OBSTETRICS & GYNECOLOGY

NOTICE: This document contains correspondence generated during peer review and subsequent revisions but before transmittal to production for composition and copyediting:

- Comments from the reviewers and editors (email to author requesting revisions)
- Response from the author (cover letter submitted with revised manuscript)*

**The corresponding author has opted to make this information publicly available.*

Personal or nonessential information may be redacted at the editor's discretion.

Questions about these materials may be directed to the *Obstetrics & Gynecology* editorial office: obgyn@greenjournal.org.

Date: May 07, 2021
To: "Jennifer A Callaghan-Koru" [redacted]
From: "The Green Journal" em@greenjournal.org
Subject: Your Submission ONG-21-510

RE: Manuscript Number ONG-21-510

Reduction in cesarean delivery rates following a state quality collaborative in Maryland

Dear Dr. Callaghan-Koru:

Your manuscript has been reviewed by the Editorial Board and by special expert referees. Although it is judged not acceptable for publication in Obstetrics & Gynecology in its present form, we would be willing to give further consideration to a revised version.

If you wish to consider revising your manuscript, you will first need to study carefully the enclosed reports submitted by the referees and editors. Each point raised requires a response, by either revising your manuscript or making a clear and convincing argument as to why no revision is needed. To facilitate our review, we prefer that the cover letter include the comments made by the reviewers and the editor followed by your response. The revised manuscript should indicate the position of all changes made. We suggest that you use the "track changes" feature in your word processing software to do so (rather than strikethrough or underline formatting).

Your paper will be maintained in active status for 21 days from the date of this letter. If we have not heard from you by May 28, 2021, we will assume you wish to withdraw the manuscript from further consideration.

REVIEWER COMMENTS:

Reviewer #1: This is a report on the reduction of cesarean deliveries following the implementation of a statewide perinatal quality improvement collaborative. The QI intervention utilized a patient safety consensus bundle that was developed through a national partnership. The lead author had previously reported on the implementation of this bundle at the 1 year mark. This study is a pre-post evaluation that makes use of web-based survey administered at 12 months and then at 30 months after the start of the initiative to assess stage of implementation. Secondary data sources were used to calculate birth rates, data from the QI collaborative's data portal was used to assess provider training.
Implementation was variable and some hospitals had increases in C-section delivery rates. However modest reductions in overall C-section delivery rates were noted with larger reductions in C-section deliveries in those nulliparous term singleton vertex deliveries via induction. Reductions were driven by changes in high volume hospitals. This study also implies that more support of staff education and of integrating doulas into the care teams is in order.

1. Line 70: 30-months versus 30-month?

2. Line 153; Consider adding (AIM).

3. Lines 162-163: Consider elaborating on "movement in labor" as it's doesn't seem to be defined in the appendix (maybe describe as 'maternal physical movement in labor?').

4. Line 185: What does "teaching states" mean?

5. Line 203: Consider "..and among NTSV births.."

6. Lines 255-259: As per Figure 3 it seems that the steepest drop in C-section rates occurred between Q1-Q2 2018 and Q1-Q2 2019. Any thoughts on why?

7. Lines 300-301: Adopting more practices was only associated with lower c-section rates for the Response domain. It may be useful to point out here that most of the implementation of the Response bundle occurred prior to the Collaborative (Table 2).

Reviewer #2: General Comments:
Although, the authors do state in the beginning of the discussion section that this study is the second evaluation of the State's perinatal quality improvement collaborative to reduce NTSV cesarean delivery rates, this manuscript does not provide enough detail in the current study design to stand alone in making sense to readers that have not read the manuscript for the first evaluation of the collaborative. In addition, after pulling the original manuscript, there are a number of changes in the current reporting (e.g. respondents, responses, hospitals included, etc.) that needs some explanation in this manuscript.

Specific Comments:
Page 9, Lines 175 - 180: was the original study design to be collect reported data through web-based surveys in Qualtrics at months 12 and 30 months after the collaborative was started. To be clear, is this study a follow-up with an additional survey for expanded analysis? Were the respondents predetermined (e.g. at time of study data collection, following study)? Could the data be affected by who responded? How many attempts were made to get surveys from these participants and what sort of time frame was given for the response? What was your overall response rate?

Page 11, Lines 220-221: hospitals in collaborative included 7 Level 1…, 8 Level II… and 16 Level II/IV hospitals. Why are the numbers different in Table 1 and Figures 4 and 5?

Page 23, Table 1: Were there differences in ability to implement collaboratives/c-section rates in hospitals with higher proportions of Medicaid patients as compared to non-Medicaid patients?

Pages 29-30, Figures 4 and 5: The figures need better clarification in the footnote so that interpretation of these figures can stand alone. It is not self-evident what ABCDEFGHIJK…. in the first column represents (my assumption is that they represent the individual hospitals, although in the previous manuscript the hospitals were labeled H1, H2, etc.). Why are the level of care hospitals ordered differently in Figure 5? Why are there only 10 hospitals listed in Level II versus 11 listed in table 1?

