PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Linking World Bank Development Indicators and Outcomes of Congenital Heart Surgery in Low- and Middle-Income Countries: a Retrospective Analysis of Quality Improvement Data
AUTHORS	Rahman, Sarah; Zheleva, Bistra; Cherian, K.M.; Christenson, Jan; Doherty, Kaitlin; de Ferranti, David; Gauvreau, Kimberlee; Hickey, Patricia; Kumar, Raman; Kupiec, Jennifer; Novick, William; Sandoval, Nestor; Jenkins, Kathy

VERSION 1 - REVIEW

REVIEWER	Hao Zhang National Center for Cardiovascular Diseases and Fuwai Hospital, Beijing, China
REVIEW RETURNED	30-Dec-2018

GENERAL COMMENTS	The authors conducted a retrospective analysis using an international multicenter-database of congenital heart surgery outcomes and explored the relationship between country-level development/health indicators and risk-adjusted in-hospital surgical mortality of congenital heart disease in LMICs. The results were able to show that several development/health indicators are correlated with the surgical outcomes of CHD. The data presented in this study is interesting and provides important insights into health investment and outcomes of CHD in lower resource settings. Overall, this study is an important contribution to the literature because of high-quality datasets and well-performed analysis. I have some minor concerns regarding this manuscript. 1. Regarding the many tables and figures presented in this study, the description of findings in the results section is too brief and over-simplified. Could the authors provide a more detailed and summarizing presentation of their study findings? 2. The data of age distribution of CHD surgery by country and RACHS category by country are presented in numbers (Online Supplemental Table 1 and Table 2), which is not intuitive and straightforward for comparison between countries. Transforming the tables into stacked columns with different colors for the categorizing factor may have a better visual effect. 3. The authors adjusted for GDP per capita in the linear regression analysis and found that regression coefficients decrease in magnitude for most indicators, which led to the conclusion that

	GDP per capita is the main driver of the association between development indicators and surgical mortality. Why did the authors adjust for GDP in the first place? Is this adjustment a part of their original hypothesis or is it based on previous studies? Did they adjust for other indicators and found no significant changes in regression coefficients? I think the authors should describe the rationale in the methods section. 4. The authors found that low RACHS complexity and high RACHS complexity cases have different weights to the contribution of the association between SMR and development indicators. This finding provides important insight to their main conclusions, but the data is present in online supplement table 4. Could they remove this supplemental table into the main manuscript? 5. There are some grammatical issues. For example, in line 25 and 30 of page 9, “Graph showing” should be removed.

REVIEWER	Kriti Puri, Amy E Sanyahumbi Division of Pediatric Cardiology, Department of Pediatrics, Baylor College of Medicine, Texas Children's Hospital, Houston, TX, USA
REVIEW RETURNED	31-Dec-2018

GENERAL COMMENTS	This is an interesting study with a great analysis. It responds to a growing need for literature in the field of global pediatric cardiology. The analysis is thorough and the results are well-reported. The analysis suggests that World Bank indicators for economic development are associated with outcomes after congenital heart surgery. However, since statistical significance was not achieved for many of the analyses, while the data suggests many things, it is difficult to conclusively state a major finding. Hence it is difficult to state if one factor is more important than the other. Also, a major limitation is the bias of the study towards centers that agreed to participate in this collaborative effort, and a lack of an estimation of what percent of their country’s surgical volume is represented by these centers. Depending on whether the sampled population is representative or not, a national indictor is being used for an extremely limited population. This is potentially a major player in why the analysis may not be achieving significance. Finally, in addition to the interesting results, the study does find some actionable items for local health systems as well as for future research that may be discussed. With the association between general pediatric health indicators and surgical outcomes, it may help target interventions for heart surgery patients as well as for the general patient population in health systems. The trend towards an association of economic factors with outcomes suggests that it may be interesting to see if a single center analysis using socioeconomic data from individual patients also bears out these same results. This may be help centers in identifying higher risk populations. The study also generates some support for specialty training even in LMIC settings. A majority of the criticisms listed below are about the language of the discussion of the results which while very interesting, were limited by the lack of statistical significance and by the biased population included in this analysis. • Technical soundness of the work – retrospective study, robust database

