PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Intention-to-Treat Analyses and Missing Data Approaches in Pharmacotherapy Trials for Alcohol Use Disorders
AUTHORS	Del Re, AC; Maisel, Natalya; Blodgett, Janet; Finney, John

VERSION 1 - REVIEW

REVIEWER	Nich, Charla Yale University
REVIEW RETURNED	02-Aug-2013

GENERAL COMMENTS	"Intention to Treat Analyses and Missing Data Approaches in Pharmacotherapy Trials for Alcohol Use Disorders?" addresses the problem of lack of uniformity in defining and reporting on intention to treat using a literature review of a decade of alcohol disorder treatment publications. This work demonstrates an impressive undertaking in an area of great ambiguity. The basic questions involved in ITT are both in the analysis and the reporting. (1) Does the analysis include at least one set of data in which all randomized study participants are included? And, (2) Does the report declare the number of randomized participants included in the analysis? I. My major concern is in the description of mixed effects models. The mixed effect regression model (or random effect regression model or hierarchical linear model) includes the possibility that all study participants can be included if they have 1 or more data points. For a study with either monotone or intermittent missingness, the mixed effect regression model interpolates values by creating a longitudinal trajectory for all participants based on an iterative process that utilizes the participant level data and an aggregate level data, such as treatment condition. The models are very flexible, and do not require imputation, a process in which the researcher actively assigns values to replace missing data. a) On page 9, the authors discuss various means of handling missing data, but neglect to mention interpolation. b) On page 13, the authors refer to "using a statistical imputation strategy…, such as a mixed effects model." As stated above, there is a notable difference between imputation and interpolation. c) In Table 2, the description for column 3 is "…imputed missing data, e.g. Mixed model." See above. II. Table 3 presents the findings for an evaluation of change in true ITT over time. It would be great to see a visual representation of either the raw data or the estimated values for this analysis. For consistency, the sample size should be reported as well. II. Also in the footnote of table 2, the authors state that "it was unclear whether an ITT analysis was conducted or not for 15 analysis." This is not consistent with the information in the table. Finally, it might be helpful to evaluate the studies in the context of the overall study design. How many of these studies were

	longitudinal? If some studies were not longitudinal, it would be interesting to look at the tables for the subset of longitudinal studies. The LOCF, for instance, might be showing an increase in the probability of use over time simply because there are more studies with a longitudinal design.

REVIEWER	Liu, Siwei UC Davis, Human Ecology
REVIEW RETURNED	06-Aug-2013

GENERAL COMMENTS	bmjopen-2013-003464 Intention-to-Treat Analyses and Missing Data Approaches in Pharmacotherapy Trials for Alcohol Use Disorders In this study, the authors investigated 165 articles published between 1970 and 2009 on randomized controlled trials for alcohol use disorders, and examined how intention-to-treat analyses (ITT) were used and how missing data were handled in these studies. They found that many studies claiming to use ITT did not actually use this strategy, but a decent proportion of studies making no claim of using ITT actually did perform ITT analyses, reflecting a large discrepancy in the definition of ITT. No change was found in ITT used over time. However, the use of sophisticated missing data strategies, such as multiple imputation, increased over time. The authors are to be commended for providing a clear introduction to ITT analyses and alternative strategies, and examining a large body of literature in alcohol use disorders. They conducted a thorough literature search, which allows evaluation of ITT use in alcohol related research over the past few decades. The biggest limitation of this study is a lack of comparison in effect sizes between studies using ITT and those not using ITT. This piece of information would greatly enhance our understanding of whether a certain strategy is related to bias in the results. This manuscript would also be strengthened if the authors could highlight the importance of their investigation early in the manuscript. For example, the authors should state up front what the purposes are for the current investigation and how this manuscript contributes to research on alcohol use disorders. Minor points: Page 6-7: It would be helpful to provide a list of all reasons for excluding a study (e.g., open-label trials). This can be included in an appendix.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

My major concern is in the description of mixed effects models. The mixed effect regression model (or random effect regression model or hierarchical linear model) includes the possibility that all study participants can be included if they have 1 or more data points. For a study with either monotone or intermittent missingness, the mixed effect regression model interpolates values by creating a longitudinal trajectory for all participants based on an iterative process that utilizes the participant level data and an aggregate level data, such as treatment condition. The models are very flexible, and do not require imputation, a process in which the researcher actively assigns values to replace missing data.

a) On page 9, the authors discuss various means of handling missing data, but neglect to mention interpolation. b) On page 13, the authors refer to “using a statistical imputation strategy…, such as a

mixed effects model." As stated above, there is a notable difference between imputation and interpolation. c) In Table 2, the description for column 3 is "…imputed missing data, e.g. Mixed model." See above.

