## PEER REVIEW HISTORY



## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | Effects of therapeutic play on children undergoing cast-removal procedures: a randomised controlled trial |
| **AUTHORS** | Wong, Cho Lee; Ip, Wan Yim; Kwok, Blondi Ming Chau; Choi, Kai Chow; Ng, Bobby King Wah; Chan, Carmen Wing Han |

## VERSION 1 – REVIEW

| | |
|---|---|
| **REVIEWER** | Alberto Dionigi<br>Federazione Nazionale Clowndottori - ITALY |
| **REVIEW RETURNED** | 08-Jan-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | I have now had a chance to read and review the paper you recently sent to me. Overall, it is a generally well-written and well-conducted study on a therapeutic play in children undergoing cast removal. The paper does make a contribution and I would recommend that it be accepted for publication.<br>I also have a number of specific comments about the paper. Thec omments are mostly suggestions about how to make certain points in the manuscript clearer for the reader.<br>Perhaps the most important of these is to have the authors add more information in the introduction about the specific aspects of the medical risks and the structured activities therapeutic play is made of.<br><br>I would suggest to enlarge the part of the Theoretical Framework where the authors state that children feel more stress if they do not control the situation they are living.<br>At pag. 10 I reccond to add some information about the experience of the play specialist.<br><br>Pag. 11: please clarify whether children know that they will receive a reward, as it can influence the results.<br><br>pag. 12 It is not clear to this reviewer why two alphas are reported for the CSAS_C<br><br>Pag. 13 A major concern about the paragraph on HEART RATE MONITORING. No cut off and general info are reported. I do reccomend to provide more information to the reader.<br><br>Pag. 14<br>Not clear how the Experimantal and the Control Group were identified: Did the children stay in the same room?<br><br>Please clarify why VAS was filled in by parents for patient more than 5, if researchers state that this is an istrument 3-7. It doesn't make sense. |

| | |
|---|---|
| | Pag 15: A major concern is about the reserach nurse who filled in the CEMS: she is not blind and the results may have been influenced by her, as she knew which patients belong to each group...<br><br>Pag. 16: Authors state that VAS showed moderate (.36) to high (.50) correlations but no significance is reported. |

| | |
|---|---|
| **REVIEWER** | Marta Tremolada<br>Department of Developmental and Social Psychology, University of Padua, Italy |
| **REVIEW RETURNED** | 11-Jan-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | This manuscript is really original and innovative highlighting the importance of<br>integrating therapeutic play into standard care. I suggest some minor revisions to make the manuscript more fully scientific valuable.<br><br>Introduction<br>1. I suggest to insert some quotations on coping with pain in children, even if dealing with other chronic conditions such as cancer, stressing the importance of parental perception on child's coping with procedures and hospitalization. A possible quotation could be: Tremolada, M., Bonichini, S., Basso, G., Pillon, M. (2015). Coping with pain in children with leukemia, International Journal of Cancer Research and Prevention, 8, 451-466. I suggest to insert this concept at page 19 line 36.<br>Participants<br>2. The exclusion criteria showed that children with neurologic and developmental problems were not taken into consideration in the therapeutic play. Why this decision? Clarify this point<br>Measures<br>3. Parents should be involved also adopting some self- or proxy-report questionnaires, not only satisfaction with the therapeutic play. The involvement of parents during the distraction should be considered as an important mediator/predictor. You could insert this concept as a limit or as a recommendation for future research.<br>Results<br>4. No possible gender differences were checked.<br>Limitations<br>5. Limit shown at page 20, lines 36-40: children could not be influenced by knowing that they are in intervention group, but parents yes…anxiety perceptions in children under 5 years old reported by parents and parent's satisfaction could be influenced by parent's symptomatology.<br>6. No parents' own anxiety was assessed and this variable that could influence also child's anxiety perceptions.<br>7. Another possible future direction could be to assess also children's coping styles or temperament, possible key factors associated with the anxiety and play predisposition<br>8. Another suggestion could be to propose specific occupational and play therapy in the hospitals that can reduce anxiety but also implement motor abilities of children that underwent more invasive procedures such as cancer (at this purpose you can cite this paper: Taverna L, Tremolada M, Bonichini S, Tosetto B, Basso G, Messina C, et al. (2017) Motor skill delays in pre-school children with leukemia one year after treatment: Hematopoietic stem cell transplantation therapy as an important risk factor. PLoS ONE 12(10): e0186787. https://doi.org/10.1371/journal.pone.0186787 |

