PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

Title (Provisional)

Investigating outcomes in a substance use treatment provider: A cross-sectional comparison of Long-Acting Injectable Buprenorphine and oral Medication for Opioid Use Disorder.

Authors

Montgomery, Catharine; Abbasi, Yasir; De Silva, Devon; Gittins, Rosalind; Jones, Andrew; Van Hout, Marie-Claire

VERSION 1 - REVIEW

Reviewer	**1**
Name	**Golan, Olivia K**
Affiliation	**NORC at the University of Chicago, Chicago, IL, USA**
Date	**22-Jul-2024**
COI	**N/a**

This paper explores a very important topic, differences in outcomes between extended-release buprenorphine (XR-BUP) and oral medication for opioid use disorder (MOUD) treatments. Extended-release formulations remove the need for taking daily oral medication, with potential to improve buprenorphine treatment satisfaction, adherence, and retention. Given the low retention rates of oral MOUD treatment and the limited body of research around extended-release formulations, research on XR-BUP is urgently needed. While I applaud the authors for exploring this important topic, I have concerns about the data and approach, and the organization/clarity of the paper could be improved. My specific recommendations are outlined below.

Introduction

- The introduction could be streamlined. I recommend focusing on: 1) the effectiveness of oral MOUD in reducing overdose deaths, etc.; 2) why extended-release formulations may be advantageous (e.g., potential to improve retention, people don't have to remember to take medication daily); 3) existing research on pros & cons of XR-BUP; and 4) what is missing in existing research/aims of the study. Much of this is currently included in the introduction; I suggest removing unnecessary details and making it more concise.

Methods

- The TOPS scores are very interesting, patient-centered outcome measures! Examining changes in psychological health, physical health, quality of life, and substance use is a major strength of this study.

- The approach could be significantly improved. I believe that the authors should focus on changes in TOPS scores as the only outcomes in analytic models (i.e., not summary TOPS scores), because the findings are far less useful without considering baseline scores. However, I am not sure if this is possible with the existing data (are there enough participants with multiple assessments over time?). If authors decide to keep TOPS summary scores as outcome measures, they will need to provide better rationale to convince readers why this is a worthwhile outcome measure to study.

- Please explain how matched controls were selected. How did you decide to stratify based on sex, ethnicity, and primary substance? Also, please describe the types of oral MOUD used by the matched controls. If it is just oral buprenorphine, I'd recommend referring to this as oral buprenorphine throughout, rather than oral MOUD.

- Please provide rationale for the covariates you included in the adjusted models.

- How was treatment duration considered in analyses? (e.g., do you differentiate between people who have used XR-BUP for 1 month versus 9 months?)

- How did you handle missing data?

- The organization of the methods could be improved. I'd recommend describing the variables in paragraph form. Also, I'd recommend describing how you created the TOPs variables in the variables section, rather than data analysis. Please make sure that the headings are correct (e.g., the heading of "outcome variables of interest" seems to include other variables as well [e.g., sociodemographic characteristics]). In addition, I suggest explaining the data analysis methods for examining the TOPS substance use variables in the methods section, rather than with the results.

Results

- A large portion of the results focuses on predictors of Buvidal prescribing, which I think is much less relevant/interesting than changes in TOPs scores. Unless the authors can provide better rationale, I would remove the predictors of Buvidal prescribing from the results (Table 2, Figure 1) or just describe it briefly to provide context for the other results (it could potentially go in an Appendix). I would also describe summary TOPS scores in the description of participants, but not use it as an outcome measure in analytic models.

- A table should be provided with regression results for the adjusted analyses with changes in TOPS scores as the outcome, including sample sizes used in each model.

Discussion

- Similar to the introduction, the discussion section could be streamlined, especially if the authors make the changes to the approach suggested above. The finding that Buvidal was associated with positive changes in QoL is very interesting! I would give this finding more attention in the discussion section.

Throughout

- Please review the full paper for copyedits (e.g., spelling, consistent capitalization, remove contractions, define acronyms at first use and use the acronyms in all subsequent mentions).

Table 1

- Please include type of oral MOUD in Table 1. Including information about average dosages for oral MOUD controls could also provide important context, as doses of oral MOUD (e.g., 4mg vs 16 mg of BUP) can have a big impact on its effectiveness.

- What was included in “Other” for ethnicity, primary substance, and secondary substance?