	• Rigor of the analysis - standard analysis done • Clear use of English language - yes Title – since the study uses a congenital heart surgery hospital outcomes database, it may be more reasonable to title it as "…outcomes of congenital heart surgery in…" Page 2, Line 52 – the study reports center specific outcomes from certain countries. How many of the centers performing cardiac surgery from each country were included is not reported. So we could not call them country-level outcomes. Page 3, line 47 – It may be better to state that 'centers from Argentina' chose to be included in the QI consortium, as it was a voluntary center-level decision to participate or not, and not a national policy-level decision. Page 4, line 29-33 – again it would me more appropriate to omit the mention of 'country level' and 'countries choosing to participate', as this was a center-specific decision to participate in this collaborative, and in fact for a few of the countries the participating centers represent a minority of the centers. Page 8, line 16 – If the authors could additionally report the total numbers of surgeries from each country and the percentages out of the total of 24,917, it would be easier to get a sense of how many cases each country accounted for. Page 9, line 38 - If the authors could more explicitly state the cumulative n's for the low complexity and high complexity cases, it would help lend strength to the discussion about how modifiable risk factors for the high complexity surgeries are worth pursuing for LMICs, if they account for a reasonable percent of the case burden. Page 9, line 51 – Discussion – The authors may consider mentioning the statistically significant and those tending towards statistical significance first and then the associations later. While the r-coefficients of GDP and health expenditure per capita show a difference, since neither were statistically significant, it may be difficult to really make a speculation about which factor is more important. Page 10, line 24 – It may be reasonable to slightly temper the first line of this paragraph. None of the associations discussed in this paragraph were statistically significant, and it may be more appropriate to discuss the potential variation in SMRs between countries at similar developmental levels as an interesting finding suggested by the data. Page 10, line 31 – India, Pakistan, and Afghanistan have dramatically different GDPs, which is likely what explains the different SMRs, in addition to the burden of wars. Page 10, line 44 – It may be reasonable to include this ancillary analysis in the Methods section. Looking at the figures the readers may wonder if a separate analysis excluding Peru was done, and it would be impactful to mention that upfront. Page 10, line 49 – The first line of this paragraph may be an overstatement. The table discussed only 7 indicators, of which 3 seem to increase while 3 decrease after adjustment. The association of the health expenditure per capita seems to be intuitive in the unadjusted analysis, but after adjusting for GDP per capita, it is counter-intuitive (as to the plausibility of a positive correlation between health expenditure per capita and SMR?). It is difficult to say that GDP per capita is the 'main driver' of the associations in an analysis in which statistical significance was not achieved.

