric (&lt;18 years old) population</td>
</tr>
<tr>
<td>- Published observational studies</td>
<td>- Operator volume and postoperative (in-hospital/30-day) mortality.</td>
</tr>
<tr>
<td>- Study uses primary data</td>
<td>- hospital cluster effect</td>
</tr>
<tr>
<td>- Study reports at least one of the predefined primary outcomes</td>
<td>- hospital characteristics (for example presence of on-site surgical cover)</td>
</tr>
<tr>
<td>- Study reports adjusted rates</td>
<td>- severity of disease (for example cardiogenic shock)</td>
</tr>
<tr>
<td>- For hospital volume studies only: does not describe the results obtained at a</td>
<td>- treatment differences (for example salvage PCI)</td>
</tr>
<tr>
<td>single centre</td>
<td>- difference in findings between middle groups and highest/lowest groups.</td>
</tr>
<tr>
<td>- For operator studies only: does not describe the results of a single operator</td>
<td>- The primary outcomes:</td>
</tr>
<tr>
<td></td>
<td>- mortality</td>
</tr>
<tr>
<td></td>
<td>The secondary outcomes:</td>
</tr>
<tr>
<td></td>
<td>- complications of PCI (for example major adverse cardiac events (MACE)/ major</td>
</tr>
<tr>
<td></td>
<td>adverse cardiac and cardiovascular events (MACCE), emergency coronary artery</td>
</tr>
<tr>
<td></td>
<td>bypass graft (CABG), bleeding, peri-procedural myocardial infarction, vascular</td>
</tr>
<tr>
<td></td>
<td>complications, stroke, contrast-induced nephropathy and stent thrombosis)</td>
</tr>
<tr>
<td></td>
<td>- process outcomes (for example time-to-treatment and appropriateness of PCI)</td>
</tr>
<tr>
<td></td>
<td>- healthcare utilisation outcomes (for example hospital readmission, hospital</td>
</tr>
<tr>
<td></td>
<td>length of stay, unplanned repeat vascularisations).</td>
</tr>
<tr>
<td></td>
<td>A modified version of the Critical Appraisal Skills Programme (CASP) tool for cohort</td>
</tr>
<tr>
<td></td>
<td>studies was used for quality appraisal at the study-level\textsuperscript{15}.</td>
</tr>
<tr>
<td></td>
<td>The tool was piloted by two reviewers initially.</td>
</tr>
<tr>
<td></td>
<td>Data synthesis</td>
</tr>
<tr>
<td></td>
<td>The methods for data synthesis are summarised here, with greater detail online</td>
</tr>
<tr>
<td></td>
<td>(Extended data: S4\textsuperscript{15}). Meta-analysis was performed for the primary</td>
</tr>
<tr>
<td></td>
<td>outcome, if appropriate to determine the relationship between:</td>
</tr>
<tr>
<td></td>
<td>- Hospital volume and postoperative (in-hospital/30-day) mortality</td>
</tr>
<tr>
<td></td>
<td>Meta-analyses of adjusted odds ratios (aOR) were conducted separately for studies</td>
</tr>
<tr>
<td></td>
<td>reporting outcomes for total PCI procedures and for studies reporting outcomes</td>
</tr>
<tr>
<td></td>
<td>for PPCI procedures only. Pooled estimated effect sizes were calculated using</td>
</tr>
<tr>
<td></td>
<td>the adjusted outcomes of the highest volume group compared with the lowest</td>
</tr>
<tr>
<td></td>
<td>volume group. RevMan version 5.3 was used to conduct the random-effects, inverse-</td>
</tr>
<tr>
<td></td>
<td>variance meta-analysis.</td>
</tr>
<tr>
<td></td>
<td>Pre-planned, sensitivity, subgroup and random-effects meta-regression analyses</td>
</tr>
<tr>
<td></td>
<td>(STATA version 13, StataCorp, College Station, TX, USA) were conducted to explore</td>
</tr>
<tr>
<td></td>
<td>anticipated heterogeneity and to assess the effect of various studies, sub-</td>
</tr>
<tr>
<td></td>
<td>groups and quality on the overall outcome. Where a temporal trend was observed,</td>
</tr>
<tr>
<td></td>
<td>a random-effects cumulative meta-analysis was conducted using the median year of</td>
</tr>
<tr>
<td></td>
<td>study data. A narrative synthesis was undertaken for the findings not included in</td>
</tr>
<tr>
<td></td>
<td>the meta-analysis.</td>
</tr>
<tr>
<td></td>
<td>Certainty of the evidence</td>
</tr>
<tr>
<td></td>
<td>The certainty of the evidence for each primary outcome was assessed independently by</td>
</tr>
<tr>
<td></td>
<td>two reviewers using the ‘Grading of Recommendations, Assessment, Development and</td>
</tr>
<tr>
<td></td>
<td>Evaluations’ (GRADE) approach\textsuperscript{16}. A summary of findings table</td>
</tr>
<tr>
<td></td>
<td>using the GRADEpro software were generated for the primary outcomes\textsuperscript{17}.</td>
</tr>
<tr>
<td></td>
<td>Results</td>
</tr>
<tr>
<td></td>
<td>Search results</td>
</tr>
<tr>
<td></td>
<td>Searching electronic databases identified 1,730 records; searching grey literature</td>
</tr>
<tr>
<td></td>
<td>and other sources identified another 55. After removal of duplicates, 1,154</td>
</tr>
<tr>
<td></td>
<td>records were screened, with 1,017 excluded based on titles and abstracts, leaving</td>
</tr>
<tr>
<td></td>
<td>a total of 137 full-text articles to be assessed for eligibility. Of these, 115</td>
</tr>
<tr>
<td></td>
<td>references were excluded (Extended data: S5\textsuperscript{15}). This resulted</td>
</tr>
<tr>
<td></td>
<td>in 22 studies being included, of which 16 were included in meta-analyses (Figure</td>
</tr>
<tr>
<td></td>
<td>1).</td>
</tr>
</tbody>
</table>
Characteristics of included studies

