Author’s response to reviews

Title: Treatment received and treatment adequacy of depressive disorders among young adults in Finland

Authors:

Teija Kasteenpohja (teija.kasteenpohja@thl.fi)
Mauri Marttunen (mauri.marttunen@thl.fi)
Terhi Aalto-Setälä (terhi.alto-setala@hus.fi)
Jonna Perälä (jonna.perala@thl.fi)
Samuli I Saarni (samuli.saarni@thl.fi)
Jaana Suvisaari (jaana.suvisaari@thl.fi)

Version: 3 Date: 12 February 2015

Author’s response to reviews: see over
To:
Dr. Vicent Balanza, Editor in Chief, BMCPsychiatry

Reference: manuscript 1368659163134028, “Treatment received and treatment adequacy of depressive disorders among young adults in Finland”

February 12, 2015

Dear Editor,

Thank you for your kind email and for the valuable comments by the reviewer. Please find enclosed our manuscript “Treatment received and treatment adequacy of depressive disorders among young adults in Finland” (1368659163134028), which we resubmit for consideration for publication as an original article in the BMCPsychiatry. We have taken into account the comments by reviewers Ronald Kessler, Ana Fernandez, Arnaud Duhoux and Frank Jacobi and revised the paper accordingly. We thank all reviewers for their excellent comments. In our response to reviewers, we reply to the comments of the reviewers and explain the revisions we have made. Additions in the manuscript have been marked in boldface. If needed, we will be happy to modify the manuscript further.

Thank you for your attention and consideration. We look forward to hearing from you.

Yours sincerely,
Teija Kasteenpohja, MD
Department of Health, Mental Health Unit, National Institute for Health and Welfare, Helsinki, Finland
Postal address: P.O. Box 30, FIN-00271 Helsinki, Finland
Tel.: +358 29 524 6000
fax.: +358 29 524 7155
E-mail: teija.kasteenpohja@thl.fi
www.thl.fi
RESPONSE TO THE REVIEWERS:

Title: Treatment received and treatment adequacy of depressive disorders among young adults in Finland

Reviewer: Ronald Kessler

1. I'm not sure from the description that the sample was weighted properly to reconstruct the population. Was that done? Where is it described in the paper? I probably missed it, but I want to make sure.
   1. Response: We did not use the weights, because the analysis was restricted to people with a lifetime depressive disorders. We thought that adjusting for the sampling design (with 80 primary sampling units) would not be meaningful when the sample size of people with depressive disorder is 142, and this was the reason for not using SUDAAN. However, poststratification weights that take nonresponse into account could have a larger effect on the results. We checked whether the results would be different if the weights are used, using SUDAAN and restricting the analysis to people with depressive disorders with the subpopn command. As an example, the proportion of persons having received minimally adequate treatment in the most recent episode when the weights were used would be 37.8% and for the most intensively treated episode 43.1%, when the unweighted results were 40.9% and 45.1%. Therefore, the results were quite similar when the weights were used. One person (who had received adequate treatment) had not been assigned a weight (he had not participated in any phase of the survey due to failure to locate him but we had his hospital records) and was dropped from the analysis in SUDAAN, and this slightly affected the results. Screen-negatives who had received a lifetime diagnosis of depressive disorder (4 cases) did not have any treatment contacts (as a treatment contact was a screening criterion), and these were the other persons who influenced the results (they had larger post-stratification weights because only a random sample of screen-negatives were invited).

If it is considered more appropriate, we can recalculate the results using survey weights, as was done in the article reporting the lifetime prevalences of mental disorders (Suvisaari J, Aalto-Setälä T, Tuulio-Henriksson A, Härkänen T, Saarni SI, Perälä J, Schreck M, Castaneda A, Hintikka J, Kestilä L, Lähteenmäki S, Latvala A, Koskinen S, Marttunen M, Aro H, Lönnqvist J. Mental disorders in young adulthood. Psychol Med. 2009 Feb;39(2):287-99.). We have not previously done this in studies focusing on specific subgroups, for three reasons: (1) adjustment for the original two-stage stratified cluster sampling design, although technically possible, would not be meaningful because there are too few participants per area, (2) weighted numbers are more difficult for the reader to interpret, and the combination of using original numbers but weighted prevalences is confusing, and (3) at least the older versions of SUDAAN did not allow Fisher’s exact test, which sometimes is needed in these subgroup analyses.