	Page 11, line 3 – the association of undernourishment with SMR is interesting and intuitive. Poorer nutritional status is known to be associated with surgical mortality. It further suggests that perhaps the factors known to be associated with general pediatric health in a country may also have an association with specialty surgery mortality rates and may give us targets for future improvement. Page 11, line 8 – It is difficult to talk about factors being 'independently' associated in an analysis in which statistical significance was not achieved. Page 11, line 26 - It is interesting that having a specialist workforce is associated with lower SMR in the high complexity cases. It may be a point of discussion that more specialist surgical training may help improve outcomes for the high complexity cases, and it may be a more directly modifiable factor (while the other associations discussed in this paper are not readily modified). It also lends more strength to policies to train specialists while trying to build up the generalist workforce in the LMICs. Page 12, line 17 – An additional facet of this limitation is that we don't know what percent of a country's congenital heart surgery volume was captured by the centers included in the study. So it is difficult to say how much or how little weight to put into an association, since we don't know how reflective these surgical results are of the whole country. However with the signals seen in this study, it may be a future direction to analyze by center and see if there is a stronger signal of zipcode or annual income of the families treated. Page 12, line 19 – this is not really a limitation of the study, since the study aim was to look at surgical mortality. This may be more appropriate for a future direction, to look at other measures of success in congenital heart surgery including longer term outcomes. Page 12, line 26 – It is difficult to know what is classified as 'large academic' centers in the countries included in this study. For instance, among the centers included from India, only a minority are affiliated with a university. It may be more reasonable to say that this included tertiary level centers with large referral base and academic intent towards quality improvement. Page 13, line 17 – This is also not a limitation of the study, but more of a future direction, since the study used a hospital surgical outcomes database. Page 13, line 28 – The variation being discussed is in countries with similar 'GDP per capita' not similar GDPs. Also, most of the discussion has focused on trend towards higher SMR in countries with lower World Bank development indices – if there were countries found to have lower SMRs in the setting of lower development indices, that weakens the trend of the results and discussion. Page 13, line 31-40 – This analysis itself actually gives very few actionable items for governments to act upon. It does however generate a signal for local centers to see if they can find an association with their patients' financial and nutritional well-being and their outcomes, and may identify a higher risk population for them to focus on. Perhaps the one actionable item for policy-level change may be that there is an important role of training specialists even in LMICs which are known to have a worse patient-provider ratio even for general providers, since a good proportion of their cardiac disease burden is medium-high complexity and may benefit from highly trained professionals.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1
Reviewer Name: Hao Zhang

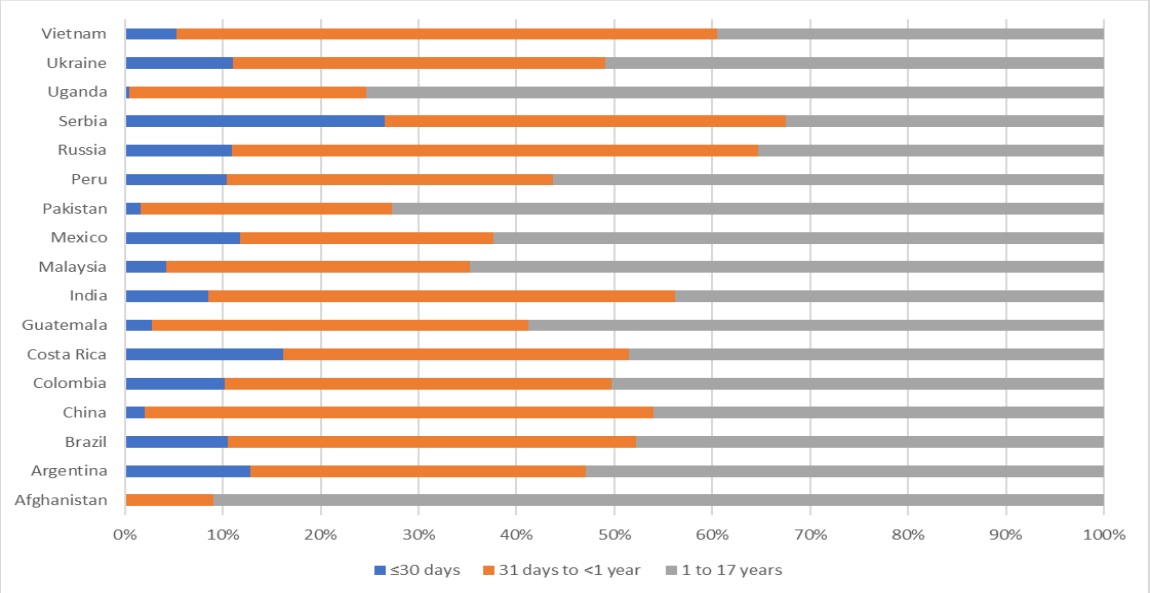
Institution and Country: National Center for Cardiovascular Diseases and Fuwai Hospital, Beijing, China

