Author's response to reviews

Title: Implementing an Outreaching, Preference-led Stepped Care Intervention Programme to Reduce Late Life Depressive Symptoms: Results of a Mixed-methods Study

Authors:

Ilse M.J. van Beljouw (imjbeljouw@gmail.com)
Miranda G.H. Laurant (m.laurant@iq.umcn.nl)
Marjolijn Heerings (m.laurant@iq.umcn.nl)
Max L. Stek (m.stek@ggzingeest.nl)
Harm W.J. van Marwijk (hwi.vanmarwijk@vumc.nl)
Eric van Exel (e.vanexel@ggzingeest.nl)

Version: 2 Date: 24 June 2014

Author's response to reviews: see over
Dear Dr. Straus,

Please find for your consideration a revised version of our paper “Implementing an Outreaching, Preference-led Stepped Care Intervention Programme to Reduce Late Life Depressive Symptoms: Results of a Mixed-methods Study”. We thank the reviewers and yourself for the careful review of our paper and we are pleased that the reviewers felt our study addresses important issues. The comments were very helpful and we think that by covering the issues raised we have strengthened our manuscript. We agree with nearly all of the points raised by the referees and we have worked hard to try to carefully address them.

We indicate our point of view on a comment between [ ] and then describe the changes we propose in the new manuscript. We have addressed these as follows:

**Reviewer 1: Peter Bower**

1. The introduction talks about preference led and stepped care, but the preference model (to me) seems to be a version of stepped care (choice within steps) rather than the more usual comparison of stepped and stratified care. This might be noted in the introduction. It is interesting that the current ‘stepped versus stratified’ debate misses one of the key points raised by the current paper – that certain populations of depressed patients may not perceive a high need for help.

   **Comment: [We agree with the reviewer]**

   **Proposed change:** The last paragraph on page 5 now reads: “Yet, current randomised clinical trials have compared stepped care models to stratified care, and thereby missed out on participants’ needs for care. It is unknown how a preference-led model – integrated in the provision of stepped care – would impact on the clinical effectiveness of such a programme. Also, it is uncertain if these models are also effective when implemented in routine primary care services.”

2. The use of the stepped wedge design is interesting, although may benefit from a more detailed introduction as they are not that conventional in this area. I know readers can refer to another paper, but it would be helpful if they didn’t have to. A description in a separate box might be best to provide detail without ruining the narrative flow.

   **Comment: [We agree with the reviewer]**

   **Proposed change:** We added Box 1 with a more detailed description of the design: “A stepped-wedge design is a type of crossover design in which the intervention is rolled out sequentially to participants over a number of time periods. All participants are recruited at the start of the study and assigned to several clusters. Starting moments are determined by (cluster-) randomisation and by the end of the study, all participants will have received the intervention. That is, clusters cross over from control to intervention condition. During the control condition, usual care is provided.”

3. From the reading of the description of the implementation of the programme, the reader might forget that the study showed a positive effect! It might be useful if the overall effect could be provided in a more standardised form (as an effect size) so as to encourage comparison with other depression interventions. The low rates of take up ARE noteworthy and a genuine issue for implementation science researchers, but many depression interventions show small effects. Is Lust for Life any different in terms of the overall effects, even in the context of poor take up? Is it necessarily an issue if the overall effect is reasonable? The data on uptake could also be compared with some recent studies in other contexts

   **Comment: [We agree with the reviewer]**

   **Proposed change:**

   - At the bottom of page 7, we added a sentence to help the reader interpret the size of the effect. Statistically, no formal effect measure can be calculated due to the nature of the analysis (GEE).
We agree with the reviewer that we were perhaps overly critical of our own results. We therefore removed the two (critical) sentences below that.

- We added a sentence to the discussion about uptake of interventions and added references and rewrote this paragraph to reflect the comments of the reviewer.

4. The authors clearly undertake a major data collection exercise and the description of it is very long and could be cut down – Figure 2 is very much more useful in this regard. The text was hard to read through and could detract from the much more interesting core issues.

