Sunday 1st June, 2014

Dear Dr. Fang-Fang Ji, Dr. Anderson, Dr. Biondi-Zoccai and other editors other editorial team members and reviewers,

Revised Title: Effect of institutional volume on laparoscopic choleystectomy outcomes: systematic review and meta-analysis

Authors: Muireann Murray, Donagh A Healy, Khalid Bashar, Seamus McHugh, Mary Clarke Moloney, Stewart R Walsh.

Name of Journal: *World Journal of Meta-Analysis*

ESPS Manuscript NO: 10657

We thank the editors and reviewers for the time taken to review our manuscript. We think that this revision has improved the quality. Here is a response to the queries raised by the reviewers.

Reviewer 00742502

Point 1: The authors should give definitions of the high-volume and low-volume centres.

Response: We agree on the importance of including this. Definitions have been added to the introduction. It reads as follows:

"An expanding body of evidence suggests that outcomes in a variety of conditions are improved when patients are managed in high-volume centres or by high-volume healthcare providers [6]. High-volume centres dramatically improve the management of pancreatic cancer (≥20 cases per year), oesophageal cancer (≥30 cases per year), paediatric cardiac conditions (≥300 cases per year), unruptured abdominal aortic aneurysms (AAA) (≥36 cases per year) and acquired immune deficiency syndrome (AIDS) (≥100 cases per year) [6]. Similarly, high-volume surgeons or physicians dramatically improve the management of pancreatic cancer (10-42 cases per year), ruptured AAAs (≥10 cases per year), paediatric cardiac conditions (≥75 cases per year), colorectal cancer (≥22 cases per year), carotid endarterectomy (≥30 cases per year) and coronary artery bypass grafting (CABG) (≥150 cases per year) [6]. In contrast, no proven volume-outcome relationships exist for conditions such as diabetes, cystic fibrosis, rheumatoid arthritis, appendicitis and hernias [7,8].

Recently, data have emerged confirming that high-volume surgeons improve outcomes following LC (2, 4, 5, 9-12). Giger et al found improved results with surgeons who performed >100 LCs per year (5), Nuzzo et al found improved results with surgical teams who performed >450 LCs in three years (10), Csikesz et al found improved results with surgeons who performed >15 LCs per year (11) and McMahon et al found improved results for surgeons who had performed more than 200 cases (12). Andrews et al (2) and Hobbs et al (4) did not specify thresholds although they identified significantly reduced complications with increasing surgeon volume. However, it is unclear whether a volume-outcome relationship exists for LC at institutional level. If such an institutional relationship can be proven and understood, the creation of high-volume LC centres may become a priority. Therefore we performed a systematic review and meta-analysis focusing on institutional volume/outcome relationships for LC. The aim was to determine whether institutional LC

volume affects rates of mortality, conversion to open surgery, bile leakage and bile duct injury."

Point 2: Only one author performed the search and data extraction. I recommend that at least 2 authors are required to carry out these procedures to enhance accuracy.

Response: We accept this point. A second author has independently performed data extraction. However, due to resource constraints only one reviewer identified eligible studies. This remains a limitation although it is reported clearly.

The revised methods section reads as follows:

"One author (MM) identified eligible studies. Firstly, titles and abstracts were screened. Full-text manuscripts of potentially relevant studies were examined to finalise eligibility. Uncertainties regarding eligibility were discussed with a second author (DH). For each included study, the following data were extracted independently by two authors (MM & DH)): author, publication date, study design, the institution's name, start and finish dates, duration, number of LCs, number of mortalities, number of conversions to open surgery, number of bile leaks and the number of cases of BDI. Percentage complication rates were calculated for each outcome. Disagreements regarding data extraction were resolved by discussion with a third author (SRW). Data were entered into a computerised spreadsheet for analysis."

Reviewer 00535896

Point 1: The introduction and method section talk about high and low volume centres in laparoscopic cholecystectomy. A definition is needed about what is high and low volume.

Response: We accept that it is important to define high and low volume in relation to LC where mentioned in the introduction. The revised paragraph now reads:

"Recently, data have emerged confirming that high-volume surgeons improve outcomes following LC (2, 4, 5, 9-12). Giger et al found improved results with surgeons who performed >100 LCs per year (5), Nuzzo et al found improved results with surgical teams who performed >450 LCs in three years (10), Csikesz et al found improved results with surgeons who performed >15 LCs per year (11) and McMahon et al found improved results for surgeons who had performed more than 200 cases (12). Andrews et al (2) and Hobbs et al (4) did not specify thresholds although they identified significantly reduced complications with increasing surgeon volume. However, it is unclear whether a volume-outcome relationship exists for LC at institutional level. If such an institutional relationship can be proven and understood, the creation of high-volume LC centres may become a priority. Therefore we performed a systematic review and meta-analysis focusing on institutional volume/outcome relationships for LC. The aim was to determine whether institutional LC volume affects rates of mortality, conversion to open surgery, bile leakage and bile duct injury."

Point 2: An explanation is needed regarding the PRISMA guidelines.

Response: This explanation has been added. The text reads as follows:

"This systematic review was performed in accordance with the PRISMA guidelines (13). These guidelines are an evidence-based set of items that aim to enhance methodology and reporting clarity."

Point 3: A meta-analysis will be a great method for trying to define some thresholds. There is no exact science regarding the definition of thresholds, maybe a classification into 3 groups high-, low and mid- volume will be of sense. But after an analysis of so many studies an effort to define such threshold value is important.

Response: We agree with the reviewer that defining institutional volume thresholds to ensure optimal safety is an important. However, our review was not designed to generate specific cut-off points – our aim was to measure the effect of institutional volume on outcomes using a regression analysis. Our significant regression equations are reported and individual institutions may use them in order to estimate complication rates based upon projected institutional volume. We declined from providing some sample thresholds due to the limitations of the review (namely the absence of case-mix data, the unavailability of multivariable analysis, the temporal and geographic variations in included studies).

This information is conveyed in the final paragraph of the discussion:

"The chief strength of the current study relates to the inclusion of a large number of studies, including both small and large cohorts. Furthermore, we used an extensive search strategy and we focused on patient-important outcomes that are simply defined and easily diagnosed and are thus likely to be accurate even in retrospective studies. The external validity of the study is further enhanced by the finding of average complication rates that are quite similar to accepted published rates. The main limitation is the lack of data on case mix. Furthermore, as we included studies that spanned a twenty year period across all areas of the world, undoubtedly temporal and geographical variations in care would have existed. Finally we were limited to univariate analyses, thereby restricting conclusions on other factors that influence safety. We also wish to highlight that we did not aim to estimate specific optimal volume thresholds but rather we aimed to measure the effect of institutional volume on outcomes. Provided that the above limitations are accepted, individual institutions could use the significant regression equations to estimate conversion or bile leakage rates; however we would advise caution with such an approach until large scale registry data are available. Overall, we think that the results of our review are striking. We wish to encourage research on volume-outcome relationships in surgery, particularly through the use of large scale registries."

Point 4: The figures shows values from 0- 1200. I think valuation of more than 1000 cholecystectomies per year is required.

Response: We have chosen to use scatterplots. In this way, for each outcome, each included study is clearly represented on the graph. We think that this enhances clarity.

Point 5: In the Result part the authors says: "…56 cohorts (113526 patients) provided data on bile duct injury rates. Figure 5 displays the relationship between average annual number of LC procedures and institutional percentage bile duct injury rate. The linear regression

equation was non-significant (p=0.176). When only those studies that were published after 1995 were included (42 cohorts, 105570 patients) the regression equation was non-significant (p=0.248)……." It will be also useful to show/ formulate trends.

Response: We agree with the reviewer that it would be very useful and interesting to show/formulate trends in incidences of outcomes over time. Unfortunately, we were not able extract data according to defined time periods – we had to deal solely with the reported data of each individual study based upon reported start and finish dates. In order to show trends we think that institutional outcomes per consecutive defined time period would be needed and our study design does not allow this. This may be achievable with future large scale registries.

The final paragraph communicated this and reads as follows:

"The chief strength of the current study relates to the inclusion of a large number of studies, including both small and large cohorts. Furthermore, we used an extensive search strategy and we focused on patient-important outcomes that are simply defined and easily diagnosed and are thus likely to be accurate even in retrospective studies. The external validity of the study is further enhanced by the finding of average complication rates that are quite similar to accepted published rates. The main limitation is the lack of data on case mix. Furthermore, as we included studies that spanned a twenty year period across all areas of the world, undoubtedly temporal and geographical variations in care would have existed. Notably, we declined to evaluate trends in outcomes over time as study inclusion periods were heterogenous (table 1) and results were not provided by year but rather for entire study inclusion periods. Finally we were limited to univariate analyses, thereby restricting conclusions on other factors that influence safety. We also wish to highlight that we did not aim to estimate specific optimal volume thresholds but rather we aimed to measure the effect of institutional volume on outcomes using a regression analysis. Provided that the above limitations are accepted, individual institutions could use the significant regression equations to estimate conversion or bile leakage rates; however we would advise caution with such an approach until large scale registry data are available. Overall, we think that the results of our review are striking. We wish to encourage research on volume-outcome relationships in surgery, particularly through the use of large scale registries."

Reviewer 00504581

Point 1: I would like some commentary by the authors: The authors said "The external validity of the study is further enhanced by the finding of average complication rates that are quite similar to accepted published rates". Would it be possible to know a summary of this figures?

Response: We agree that it is important to include average complication rates for the outcomes that we evaluated. We thank the reviewer for noticing this omission.

We have added this information to the results section. We have reported number of studies, number of patients, number of events and average percentage complication rate for each of the outcomes:

".....43 studies (71305 patients) provided data on mortality (43 cases of mortality; average mortality was 0.06%). .....

58 studies (87840 patients) provided data on conversion rates (2835 cases of conversion; average conversion rate was 3.23%) . .....

44 studies (86025 patients) provided data on bile leak rates (381 cases of bile leakage; average bile leak rate was 0.44%). .....

56 cohorts (113526 patients) provided data on bile duct injury rates (316 cases of bile duct injury; average bile duct injury rate was 0.28%). ....."

We hope that the changes that we have made are satisfactory and that we may look forward to publication in *World Journal of Meta-Analysis*.

Yours sincerely,

Muireann Murray,

Donagh A Healy,

Khalid Bashar,

Seamus McHugh,

Mary Clarke Moloney,

Stewart R Walsh.