nature portfolio

Peer Review File

Cost-Effectiveness of Community-based Integrated Care Model for Patients with T2DM and Depressive Symptoms

Corresponding Author: Dr Ping He

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have conducted a within-trial economic evaluation of integrated care for people with comorbid diabetes and depression. With multi-morbidity becoming more prevalent and calls being made for integrated care in many areas of health care for many years, this study provides important quantitative insights on the health and economic benefits of this model of care.

Abstract

Suggest adding the lower threshold in brackets to line 44 so that international readers can compare the reported ICERs to this without having to refer to the main manuscript.

Consider adding another sentence at line 46 converting the ICERs to international dollars or US dollars to assist the interpretation of the findings by international readers.

Suggest adding the time horizon of the analysis to the abstract.

Otherwise the abstract is clearly written and contains the key information for an economic evaluation.

Main manuscript

Page 8, lines 288 to 299 – More explanation is warranted about the similarity and contrasts of baseline characteristics beyond “both groups are comparable”. Although further detail is provided about baseline depressive symptoms and utility, referring to Appendix Table 2 there are important differences between groups regarding household income (61% below 26,000 for usual care vs. 48% for integrated care) and rurality (78% for usual care vs. 63% for integrated care). These may have implications for accessing health care services and adherence with treatment protocols, and could even explain some of the differences between groups in depressive symptoms and utilities directly. Consider adding these to the sensitivity analysis ‘Baseline Adjustment’. Also worth checking are the baseline utilities for each group that appear in Appendix Table 2 – these are not just similar as stated in the body of the manuscript but exactly the same.

Page 10, lines 366 to 369 – it’s not immediately clear what the proportions are referring to. Suggest adding a few words to this sentence that they are for the upper and lower thresholds. Also consider adding a sentence reporting these figures for the other two perspectives.

Page 10, lines 372 to 374. It seems a little strange to use 25,000 RMB (should this be 250,000RMB?) as the reference point for the CEAC considering you’ve already defined the upper and lower thresholds based on GDP. I think this sentence could be deleted, unless the figure has a historical/conventional role in health economic decision making in China, in which case this should be explained here. Alternatively, if you want to say something else about the CEAC that isn’t already covered by

the explanation of the cost-effectiveness plane, you could identify the willingness to pay amounts that intersect with 50% probability.

Page 11, line 324. I think the comparison with existing cost-effectiveness analyses in this area could better recognise that all the ICERs reported here are actually cost effective so there appears to be more consistent findings across the literature on integrated/collaborative care than is currently suggested by “mixed results” in the first sentence of the paragraph. Could also note whether all these other studies are within-trial economic evaluations like the present analysis (or modelled cost-effectiveness analysis).

Page 13, line 496. Is another limitation that although the societal perspective has included productivity impacts due to time costs, productivity impacts due to presenteeism and absenteeism associated with depression or diabetes have not been included?

Figure 1 - The cost-effectiveness plane is difficult to interpret with the 5000 simulations of all three perspectives overlapping each other. Suggest separating this out into 3 separate panels for figure 1, one for each perspective. This would also match the approach taken for Figure 2. Otherwise this is a clearly labelled, useful figure.

Table 1 - Subtotals have been provided for each perspective. However it’s difficult to work out which costs are included in which perspective from looking at Table 1 alone. Suggest adding a note to the table footer pointing the reader to Appendix Table 1 for additional information on allocation of cost items to perspectives.

Supplementary material

Page 9, lines 101 to 103 - Although a reference to the original source is provided, it would be useful to explain what the utility values are based on here considering how integral they are the calculation of QALYs.

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)
This cost-effectiveness analysis, conducted within a randomized controlled trial (RCT), has both strengths and weaknesses.

The economic analysis considered three perspectives, as well as both direct and indirect costs. Sensitivity analyses addressed key parameter variability for robustness. However, the bootstrap method used did not account for clustered data. A clustered bootstrap approach, as discussed in Bachman et al. (2007), would provide more robust estimates. The authors should clarify their method accordingly.

The protocol is correctly cited, ensuring methodological transparency. To clarify the processes involved for readers, consider expanding the "data collection and methods" section around line 265.

A few additional points require attention:

The measurement methods for healthcare utilization costs and indirect costs are not specified in detail, including in "Appendix Table 3."

