PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Meta-analysis of anti- Saccharomyces cerevisiae antibodies as diagnostic markers of Behçet's disease with gastrointestinal involvement
AUTHORS	Cheng, Linlin; Li, Liubing; Liu, Chenxi; Yan, Songxin; Li, Yongzhe

VERSION 1 – REVIEW

REVIEWER	Jin Lin Department of Rheumatology, First Affliated Hospital, College of Medicine, Zhejiang University
REVIEW RETURNED	22-Nov-2019

GENERAL COMMENTS	1 The English language needs linguistic modification. 2 The intruction and discussion parts need refine.

REVIEWER	Luis J Jara Quezada HOSPITAL DE ESPECIALIDADES "DR ANTONIO FRAGA MOURET", INSTITUTO MEXICANO DEL SEGURO SOCIAL. MEXICO CITY, MEXICO
REVIEW RETURNED	08-Dec-2019

GENERAL COMMENTS	This is an interesting meta analysis that aims to assess the diagnostic value of anti-saccharomyces cerevisiae antibodies (ASCA) in Behçet's disease (BD) patients and explore their relationship with other autoimmune diseases(AID).In relation to this topic, I have the following observations and questions: INTRODUCTION. This section should be reduced and emphasis should be placed on studies related to this research. 2. METHODS. Language can be a barrier to knowledge. Why did the authors exclude studies that were not written in English?. This is a study of antibodies, therefore it is important to define in this section which isotype of antibodies were determined in the studies: IgG, igA or both. Is the method to detect these antibodies standardized or are there variations?. What is the sensitivity and specificity of the method to detect these antibodies? Is the same method used in all studies?. With what method did they evaluate the publication bias of the articles?. DISCUSSION. Standardization of tests to detect antibodies should be discussed

REVIEWER	Roger Ho National University of Singapore Singapore
REVIEW RETURNED	12-Jan-2020

GENERAL COMMENTS	Thank you for inviting me to provide statistical review for the paper

	"Meta-analysis of Anti- Saccharomyces Cerevisiae Antibodies as diagnostic markers of Behçet's disease. I have the following recommendation: 1. Under line 130-131, the authors stated, "the random effects models (REM) was used for analysis." The authors need to provide more explanation on why random-effects model was used. Please add the following explanation: models (REM) was used for analysis. REM attempted to generalize findings beyond the included studies by assuming that the selected studies are random samples from a larger population (Cheung MW et al 2012). Reference: Cheung MW et al. Conducting a meta-analysis: basics and good practices. Int J Rheum Dis. 2012 Apr;15(2):129-35. PMID:22462415 2. Under Figures 3A and 3B, the p-value refers to heterogeneity or I2. Can the authors show p-value for the difference in antibodies between two groups?

REVIEWER	Giovanni Di Leo IRCCS Policlinico San Donato (Italy)
REVIEW RETURNED	15-Jan-2020

GENERAL COMMENTS	Manuscript ID bmjopen-2019-033880, entitled "Meta-analysis of Anti- Saccharomyces Cerevisiae Antibodies as diagnostic markers of Behçet's disease" GENERAL COMMENTS Authors have performed a systematic review and meta-analysis of the diagnostic accuracy of Anti-Saccharomyces Cerevisiae Antibodies (ASCA) as a marker of Behçet's disease. Overall, the methods are adequate and the study is well conducted. However, it is quite hard to follow as it reports a bulk of data due to the combinations of subgroup analyses (BD vs healthy controls vs gastrointestinal BD etc.) and markers (IgG, IgA, IgG+IgA etc.). Maybe, authors could divide this manuscript in two or more separate manuscripts, although I cannot see an obvious division. Alternatively, authors could limit the statistical analysis to the diagnostic odds ratio and sROC analysis. Finally, the Results section is rather "delegated" to the reading of the Tables while I would suggest a "more systematic" presentation. SPECIFIC COMMENTS 1. Row 80: reporting the mean delay as 3.77±4.43 years is not appropriate. Better to provide a median (if available) or interval min-max. 2. Row 90: diagnostic accuracy instead of relevance. 3. Literature search: SOCPUS is SCOPUS (both in the text and abstract). Can the authors include the search string used, for example, in EMBASE? Have the authors manually searched the references of the analyzed articles? Authors seem to have followed the PRISMA statement but I would suggest to specify this in the text. 4. Eligibility and exclusion criteria: the fifth inclusion criterion is rather the fourth. 5. Statistical analysis: please, specify the model used for data

	pooling. 6. Row 168: prognostic should be prospective. 7. Rows 176-178: these are methods, not results.