Please state any competing interests or state ‘None declared’: None declared

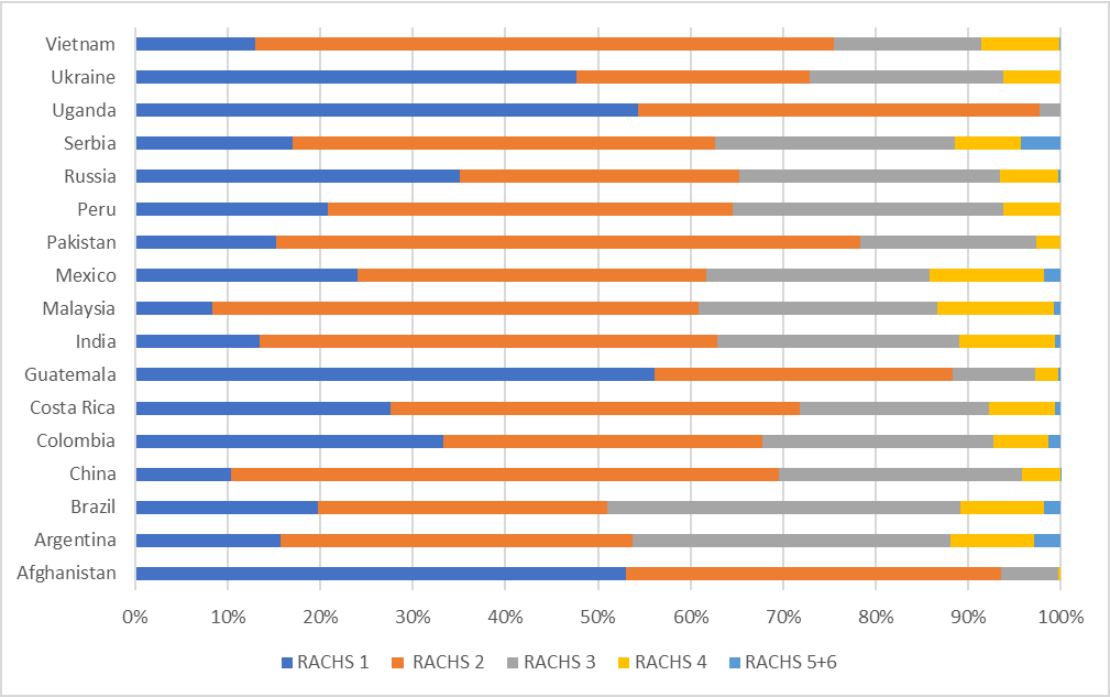
Please leave your comments for the authors below
The authors conducted a retrospective analysis using an international multicenter-database of congenital heart surgery outcomes and explored the relationship between country-level development/health indicators and risk-adjusted in-hospital surgical mortality of congenital heart disease in LMICs. The results were able to show that several development/health indicators are correlated with the surgical outcomes of CHD. The data presented in this study is interesting and provides important insights into health investment and outcomes of CHD in lower resource settings. Overall, this study is an important contribution to the literature because of high-quality datasets and well-performed analysis. I have some minor concerns regarding this manuscript.

1. Regarding the many tables and figures presented in this study, the description of findings in the results section is too brief and over-simplified. Could the authors provide a more detailed and summarizing presentation of their study findings? We added more detailed descriptions of the proportion of cases by country, the age, and RACHs category distributions. We added descriptions of some figures to point out trends and outliers in the findings.
2. The data of age distribution of CHD surgery by country and RACHS category by country are presented in numbers (Online Supplemental Table 1 and Table 2), which is not intuitive and straightforward for comparison between countries. Transforming the tables into stacked columns with different colors for the categorizing factor may have a better visual effect. We transformed the tables to stacked columns with the percentage of each age group or RACHS category as online supplemental Figures 2 and 3.

Online Supplemental Figure 2. Age Distribution of Cases by Country



Online Supplemental Figure 3. RACHS Category Distribution by Country



Online Supplemental Figure 2. This graph shows the age distribution of congenital heart surgery cases by country. Each row adds to 100%, and consists of each age category as a unique color.

Online Supplemental Figure 3. This graph shows the RACHS category distribution of congenital heart surgery cases by country. Each row adds to 100%, and consists of each RACHS category as a unique color.