Comment: [We agree with the reviewer]
Proposed change: We significantly shortened the section on qualitative data collection by referring to Figure 2.

5. The limitations of the use of proxies to collect data on non-respondents might be highlighted in the discussion

Comment: [We agree with the reviewer]
Proposed change: The following sentence was added to the section on strengths and limitations of the study: "Also, we could only use proxy-data to gather information on perspectives from non-respondents given the absence of their informed consent."

6. The themes from the qualitative interviews are relevant and interesting and I had no major concerns with the analyses. I felt the section on ‘illness perceptions’ (page 15) may have conflated beliefs about the existence of a ‘mental health problem’ with beliefs about the cause of the problems, which might be distinguished.

Comment: [We agree with the reviewer]
Proposed change: We separated the discussion of the way respondents labelled their emotional distress and the self-perceived causes of their distress. This paragraph now reads: “Qualitative results confirmed this finding: Approximately half of all subjects did not feel that (medical) terms such as having a depressed mood or being depressed were applicable to their situation, especially those who declined the intervention offer. Moreover [...] Next, the qualitative interviews shed a light on self-perceived causes of emotional distress. [...]"

7. The general lack of belief in the need for care is an important finding. I felt the authors could explore in more detail whether this reflects low expectations about the effectiveness of interventions, or and a belief that they are not deserving ‘candidates’ for care, or something different. These different explanations might have very different policy implications.

Comment: [We agree with the reviewer] The reviewer mentions an important distinction between doubts about the effectiveness of the interventions on one hand, and the appropriateness of their candidacy on the other hand. While some respondents reflected on the effectiveness of the interventions, for almost all respondents, the last reason was applicable.

Proposed change: The second paragraph on results on self-perceived need for care now reads: “Among those who perceived depressive symptoms as hindering, many questioned whether they would be deserving candidates to receive the intervention. For instance because they preferred to manage problems themselves or they perceived their problems as a status-quo that cannot or do not have to be changed: [...]”

8. The authors make some statements about the effectiveness of the mental health and home care nurses. Are those comparisons potentially confounded – presumably allocation was not randomised? I think the interpretation is probably correct, but it would be useful to be clear about other potential causes of those differences.

Comment: [We agree with the reviewer]
Proposed change:

- The last paragraph on stepped care preference-led intervention in the discussion section now reads: “Although it therefore seems that nurses must absolutely have sufficient background knowledge of and skills in treating older persons with depressive symptoms in order to adequately provide the interventions, these results should be interpreted with care since allocation to receive treatment from mental health care or home care nurses was not randomised.”

- Our statement in the conclusion section was put less firmly: “Lastly, results suggest that adequate provision of the programme requires that nurses have significant knowledge of and skills about treating persons with emotional distress. Although more research is needed to confirm these findings, it might imply that implementation of the programme is more suitable in (generic) mental health care settings or primary care than in home care organizations.

- The same accounts for the conclusion section in the abstract: Lastly, since the provision of the interventions seems to ask for specific skills and experiences, it might require mental health care nurses to offer the programme.

9. Similarly, there needs to be some consideration of threats to the validity of the comparisons of the different interventions (page 23), including non randomised allocation and power. I felt it would be useful to more clearly distinguish findings which benefit from the protection of randomisation, and those that did not.

Comment: [We agree with the reviewer]
Proposed change: We have added a statement to this effect to the paragraph on limitations (page 31): “A stepped wedge design is attractive for implementation research as it allows both an estimate of effect and a flexible, hybrid design, making it possible to improve implementation. This type of design has more threats to the validity of the comparisons of the different interventions than for instance cluster randomised trials, including the non-randomised allocation and power issues.”

10. In a similar vein, the strong statement that ‘offering one single preference led intervention is effective by itself and might therefore be just as effective as providing a stepped care intervention programme’ might need more justification. Do they have the data from this study to make a strong statement of that? I wondered if that needed to be presented with more caution.

Comment: [We agree with the reviewer]
Proposed change: We have rewritten the top paragraph of page 28 more cautiously. It now reads: “Finally, our data suggest that offering one single preference-led intervention is perhaps as effective as providing a whole stepped care intervention programme.”