In Table 1, the "cost types and components" section is labelled as "mean (SD)," which suggests that the values represent means and standard deviations. However, the values provided do not appear to be means, and no standard deviations are presented in brackets.

Lastly, the study acknowledges several limitations in the discussion section, which the authors openly address. While these limitations may impact the generalizability and interpretation of the results, they are appropriately considered in the study context.

Overall, this study is well-conducted and provides valuable insights. Addressing the bootstrap and protocol issues and minor issues would improve the clarity and appropriateness of the report's methods.

REFERENCE:

Bachman et al. 2007, Methods for analyzing cost effectiveness data from cluster randomized trials; doi:10.1186/1478-7547-

5-12

Reviewer #4

(Remarks to the Author)
see uploaded file

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
No further comments following revision

Reviewer #3

(Remarks to the Author)
Thank you for your revised manuscript and for addressing the comments provided.

The corrections for points Q2 and Q3 are satisfactory, and I appreciate your effort in implementing the changes. However, the revisions for points Q1 and Q4 do not fully address my initial requests:

Q1: The clustered bootstrap method should be applied to the main analysis, rather than as an additional sensitivity analysis. Please revise accordingly.

Q4: Since these are fixed costs, calculating a mean and standard deviation is not appropriate. Instead, you should simply report the fixed cost applied without any statistical analysis.

Reviewer #4

(Remarks to the Author)
Overall: hard to assess what was actually changes as no track changes or anything was used and the mentioned line number do not match those in manuscript.

Q1: there is still a mix of diabetes en T2DM found throughout the manuscript.

Q9-11: introduction is revised sufficiently.

Q16: question not addressed, authors reference a paper, but in the paper no details on standard services are provided either. Besides, it is important that a reader can understand the intervention without reading another article by briefly describing the main elements. This also apply for Q17, were reference is made to a previous publication.

Q19: thanks for adding the sample size calculation. This part could benefit from specifying what the primary outcome entails.

Q22: thanks for the explanation. We would suggest to remove the sentence that baseline characteristics are comparable, as this is misleading. It might be good to shortly mention the similarities and differences.

Q23: this might be relevant to reflect on in the discussion. And we do not agree that a statistical difference suggest clinical relevance, as a statistical difference only indicates the likelihood of an actual difference but does not specify anything in terms of clinical relevance. So good to reflect on the clinical relevance of the difference in the discussion.

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Dear Reviewers,

Thank you so much for taking your time and effort to review our manuscript entitled *“Cost-Effectiveness of Community-based Integrated Care Model for Patients with Diabetes and Depressive Symptoms”*. We appreciate all your valuable comments and suggestions, which are very helpful in improving the quality of our paper. All the comments raised by reviewers have been addressed properly in the revised manuscript, which have been clearly marked in red font color in the paper. Please find our itemized responses and revised manuscript in the re-submitted files.

Reviewer #1:

The authors have conducted a within-trial economic evaluation of integrated care for people with comorbid diabetes and depression. With multi-morbidity becoming more prevalent and calls being made for integrated care in many areas of health care for many years, this study provides important quantitative insights on the health and economic benefits of this model of care.

Abstract

Q1:Suggest adding the lower threshold in brackets to line 44 so that international readers can compare the reported ICERs to this without having to refer to the main manuscript.

Response: Thank you for your suggestion. Firstly, we have converted all monetary values to US dollars to facilitate international comparisons. Additionally, we have included the specific values of 1 to 3 times the per capita GDP, measured in US dollars (Line 48-49).

Q2:Consider adding another sentence at line 46 converting the ICERs to international dollars or US dollars to assist the interpretation of the findings by international readers.

Response: Thanks for your point. We have converted all monetary values in this study to US dollars.

Q3:Suggest adding the time horizon of the analysis to the abstract.

Response: Thank you for your suggestion. We have added this information (Line 38).

Main manuscript

Q4:Page 8, lines 288 to 299 – More explanation is warranted about the similarity and contrasts of baseline characteristics beyond “both groups are comparable”. Although further detail is provided about baseline depressive symptoms and utility, referring to Appendix Table 2 there are important differences between groups regarding household income (61% below 26,000 for usual care vs. 48% for integrated care) and rurality (78% for usual care vs. 63% for integrated care). These may have implications for accessing health care services and adherence with treatment protocols, and could even explain some of the differences between groups in depressive symptoms and utilities directly. Consider adding these to the sensitivity analysis ‘Baseline Adjustment’. Also worth checking are the baseline utilities for each group that appear in Appendix Table 2 – these are not just similar as stated in the body of the manuscript but exactly the same.

