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)	Demographic and clinical factors associated with different antidepressant treatments: a retrospective cohort study design in a UK psychiatric healthcare setting
AUTHORS	Fernandes, Andrea; Chandran, David; Khondoker, Mizanur; Dewey, Michael; Shetty, Hitesh; Dutta, Rina; Stewart, Robert

VERSION 1 – REVIEW

REVIEWER	Mythily Subramaniam Institute of Mental Health, Singapore
REVIEW RETURNED	30-Mar-2018

GENERAL COMMENTS	The authors have investigated the association between clinical correlates and antidepressant use in a psychiatric healthcare setting using a retrospective cohort study design. The study employs interesting methods to extract the data . Few points for consideration: 1. While the authors state in the title that have investigated the association between clinical correlates and antidepressant use, to be precise they have investigated the factors that are associated with different antidepressant treatments in a specific setting. 2. Terms like secondary care are not immediately clear to someone who is not part of the system and need to be clarified further. 3. In the introduction the authors state that "Studies profiling antidepressant prescription in secondary care can highlight factors that play a key role in managing treatment-resistant depression and individuals with exacerbated symptom profiles". However in the methodology it become apparent that data on treatment resistance and even symptom severity are not well captured. 4. How does this study intend to inform policy and practice as stated in the introduction? I suggest that the authors rewrite the introduction keeping in mind what this study intends to do. 5. Many of the factors identified as playing an important role in prescription of anti-depressants cannot be explored using the current design and thus, I again suggest that the introduction be rewritten to make it relevant to this study. 6. Were there any patients who did not meet any of the diagnostic criteria specified by the authors who were prescribed anti-depressants? Why were these diagnoses chosen? 7. How many patients were prescribed more than 2 antidepressants? 8. What are IAPT services? 9. How was area level deprivation score determined? 10. How did diagnostic codes serve as proxy indicators of severity of depression? How did this correlate with mention of symptoms as stated in the last paragraph of page 6? 11. As regards symptom severity do the authors have any reference

	for using the criteria as specified by them - "Individuals with zero or one mention of a symptom in the 12-month period prior to the observation frame were considered to be undergoing a milder experience of the symptom compared to individuals who had two or more mentions of the symptom in the same period". In table 2 they there is an Others category which is not specified in any manner. 12. In the discussion the authors state, " patients on past non-antidepressant and antidepressant medication are more likely to be on newer antidepressants such as mirtazapine and venlafaxine or a combination of low toxicity antidepressants". Is this correct? 13. The discussion needs to be more robust and the authors should explain their findings in further detail.

REVIEWER	Rosalinda V. Ignacio Department of Psychiatry, University of Michigan and Center for Clinical Management Research, Ann Arbor Veterans Affairs Healthcare System Ann Arbor, Michigan United States
REVIEW RETURNED	19-Apr-2018

GENERAL COMMENTS	Overall Comments This is an important study that examines predictors of antidepressant use in patients treated in a secondary mental health setting. The use of text mining algorithms to gather data on antidepressant use from electronic health records is innovative. I have some reservations regarding the statistical method used in building the model, and other comments described below. If the authors can address these, I feel this is a worthwhile publication that will contribute to health services research. Abstract State setting (psychiatric healthcare) in conclusion for context. Association with suicidality was one of many significant results and only among adults in the study. Include association with past medication use and symptoms in statement of conclusion. Line 10 - insert "(RRR)" after Relative risk ratio to introduce acronym. Inclusion Criteria Indicate how many patients were excluded, that is, N (%) of patients who did not receive any antidepressants, N (%) of patients who received 3 or more antidepressants. Describe if the excluded patients are similar in characteristics to patients in the study. Discuss implications of excluding these group of patients. Line 50 - delete "on or before the 31st of August 2015", already stated in line 44 Clinical Determinants included in the Study Covariate data pertains to data recorded in the 12 months prior to the observation window. Were there patients in the study who did not have data in the 12 months prior? Were they excluded in the study? If so, describe in the inclusion criteria section. These patients will have missing data in the covariates. How were missing data handled? Demographic variables Race instead of ethnicity, for categories used in the study: White, Other What is area-level deprivation score? Please define. Explain how latest diagnostic code served as a proxy indicator of

severity of depression.
What is IAPT services? What does IAPT acronym stand for? Introduce in first use.