2. The demo predictors: Were these always temporally prior to the treatment? If “most recent treatment was 5 years ago and the respondent is currently at university, the respondent would presumably have been in secondary school at the time of the most recent treatment.
   2. Response: The sociodemographic “predictors” were not always temporally prior to the treatment. Whereas basic education had almost always been achieved prior to the most recent episode, current employment was employment during the interview. Therefore, we replaced the word “predictor” with the expression “factors associated with”. Because of this problem, we took away the discussion related to current employment.
Reviewer: Ana Fernandez

- I encourage authors to justify why it is important to study the proportion of minimally adequate treatment. What are the outcomes associated to this? Do people who receive a minimally adequate treatment recover faster? Do they have better outcomes? What the evidence says? The paper by Hepner et al is a good starting point: Hepner KA, Rowe M, Rost K, Hickey SC, Sherbourne CD, Ford DE, et al. The effect of adherence to practice guidelines on depression outcomes. Ann Intern Med 2007 Sep;147(5):320-9.

Response: The reviewer is correct, and we have now modified the Introduction accordingly.

Method: Overall, the method is clear. I just need some more clarification with regard:

- Figure 1: What happened with the 334 (1316-982) people that were not invited to participate? I have not been able to find this information in the body of the paper.

Response: We invited all screen-positive people but only a random sample of screen-negative people. Weighting methods were used in the original prevalence publication of this study to adjust for the fact that only a random sample of screen-negatives had been invited to participate (Suvisaari J, Aalto-Setälä T, Tuulio-Henriksson A, Härkänen T, Saarni SI, Perälä J, Schreck M, Castaneda A, Hintikka J, Kestilä L, Lähteenmäki S, Latvala A, Koskinen S, Marttunen M, Aro H, Lönnqvist J. Mental disorders in young adulthood. Psychol Med. 2009 Feb;39(2):287-99.). In practice, our screening method for possible mental disorders was so broad that 58% were screen-positive, and the lifetime prevalence of mental disorders was very low in the screen-negatives. We now describe the procedure in more detail in Supplementary methods.

- Table 1: I recommend the authors to include the references that justify the use of these cut-offs (in addition to the reference to the main paper; as not all the readers would have access to the source)

Response: We have now added a supplementary method description that explains these issues in detail. As a further clarification: only a random sample of screen negatives was asked to participate, but our screening for possible lifetime mental disorders was very inclusive, capturing 58% of the participants.

- Definitions of minimally adequate treatment: please include in the definitions (lines 132-136) the time frame (i.e. within one year).

Response: The time frame (within 12 months) is included in the definitions.

- I suggest the authors to move the limitations associated to the definition of adequate treatment to the discussion.

Response: The limitations associated with the definition of a psychotherapeutic session were moved to the section of strengths and limitations.

Results: are well explained and easy to follow. However, I would avoid to use the word “determinants” as it implies some type of causality. I think it is better to change this and use “factors associated to treatment and drop-out”.

Response: The word “determinants” was replaced by the suggested expression.

On the other hand, I was not able to find in the table the 76.1% who have use any type of contact. I think it would be worth to include this information.

Response: We aimed to study especially minimally adequate treatment and its components. Because this already meant a lot of different analyses, we did not analyze separately the features related to treatment seeking.
However, we wanted to report the proportion of participants who had sought treatment. Because our tables are already very large, this figure appears only in the text.

I would like to see how “age at the onset of depression” (treated as a continuous variable) was associated to the care received. The authors included this as a continuous variable in the logistic models, but this is the first time we know about this variable. I recommend the authors to include this variable in table 4.

Response: We included age at the onset of depressive episode in the analysis because younger people might have better access to e.g. psychotherapy through student health services. There were no statistically significant differences in age at onset of depressive disorders in any of the treatment-related variables, and we did not add these results on the already crowded table. Nevertheless, for example in minimally adequate psychotherapy the mean age at onset had been 24.1 years for those who received it vs. 25.4 for those who had not, and therefore we decided to keep this variable, which was chosen on theoretical grounds, in the model.

Related to this, I also encourage them to explain why they have selected these variables to be included in the final models. Is this selection theory-driven or data-driven (I mean, you included the variables that were statistically significant associated)? This should be explained.