**Study country, population and design.** Of the 22 included studies, ten were conducted in the US\(^{18-27}\), with the remainder conducted in Japan (n=5)\(^{28-32}\), England and Wales (n=2)\(^{33,34}\), South Korea (n=1)\(^{35}\), Italy (n=1)\(^{36}\), China (n=1)\(^{37}\), Taiwan (n=1)\(^{38}\) and Germany (n=1)\(^{39}\). In total, 6,432,265 patients or procedures were included across the 22 studies. Median or mean age of the study populations ranged from 59 years\(^{21}\) to 74 years\(^{23}\). The majority of included patients were male, ranging from 62%\(^{23}\) to 97.9%\(^{21}\). All included studies were observational; with 20 cross-sectional studies\(^{18-20,22-26,36,38,39}\) and two cohort studies\(^{11,17}\). Details of the individual study characteristics and outcomes, and the level of analysis of included studies (i.e. hospital-level vs. operator-level) are outlined in the Extended data (S6 and S7, respectively)\(^{13}\).

**Definition of high- and low-volume.** Definitions of high and low volume varied between studies. Some studies developed thresholds that were predominantly data-driven\(^{19,20,26,29,30,32,36,39}\) (for example, dividing the population into two or more equal sized groups), some developed thresholds that were predominantly guideline-driven\(^{21,22,27,28,33-35,38}\), and others did not provide any clear explanation\(^{18,37}\).

How the volumes were grouped (tertiles, quartiles, etc) varied substantially between studies, with the number of quantiles...
ranging from two21–26,28,31,33,36–39 to 1029. Studies also tested the effect of changing the threshold, or how the threshold/volume was calculated, on the overall outcome.22,25–27,30,33,38,39 Some of the alternative thresholds changed the results of the study.26,27,39

In general, the proportion of PCI procedures provided in low-volume settings or by low-volume providers decreased over time. When comparing the oldest with the newest studies, the proportion of procedures provided in low-volume settings decreased from: 20.6%39 to 2.2%27 for total PCI and 57.2%31 to 0.07%34 for PPCI; while for low-volume operators the proportion of procedures decreased from 25%30 to 4.8%27 for total PCI. The lowest volume groupings were found to perform a disproportionately higher number of emergent procedures than their higher volume counterparts in six of eight studies that reported this information.14,20,22,23,28,39

Primary outcomes
Postoperative mortality rates (aggregated at the study level) ranged from 0.9%29 to 2.6%31 following total PCI procedures, with a mean mortality rate of 1.5%. For patients undergoing PPCI procedures, postoperative mortality rates ranged from 3.2%25 to 10.1%31, with a mean mortality rate of 5.3% (Extended data: S613).