Statistical Analysis
Line 35 - Chi-square instead of Chi-squared
Line 48 - association instead of contribution
The statistical approach to build the model was based on the significance of the predictors in a sex- and age-adjusted model using p<.25. This approach may miss important covariates that may turn out to be significant covariates in a multivariable model. For instance, using the present approach, race, marital status, depression severity, deprivation score, and length of spell were not included as predictors in the final model. These are important covariates, especially depression severity, for predicting antidepressant use.
I suggest to run a multivariable model with all predictors in the model to begin with, then maybe drop covariates that are not significant or those that may not be clinically important in the study. You may describe the results of this alternative approach as a sensitivity analysis to strengthen the paper, especially if you end up with the same set of predictors.
Explain why 26 years old was used as cut-off for adult group.

Results
Line 23 - race instead of ethnicity
Line 30 - more instead of majority
Age was also a significant predictor in both models - all patients, adults group. Include this in the results paragraph from table 3 and table 4.

Discussion
Line 41 – consists instead of consistent; individuals instead of individuals'
Last paragraph, line 30 – drop individual

Table 1
Include median age for each group
Race instead of Ethnicity
Line 18 - Insert line to say "Area Deprivation Score"
In Total column, add column %'s in parentheses for easier comparison with the groups.

Table 2
Depression severity = Other – what does this mean? Need to define this.
Lines 15, 18, 21, 25, 29, 31 - Drop the word "Past" since the heading already indicates this.

Table 3
Lines 23, 25, 27, 29 - change to ">=2 mentions" instead of ">2 mentions"

Table 4
Lines 27, 29, 31, 32 - change to ">=2 mentions" instead of ">2 mentions"

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1 COMMENTS

Reviewer 1 Comment 1: While the authors state in the title that have investigated the association between clinical correlates and antidepressant use, to be precise they have investigated the factors that are associated with different antidepressant treatments in a specific setting.

Response: Thank you for this comment. We have now changed the title of the manuscript to reflect what the analysis is set out to do and in accordance to editorial requirements (see above).

Reviewer 1 Comment 2: Terms like secondary care are not immediately clear to someone who is not part of the system and need to be clarified further.

Response: Thank you. We have now followed this up with a definition of secondary care.

The following comments have been answered in one response

Reviewer 1 Comment 3: In the introduction the authors state that "Studies profiling antidepressant prescription in secondary care can highlight factors that play a key role in managing treatment-resistant depression and individuals with exacerbated symptom profiles". (see highlighted sentence) However, in the methodology it become apparent that data on treatment resistance and even symptom severity are not well captured.

Reviewer 1 Comment 4: How does this study intend to inform policy and practice as stated in the introduction? I suggest that the authors rewrite the introduction keeping in mind what this study intends to do.

Reviewer 1 Comment 5: Many of the factors identified as playing an important role in prescription of anti-depressants cannot be explored using the current design and thus, I again suggest that the introduction be rewritten to make it relevant to this study.

Response: Thank you for these comments which are very helpful. We accept these criticisms and have extensively re-written the introduction to be clearer on the aims and rationale behind the study.

Reviewer 1 Comment 6: Were there any patients who did not meet any of the diagnostic criteria specified by the authors who were prescribed anti-depressants? Why were these diagnoses chosen?

Response: Research in the field of profiling antidepressant use in secondary psychiatric care is sparse. While antidepressants are prescribed in some non-depressive disorders, we feel that profiling antidepressant use including these other disorders would be another research question as symptoms may differ, as may indications for antidepressant use and co-prescribing of other non-antidepressant drugs. Hence, we selected patients with depressive disorders as what we felt would be a recognised and generalisable category (e.g. as defined by NICE guidance for treatment of depressive disorders). Considering reviewer comments, we have now added this detail in the introduction and discussion section.

Reviewer 1 Comment 7: How many patients were prescribed more than 2 antidepressants?

Response: There were 3991 patients who met the inclusion criteria. According to the medication extraction algorithm, 1936 patients were not on antidepressants during the evaluation period, while

1561 were on single or dual antidepressant therapy, and 494 received three or more antidepressants. We have now added text in response to this comment.

Reviewer 1 Comment 8: How was area level deprivation score determined?

Response: Thank you for your comment. We have now defined area-level deprivation score and how it was determined.

Reviewer 1 Comment 9: What are IAPT services?

Response: This has now been defined.

Reviewer 1 Comment 10: How did diagnostic codes serve as proxy indicators of severity of depression? How did this correlate with mention of symptoms as stated in the last paragraph of page 6? In table 2 they there is an Others category which is not specified in any manner.