Reviewer #3: General Comments: This is a high-quality study examining the impact of a state-wide initiative to reduce NTSV cesarean section (cs) rates in Maryland. This study's findings dovetail with the recent findings from Rosenstein et al in JAMA (see below). A large volume of information is presented in the paper, which includes survey data, rates of hospital adoption of bundle elements, and changes in NTSV cs rates over time. It seems the authors focus most of their attention on the changes in NTSV rates over time (ie primary aims). Therefore, I would recommend packaging the survey data on domain uptake as secondary aims/outcomes.

Specific Comments:
L113 - The Kozhimannil study only reported variation in cesarean rates. It may be an overreach to state that standardizing care can impact cesarean rates based only from this study's finding. At this point, you may want to cite the recent Rosenstein JAMA paper as evidence of how a multifaceted state-wide approach may result in a reduction in the NTSV cesarean rate in California over time (PMID: 33904868). I appreciate that this study came out whilst your paper was under peer review.
L143/145/214/Table 1: Any information on the team member or representatives sent from each hospital - e.g. clinical or nursing lead or QI representative? This may speak to whether the member or representative can be considered a local 'champion' of bundle implementation and its potential impact. Also, were these representatives the same as those who completed the endline survey (Table 1)?
L176: Can you point readers to Appendix 3 which describes each of these bundle practices? Else, readers have to dig to find them.
L220: When describing Level I-III hospitals, are you referring to ACOG Levels of Maternal Care? If so, this needs describing in the methods &/or appendix.
L228/Table 2: I'm unsure about the clinical value of summarizing data from the domains - currently presented in Table 2. Also, you present means but, from a practical standpoint, it is not possible to implement a fraction of a bundle element e.g., 0.8 for R1. I would suggest dropping table 2 and simply commenting on the findings from Appendix 3 which provides more granular and relevant information about the degree of implementation of each practice across all the hospitals. Consider highlighting practices that had the highest rates of 'not starting' and fully implemented. For example, R1.2, R3.6, and R.4.3 had high rates of not being started. The high rate (74%) for not integrating doulas stands out.
Figures 1 and 2: Clarify the legend details in the text and figures. What does 20-39%, 40-59%, etc..refer to? Do these data refer to the proportion of all practices in the bundle completed? For example, does 20-39% refer to 5-10 of the 23 practices? Also, I wonder whether a bar chart would be a better visual approach for conveying changes in coverage over time.
Figure 3: As you have data for CS rates by quarter, consider presenting quarterly as opposed to semi-annual rates. More granular information would be valuable in assessing changes over time than semi-annual rates. Regression modeling would help in better examining the degree of change over time as opposed to simply doing chi-squared tests comparing the 1st quarter to the last.
Figures 4 and 5: Could you reorganize the chart to rank hospitals from those with the greatest reduction in NTSV cesarean rates at one end of the Y axis and hospitals with the greatest increase in rates at the other end (ie a caterpillar plot). The

data could be presented as % change or absolute % increase/decrease as opposed to dots with arrows to depict differences that are harder to interpret.
Table 3: Instead of a 'significance' column - I'd recommend calculating RRs comparing hospitals above the median to those below (reference group) and presenting absolute risk differences also. Also, instead of medians (arbitrary dichotomization), consider using tertiles (or other approaches) to obtain more detailed comparative data on the effect of domain uptake on NTSV CS rates. Please also consider regression modeling comparing hospitals by the level of maternal care. Might there be a way to model the use of practice elements (independent variable) against NTSV cs rates (dependent variable) in a mixed-effects model with hospitals as the random effect? This would provide some insight into the potential effect of each practice element on the NTSV rate. If not, highlight as a limitation, with a call for more studies to examine this.
L268 - I didn't see the statistical analysis for how you compared cs rates across Levels of hospitals.
L331 - can you discuss or comment on what barriers or reasons there might be for implementing and affecting QI? I'm unclear what 'quality improvement capacity and motivation' and 'engaging clinicians' mean in this context. Some specific examples would be helpful.
L348 - A key limitation in your study and that of Rosenstein et al in JAMA is that there are no maternal morbidity data to determine whether reducing NTSV CS rates is associated with less, more, or the same risk of maternal morbidity. This is a problem if there is greater pressure on obstetric providers to deliver patients vaginally instead of by cs. Is there an unintended consequence of this? Highlight the fact that we need more data to examine changes in specific morbidities (not just SMM) e.g., infections and PPHs from longer labor times before a vaginal delivery or even a cesarean, especially in women who have a prolonged and/or complicated labor. Although Rosenstein et al reported no change in the rate of neonatal morbidity (as a composite), this finding needs to be confirmed in other populations also. The ultimate goal should be a reduction in NSTV cesarean rates coupled with no increases in maternal and perinatal morbidity. The pressure on obstetric providers to weigh the risk of maternal/perinatal morbidity in a woman with abnormal labor against the subjective pressure to perform a vaginal delivery and 'hitting their NSTV cs metric' should not be underappreciated.

STATISTICAL EDITOR COMMENTS:

The Statistical Editor makes the following points that need to be addressed:

Table 1: Since the n = 31, should round the %s to nearest integer %, not cite to 0.1% precision. Given the sample size, should cite the years experience, annual number of births and proportion of pts covered by Medicaid as median (Range).

Table 2: It is not clear what the Authors meant by "Average". ? mean. Given the ranges and counts, a better metric would be to summarize as median(range) for each time period. Should also enumerate all missing data.