We have now updated the manuscript in several places by changing "imputation" to "imputation or interpolation" to more accurately describe the handling of missing data from mixed models.

II. Table 3 presents the findings for an evaluation of change in true ITT over time. It would be great to see a visual representation of either the raw data or the estimated values for this analysis. For consistency, the sample size should be reported as well.

We have now included a Figure (Figure 1) to convey this relationship graphically.

II. Also in the footnote of table 2, the authors state that "it was unclear whether an ITT analysis was conducted or not for 15 analysis." This is not consistent with the information in the table.
This sentence has now been deleted from the footnote. It was placed there in an earlier version of the Table, prior to finalizing the analyses.

Finally, it might be helpful to evaluate the studies in the context of the overall study design. How many of these studies were longitudinal? If some studies were not longitudinal, it would be interesting to look at the tables for the subset of longitudinal studies. The LOCF, for instance, might be showing an increase in the probability of use over time simply because there are more studies with a longitudinal design.

Although we had not specifically coded for type of statistical method used, we were able to determine whether the study had assessment points beyond treatment termination. When added as a control variable, the results were virtually unchanged. We have added the following sentence to the Results section (pg 15) that describes longitudinal methods not being confounded with increasing use of certain missing data strategies over time : " To explore whether increasing use of certain missing data strategies over time was confounded with longitudinal methods being increasingly employed, a proxy dummy control variable (0=only end-of treatment assessment, 1= posttreatment and follow-up assessment(s))s was added to the analyses; the results were virtually unchanged."

Reviewer 2
The biggest limitation of this study is a lack of comparison in effect sizes between studies using ITT and those not using ITT. This piece of information would greatly enhance our understanding of whether a certain strategy is related to bias in the results.

We agree with the Reviewer that having effect size (ES) information would valuable. However, the focus of this study was on describing the existing state of ITT strategies used in pharmacotherapy trials for AUD, which is of substantial interest in its own right. We believe that including ES information in this study would detract from that focus. Further, ES information could not be adequately addressed in a manuscript of the current length, as there are several complexities involved that would require substantially more journal space and discussion. For example, several studies conducted ITT for some outcomes (imputing "failure" for missing data with dichotomous outcomes) and not others (sample followed-up analyses for continuous outcomes). Thus, a straightforward moderator analysis comparing studies with and without ITT will not give any definitive conclusions, given that between-study differences in type of outcome variables (e.g., dichotomous versus continuous), samples, setting, treatment duration and other study characteristics will influence between-study outcomes.

This manuscript would also be strengthened if the authors could highlight the importance of their investigation early in the manuscript. For example, the authors should state up front what the purposes are for the current investigation and how this manuscript contributes to research on alcohol use disorders.

We have now updated the final paragraph at the end of the Introduction: " It is unknown to what degree ITT strategies are being employed in pharmacotherapy for alcohol use disorders. One aim of this review was to determine if there are discrepancies between the types of analyses that researchers stated they used and those they actually used, based on information in reports of a large

pool of randomized controlled trials of pharmacotherapy for alcohol use disorders published between 1970 to 2009. A second aim was to describe the use of different missing data strategies in studies in which true and modified ITT analyses were and were not conducted. The final aim was to determine whether the use of different data analytic approaches and certain types of missing data approaches (e.g., multiple imputation) has increased over time while the use of others has decreased."

Minor points:
Page 6-7: It would be helpful to provide a list of all reasons for excluding a study (e.g., open-label trials). This can be included in an appendix.

To conserve journal space, we now refer readers to Maisel et al. (2013) for a descrption of the exact exclusion criteria, as the same criteria were used the present review. On page 7: " The details of inclusion/exclusion criteria can be found in Maisel et al.12 "

VERSION 2 – REVIEW

REVIEWER	Nich, Charla Yale University
REVIEW RETURNED	19-Sep-2013

GENERAL COMMENTS	"Intention to Treat Analyses and Missing Data Approaches in Pharmacotherapy Trials for Alcohol Use Disorders" is an impressive and timely manuscript that details the current use and definition of the "intention to treat" strategy in randomized clinical trials by summarizing the techniques used in 39 years of pharmacotherapies publications in alcohol treatment research. Overall, the piece is succinct and covers the essential material. However, the authors use the terms "imputed" and "interpolated" interchangeably. There are many types of imputations, including last observation carried forward (LOCF), that are fraught with error. Manually imputing a value, such as in LOCF, artificially reduces the variance and introduces bias that may be associated with between group differences in missingness. Conversely, interpolation, through methods such as hierarchical linear modeling or random effects regression modeling does not introduce this bias, and is an acceptable form of intention to treat, when intermittent or monotone missing data appear in randomized clinical trials (RCTs). There are other methods of addressing missingness, such as multiple imputation for missingness strategies or Expectation-Maximization imputation, that "fill in" missing values using the strength and values of the existing dataset. Both interpolation and multiple imputation for missing data strategies are state of the art in randomized clinical trials research – and should notcategorized with "imputation," which may include casual insertion of values to replace missing values. The primary concern for addressing intention to treat in RCTs is: 1) Clearly stating the sample size used for each analysis a. Stating whether there was any analysis that used all randomized subjects without (manual) imputation of missing data