Response: Thank you for your good suggestion. We have further reported the

statistical test results of baseline characteristic differences between the two groups (Appendix Table 2). Due to the nature of the cluster RCT design, some differences were observed between the groups in terms of gender, residence, and income. We conducted an additional sensitivity analysis by including baseline characteristics as covariates to obtain adjusted results accordingly (Line 302-303; Line 430-434; Appendix Table 6 and Appendix Table 7). The findings remained consistent with the original analysis.

Q5:Page 10, lines 366 to 369 – it's not immediately clear what the proportions are referring to. Suggest adding a few words to this sentence that they are for the upper and lower thresholds. Also consider adding a sentence reporting these figures for the other two perspectives.

Response: Thank you for your suggestion. We have clarified the meaning of these proportions (Figure 1) and included the probabilities from the other two perspectives (Line 411-419).

Q6:Page 10, lines 372 to 374. It seems a little strange to use 25,000 RMB (should this be 250,000RMB?) as the reference point for the CEAC considering you've already defined the upper and lower thresholds based on GDP. I think this sentence could be deleted, unless the figure has a historical/conventional role in health economic decision making in China, in which case this should be explained here. Alternatively, if you want to say something else about the CEAC that isn't already covered by the explanation of the cost-effectiveness plane, you could identify the willingness to pay amounts that intersect with 50% probability.

Response: Thanks for your point. We apologize for this oversight; the correct value should indeed be 250,000. We have removed the statement.

Q7:Page 11, line 324. I think the comparison with existing cost-effectiveness analyses in this area could better recognise that all the ICERs reported here are actually cost effective so there appears to be more consistent findings across the literature on integrated/collaborative care than is currently suggested by "mixed results" in the first sentence of the paragraph. Could also note whether all these other studies are within-trial economic evaluations like the present analysis (or modelled cost-effectiveness analysis).

Response: Thank you for your suggestion. Firstly, we have revised our wording (Line 12). Additionally, we carefully reviewed other studies and found that they are all trial-based cost-effectiveness studies, so a direct comparison with our research is appropriate.

Q8:Page 13, line 496. Is another limitation that although the societal perspective has included productivity impacts due to time costs, productivity impacts due to

presenteeism and absenteeism associated with depression or diabetes have not been included?

Response: Thanks for your point. We have added a description of this limitation in the limitations section (Line 582-585).

Q9:Figure 1 - The cost-effectiveness plane is difficult to interpret with the 5000 simulations of all three perspectives overlapping each other. Suggest separating this out into 3 separate panels for figure 1, one for each perspective. This would also match the approach taken for Figure 2. Otherwise this is a clearly labelled, useful figure.

Response: Thanks for your point. We have presented these figures separately (Figure 1).

Q10: Table 1 - Subtotals have been provided for each perspective. However it's difficult to work out which costs are included in which perspective from looking at Table 1 alone. Suggest adding a note to the table footer pointing the reader to Appendix Table 1 for additional information on allocation of cost items to perspectives.

Response: Thanks for your point. We have added a note in Table 1 to help readers better understand the content.

Supplementary material

Q11: Page 9, lines 101 to 103 - Although a reference to the original source is provided, it would be useful to explain what the utility values are based on here considering how integral they are the calculation of QALYs.

Response: Thanks for your point. We have added the relevant literature and statement (Supplement 5. Overview of sensitivity analyses).

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3 (Remarks to the Author):

This cost-effectiveness analysis, conducted within a randomized controlled trial (RCT), has both strengths and weaknesses.

Q1: The economic analysis considered three perspectives, as well as both direct and indirect costs. Sensitivity analyses addressed key parameter variability for robustness. However, the bootstrap method used did not account for clustered data. A clustered bootstrap approach, as discussed in Bachman et al. (2007), would provide more robust estimates. The authors should clarify their method accordingly.

Response: Thank you for your helpful suggestion and reference. We carefully reviewed this article and reanalyzed our results using cluster bootstrap. We found that the results remain robust (Appendix Figure 4). We have highlighted this in the robustness checks and results sections (Line 306-308; Supplement 5. Overview of sensitivity analyses).

Q2: The protocol is correctly cited, ensuring methodological transparency. To clarify the processes involved for readers, consider expanding the "data collection and methods" section around line 265.