REVIEWER	Mariska Leeflang Amsterdam UMC
REVIEW RETURNED	22-Jan-2020

GENERAL COMMENTS	This meta-analysis is a mix of objectives, methods, outcomes and results. This might work well in a review, but only if the authors are capable of strict structuring and clear reporting. That is not really the case here. Moreover, the methods they use for the meta-analysis (at least for the diagnostic accuracy) is outdated. 1. I understand that an abstract may not be the place to elaborate on the rationale of a certain research question, but for me as a relative lay person, the association between Saccharomyces and behcet's disease is completely unknown and unexpected. So some explanation here may be helpful. 2. The abstract contains a lot of abbreviations, which are really not necessary and do not improve readability. 3. the section 'Strengths and limitations of the study' on the top of page 3 does not contain any strangths or limitations. Please rewrite. 4. Page 4, first line: "the third and fourth decade" is difficult language. I would rather see "between 20 and 40 years of age". 5. Most on page 4 is not really relevant for the research question: a lot about the genetic origin of the disease, but the rest of the article does not address this genetic origin. 6. Would it be possible to explain why this antibody is linked to this disease? 7. A clear and complete description of the search strategy is lacking. Please add the full search strategy as an appendix, or otherwise verbatim in the text. 8. The search was done a year agon. Maybe an update is appropriate. 9. Why were non-English articles excluded. this should also be reported as a limitation. And it would be good to know how many non-English papers were excluded. 10. MetaDisc is obsolete for diagnostic test accuracy meta-analyses. Please use another program (or report the use of MetaDisc as a limitation) 11. For diagnostic test accuracy, random effects model is standard. FEM is not recommended. Also, the I-square is not recommended, as its interpretation in diagnostic accuracy is difficult. 12. Even if the authors would use I-square to indicate heterogeneity, then this should be better explained and reporrted. I-square for what (sensitivity, specifciity, DOR)? And what was the I-square then for the meta-analyses reported? Please report.

	13. Why do sensitivity analysis in STATA? and what code was used for it? (metandi? midas?) 14. No methods explained for the Association with autoimmune disease. Please add. 15. Page 9, line 154, first word: we identified should be we included in the review? 16. Page 10, line 168: Prognostic should be propective? Also, prospective or retrospective is not part of QUADAS-2. And all studies seem to be case-control design, is that true? In that case all studies should have high risk of bias for the patient selection domain and probably also high concerns regarding applicability. 17. The results section is very difficult to read and to make sense out of it. It is a mix of methods and results. Please rewrite. 18. Results for the relationship between ASCA and AID make no sense. Please remove or write a nice summary of what can be seen in Table 4. Methods should be reported in the Methods section. 19. The Discussion section is largely a repitition of the introduction. Please shorten. 20 Page 17, line 286-287: Using three different software packages really does not strengthen the credibility of results. Especially not if one of these is outdated.

VERSION 1 – AUTHOR RESPONSE

Replies to Reviewer 1
1 The English language needs linguistic modification.
Answers: This manuscript has been edited and proofread by Medjaden Bioscience Limited in the first submission and the revised manuscript submission.
2 The instruction and discussion parts need refine.
Answers: Thanks for the suggestion. We have re-written this part according to your suggestion. We have refined the introduction and discussion sections and also modified some details according to the comments of all reviewers and editorial requests.
Finally, special thanks to you for your good and helpful comments.

Replies to Reviewer 2
This is an interesting meta analysis that aims to assess the diagnostic value of anti-saccharomyces cerevisiae antibodies (ASCA) in Behçet's disease (BD) patients and explore their relationship with other autoimmune diseases(AID).In relation to this topic, I have the following observations and questions:
1.INTRODUCTION. This section should be reduced and emphasis should be placed on studies related to this research.
Answers: We are very sorry for our inappropriate writing in introduction. We have deleted unrelated contents to reduce the introduction like too much background description of BD, gene markers relevance of BD and diagnostic criteria introduction, etc. And we have placed more emphasis on related part to the research.
2. METHODS. Language can be a barrier to knowledge. Why did the authors exclude studies that were not written in English?.

Answers: We considered this when searching the literature. At first, we didn't restrict the language when searching the five database, but finally, we found that there were few (roughly less than 3) retrieved non-English studies related to this meta-analysis, but they couldn't be obtained the full text due to its publication in its domestic magazine, or there were no useful data found for this meta-analysis after the translation (using Google translate). We didn't remember clearly the specific number of the related non-English studies, but we am sure it is few and we tried but we didn't make it to be included in this meta-analysis due to the two reasons.

3.This is a study of antibodies; therefore it is important to define in this section which isotype of antibodies were determined in the studies: IgG, igA or both. Is the method to detect these antibodies standardized or are there variations?. What is the sensitivity and specificity of the method to detect these antibodies? Is the same method used in all studies?. With what method did they evaluate the publication bias of the articles?.

Answers: Thanks a lot for the comments and question pertaining to standardization of methods. The concern raised is very important for antibody detection and analysis.