3. The authors adjusted for GDP per capita in the linear regression analysis and found that regression coefficients decrease in magnitude for most indicators, which led to the conclusion that GDP per capita is the main driver of the association between development indicators and surgical mortality. Why did the authors adjust for GDP in the first place? Is this adjustment a part of their original hypothesis or is it based on previous studies? Did they adjust for other indicators and found no significant changes in regression coefficients? I think the authors should describe the rationale in the methods section. We adjusted for GDP per capita at the suggestion of one of our authors, Dr. De Ferranti, an expert on development, based on the concept that GDP per capita is a general indicator of resources available to a country that is widely used in literature. Therefore we only adjusted for GDP per capita, not any other indicators. We added a sentence with this reasoning in the manuscript and included a citation describing GDP per capita as a main indicator in the discussion of economic resources and social determinants of health. (Page 7, bottom of first paragraph in new manuscript)
4. The authors found that low RACHS complexity and high RACHS complexity cases have different weights to the contribution of the association between SMR and development indicators. This finding provides important insight to their main conclusions, but the data is present in online supplement table 4. Could they remove this supplemental table into the main manuscript? We moved the online supplemental table 4 to the main manuscript as Table 2.
5. There are some grammatical issues. For example, in line 25 and 30 of page 9, “Graph showing” should be removed. We removed “graph showing” and completed a grammar check of the manuscript.

Reviewer: 2
Reviewer Name: Kriti Puri, Amy E Sanyahumbi

Institution and Country: Division of Pediatric Cardiology, Department of Pediatrics, Baylor College of Medicine, Texas Children's Hospital, Houston, TX, USA

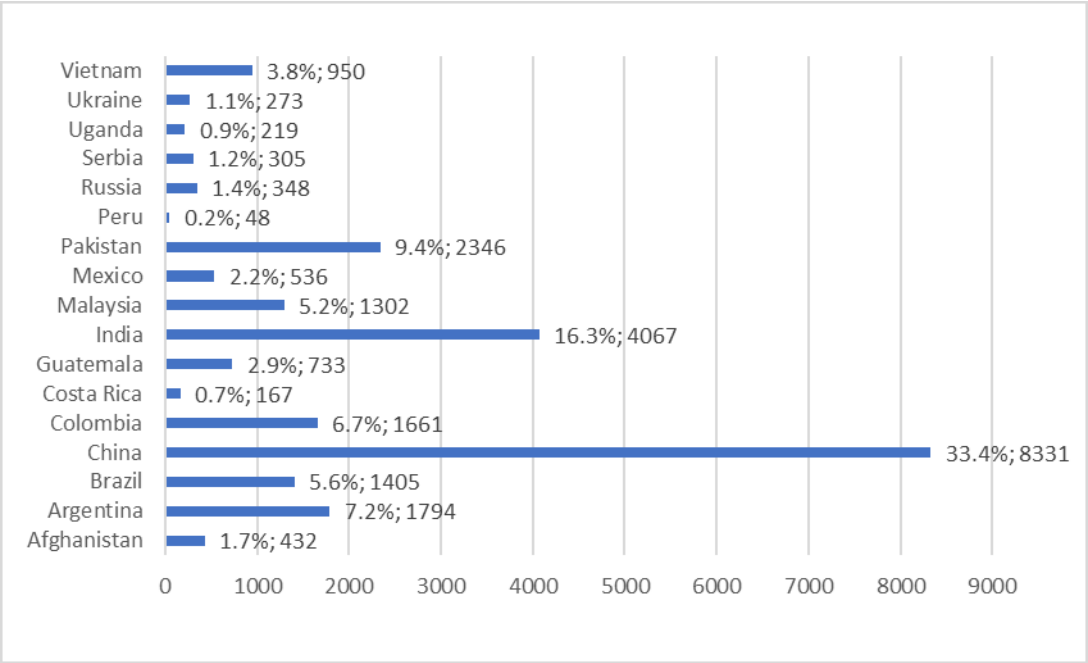
Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below
This is an interesting study with a great analysis. It responds to a growing need for literature in the field of global pediatric cardiology. The analysis is thorough and the results are well-reported. The analysis suggests that World Bank indicators for economic development are associated with outcomes after congenital heart surgery.
However, since statistical significance was not achieved for many of the analyses, while the data suggests many things, it is difficult to conclusively state a major finding. Hence it is difficult to state if one factor is more important than the other. While we did not achieve statistical significance, the overall results may show meaningful links between development and outcomes, and the regression analysis controlling for GDP per capita with all other indicators reveals that GDP per capita may be a particularly important indicator. We understand statistical significance is a major limitation of the study and describe this as a limitation. Per your comments, we have reduced statements that may be overstating the magnitude of our results or overstate one indicator over others from the correlations.
Also, a major limitation is the bias of the study towards centers that agreed to participate in this collaborative effort, and a lack of an estimation of what percent of their country’s surgical volume is represented by these centers. Depending on whether the sampled population is representative or not, a national indictor is being used for an extremely limited population. This is potentially a major player in why the analysis may not be achieving significance. We agree that the potential lack of representation of the nation by the centers is a major limitation, and describe this as our second limitation on the bottom of page 14 of the new manuscript.
Finally, in addition to the interesting results, the study does find some actionable items for local health systems as well as for future research that may be discussed. With the association between general pediatric health indicators and surgical outcomes, it may help target interventions for heart surgery patients as well as for the general patient population in health systems. The trend towards an association of economic factors with outcomes suggests that it may be interesting to see if a single center analysis using socioeconomic data from individual patients also bears out these same results. This may be help centers in identifying higher risk populations. The study also generates some support for specialty training even in LMIC settings. We added a summary of policy recommendations in the conclusion on page 16 of the new manuscript, and included that the trend can be investigated by single centers using socioeconomic data of individual patients to see if a similar link arises, potentially helping to identify higher risk populations.
A majority of the criticisms listed below are about the language of the discussion of the results which while very interesting, were limited by the lack of statistical significance and by the biased population included in this analysis.
• Technical soundness of the work – retrospective study, robust database
• Rigor of the analysis - standard analysis done
• Clear use of English language - yes