Response: The selection of variables was mostly theory driven, but we left out some theoretically interesting variables, e.g. comorbid eating disorders, because they did not associate with any treatment-related variables.

Discussion: Discussion is also good and well written. However, I feel the authors are given to much importance to bivariate associations.

Response: We mostly discuss results from multivariate analyses. When bivariate associations are discussed, we always mention it separately. Because of the relatively small sample size, there is a risk of type 2 error in multivariate analysis, and therefore we felt it was important to mention some of the bivariate results as well.

I feel also that they should acknowledge that, from an equity point of view, it is good that none of the sociodemographic factors were associated to receive an minimally adequate treatment.

Response: The reviewer is correct, and we have added this point of view into the Discussion.

I am not sure authors are right when they conclude that “services for the treatment of depression are not functioning efficiently because people with comorbidities do not have a higher probability of receiving adequate care”. Authors are assuming that comorbidity is a proxy of severity when it may not be always the case. I recommend the authors to develop this conclusion and to justify it with additional evidence.

Finally, limitations related to the definition of "minimally adequate treatment" used should be acknowledged. The reviewer is correct in that minimally adequate treatment only signifies that the minimum criterion for treatment adequacy is reached, it does not mean optimal treatment. Optimal treatment would mean that every effort according to depression treatment algorithms (Katzman MA, Anand L, Furtado M, Chokka P. Food for thought: understanding the value, variety and usage of management algorithms for major depressive disorder. Psychiatry Res. 2014 Dec;220S1:S3-S14) was made in order to achieve full symptomatic remission. Also we were not able to examine actual psychotherapy, since the data we had on psychosocial treatments was not detailed enough. We now mention these issues in the Discussion.
Reviewer: Arnaud Duhoux

Abstract:
1. the abstract should be revised according to revisions suggested in the next sections
1. Response: We have modified the abstract according to the comments made by the reviewers.

Methods:
2. the construction of the sample is not easy to understand, and more details should be given to make it clearer.
2. Response: We now provide this information in Supplementary methods.

Line 84 “Information from the Finnish National Hospital Discharge Register (NHDR) was used to identify all persons who had received hospital treatment due to any mental disorder” Do the authors mean: all persons from the original survey? Because later, they wrote: “Persons selected via NHDR who had not returned the MEAF questionnaire were contacted through the person responsible for the treatment”. This is not clear to me.
2. Line 84. Response: This is correct: if a person had not returned the questionnaire but we had the register information on treatment, we attempted contact through the treating personnel. This applied to all people belonging to the original survey except for those who explicitly had denied any further contacts (N=26) or who had died (N=5).

Line 87: “In addition, a random subsample of Health 2000 young adults was invited to the interview regardless of their answers to the screening questionnaire”. Please explain why it was necessary to do so.
2. Line 87. Response: When designing the study, this was thought as the best method of achieving a representative random sample of people who are screen-negative. This random sample obviously included both screen-negative and screen-positive persons in the same proportion as the total sample.

Figure 1: What were the selection criteria when the sample was reduced from 1316 to 982?
2. Figure 1. Response: All screen positives but only a random sample of screen-negatives was invited to the interview.

3. Line 96: was the Mental health assessment done on face-to-face or on the phone?
3. Response: The interview was a face-to-face interview, and also neuropsychological testing was included in the assessment.

4. Line 110: “all case records from hospital and outpatient treatment were obtained”. Could the authors provide more details on the Finnish data system, this seems impossible to do in other countries such as Canada where data are not centralized.
4. Response: The actual case record data were not centralized in Finland at the time this study was done, although currently such system is being developed in Finland. In practice, collecting case records was a huge effort. The participants had told us where they had been treated, but we had to contact all places that had provided treatment case-by-case, finding the right person to send us the case records. Hospital case records were easier, because they were ordered based on the hospital discharge register data, and hospitals have central archives which were easy to locate.

5. Line 129 and table 2 and 4: Specific is misspelled.
5. Response: Misspellings have been corrected.

6. Line 128: “Therefore, the present paper focuses on treatment received during the most recent depressive episode.” It is not clear for me how the episodes were defined. This is important to understand, because the adequacy of treatment is dependant on the time the studied episode started and lasted. Please give more details.