Response: Formal ICD-10 diagnoses of depression ascertain this as mild, moderate, severe with psychotic symptoms and severe without psychotic symptoms. Hence we classified milder and more severe diagnoses accordingly. We term this "diagnosis-derived depression severity" and have added text to the Methods so that its derivation is described in more detail and more explicitly. In the event, although diagnostic severity was highly correlated with the depressive symptoms included in this analysis, unadjusted analysis of depressive severity with antidepressant use did not approach statistical significance and this variable was therefore not taken forward to the multinomial analysis.

Reviewer 1 Comment 11: As regards symptom severity do the authors have any reference for using the criteria as specified by them - "Individuals with zero or one mention of a symptom in the 12-month period prior to the observation frame were considered to be undergoing a milder experience of the symptom compared to individuals who had two or more mentions of the symptom in the same period".

Response: We accept that the wording for this explanation was not sufficiently clear and we have amended this in the relevant part of the Methods. Essentially, we defined the presence of symptoms on the basis of at least two mentions simply in order to avoid false positive classifications arising from the natural language processing algorithms.

Reviewer 1 Comment 12: In the discussion the authors state, "patients on past non-antidepressant and antidepressant medication are more likely to be on newer antidepressants such as mirtazapine and venlafaxine or a combination of low toxicity antidepressants". Is this correct?

Response: This is technically correct; however, we realise, after reading the comment, that it was not clearly worded in the manuscript. We have now sought to clarify the text.

Reviewer 1 Comment 13: The discussion needs to be more robust and the authors should explain their findings in further detail.

Response: Thank you for this comment. We have extensively revised and extended the Discussion section in response to this and other comments and hope that it now reads more clearly.

REVIEWER 2 COMMENTS

Reviewer 2 Comment 1: State setting (psychiatric healthcare) in conclusion for context.

Response: This change has been made.

Reviewer 2 Comment 2: Association with suicidality was one of many significant results and only among adults in the study. Include association with past medication use and symptoms in statement of conclusion.

Response: This change has been made.

Reviewer 2 Comment 3: insert "(RRR)" after Relative risk ratio to introduce acronym.

Response: This change has been made.

Reviewer 2 Comment 4: Indicate how many patients were excluded, that is, N (%) of patients who did not receive any antidepressants, N (%) of patients who received 3 or more antidepressants. Describe if the excluded patients are similar in characteristics to patients in the study. Discuss implications of excluding these group of patients.

Response: See response to **Reviewer 1 Comment 7** regarding the patients who received 3 or more antidepressants.

With regards to patients who did not receive any antidepressants we have now included the following table, comparing patients on antidepressants with patients not on antidepressants, in the manuscript as part of the results.

Patient Characteristics	On single or dual therapy antidepressants 1561	Not on antidepressants 1936	
Gender			
Female	1229	1295	$x^2 = 0.07$, df = 1, p = 0.80
Male	706	759	

Mean age	44.3 years	40.1 years	t = -7.6, df = 3908, p < 0.001
Marital status			
Single	1091 (56.3%%)	1302 (63.4%)	x^2 = 0.02, df = 2, p = 0.999
Married	456 (23.5%)	319 (15.5%)	
Other	389 (20.0%)	434 (21.1%)	
Area-level deprivation score			
2.25 – 22.3 (least deprived)	631 (33.0%)	580 (28.7%)	x^2 = 9.1, df = 2, p < 0.05
22.4 – 42.3	1099 (57.6%)	1224 (60.6%)	
42.4 – 62.3 (most deprived)	179 (9.4%)	215 (10.6%)	
Race			
White	1220 (63.9%)	864 (42.1%)	x^2 = 10.5, df = 1, p < 0.01
Other	716 (36.1%)	1191 (57.9%)	
Severity of depression			
Mild	383 (24.5%)	460 (28.3%)	x^2 = 0.007, df = 2, p = 0.99
Moderate-Severe	845 (54.1%)	845 (48.4%)	
Unspecified	333 (21.3%)	379 (23.3%)	

Reviewer 2 Comment 5: Were there patients in the study who did not have data in the 12 months prior? Were they excluded in the study? If so, describe in the inclusion criteria section. These patients will have missing data in the covariates. How were missing data handled?

Response: Missing data were minimal due to the text mining techniques used to data extract. However, we have added text to the statistical analysis section to describe this issue in more detail.

Reviewer 2 Comment 6: What is area-level deprivation score? Please define.

Response: Thank you for your comment. We have now defined area-level deprivation score and how it was determined.