Title – since the study uses a congenital heart surgery hospital outcomes database, it may be more reasonable to title it as “…outcomes of congenital heart surgery in…” We edited the title to reflect this change, reflecting our use of surgical outcomes data.
Page 2, Line 52 – the study reports center specific outcomes from certain countries. How many of the centers performing cardiac surgery from each country were included is not reported. So we could not

call them country-level outcomes. We revised this line to describe that the study may miss regional variability, without referring to outcomes as country-level outcomes. However, since we calculate SMRs by country (multiple centers in one country are analyzed together as one SMR rather than separately) to compare with national development indicators, we occasionally refer to them as country SMRs in the results. We cannot find a better phrase to refer to these SMRs and welcome suggestions.
Page 3, line 47 – It may be better to state that 'centers from Argentina' chose to be included in the QI consortium, as it was a voluntary center-level decision to participate or not, and not a national policy-level decision. We have changed the language to 'centers from Argentina'.
Page 4, line 29-33 – again it would me more appropriate to omit the mention of 'country level' and 'countries choosing to participate', as this was a center-specific decision to participate in this collaborative, and in fact for a few of the countries the participating centers represent a minority of the centers. We removed the 'country-level' language in these lines and refer instead to centers in the countries.
Page 8, line 16 – If the authors could additionally report the total numbers of surgeries from each country and the percentages out of the total of 24,917, it would be easier to get a sense of how many cases each country accounted for. We added Online Supplementary Figure 1, reporting the total number of surgery cases per country along with the percentages out of the total.

Online Supplementary Figure 1. Congenital Heart Surgery Cases by Country



Online Supplemental Figure 1. This bar graph shows the number of congenital heart surgery cases in each country. The label lists the number of cases and the percentage of the total cases for each country.