11. Does the statement on page 28 that ‘many older patients in the target group’ were missed have a strong empirical backing, or is it based on staff perceptions? If the latter, that might be made explicit.

Comment: [We agree with the reviewer]
Proposed change: Instead of stating that “General practitioners and nurses moreover concluded that” we now state that they ‘perceived that’.

12. Our Dutch colleagues always write beautifully in English, but there were terms (‘quitted’) which need editing. I would prefer the term ‘Proactive case finding’ to Outreach but that is a personal preference.

Comment: [We agree with the reviewer]
Proposed change: ‘quitted’ was replaced by ‘quit’

Proposed change: Outreaching approach was replaced by proactive case finding
Reviewer 2: Gerdien Franx

1. Is the question posed by the authors new and well defined? The question is not new but well defined, although I do have a concern about the stepped-care part of the title, see under 2.

Comment: [We partially agree with the reviewer] We agree with the reviewer that research on stepped care models is not new. However, unique to our study is the provision of a stepped care model that offers multiple treatment options in every step (preference led care). Also, the target group of older adults is relatively new. See also point 2 for our response to concerns about the stepped care approach.

Proposed change: Therefore, we changed the first sentences of the last paragraph of the discussion section to put more emphasis on this unique aspect of our study: “Yet, these current randomised clinical trials have compared stepped care models to stratified care, and thereby missed out on participants’ needs for care. It is unknown how a preference-led model –integrated in the provision of stepped care- would impact on the clinical effectiveness of such a programme in older persons.”

2. Are the methods appropriate and well described, and are sufficient details provided to replicate the work? Discretionary Revisions. Yes, I think the description of the method is done thoroughly. One reflection concerning the title and focus of the study. The staged implementation is very appealing but in this case it incurred a total change of the protocol. The stepped care approach was (rightly) abandoned half way through the study, because of high drop-out rates. Considering this crucial change, I wonder if the authors can still say that they examined a stepped care approach in the latter clusters? Perhaps the focus of the study gradually became the implementation of a screening and patient centered (because of the explicit match with the patient’s needs) early intervention programme for the elderly?

Comment: [We partially agree with the reviewer] We agree with the reviewer that we made an important change to the study protocol in the last two clusters. However, we still consider it appropriate to speak of a stepped care approach for two reasons: 1) the study was designed as a stepped care study, and changes to the protocol were made based on hindering factors to the implementation of this stepped care approach. 2) Only 5 persons (4% of all cluster 3 and 4 respondents and only 2% of all respondents) did not follow the interventions according to the stepped care approach and are therefore expected not to be of significant influence. Also, all respondents who participated in the in-depth interviews followed interventions according to the stepped care principles.

Proposed change: None

5. Are the discussion and conclusions well balanced and adequately supported by the data? Minor Essential Revisions. Partly. Please check the first sentence of the conclusion which states that the intervention was effective compared to usual care, which was not the case since the comparison condition was no care at all.

Comment: [We agree with the reviewer] And also since we compared conditions within the same persons (each person switched from control to intervention group), we decided to leave out the words ‘usual care’

Proposed change: the words ‘compared to usual care in the first sentence of the conclusion section were deleted.

6. In the conclusion I do miss some acknowledgment of a wider range of studies in this area, which all conclude that universal prevention of depression programme, with a screening component addressing an entire population, is not effective and should not be rolled out in daily practice. This is an important finding in the Dutch context of insurers gradually making screening procedures mandatory in primary care.

Comment: [We agree with the reviewer] We agree and thank the reviewer as we have now added these last sentences to our conclusions. See also: Thombs BD, Ziegelstein RC, Roseman M, Kloda LA,

**Proposed change:** We added the following sentence to the first paragraph of the conclusion section: “Our findings add to the notion that universal prevention by depression programmes, with a screening component addressing an entire population, is not effective and should not be rolled out in daily practice. This is an important finding in the Dutch context of insurers gradually making screening procedures mandatory in primary care.”

7. In their recommendations, the authors understandably state that new approaches to reach the depressed elderly population need to be found. Can they do some suggestions for this based on the rich qualitative information they gathered?

*Comment:* [We agree with the reviewer]

**Proposed change:** The last sentence of the paragraph on routine screening in the discussion section now reads: “Future studies are necessary to explore possible ways of detecting older persons with depressive symptoms (for instance by integrating depression screening with a general health screening visit from a practice nurse that most older persons in the Netherlands receive after their 75th birthday), taking into account their self-perceived need for care.”

8. Do the title and abstract accurately convey what has been found?

*Discretionary Revisions.* See remarks under 2.

*Comment:* See our response under 2

**Proposed change:**


Yes, except for some citations which could be made less Double Dutch. Of more importance is the repetitive use of the words ‘older people’ or ‘elderly people’, underscoring a strangeness of that particular patient group. Please replace by more neutral terms as respondents or participants, as one would do in other adult research.

*Comment:* [we agree with the reviewer]

**Proposed change:**

- Quotes were checked by a translator and adjusted
- Most of the terms ‘Older people’ and ‘elderly people’ were replaced by ‘respondents’ or ‘participants’ or ‘people’, accept when referring to the target group of older persons with depressive symptoms in general (e.g. “Future studies are necessary to explore possible ways of detecting older persons with depressive symptoms [...]” since our findings are specifically related to older persons compared to younger adults.
Reviewer 3: Sarah Alderson

This study aims to look at facilitating and hindering factors in implementing a depression stepped care management programme for older adults. Overall it has some interesting results that help explain why routine depression screening does not improve outcomes and some of the difficulties in implementing stepped care approaches to treating depression. It does however have some problems in data collection or presentation, in that facilitating factors and contextualising factors do not appear to have been collected during data gathering, despite this being part of the aim of the study.

Major compulsory revisions

1. The authors mention in several places that they aimed to look for facilitating and hindering factors in implementing the Lust for Life programme, however facilitating factors are rarely mentioned in the article. The majority of the results focus on the hindering factors and problems identified. The interview guide used a framework so facilitating factors should have been identified during the interviews and focus groups. The results need to be reanalysed to show this data as well as the hindering factors. The authors do acknowledge the lack of facilitating factors identified in the limitations of the study but do not explain adequately enough why these were either not identified or reported.

Comment: [We agree with the reviewer] We agree with the reviewer that facilitating factors are underrepresented in our paper. During the interviews, the focus of most respondents was on limiting factors to the implementation of the programme and although the researchers also enquired about facilitating factors as mentioned in the interview guide, they could have made more effort to explore these facilitating factors in more detail. Also, during the analyses of the data, we focused to strongly on hindering factors to the implementation. Therefore, we reanalysed the data and identified and described more facilitating factors in this revision.

Proposed change:

- We added the following sentence to limitations to our study to explain why data on facilitating factors was underrepresented (see underlined): “Although our open questions also enquired about facilitating factors, these came up less spontaneously since stakeholders mainly focused on factors that limited the implementation.”

We added the following facilitating factors to the paper:

- The case finding procedure reached persons who had not yet expressed a need for care to their general practitioner (page 18): “On the other hand, nurses and general practitioners appreciated that the case finding procedure facilitated the inclusion of respondents who had not yet expressed a need for care to their physician.”
- The ability to work in the same patient medical file facilitated professional’s interactions (page 22): “An important factor that facilitated professional’s interactions in West-Friesland (compared to the other locations) was that nurses and physiotherapists were able to work in the same electronic medical file, which enhanced the exchange of information.”
- We put more emphasis on the most important facilitating factor to the implementation of the programme mentioned by almost all participants with depressive symptoms: the interaction with the nurses/physiotherapists (page 23): “The most important facilitating factor that emerged from interviews with almost all participants was the interaction with the nurses and/or physiotherapists. Irrespective of which interventions was chosen, participants placed high (some even most) value on their interaction with these professionals”
- Participants highly valued the possibility of receiving the nurse at home (page 25): “Irrespective of which intervention was chosen, participants highly valued the possibility to receive the nurse at home for the intake session and (most of) the interventions, instead of having to leave the house.”
• Group interactions were valued by participants and physiotherapists during the exercise programme (see page 24): “Participants and physiotherapists were very enthusiastic about the group interactions during the exercise programme.”
• We added facilitating factors about the PST-intervention (page 24): “On the other hand, nurses valued that PST could easily be learned by professionals and that it provides good opportunities to adjust to participants’ needs since participants are given the choice which problems they want to work on.”
• We added facilitating factors about the self-help course (page 25): “Despite these barriers, nurses generally valued the content of the book and the opportunity to choose from several exercises and information to adjust the course to participants’ individual situation and needs.”

2. The change in protocol to reduce the high dropout rates, where patients could choose to start a step-two intervention if desired rather than step one intervention as planned may mean the results are less applicable to ‘real life’ situations of a stepped care approach and this needs discussing as a limitation in the discussion.

Comment: [We partially agree with the reviewer]: This indeed should be mentioned as a limitation of our study. However, we feel that the impact of this decision remains limited since only 4% of respondents made use of this option. Further, we feel that this decision made our protocol more applicable to real life situations since we responded to preferences from participants themselves.

Proposed change: We added the following limitation to the discussion section: “Second, in order to prevent drop-out from the interventions and adjust the programme to participants’ preferences, we changed the protocol during the trial by providing the opportunity to choose from all interventions at the same time when required, which should be taken into consideration when interpreting our results. However, the impact of this decision remains limited since only 5 persons (4%) made use of this opportunity.”

3. Under the qualitative data collection section the authors need to make it clear who was interviewed and whether those who dropped out were included in the interviews as their views are important to this type of study.

Comment: [We agree with the reviewer] We added information on which respondents with depressive symptoms were interviewed. We refer to Figure 2 for an overview of interviewed respondents who did and did not drop-out of the interventions.

Proposed change: The second sentence of the paragraph on qualitative data collection now reads: “Individual interviews were conducted with respondents with depressive symptoms (all cluster two respondents were invited to participate in the interviews and were followed-up, irrespective of their decision to accept or refuse the intervention offer and possible drop-out of the interventions) [...]”

4. The authors also mention in this section that they gathered information from returned screening questionnaires and telephone conversations with those who did not respond or declined participation. Was consent obtained from the older adults regarding this, particularly from telephone conversations? Was this data transcribed and analysed with the rest of the interview data or treated separately? How did this data influence the results?

• Comment: [We agree with the reviewer]
Obtained consent: The Medical Ethical Review Board considered that informed consent was not required to collect data on reasons for refusing participation in the study reported on the screening questionnaire and during telephone interviews with research staff given the importance of collecting this information and the low burden it put on respondents.

• Analyses of data: The data (notes of the telephone conversations and remarks on the screening questionnaire) were coded separately from the data from the interviews and focus groups. The first data were analysed for similarities and differences with the data of the interviews and focus groups with stakeholders. Furthermore we quantified reasons for refusing our intervention offer
(provided by participants on the screening questionnaire) by calculating the number of times a certain code (reason for declining) was mentioned within the group of older persons that provided reasons for refusal. See for instance section 1.1: ‘Many of them declined participation for this reason: 69 persons (21.5% of those who provided reasons for refusal) explained not to perceive depressive symptoms.’ We agree with the reviewer that the way we analysed the reasons for declining on the screening instrument and in the telephone conversations can gain in clarity. We therefore added this point to the data analysis section.

The data from the telephone interviews and screening questionnaire generally led to similar conclusions compared to the data of the interviews and focus groups (regarding the themes covered in both data). We agree with the reviewer that this can gain in clarity. Therefore we made this more clear in the result section.

**Proposed change:**

- We added the following paragraph to the section on qualitative data analysis: “Data on reasons for declining the intervention offer provided by non-respondents on the screening questionnaire during the telephone conversation with research staff were coded separately and consequently compared to data derived from the interviews and focus group discussions.”

- We clarified the origin of the data derived from the screening questionnaire by adding the underlined words:
  - ‘Many of them declined participation for this reason: 69 persons (21.5% of those who provided reasons for refusal on the screening questionnaire) explained not to perceive depressive symptoms.
  - ‘Most not-included persons with depressive symptoms who provided reasons for their refusal on the screening questionnaire stated that they perceived no need for care (27.4%, n=88).

5. In the results, section 1.2 (factors relating to health care professionals) needs to be expanded to include more analysis of the data rather than just descriptive findings of the qualitative data. Particularly the attitudes of nurses towards screening – in what sense were they critical?

**Comment:** [We agree with the reviewer] We have included more analysis of the data to the result section

**Proposed change:**

We added the following paragraph to the section 1.2 on page 17: “Nurses felt that proactive case finding failed to include the target group of people with depressive symptoms who would benefit from the interventions for several reasons. First, persons with a depressed mood who had little insight into their emotional problems would be less inclined to fill in the screening questionnaire compared to people who are more self-reflexive: “Those people who think ‘oh, it is not that bad’ or ‘no, I don’t have mental but physical complaints’. I think that if the general practitioner does not refer them, you cannot catch these people.” (nurse, MHC2). Second, they were concerned that the procedure might have a ‘honeypot effect’ on persons with certain personality related problems who would not benefit from the interventions. Third, a screening procedure was considered to be too insensitive to the fluctuation of people’s mood and would therefore lead to the unwanted exclusion of many eligible persons.”

6. Results section 1.4 (context-related factors) again seems to be underrepresented in data given that a framework was used to design the interview guide and so should have been covered in the interviews and focus groups.

**Comment:** [We agree with the reviewer] We agree with the reviewer that, compared to other factors that influenced the implementation of the intervention programme, context-related factors seemed underrepresented (although, most levels were comprised of 1-2 factors, and on the context
level, at least 1 factors was mentioned within each theme). Given that stakeholders extensively commented on the intervention programme and most of them had little insight in context-related factors that were of influence on the implementation, we were not able to extract additional context-related information from our data. We therefore added this limitation to the discussion section.

**Proposed change:** Added to the discussion section: “Also, context-related factors were slightly underrepresented since respondents reported extensively on the intervention programme (and less focused on context-related factors) and many stakeholders had little insight into context-related factors that influenced the implementation of the programme.”

7. Results section 2.2 (health care professionals’ backgrounds) mentions that the lust for life work may have been perceived as burdening and they doubted the suitability of this but no evidence from the data is reported.

**Comment:** [We agree with the reviewer] We added a quote from the data as evidence for this point.

**Proposed change:** quote added: Page 23: “We see now how much time it costs [to conduct the programme] but most of all how much psychological burden [it puts on us; R3]. [...] It is quite a different role we have [in this program] [...] our [usual] work is more goal-oriented to put it like that. Then there are problems, and you can help to solve those problems [R2].” “I am really struggling [...] It takes so much effort and I keep on trying but it takes so much energy from me, and I just do not see any results. I do not see any progress at all [with my clients; R1].”

Minor essential revisions

8. The authors describe the PHQ9 as a depression screening tool, however it is designed as a depression severity assessment tool rather than for screening and this needs to be amended in the manuscript.

**Comment:** [We partially agree with the reviewer.] We understand from the original papers on the PHQ that this instrument is designed as a screening tool to establish a depressive disorder as well as severity measurement to monitor outcome. See for instance Kroenke et al (2001): PHQ-9: validity of a brief depression severity measure. J Gen Intern Med, 16: 606-613: “The PHQ-9 has the potential of being a dual-purpose instrument that, with the same 9 items, can establish depressive disorder diagnoses as well as grade depressive symptom severity”. Or Spitzer RL, Kroenke K, Williams JB: Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999, 282:1737-1744.

**Proposed change:** We added a sentence in methods that the PHQ 9 was designed as a dual-purpose instrument to diagnose a depressive disorders as well as monitor depression severity (see page 6/7).