Reviewer	**2**
Name	**Oesterle, Tyler**
Affiliation	**Mayo Clinic Rochester**
Date	**29-Oct-2024**
COI	**none**

I have significant concerns about the quality of the writing with many long run on sentences that make this paper difficult to read. I think this paper would benefit from a good editorial review. The "TOP" score is very vague and the results appear contradictory. However the conclusions appear to be excessively flattering to the product. While they appear very open about their financial support in conducting this research, their "improved outcomes over a 1-year period" is a stretch.

VERSION 1 - AUTHOR RESPONSE

Reviewer: 1		
This paper explores a very important topic, differences in outcomes between extended-release buprenorphine (XR-BUP) and oral medication for opioid use disorder (MOUD) treatments. Extended-release formulations remove the need for taking daily oral medication, with potential to improve buprenorphine	Thank you – we are pleased that Reviewer 1 sees this as an important topic and have answered the queries and amended the manuscript accordingly.	

treatment satisfaction, adherence, and retention. Given the low retention rates of oral MOUD treatment and the limited body of research around extended-release formulations, research on XR-BUP is urgently needed. While I applaud the authors for exploring this important topic, I have concerns about the data and approach, and the organization/clarity of the paper could be improved. My specific recommendations are outlined below.		
Introduction The introduction could be streamlined. I recommend focusing on: 1) the effectiveness of oral MOUD in reducing overdose deaths, etc.; 2) why extended-release formulations may be advantageous (e.g., potential to improve retention, people don't have to remember to take medication daily); 3) existing research on pros & cons of XR-BUP; and 4) what is missing in existing research/aims of the study. Much of this is currently included in the introduction; I suggest removing unnecessary details and making it more concise.	We agree with Reviewer 1 that the introduction could be more streamlined. We have significantly revised the introduction in line with Reviewer 1's suggestions and have deleted unnecessary content that does not add to the rationale for the study. We have included the additional citations and references and removed any that were no longer in the paper. We think his has made the introduction much more robust and we hope this is now acceptable for Reviewer 1.	5-9
Methods The TOPS scores are very interesting, patient-centered outcome measures! Examining changes in psychological health, physical health, quality of life, and substance use is a major strength of this study.	Thank you – we have included some more text on the reliability and validity of the TOPS after a comment from Reviewer 2. The changes can be found here:	12-13
The approach could be significantly improved. I believe that the authors should focus on changes in TOPS scores as the only outcomes in analytic models (i.e., not summary TOPS scores),	We originally included the summary scores because we felt that this gave us a better overall picture of the psychological state overall during the 1-year period. On reflection, we agree with Reviewer 1 that the summary TOPs scores are not as useful as the change	S1 & S2

because the findings are far less useful without considering baseline scores. However, I am not sure if this is possible with the existing data (are there enough participants with multiple assessments over time?). If authors decide to keep TOPS summary scores as outcome measures, they will need to provide better rationale to convince readers why this is a worthwhile outcome measure to study.	scores. As such we have moved these analyses to supplementary file S1 and we have referred the reader to this file when discussing TOP scores. For the change score analysis this reduced the N to 383 people.	
Please explain how matched controls were selected. How did you decide to stratify based on sex, ethnicity, and primary substance? Also, please describe the types of oral MOUD used by the matched controls. If it is just oral buprenorphine, I'd recommend referring to this as oral buprenorphine throughout, rather than oral MOUD.	Matched controls were selected using the following procedure. We were provided with the anonymised patient identifier (to allow future data linkage) and demographic information of 2,048 individuals who received oral MOUD. We used gender, ethnicity and primary substance of use information as stratifiers to obtain a smaller sample (which reflected the balance of these stratifiers), using the 'stratified' function from the 'splitstackshape' package in R Studio (as used in previous research. See e.g. [51]). We aimed for a similar sample size to our Buvidal sample, which would still provide us with appropriate statistical power. We then provided the patient identifiers of the stratified sample to the data controller in Via, who provided us with the TOPs data for these individuals. We were unable to request data from all 2,048 individuals due to limited resources in the substance use treatment provider data and performance team. We have changed the text on Page 10 to clarify how and why we stratified the sample. In terms of other MOUD, people were prescribed SL buprenorphine, oral buprenorphine or methadone. Methadone was the most commonly prescribed other MOUD, and we have included the percentages of the group prescribed these medications and the most common dosage for each medication.	10-11
Please provide rationale for the covariates you included in the adjusted models.	We included available demographic information in our adjusted models as this data was available to us, and we were interested in how there may be inequalities in initiation of,	8-9

