Reviewer's report

Title: Impact of a mobile phone and web psychological program on depression, anxiety, stress, work and social functioning: A randomised controlled trial

Version: 1 Date: 30 September 2013

Reviewer: Jeroen Ruwaard

Reviewer's report:

This manuscript describes a large randomized controlled efficacy trial (N = 720) of a web-based / phone-based intervention for mild-to-moderate mental health problems, targeting depression, anxiety, stress, work and social functioning. On the short term (after seven weeks), the intervention was found to be more efficacious in reducing symptoms compared to an attention control condition and a waiting list condition. At 3-month follow-up, however, no additional benefits of the intervention were found in comparison to the (assumed inactive) attention control.

The subject of study is important, because the e-mental health field is embracing mobile technology to deliver psychotherapeutic interventions while there is little evidence with regard to the acceptance and efficacy of this strategy. This work provides one of the first indications in the form of a large-scale RCT.

The design of the study does not permit firm statements about the added benefit of mobile technology (which is acknowledged in the discussion). Also, the included sample size was lower than originally planned: N = 720 instead of N = 1200 (N = 2000 in the original trial register entry). Finally, results are mixed (i.e., short-term, but no long-term between-group effects). Overall, however, I feel this manuscript is of interest to those interested in e-mental health development, and the trial is well-executed. Therefore, I recommend publication, although a number of issues should be clarified. Comments and suggestions for possible improvement are listed below.

- Major Compulsory Revisions

1. Please be more explicit about the short-term efficacy of the program (i.e. by clearly stating that no differences between AC and MyCompass were found at the 3-month follow-up. Please do so in the abstract (in the results section) as well as in the discussion (first paragraph).

2. The effect sizes in Table 5 appear to be somewhat inflated, based on the results presented in Table 2 and 3. For instance, according to Table 2, participants in the myCompass group reported an increase of about 4 scale points on the WSAS. With a pooled standard deviation of - roughly - 8 [ (8.07 + 8.11 + 8.77) / 3 ], I
would expect a \( d \) of around .5, while the reported \( d \) is .84. Similar inflations appear to be present for the other measures. Also, between-group effect sizes are higher than one would expect based on the parameter results in Table 3. Please check. Did estimated means reflect more favourable outcomes than the raw observed means? Were pooled standard deviations calculated correctly?

- Minor Essential Revisions

Title
------

3.
Please consider to add targeted symptom severity to the title (i.e., mild-to-moderate depression, anxiety, stress, and work and social functioning).

Abstract:
-------------

4.
Please add effect sizes (point estimates and 95% confidence intervals) to the results section of the abstract, as well as adherence rates. These are key outcomes that readers would expect to find in the abstract.

5.
The last sentence of the conclusion ("Mobile phones offer …") is not necessarily grounded in the design and reported results of the study. In my opinion, this statement is too generic and should be best removed.

Introduction
----------------

5.
Please provide references to any previous published research or descriptions of myCompass and its theoretical basis here (e.g., Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: Review of the emerging field and proof of concept study. J Ment Health 2011; 20: 509-24.). The introduction provides a little too less information on the specifics of the program (is it CBT? solution-focused? eclectic, etc), which makes it hard to follow the authors in the hypothesis that the program will have beneficial effects. While this information is provided in the Methods section, I feel the manuscript would benefit from more background information at this point.

Material and Methods:
-------------------------------

6.
Please provide a short explanation for not reporting the outcomes on three secondary outcome measures that were included in the trial register (Subjective
Units of Distress Scale, Self-efficacy and health service utilization). What results were obtained with these measures? If these measures did not show any effect, please report so as well. Please also explain briefly why the number of measurements in the trial register entry and the final report differ (the trial register mentions repeated measurements at baseline and at 3, 6, 9, 12, 15 and 18 weeks after intervention commencement).

7. In the description of the intervention, please provide the exact number of psychological modules available to participants (especially since usage is reported in the results section). Any reference to a more extensive description of the treatment elements would also be helpful to readers who consider a replication study. Please also provide somewhat more detailed information with regard to the tailored suggestions that are made the myCompass program (and by what rule sets). For instance, what profile questionnaire was used?