| | |
|---|---|
| **REVIEWER** | Peter Herbison<br>University of Otago, New Zealand |
| **REVIEW RETURNED** | 02-Feb-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | This is a worthwhile study that appears to have been well conducted. But I have a few issues with the analysis and reporting.<br><br>Most of the outcomes are continuous and have been recorded at baseline and the end of the study. Some have been measured during the study as well and so have three values. The authors say they used GEE models to analyse this data. When you have continuous baseline and follow up data the most common method of analysis is to use ANCOVA. This looks at the differences at follow up adjusted for the baseline values. Now GEE models will take the baseline values into account and any differences at baseline will count as differences between the treatments. When there are more than the two measurements GEE models are more appropriate but the baseline values should be used as covariates rather than as another measurement of the outcome.<br><br>Table one has a column where a p-value is given for the differences in baseline variables between treatment groups. This is inappropriate as this is a randomised trial and any differences are certain to be there because of chance. There are other, better, ways of seeing if the baseline differences affect the outcome.<br><br>The results for the outcomes start with the within group results. This is a randomised trial so what is of interest is the difference between treatments. ANCOVA does give an estimate of the difference between treatments and this should be reported. Any within treatment differences can be left to readers to work out from the data in table 2.<br><br>The results duplicate a lot of what is in table 2.<br><br>The abstract appears to only report on the younger age group, as does the first paragraph of the discussion which should have a brief summary of all the results.<br><br>It would be preferable if the comments on power used an important difference in the outcome variables, rather than differences others have found.<br><br>On page 13 the abbreviation CVI is used without explanation.<br><br>The section on satisfaction levels of parents and cast technicians does not makes sense. It talks about "higher" without saying higher than what.<br><br>In table 2 at the bottom of the table it says that superscript a is "P-value testing for differential change of heart rate at the underlying time point with respect to T1 by using GEE<br>model;" but many of the results labelled a have nothing to do with heart rate. |

| | |
|---|---|
| **REVIEWER** | Kush Kapur<br>Harvard Medical School, USA |
| **REVIEW RETURNED** | 07-Feb-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors have clearly defined their study design and the statistical methods employed by them are mostly adequate. However, I believe the authors should address my following minor comments in order to improve the overall presentation of their work:<br><br>1. Sample size calculation on Page 9 does not correspond to the GEE modeling technique used to for the analysis. Also, this section does not contain any detail of the clinical implication of chosen effect size of 0.6 SD.<br>2. Based on the sample size calculation on Page 9, the study should have recruited 53 subjects per arm. The final sample contains 55 in control group and 52 in the intervention group. The reason for the imbalance should be clarified in the discussion.<br>3. Generalized Estimating Equations do not provide estimates under Missing at random. This should be corrected in the data analysis section.<br>4. The issue regarding the multiple comparisons hasn't been addressed at all in this manuscript (Table 2).<br>5. The results on Page 16-17 should contain the observed effect estimates of the correlations and effect sizes along with standard error and/or 95% confidence intervals.<br>6. Table 2: Clarification of the effect size. Are the VAS, CSAS-C and Heart rate effect sizes denote the standardized difference of the change in the control and the intervention group? Please clarify this in the footnote.<br>7. I encourage the authors to read the following articles and modify their results and discussion section accordingly– https://peerj.com/articles/3544/ "The widespread use of 'statistical significance' (generally interpreted as 'p ≤ 0.05') as a license for making a claim of a scientific finding (or implied truth) leads to considerable distortion of the scientific process" (Wasserstein & Lazar, 2016)." |

**VERSION 1 – AUTHOR RESPONSE**

**BMJ Open – Response to editorial team and reviewers' comments**

Thank you very much for your valuable comment. I have revised the manuscript accordingly and please find my response to reviewer's comments below.