Table 3: Again, the issue of precision for number of hospitals, which should be rounded to nearest integer %, not to 0.1 % precision. For the cesarean rates, should round to nearest 0.1%, not to 0.01% precision.

Figs 4 and 5 should be in supplemental material.

lines 81-83: Should clarify that these % changes are relative, not absolute changes in proportion cesarean delivery rates.

Editor: Please throughout emphasize absolute as opposed to relative risk reductions.

EDITOR COMMENTS:

1. The Editors of Obstetrics & Gynecology are seeking to increase transparency around its peer-review process, in line with efforts to do so in international biomedical peer review publishing. If your article is accepted, we will be posting this revision letter as supplemental digital content to the published article online. Additionally, unless you choose to opt out, we will also be including your point-by-point response to the revision letter. If you opt out of including your response, only the revision letter will be posted. Please reply to this letter with one of two responses:
A. OPT-IN: Yes, please publish my point-by-point response letter.
B. OPT-OUT: No, please do not publish my point-by-point response letter.

2. Obstetrics & Gynecology uses an "electronic Copyright Transfer Agreement" (eCTA). Please check with your coauthors to confirm that the disclosures listed in their eCTA forms are correctly disclosed on the manuscript's title page. Each of your coauthors received an email from the system, titled "Please verify your authorship for a submission to Obstetrics & Gynecology." Each author should complete the eCTA if they have no yet done so.

The following authors need to complete the form:

Ann B Burke (burkean@holycrosshealth.org)
Geoffrey Curran (Geoffrey Curran)

3. If your study is based on data obtained from the National Center for Health Statistics, please review the Data Use Agreement (DUA) for Vital Statistics Data Files that you or one of your coauthors signed. If your manuscript is accepted for publication and it is subsequently found to have violated any of the terms of the DUA, the journal will retract your article. The National Center for Health Statistics may also terminate your access to any future vital statistics data.

4. Standard obstetric and gynecology data definitions have been developed through the reVITALize initiative, which was convened by the American College of Obstetricians and Gynecologists and the members of the Women's Health Registry Alliance. Obstetrics & Gynecology has adopted the use of the reVITALize definitions. Please access the obstetric data definitions at https://www.acog.org/practice-management/health-it-and-clinical-informatics/revitalize-obstetrics-data-definitions and the gynecology data definitions at https://www.acog.org/practice-management/health-it-and-clinical-informatics/revitalize-gynecology-data-definitions. If use of the reVITALize definitions is problematic, please discuss this in your point-by-point response to this letter.

5. Because of space limitations, it is important that your revised manuscript adhere to the following length restrictions by manuscript type: Original Research reports should not exceed 5,500 words. Stated word limits include the title page, précis, abstract, text, tables, boxes, and figure legends, but exclude references.

6. Specific rules govern the use of acknowledgments in the journal. Please note the following guidelines:

* All financial support of the study must be acknowledged.
* Any and all manuscript preparation assistance, including but not limited to topic development, data collection, analysis, writing, or editorial assistance, must be disclosed in the acknowledgments. Such acknowledgments must identify the entities that provided and paid for this assistance, whether directly or indirectly.
* All persons who contributed to the work reported in the manuscript, but not sufficiently to be authors, must be acknowledged. Written permission must be obtained from all individuals named in the acknowledgments, as readers may infer their endorsement of the data and conclusions. Please note that your response in the journal's electronic author form verifies that permission has been obtained from all named persons.
* If all or part of the paper was presented at the Annual Clinical and Scientific Meeting of the American College of Obstetricians and Gynecologists or at any other organizational meeting, that presentation should be noted (include the exact dates and location of the meeting).
* If your manuscript was uploaded to a preprint server prior to submitting your manuscript to Obstetrics & Gynecology, add the following statement to your title page: "Before submission to Obstetrics & Gynecology, this article was posted to a preprint server at: [URL]."

7. The most common deficiency in revised manuscripts involves the abstract. Be sure there are no inconsistencies between the Abstract and the manuscript, and that the Abstract has a clear conclusion statement based on the results found in the paper. Make sure that the abstract does not contain information that does not appear in the body text. If you submit a revision, please check the abstract carefully.

In addition, the abstract length should follow journal guidelines. The word limit for Original Research articles is 300 words. Please provide a word count.

8. ACOG is moving toward discontinuing the use of "provider." Please replace "provider" throughout your paper with either a specific term that defines the group to which are referring (for example, "physicians," "nurses," etc.), or use "health care professional" if a specific term is not applicable.

9. In your Abstract, manuscript Results sections, and tables, the preferred citation should be in terms of an effect size, such as odds ratio or relative risk or the mean difference of a variable between two groups, expressed with appropriate confidence intervals. When such syntax is used, the P value has only secondary importance and often can be omitted or noted as footnotes in a Table format. Putting the results in the form of an effect size makes the result of the statistical test more clinically relevant and gives better context than citing P values alone.

Please standardize the presentation of your data throughout the manuscript submission. For P values, do not exceed three decimal places (for example, "P = .001"). For percentages, do not exceed one decimal place (for example, 11.1%").

10. Please review the journal's Table Checklist to make sure that your tables conform to journal style. The Table Checklist is available online here: http://edmgr.ovid.com/ong/accounts/table_checklist.pdf.