Response: Thanks for your point. We expanded the data collection and methods (Line 283-299).

Q3: The measurement methods for healthcare utilization costs and indirect costs are not specified in detail, including in "Appendix Table 3."

Response: Thanks for your point. We have revised the section on cost calculation (Supplement 3).

Q4: In Table 1, the "cost types and components" section is labelled as "mean (SD)," which suggests that the values represent means and standard deviations. However, the values provided do not appear to be means, and no standard deviations are presented in brackets.

Response: Thank you for your suggestion. We apologize for this oversight. Since some costs are fixed, it was not possible to calculate the SD. We have now corrected this table (Table 1).

Reviewer #4 (Remarks to the Author):

General comments

Q1:We would recommend to be consistent in the key words/ concepts like T2DM vs diabetes and primary health care vs primary healthcare, US vs usual care.

Response: Thanks for your point. We have thoroughly revised the entire manuscript to ensure consistency.

Q2:Superscript refences is placed behind the full stop instead of before.

Response: Thanks for your point. We have thoroughly revised the placement of the superscript references.

Q3:Grammar, spelling layout including the reference list check is recommended. Also check correct use of past, present voice in the discussion section.

Response: Thank you for your suggestion. We have thoroughly checked and revised the grammar and wording throughout the manuscript.

Abstract

Q4:How many CHC were randomized considering the cluster RCT design.

Response: Thank you for your suggestion. We have added the number of CHCs in the abstract (Line 33).

Q5:Was an ITT analysis performed

Response: Thanks for your point. We have added this statement in the abstract (Line 37).

Q6:We would advise to mention the costs in US dollars for interpretation purpose.

Response: Thanks for your point. We have converted all the currencies to US dollars.

Q7:Is there a national cut-off for a QUALY in China, good to reflect on that

Response: Thanks for your point. We added the threshold (Line 48-49).

Q8:We prefer if subheadings like background, methods, results, discussion are used.

Response: Thanks for your suggestion. The content of our abstract is organized in this way, but to comply with the journal's format, we did not include subheadings.

Introduction:

Q9:Could be written more concisely by removing some repeated/obvious statements and clarifying some claims. For example, it seems obvious that China is the country with the largest number of diabetes patients globally, as it is the country with the largest population in general. What about proportionally?

Response: Thanks for your suggestion. We have revised these statements and completely revised the introduction.

Q10: Potentially shortly elaborate on why the CIC-PDD is effective/why you would

use an integrated care initiative to improve care.

Response: Thanks for your good suggestion. We have added content regarding the effectiveness of CIC-PDD (Line 94-96; Line 109).

Q11: The sentence “Moreover, the outcomes of diabetes and depression multimorbidity are worse when they appear together’ is not clear, do the authors mean that outcomes of diabetes and depression are separately worse than outcomes when diabetes and depression co-exists? I assume the authors mean the other way around.

Response: Thanks for your good point. We have reorganized this content (Line 75-76).

Methods:

Q12: Explanation for significantly larger group opting out of the intervention group and possibility of this causing selection bias?

Response: Thanks for your good suggestion. We conducted a differential analysis of the follow-up loss between the two groups and found no statistically significant differences (Line 330; Appendix Table 3).

Q13: In healthcare utilization it is unclear what costs are included, specifically the inpatient visits: “[…] with a specific focus on those resulting from diabetes-related complications”.

Response: Thanks for your point. We have added this description in the Supplement 3, along with the relevant details.

Q14: Unclear how the distinction between societal costs and multipayer perspective is made and what this distinction actually reflects.

Relevance of both is questionable as they only include costs directly related to care, and not costs such as diabetes/depression related sick leave? It would be understandable that this information is left out if it wasn’t gathered, but than it might be better to leave out this category completely as it might not reflect actual societal costs well.

Response: Thanks for your suggestion. We have added the details of these two perspectives in the appendix (Supplement 4).

Q15: Are sex-differences explored?

Response: Thanks for your point. We explored sex differences and found no significant differences (Line 389; Appendix Table 4).

Q16: Did the usual care group receive no depression-related treatment? As in the discussion it is mentioned that there’s a shortage. Cause then are the authors really measuring (cost-) effectiveness of the integrated care approach or simply some

sort of mental health care?