	b. Stating the amount (either raw number of %) of missing data Specific and minor comments: 1) In the introduction (on page 4, in the second full paragraph), the authors state "…ITT requires either no attrition or some imputation procedure to account for any missing data." Rather than "some imputation procedure," "ITT requires either no attrition or a strategy to handle missing data" is more valid. (Preferably, interpolation or multiple imputation would be used to reduce bias.) 2) On page 6, in the second full paragraph, that authors state "… " PubMed complete cases" approach also referred to as a 'complete cases' to approach." This must be an oversight. 3) The categories of "sample analyzed" reported on page 8 and in table 1 are useful, but there should be clarification of the overlap, if any, between (2) full random sample and (3) random sample followed. By way of example, in some RCTS, attempts are made to follow up on all randomized participants, but not all participants are actually interviewed. With a hierarchical linear model of repeated measures that evaluates slopes from baseline, all participants who provided any data are included. Therefore, even if the participant is not followed, as long as there is a baseline value, the case is included in the analysis. This qualifies as intention to treat, but might fall under the current category (3) and be deemed"not true ITT."

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:

"Intention to Treat Analyses and Missing Data Approaches in Pharmacotherapy Trials for Alcohol Use Disorders" is an impressive and timely manuscript that details the current use and definition of the "intention to treat" strategy in randomized clinical trials by summarizing the techniques used in 39 years of pharmacotherapies publications in alcohol treatment research.
Overall, the piece is succinct and covers the essential material. However, the authors use the terms "imputed" and "interpolated" interchangeably. There are many types of imputations, including last observation carried forward (LOCF), that are fraught with error. Manually imputing a value, such as in LOCF, artificially reduces the variance and introduces bias that may be associated with between

group differences in missingness. Conversely, interpolation, through methods such as hierarchical linear modeling or random effects regression modeling does not introduce this bias, and is an acceptable form of intention to treat, when intermittent or monotone missing data appear in randomized clinical trials (RCTs). There are other methods of addressing missingness, such as multiple imputation for missingness strategies or Expectation-Maximization imputation, that “fill in” missing values using the strength and values of the existing dataset. Both interpolation and multiple imputation for missing data strategies are state of the art in randomized clinical trials research – and should not categorized with “imputation,” which may include casual insertion of values to replace missing values.
The primary concern for addressing intention to treat in RCTs is:
1) Clearly stating the sample size used for each analysis
a. Stating whether there was any analysis that used all randomized subjects without (manual) imputation of missing data
b. Stating the amount (either raw number of %) of missing data

We appreciate the reviewers detailed clarification of imputation versus interpolation. We have now updated the manuscript in several places by appropriately changing "imputation or interpolation" to "interpolation".

Specific and minor comments:
1) In the introduction (on page 4, in the second full paragraph), the authors state “…ITT requires either no attrition or some imputation procedure to account for any missing data.” Rather than “some imputation procedure,” “ITT requires either no attrition or a strategy to handle missing data” is more valid. (Preferably, interpolation or multiple imputation would be used to reduce bias.)

This sentence has been updated on pg 4: " Currently, no universally accepted definition of ITT exists, although many researchers consider it to require either no attrition or a strategy to handle missing data."

2) On page 6, in the second full paragraph, that authors state “… “complete cases” approach also referred to as a ‘complete cases’ to approach.” This must be an oversight.

Indeed, this is an oversight. The duplicated part of the sentence has now been removed and now reads: "...“sufficient dose” of treatment are used, to those in which only participants who fully completed treatment are included [also referred to as a ‘complete cases’ approach; 2]."

3) The categories of “sample analyzed” reported on page 8 and in table 1 are useful, but there should be clarification of the overlap, if any, between (2) full random sample and (3) random sample followed. By way of example, in some RCTS, attempts are made to follow up on all randomized participants, but not all participants are actually interviewed. With a hierarchical linear model of repeated measures that evaluates slopes from baseline, all participants who provided any data are included. Therefore, even if the participant is not followed, as long as there is a baseline value, the case is included in the analysis. This qualifies as intention to treat, but might fall under the current category (3) and be deemed “not true ITT.”

This is an important clarification. There is no overlap between the "Full random sample" and "Random sample followed up" categories. We have now included the following sentence on pg 9: " Note there is

no overlap between categories 1 ("Full random sample") or 2 ("Full random sample (likely)") and "Random sample followed-up." "