

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Socioeconomic Gradients of Adverse Birth Outcomes and Related Maternal Factors in Rural and Urban Alberta, Canada: A Concentration Index Approach.
AUTHORS	Ospina, Maria; Osornio-Vargas, Álvaro Román; Nielsen, Charlene; Crawford, Susan; Kumar, Manoj; Aziz, Khalid; Serrano-Lomelin, Jesus

VERSION 1 – REVIEW

REVIEWER	Hamid Baradaran Iran University of Medical Sciences
REVIEW RETURNED	09-Sep-2019

GENERAL COMMENTS	<p>Abstract:</p> <ul style="list-style-type: none">· Mention the time of study in the setting.· The used data is old. 2006-2012. Socioeconomic status and maternal health problems were changed during 2012-2019. <p>Methods:</p> <ul style="list-style-type: none">· Use Stata instead of STATA.· The authors assumed that socioeconomic status was not changed during 2006-2012. It's an important assumption. <p>Results:</p> <ul style="list-style-type: none">· The comparison of characteristics between rural and urban samples should be presented in a table.
-------------------------	--

REVIEWER	Tom Clemens University of Edinburgh, UK
REVIEW RETURNED	18-Sep-2019

GENERAL COMMENTS	<p>Comments:</p> <p>1. On page 5/6 I think the text needs to spell out in more detail why precisely the study is important. At the moment it isn't clear why the specific question of SES gradients differing between urban and rural areas is being asked and I think some more theoretical justification is needed. Was there an a priori hypothesis about what ways the gradients might differ between urban and rural areas? This might help to frame and provide more justification for the study. It is mentioned that the study deals with a relevant health policy issue in</p>
-------------------------	---

	<p>countries with universal healthcare access but I don't think this is clear.</p> <p>2. There is a lack of detail in the urban and rural classification which I think is a limitation. It seems to me that the definition of rural will contain some significant heterogeneity and areas within the designation of rural may not be especially comparable. Are there other more detailed categorisations that could be used? Looking at the stats Canada page I can see measures of "Urban Core", "urban fringe" and "rural fringe" for example. Linked to the point above though, without detailing what the theoretical basis for expecting differences in SES gradients between urban and rural areas it is difficult to determine the effect this urban and rural classification issue has in terms of interpretation of the results.</p> <p>3. There is a lack of distinction or conflation throughout the text between "area-level" deprivation, area level SES and individual level SES. Many of these terms are used interchangeably and it is unclear which of these the study is most interested in (they are quite different e.g. one is measuring the aggregate level of SES in an area (i.e. an "area effect") and the other is measuring SES at the individual level). If the study is interested in individual level SES (which I think it is...) and area SES measures are being used to proxy this (in the absence of individual level SES information) then it needs to be stated much more clearly but also the potential problems of the ecological fallacy needs to be acknowledged.</p> <p>4. Linked to the point above, I think the measure of area SES may be problematic because it may not be measuring things consistently between urban and rural areas. The meaning of variables like e.g. home ownership, transport mode, year of home construction are unlikely to be measuring the same underlying phenomena (SES) in urban areas as they are in rural areas and vice versa. For example renting is likely to be higher in cities among young professionals and car ownership is likely to be higher in rural areas out of necessity rather than material privilege. I can see the issue of misclassification is addressed to some degree in the discussion but I think this is a much wider problem that introduces bias and the study will need to acknowledge this. One thing that might also be considered is reworking the area SES measure by dropping some of the individual components and re-running the analysis. Again though, this depends on the precise aims that the study is addressing and the theoretical basis for these.</p> <p>Minor comments:</p> <p>1. I couldn't see anywhere the acronym AHP being defined before page 7 (apart from in the abstract). Please specify what this acronym stands for in the main body of the text.</p> <p>2. More discussion on missing data is needed. Was there any? Particularly for variables like pre-pregnancy weight, smoking and drug dependency etc which can often have high rates of missing data.</p> <p>3. Figure 2 was very difficult to read as the text is too small. It might be worth splitting into two tables, one for clinical outcomes e.g. PTB, SGS, Gestational diabetes etc and another for other outcomes to make it more readable.</p>
--	---

REVIEWER	<p>Dusan Petrovic Unisanté Centre universitaire de médecine générale</p>
-----------------	--

	et santé publique - Lausanne Département Epidémiologie et systèmes de santé Secteur Maladies Chronique Lausanne Switzerland
REVIEW RETURNED	11-Nov-2019