11. Please review examples of our current reference style at http://ong.editorialmanager.com (click on the Home button in the Menu bar and then "Reference Formatting Instructions" document under "Files and Resources). Include the digital object identifier (DOI) with any journal article references and an accessed date with website references. Unpublished data, in-press items, personal communications, letters to the editor, theses, package inserts, submissions, meeting presentations, and abstracts may be included in the text but not in the reference list.

In addition, the American College of Obstetricians and Gynecologists' (ACOG) documents are frequently updated. These documents may be withdrawn and replaced with newer, revised versions. If you cite ACOG documents in your manuscript, be sure the reference you are citing is still current and available. If the reference you are citing has been updated (ie,

replaced by a newer version), please ensure that the new version supports whatever statement you are making in your manuscript and then update your reference list accordingly (exceptions could include manuscripts that address items of historical interest). If the reference you are citing has been withdrawn with no clear replacement, please contact the editorial office for assistance (obgyn@greenjournal.org). In most cases, if an ACOG document has been withdrawn, it should not be referenced in your manuscript (exceptions could include manuscripts that address items of historical interest). All ACOG documents (eg, Committee Opinions and Practice Bulletins) may be found at the Clinical Guidance page at https://www.acog.org/clinical (click on "Clinical Guidance"

12. When you submit your revision, art saved in a digital format should accompany it. If your figure was created in Microsoft Word, Microsoft Excel, or Microsoft PowerPoint formats, please submit your original source file. Image files should not be copied and pasted into Microsoft Word or Microsoft PowerPoint.

Figures 1-2: Please add tick marks along the axes
Figure 3: Okay
Figures 4-5: Please add tick marks along the y-axis
All figures: Please upload as figure files on Editorial Manager.

13. Authors whose manuscripts have been accepted for publication have the option to pay an article processing charge and publish open access. With this choice, articles are made freely available online immediately upon publication. An information sheet is available at http://links.lww.com/LWW-ES/A48. The cost for publishing an article as open access can be found at https://wkauthorservices.editage.com/open-access/hybrid.html.

Please note that if your article is accepted, you will receive an email from the editorial office asking you to choose a publication route (traditional or open access). Please keep an eye out for that future email and be sure to respond to it promptly.

You will be receiving an Open Access Publication Charge letter from the Journal's Publisher, Wolters Kluwer, and instructions on how to submit any open access charges. The email will be from publicationservices@copyright.com with the subject line 'Please Submit Your Open Access Article Publication Charge(s)'. Please complete payment of the Open Access charges within 48 hours of receipt.

***

If you choose to revise your manuscript, please submit your revision through Editorial Manager at http://ong.editorialmanager.com. Your manuscript should be uploaded in a word processing format such as Microsoft Word. Your revision's cover letter should include the following:

* A confirmation that you have read the Instructions for Authors (http://edmgr.ovid.com/ong/accounts/authors.pdf), and
* A point-by-point response to each of the received comments in this letter. Do not omit your responses to the Editorial Office or Editors' comments.

If you submit a revision, we will assume that it has been developed in consultation with your co-authors and that each author has given approval to the final form of the revision.

Again, your paper will be maintained in active status for 21 days from the date of this letter. If we have not heard from you by May 28, 2021, we will assume you wish to withdraw the manuscript from further consideration.

Sincerely,

The Editors of Obstetrics & Gynecology

2019 IMPACT FACTOR: 5.524
2019 IMPACT FACTOR RANKING: 6th out of 82 ob/gyn journals

Manuscript Number ONG-21-510

To the Editors and Reviewers:

Thank you for the opportunity to revise this manuscript for further consideration by Obstetrics & Gynecology. We appreciate the reviewer's thoughtful comments which identified several areas where we could strengthen the clarity of the text and the presentation of findings. Our response to each of the reviewer's detailed comments is presented in the table below. Please note that the line numbers referenced in our response refer to the clean version of the revised manuscript.

No.	Comment	Response
Reviewer 1's Comments		
1.1	Line 70: 30-months versus 30-month?	Thank you for this correction. We agree that the number-noun modifier should be singular.
1.2	Line 153; Consider adding (AIM).	We added the acronym.
1.3	Lines 162-163: Consider elaborating on "movement in labor" as it's doesn't seem to be defined in the appendix (maybe describe as 'maternal physical movement in labor?').	Thank you for this clarification request. We changed the text to the term to be aligned with the terminology in the AIM metrics definitions (Appendix 2, Measure S2) which is "freedom of movement."
1.4	Line 185: What does "teaching states" mean?	Thank you, we have corrected this typo, which now reads "teaching status."
1.5	Line 203: Consider "..and among NTSV births.."	Thank you, we added the suggested text. The sentence now reads "Half-year rates were calculated for three outcomes: cesarean deliveries among all births and among NTSV births, and cesarean deliveries among NTSV births that were induced."
1.6	Lines 255-259: As per Figure 3 it seems that the steepest drop in C-section rates occurred between Q1-Q2 2018 and Q1-Q2 2019. Any thoughts on why?	Thank you for this question. Many hospitals in the collaborative had a slow start implementing the bundle. In the last six months of the collaborative, training coverage reached higher levels at the majority of hospitals (figure 1 and 2) and the collaborative offered a banner of recognition, which increased the motivation of some hospitals to complete practice changes.
1.7	Lines 300-301: Adopting more practices was only associated with lower c-section rates for the Response	Thank you, this is an important point, and we agree it should be noted. We have added the following text to the discussion: "The majority of