Reviewer 2 Comment 7: Explain how latest diagnostic code served as a proxy indicator of severity of depression.

Response: See Response to **Reviewer 1 Comment 10**. As mentioned, we have added text to the Methods to clarify this issue.

Reviewer 2 Comment 8: What is IAPT services? What does IAPT acronym stand for? Introduce in first use.

Response: See response to **Reviewer 1 Comment 8**. We have now explained this.

Reviewer 2 Comment 9: The statistical approach to build the model was based on the significance of the predictors in a sex- and age-adjusted model using $p < 0.25$. This approach may miss important covariates that may turn out to be significant covariates in a multivariable model. For instance, using the present approach, race, marital status, depression severity, deprivation score, and length of spell were not included as predictors in the final model. These are important covariates, especially depression severity, for predicting antidepressant use.

I suggest to run a multivariable model with all predictors in the model to begin with, then maybe drop covariates that are not significant or those that may not be clinically important in the study. You may describe the results of this alternative approach as a sensitivity analysis to strengthen the paper, especially if you end up with the same set of predictors.

Response: We accept this helpful comment. We have now run a sensitivity analysis as suggested. The results are not different to our current results. Results from the sensitivity analysis are presented in Appendix tables. Depression severity was excluded from initial analysis as it is highly correlated with the symptoms data variables (i.e. psychotic, somatic, affective and cognitive symptoms) and we are more interested in how these symptoms are associated with antidepressant use.

Reviewer 2 Comment 10: Explain why 26 years old was used as cut-off for adult group.

Response: We have added text to describe this.

Reviewer 2 Comment 11: Age was also a significant predictor in both models - all patients, adults group. Include this in the results paragraph from table 3 and table 4.

Response: Thank you for your comment. We have now added text to the Results on these associations.

Reviewer 2 Comment 12: Depression severity = Other – what does this mean? Need to define this.

Response: See response to **Reviewer 1 Comment 10**. We have added text to the Methods on this variable. Also considering this comment Depression Severity: Other is now explained and also renamed to Depression Severity: Unspecified.

In response to further comments from Reviewer 2 below, all requested text changes have been made. The only exception concerns the request to change 'ethnicity' to 'race' (in italics below). We believe 'ethnicity' is the more correct and acceptable term for the categories we have applied in this analysis (e.g. the source field in the electronic health record refers to ethnic group); however, we are happy to accept an editorial decision if this is contrary to journal house style.

Methods
Line 50 - delete "on or before the 31st of August 2015", already stated in line 44

Results
Line 23 - race instead of ethnicity
Line 30 - more instead of majority
Age was also a significant predictor in both models - all patients, adults group. Include this in the results paragraph from table 3 and table 4.

Discussion
Line 41 – consists instead of consistent; individuals instead of individuals'
Last paragraph, line 30 – drop individual

Table 1
Include median age for each group
Race instead of Ethnicity
Line 18 - Insert line to say "Area Deprivation Score"
In Total column, add column %'s in parentheses for easier comparison with the groups.

Table 2
Depression severity = Other – what does this mean? Need to define this.
Lines 15, 18, 21, 25, 29, 31 - Drop the word "Past" since the heading already indicates this.
Lines 23, 25, 27, 29 - change to ">=2 mentions" instead of ">2 mentions"

Table 4
Lines 27, 29, 31, 32 - change to ">=2 mentions" instead of ">2 mentions"

VERSION 2 – REVIEW

REVIEWER	Mythily Subramaniam Institute of Mental Health, Singapore
REVIEW RETURNED	08-Jul-2018

GENERAL COMMENTS	The authors have adequately addressed my comments and the article reads well now.

REVIEWER	Rosalinda V. Ignacio Department of Psychiatry, University of Michigan and Center for Clinical Management Research, Ann Arbor Veterans Affairs Healthcare System, Ann Arbor, Michigan, United States

REVIEW RETURNED	29-Jun-2018

GENERAL COMMENTS	Overall Comments The authors have addressed my major concerns and suggestions other than minor comments below. I believe the paper has been strengthened and will be a worthwhile publication that will contribute to health services research Statistical Analysis, 4th sentence In order to build a representative model of which correlates predict antidepressant use in secondary care, decisions to include variables in the final model were guided by the association (instead of contribution which has a specific meaning in statistics) of each variable in an initial model including age and gender as covariates. Table 2 and Table 3 In Total column, add column %'s in parentheses for easier comparison between the sub-groups. This was also suggested in the first review.