6. Response: We gathered information about depressive episodes from interviews (including SCID) and case records. Based on this information we could estimate the duration of depressive episodes and how they were treated. Episodes were considered separate if they were separated by at least 6 months when the participant’s ability to function was normal and there were at most only minor residual symptoms. The treatment periods were separate if there was at least 1 year time between them. Based on all this information we decided which depressive episode was the last one and if there was another episode which was treated more intensively. Because there may have been a large variation between durations of depressive episodes as well as treatment periods we decided a timeframe of one year to count the visits. If someone had been treated more than one year we took into account the most intensively treated year from this treatment period. We decided to use the one year time frame because we wanted it to be comparable to previous studies of this topic which have studied treatment in the previous 12 months.

7. Line 136: “We also defined at least four days of hospitalizations for depression as adequate care.” Could the authors explain why they choose the limit of 4 days? Is this limit derived from a practice guideline or a previous study?

7. Response: Hospitalizations have not been used as an adequacy criterion in previous surveys, possibly because detailed information on them has not been available. Because we had accurate information about hospitalizations based on the national hospital discharge register and corresponding case records, we wanted to use it. We chose the four days’ time limit because we did not want to take account for example short visits in the emergency department. Based on the expression we got from the case records and the fact that evaluation period in Finland is four days, we decided that this time is enough to assess a diagnosis as well as plan and start a treatment. This information is added to the supplement.

8. Line 132: “Based on this information, receiving antidepressant pharmacotherapy for at least two months with at least four visits with any type of physician or at least eight sessions of psychotherapy were regarded as minimally adequate treatment.” The authors should specify the period of time observed for the reception of the minimum number of psychotherapy sessions (over 6 months, 12 months?). In the literature, there is a great variation.

8. Response: A period of 12 months time has been specified to the manuscript.

9. Line 137: “Visits with physicians were assessed from case records and interviews and divided into three categories: none, at least one visit or at least four visits a year.” Why did the authors choose to measure the number of visit on a one-year period if they can determine the beginning of the episode? (Same question for psychotherapy). This can represent a problem, because the sample is constituted of people with lifetime disorder, but the number of visit is measured on a one-year period. Is it the year after the beginning of the episode? Please give more details.

9. Response: As mentioned before there was a large variation between duration of depressive episodes and treatment periods. In long treatments, there typically was a period of intensive treatment followed by less frequent visits when the situation had improved. Therefore, we chose a timeframe of 12 months often used in studies about this topic, and counted the visits from the year the person had received the most intensive treatment.
Also, were only the visits related to mental health considered or every visit? Please make this point clear, as many other study considered visits for mental health reasons only. This can have a large impact on the results. In a previous study, we showed that people with depression had visits in primary care but were not detected and didn’t consult at all for mental health reasons (even if they consulted in primary care for other reasons).


9. Response: We took account only the visits related to mental health and when possible even more specifically visits related to depressive symptoms of participants. So if depressive symptoms were not detected in the primary health care center when consulting for other reasons we did not count those visits. The information about visits has been specified in the manuscript.

10: Line 141: “A psychotherapeutic session was defined as a visit with a psychiatrist, psychologist, psychotherapist or other professional in a psychiatric clinic.” Is it possible in Finland to receive psychotherapy in other settings? If so, why did the authors consider only psychiatric clinic?

10. We considered the visits of psychiatrists, psychologists and psychotherapists in all settings and other professionals only in a psychiatric clinic. This information has been specified in the manuscript.

We accepted any professional in a psychiatric clinic because we thought that for example a social worker in a psychiatric clinic is usually trained to give a psychosocial support to persons with mental disorders. Furthermore, as mentioned in the limitations-part, we did not study actual psychotherapy, but more broadly psychosocial support.

11: Line 179:“Logistic regression analysis was used to identify variables (i.e. gender, age, age at the onset of depression, basic education, MDD, suicidality, anxiety disorder, and substance use disorder) that were independently associated with the use of mental health services.” The regression analyses were used with many dependant variables, not only use of mental health services. Please specify.

11. We have specified this as: “the use of mental health services and the type and adequacy of treatment.”

12. The authors performed many bivariate analyses. However, they didn’t use a Bonferroni correction. This should be added in the method section, or at least discussed in the limits.