	domain. It may be useful to point out here that most of the implementation of the Response bundle occurred prior to the Collaborative (Table 2).	the implementation of practices in the Response domain at Maryland hospitals took place prior to the collaborative. The limited progress in the Response domain during the collaborative may help explain why the reduction in cesarean rates was smaller than that observed in California."
Reviewer 2's Comments		
2.1	Although, the authors do state in the beginning of the discussion section that this study is the second evaluation of the State's perinatal quality improvement collaborative to reduce NTSV cesarean delivery rates, this manuscript does not provide enough detail in the current study design to stand alone in making sense to readers that have not read the manuscript for the first evaluation of the collaborative. In addition, after pulling the original manuscript, there are a number of changes in the current reporting (e.g. respondents, responses, hospitals included, etc.) that needs some explanation in this manuscript	Thank you for pointing out areas where we could improve the clarity of the manuscript. We have made revisions to provide more detail on the methods in response to specific comments from the reviewer, as described below. Also, the Appendix provides granular information for readers, including specific collaborative activities, webinar topics, hospital recognition materials. If the reviewer and editor would like us to add further details in any area, we are happy to do so.
2.2	Page 9, Lines 175 - 180: was the original study design to be collect reported data through web-based surveys in Qualtrics at months 12 and 30 months after the collaborative was started. To be clear, is this study a follow-up with an additional survey for expanded analysis?	Thank you for this question. The two papers have distinct goals and analyses. The goal of the first paper, published in 2019 in this journal, was to evaluate the influence of implementation strategies on bundle practice adoption. Specifically, we assessed the association between discrete implementation strategies and the number and type of new practices adopted in the first year. That paper did not report any patient outcomes (e.g., cesarean delivery rates). The goal of the present paper is to assess the effect of the adoption of new bundle practices on patient outcomes. Specifically, we report the change in cesarean delivery rates, and assess whether the number of practices adopted is associated with cesarean delivery rates at the end of the collaborative.
2.3	Were the respondents predetermined (e.g. at time of study data collection, following study)?	We have clarified in the revised version that the survey invitation was sent to the lead of collaborative activities as designated by each

		hospital's leadership (lines 178-179). This was the person who represented the hospital at collaborative events (e.g., meetings, conference calls) and was responsible for submission of required data.
2.4	Could the data be affected by who responded?	It is possible that the data could be influenced by who responded, and we note the limitations of self-reported data in the discussion (lines 363-365). We did expect that the person responsible for leading collaborative activities will have the most comprehensive knowledge of bundle implementation, hence our inviting their participation in the surveys. In a minority of hospitals (23%), another team member completed the survey, in some cases because the person leading the collaborative activities left their position. We have added this information about the respondents to Table 1 and in the results (line 227-228).
2.5	How many attempts were made to get surveys from these participants and what sort of time frame was given for the response? What was your overall response rate?	Thank you for these questions. We have revised the first sentence of the results to highlight the 100% response rate as follows: "Each of the 31 hospitals in the Maryland cesarean collaborative completed the endline implementation survey (100% response rate; Table 1)." We have also clarified our efforts to achieve a high response rate in the methods (lines 184-186): "The surveys were first distributed in September 2018, and hospital leads were asked to respond within two months, with up to five e-mail reminders sent to encourage participation ."
2.6	Page 11, Lines 220-221: hospitals in collaborative included 7 Level 1…, 8 Level II… and 16 Level II/IV hospitals. Why are the numbers different in Table 1 and Figures 4 and 5?	We understand the reviewer's confusion and have added additional clarification to the paper. In brief, while all hospitals responded to the survey, one of the 31 hospitals did not submit a study participation agreement to provide their vital statistics and portal data. We have added the following clarification to the methods (line 203-205): "Of note, one Level II hospital responded to the implementation survey but declined to provide vital statistics and portal data for this evaluation." We have also added notes to Table 3 and all figures clarifying that one of the 31 hospitals is not included in analyses because their data were not provided.

2.7	Page 23, Table 1: Were there differences in ability to implement collaboratives/c-section rates in hospitals with higher proportions of Medicaid patients as compared to non-Medicaid patients?	Thank you for this question. We assessed differences in early implementation in the 2019 paper and found no differences by the proportion of Medicaid patients. During the first year of the collaborative, hospitals with >20% of patients covered by Medicaid implemented nearly the same number of new practices (difference in mean number of practices adopted: 0.04; p<0.898). We repeated that analysis in exploratory analysis for this paper and found similar results. By the end of the collaborative, hospitals with <20% Medicaid patients implemented an average of 1.3 more new practices, but the differences were not statistically significant in our sample (p=0.589). Due to space limitations, we did not include this analysis, or differences by other hospital characteristics, in the current paper.
2.8	Pages 29-30, Figures 4 and 5: The figures need better clarification in the footnote so that interpretation of these figures can stand alone. It is not self-evident what ABCDEFGHIJK…. in the first column represents (my assumption is that they represent the individual hospitals, although in the previous manuscript the hospitals were labeled H1, H2, etc.). Why are the level of care hospitals ordered differently in Figure 5? Why are there only 10 hospitals listed in Level II versus 11 listed in table 1?	Thank you for pointing out this need -- we agree that more clarification is needed to avoid confusion. In the original submission versions of the figures, hospitals were first ordered by care level, and then within each level, arranged from greatest decrease to greatest increase in cesarean delivery rates. Hospital code names were assigned in order in Figure 4, and because some hospitals had a different result for NTSV inductions then for all NTSV births, the order of hospitals changed accordingly in Figure 5. Figures 4 and 5 are missing one Level II hospital because one hospital responded to the survey, but did not share their vital statistics data – as also noted before. In the revised manuscript, we have clarified the hospital ordering and labeling in the footnote for all the figures, which now reads: "Hospitals are ordered from greatest decrease in rates to greatest increase in rates between baseline and endline. The baseline cesarean rate is indicated by a dot, and the endline cesarean rate is indicated by an arrowhead. Hospitals with decreasing rates are plotted in blue and hospitals with increasing rates are plotted in orange. Hospital code names reflect the level of care (L1 to L3/4) as well as a code letter.

		One of the 31 hospitals in the collaborative is not included because vital statistics data were not available." Please also note that in response to comment 3.10 from reviewer 3, we have also reorganized figures 4 and 5 to order hospitals from greatest decrease to greatest increase in cesarean rates, removing the stratification by level of care.
Reviewer 3's Comments		
3.1	General Comments: This is a high-quality study examining the impact of a state-wide initiative to reduce NTSV cesarean section (cs) rates in Maryland. This study's findings dovetail with the recent findings from Rosenstein et al in JAMA (see below). A large volume of information is presented in the paper, which includes survey data, rates of hospital adoption of bundle elements, and changes in NTSV cs rates over time.	Thank you.
3.2	It seems the authors focus most of their attention on the changes in NTSV rates over time (ie primary aims). Therefore, I would recommend packaging the survey data on domain uptake as secondary aims/outcomes.	Thank you for this suggestion. We have revised the description of study aims as follows (lines 128-132): "The primary aim of this paper is to evaluate the impact of Maryland's collaborative on state-level NTSV cesarean rates. The secondary aim is to assess the extent to which hospitals participating in Maryland's collaborative adopted policy and practice changes, and whether adoption of more practices was associated with lower cesarean delivery rates."
3.3	L113 - The Kozhimannil study only reported variation in cesarean rates. It may be an overreach to state that standardizing care can impact cesarean rates based only from this study's finding. At this point, you may want to cite the recent Rosenstein JAMA paper as evidence of how a multifaceted state-wide approach may result in a reduction in the NTSV cesarean rate in California over time (PMID: 33904868). I appreciate that this study came out whilst your paper was under peer review.	Thank you for this suggestion, we softened the language around the implications of the Kozhimannil study. We agree the recently published Rosenstein study is important to reference. We updated our references to CMQCC's work in both the introduction and the discussion.

3.4	L143/145/214/Table 1: Any information on the team member or representatives sent from each hospital - e.g. clinical or nursing lead or QI representative? This may speak to whether the member or representative can be considered a local 'champion' of bundle implementation and its potential impact. Also, were these representatives the same as those who completed the endline survey (Table 1)?	The person who was delegated responsibility for the leading the bundle implementation at each hospital also attended the majority of collaborative events, yet additional hospital representatives participated from most hospitals. In most cases, a member of the nursing staff (director, manager, or quality/safety manager) was responsible for leading bundle implementation and designated as "lead" by their hospital leadership. We updated Table 1 to include whether the respondents were responsible for leading collaborative activities (also in response to reviewer comment 2.4).
3.5	L176: Can you point readers to Appendix 3 which describes each of these bundle practices? Else, readers have to dig to find them.	Thank you for this suggestion. We agree and have added a parenthetical notation guiding readers to Appendix 3.
3.6	L220: When describing Level I-III hospitals, are you referring to ACOG Levels of Maternal Care? If so, this needs describing in the methods &/or appendix.	We clarified in the methods that we classified hospitals by the levels of care in the ACOG/SMFM consensus statement[1] and added the reference (line 194).
3.7	L228/Table 2: I'm unsure about the clinical value of summarizing data from the domains - currently presented in Table 2. Also, you present means but, from a practical standpoint, it is not possible to implement a fraction of a bundle element e.g., 0.8 for R1. I would suggest dropping table 2 and simply commenting on the findings from Appendix 3 which provides more granular and relevant information about the degree of implementation of each practice across all the hospitals. Consider highlighting practices that had the highest rates of 'not starting' and fully implemented. For example, R1.2, R3.6, and R.4.3 had high rates of not being started. The high rate (74%) for not integrating doulas stands out.	Thank you for this suggestion. The statistical editor also requested that we present the median value rather than mean value, and do so in the revised manuscript in Table 2 to respond to both comments. We also understand the reviewer's suggestion to replace Table 2 with Appendix 3, which describes adoption for each practice – we can do so if the reviewer and editor agree that this is needed. We strongly agree with the reviewer's sentiment that the practices with particularly high and low adoption should be noted. We previously highlighted practices with high adoption in the text (lines 244-254). In the revised manuscript, we have also added a sentence to highlight practices with low adoption noted by the reviewer (lines 254-256): "The practices with the lowest overall adoption include implementing a policy to integrate doulas in the birth care team (3,

[1] Levels of Maternal Care: Obstetric Care Consensus No, 9 Summary. *Obstet Gynecol*. 2019;134(2):428-434.

