PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<u>http://bmjopen.bmj.com/site/about/resources/checklist.pdf</u>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

(This paper received three reviews from its previous journal but only two reviewers agreed to published their review.)

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effectiveness of online interventions in preventing depression: A protocol for systematic review and meta-analysis of randomized controlled trials.
AUTHORS	Rigabert, Alina; Motrico, Emma; Moreno-Peral, Patricia; Resurrección, Davinia M.; Conejo-Cerón, Sonia; Navas-Campaña, Desirée; Bellón, Juan A.

VERSION 1 – REVIEW

REVIEWER	Hiran Thabrew
	University of Auckland, New Zealand
REVIEW RETURNED	12-Feb-2018

GENERAL COMMENTS	 This article outlines the protocol for a systematic review of online interventions for preventing depression. The article is well written and the PRISMA checklist is complete. Key points to address prior to publication are: 1. Reference to other relevant existing reviews in this area (such as the 2010 review of internet-based interventions for preventing depression in children and adolescents by Calear and Christiensen). 2. An explanation of the 'shared mechanism of action' mentioned in the paragraph titled "Type of interventions". 3. A description of how scales will be ranked from lowest to highest validity. It would help if the 5-10 most commonly used scales were explicitly ranked. If this is not done upfront, it could be a potential area of bias during the review. 4. An explanation regarding the timeframe of post-intervention data collection. Will it only include immediate post-intervention measures or will it additionally include 3,6,12 or longer-term measures. Otherwise, what happens if more than one post-intervention outcome has been collected. Will one be prioritised over the other(s)?
REVIEWER	Conceição Granja

REVIEWER	Conceição Granja
	Norwegian Centre for E-health Research
REVIEW RETURNED	15-May-2018

GENERAL COMMENTS	In the manuscript the authors frame their question very well using
	the PICO framework. I would like to have seen the keywords and
	index terms used in the searches structured as a PICO query.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1#:

Thank you for your time and for your comments, we appreciate the positive feedback from the reviewer.

Reference to other relevant existing reviews in this area (such as the 2010 review of internet-based interventions for preventing depression in children and adolescents by Calear and Christiensen).

Thank you for the reference. The research group of the Black Dog Institute and other researchers from Australia and New Zealand have contributed with numerous and relevant trials on the prevention of depression, some of them through online interventions. They have also contributed with some systematic reviews and meta-analyzes on the prevention of depression that are also relevant. The meta-analysis to which you refer [Calear AL, Christensen H. Systematic review of school-based prevention and early intervention programs for depression. J Adolesc. 2010 Jun; 33 (3): 429-38], does not focus on online interventions and additionally it included trials that did not rule out depression at baseline, so in this case they would not be focused on prevention (they include people already suffering from depression). In the last meta-analysis on this same topic that these authors have published (Werner-Seidler A, Perry Y, Calear AL, Newby JM, Christensen H. School-based depression and anxiety prevention programs for young people: A systematic review and goal -analysis, Clin Psychol Rev. 2017 Feb; 51: 30-47) they do not exclude depression at baseline nor focus exclusively on online interventions. For these reasons we have not cited them in our manuscript.

2. An explanation of the 'shared mechanism of action' mentioned in the paragraph titled "Type of interventions".

We have included an explanation of the 'shared mechanism of action': "We will only include RCTs assessing the effectiveness of *psychosocial and/or educational*, since they share the same mechanism of action that facilitates changes in attitudes and behaviours and because most interventions to prevent depression are of this type. *Educational interventions provide information sessions or fact sheets, whereas psychosocial interventions attempt to change how people* think and behave by using a variety of strategies (e.g. cognitive-behavioural or interpersonal). However, in real practice psychosocial and educational interventions can overlap, being difficult to distinguish them." (p.6).

3. A description of how scales will be ranked from lowest to highest validity. It would help if the 5-10 most commonly used scales were explicitly ranked. If this is not done upfront, it could be a potential area of bias during the review.

The validity and reliability parameters of the questionnaires vary according to the characteristics of the population to which they apply. For instance, PHQ has an acceptable sensitivity and specificity in primary care patients in the USA; however, in primary care patients in the Netherlands has a very low sensitivity. That is, if we gave a certain hierarchical position to this questionnaire based on the USA data and the trial that we include in our metaanalysis takes place in the Netherlands, we would be committing a bias. Other variables of the patients, as well as the country of origin, such as age (adolescents versus adults or the elderly), setting (primary care patients versus hospital patients), clinical characteristics (eg patients with cancer), etc. can also modify the validity and reliability parameters of the questionnaires. Therefore, when two or more instruments have been used to measure the outcome in the trials, we propose the following rule that we include in the text of the manuscript:

When more than a symptom scale has been used to measure outcomes in a RCT, the data from the highest validity scale will be employed. *"If the validation data of the scales, in the country and setting where the study was conducted, are not reported in the article, they will be searched in the literature and other sources. The parameters that will be used to select the scale of symptoms are: higher Youden's J statistic (J=Sensitivity+Specificity-1), Cronbach alpha, and Intraclass Correlation Coefficient (test-retest), and sensitivity to change over time (Yes/no/not available). For each trial, the scale of symptoms that provide more validation data and of higher quality will be chosen." (p.6).*

4. An explanation regarding the timeframe of post-intervention data collection. Will it only include immediate post-intervention measures or will it additionally include 3,6,12 or longer-term measures. Otherwise, what happens if more than one post-intervention outcome has been collected. Will one be prioritised over the other(s)?

It has been included the following sentence "and all follow-up provided from the RCTs"

Reviewer 2#:

Thank you for your time and for your comments, we appreciate the positive feedback from the reviewer.

In the manuscript the authors frame their question very well using the PICO framework. I would like to have seen the keywords and index terms used in the searches structured as a PICO query.

It has been modified the supplementary file 1 addressing this question.

Additionally, we included in the text: "A draft MEDLINE search strategy in PICOS format is included in *supplementary file." (p.7).*

VERSION 2 – REVIEW

REVIEWER	Hiran Thabrew University of Auckland, New Zealand
REVIEW RETURNED	10-Aug-2018
GENERAL COMMENTS	Thanks. All of my previous queries have been addressed by the
	authors.