Author’s response to reviews

Title: A randomized controlled trial of mental health interventions for survivors of systematic violence in Kurdistan, Northern Iraq

Authors:

Paul Bolton (pbolton1@jhu.edu)
Judith Bass (jbass1@jhu.edu)
Goran Zangana (gsabir@heartlandalliance.org)
Talar Kamal (talar_kamal@yahoo.com)
Sarah Murray (smmuray9@jhu.edu)
Debra Kaysen (dkaysen@uw.edu)
Carl Lejuez (clejuez@umd.edu)
Kristen Lindgren (kpl9716@uw.edu)
Sherry Pagoto (Sherry.Pagoto@umassmed.edu)
Laura Murray (lmurra15@jhu.edu)
Stephanie Skavenski (sskaven1@jhu.edu)
Nazar Amin (nazar.amin@gmail.com)
Michael Rosenblum (mrosen@jhu.edu)

Version: 2
Date: 20 November 2014

Author’s response to reviews: see over
Dear BMC Psychiatry Editors,

My co-authors and I are submitting a revised version of the manuscript titled ‘A randomized controlled trial of mental health interventions for survivors of systematic violence in Kurdistan, Northern Iraq’, based on your comments and those of the reviewers. Each of these comments is copied below, followed by our description, in italics, of how we have addressed that comment.

EDITOR’S REQUEST

Date of Trail Registration

The last section of the abstract should be Trial Registration: listing the trial registry and the unique identifying number and the date of registration, e.g. Trial registration: Current Controlled Trials ISRCTN73824458. Registered 28 September 2004. Please note that there should be no space between the letters and numbers of the trial registration number.

Trial Registration has been moved to the end of the abstract (after keywords). The date of registration has been added.

REVIEWERS

Reviewer 1:

1. The introduction was well-written and provided a good introduction to the literature, and strong background to the current study. A couple of sentences in the introduction could be clarified. On page five (line 80), the authors state that "Iraqi refugees in the US who reported torture had 4.32 and 2.08 times the odds of mental distress and physical health problems respectively". It would be helpful if the authors clarify what the reference for these odds ratios were (e.g. compared to those who had not been tortured?)
The reference group was indeed those who had not experienced torture. This has been added to the text.

2. On page 6, line 101 the authors state "At the time of the study, Kurdistan had experienced relatively little violence". This seems to contradict the next statement that there was significant trauma in the 1980s. Do the authors mean that, at the time of the study, Kurdistan was experiencing relatively little violence?

Yes. We have reworded the sentence as the reviewer suggested.

3. Could the authors clarify what they mean by "their perception of how they are treated by others" page 6, line 104?

Many respondents complained of being treated poorly by the community and even family; that the suffering and efforts of those who had experienced violence under the Saddam regime were not respected. We have added this sentence to the text to explain this point.

4. On page 9, line 179, the authors state that "instrument reliability and validity were tested for all outcomes among local survivors of systematic violence". Could the authors clarify if these individuals were part of the study sample? What was the interval between instrument administration for the indices of test-retest reliability?

Both studies were done in the same area using the same recruitment criteria but at different times. Therefore, some persons were recruited to both studies. This has been noted in the text. The interval between test and retest interviews was 1-3 days. We have added this information to the text. Also, the text mistakenly referred to combined test retest and inter rater reliability (repeating interviews by different interviewers). However, only test retest reliability was assessed (repeat interview by same interviewer). This has been corrected in the text.

5. The authors state on page 9, line 184 "In previous studies using this approach, we have found that instruments tend to perform either well or poorly across most tests". I'm not sure how useful this is as an argument for retaining the PTSD scale, as the authors report that, in this case, the instrument did not perform well across tests. I would suggest leaving this out, and just stating that this scale was retained despite psychometric limitations.

We feel the reviewer has misread the text: the PTSD scale performed adequately for both sexes on all tests, so removing it was not an issue. We also wrote that the other scales also performed at least adequately on all tests except for criterion validity, where it the other scales failed in women. Given that the other scales had problems on only one this one measure we decided to retain them.

Our point about scales usually performing similarly across measures was a suggestion that the problem may have been with the criterion validity testing process and not the measure. We have tried to make this clearer in the text.
6. If possible, it would be helpful to outline some of the basic principles and components of BATD and CPT. This would be useful for readers who are not familiar with these interventions. I also wasn’t sure what the authors meant by "aspects of the treatment with greater focus on the underlying behavior principles in BATDS were streamlined or removed". Could the authors provide examples of this?

We have added summaries of the principles and components of CPT and BATD in the Intervention section. Regarding the phrase in the text that "aspects of the treatment with greater focus on the underlying behavior principles in BATDS were streamlined or removed": The BATD instructors did not train the CMHWs on the psychological principles underlying BATD because understanding of the principles is not necessary to correct conduct of BATD. However, on sharing this reviewer comment with the BATD trainers, they pointed out to the lead author that this is true of basic BATD everywhere and so was not a modification in the approach for Kurdistan. We have therefore removed this from the text.

7. On page 14, line 283, the authors reported that, in addition to endorsement of key depression symptoms, participants must have a score of at least 29 to be eligible for the intervention. Could the authors provide detail on how they chose this number?

The cutoff score was 20 rather than 29. The choice was arbitrary because we had no pre-existing data for this population. We chose 20 because the depression instrument had 20 questions and therefore a score of 20 or more would indicate symptomatology on a significant proportion of the instrument – at least 7 symptoms. This explanation has been added to the text.

8. On page 16, line 346, the authors state that "models with and without governate level were not qualitatively different". Could they clarify what this means? Were intraclass correlation coefficients used to test for the impact of data clustering?

We generated models with and without governorate as a cluster variable and with and without it as a covariate and all 4 models gave very similar results for the outcomes. ICC were not used. We have inserted text to this effect. Also, we moved the explanation from the analysis section to the Results (Treatment Effects) as being the more appropriate place since it dealt with the results of analyses.

9. I thought the results and discussion were well written and clear. On page 22, line 482, the authors state that "A similar trial of CPT in Southern Iraq (unpublished data) showed roughly similar results." I'm not sure how much this adds to the paper given the readers do not have access to this study.

We have removed this from the text.

Reviewer 2

Suggestions for major compulsory revision:
The trial is adequately described according to the CONSORT guidelines. In the penultimate
paragraph of the introduction (line 114), the authors noted changes in the trial design (due to linguistic and other differences across populations in the selected study areas) that became evident after the trial had commenced. This section may be better placed under the method section, changes to trial design.

*We have incorporated this paragraph into the Methods section as requested.*

The authors stated they applied locally adapted outcome measures for depression, disability, PTSD and grief with locally understood syndromes included in the measures. Depression and disability were selected as primary outcomes. I suggest the authors provide a brief account of the additional local syndromes in text and dispense with Table 1a & b.

*For table 1a we feel that this is an efficient way of providing this information. Replacing it with text would either require a lot of text or the loss of a lot of information that we feel would be of interest to readers regarding of how the instruments were adapted. If the reviewer feels that the table is too long we could remove the standard items and leave only the added items. For table 1b there are no standard items from an existing instrument so the entire table would have to be represented by text. Since this is an editing question we are happy to comply with whatever the editors decide.*

Clarify “non-validated criterion for criterion validity” in line 186.

*By this we mean that the criterion was one that we expected to be a good indicator of the presence of a syndrome but was not proven to be accurate by previous studies. On reflection we consider this term to be misleading and have removed it from the text.*

The authors stated that they subjected both interventions to extensive local adaptation at various stages according to the population characteristics (with reference to a paper elsewhere). I suggest the authors provide a brief account of the core components of CPT and how they had been adapted to match the literacy levels (and other characteristics) of participants.

*We have added summaries of the principles and components of CPT in the Intervention section. We have included further information on how CPT was adapted, including adaptations to the low literacy levels.*

Additional information is needed in relation to the lay treatment providers (which has been detailed elsewhere) – e.g., age, gender, education, how they were selected.

*Information on provider education, gender, and how they were selected has been added in the section on Study Settings.*

Inclusion criteria are reasonably well defined but more information is needed on exclusion criteria (e.g., acute psychosis, social issues, etc).

*The exclusion criteria have been added to the section on study eligibility. These were inability to*
be interviewed due to a cognitive or physical disability, or severe suicidal ideation or behavior.

The authors stated that those who scored 20 (including one of the core symptoms of depression) or above on the modified HSCL were eligible for the study. This needs clarification – what is considered as positive endorsement for each symptom?

A positive endorsement for a symptom was an item score of 2 or 3, corresponding to experiencing a symptom often or always, respectively. This has been added to the text.

The range of the number of months for follow up assessment seems quite wide from 1.6 to 15.5 months (a mean of 5.5 months with perhaps a few outliers)–does this impact on the trial?

Although we did not anticipate such a wide range of post assessment times, we feel that this represents the reality of future programs using these types of interventions. Therefore, while this range may have impacted the results of the trial compared to the original design, they are probably closer to the results that would be achieved in normal programming. Notes to this effect have been added to the Limitations section.

The authors stated that according a post-hoc power calculation there was insufficient power to compare CPT and CPT-controls – how does this impact on the analysis?

We calculated that to have 80% power to detect the difference we found for CPT vs CPT controls would have required a substantially higher sample size of 385.

The authors stated that clustering at the governorate level was adjusted for and because there was little difference between the adjusted (with the governorate level) and unadjusted model (without), the most parsimonious model was selected – is this selection based on the degree of clustering in each model (or ICC)?

Reviewer 1 also asked about this, so we have inserted here the same explanation: We generated models with and without governorate as a cluster variable and with and without it as a covariate and all 4 models gave very similar results for the outcomes. ICC were not used. We have inserted text to this effect. Also, we moved the explanation from the analysis section to the Results (Treatment Effects) as being the more appropriate place since it dealt with the results of analyses.

The authors compared BADT and CPT to all controls in their initial analysis and then BADT to BADT matched controls and CPT to CPT-controls based on heterogeneity identified in the two treatment arms. The authors then did a third analysis using a permutation based method which essentially relaxes the homogeneity assumption in regression models. Could this method be applied to the initial analysis?

Unfortunately, the permutation based method that we used cannot be applied to the initial analysis. This is because we separately test the null hypothesis that BATD has no effect compared to control, and the null hypothesis that CPT has no effect compared to control. The
The authors included age, sex, marital status, and disability as covariates in their mixed effect models and stated that (line 343) "Additional variables that differed between treatment and control at baseline or predicted change in outcome were included as covariates (p<0.10)". Was education accounted for? This may be of concern given the widespread illiteracy amongst the study population.

We agree that education was a concern and therefore it was one of the variables that we tested for. Whether is was included in an analysis was based on the criteria noted by the reviewer – different at baseline or predicted change in outcome. On this basis education was included in some analyses but not others: In the CPT vs. CPT controls we controlled for both employment and education. In the BA vs. BA controls we controlled for any education vs. none (this is indicated in the notes under the tables). In the BA/CPT vs. All controls analyses we controlled for employment, not education. In that case it may be that employment was a proxy for education (or vice versa) since we would expect them to be related. The above descriptions of which covariates were included in each analysis are in the footnotes of the results tables so we have not added this material to the text.

Whether verbal and/or written consent was collected is unclear prior to intervention – this needs clarification.

*In the section on consent we have clarified that consent was verbal.*

**OTHER CHANGES**

We have made some additional changes to the text. In the Analysis section, paragraph beginning ‘we conducted a third analysis…’ we have added a reference to Rosenbaum et al’s paper as the correct source for the permutation-based method.

We have also added an author – Dr. Ahmed Ahmed. Dr. Ahmed was integral to the conduct of both the preliminary research and the trial itself and has contributed to the manuscript. His name was mistakenly omitted from the original submission.

We hope these change are satisfactory and thank you again for considering our manuscript for publication in BMC Psychiatry.

Sincerely,
Paul Bolton MB BS MPH MSc
Center for Refugee and Disaster Studies
Department of International Health
Johns Hopkins University Bloomberg School of Public Health.