**Editorial Team**

**Comment (1):** We note that this study was prospectively registered with the Chinese Clinical Trials Registry (http://www.chictr.org.cn/hvshowproject.aspx?id=7178) and also with the Centre for Clinical Research and Biostatistics. The details of the former should be included after the abstract instead of the latter. This is because we only consider trials registered with a WHO/ ICMJE approved registry and the Centre for Clinical Research and Biostatistics does not appear to be approved by the WHO/ ICMJE.

Response to comment (1): Thank you for the comment! The information regarding the registration with the Centre for Clinical Research and Biostatistics has been deleted. We have included the Chinese Clinical Trials Registry with registration number (ChiCTR-IOR-15006822) in the abstract.

**Comment (2):** Please also clarify whether the study protocol has been published anywhere.

Response to comment (2): Please kindly note that the study protocol has not been published elsewhere.

**Comment (3):** Play therapy before cast removal is an interesting research topic but the study seems rather obvious. Please provide a better explanation for doing this study in the introduction section.

Response to comment (3): Thank you for your comments! We have strengthened the introduction section by providing more information on: 1) what is this study about – Cast-removal procedure is stressful to children but previous interventions were shown not to be very effective in reducing anxiety of children; 2) why is this study needed – Although previous studies suggested that hospitalized children who were engaged in therapeutic play exhibited fewer negative emotions and experienced lower levels of anxiety than those who were not, these studies either are limited by some methodological flaws, did not present the play procedures involved clearly, or did not explore the comprehensive value of therapeutic play on the children, parents and health care provider as a whole. As such, this study is needed to provide more information in this area; 3) what questions does this study attempt to answer – To the best of our knowledge, only few studies have been conducted to examine the effects of therapeutic play in reducing anxiety and negative emotional manifestations among children undergoing cast-removal procedures. Our study addresses the limitations of previous studies mentioned above. It also aims to contribute further evidence on the benefit of therapeutic play by assessing the satisfaction ratings of parents and cast technicians in respect to the cast-removal procedures as well.

**Comment (4):** The Abstract is rather scant. You would need to explain and report the 3-7 year olds separately from the 8-12 year olds. Can you also elaborate on why children were stratified by age in this way? The Abstract also needs CIs not just p values.

Response to comment (4): Thank you for the comments! Based on your comments, we have added two tables (Table 2 and Table 3) to report the outcome measures at various time points between the intervention and control groups among children in the two age groups (3-7 years; 8-12 years). We have revised the manuscript to report the 3-7 years old separately from the 8-12 years old with both CIs and p values indicated. However, for some outcome measures such as emotional manifestation and parent/ technician satisfaction, we have reported the results for all children because similar findings were obtained among children aged 3-7 and 8-12.

The reason for stratifying children by age in this way is because children aged 3-7 belong to the pre-operational stage, while those aged 8 -12 belong to the concrete operational stage, according to Piaget's (1963) theory of cognitive development. Children in different age groups are at different stages of psychosocial development, and are therefore likely react to cast-removal procedures and therapeutic play differently. Therefore, we stratified children by age to provide more information on the effects of therapeutic play on children in different age groups.

**Comment (5):** The outcomes listed in the clinical trials registry are quite vague in comparison to what is reported in your manuscript. Can you please comment on this? Please also specify in your manuscript which outcomes are primary and which are secondary.

Response to comment (5): Thank you for your comment! We have listed the outcomes such as anxiety and emotional distress of children, and satisfaction levels of parents and cast technicians in the clinical trials registry. We admit that the outcomes listed were not clear enough. However, the outcomes listed were consistent with those reported in the manuscript.

In response to your comment, we have now specified the primary outcomes (anxiety and emotional manifestation) and secondary outcomes (e.g. parent and staff satisfaction) in the manuscript.

**Comment (6):** Please elaborate on how you calculated your sample size. There needs to enough information in the paper (page 8) to reproduce the sample size calculation. What outcome is this calculation based on? Why has a Cohen's d of 0.6 been chosen?

Response to comment (6): The sample size estimation of the study was based on the main outcomes of anxiety level and emotional manifestation with reference to previously published studies on therapeutic play to guide the selection of a minimum detectable effect. We have amended the section reporting sample size estimation accordingly.

**Comment (7):** There are quite a few limitations, notably the lack of blinding. These limitations need to be discussed more extensively in the discussion section.

Response to comment (7): We understood that the lack of blinding was one of the limitations of this study. We have therefore revised the manuscript and discussed it more extensively in the manuscript.

**Reviewer 1**

**Comment (1):** I also have a number of specific comments about the paper. The comments are mostly suggestions about how to make certain points in the manuscript clearer for the reader. Perhaps the most important of these is to have the authors add more information in the introduction about the specific aspects of the medical risks and the structured activities therapeutic play is made of.

Response to comment (1): Thank you very much for your comments! We have added more information in the introduction section on the medical risk of therapeutic play and the activities that it involves.

**Comment (2):** I would suggest to enlarge the part of the Theoretical Framework where the authors state that children feel more stress if they do not control the situation they are living.

Response to comment (2): Thank you for your comment! We have added further information to the Theoretical Framework section and revised the manuscript accordingly.

**Comment (3):** At pag. 10 Add some information about the experience of the play specialist.

Response to comment (3): The experience of the play specialist has been added to P.10. In brief, the play specialist in this study is an experienced and well-trained senior hospital play specialist (HPS). She has more than five years of experience in delivering therapeutic play—including preparation play and distraction play - to children undergoing medical treatments in various units of hospitals.

**Comment (4):** Pag. 11: please clarify whether children know that they will receive a reward, as it can influence the results.

Response to comment (4): The children were not informed that they will receive a reward upon completion of the intervention. We have revised the manuscript by including this information.

**Comment (5):** Pag. 12 It is not clear to this reviewer why two alphas are reported for the CSAS_C.

Response to comment (5): As the participants in this study were required to complete the questionnaire both before and after the cast-removal procedures, two alphas were reported in the paper which indicate the data collected at both time points.

**Comment (6):** Pag. 13 A major concern about the paragraph on HEART RATE MONITORING. No cut off and general info are reported. I do recommend to provide more information to the reader.

Response to comment (6): Thank you very much for your comments! We have now added information indicating the participants' heart rate in the manuscript. As heart rate varies across individual, we did not provide a cut off value for categorizing heart rate among the participants in the manuscript.

**Comment (7):** Pag. 14 Not clear how the Experimental and the Control Group were identified: Did the children stay in the same room?

Response to comment (7): Children with their casts removed would be requested to wait in a waiting area outside the cast room of the out-patient department in a separate timeslot. The research nurse would approach the children and their accompanying parent in the waiting area. The research nurse would then invite the parent-children pairs to an interview in a private room in order to collect the baseline data.

**Comment (8):** Please clarify why VAS was filled in by parents for patient more than 5, if researchers state that this is an instrument 3-7. It doesn't make sense.

Response to comment (8): Thanks for your comments! VAS is a valid and reliable measure to assess children's anxiety level, and it was commonly utilized in previous studies for children to self-report their anxiety. However, as children aged 3 or 4 may have limited verbal expression abilities, parents of children under 5 were also invited to fill in the VAS in order to obtain additional information on children's anxiety level perceived by their parents.

**Comment (9):** Pag 15: A major concern is about the research nurse who filled in the CEMS: she is not blind and the results may have been influenced by her, as she knew which patients belong to each group...