Total PCI at the hospital-level. For total PCI procedures, the relationship between hospital volume and mortality, was investigated in nine studies20,22,23,25–27,29,34,36–39 with no statistically significant difference found between the highest and the lowest volume hospitals (OR: 0.84, 95% CI: 0.69-1.03) (Figure 2). Of note there was considerable heterogeneity (I² = 86%).

A temporal trend is evident when a cumulative meta-analysis was conducted (Figure 3). A gradual change can be seen with the pooled effect estimate shifting from a 33% reduction in the odds of mortality (OR: 0.67, 95% CI: 0.52-0.86) when limited to the earliest study data29 to a 16% reduction (OR: 0.84, 95% CI: 0.69-1.03) when data from the most recent studies are included.27 Notably, the difference is no longer statistically significant when study data from the year 2010 onwards is included.49

From the exploratory meta-regression analyses, no covariate reached statistical significance (Extended data: S815). Subgroup analyses indicated that the overall pooled effect estimate was sensitive to bias of case-mix adjustment. Sensitivity analyses found that the overall pooled effect estimate remains non-significant under all alternative threshold scenarios, except when a threshold of 400 PCI procedures per year or higher was used in one study27, indicating the impact that shifting the threshold can have on the overall findings (Extended data: S915).

Total PCI at the operator-level. Six studies investigated the relationship between operator volume and mortality, for all PCI procedures (Figure 4).20,22,23,27,29,33,38 The pooled effect estimate was found to be significantly in favour of high-volume operators (OR: 0.77, 95% CI: 0.63-0.94). Of note there was considerable heterogeneity (I² = 93%).

From the exploratory meta-regression analyses, no covariate reached statistical significance (Extended data: S815). Subgroup analyses did not reveal any significant between-group differences. In sensitivity analyses, a study-by-study exclusion process found that the volume-outcome relationship remains significant except when the studies by Badheka et al.29 or Fanaroff et al.32 are removed. Thus, indicating the strong influence of these large population-based studies (Extended data: S915).

Primary PCI at the hospital-level. Seven studies investigated the relationship between hospital volume and mortality, specifically for PPCI procedures (Figure 5).24–26,28,31,34,36 The pooled effect estimate was found to be significantly in favour of high-volume hospitals (OR: 0.77, 95% CI: 0.62-0.94). Heterogeneity again was considerable (I² = 78%).

From the exploratory meta-regression analyses, no covariate reached statistical significance (Extended data: S815).

Figure 2. Results of the meta-analysis investigating the relationship between hospital volume and mortality, for total PCI procedures. Studies are arranged in chronological order based on date of publication. Point estimates to the left of, and confidence intervals not crossing the “line of no effect”, indicate that the effect favours high volume. CI, confidence interval; IV, inverse variance; PCI, percutaneous coronary intervention; SE, standard error.
Figure 3. Cumulative meta-analysis investigating the relationship between hospital volume and mortality, for total PCI procedures. Studies are arranged in chronological order based on median year of data collection. Point estimates to the left of, and confidence intervals not crossing the "line of no effect", indicate that the effect favours high volume. CI, confidence interval; ID, Identification.

Figure 4. Results of the meta-analysis investigating the relationship between operator volume and mortality, for total PCI procedures. Studies are arranged in chronological order based on date of publication. Point estimates to the left of, and confidence intervals not crossing the "line of no effect", indicate that the effect favours high volume. CI, confidence interval; IV, inverse variance; PCI, percutaneous coronary intervention; SE, standard error.