Response: Thanks for your suggestion. From the perspective of intervention components, the intervention group received a structured, integrated, and continuous collaborative care provided by a multidisciplinary team, whereas the usual care group did not receive such an intervention and only received standard services. Please see the our previous reference for intervention details.

-Wang Y, Guo D, Xia Y, et al. Effect of Community-Based Integrated Care for Patients With Diabetes and Depression (CIC-PDD) in China: A Pragmatic Cluster-Randomized Trial[J]. Diabetes Care, 2024: dc241593.

Q17: It is not completely clear if the aim of the intervention is to screen on depression in people with T2DM or to screen and treat depression in this population. Overall the manuscript would benefit from some explanation of the intervention including whether or not it was found to be effective.

Response: Thanks for your point. Considering this is a pragmatic trial, we initially focused on identifying and screening multimorbidity patients, before considering their management. Our effectiveness analysis has shown that this intervention has a positive impact on patients' mental health and glycemic control. Please see the reference:

-Wang Y, Guo D, Xia Y, et al. Effect of Community-Based Integrated Care for Patients With Diabetes and Depression (CIC-PDD) in China: A Pragmatic Cluster-Randomized Trial[J]. Diabetes Care, 2024: dc241593.

Q18: In which register is the trial registered and was ethical approval granted

Response: Thanks for your point. We have added this information (Line 297-299).

Q19: Was the study powered for cost-effectiveness or was this performed for another outcome.

Response: Thanks for your point. We have added the sample size calculation method (Line 288-294), and the current study has the largest sample size among similar studies.

Q20: Why was the SF-12 used instead of the EQ5D?

Response: Thanks for your point. Considering that our intervention targets comorbid patients with diabetes and depression, the SF-12 helps break down quality of life into mental and physical health components, allowing us to separately analyze the impact of the intervention on mental and physical well-being (see our study :). In addition, previous systematic reviews of intervention studies targeting comorbid mental and physical health conditions have found that a significant portion uses the SF-12 to measure quality of life. Please see the following reference:

-Wang Y, Guo D, Xia Y, et al. Effect of Community-Based Integrated Care for Patients With Diabetes and Depression (CIC-PDD) in China: A Pragmatic Cluster-Randomized Trial[J]. Diabetes Care, 2024: dc241593.

-Ell K, Katon W, Xie B, et al. Collaborative care management of major depression among low-income, predominantly Hispanic subjects with diabetes: a randomized controlled trial[J]. Diabetes care, 2010, 33(4): 706-713.

-Ell K, Katon W, Xie B, et al. One-year postcollaborative depression care trial outcomes among predominantly Hispanic diabetes safety net patients[J]. General hospital psychiatry, 2011, 33(5): 436-442.

-Wang Y, Hu M, Zhu D, et al. Effectiveness of collaborative care for depression and HbA1c in patients with depression and diabetes: a systematic review and meta-analysis[J]. International Journal of Integrated Care, 2022, 22(3).

Q21: It would be helpful to include in the flow chart how many individuals within the centers were eligible for the intervention.

Response: Thanks for your suggestion. We have added the patient numbers for each CHC in the flow chart (Appendix Figure 1).

Results:

Q22: Comparable baseline characteristics, based on what? When comparing percentages they do seem to differ extensively on gender, residency, and income. Could this be a result of selection bias as indicated earlier? (Proportionally more males, urban and affluent in the intervention group)

Response: Thanks for your point. We have provided the results of the baseline characteristic differences between the two groups. Due to the nature of the cluster RCT design, we observed certain differences between the groups in terms of gender, residence, and income. Therefore, we conducted an additional sensitivity analysis, adjusting for baseline characteristics as control variables (Appendix Table 6 and Appendix Table 7). Our results remained consistent.

Q23: Results on the DFDs are hard to interpret and unclear whether the difference in QALY's is clinically important as the CIs largely overlap.

Response: Thanks for your point. We have added the definition of DFDs in the Supplement 2. This indicator is a crucial and commonly used measure of mental health utility in studies focusing on patients with depression. Secondly, for QALY, in our previous research (*Wang Y, Guo D, Xia Y, et al. Effect of Community-Based Integrated Care for Patients With Diabetes and Depression (CIC-PDD) in China: A Pragmatic Cluster-Randomized Trial[J]. Diabetes Care, 2024: dc241593.*), we analyzed the effects of this intervention on both mental-related quality of life and physical-related quality of life. We found that the intervention significantly improved

mental QoL, while physical QoL showed a slight increase but was not statistically significant. The QALY differences calculated here actually reflect both mental and physical QoL outcomes. Despite the small effect size, it remains statistically significant, suggesting clinical relevance. Additionally, when we adjusted for baseline characteristics, we still observed a significant difference in QALYs between the two groups (Appendix Table 6 and Appendix Table 7).