Response to comment (9): Thanks for your comments! We agree with you that the research nurse who filled in the CEMS was not blinded. However, blinding the research nurse was difficult and we admit that this was one of the limitations of this study. Therefore, we adopt both objective (heart rate) and subject measures (children reported anxiety level) to assess the effects of therapeutic play on children.

**Comment (10):** Pag. 16: Authors state that VAS showed moderate (.36) to high (.50) correlations but no significance is reported.

Response to comment (10): We have revised the manuscript and added the CI and p-value.

**Reviewer 2**

**Comment (1):** I suggest to insert some quotations on coping with pain in children, even if dealing with other chronic conditions such as cancer, stressing the importance of parental perception on child's coping with procedures and hospitalization. A possible quotation could be: Tremolada, M., Bonichini, S., Basso, G., Pillon, M. (2015). Coping with pain in children with leukemia, International Journal of Cancer Research and Prevention, 8, 451-466. I suggest to insert this concept at page 19 line 36.

Response to comment (1): Thank you for your information and suggestion! We have cited this paper in our manuscript to emphasize the importance of parental perception on child's coping with medical procedures.

**Comment (2):** The exclusion criteria showed that children with neurologic and developmental problems were not taken into consideration in the therapeutic play. Why this decision? Clarify this point.

Response to comment (2): Thanks for your question. We did not include children with neurologic and developmental problems in this study because we have concern that these children may have difficulty in understanding or responding to the questionnaires used in this study.

**Comment (3):** Parents should be involved also adopting some self- or proxy-report questionnaires, not only satisfaction with the therapeutic play. The involvement of parents during the distraction should be considered as an important mediator/predictor. You could insert this concept as a limit or as a recommendation for future research.

Response to comment (3): Thank you for your suggestion! We agree with you and have included your points in the recommendation section of this manuscript.

**Comment (4):** No possible gender differences were checked.

Response to comment (4): We have checked for possible gender differences and no gender differences were found. As gender difference was not our main concern, we did not report it in the manuscript.

**Comment (5):** Limit shown at page 20, lines 36-40: children could not be influenced by knowing that they are in intervention group, but parents yes…anxiety perceptions in children under 5 years old reported by parents and parent's satisfaction could be influenced by parent's symptomatology.

Response to comment (5): Thank you for your comments! We agree with you and therefore have included your points in the limitation section of the manuscript.

**Comment (6):** No parents' own anxiety was assessed and this variable that could influence also child's anxiety perceptions.

Response to comment (6): Thank you for your comments! We agree with you and therefore have included the assessment of parents' own anxiety as a recommendation for future studies.

**Comment (7):** Another possible future direction could be to assess also children's coping styles or temperament, possible key factors associated with the anxiety and play predisposition

Response to comment (7): Thank you! We agree with you and have included the suggested parameters in the recommendation for assessment in future studies.

**Comment (8):** Another suggestion could be to propose specific occupational and play therapy in the hospitals that can reduce anxiety but also implement motor abilities of children that underwent more invasive procedures such as cancer (at this purpose you can cite this paper: Taverna L, Tremolada M, Bonichini S, Tosetto B, Basso G, Messina C, et al. (2017) Motor skill delays in pre-school children

with leukemia one year after treatment: Hematopoietic stem cell transplantation therapy as an important risk factor. PLoS ONE 12(10): e0186787. https://doi.org/10.1371/journal.pone.0186787

Response to comment (8): Thank you for your information! We have included this piece of information in our manuscript and cited the suggested paper.

**Reviewer: 3**

**Comment (1):** Most of the outcomes are continuous and have been recorded at baseline and the end of the study. Some have been measured during the study as well and so have three values. The authors say they used GEE models to analyse this data. When you have continuous baseline and follow up data the most common method of analysis is to use ANCOVA. This looks at the differences at follow up adjusted for the baseline values. Now GEE models will take the baseline values into account and any differences at baseline will count as differences between the treatments. When there are more than the two measurements GEE models are more appropriate but the baseline values should be used as covariates rather than as another measurement of the outcome.

