

Referee(s)' Comments to Author:

Referee: 1

Comments to the Author

The authors are to be congratulated on this study. It comes at a significant time, with the recent passing of the World Health Assembly Resolution on Essential and Emergency Care, the publication of the Lancet Commission on Global Surgery and the publication of the first volume of the 3rd Edition of Disease Control Priorities for Developing Countries, entirely dedicated to Surgery, all published earlier this year. All these should also be referenced in the manuscript. There is also a push to include Post Operative Mortality Ratios in the World Bank Development Indicators and how best to collect this data is a very pertinent question currently. This paper is therefore both timely and important.

The authors are also to be congratulated on managing to build such a wide and extensive network during this study. With an interest both from the Lancet and the World Bank in publishing data on surgery to look at trends over the next few years, such a network is essential to drive global improvement in surgical care across the world in the long term.

The paper is well written with good grammar.

The most striking criticism of the paper when reading it is the data pertaining to low income countries. The data is simply not typical of the majority of emergency laparotomies occurring in government district hospitals in low income countries. I'm impressed to see a senior anaesthetist was present at 794/1318 operations, for example, as qualified anaesthetists don't tend to exist outside of the big tertiary or NGO funded hospitals in many countries. I can only assume that the data has come from large well-funded NGO or private institutions in capital cities of these low income countries, often with large expatriate population of doctors – would that be correct? If so, would it be possible to have a table describing the characteristics of the hospitals involved and acknowledge this in more detail as a weakness of the study?

This weakness would potentially impact on the conclusion of the paper. The study nicely shows that mortality is related to use of a surgical safety checklist (although could this simply reflect the facilities and organisation of the hospital?). However, if the majority of people in rural areas of low income countries die at home with peritonitis as they are unable to access hospital care, or leave hospital after being faced with the costs of the operation, or are operated on by clinical or medical officers with no medical degree under ketamine anaesthesia with no muscle relaxant and no formal anaesthetist, no reliable oxygen or blood supply, poor analgesia, perioperative nursing care and no high dependency unit, then the conclusion for improving mortality may be to invest in infrastructure and training, as well as change financing strategies rather than improve patient safety alone. In addition, the 'real' post operative mortality rate is likely to be considerably higher which significant intra-country variation. This, of course, is not what this study has set up to explore, but needs to be acknowledged in considerably more detail in the discussion section with appropriate referencing.

Finally – a minor point – Table 1 – are the numbers in brackets meant to be %'s? If so, of what? 794/1318 is not 9.3% and the other numbers would also need revision.

Referee: 2

#### Comments to the Author

One of the biggest challenges in global surgery research and data analysis is that there is a severe dearth of reliable data. Some of the most quoted studies (initial GBD study saying 11% of the global disease burden, Weiser, et. al, the Lancet Commission), are studies with limited reliable data—mostly statistical modeling. This is, by definition, severely flawed data. Yet, it is the best we have. And it is damning to the global surgery community to highlight the need and the disparities. The danger we fall into is using this data and not having or advocating and developing on-the-ground data systems that can measure change, particularly at the local level. The authors of this study are to be commended for their ambitious and visionary collaborative and for a broad-reaching, but very realistic study. A few comments:

1) The time frame is short—2-week intervals. This is limited and may be subject to considerable bias in extrapolating from their results. However, their time frame is tangible and realistic in the world of limited data, and the authors adequately explain their rationale for using a short time frame to obtain more quality data.

--- I would like to see, however, a brief comment on some of the limitations (I do not think it adequately adjusts for seasonal variations, for example) and any future plans for longer studies.

2) The authors also very clearly state their objectives and primary outcome—specifically looking at the POMR vs 30-day mortality. I read a lot of research proposals in global surgery, and this study is one of the best organized and most realistic in its design that I have read. It is also highly applicable in terms of advocating for realistic markers/indices for global surgery to be included at a population level.

3) Statistical limitations are adequately described by the authors, such as limited ability to calculate adequate power analysis. As such, the results must be taken with a grain of salt. In evaluating the data, what limitations, etc, did the researchers find. I.E, does the experience suggest ways in which to create a more statistically robust study could be designed?

4) Mortality adjusted for case mix. Very different case mix and definition in different location—would like more information on how this is defined.

5) Regarding actual data collected: Relies heavily on certain data being routinely collected. Is ASA, for example, routinely collected? If not, was any training performed? Any attempts at sensitivity/internal validity?

6) Table 1 and Table 3 use percent by row. This is confusing. Particularly for Table 3, it is unclear what the denominator is, and thus difficult to extrapolate.

7) Fig 2—am surprised by the proportion of patients with appendicitis. This is a fairly western/high HDI diagnosis. What about typhoid perforation, etc? There should be some comment as to how this differs from current literature (mostly case report/retrospective analyses at district hospitals).

8) Additionally, the covariates used may also be most applicable to higher income settings. Was

any thought given to the use of other covariates that may be more applicable in LMIC. Examples might include HIV coinfection, assessment of nutrition status, BMI. (Fig 4). Any discussion of future studies to determine how best to assess case-mix, how to better risk-stratify.

9) Any assessment of time to operation/delay? This is discussed and is a major limitation of the study in terms of making comparisons. Is this data available or being included in future studies?

10) Appreciate the comment that the safety checklist itself is not necessarily a marker of improved outcomes but a proxy of safer systems.

11) The authors and study group are to be commended for the engagement and encouragement of local investigators. As they write in the discussion,

“Rather than department heads, junior clinicians were often the contact point by which a hospital became involved. Social media and technology played an important role in the recruitment and running of the study. Collaborators, particularly in LMICs, were clear in their view that those providing the clinical care can generate high quality data and lead international clinical research.”

The future of global surgical development and improvement in global surgical data and outcomes assessment lies in building the infrastructure and human resource capacity, particularly in LMIC. The authors are to be congratulated for the respect their manuscript and study shows to local providers and future surgical leaders.

12) I would like to see a section on future plans for this collaborative? Any additional studies planned from the data? What is the next step in utilizing the lessons learned from this study?

13) Typo line 3 introduction

14) Line 36 – would say “potentially” a proxy. This study is extraordinarily ambitious and well-done. However, there are many potential confounders in data quality and patient characteristics that cannot be accounted for. Using indices and proxies should not be a substitute for good safety, quality, and outcomes data that is locally applicable and relevant. Using broad proxies does risk threatening the quality of data collection, particularly in locations where data quality is currently limited.

15) What is "Variation in mortality by service?"

<\* Guidelines-for-revisions-2011.pdf>

<\* Author-statements.docx>