		10%) and integrating new tools or guidelines in the electronic health record system (4, 13%)." At the same time, we think that the variability in cumulative progress in adopting recommended practices at each hospital is a key finding in this evaluation, and important for interpreting the variability in cesarean delivery rates by implementation strength as presented in Table 3.
3.8	Figures 1 and 2: Clarify the legend details in the text and figures. What does 20-39%, 40-59%, etc..refer to? Do these data refer to the proportion of all practices in the bundle completed? For example, does 20-39% refer to 5-10 of the 23 practices? Also, I wonder whether a bar chart would be a better visual approach for conveying changes in coverage over time.	Thank you for pointing out the need for clarification. We revised the figure titles to explain the data presented as follows: "Change over time in the cumulative proportion of [position] that completed education on ACOG/SMFM labor management guidelines at hospitals participating in the collaborative." The Y axis is labelled "number of hospitals" and the legend corresponds to the proportion of providers who completed education at each hospital counted on the y axis, as clarified by the revised figure title. We chose a stack line chart to visualize this data in order to convey the progressive gains in training coverage over time.
3.9	Figure 3: As you have data for CS rates by quarter, consider presenting quarterly as opposed to semi-annual rates. More granular information would be valuable in assessing changes over time than semi-annual rates. Regression modeling would help in better examining the degree of change over time as opposed to simply doing chi-squared tests comparing the 1st quarter to the last.	Given the relatively small number of hospitals and births represented in the Maryland collaborative, quarterly variations in the cesarean delivery rates tend to obscure the overall trends that can be observed when looking at half-year results. Of note, the other published studies of cesarean collaborative outcomes also report half-year rates[2] or annual rates.[3] By also reporting half-year rates, we better facilitate comparisons between the results of cesarean collaboratives in different states.

[2] Main EK, Chang S-C, Cape V, Sakowski C, Smith H, Vasher J. Safety Assessment of a Large-Scale Improvement Collaborative to Reduce Nulliparous Cesarean Delivery Rates: *Obstetrics & Gynecology*. 2019;133(4):613-623.

[3] Rosenstein MG, Chang S-C, Sakowski C, et al. Hospital Quality Improvement Interventions, Statewide Policy Initiatives, and Rates of Cesarean Delivery for Nulliparous, Term, Singleton, Vertex Births in California. *JAMA*. 2021;325(16):1631-1639.

		We separately respond to the reviewer's detailed comments regarding regression modelling under comment 3.12.
3.10	Figures 4 and 5: Could you reorganize the chart to rank hospitals from those with the greatest reduction in NTSV cesarean rates at one end of the Y axis and hospitals with the greatest increase in rates at the other end (ie a caterpillar plot). The data could be presented as % change or absolute % increase/decrease as opposed to dots with arrows to depict differences that are harder to interpret.	Thank you for this suggestion, we agree that reordering would improve readers' ability to interpret the figure. We have reorganized the dumbbell plots from greatest decrease to greatest increase in rates between baseline and endline (similar to the approach taken with caterpillar plots). We have kept the dumbbell plot format because we believe that showing the differences in starting rates, in addition to the absolute change in rates, is important context. This reordering required us to remove the stratified presentation by level of care. Therefore, we have changed the code names for each hospital to include a reference to level of care, as well as a code letter. This nomenclature is explained in the notes for the figure.
3.11	Table 3: Instead of a 'significance' column - I'd recommend calculating RRs comparing hospitals above the median to those below (reference group) and presenting absolute risk differences also. Also, instead of medians (arbitrary dichotomization), consider using tertiles (or other approaches) to obtain more detailed comparative data on the effect of domain uptake on NTSV CS rates.	Thank you. This change was made. Table 3 now presents the relative risks with 95% confidence intervals.
3.12	Please also consider regression modeling comparing hospitals by the level of maternal care. Might there be a way to model the use of practice elements (independent variable) against NTSV cs rates (dependent variable) in a mixed-effects model with hospitals as the random effect? This would provide some insight into the potential effect of each practice element on the NTSV rate. If not,	We share the reviewer's desire to investigate the relative effect of each practice in the bundle. However, we do not have access to patient level data, rather aggregate data on birth outcomes for each of the 30 hospitals that agreed to share their vital statistics data. Therefore, the outcome data available for this study limit the analyses that we can complete, given the documented inaccuracies of regression estimation with a small number of observations per predictor variable.[4] In

[4] Courvoisier DS, Combescure C, Agoritsas T, Gayet-Ageron A, Perneger TV. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. *Journal of Clinical Epidemiology*. 2011;64(9):993-1000.