GENERAL COMMENTS	<p>In this research, the authors investigate the association between SES and Adverse birth outcomes (ABO) in rural and urban areas of the Alberta province in Canada. The paper is very clear, well written, with the main objectives being well defined and well addressed. While most of my comments are minor, I would have three more important comments that if addressed would further improve the quality of this paper.</p> <p>Discussion:</p> <ol style="list-style-type: none"> 1. The second and third chapter of the discussion addressing the rural vs. urban differences as well as the socioeconomic gradient in ABO are clear, however these two paragraphs could be somewhat more elaborated. More details about the underlying mechanisms could be given, particularly for the urban vs. rural differences (the authors mention : adverse maternal behaviors, and fewer medical resources in rural areas as being more prevalent in rural areas) however, in spite of this, the reader is left yearning for more. 2 or 3 additional sentences would be greatly appreciated. The same should be applied for the socioeconomic gradient in ABO. 2. The structure of the discussion could be improved. Particularly, it would be appreciated if there were clearly defined paragraphs for "Strengths and limitations" (p.14-15), a clearly mentioned paragraph addressing "Future perspectives" described on p.15 ("Studies about ... "), and a "Conclusion" paragraph for the very last part of the discussion ("In summary...") 3. My last "important comment" is related to Table 1. This table is very informative, however, it would be appreciated if a p-value for trends related to SES differences in ABO were added. In the results section, the authors mention that the prevalence of ABO increases as SES decreases. This is very good and in line with previous findings related to SES differences in health related outcomes (generally), however, a p-value for a increasing/decreasing trend for each ABO outcome would be appreciated (despite the fact that one observed that there is no overlapping between CI in percentages). <p>Some minor comments:</p> <ol style="list-style-type: none"> 1. References should usually go at the end of the sentence. I would personally also prefer having references for a given sentence prior to the ".", but if the journal is OK with the way authors proceeded , it is fine. 2. In the results section (text), it would be clearer if the authors would keep the brackets "[]" for the CIs, otherwise it is very hard to read. 3. I've noticed a couple of mistakes when citing the concentration index abbreviation. For both rural and urban the authors frequently put CIdxU which is confusing, particularly when the abbreviation is
-------------------------	--

mentioned for the first time in the Methods-Statistical analyses (p.9).

VERSION 1 – AUTHOR RESPONSE

Peer-Review Comments	Response to Reviewers
Reviewer # 1	
1. Abstract: Mention the time of study in the setting.	Information about study period was moved to the “Setting” section of the abstract.
2. Abstract: The used data is old. 2006-2012. Socioeconomic status and maternal health problems were changed during 2012-2019.	<p>Data supporting this manuscript was part of a larger study that started retrospective data collection in 2013 and finished recently as a result of a sophisticated interdisciplinary approach (https://sites.google.com/a/uAlberta.ca/domino/).</p> <p>Our research assumed that SES will be stable over relatively short (5 years) periods of time. Please see our reply to comment #4 about stability of area-level SES grouping over time.</p> <p>A national census in Canada is conducted every five years by Statistics Canada. Census data is used by researchers in Canada for rate standardization purposes and for calculating area-level deprivation measures. The Canadian census of 2006 still remained to be used as the reference group for rate standardization and calculation of area-level deprivation measures in Canadian studies.</p>
3. Methods: Use Stata instead of STATA	Change has been made to Stata
4. The authors assumed that socioeconomic	We used 2006 Canada census data to calculate

status was not changed during 2006-2012.

It's an important assumption

our area-level deprivation measures. Canada's national census is conducted every five years by Statistics Canada. Census data from 2006 is the standard in Canada for rate standardization purposes and for calculating area-level deprivation measures because of completeness in coverage and comprehensiveness of data collected. The methods assume no changes in area-level deprivation between 2006 and 2012. We disclosed in the section of "Strengths and limitations of this study" that potential

	<p>misclassification of the SES may occur as we assume no changes in area-level SES index between 2006 and 2012. We have added a paragraph in the Discussion session to discuss this assumption and any potential limitations. (Page 18, Paragraph 2, Lines 1-7):</p> <p><i>“We used area-level data from the 2006 Canadian census for the calculation of the socioeconomic status (SES) index. The method assumed no changes in area-level deprivation between 2006 and 2012 and therefore, potential misclassification of the SES may occur. Other studies using area-level deprivation measures have attempted to quantify changes in SES categories over time and have assumed that SES remains relatively stable over time[30, 46], and that census-based measures of deprivation can be used in larger comparative studies across decades without loss of continuity over time.[50, 51]”</i></p>
<p>Reviewer # 2</p>	
<p>5. On page 5/6 I think the text needs to spell out in more detail why precisely the study is important. At the moment it isn't clear why the specific question of SES gradients differing between urban and rural areas is being asked and I think some more theoretical justification is needed. Was there an a priori hypothesis about what ways the</p>	<p>Thank you for this suggestion. We have expanded the Introduction to address why the specific question of SES gradient differences between urban and rural settings (Page 5, Paragraph 2, Lines 6-14): <i>“Knowledge gaps remain to fully understand the interconnections between socioeconomic characteristics, area of residence, and maternal and perinatal health.</i></p>

gradients might differ between urban and rural areas? This might help to frame and provide more justification for the study. It is mentioned that the study deals with a relevant health policy issue in countries with universal healthcare access but I don't think this is clear

Exploring this association is particularly important as both urban and rural living have been also associated with adverse health outcomes.[14] However, it is unknown whether health advantages and disadvantages of living in urban and rural areas are equally distributed in all socioeconomic groups or if gradients in health exist affecting the more disadvantaged groups. On the one hand, diverse theories about urban residence posit that cities create harmful environments for human health.[15-17] Alternatively, rural areas encompassing vast extensions of land have been also associated with poor outcomes.[18]

We have also included our apriori hypothesis after the study objective (Page 7, Paragraph 1, Lines 2-4): “*We hypothesize that adverse birth*

	<p><i>outcomes in urban and rural areas are distributed differently and potentially related to socioeconomic gradients within the two areas of residence”.</i></p>
<p>6. There is a lack of detail in the urban and rural classification which I think is a limitation. It seems to me that the definition of rural will contain some significant heterogeneity and areas within the designation of rural may not be especially comparable. Are there other more detailed categorisations that could be used? Looking at the stats Canada page I can see measures of “Urban Core”, “urban fringe” and “rural fringe” for example. Linked to the point above though, without detailing what the theoretical basis for expecting differences in SES gradients between urban and rural areas it is difficult to determine the effect this urban and rural classification issue has in terms of interpretation of the results.</p>	<p>We used the standard definition provided by Statistics Canada for the 2006 geographic framework. For the subsequent Census (i.e., 2011, 2016), as the reviewer noted, Statistics Canada has provided a different classification of places (for example, large urban population centres, medium population centres, small population centres, rural areas), which do not apply for the 2006 geographic framework. We have clarified this in Methods section about Definitions of urban and rural maternal place of residence at delivery (Page 8, Paragraph 3, Lines 1-4): “<i>We used the 2006 geographic standards provided by Statistics Canada to classify areas of residence (urban, rural) and georeferenced data for postal code locations.</i>[27] The six-character postal codes of the maternal place of residence at delivery were classified as rural or urban according to population concentration and density, <i>based on the 2006 geographic framework.</i>”</p>
<p>7. There is a lack of distinction or conflation throughout the text between “area-level” deprivation, area level SES and individual level SES. Many of these terms</p>	<p>In this study, we used area-based SES indicators rather than individual level SES. We made sure to described the use of area-level SES indicators throughout the text. We mentioned in the</p>

are used interchangeably and it is unclear which of these the study is most interested in (they are quite different e.g. one is measuring the aggregate level of SES in an area (i.e. an “area effect”) and the other is measuring SES at the individual level). If the study is interested in individual level SES (which I think it is...) and area SES measures are being used to proxy this (in the absence of individual level SES information) then it needs to be stated much more clearly but also the potential problems of the ecological fallacy needs to be acknowledged.

Discussion that area-level SES indicators were used as a proxy for individual-level measures in the study. We expanded in the Discussion the implications of reporting an area-level deprivation measure for the interpretation of the results. We have added the following paragraph: (Page 17, Paragraph 2, Lines 8-15):
“Area-level measures of SES gradients are important to describe inequalities in health outcomes across populations.[47, 48] There is evidence that these aggregate measures are good proxies for individual deprivation, have similar performance than individual-level SES measures, and represent a low risk of ecological bias.[49] Furthermore, we did not use area-level data to impute individual values in the study cohort but rather used individual

	<p><i>maternal postal codes to assign cohort members to a dissemination area that shared particular features from a census perspective. Since both the exposure (maternal postal code) and outcome were measured at the individual level, the risk of ecologic fallacy is likely low.[47]”</i></p> <p>The Concentration Index is a tool to understand health disparities at population level. Results are reported at group level, acknowledging that some residual differences between the individual and areal-level exist.</p>
<p>8. Linked to the point above, I think the measure of area SES may be problematic because it may not be measuring things consistently between urban and rural areas. The meaning of variables like e.g. home ownership, transport mode, year of home construction are unlikely to be measuring the same underlying phenomena (SES) in urban areas as they are in rural areas and vice versa. For example renting is likely to be higher in cities among young professionals and car ownership is likely to be higher in rural areas out of necessity rather than material privilege. I can see the issue of misclassification is addressed to some degree in the discussion but I think</p>	<p>The area SES measure that we used in our study is a Canadian SES index that has been developed and validated independently (Chan et al 2015. BMC Public Health 2015). The index resulted from conducting principal component analysis (PCA) on more than 20 SES variables for 52,974 census dissemination areas in Canada. It is not feasible to tease out/drop off individual components of the index as this would imply building a completely new SES index with unknown measurement properties. The study by Chan et al (referenced in our study) describes the parameters and variables used in the selection for the PCA.</p> <p>We agree with the reviewer that area-based SES indexes remain sensitive to urban-rural</p>

this is a much wider problem that introduces bias and the study will need to acknowledge this. One thing that might also be considered is reworking the area SES measure by dropping some of the individual components and re-running the analysis. Again though, this depends on the precise aims that the study is addressing and the theoretical basis for these.

differences and that variables that capture deprivation and SES in cities may not perform well in rural areas. We have added the following paragraph in the Discussion to address this concern (Page 18, Paragraph 3; Lines 1-6): "*There is concern that area-based SES indexes are likely sensitive to urban-rural differences and that variables that capture deprivation and SES in cities may not perform well in rural areas. Despite these conceptual constraints, there is evidence from other studies showing that available deprivation indexes can be used legitimately used in both settings, supporting the hypothesis that the underlying relationship between areal-level SES and health gradients is the same in rural and urban areas.[35, 52, 53]*"

<p>9. I couldn't see anywhere the acronym APHP being defined before page 7 (apart from I the abstract). Please specify what this acronym stands for in the main body of the text.</p>	<p>Thank you for pointing out this error. APHP is the acronym for the Alberta Perinatal Health Program. We spelt the acronym out in the abstract but we forgot to include this in the manuscript body. We have now spelt out the acronym the first time that is cited in the text (Page 7): <i>"We used data from the Alberta Perinatal Health Program (APHP), which is a validated clinical perinatal registry."</i></p>
<p>10. More discussion on missing data is needed. Was there any? Particularly for variables like pre-pregnancy weight, smoking and drug dependency etc which can often have high rates of missing data.</p>	<p>Thank you for this comment.</p> <p>The AHPH is a high-quality clinical perinatal registry in which the percentage of missing values is generally low for most of the variables. We have indicated in the description of the Results the % of missing values for the variables included in the study (Page 10, Paragraph 4, Line 5 to 8):</p> <p><i>"Small numbers of missing values were present for maternal weight, gestational hypertension, gestational diabetes, and smoking during pregnancy in the urban (0.81%; n = 2,667) and rural areas (1.3%; n = 497). There were no missing values for PTB, SGA, and LGA categories in both urban and rural areas"</i>.</p>
<p>11. Figure 2 was very difficult to read as the text is too small. It might be worth splitting into two tables, one for clinical outcomes e.g. PTB, SGS, Gestational diabetes etc.</p>	<p>Thank you for this suggestion.</p> <p>We think that by splitting the results into two tables and figures, we cannot provide a fine description of the overall pattern of the</p>

<p>and another for other outcomes to make it more readable.</p>	<p>inequality-gradient across all variables for urban and rural settings. Therefore, we considered a new improved Figure 2 (now Figure 4). The new Figure 4 has a larger font to facilitate its reading. Figure 4 has been referenced in the manuscript (Page 14, Paragraph 1, Line 1).</p>
<p>Reviewer # 3</p>	
<p>12. In this research, the authors investigate the association between SES and Adverse birth outcomes (ABO) in rural and urban areas of the Alberta province in Canada. The paper is very clear, well written, with the main objectives being well defined and well addressed. While most of my comments are minor, I would have three more important comments that if addressed</p>	<p>Thank you for this positive feedback. We have made our best efforts to address the reviewer's comments.</p>

<p>would further improve the quality of this paper.</p>	
<p>13. The second and third chapter of the discussion addressing the rural vs. urban differences as well as the socioeconomic gradient in ABO are clear, however these two paragraphs could be somewhat more elaborated. More details about the underlying mechanisms could be given, particularly for the urban vs. rural differences (the authors mention : adverse maternal behaviors, and fewer medical resources in rural areas as being more prevalent in rural areas) however, in spite of this, the reader is left yearning for more. 2 or 3 additional sentences would be greatly appreciated. The same should be applied for the socioeconomic gradient in ABO.</p>	<p>Thank you for this suggestion. We have expanded the discussion to provide more details about potential underlying mechanisms of urban/rural and SES gradient in ABO (Page 15; Paragraph 2; Lines 1-4).</p> <p><i>“The pathways for the associations among area-level deprivation, maternal health, and adverse birth outcomes are complex and likely multifactorial. We found that area-level deprivation and geographic area of residence differentially associate with fetal growth and duration of gestation. One potential explanation for these results is that women residing in rural areas are more vulnerable to neighbourhood deprivation.[35]”</i></p> <p>We also added (Page 15; Paragraph 2; Lines 11-19):</p> <p><i>“Other potential explanations may be linked to low health literacy in rural populations about the effects of lifestyle behaviours in childbearing age and the impact on birth outcomes, and shortages in resources to stay better informed than women living in more urbanized areas.[40] Systemic and structural influences such as food security, health services</i></p>

	<p><i>access may also account for the socioeconomic gradient in the urban-rural divide. Lastly, the “healthy migration” effect[41] can contribute to our study results. It is possible that healthy women living in rural and remote areas are most likely to migrate to more urbanized areas, leaving behind their counterparts at a higher risk of experiencing adverse birth outcomes.”</i></p>
<p>14. The structure of the discussion could be improved. Particularly, it would be appreciated if there were clearly defined paragraphs for "Strengths and limitations" (p.14-15), a clearly mentioned paragraph addressing "Future perspectives" described on p.15 ("Studies about ... "), and a "Conclusion" paragraph for the very last part of the discussion ("In summary...")</p>	<p>Thank you for this suggestion. We have inserted appropriate subheadings for “Strengths and limitations of the study” (Page 16), “Future perspectives” (Page 18) and “Conclusion” (Page 19).</p>

<p>15. My last "important comment" is related to Table 1. This table is very informative, however, it would be appreciated if a p-value for trends related to SES differences in ABO were added. In the results section, the authors mention that the prevalence of ABO increases as SES decreases. This is very good and in line with previous findings related to SES differences in health related outcomes (generally), however, a p-value for a increasing/decreasing trend for each ABO outcome would be appreciated (despite the fact that one observed that there is no overlapping between CI in percentages).</p>	<p>Thank you for this suggestion.</p> <p>We have prepared panel-graph figures (Figures 2 and 3) to show the gradients of prevalence by health outcomes across SES quintiles. We hope these graphs will facilitate the interpretation of the results presented in Table 1. We have incorporated in these graphs the p-value for the linear trend when it was statistically significant ($p < 0.05$). We have added notes to Figures 2 and 3 to explain the test (regression for linear trends). We only tested for linear trends and not for all-possible pairwise comparisons as our objective was not to estimate all possible differences across SES-quintiles. Our primary purpose was to assess health inequalities across a SES gradient. We have also included a reference to these new figures in the manuscript body (Page 11, Paragraph 2, Line 1).</p>
<p>16. References should usually go at the end of the sentence. I would personally also prefer having references for a given sentence prior to the ".", but if the journal is OK with the way authors proceeded , it is fine.</p>	<p>Thank you for this suggestion. We followed the BMJ Open formatting guidelines (Reference numbers in the text should be inserted immediately after punctuation (with no word spacing). No changes were made.</p>
<p>17. In the results section (text), it would be clearer if the authors would keep the brackets "[]" for the CIs, otherwise it is very hard to read.</p>	<p>We have included brackets "[]" for the confidence intervals reported in the study results.</p>
<p>18. I've noticed a couple of mistakes when citing the concentration index abbreviation.</p>	<p>We apologize for these typos. We have revised the manuscript and used the correct</p>

For both rural and urban the authors frequently put CldxU which is confusing, particularly when the abbreviation is mentioned for the first time in the Methods-Statistical analyses (p.9).	nomenclature for rural (CldxR) and urban concentration indexes (CldxU) throughout the text.
---	---

VERSION 2 – REVIEW

REVIEWER	Dusan Petrovic UNISANTE, Route de la corniche 10, 1010 Lausanne
REVIEW RETURNED	22-Dec-2019
GENERAL COMMENTS	The authors have addressed my comments in a complete and appropriate way.