Q24: Table 1 and table 2 might be better swoped

Response: Thanks for your point. Our Table 1 primarily presents the detailed costs for the intervention and control groups post-intervention, from a descriptive analysis perspective, while Table 2 is used to analyze the differences between the two groups, focusing on both direct and indirect costs. We haven't made any changes yet, but we would be happy to hear any further suggestions you may have.

Q25: Table 1 does not include SDs

Response: Thanks for your point. We have revised Table 1 accordingly.

Q26: Please mention in table 2 the measures, what do the numbers indicate.

Response: Thanks for your point. We have added the measurement details in Table 2.

Discussion:

Q27: Could be helpful to convert all costs to one currency in the discussion to improve comparability of the findings.

Response: Thanks for your suggestion. We have converted all currencies to US dollars.

Q28: Unclear what gaps and flaws are addressed in this paper ("This study strengthens the evidence base for managing patients with diabetes 462 and depression, addressing gaps identified in systematic review 18. The review points out 463 limitations in experimental study evidence and methodological flaws hindering 464 comprehensive analysis from various perspectives 18.")

Response: Thanks for your point. We have revised this statement (Line 527-530).

Q29: What was the study duration of the previous studies referred to in paragraph 423-438?

Response: Thanks for your point. We reviewed the previous studies and found that the duration ranged from 6 months to 24 months.

-Cost-effectiveness of a Multicondition Collaborative Care Intervention: 24 months

-Cost-Effectiveness Analysis of Collaborative Care Management of Major Depression among Low-Income: 6 months

-Long-term clinical and cost-effectiveness of collaborative care (versus usual care)

for people with mental–physical multimorbidity: cluster-randomised trial: 24 months

-Cost-Effectiveness Evaluation of Collaborative Care for Diabetes and Depression in Primary Care: 12 months

Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Dear Reviewers,

Thank you so much for taking your time and effort to review our manuscript entitled *"Cost-Effectiveness of Community-based Integrated Care Model for Patients with Diabetes and Depressive Symptoms"*. We appreciate all your valuable comments and suggestions, which are very helpful in improving the quality of our paper. All the comments raised by reviewers have been addressed properly in the revised manuscript, which have been clearly marked in red font color in the paper. Please find our itemized responses and revised manuscript in the re-submitted files.

Reviewer #3 (Remarks to the Author):

Thank you for your revised manuscript and for addressing the comments provided. The corrections for points Q2 and Q3 are satisfactory, and I appreciate your effort in implementing the changes. However, the revisions for points Q1 and Q4 do not fully address my initial requests:

Q1: The clustered bootstrap method should be applied to the main analysis, rather than as an additional sensitivity analysis. Please revise accordingly.

Response: Thank you for your suggestion. We have moved these bootstrap results to main analysis.

Q2: Since these are fixed costs, calculating a mean and standard deviation is not appropriate. Instead, you should simply report the fixed cost applied without any statistical analysis.

Response: Thank you for your suggestion. Since we need to calculate the cost per person, we report these fixed costs using the mean value.

Reviewer #4 (Remarks to the Author):

Overall: hard to assess what was actually changes as no track changes or anything was used and the mentioned line number do not match those in manuscript.

Q1:there is still a mix of diabetes en T2DM found throughout the manuscript.

Response: Thanks for your point. We have revised that to T2DM thoroughly.

Q2:question not addressed, authors reference a paper, but in the paper no details on standard services are provided either. Besides, it is important that a reader can understand the intervention without reading another article by briefly describing the main elements.

Response: Thanks for your suggestion. We added the intervention details into “Comparators” section (Line146-148).

Q3:This also apply for Q17, were reference is made to a previous publication.

Response: Thanks for the suggestion. We added the details of screening and recruitment into the “Participants” section (Line 110-113).

Q4:We would suggest to remove the sentence that baseline characteristics are comparable, as this is misleading. It might be good to shortly mention the similarities and differences.

Response: Thanks for the point. We made revisions (Line 302-307).