Page 9, line 38 - If the authors could more explicitly state the cumulative n's for the low complexity and high complexity cases, it would help lend strength to the discussion about how modifiable risk factors for the high complexity surgeries are worth pursuing for LMICs, if they account for a reasonable percent of the case burden. We added a line on page 10 in new manuscript stating the cumulative n's for low and high complexity, "The low RACHS categories analysis includes a total of 16,870 cases, and the high RACHS categories analysis includes a total of 7,997 cases."
Page 9, line 51 – Discussion – The authors may consider mentioning the statistically significant and

those tending towards statistical significance first and then the associations later. While the r-coefficients of GDP and health expenditure per capita show a difference, since neither were statistically significant, it may be difficult to really make a speculation about which factor is more important. We restructured the discussion for clarity to first include the statistically significant under-5 mortality correlation, then the borderline significant specialist surgical workforce correlation, then the remaining non-significant associations. We understand that we may have over speculated, so we removed the statement that GDP per capita could be more important. Page 11-12 in new manuscript.
Page 10, line 24 – It may be reasonable to slightly temper the first line of this paragraph. None of the associations discussed in this paragraph were statistically significant, and it may be more appropriate to discuss the potential variation in SMRs between countries at similar developmental levels as an interesting finding suggested by the data. We rephrased the line to "Although it is difficult to be conclusive, an interesting finding of this study is that there is still substantial variation between countries in SMRs even at similar development levels". Bottom of page 12 in new manuscript
Page 10, line 31 – India, Pakistan, and Afghanistan have dramatically different GDPs, which is likely what explains the different SMRs, in addition to the burden of wars. We added a line to describe the different GDPs in the countries: "Also, while Afghanistan, Pakistan, and India have similar GDP per capita levels, they have very different overall GDP levels, which may access to surgery for patients and impact congenital heart surgery SMR variation." Bottom of page 12 in new manuscript
Page 10, line 44 – It may be reasonable to include this ancillary analysis in the Methods section. Looking at the figures the readers may wonder if a separate analysis excluding Peru was done, and it would be impactful to mention that upfront. We added a line in the methods section on page 7 in new manuscript: "An ancillary analysis without Peru was also completed after it appeared to be an outlier in figures generated."
Page 10, line 49 – The first line of this paragraph may be an overstatement. The table discussed only 7 indicators, of which 3 seem to increase while 3 decrease after adjustment. The association of the health expenditure per capita seems to be intuitive in the unadjusted analysis, but after adjusting for GDP per capita, it is counter-intuitive (as to the plausibility of a positive correlation between health expenditure per capita and SMR?). It is difficult to say that GDP per capita is the 'main driver' of the associations in an analysis in which statistical significance was not achieved. We tempered the line by replacing 'most' with 'some' indicators. We no longer describe GDP per capita as a 'main' driver. (Page 13 middle paragraph in new manuscript.)
We agree that the health expenditure adjusted positive correlation is counter-intuitive. It remains unexplained, but it is close to 0. We removed health expenditure from the list of indicators that have decreased in magnitude correlations after adjusting for GDP per capita.

Page 11, line 3 – the association of undernourishment with SMR is interesting and intuitive. Poorer nutritional status is known to be associated with surgical mortality. It further suggests that perhaps the factors known to be associated with general pediatric health in a country may also have an association with specialty surgery mortality rates and may give us targets for future improvement. The link between undernourishment and SMR is certainly and important finding in this manuscript and we have added a line about the connection to general pediatric health: "This also suggests that factors that correlate with general pediatric health, such as undernourishment, impact surgical mortality, offering a target for further improvement of health and surgical outcomes."
Page 11, line 8 – It is difficult to talk about factors being 'independently' associated in an analysis in which statistical significance was not achieved. We removed 'independently' and replaced the description as "suggesting that population-wide undernourishment affects mortality after congenital heart surgery 'outside the effect' of GDP".
Page 11, line 26 - It is interesting that having a specialist workforce is associated with lower SMR in the high complexity cases. It may be a point of discussion that more specialist surgical training may help improve outcomes for the high complexity cases, and it may be a more directly modifiable factor (while the other associations discussed in this paper are not readily modified). It also lends more strength to policies to train specialists while trying to build up the generalist workforce in the LMICs.
Definitely. We added a line to emphasize this, "This potentially suggests that increasing training for a