Response to comment 1: We do not agree that ANCOVA is superior to GEE even when only two measurement time points were involved in the study, as the autocorrelation of outcome measures within each subject would be completely ignored if ANCOVA is used. We used the GEE model to estimate the mean change on each outcome between groups with adjustment for the baseline group difference and accounting for autocorrelation of the outcome across time points.

**Comment (2):** Table one has a column where a p-value is given for the differences in baseline variables between treatment groups. This is inappropriate as this is a randomised trial and any differences are certain to be there because of chance. There are other, better, ways of seeing if the baseline differences affect the outcome.

Response to comment (2): We agree that it is unnecessary and indeed inappropriate to test the baseline differences between groups. We have now removed the p-value column in Table 1.

**Comment (3):** The results for the outcomes start with the within group results. This is a randomized trial so what is of interest is the difference between treatments. ANCOVA does give an estimate of the difference between treatments and this should be reported. Any within treatment differences can be left to readers to work out from the data in table 2.

Response to comment (3): Sorry for the confusion. We actually have not compared the within group differences for the outcomes. We only described their mean changes in each group but the comparisons were made on a between-group basis.

**Comment (4):** The abstract appears to only report on the younger age group, as does the first paragraph of the discussion which should have a brief summary of all the results.

Response to comment (4): Thank you for your comments! We have revised the abstract and provide a brief summary of all the results.

**Comment (5):** It would be preferable if the comments on power used an important difference in the outcome variables, rather than differences others have found.

Response to comment (5): Thank you for your suggestion! We have included difference in the outcome variables rather than differences others have found.

**Comment (6):** On page 13 the abbreviation CVI is used without explanation.

Response to comment (6): Thanks for your comment! We have deleted the abbreviation and CVI is now written in full instead.

**Comment (7):** The section on satisfaction levels of parents and cast technicians does not makes sense. It talks about "higher" without saying higher than what.

Response to comment (7): Thanks for your comment! In fact, the term “higher” refers to the higher satisfaction level of parents and cast-technicians in the intervention group toward the cast-removal procedure, compared to those in the control group. We have revised the sentence to make the statement clearer.

**Comment (8):** In table 2 at the bottom of the table it says that superscript a is "P-value testing for differential change of heart rate at the underlying time point with respect to T1 by using GEE model;" but many of the results labelled a have nothing to do with heart rate.

Response to comment (8): Thanks for pointing this out. We have revised superscript ‘a’ accordingly.

**Reviewer: 4**

**Comment (1):** Sample size calculation on Page 9 does not correspond to the GEE modeling technique used to for the analysis. Also, this section does not contain any detail of the clinical implication of chosen effect size of 0.6 SD.

Response to comment (1): To the best of our knowledge, there is no sample size estimation algorithm for GEE modeling available in any power analysis software. Our sample size estimation was based on the results of independent t-test, which should yield a greater required sample size as GEE is generally more powerful. Since there are no clinically relevant differences on our main outcomes of anxiety level and emotional manifestation, we made reference to previous studies on therapeutic play to guide the selection of a minimum detectable effect of 0.6 SD in both outcomes. Please refer to the amended sample size.

**Comment (2):** Based on the sample size calculation on Page 9, the study should have recruited 53 subjects per arm. The final sample contains 55 in control group and 52 in the intervention group.

Response to comment (2): Thank you for your comments. The reason for the difference in the number of subjects in the two groups is likely to be the stratification of participants by two age groups (3–7 and 8–12 years) during randomization. Further, the difference was further contributed by the difference in the number of patients in both age groups on the last day of data collection.

**Comment (3):** Generalized Estimating Equations do not provide estimates under Missing at random. This should be corrected in the data analysis section.

Response to comment (3): Thank you for pointing this out. We have revised the data analysis section accordingly.