Q5:this might be relevant to reflect on in the discussion. And we do not agree that a statistical difference suggest clinical relevance, as a statistical difference only indicates the likelihood of an actual difference but does not specify anything in terms of clinical relevance. So good to reflect on the clinical relevance of the difference in the discussion.

Response: Thanks for the suggestion. We have addressed the clinical relevance in “Discussion” section (Line491-495).

Review: Cost-Effectiveness of Community-based Integrated Care Model for Patients with Diabetes and Depressive Symptoms

Key results

The authors assessed cost-effectiveness of a community-based integrated care model for patient with type 2 diabetes and depressive symptoms using a cluster RCT. They reported higher health care costs in the intervention group compared to usual care, but with a high likelihood of cost-effectiveness when exploring gains in QALYs and DFDs.

Assessing the costs-effectiveness of an integrated care initiative is an important step of implementation and adds valuable information. However, we have some questions/ comments for the authors:

This manuscript was co-reviewed.

General comments

- We would recommend to be consistent in the key words/ concepts like T2DM vs diabetes and primary health care vs primary healthcare, US vs usual care.
- Superscript refences is placed behind the full stop instead of before.
- Grammar, spelling en layout including the reference list check is recommended. Also check correct use of past, present voice in the discussion section

Abstract

- How many CHC were randomized considering the cluster RCT design.
- Was an ITT analysis performed
- We would advise to mention the costs in US dollars for interpretation purpose.
- Is there a national cut-off for a QUALY in China, good to reflect on that
- We prefer if subheadings like background, methods, results, discussion are used.

Introduction:

- Could be written more concisely by removing some repeated/obvious statements and clarifying some claims. For example, it seems obvious that China is the country with the largest number of diabetes patients globally, as it is the country with the largest population in general. What about proportionally?
- Potentially shortly elaborate on why the CIC-PDD is effective/why you would use an integrated care initiative to improve care.
- The sentence “Moreover, the outcomes of diabetes and depression multimorbidity are worse when they appear together’ is not clear, do the authors mean that outcomes of diabetes and depression are separately worse than outcomes when diabetes and depression co-exists? I assume the authors mean the other way around.

Methods:

- Explanation for significantly larger group opting out of the intervention group and possibility of this causing selection bias?
- In healthcare utilization it is unclear what costs are included, specifically the inpatient visits: “[…] with a specific focus on those resulting from diabetes-related complications”.

- Unclear how the distinction between societal costs and multipayer perspective is made and what this distinction actually reflects.
    - Relevance of both is questionable as they only include costs directly related to care, and not costs such as diabetes/depression related sick leave? It would be understandable that this information is left out if it wasn't gathered, but than it might be better to leave out this category completely as it might not reflect actual societal costs well.
- Are sex-differences explored?
- Did the usual care group receive no depression-related treatment? As in the discussion it is mentioned that there's a shortage. Cause then are the authors really measuring (cost-)effectiveness of the integrated care approach or simply some sort of mental health care?
- It is not completely clear if the aim of the intervention is to screen on depression in people with T2DM or to screen and treat depression in this population. Overall the manuscript would benefit from some explanation of the intervention including whether or not it was found to be effective.
- In which register is the trial registered and was ethical approval granted
- Was the study powered for cost-effectiveness or was this performed for another outcome.
- Why was the SF-12 used instead of the EQ5D?
- It would be helpful to include in the flow chart how many individuals within the centers were eligible for the intervention.

Results:

- Comparable baseline characteristics, based on what? When comparing percentages they do seem to differ extensively on gender, residency, and income. Could this be a result of selection bias as indicated earlier? (Proportionally more males, urban and affluent in the intervention group)
- Results on the DFDs are hard to interpret and unclear whether the difference in QALY's is clinically important as the CIs largely overlap.
- Table 1 and table 2 might be better swoped
- Table 1 does not include SDs
- Please mention in table 2 the measures, what do the numbers indicate.

Discussion:

- Could be helpful to convert all costs to one currency in the discussion to improve comparability of the findings.
- Unclear what gaps and flaws are addressed in this paper ("This study strengthens the evidence base for managing patients with diabetes 462 and depression, addressing gaps identified in systematic review 18. The review points out 463 limitations in experimental study evidence and methodological flaws hindering 464 comprehensive analysis from various perspectives 18.")
- What was the study duration of the previous studies referred to in paragraph 423-438?