	and outcomes from, treatment. We have provided a better rationale for this in the introduction (see also point below on predictors analysis).	
How was treatment duration considered in analyses? (e.g., do you differentiate between people who have used XR-BUP for 1 month versus 9 months?)	Unfortunately, we did not have any information on treatment duration for our sample, largely because we were only able to get a snapshot of one-year of data. We agree with Reviewer 1 that this was an important point and we have included this as a limitation of the analysis in the discussion.	22
How did you handle missing data?	As mentioned on Page 14, we had some missing data on Indices of Multiple Deprivation (IMD) decile, which is because some individuals included in the analyses did not have a fixed residence (due to e.g. street homelessness) and as such they did not have a valid UK post code which is used to calculate the IMD decile. Given that this data was systematically missing (i.e. it is missing as a result of measurement) it is not suitable for imputation methods ('When data are missing not at random, bias in analyses based on multiple imputation may be as big as or bigger than the bias in analyses of complete cases.' See Sterne et al., 2009 below) However, we now conduct sensitivity analyses whereby we do not include those without an IMD score in our adjusted models to allow them to be represented. Similarly, there were also small amounts of missing data for age of first use (~4%). As this was only a negligible amount of missing data on one predictor variable we decided against multiple imputation. We explain this approach on Page 13-14 and hope this clarifies how missing data was handled. Sterne J A C, White I R, Carlin J B, Spratt M, RoystonP, Kenward M G *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ,* 2009; 338 :b2393 doi:10.1136/bmj.b2393	13-14

The organization of the methods could be improved. I'd recommend describing the variables in paragraph form. Also, I'd recommend describing how you created the TOPs variables in the variables section, rather than data analysis. Please make sure that the headings are correct (e.g., the heading of "outcome variables of interest" seems to include other variables as well [e.g., sociodemographic characteristics]). In addition, I suggest explaining the data analysis methods for examining the TOPS substance use variables in the methods section, rather than with the results.	We have rearranged the methods as suggested by Reviewer 1 and we think that this section is much clearer now.	10-14
Results A large portion of the results focuses on predictors of Buvidal prescribing, which I think is much less relevant/interesting than changes in TOPs scores. Unless the authors can provide better rationale, I would remove the predictors of Buvidal prescribing from the results (Table 2, Figure 1) or just describe it briefly to provide context for the other results (it could potentially go in an Appendix). I would also describe summary TOPS scores in the description of participants, but not use it as an outcome measure in analytic models.	As suggested by Reviewer 1 above, we have moved the summary TOPs scores to the supplementary file S1. We have described how the summary TOPs scores were calculated in the methods but referred the reader to the S1 file after this (including what was formerly Figure 1 but is now supplementary Figure 2 now). For predictors, we feel that it is important to keep this analysis in the manuscript. Our rationale for this is that funding is very limited in UK treatment services and there are a limited number of people offered LAIB due to budget constraints. As such each service may prioritise certain people or groups of people for LAIB based on how successful they think that treatment may be for a particular person. Indeed, in our conversations with clinicians, they indicate that if someone is stable, they may be more likely to be offered LAIB because for the service this is a significant cost and they want to maximise the chances of recovery. However, we know that LAIB may be most cost effective when treatment stay is over 28 days, or when OUD is more severe (Marsden et al., 2023). For us this creates an inequality where people who are "responders" (i.e. they are engaged in treatment anyway), those who live in more affluent areas and have less chaotic lives, and those from ethnic majority backgrounds may be prioritised for treatment.	8-9 and files S1/S2