8. Treatment satisfaction was assessed for myCompass, but not for the attention control. Please briefly explain why not.

9. Please provide score ranges for the WSAS as well.

10. Please consider to correct statistical results for chance findings resulting from multiple testing (i.e., by using Bonferroni corrections or similar procedures).

11. Satterthwaite’s approximation appears to have been used to estimate the degrees of freedom in the mixed model analyses? Please provide the method. In addition, it may not be necessary to report degrees of freedom up to 2 decimal points precision in the Tables (as there is little practical difference in df = 710 or df = 710.44).

12. The definition of Adherence (last paragraph in the statistical analyses subsection, just before the results section) should perhaps be moved to the Secondary Outcomes section, as it describes three measures, not an analysis.

Results
-------

13. Fifteen participants withdrew after randomization (n = 11 from the myCompass group). Their data was removed from analysis. Please explain why. I would have expected these data to be included in the intention-to-treat analyses.
14. Table 1: Would it be better not to use abbreviations in the table headers, or provide a list of terms in Table notes (e.g., AC: attention control)?

15. Table 1/2/3/4/5: Would it be wise to explain abbreviations used in each Table using a Table note? This would help readers a lot to quickly assess the data.

16. Table 2: Please provide N's for each group and assessment moment.

17. Table 2: "mCompass" -> myCompass.

18. Table 3 and Table 4: Please add p-values and parameter confidence intervals for all tests in the Tables, including those that were found to be non-significant.

19. Table 3/4: Please check denominator degrees of freedom. It seems odd that these exceed the sample size (n = 720). I would expect the df's to reflect the number of participants?

20. Table 4: please add explicit reference to the two groups that were compared in the analyses in the title (myCompass vs AC).

21. Table 4: the contrasts comparing AC to myCompass during the intervention phase are essentially identically to those presented in Table 3, if I understand correctly? If so, would it be better to remove these from Table 4, to prevent redundancy and confusion?

22. Table 5: In the method section, between-group and within-group Cohen’s d appears to be calculated using different standard deviations. As a result, there is a slightly confusing inconsistency between the within-group and between-group effects. For instance, if myCompass vs. WL is d = -.50 (in favour of myCompass), then readers could expect to find that if the within-effect in the myCompass group is .84, that the effect in the WL group would be .34. However, in the table, this effect is .38. Using the same pooled standard deviation for both between and within-group effect size calculations would perhaps provide a more clear and more consistent presentation?
23. Results/last paragraph before waitlist group results ("Moderate- to- large (d range = .44 to 1.04) 312 within-group effects were seen in the myCompass group"). Do you mean d = .84 instead of 1.04 as the maximum effect size (as listed in Table 5)?

Discussion: 
-------------

23. Please discuss possible negative effects of drop-out more. Mixed Modelling provides estimates that are based on observed values. It statistically corrects for attrition by increasing the uncertainty of the regression estimates. However, it can not be ruled out that study drop-outs would have reported more negative outcomes. With a 50% attrition rate in the MyCompass group (significantly higher than in the control groups), this is a major concern.

25. Statistical results were replicated in the wait list group, although within-group effect sizes were considerably smaller in this group after treatment. These results seem to contradict the statement that "The WL group achieved outcomes comparable to those observed for the myCompass group after using the program for seven weeks" (discussion/first paragraph). Please explain.

24. While the effects are quite large for a self-help program, I do not necessarily agree with the suggestion that the effects of MyCompass are as good as guided programs (second paragraph). Although there certainly are guided programs with similar effect sizes, there are also guided programs with which effect sizes have been obtained that are substantially greater (see, for example, the individual study results in the review by Andersson and Cuijpers, 2009). Therapist input can make a difference, and readers should be informed about this.

- Discretionary Revisions

26. While the authors announce future dismantling studies in the discussion, any information relating to the added benefit of the therapeutic elements delivered through the mobile phone would be helpful, since this is one the primary contributions of the study. For instance, it would be informative to learn whether usage of the mobile phone elements was related to better outcomes. Such exploratory analyses would perhaps increase the impact of the study?

27. Please consider to change signs of effect sizes so that positive effects indicate effects in the hypothesized direction and negative effects indicate effects in the
opposite direction.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests