Article details: 2022-0239
Title: Utilization of physician mental health services by mothers with young children before and during the COVID-19 pandemic: a population-based study in Ontario, Canada
Authors: John Moin, Simone Vigod, Lesley Plumptre, Natalie Troke, Irene Papanicolas, Walter Wodchis, Geoff Anderson

Reviewer 1: Dr. Amelie Boutin
Institution: Universite Laval
General comments (author response in bold)

Summary
The Authors conducted a population-based repeated cross-sectional study of mental health care utilisation among women of young children (0-5 and 6-12 years old) before and after the onset of the covid-19 pandemic. They observed higher rates of mental health care utilisation at the onset of the pandemic, which remained high up until the end of the studied period (November 2021). Most services used were related to mood and anxiety disorders diagnoses.

Thank you for the opportunity to review this very interesting manuscript.

COMMENTS

1. The abbreviation NPI in the abstract needs to be defined.
NPI has been defined in the abstract (page 1).

2. At the end of the introduction, the Authors mention that they examined mental health visits rates by provider type, variations in outcomes (which outcomes are we talking about), and child (age?), which I did not see in the manuscript.
"Variations in outcomes" (refers to the four mental health diagnostic categories which are based on a series of related individual diagnostic codes with details (page 2 - supplemental document). The age for mothers is based on approximate 10-year age bands and for the children they correspond to either cohort 1 (0-5 years of age) or cohort 2 (6-12 years of age). The end of the introduction has been rewritten. We also included more detail about study outcomes (pages 5-6).

3. Would it be possible to provide information on the type of care seeked? (e.g., were hospitalizations for mental health disorders included in the "visits" considered for this study, did their rate change after the onset of the pandemic and over its course? Was there a change in the type of care services seeked/utilized?)
We do not have information as to the specific type of care that was provided only the diagnostic code provided based on the main reason a person visited a GP/FP or psychiatrist. We saw similar utilization patterns between GP/FP and psychiatrists and to be concise, reported them together. We have now provided more details regarding provider type (visits to GP/FPs and psychiatrists) (page 11). Hospital and ED visits were beyond the scope of this paper.

4. In the results section, on page 7, line 52, there seems to be a word missing: "Mothers with [older?] children".
This sentence was no longer required in this draft and was removed (page 7).

5. In Figure 1's table, please clarify the measure reported as this is not a relative risk (those cannot be expressed as a percentage).
This was an error on our end and has been corrected. Please see page 2 - 'Tables and Figures' document.

6. In Figure 1's legend, it is mentioned that a dotted line refers to "expected", but all lines are solid.
Thank you for catching this, we have made the necessary changes to the legend as suggested and made minor visual changes to improve the figure. See page 2 – 'Tables and Figures' document.

7. Could the Authors clarify the results of observed vs expected prior to the onset of the pandemic, especially where the adjusted measures are significant (i.e., the confidence interval do not cross 0%)? This raises doubts about models for the pre-pandemic months being properly specified.
We examined the goodness of fit by calculating the proportion of observed values that fall outside the 95% CI for the fitted values. In nearly all models 100% of the observations used fell within the 95% CIs. In rare cases at most we saw up to 1.25% of the observations falling outside the 95% CI.

8. Did the Authors considered using interrupted time series analyses? I would suggest to conduct this type of analyses in the context of your study question.
Regarding the use of interrupted time series analysis, we agree that this is an alternative approach, however, we used the methodology of observed versus expected to get a better idea on how mental health utilization varied over specific points in time by comparing observed with expected over the study period.

9. Please consider adding CI bands to the figures to help the visual interpretation of the results.
We had initially included CI bands in Figure 2, however, the figure looked less legible with many overlapping lines. However, that figure has now been replaced with forest plots which include the CI bands requested, therefore this issue has been resolved. See page 3 - 'Tables and Figures' document.

Reviewer 2: Ms. Erin Graves
General comments (author response in bold)

Questions:

1. Inclusion criteria: why are women who were institutionalized (assuming this is for inpatient mental health?) be excluded if they were admitted during the study period? Wouldn't this confound what you're look at as it could be an escalation of the initial primary care visit for mental health dx? I would understand excluding women who were institutionalized in a look-back period, but I don't think it's appropriate to exclude them during the same period in which you're assessing the outcome.
The term 'institutionalized' has been replaced with more accurate terminology 'long-term care facilities' to better reflect the data protocol used to create the dataset and eliminate possible confusion with institutional psychiatric facilities, Eds and hospitals. We did not exclude persons who may have visited or been institutionalized at psychiatric facilities as long as they had a residential record in the RPDB (registered persons database). We used this definition to define what a

long-term care facility is: according to the Ministry of Health and Ministry of Long-Term Care, a long-term care home provides care and services for people who are no longer able to live independently or who require nursing care, 24-hour supervision or personal support. To be eligible to live in a long-term care home, an individual must be 18 years of age or older, have a valid Ontario Health Insurance Plan (OHIP) card, and have these care needs.

2. Why was only the material deprivation of the ON-MARG used? Given the literature on the topic, I would have thought that social deprivation (the other dimension of the scale) would also be important and of interest to the study.
Given the need for timely evidence on this topic we were unable to rely on more granular level predictors for example from the 2021 census, which to this day is not available inside the ICES data environment for this kind of analysis. Of the four dimensions of ONMARG, material deprivation was the most appropriate with respect to critical SES indicators. We also wanted to keep the study relatively streamlined and decided the other indicators would not add sufficient value to the study to be included.

3. Study population and timelines: this is unclear. You were provided with quarterly aggregated estimates? That would not result in 23 during your time period. Or you were provided with weekly rates during the 3 quarter period you're using to define your 'pre-covid' baseline period? I don't think this is a methodological concern, just additional clarity in the writing would be helpful to understand what was done. Also, the inclusion criteria for mothers does not need to be repeated here; it was well described in the earlier section - suggest removing here.
With regards to quarters, they were provided to us that way with aggregated estimates, for a total of 23 (16 pre-pandemic, 7 post-pandemic). With respect to clarity within the manuscript, we have made the changes as you suggested and removed repeated statements. See page 5 and 6 for these changes.

4. Study Outcomes: I thought OHIP used ICD-9 dx codes, not ICD-8? Please confirm.
We have removed the ICD-8 wording from the manuscript as one of the coauthors confirmed that OHIP does not use neither ICD-, 8, 9, or 10 consistently but rather a mixture of codes which are slowly updated over time. It's probably most accurate not to label them as such. All references to ICD-8 have been removed.

5. Statistical Analysis: what was your denominator for visit rates out of 1,000? Because your focus is comparing COVID to pre-COVID time periods, and the data you're using to do this changed as a function of the event, I think this needs to be really clear how your rates were calculated to ensure they address potential bias introduced by artifacts of how this data was collected while people were advised to stay home and avoid things like physician visits. Also, I would like more detail on how you coded your outcome; by nature of the regression technique described and the table of codes, I think that it's binary (y/n); however, I'd urge you think further on if a physician visit for 'marital difficulties' or 'problems with aged inlaws' is really equivalent to a physician diagnosis for bi-polar disorder? Would it be possible to use an ordinal or multinomial outcome to capture the broader 'mental health' nature of these conditions while also acknowledging the very heterogenous nature of the codes used to define your outcome? Is it reasonable that patterns could vary by severity of the disorder/type of visit, in particular in your chose population of interest: new mothers?

The manuscript should have been clearer on this point. The denominator in this study was the total number of mothers with children eligible at each quarter during the study period and numerator the number of visits. Because we used the actual number of eligible birthing mothers at each quarter, our findings reflect the true rate of utilization by this specific population within Ontario. The outcome was the number of quarterly visits by eligible birthing parents. The outcome of participants with no visits were assigned as zero and contributed to the denominator. Severity with respect to these individual conditions and the nature of the visits is beyond the scope of this paper.

6. Results: Suggest re-wording portion relating to 'mothers of younger children were younger' as this is not novel; consider the relevance and importance of your results towards clinical care or population health and frame accordingly. Further, younger mothers being more likely to live in lower income areas, is not a novel finding. It would better position your paper to position the introduction around points like this (maternal age, material deprivation, relationships that exist between both and mental health, ideally on a scale from severe to relatively benign reasons, so that the result section can be use to parse out the most relative, novel contributions of this study to the literature.
The results section has been edited and mostly rewritten to focus on the main findings in this study (pages 7-12).

7. In reading further into the results, I understand that the unit of aggregate analysis was a quarter. However, it would be helpful to understand the logic behind choosing such a large time frame, coupled with why seasonality, typically seen at more granular levels of analysis was thought to be important in the analysis? Suggest also not referring to your results by 'cohort 1 & cohort 2'. Please reframe based on the inclusion criteria or specific focus of each cohort in this section, so that the full weight of your results can be easily understood by the reader.
We agree with you that monthly intervals for the data would have been ideal, however, due to issues of privacy and remote data access, the data had to be sufficiently anonymized. ICES flagged potential privacy issues of persons becoming identifiable due to small cell counts in certain situations, when switching from one cohort to the other over time, and other similar issues. One of their solutions was to provide the data in quarters. We agree with your comments around better reframing the results/discussion around the population under study, therefore we have ensured that we used more appropriate terminology than 'cohort 1 & cohort 2'.