In the revised manuscript, we have defined the isotypes of the analyzed antibodies in the statistical analysis subsection: IgG, IgA, IgG/IgA (positive result of either IgG or IgA), and IgG+IgA (positive result of both IgG and IgA) (Page 43, Lines 162-164).

In addition, the methods used and the corresponding details like cut-off values used in the studies are summarized in Table 1 (Page 61). There was variability with respect to the use of two methods (ELISA and IIF). We agree that it would be better to perform stratified analysis based on the detection methods for all subgroups. However, unfortunately, only 3 studies had used IIF; therefore, we could not perform a detailed stratified analysis by method. However, in the section of meta-analysis (Diagnostic ability of ASCA for GIBD) (Page 46, Line 233-238), we have calculated and compared the odds ratios, sensitivity, and specificity of ASCA-IgG for diagnosis of GIBD using the two methods; the results show a strong association using both the methods. We have briefly discussed this issue in the Discussion section of the revised manuscript.

As for publication bias, we are sorry that could not perform this analysis owing to an average of 2–3 studies included in each subgroup. It seems that more than 10 studies are required to perform a meaningful analysis of publication bias. In case of any misunderstanding at our end, we will appreciate your correction.

4.DISCUSSION. Standardization of tests to detect antibodies should be discussed

Answers: Considering the Reviewer's suggestion, we have discussed the different condition in included studies, and explained the heterogeneity. And besides, based on the results of meta-analysis (Diagnostic ability of ASCA for GIBD), we have made a brief discussion and gave our point in the discussion. (Page 51, Lines 326-336)

Finally, special thanks to you for your good and helpful comments.

Replies to Reviewer 3

1. Under line 130-131, the authors stated, "the random effects models (REM) was used for analysis." The authors need to provide more explanation on why random-effects model was used. Please add the following explanation:

models (REM) was used for analysis. REM attempted to generalize findings beyond the included studies by assuming that the selected studies are random samples from a larger population (Cheung MW et al 2012).

Reference:

Cheung MW et al. Conducting a meta-analysis: basics and good practices. Int J Rheum Dis. 2012 Apr;15(2):129-35. PMID:22462415

Answers: Thanks very much for professional help in this explanation of REM! And we have added this explanation and cited the reference in the section of statistical analysis under your suggestion (Page 42, Lines 158-161).

2. Under Figures 3A and 3B, the p-value refers to heterogeneity or I2. Can the authors show p-value for the difference in antibodies between two groups?

Answers: We have made supplement according to your comments. We reanalyzed and got the specific p value (z test) just as follows, and we have added these p values into the results (Difference in serum levels of ASCA in GIBD and other intestinal diseases) for more clear presentation (Page 47, Lines 239-245):
BD vs HC ASCA-IgA p = 0.817
BD vs HC ASCA-IgG p = 0.214
GIBD vs HC ASCA-IgA p = 0.004
GIBD vs HC ASCA-IgG p = 0.092
GIBD vs CD ASCA-IgA p = 0.634
GIBD vs CD ASCA-IgG p = 0.000
GIBD vs UC ASCA-IgA p = 0.034
GIBD vs UC ASCA-IgG p = 0.335
Besides, considering your suggest, we have added p values of significant test into the association analysis between ASCA and diseases (Table 3) (Page 63).
Finally, special thanks to you for your good and helpful comments.

Replies to Reviewer 4
GENERAL COMMENTS
Authors have performed a systematic review and meta-analysis of the diagnostic accuracy of Anti-Saccharomyces Cerevisiae Antibodies (ASCA) as a marker of Behçet's disease. Overall, the methods are adequate and the study is well conducted. However, it is quite hard to follow as it reports a bulk of data due to the combinations of subgroup analyses (BD vs healthy controls vs gastrointestinal BD etc.) and markers (IgG, IgA, IgG+IgA etc.). Maybe, authors could divide this manuscript in two or more separate manuscripts, although I cannot see an obvious division. Alternatively, authors could limit the statistical analysis to the diagnostic odds ratio and sROC analysis. Finally, the Results section is rather "delegated" to the reading of the Tables while I would suggest a "more systematic" presentation.
Answers: We are very sorry for the inappropriate writing. 1) There are indeed much data and analysis. In the revised manuscript, we have re-written the Results section. As suggested, we have now limited the statistical analysis to odds ratio, in order to assess the association between ASCA and BD/GIBD (Table 3) (Page 46, Lines 215-232; Page 63). Further, we made analyzed the sensitivity and specificity stratified by detection method (Table 4) (Page 46, Lines 233-238; Page 64). In addition, we removed the analysis performed using Revman 5.3, as there were no overall results after pooled analysis. Considering that only a few studies (average 2–3) were included in these subgroups, Meta-DiSc cannot perform sROC analysis under this condition. Therefore, the figure of sROC analysis has also been deleted, and only the OR results have been retained for preliminary analysis.
3) Considering your comments on the unclear presentation of results due to the bulk of data, we have rewritten the results into three parts (Meta-analysis 1, 2, 3). We have deleted the relatively less important parts and have focused on the important data, in order to present a more systematic description. We found that the results section also included some description of methods which may cause confusion. We have corrected this in the revised manuscript. We will be happy to address any other concerns.
SPECIFIC COMMENTS
1. Row 80: reporting the mean delay as 3.77±4.43 years is not appropriate. Better to provide a median (if available) or interval min-max.
Answers: We agreed that it's better using a median or min-max, and we tried but the cited article [Alpsoy E, et al. Review of the chronology of clinical manifestations in 60 patients with Behçet's disease. Dermatology. 2003;207(4):354-6)] doesn't present original data but only the final result (3.77±4.43), which inhibit the conversion of data presentation. So, we have modified into a general argument "The estimated duration between 78 the onset of symptoms and the fulfilment of diagnostic criteria is approximately 4 years " according to the expression in discussion section of the cited article, which may be also easy for understanding (Page 38, Lines 77-78).