	The analysis on predictors of being prescribed Buvidal was intended to give us an indication of if any such inequalities exist. We have made a stronger case for this in the introduction, and also discussed this in more detail in the discussion.	
A table should be provided with regression results for the adjusted analyses with changes in TOPS scores as the outcome, including sample sizes used in each model.	We agree that including this information is important and we have now included Table 3 in the manuscript. This includes the results for adjusted analyses on TOP change scores.	18
Discussion Similar to the introduction, the discussion section could be streamlined, especially if the authors make the changes to the approach suggested above. The finding that Buvidal was associated with positive changes in QoL is very interesting! I would give this finding more attention in the discussion section.	We have reviewed and edited the discussion to make it more streamlined. This includes more focus on the improvements in quality of life, and also more of a discussion on health inequalities in initiation of LAIB. We hope this section is now clearer.	19-23
Throughout Please review the full paper for copyedits (e.g., spelling, consistent capitalization, remove contractions, define acronyms at first use and use the acronyms in all subsequent mentions).	We have been through the paper and thoroughly reviewed our language, spelling and acronyms. As part of this we changed the term and acronym we use from extended-release buprenorphine to long acting injectable buprenorphine (LAIB) throughout as this better reflects the product that is used in UK treatment services. Acronyms are not used in the abstract and article summary but are defined and used from the introduction onwards.	Throughout
Table 1 Please include type of oral MOUD in Table 1. Including information about average dosages for oral MOUD controls could also provide important context, as doses of oral MOUD (e.g., 4mg vs 16 mg of BUP) can have a big impact on its effectiveness.	We have attempted to summarise the information on treatment more concisely. However this is compounded by the complexity of the sample and their needs, which makes it difficult for us to display the data in a meaningful way. For the majority of individuals in the dataset, they were not on one-specific dose (or in the MOUD group, one type of MOUD) over the course of the year, as such it isn't possible to provide average doses. Therefore, we believe this information could be misleading, rather than adding to our understanding. We have attempted to include information on the most commonly prescribed	11

	medications over the year in each group, which we think is the best way to summarise this.	
What was included in "Other" for ethnicity, primary substance, and secondary substance?	The ethnicity data collected in routine data in UK National Health Service (NHS) and substance use treatment services is required to conform to the very broad categories of the UK census (2021). These 5 categories are: Asian or Asian British (including Indian, Pakistani, Bangladeshi, Chinese and any other Asian background); Black, Black British, Caribbean or African)including Caribbean, African and any other Black, Black British, or Caribbean background); Mixed or multiple ethnic groups (including White and Black Caribbean, White and Black African, White and Asian and any other mixed or multiple ethnic background); White (including English, Welsh, Scottish, Northern Irish, British, Irish, Gypsy or Irish Traveller, Roma and any other White background); and Other ethnic group (including Arab and any other ethnic group). We have included this information in Table 1 to show the numbers of people in the other ethnic groups, though due to the small number of people in these groups we still use the binary "White British vs. other" variable in our analyses.	15
Reviewer: 2		
I have significant concerns about the quality of the writing with many long run on sentences that make this paper difficult to read.	We have significantly revised all block text sections of the main manuscript file (strengths and limitations, introduction, methods, discussion) and removed any long or vague sentences. We hope this is now suitable for reviewer 2.	Throughout
I think this paper would benefit from a good editorial review.	We have revised the paper to ensure rigorous editorial review. This has included revising the strengths and limitations in line with comments from the editor, restructuring and revising the introduction in line with suggestions made by Reviewer 1, revising the methods to include more detail on the measures we used and revising the discussion to ensure clarity in statement of implications. In addition to this, we have proofed the whole manuscript and believe it flows better after these revisions.	Throughout

The "TOP" score is very vague and the results appear contradictory.	We have included more description of the TOP tool in the method section. The TOP is a tool used in all UK treatment settings since 2007 and forms the basis of the UK National Drug Treatment Monitoring System database. It has been shown to be a reliable and valid tool for assessing treatment outcomes and is a rich data source because its use is mandatory so every person with OUD engaged with treatment services will have TOP data (usually multiple assessments depending on stage and length of treatment). In addition to the amendments made to the method, we have also included a link to the TOP on the NDTMS website in this section. We hope this is now less vague for Reviewer 2.	12-13
However the conclusions appear to be excessively flattering to the product. While they appear very open about their financial support in conducting this research, their "improved outcomes over a 1-year period" is a stretch.	We have reviewed and changed the language throughout to reflect that there were significant changes in self-reported quality of life rather than significantly improved outcomes. We have also moved one of the analyses that was investigating summary scores for the TOP to a supplementary file, so this has made any results relating to potential effects of Buvidal (i.e. changes rather than summaries) clearer. We hope this is now acceptable for Reviewer 2.	Throughout and S1-S2

VERSION 2 - REVIEW

Reviewer	**1**
Name	**Golan, Olivia K**
Affiliation	**NORC at the University of Chicago, Chicago, IL, USA**
Date	**10-Dec-2024**
COI	

The authors have addressed my concerns. My only remaining recommendation is to further streamline the introduction. Some paragraphs contain up to 11 sentences, which may be difficult for readers to digest.

VERSION 2 - AUTHOR RESPONSE

We have split up the two 10-sentence paragraphs in to 2 paragraphs and highlighted yellow in the tracked file where the paragraphs have been split.