**Comment (4):** The issue regarding the multiple comparisons hasn't been addressed at all in this manuscript (Table 2).

Response to comment (4): Thank you for your comment! One of the primary aims of this study was to estimate the effects of the therapeutic play intervention on relevant stress-related outcomes in children undergoing cast-removal procedures. We believe we should place more emphasis on the intervention effects, not their statistical significance, in the reporting of findings. Potentially inflated overall false positive rate should therefore not be a main concern in our study.

**Comment (5):** The results on Page 16-17 should contain the observed effect estimates of the correlations and effect sizes along with standard error and/or 95% confidence intervals.

Response to comment (5): We have added the effect estimates and their 95% confidence intervals to the results section.

**Comment (6):** Table 2: Clarification of the effect size. Are the VAS, CSAS-C and Heart rate effect sizes denote the standardized difference of the change in the control and the intervention group? Please clarify this in the footnote.

Response to comment (6): Yes, they are the standardized differences of the changes in the control and intervention groups. Footnotes have been added to Table 2 to clarify the effect size.

**Comment (7):** I encourage the authors to read the following articles and modify their results and discussion section accordingly– https://peerj.com/articles/3544/ "The widespread use of 'statistical significance' (generally interpreted as 'p ≤ 0.05') as a license for making a claim of a scientific finding (or implied truth) leads to considerable distortion of the scientific process" (Wasserstein & Lazar, 2016)."

Response to comment (7): We understand that clinical relevance of our intervention effects should not be confused with statistical significance. However, in view of the lack of studies assessing the clinical relevance of the intervention effects on the study outcomes, we prefer to present our study findings by providing the effects estimates together with CIs and p values.

## VERSION 2 – REVIEW

| | |
|---|---|
| **REVIEWER** | Marta Tremolada<br>Department of Developmental and Social Psychology, University of Padua, Italy |
| **REVIEW RETURNED** | 23-Mar-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | The Authors fullfilled the requested modifications and especially they added some recommendations for future research. An appropriate review of the manuscript according to the requested issues. |

| | |
|---|---|
| **REVIEWER** | Alberto Dionigi<br>Federazione Nazionale Clowndottori, Italy |
| **REVIEW RETURNED** | 29-Mar-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | Thank you for having revised the manuscript.<br>I have no additional suggestions. |

| | |
|---|---|
| **REVIEWER** | Kush Kapur<br>Harvard Medical School, USA |
| **REVIEW RETURNED** | 30-Mar-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors have satisfactorily responded to all my comments. |

| | |
|---|---|
| **REVIEWER** | Peter Herbison<br>University of Otago, New Zealand |
| **REVIEW RETURNED** | 31-Mar-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | While from the authors response it appears that they have used the GEE analysis correctly this is not reflected in the methods. It should be clear that baseline measurements were included as co-variates in the model, rather than just being another measure at a different time.<br><br>Incidentally if there are just baseline and one follow up measurement then ANCOVA is the correct method of analysis and it does take the correlation between the baseline and follow up measurements into |

| | |
|---|---|
| | account.<br><br>I am happy with the other changes. |

## VERSION 2 – AUTHOR RESPONSE

BMJ Open – Response to reviewer's comment

Comment of Reviewer 3: While from the authors response it appears that they have used the GEE analysis correctly this is not reflected in the methods. It should be clear that baseline measurements were included as co-variates in the model, rather than just being another measure at a different time.

Response to comment: Thank you very much for your valuable comment. We have clarified that in the revised statistical analysis part: "Specifically, the GEE model was used to estimate the mean change on each outcome between group with adjustment for the baseline group difference and accounting for autocorrelation of the outcome across time".

## VERSION 3 – REVIEW

| | |
|---|---|
| **REVIEWER** | Peter Herbison<br>University of Otago, New Zealand |
| **REVIEW RETURNED** | 27-Apr-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | I have no further comments on this paper. |