Subgroup analyses indicated that the overall pooled effect estimate was sensitive to the mortality outcome used. As only one study in this meta-analysis used 30-day mortality rates, caution is needed when interpreting this finding. Sensitivity analyses did not reveal any factor that changed the significance of the overall pooled effect estimate (Extended data: S9).
Two studies investigated the relationship between hospital volume and mortality, for primary PCI procedures. Studies are arranged in chronological order based on date of publication. Point estimates to the left of, and confidence intervals not crossing the “line of no effect”, indicate that the effect favours high volume. CI, confidence interval; IV, inverse variance; PCI, percutaneous coronary intervention; SE, standard error.

**Figure 5.** Results of the meta-analysis investigating the relationship between hospital volume and mortality, for primary PCI procedures. Studies are arranged in chronological order based on date of publication. Point estimates to the left of, and confidence intervals not crossing the “line of no effect”, indicate that the effect favours high volume. CI, confidence interval; IV, inverse variance; PCI, percutaneous coronary intervention; SE, standard error.

**Primary PCI at the operator-level.** Only two studies investigated the relationship between PPCI at the operator level and mortality; hence a meta-analysis was not conducted. A 2009 study based on procedures conducted on 7,321 patients between 2000 and 2002, found the odds of mortality were 34% lower (OR: 0.66, 95% CI: 0.48-0.91) for procedures undertaken by operators completing more than 10 PPCI procedures annually compared with those undertaken by operators performing 10 or fewer procedures. However, a larger (n=133,970) study based on more recent data (2013–2014) found no significant association between operator volume (at a threshold of 75 total PCI procedures per year) and mortality, following PPCI (OR: 0.93, 95% CI: 0.72-1.20).

**Minimum volume threshold.** Ten studies reported annual thresholds above which the adjusted odds ratio for mortality became non-significant. These cut-points ranged from 208 to 400 procedures for total PCI hospital volume; from 150 to 100 for total PCI operator volume; from 364 to 66 for PPCI hospital volume; and 10 in the sole study which used PPCI operator cut-points. However, in three studies that had more than two quantiles, some intermediate groups were found to have better outcomes than adjacent higher volume groups. Six studies conducted spline analysis to investigate the dynamic relationship between volume and outcome, the estimated optimum annual thresholds varied between studies, ranging from 100 to 1,000 total PCI hospital procedures.

**Long-term mortality outcomes.** Two studies investigated the relationship between volume and long-term (i.e., greater than 30 days) primary outcomes. After adjustment for confounders, no significant difference in mortality between high- and low-volume operators was found at one year (Hazard ratio (HR): 1.04, 95% CI: 1.00-1.08) or three years (HR: 0.70, 95% CI: 0.45–1.11) post-procedure. Notably, both studies had found significant differences in mortality between high- and low-volume groups at-hospital and 30-day time-points.

**Secondary outcomes**
A number of secondary outcomes were reported across twelve studies. These included healthcare utilisation or process lengths (length of stay, door-to-balloon (DTB) time, re-admission, and inappropriate use of PCI) and PCI complication outcomes (bleeding, dialysis, recurrent MI, unplanned revascularisations, and a range of composite outcomes including MACE). A consistently significant relationship between procedural volume and healthcare utilisation or process outcomes, was found in favour of high-volume operators and hospitals, with the exception of PCI procedure appropriateness, which had conflicting findings, and PCI procedural volume and PCI complications (Extended data: S10).

**Quality appraisal**
Using the CASP quality appraisal tool, eight studies were judged to have an overall low risk of bias and five a high risk of bias (Figure 6) and five an unclear risk of bias across many included studies, considerable levels of heterogeneity and some concerns regarding the imprecision of results (Figure 7).

**Certainty of the evidence**
Overall, the certainty of the evidence is ‘very low’ owing to the observational nature of included studies, a high or unclear risk of bias across many included studies, considerable levels of heterogeneity and some concerns regarding the imprecision of results (Figure 7).