2. Row 90: diagnostic accuracy instead of relevance.
Answers: We are very sorry for our incorrect writing. We have modified this sentence to “to perform a meta-analysis to assess its diagnostic accuracy for BD.” (Page 39, Line 98)
3. Literature search: SOCPUS is SCOPUS (both in the text and abstract). Can the authors include the search string used, for example, in EMBASE? Have the authors manually searched the references of the analyzed articles? Authors seem to have followed the PRISMA statement but I would suggest to specify this in the text.
Answers: 1) Thank you for correction. We have fixed the wrong expression to SCOPUS both in the text and the abstract.
2) We combined the synonyms searched in Mesh (PubMed) and Emtree (EMBASE) to form a comprehensive search string. The search string (no restricts on field) used in EMBASE is as follows. Under your and reviewer’s suggestion, we have added the search strategy into the Methods section (Page 40, 112-119). The search string was run on July 12, 2019 and found 38 studies:
#3 #1 AND #2
#2 behcet? OR ('triple symptom' AND complex) OR (triple AND symptom AND complex) OR (complex, AND triple AND symptom) OR (complices, AND triple AND symptom) OR (symptom AND complex, AND triple) OR (symptom AND complices, AND triple) OR (triple AND symptom AND complices) OR 'adamantiades behcet' OR (old AND silk AND route AND disease) OR behçet
#1 yeast?, AND baker? OR (baker? AND yeast?) OR (yeast?, AND brewer?) OR (brewer? AND yeast?) OR (s AND cerevisiae) OR (s. AND cerevisiae) OR (saccaromyces AND cerevisiae) OR (saccharomyces AND capensis) OR (saccharomyces AND cerevisia) OR (saccharomyces AND cerevisiae) OR (saccharomyces AND cerevisial) OR (saccharomyces AND cervisiae) OR (saccharomyces AND diastaticus) OR (saccharomyces AND italicus) OR (saccharomyces AND oviformis) OR (saccharomyces AND uvarum AND var. AND melibiosus) OR asca
3) As for the manual search, we have manually searched the references of the obtained articles but didn’t found any other new studies except the analyzed articles.
4) It is true that we checked all the PRISMA statement (Table S1) and we have specified this in the abstract (Page 35, Line 27) and methods section (Page 39, Lines 102-103) under your suggestion.
4. Eligibility and exclusion criteria: the fifth inclusion criterion is rather the fourth.
Answers: Thanks for correcting, the number has been corrected (Page 41, Line 127).
5. Statistical analysis: please, specify the model used for data pooling.
Answers: As your suggested, we have specified the model used for data pooling in the statistical analysis, “We chose the random 129 effects models (REM) since REM tends to generalize findings beyond the included studies by assuming that the selected studies are random samples from a larger population”. Besides, we have limited the model to REM and added corresponding explanation according to the suggestion of another two reviewers.(Page 42, Lines 158-161)
6. Row 168: prognostic should be prospective.
Answers: So sorry for this error. we had planned to use prospective instead of prognostic. But in order to correspond to Table 1, we have modified correctly to “case-control studies” instead of prospective (Page 45, Line206).
7. Rows 176-178: these are methods, not results.
Answers: We are sorry for our unclear writing. It is truly methods instead of results. And considering the repeat in Methods, so we deleted this part of Rows 176-178.
Finally, special thanks to you for your good and detailed comments.

Replies to Reviewer 5
This meta-analysis is a mix of objectives, methods, outcomes and results. This might work well in a review, but only if the authors are capable of strict structuring and clear reporting. That is not really the case here. Moreover, the methods they use for the meta-analysis (at least for the diagnostic accuracy) is outdated.