8. Changes by diagnosis: Were the aggregated rates denominators from the population (ie database independent) or were they out of the total visits seen in the period? Was this stratified analysis by dx code or by dx code grouping outlined in supplementary table 1? Can you comment on not seeing a change in alcohol/substance use disorder visits? This is counter to other literature on the topic; do the authors have a hypothesis on why this might be or at least why this doesn't shed doubt on the validity of their other results given we know both alcohol and illicit substance use exploded during the pandemic period?
The aggregated rates denominator were from the total population of birthing mothers under study and 'at risk' of experiencing the 'event' (e.g. mental health service use) not total visits in the period. The stratified analysis is by dx code groupings that were in the supplemental table 1. They were grouped together by ICES data analysists who curated and provided us with the final dataset for

analysis. We did see some increase in alcohol and substance abuse among birthing parents with children (6-12 years), however, the magnitude was not such that warranted extensive discussion. It is important to keep in mind that many studies on increased alcohol and substance use during covid19 were for the whole adult population and often based on surveys collected early in the pandemic and other cross-sectional data which don't give us trends over time. There was an increase from September 2020 to about May of 2021 (only Dec-2020/Feb-2021-quarter was statistically significant), which makes sense as you mentioned alcohol and substance use exploded in the first few months to about a year into the pandemic and therefore these numbers would suggest the delayed impact on seeking care for those 'excessive months' prior. However, it does not seem based on our results that there was a major sustained increase with respect to needing alcohol and substance use supports from doctors and psychiatrists and that the increase observed was not sustained over the period of the study and reached expected levels by the end of the study window, unlike other mental health outcomes such as anxiety, depression and other mood disorders.

9. Discussion: can you cite literature on why you might see think you might see a marked increase in anxiety/mood disorders in the younger and older mothers, but less in the middle years?
We are no longer reporting diagnosis types stratified by age groups. Overall mental health increases stratified by age showed increases in nearly all age groups.

10. Table 2: what are these numbers? Rates? Ns? Please clarify in the title; it's hard to understand how the trend for older moms is potentially an artifact of small sample size when your table has 64,050 under age. Label clarity could fix this, however.
We have done our best to make the table more legible. The purpose of this table was to be transparent with respect to the distribution of age (expressed within study age-bands) and material deprivation (expressed as quintiles) for birthing mothers. As for your second comment, all the potential 'artifacts of small sample size' for older moms was impacting 'cohort 1' where the sample size for older moms was n=3,161. We did not have any issues with birthing mothers 48+ within 'cohort 2' since the sample size as you mentioned (n=64,050) was sufficient for the analysis. We are no longer reporting study outcomes by age-groups stratified by material deprivation.

11. Discussion: As this data and study didn't look at ethnicity, I don't think that this should be referenced in your conclusions - there is no indication that the women in your study who used mental health resources were (or were not) from different backgrounds than those who did not - suggest sticking to the analysis and data which were completed and used in the study.
Been removed from the manuscript.

12. Methods: were the OHIP physician records used limited by physician specialty? It says earlier in the paper that visits to family physicians or psychiatrists were included, but how different physician specialties were defined is not carried through.
We have now reported findings by physician type (page 11 – manuscript) and (page 4 - tables and figures).

13. Introduction: Suggest reconsidering how your results are framed, and revising the introduction accordingly. There is quite a bit of research on mothers of younger kids and mental health, and positioning your work within that would help make the results and importance clearer.

The introduction and discussion have been updated to better situate this study within the larger context of other literature. While there is research on mothers with young children and mental health, this paper is unique in that it is the first we could identify to have looked at the mental health of birthing parents of younger and school-aged children at the population level during the first 18-months of the pandemic using this methodological approach.

Overall: This is important and well-done work. Suggest re-thinking which analysis is your primary analysis considering your most important and impactful results, and then reframing the write up to focus there. Further, more explanation of your methods and how you created the datasets/units of analysis are needed to assess the validity of your results considering the methods. The chosen methodology is sound and rigorous, but your outcome definition for the primary analysis leaves questions. Further, it is unclear (due to lack of details) on how you accounted for the decrease in overall visits due to COVID in your analysis, and if that may be introducing bias.

Thank you so much for your detailed review and constructive criticism of the manuscript. Clearly great attention and effort was made to review our work and improve the rigour and clarity of the work. We greatly appreciate your comments and suggestions, and have done our utmost best to address them. As a result, we feel the manuscript is vastly improved for reconsideration by CMAJ-open for publication.