**Discussion**
This systematic review and meta-analysis of contemporary research suggests that a significant inverse relationship between PCI procedural volume and patient outcomes may exist in certain situations, however results must be interpreted with caution due to the very high levels of heterogeneity and the ‘very low’ certainty of the evidence. Specifically, no significant association was found between total PCI hospital volume and postoperative mortality, with a temporal trend observed from significant to non-significant pooled effect estimates. Conversely, a significant inverse volume-outcome relationship was found between total PCI operator volume and postoperative mortality, as well as between PPCI hospital volume and postoperative mortality. There is some evidence to suggest that high-volume hospitals offer other benefits in terms of association with reduced length of stay and a greater likelihood of achieving...
target DTB times. However, there is also evidence to suggest that most, if not all, of the benefits conferred by high-volume operators are attenuated after 30 days\textsuperscript{23,37}. Due to the huge variability in how studies defined low- and high-volume and differences in how they analysed the data, it is not possible to determine with any degree of certainty, a threshold above which the volume-outcome relationship becomes non-significant.

A key finding of our study was the temporal trend from significant to non-significant pooled effect estimates observed for total PCI hospital volume. With advances in interventional cardiology including increasingly sophisticated operating techniques, more effective drug-eluting stents and improvements in medical management, it is likely that some of these factors may have mitigated the importance of volume on

Figure 6. Quality Appraisal of Included Studies using the CASP tool. A modified version of the CASP tool for cohort studies was used for quality appraisal at the study-level. CASP, Critical Appraisal Skills Programme.
Table 7. GRADE Summary of findings table. A summary of findings table using the GRADEpro software were generated for the primary outcome. GRADE, Grading of Recommendations, Assessment, Development and Evaluations.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of participants (studies Follow-up)</th>
<th>Certainty of the evidence (GRADE)</th>
<th>Relative effect (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality in total PCI at hospital level</td>
<td>(9 observational studies)</td>
<td>⊖⊖⊖⊖⊖ VERY LOW a,b</td>
<td>OR 0.84 (0.69 to 1.03)</td>
</tr>
<tr>
<td>Mortality in total PCI at operator level</td>
<td>(6 observational studies)</td>
<td>⊖⊖⊖⊖⊖ VERY LOW c,d</td>
<td>OR 0.77 (0.63 to 0.94)</td>
</tr>
<tr>
<td>Mortality in primary PCI at hospital level</td>
<td>(7 observational studies)</td>
<td>⊖⊖⊖⊖⊖ VERY LOW e,f</td>
<td>OR 0.77 (0.62 to 0.94)</td>
</tr>
<tr>
<td>Mortality in primary PCI at operator level</td>
<td>(2 observational studies)</td>
<td>⊖⊖⊖⊖⊖ VERY LOW f,h</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*aThe risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

**Certainty of evidence:**
- **High certainty:** We are very confident that the true effect lies close to that of the estimate of the effect.
- **Moderate certainty:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
- **Low certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
- **Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of the effect.

**Explanations**

- a. Of the 9 studies, 3 were low risk of bias, 5 were unclear risk of bias and 1 was high risk of bias
- b. I² = 86% hence there is substantial heterogeneity
- c. Of the 6 studies, 4 were low risk of bias and 2 were unclear risk
- d. I² = 93% hence there is considerable heterogeneity
- e. Of the 7 studies, 3 were low risk, 3 were unclear risk and 1 was high risk of bias
- f. I² = 78% hence there is substantial heterogeneity
- g. Only 2 studies and both had conflicting results
- h. Wide confidence intervals for both studies
mortality\textsuperscript{12}. Furthermore, introduction of advanced systems of care, streamlined processes and governance structures may also have improved standards across the board\textsuperscript{7}. Due to the implementation of minimum volume standards, it is possible that the observed decrease in the proportion of low-volume hospitals and operators over time may also have moderated the volume-outcome relationship\textsuperscript{14}. What constitutes ‘low-volume’ appears to have changed over time, and the use of traditional cut-points may no longer be sufficiently sensitive to detect a significant difference in outcomes. However, no temporal change was apparent for PPCI hospital volume. Sensitivity analyses conducted for this outcome suggests that there is robustness around this particular pooled effect estimate. Differences in PPCI and total PCI findings may reflect differences in clinical complexity and decision-making in emergent compared with elective cases.