specialist surgery workforce is necessary in LMICs." (page 13 second paragraph in new manuscript)
Page 12, line 17 – An additional facet of this limitation is that we don't know what percent of a country's congenital heart surgery volume was captured by the centers included in the study. So it is difficult to say how much or how little weight to put into an association, since we don't know how reflective these surgical results are of the whole country. However with the signals seen in this study, it may be a future direction to analyze by center and see if there is a stronger signal of zipcode or annual income of the families treated. We added this concept of the limitation to the manuscript, page 14 second paragraph in new manuscript: "..and vary in the proportion of country cases that they represent, making firm conclusions difficult, but suggesting that country-specific analyses about patients at different poverty levels may be warranted."
Page 12, line 19 – this is not really a limitation of the study, since the study aim was to look at surgical mortality. This may be more appropriate for a future direction, to look at other measures of success in congenital heart surgery including longer term outcomes. We removed this from limitations and added it to future directions, in the final paragraph on page 16 in new manuscript.
Page 12, line 26 – It is difficult to know what is classified as 'large academic' centers in the countries included in this study. For instance, among the centers included from India, only a minority are affiliated with a university. It may be more reasonable to say that this included tertiary level centers with large referral base and academic intent towards quality improvement. We rephrased this to "IQIC sites are often tertiary level hospitals with large referral bases and academic intent towards quality improvement; IQIC members are willing and able to collect data and participate in quality improvement activities." (bottom of page 14, new manuscript).
Page 13, line 17 – This is also not a limitation of the study, but more of a future direction, since the study used a hospital surgical outcomes database. We moved this to future directions, final paragraph of page 16 in new manuscript.

Page 13, line 28 – The variation being discussed is in countries with similar 'GDP per capita' not similar GDPs. Also, most of the discussion has focused on trend towards higher SMR in countries with lower World Bank development indices – if there were countries found to have lower SMRs in the setting of lower development indices, that weakens the trend of the results and discussion. We replaced GDP with 'GDP per capita' (bottom of page 15, new manuscript). We document a trend between low development and high SMR, but some notable exceptions of low development and low SMR are important findings countering the trend that offer hope for investment in very low resource programs.

Page 13, line 31-40 – This analysis itself actually gives very few actionable items for governments to act upon. It does however generate a signal for local centers to see if they can find an association with their patients' financial and nutritional well-being and their outcomes, and may identify a higher risk population for them to focus on. Perhaps the one actionable item for policy-level change may be that there is an important role of training specialists even in LMICs which are known to have a worse patient-provider ratio even for general providers, since a good proportion of their cardiac disease burden is medium-high complexity and may benefit from highly trained professionals. We added a summary of actionable policy level changes in the conclusion, page 16 in new manuscript.

VERSION 2 – REVIEW

REVIEWER	Zhanhao Su, Hao Zhang National Center for Cardiovascular Diseases and Fuwai Hospital, Beijing, China
REVIEW RETURNED	17-Feb-2019

GENERAL COMMENTS	The authors have appropriately responded to our concerns. We have no further comments.

REVIEWER	Kriti Puri Baylor College of Medicine, Texas Children's Hospital, Houston, TX, USA
REVIEW RETURNED	20-Feb-2019

GENERAL COMMENTS	The authors have very thoroughly addressed the queries and recommendations. The title is more appropriate, the discussion has been modified, and the additional figures are very helpful, overall rendering a more complete picture of the results. The analysis suggests that World Bank indicators for economic development are associated with outcomes after congenital heart surgery. Although only a minority of the total annual CHD surgical cases from each country are included in the database, it generates interesting data for further study on a local and national level. • Technical soundness of the work – retrospective study, novel database • Rigor of the analysis - standard analysis done, adequate additions made in response to reviewer comments • Clear use of English language - yes