1. I understand that an abstract may not be the place to elaborate on the rationale of a certain research question, but for me as a relative lay person, the association between Saccharomyces and behcet's disease is completely unknown and unexpected. So some explanation here may be helpful.
Answers: Thanks for your suggestion. We quite understand your question since there are not much works on BD especially antibodies research. And under your suggestion, we have added related background in the abstract between BD and ASCA or S. cerevisiae to improve the understanding of the rationale for readers, which we hope is helpful. (Page 35, Lines 21-24)
And as far as I know, ASCA is detected in our hospital during the diagnosis of BD and other gastrointestinal diseases especially in Department of rheumatology, which is however less often compared with other antibody tests. And considering the reported researches on ASCA in BD and the laboratory works, we made this review in order to get a comprehensive understand between ASCA, BD and other chronic gastrointestinal diseases.
2. The abstract contains a lot of abbreviations, which are really not necessary and do not improve readability.
Answers: It is really true as you suggested that too many abbreviations increase difficult in reading. We have reduced the use of acronyms in the abstract to improve readability, but retained the BD/GIBD and ASCA, which are the important keywords and are frequently used in the text.
3. the section 'Strengths and limitations of the study' on the top of Page 3 does not contain any strangths or limitations. Please rewrite.
Answers: Sorry for incorrect writing. We have made correction according to your comments. And we have rewrote this section including four strengths and one limitations. (Page 37, Lines 54-67)
4. Page 4, first line: "the third and fourth decade" is difficult language. I would rather see "between 20 and 40 years of age".
Answers: It's true that the third and the fourth decade is really difficult to understand which we cited from a review. And now we have changed into "between 20 and 40 years of age". But finally, we have deleted this sentence to simplify the introduction according to your comments (5.) and other reviewers suggestion.
5. Most on Page 4 is not really relevant for the research question: a lot about the genetic origin of the disease, but the rest of the article does not address this genetic origin.
Answers: At the beginning, we have thought that considering the less works on antibody markers and the limitation of HLA-B1, genetic background may introduce the awareness of investigation of novel serological markers and autoantigens merits further investigations. But it's really true that we didn't address this genetic origin in the text, and the genetic relationship is not really about the theme of this meta-analysis. Considering this, we deleted this description of genetic origin in order to refine the introduction. And we have already put more emphasis on ASCA and its mechanism of the relationship with BD as you have suggested in the next point (6) (Page 39, Lines 85-92)
6. Would it be possible to explain why this antibody is linked to this disease?
Answers: Thanks a lot for the helpful suggestion. We have briefly introduced the link between BD and ASCA in the introduction (and a short description in abstract). (Page 38, Lines 86-89) Then we have made a detailed explanation of the possible mechanism in Discussion based on review of several studies, which we think will promote a better and deeper understanding of this relationship. (Page 49, Lines 270-282)
7. A clear and complete description of the search strategy is lacking. Please add the full search strategy as an appendix, or otherwise verbatim in the text.
Answers: We have detailed the search strategy in the section of Literature search with the suggestion of you and another reviewer. (Page 40, Lines 112-120)
8. The search was done a year ago. Maybe an update is appropriate.
Answers: We have rerun the search stragety on February 12, 2020. We found more studies in each of the five database, but finally after the screening process there are still 9 studies included.
9. Why were non-English articles excluded. this should also be reported as a limitation. And it would be good to know how many non-English papers were excluded.