Our findings contrast with those of the earlier systematic reviews\textsuperscript{4-6}. Our systematic review does however indicate that a significant volume-outcome relationship may exist at the hospital-level for PPCI procedures, and at the operator-level for total PCI. Outcomes from studies evaluating primary and total PCI procedures were combined in previous systematic reviews\textsuperscript{4-6}, which may have introduced bias into the overall findings due to the inherently higher risk of mortality in STEMI patients and the likely significant confounding due to case-mix in all of these studies. Furthermore, these previous systematic reviews included studies based on PCI data as old as 1984\textsuperscript{4}, 1990\textsuperscript{5} and 1996\textsuperscript{5}. In contrast, our systematic review which is based on PCI data ranging from 2000 to 2016, provides a more contemporary evidence base.

The main strength of this review was the comprehensive search, in-depth analysis and confirmatory methods including meta-regression, adopted by a team of reviewers experienced in the conduct of systematic reviews. Furthermore, by focusing on more recent data our pooled effect estimates likely better reflect current practice than those calculated by previous systematic reviews\textsuperscript{4-6}. One of the main limitations of this study was the evidence of a considerable level of heterogeneity in the quality and design of studies; hence the calculated pooled effect estimates must be viewed with caution. Moreover, due to the observational nature of included studies, there are inherent issues surrounding unknown confounding and imbalanced case-mix which may have introduced bias into the results. To address these issues meta-regression, subgroup analyses and sensitivity analyses were conducted. However, no definitive cause of heterogeneity could be determined. Due to the limited number of included studies, the meta-regression analyses should be treated as exploratory and we urge caution in their interpretation.

Our findings have implications for policy and practice. As health care systems aim to create STEMI networks based on hub-and-spoke models\textsuperscript{30}, our findings suggest there may be a potential mortality benefit of high-volume centres for PPCI procedures in particular. However, an important observation among included studies was that a disproportionately higher number of emergent procedures tended to be conducted by low-volume hospitals and operators\textsuperscript{18,20,22,23,28,39}. This finding indicates the important role these low-volume hospitals and operators may have in terms of serving sparsely populated regions. Although enforcing minimum volume standards has been argued as a means of improving standards of care across the board\textsuperscript{14}, it has also been argued that this may unintentionally incentivise some operators and hospitals to perform PCI in patients who have a lower capacity to benefit from the procedure, in order to meet minimum volume requirements\textsuperscript{12}. Furthermore, some have argued that volume in isolation should no longer be prioritised as a key metric for PCI service delivery, and instead a more holistic and multifaceted approach to quality assessment should be adopted\textsuperscript{20,12}. Therefore, service planners need to carefully balance the need to organise STEMI networks around high-volume centres and operators, while meeting the needs of the entire population, in order to achieve a high-quality, efficient and equitable system with good patient outcomes. Further research is required to inform the optimum configuration of such a network.

In conclusion, this systematic review and meta-analysis suggests a significant inverse relationship between PCI procedural volume and patient outcomes may exist in certain situations; however, results must be interpreted with caution due to the very high levels of heterogeneity and the ‘very low’ certainty of the evidence. A temporal trend was observed, indicating that the relationship between total PCI hospital volume and mortality may be attenuating over time. While a volume-outcome relationship may exist under certain circumstances and might be important, volume should not be the only standard used to define an acceptable PCI service and a broader evaluation of quality metrics should be considered that encompass patient experience and clinical outcomes.

**Data availability**

**Underlying data**

All data underlying the results are available as part of the article and no additional source data are required.

**Extended data**

Figshare: Extended Data: The relationship between procedural volume and patient outcomes for PCI, [https://doi.org/10.6084/m9.figshare.13388060.v1\textsuperscript{13}]

This project contains the following extended data:

- Grey literature sources (S2)
- Search Terms (S3)
- Detailed Methods (S4)
- Studies Excluded After Full Text Review (S5)
- Characteristics of Included Studies (S6)
- Level of Analysis in Included Studies (S7)
- Meta-regression Analyses (S8)
- Sub-group and Sensitivity Analyses (S9)
- Secondary Outcomes (S10)

Reporting guidelines


Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgements

We wish to thank the following for their helpful comments on the protocol and manuscript:
- The HTA Directorate of the Health Information and Quality Authority (HIQA) in Ireland
- Dr Sarah Condell and Dr Alan Smith from the Department of Health in Ireland
- Prof. Philip Nolan and the Steering Group for the National Review of Specialist Cardiac Services.

The abstract of this article has been previously published: https://doi.org/10.1016/j.jval.2019.09.824

References

27. Qian F, Zhong Y, Haninan EL: Relationship between operator and hospital volumes and short-term mortality for percutaneous coronary intervention


Open Peer Review

Are the rationale for, and objectives of, the Systematic Review clearly stated?

The rationale behind this systematic review is sound. However, with regard to objectives, I felt that these were not as clear. The primary outcome was mortality. I note that from the article, this primarily seems to be in hospital/30 day mortality with only a minority of studies reporting on mortality at 1 year. This is an important aspect of this analysis. In hospital/30 day mortality post elective PCI procedures is quite rare and generally estimated to be under 1%. In hospital/30 day mortality post STEMI treated with PPCI is generally higher (~5% depending on the source).

I think the question of whether In hospital/30 day mortality is affected by volume is perhaps more relevant for ACS or PPCI for STEMI rather than elective PCI.

The review would be more useful if the objectives were more clear. Or if, for example the authors specify in their methods the scenarios in which they are going to examine the impact of PCI volume (All PCI, PPCI, ACS, etc), what they are referring to by volume (hospital or operator) and then specify the time frame they are referring to for mortality (it appears to be in hospital/30 day mortality in this study but only one study showed 30 day mortality? It was slightly challenging for me to figure this out when reading the article).

The issue with these situations is that a patient may suffer a complication during a PPCI procedure or have delay in opening the infarct related artery, resulting in a significant infarct. They may survive initially but then die from heart failure some time later.

I also think that in many ways, operator volume is more relevant than hospital volume and perhaps this should be given primacy when considering the results.
So in summary, I think that the authors could perhaps be a bit more specific with regard to their objectives and define more exactly the scenarios in which they are examining this.

**Are sufficient details of the methods and analysis provided to allow replication by others?**

Yes, the authors have done an excellent job in detailing their methods and analysis which are clear.

**Is the statistical analysis and its interpretation appropriate?**

The authors are again to be congratulated for carrying out a high quality analysis. However, a limitation is the significant heterogeneity in the data.

Looking at their results in general I would say that overall they seem to indicate that volume does impact mortality. This is significant for operator volume and mortality. For hospital volume for PPCI. It is non significant for overall PPCI and hospital volume but the point estimate is indicating this also favours higher volume.

A significant issue here are the definitions of what is considered 'high volume' and this likely contributes to the heterogeneity of the data.

I note also that the secondary outcomes were generally in favour of higher volume operators and hospitals.

**Are the conclusions drawn adequately supported by the results presented in the review?**

Overall, the authors have a well balanced discussion and their results are expressed thoroughly and cautiously.

I think the topic of PCI complexity must also be taken into account when considering changes in patterns over time. For example, a complex CTO procedure may take several hours and carry a far higher risk of complications and mortality than 5 simple PCIs carried out during this time frame.

Higher volume centres may also take on more complex cases, which may mask their overall improved outcomes and it is important to make this clear to the reader who may not be familiar with the specifics of PCI. This may also explain the attenuation over time, as lower volume centres became less common and operators begin to take on more complex cases. For example, it is common for cardiologists to now perform multi vessel PCI for patients who are surgical turn downs for CABG. In the past, these patients may have been medically managed.

I think that the most important point made by the authors is that volume should not be seen as a marker of quality in isolation and they should emphasise this.

For patients with STEMI, organised systems of care and timely opening of the artery are likely more important for the patient than volume per se.