	highlight as a limitation, with a call for more studies to examine this.	addition to the large number of practices in the bundle (26), there is also considerable heterogeneity in the characteristics of the hospitals in the collaborative. We agree with the reviewer that this is a limitation of our study, and have added the following sentence in the study limitations section of the revised manuscript (line 365-367): "Further studies are needed to assess the effectiveness of the individual practices in the cesarean bundle that are not already well studied."
3.13	L268 - I didn't see the statistical analysis for how you compared cs rates across Levels of hospitals.	We have clarified in the Methods section that the one-sided Fisher's exact test was used to test the significance of changes in cesarean delivery rates for individual hospitals (lines 214-216). The changes in rates by hospital are presented in figures 4 and 5, where significant differences are noted with an asterisk in the code name of the hospital. The revised presentation of figure 4 and 5 improves the clarity of the figures. We hope the reviewer agrees.
3.14	L331 - can you discuss or comment on what barriers or reasons there might be for implementing and affecting QI? I'm unclear what 'quality improvement capacity and motivation' and 'engaging clinicians' mean in this context. Some specific examples would be helpful.	We appreciate this suggestion and added the following examples to the text (line 350-353): "For example, labor and delivery units that have experienced interdisciplinary quality improvement teams, support from hospital-wide quality improvement offices, and strong patient safety cultures, are likely better prepared to implement maternal safety bundles." We also added a reference to a recently published (May 2021) mixed-methods study of determinants of implementation in California's cesarean collaborative.
3.15	L348 - A key limitation in your study and that of Rosenstein et al in JAMA is that there are no maternal morbidity data to determine whether reducing NTSV CS rates is associated with less, more, or the same risk of maternal morbidity. This is a problem if there is greater pressure on obstetric providers to deliver patients vaginally instead of by cs. Is there an unintended consequence of this? Highlight the fact that we need more data to examine changes in specific morbidities (not just	This is an important point and we agree with the reviewer. The California collaborative did report an earlier safety assessment of the first two rounds of their cesarean collaborative that tracked some SMM variables (chorioamnionitis, blood transfusions, third- or fourth-degree lacerations, and operative vaginal delivery) and reported that no measure was statistically worse following reductions in the cesarean rate. In Maryland, we do not benefit from linked hospital discharge-birth certificate data to be able to examine SMM rates in NTSV births. We have added the following discussion, as suggested

	SMM) e.g., infections and PPHs from longer labor times before a vaginal delivery or even a cesarean, especially in women who have a prolonged and/or complicated labor. Although Rosenstein et al reported no change in the rate of neonatal morbidity (as a composite), this finding needs to be confirmed in other populations also. The ultimate goal should be a reduction in NSTV cesarean rates coupled with no increases in maternal and perinatal morbidity. The pressure on obstetric providers to weigh the risk of maternal/perinatal morbidity in a woman with abnormal labor against the subjective pressure to perform a vaginal delivery and 'hitting their NSTV cs metric' should not be underappreciated.	(lines 371-373): "In Maryland and other states, data systems that link hospital discharge data with birth certificates are needed to accurately assess maternal morbidity among NTSV births. While California's cesarean collaborative reported no change in some maternal morbidities (e.g., transfusion, 3rd and 4th degree lacerations), strengthening the ability of all states to monitor maternal morbidities is important for ensuring maternal safety while working to reduce cesarean delivery rates."
Statistical Editor's Comments		
4.1	Table 1: Since the n = 31, should round the %s to nearest integer %, not cite to 0.1% precision. Given the sample size, should cite the years experience, annual number of births and proportion of pts covered by Medicaid as median (Range).	Thank you. This change was made as suggested. We also report the median and range for years of experience, number of births, and patients covered by Medicaid.
4.2	Table 2: It is not clear what the Authors meant by "Average". ? mean. Given the ranges and counts, a better metric would be to summarize as median(range) for each time period. Should also enumerate all missing data.	We have implemented the recommended change in Table 2, reporting medians and ranges. There are no missing data for this table.
4.4	Table 3: Again, the issue of precision for number of hospitals, which should be rounded to nearest integer %, not to 0.1 % precision. For the cesarean rates, should round to nearest 0.1%, not to 0.01% precision.	We have revised the presentation of percentages as requested.
4.5	Figs 4 and 5 should be in supplemental material.	While we understand this suggestion by the statistical editor, we think that the variability in outcomes between hospitals is a key finding of

		this paper and deserves more attention in studies of perinatal quality improvement collaboratives (PQCs). If allowed by the editors, we would like to maintain figures 4 and 5 in the main body of the paper. As recommended by the Editor (comment 5.2. below), we have added tick marks on the y-axis. We will wait for the final editorial decision regarding inclusion of these 2 figures in the body of the paper or as supplemental material.
4.6	lines 81-83: Should clarify that these % changes are relative, not absolute changes in proportion cesarean delivery rates.	We revised the text to present only absolute changes, as suggested by the Editor in comment 5.1.
Editor's Comments		
5.1	Please throughout emphasize absolute as opposed to relative risk reductions.	We have replaced any reference to percent/relative change in cesarean delivery rates with absolute changes in terms of percentage point differences.
5.2	Figures 1-2: Please add tick marks along the axes Figure 3: Okay Figures 4-5: Please add tick marks along the y-axis All figures: Please upload as figure files on Editorial Manager.	We have added tick marks to the figures as recommended. Figures are uploaded as separate files with the revised manuscript submission.