Answers: At first, we don't restrict the language when searching the five database, but finally, there are few (roughly less than 3) retrieved studies related to this meta-analysis, but they can't be obtained the full text due to its publication in its domestic magazine, or there are no useful data found for this meta-analysis after the translation (using Google translate). We am not sure about the specific number of the related non-English studies, but we are sure it is few and we didn't make it to be included in this meta-analysis due to the reasons.
Under your suggestion, we have reported this limitation in the Discussion section (Page 52, Lines 340-342).
10. MetaDisc is obsolete for diagnostic test accuracy meta-analyses. Please use another program (or report the use of MetaDisc as a limitation)
Answers: We agree that there are indeed several shortcomings of Meta-DiSc. We have explained this in the Discussion section as per your suggestion. (Page 52, Lines 347-349) However, because of subgroup analysis disaggregated by the isotypes of ASCA and different disease controls, there were less than four studies each in all subgroups (Table 3). Unfortunately, Stata cannot do pooled analysis of sensitivity and specificity when there are less than four studies included. Therefore, we used Stata for pooled analysis of ORs (Table 3) and continuous data (serum levels of ASCA, Figure 3). Besides, RevMan only provides limited analyses for diagnostic accuracy studies. The sensitivity and specificity are separately plotted in forest plot only for the purposes of exploration and presentation; no summary points or heterogeneity measures are provided. Therefore, we used Stata for analysis of diagnostic odds ratio and Meta-DiSc for analysis of diagnostic accuracy.
We agree with you that Meta-DiSc is an obsolete tool for meta-analysis. But for this meta-analysis, Meta-DiSc seem not to be a bad choice. Then we report the use of Meta-DiSc as a limitation in the Discussion.
Maybe in the future, studying R to solve this problem may be a better choice.
11. For diagnostic test accuracy, random effects model is standard. FEM is not recommended. Also, the I-square is not recommended, as its interpretation in diagnostic accuracy is difficult.
Answers: Thanks a lot for the professional suggestion pertaining to the effects models and I-square. The results of REM results are more conservative compared with FEM. Then we have re-analyzed the data all using REM and we have deleted the I-square under your suggestion. (Page 42, Lines 157-161)
12. Even if the authors would use I-square to indicate heterogeneity, then this should be better explained and reporrted. I-square for what (sensitivity, specifciity, DOR)? And what was the I-square then for the meta-analyses reported? Please report.
Answers: We have deleted I-square under your suggestion. And in this meta-analysis, heterogeneity among the included studies was evaluated using the Cochran's Q-statistic. P value > 0.10 was considered indicative of lack of significant heterogeneity. (Page 42, Lines 157-158)
13. Why do sensitivity analysis in STATA? and what code was used for it? (metandi? midas?)
Answers: 1) We performed the sensitivity analysis by successively excluding studies in order to assess the stability of the outcomes, since p value suggest the presence of heterogeneity during the meta-analysis.
2) We made the analysis in Stata through the meta-analysis menu operation (the code is metan).
14. No methods explained for the Association with autoimmune disease. Please add.
Answers: We are sorry for this omission. We have added this part of methods in the last of Methods - "Relationship between ASCA and autoimmune disease" (Page 43, Lines 178-183), including a brief description of the search strategy and a method of summarizing data in those studies.
15. Page 9, line 154, first word: we identified should be we included in the review?
Answers: Thanks for correction. We have modified "identified" to "included". (Page 44, Line 193)
16. Page 10, line 168: Prognostic should be propective? Also, prospective or retrospective is not part of QUADAS-2. And all studies seem to be case-control design, is that true? In that case all studies should have high risk of bias for the patient selection domain and probably also high concerns regarding applicability.

Answers: 1) Thanks for your correction. What we wanted to express at first was prospective instead of prognostic. And with your reminder, we think it's accurate that all studies are case-control designed studies except that 2017 Shulan Zhang is a retrospective study (Table 1). (Page 45, Line 106)
2) And it is true that almost all studies have high risk of bias for the patient selection considering the three questions of Domain 1(Patient selection), just as the Figure S1 shows.
3) As the QUADAS-2 says in the Applicability of Patient Selection: Concerns about applicability may exist if patients included in the study differ from those targeted by the review question in terms of severity of the target condition, demographic features, presence of differential diagnosis or comorbid conditions, setting of the study, and previous testing protocols. I think the concerns are not that high regarding patients included and the target of the review question. My reasons are as follows: ①The severity of patients was not mentioned in any of the studies; this implies that it is likely that no patients with specific severity were preferentially selected; ② the presence of differential diagnosis: the differential diagnosis diseases like CD/UC/iTB are included in the controls groups; in addition, the spectrum of patients included in the studies was representative of those in whom the test will be used in clinical practice; ③ all the studies were case-control studies except one (2017 Shulan Zhang); the patient selection criteria were clear (1990 ISG criteria or 1987Japan criteria) except in one study (2010 B. Kocazeybek); In addition, all criteria were based on the presence of several symptoms; the protocol was repetitive; the design of control groups was reasonable since almost all studies (except 2011 George Vaiopoulos and 2006 Chang Hwan Choi) had included diseases that are considered in the differential diagnosis… ; ④the testing protocols were from either Inova Diagnostic or Euroimmun, Luebeck except for one study (2002 Byeong Gwan Kim) used self-coated ELISA to test ASCA, which are reproducible. ⑤As for demographic features, there are indeed some limitations which have been acknowledged in the discussion section. ⑥As for the comorbid conditions, it seems that almost all studies did not mention that all patients with other other comorbid conditions were excluded except two studies (2017 Shulan Zhang and 2018 Shulan Zhang). However, to the best of our knowledge, the patients were likely to be excluded from other autoimmune diseases before inclusion in the studies. Thus, generally speaking, we think the concerns about the applicability of patient selection is relatively low. If there is misunderstanding, thank you very much for your correction.
4) More importantly, with your correction, we think it may be better using the modified Newcastle-Ottawa Scale (NOS) for the Risk of bias assessment of the included studies, since it is may not be suitable using QUADAS-2 considering that there is always high risk of bias for the patient selection domain of QUADAS-2 for case-control studies. Therefore, we re-evaluated the risk of bias of included studies using NOS (Table 2). (Page 42, Lines 137-138; Page 45, Lines 206-207)
17. The results section is very difficult to read and to make sense out of it. It is a mix of methods and results. Please rewrite.
Answers: We am sorry for the unclear structure of results. We have rewritten and made the methods and results more well-defined to present a clear description of the results.
18. Results for the relationship between ASCA and AID make no sense. Please remove or write a nice summary of what can be seen in Table 4. Methods should be reported in the Methods section.
Answers: We have rewritten the result of this section (Page 48, Lines 254-261) and reported the methods in the last part of Methods section (Page 43, Lines 178-183).
19. The Discussion section is largely a repetition of the introduction. Please shorten.
Answers: We am sorry for that. Under your suggestion, we have modified the introduction and discussion, deleted the repetition and added more explanation of mechanism between ASCA and BD/autoimmune diseases in discussion. And we have rewrote other parts of discussion for a clear presentation and argument
20 Page 17, line 286-287: Using three different software packages really does not strengthen the credibility of results. Especially not if one of these is outdated.
Answers: We agree that three software can't strengthen the credibility of results, and we have deleted this sentence. However, the two types of meta-analysis like continuous data (sensitivity and specificity etc.) and discontinuous data (serum levels) to evaluate the diagnostic potential of ASCA for BD could

help increase the credibility of our results. Besides, considering your and other reviewer's comments, we have removed Revman. Under your suggestion, we modified the original expression.
Finally, thanks for your taking time to provide the professional and detailed comments. They really help to improve the manuscript.

We tried our best to improve the manuscript and made some changes in the manuscript. These changes will not influence the content and framework of the paper but present a clearer data analysis and framework of results and discussion. And here we did not list the changes but marked in revised paper.
We appreciate for Editors/Reviewers' warm work earnestly and hope that the correction will meet with approval.
Once again, thank you very much for your comments and suggestions.

VERSION 2 – REVIEW

REVIEWER	Luis J Jara Quezada Hospital de Especialidades "Dr. Antonio Fraga Mouret" Centro Médico Nacional La Raza Instituto Mexicano del Seguro Social Mexico City Mexico
REVIEW RETURNED	30-Mar-2020

GENERAL COMMENTS	The authors have satisfactorily answered my questions and observations. This study may be accepted for publication in BMJ Open

REVIEWER	Giovanni Di Leo IRCCS Policlinico San Donato, Italy
REVIEW RETURNED	26-Mar-2020

GENERAL COMMENTS	The paper has greatly improved and I now think it is worth of publication. As a last suggestion, I would use

REVIEWER	Mariska Leeflang Amsterdam UMC, location AMC, University of Amsterdam
REVIEW RETURNED	16-Mar-2020

GENERAL COMMENTS	The authors have addressed most of my comments. However, I am very sorry if my comments led to confusion, but there are still some issues remaining. 1. The main problem is that the authors seem to have tried to adjust the manuscript according to the comments of the peer-reviewers, but without considering the real problems. For example, the authors stick to MetaDisc, because it can handle a meta-analysis of fewer than four studies. To me, this is not a valid argument to turn to an outdated method. MetaDisc does not address the bivariate nature of sensitivity/specificity data and it does not handle between-study variability in a correct way (and as far as I know, but I may be wrong, MetaDisc pools DORs in a fixed-effect-way, which may not be appropriate at all). There is noting wrong with refraining from doing a meta-analysis if there are not enough studies....

	2. Authors state that they have removed RevMan, which is absolutely fine, but there is no reason to remove RevMan at all if you used it, for example, to draw figures or to check heterogeneity visually. 3. Authors state that they used the Newcastle-Ottawa scale to assess the quality of the included studies. (1) the NOS is not suitable for diagnostic accuracy studies; and (2) using a different scale does not remove the true risk of bias... QUADAS-2 remains the appropriate tool for quality assessment of these studies and if that leads to high risk of bias, then authors should think about the implications this has for their results (and it may mean that the results from the studies are just not that trustworthy...so be it) 4. Authors conclude that their "meta-analysis results, together with the review of ASCA in autoimmune diseases strongly suggest that ASCA (especially its certain isotypes) may be helpful biomarkers for GIBD, especially with respect to their possible predictive/pathogenic/diagnostic role in clinical settings". I disagree with this conclusion, taking into account that the sensitivity is below 50% and that the DOR is not higher than 5 or so. This, in combination with the high risk of bias, the limited number of studies and the obsolete methodology, makes this an inappropriate conclusion.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 5

Reviewer Name: Mariska Leeflang

Institution and Country: Amsterdam UMC, location AMC, University of Amsterdam

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

The authors have addressed most of my comments. However, I am very sorry if my comments led to confusion, but there are still some issues remaining.

1. The main problem is that the authors seem to have tried to adjust the manuscript according to the comments of the peer-reviewers, but without considering the real problems. For example, the authors stick to MetaDisc, because it can handle a meta-analysis of fewer than four studies. To me, this is not a valid argument to turn to an outdated method. MetaDisc does not address the bivariate nature of sensitivity/specificity data and it does not handle between-study variability in a correct way (and as far as I know, but I may be wrong, MetaDisc pools DORs in a fixed-effect-way, which may not be appropriate at all). There is noting wrong with refraining from doing a meta-analysis if there are not enough studies....

Response: Considering the limitations in Meta-Disc, we have lessened the use of Meta-Disc to obtain pooled DOR results.

Besides, as you suggested in your first review, random effects model is standards for diagnostic test accuracy, we greatly appreciate your suggestion and have followed the principle in our other works with meta-analysis. Therefore, we use the pooling method of Dersimonian-Laird (REM) – a random effect model according to DerSimonian R (DerSimonian R, Laird N: Meta-analysis in clinical trials.

Control Clin Trials 1986, 7(3):177-188), to pool the results in Meta-Disc, so the method, we think, could be appropriate.

BD is a rare complex syndrome with unknown pathogenesis. There are still no specific biomarkers for BD. So far, there have been few studies on autoantibodies in BD. In this meta-analysis, 9 studies were included, but due to the stratification by isotype of ASCA, it is a pity that not very much studies were included in every subgroup. To the best of our knowledge, this is the second meta-analysis of evidence pertaining to autoantibodies in patients with BD after anticardiolipin antibodies. I hope the meta-analysis provides the insight of the relationship of ASCA with BD and the pathogenesis of ASCA or Saccharomyces cerevisiae in BD.

2. Authors state that they have removed RevMan, which is absolutely fine, but there is no reason to remove RevMan at all if you used it, for example, to draw figures or to check heterogeneity visually.

Response: Thanks for your suggestion. In the preliminary manuscript, we only drew one figure by RevMan, showing the AUC of diagnostic accuracy of ASCA comparing BD and HC (A), BD and CD (B), BD and UC (C), GIBD and HC (D), GIBD and CD (E), GIBD and UC (F), and GIBD and iTB (G). Considering that we finally refined our data analysis to DOR, and there were no any pooled effects in the results of AUC, so RevMan was finally removed after consideration.

3. Authors state that they used the Newcastle-Ottawa scale to assess the quality of the included studies. (1) the NOS is not suitable for diagnostic accuracy studies; and (2) using a different scale does not remove the true risk of bias... QUADAS-2 remains the appropriate tool for quality assessment of these studies and if that leads to high risk of bias, then authors should think about the implications this has for their results (and it may mean that the results from the studies are just not that trustworthy...so be it)

Response: We have removed the Newcastle-Ottawa scale and retained the QUADAS-2 scale (Page 40, line 137-140). And we included the discussion as per the risk of bias by QUADAS-2 in the discussion “According to the QUADAS-2, there are certain concerns that most studies have risk of bias (internal validity) in patient selection, which, to some extent, would cause the distorted estimation in diagnostic accuracy [20].” (Page 51, line 333-336)

4. Authors conclude that their "meta-analysis results, together with the review of ASCA in autoimmune diseases strongly suggest that ASCA (especially its certain isotypes) may be helpful biomarkers for GIBD, especially with respect to their possible predictive/pathogenic/diagnostic role in clinical settings". I disagree with this conclusion, taking into account that the sensitivity is below 50% and that the DOR is not higher than 5 or so. This, in combination with the high risk of bias, the limited number of studies and the obsolete methodology, makes this an inappropriate conclusion.

Response: Thanks for your suggestion, we have modified the sentence as “Our study demonstrated the relationship between ASCA/Saccharomyces cerevisiae and gastrointestinal involvement in BD. … However, detection of only ASCA may have a limited value for clinical diagnosis due to its moderate sensitivity and the presence in several other autoimmune diseases. In the future, further studies are needed to explore the role of ASCA and Saccharomyces cerevisiae in BD.” (Page 52, line 350-357)

Reviewer: 4

Reviewer Name: Giovanni Di Leo

Institution and Country: IRCCS Policlinico San Donato, Italy

Please state any competing interests or state ‘None declared’: Non declared

Please leave your comments for the authors below

The paper has greatly improved and I now think it is worth of publication. As a last suggestion, I would use

Response: Thanks very much for your suggestion to help improve the meta-analysis. And I appreciate your recognition of our work very much.

Reviewer: 2

Reviewer Name: Luis J Jara Quezada

Institution and Country:

Hospital de Especialidades "Dr. Antonio Fraga Mouret"

Centro Médico Nacional La Raza

Instituto Mexicano del Seguro Social

Mexico City

Mexico

Please state any competing interests or state ‘None declared’: NONE DECLARED

Please leave your comments for the authors below

The authors have satisfactorily answered my questions and observations. This study may be accepted for publication in BMJ Open

Response: Thanks very much for your suggestion to help improve the meta-analysis. It is worth mentioning that with your suggestion we have had a “more systematic” presentation of results and conclusion. I appreciate your recognition of our work very much.