Overall though, I congratulate the authors for taking on this challenging topic and for the comprehensive nature of the review and meta-analysis they have performed.
Are the rationale for, and objectives of, the Systematic Review clearly stated?
Partly

Are sufficient details of the methods and analysis provided to allow replication by others?
Yes

Is the statistical analysis and its interpretation appropriate?
Partly

Are the conclusions drawn adequately supported by the results presented in the review?
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Interventional Cardiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 10 February 2021

https://doi.org/10.21956/hrbopenres.14352.r28791

© 2021 Fanaroff A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Alexander Fanaroff**
Cardiovascular Division, Hospital of the University of Pennsylvania, Philadelphia, PA, USA

In this systematic review and meta-analysis of observational studies examining the association between PCI operator and hospital volume and outcomes, Walsh et al. found a significant association between operator volume and short-term post-PCI mortality, but no significant association between hospital volume and short-term post-PCI mortality. This systematic review and meta-analysis was obviously a considerable undertaking, the care the authors took in completing a comprehensive and appropriate survey of the literature is apparent. The choices the authors made with respect to study inclusion are reasonable and justified well. The meta-analysis is reported very well, enough to allow replication of the authors methods, and the authors multiple sensitivity analyses and meta-regressions are well-done.

There have been several meta-analyses of conducted fairly recently (as the authors discuss in their introduction). Interestingly, the results of this meta-analysis contradict the prior results; a 2016 meta-analysis showed a significant inverse association between hospital PCI volume and mortality, and a 2014 meta-analysis showed no significant association between operator PCI volume and mortality. Compared with these prior meta-analyses, the present one has several strengths. First, there have been several analyses of the association between operator volume and mortality, using
registry (rather than administrative) data sources that allow for appropriate adjustment for case mix published since 2014. Second, the authors limited their analysis to studies published since 2008 (with resultant limitation of data collection to PCIs performed after 2000). PCI is a rapidly evolving field, and studies of the association between hospital volume and outcomes performed prior to 2008 may not adequately represent the current state of the field.

The primary limitation of this meta-analysis is heterogeneity between included studies, which is captured in the authors' $I^2$ estimates and, as they note, reduces the certainty of the evidence they've gathered. In the case of the volume-outcome relationship for PCI, heterogeneity would seem to come, a priori, from multiple sources: 1) volume thresholds selected; 2) completeness of adjustment for case mix; 3) regional effects. The authors appropriately did subgroup analyses to look at these subgroups. Notably, they did find a significant difference in their estimate of the association between hospital volume and mortality, with studies with complete case-mix adjustment showing an inverse relationship between hospital PCI volume and mortality, and studies with incomplete case-mix adjustment showing no difference. High and low volume centers take care of different types of patients (in the U.S., for example, high volume centers do more complex PCI and fewer STEMIs, as a proportion of total PCI) and incomplete case mix adjustment (for clinical presentation, but not PCI complexity, for example), may mask true differences in risk-adjusted outcomes between high- and low-volume centers. The likelihood of bias from incomplete case mix adjustment is substantial enough that the authors should consider presenting this data in the main manuscript.

One of the challenges of the meta-analytic framework is that it forces the authors to consider every study equally without regard to quality or design. In the case of studies of this type, whether a study uses rich clinical data or administrative data is likely to have a large effect on the completeness of case adjustment, which will have a large effect on any observed outcome differences between operator categories, as case mix differs substantially by operator and hospital volume categories. Incomplete case mix adjustment (for indication but not severity of illness or anatomic complexity) may lead to misleading results. The authors may consider a sensitivity analysis examining results from studies using registry versus administrative data, as those using registry data may afford a more complete case mix adjustment.

The authors' overall assessment of certainty of the evidence is "very low," largely due to heterogeneity. Looking at large, contemporary studies using detailed clinical/registry data, this does seem to be the case for the association between hospital volume and mortality. But looking only at large, highly contemporary analyses of the association between operator volume and outcomes (Inohara$^1$, Fanaroff$^2$, Hulme$^3$), there are fairly narrow and overlapping confidence intervals suggesting a small benefit with a reasonably high level of certainty. This is not consistent with the authors meta-analytic methods, but considering the varying quality of studies in this field, including studies without regard to quality (especially richness of data) may be misleading.

Overall, however, when dealing with such substantial heterogeneity, it is important to be circumspect in the conclusions that one draws, and the authors are appropriately circumspect.

References


**Are the rationale for, and objectives of, the Systematic Review clearly stated?**
Yes

**Are sufficient details of the methods and analysis provided to allow replication by others?**
Yes

**Is the statistical analysis and its interpretation appropriate?**
Yes

**Are the conclusions drawn adequately supported by the results presented in the review?**
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** cardiovascular quality and outcomes research

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.