12. We have now added in the limitations that “Furthermore, correction for multiple testing was not done.”

We present several descriptive results concerning different aspects of the type and quality of care people with depressive episodes had received, and we consider each of these aspects important. The study was descriptive, not hypothesis driven. Bonferroni adjustment suits well for hypothesis driven studies and for studies in which there is great risk of false positives, as there is e.g. in genetic studies. However, Bonferroni adjustment is very conservative and has been criticized (e.g. Perneger TV. What's wrong with Bonferroni adjustments. BMJ. 1998;316(7139):1236-8), because it makes it tempting to report your work only partially. Our study sample is population-based and the sampling frame was carefully designed to make it representative of the Finnish general population (we added a description of the sampling in the Supplementary methods). Therefore, we believe that the results represent well the average care that people in Finland receive for depressive disorders, and reporting all the different aspects of care are important.

Results:

13. Line 219:“Participants who discontinued their visits despite appropriate treatment plan were more likely to have had medication, visited a physician, and attended psychotherapy than those who continued their visits. However, differences
were no longer statistically significant when medication for at least two months and psychotherapy at least eight times a year were taken into consideration (Table 4)."

The dropout variable is not clearly defined. I don’t really understand how you can be considered dropped out, and received adequate care. For example, if you received minimally adequate psychotherapy (at least 8 sessions), it may be possible that your treatment was effective and is over. So, the patient is getting better and stopped his treatment, in this case it is not a dropout. There is something with the definition of dropout that I don’t understand.

In Table 4, it is quite strange to see dropout as an independent variable. It is difficult to interpret the results. For example, dropout is associated with pharmacotherapy but you need to start a treatment to dropout. I suggest removing dropout from the independent variables list.

Also, I don’t think we can say that the association disappears when we take into consideration medication and psychotherapy for a least 8 sessions because we are not in a multivariate analysis. So, we don’t take into consideration any variable. You can just say that there is no association. This absence of association is certainly related with the definition of dropout. People who received medication +4 visits or 8 sessions of psychotherapy didn’t dropout, they just received an effective treatment and are getting better. This is a good argument to revise the definition of dropout. There is a difference between a dropout at the beginning of the treatment (which can be due to many different reasons, but the patient is still symptomatic) and a dropout after a longer time when the treatment can be effective and the patient getting better.

More details on the definition of dropout should be given. The last paragraph of the discussion on dropout should be revised.

13. Response: We excluded a dropout as an independent variable because of weaknesses you mentioned. Dropout is also defined more detailed in the supplement now. Dropout was defined when we could see in the case records that there had been an adequate treatment plan (for example related to medication use or psychotherapy) but a patient had discontinued the visits by his own decision.

People with a history of suicide attempts and people with low education dropped out from treatment more often than others. This explains the contradictory finding that dropout was associated with receiving pharmacotherapy, having at least 1 session of psychotherapy and having at least 1 as well as 4 visits with a physician: there had been a plan of intensive treatment, but although these people had been in treatment long enough to e.g. visit a physician for 4 times, they had then discontinued their visits. This reflects the clinical challenges of engaging young people with self-harm or low education to long-term mental health treatment.

14. Line 227: “Gender, age, age at the onset of depression (continuous), basic education, history of suicidal attempts, diagnosis of MDD, comorbid anxiety disorder and substance use disorder were entered simultaneously into a logistic regression model to explore the factors affecting treatment.” This sentence should be moved to the Analysis section.

14. Response: The sentence was moved to the section of “Statistical analysis”.

Discussion:

15. Line 244: “In our study of young adults with depressive disorders, 76.1% had had some kind of contact with the health care system during the most recent depressive episode”. It is not clear if the counted visits were during the last year, the last episode or the year of the last episode, and if those visits were for mental health reasons or not. Please make it clear in the methods section and revise if necessary.

15. Response: As mentioned before, we counted only the visits because of depressive symptoms, and this information has been added to the manuscript. A contact had been taken during the most recent depressive episode meaning the last depressive episode the participant had suffered.
16. In the first paragraph of the discussion, the author report results of only 3 studies with lower adequate care. However, in a systematic literature review we performed, we found a larger range (14 – 56 %) in 30 studies which included measure of minimally adequate treatment depending on the population studied and the definition of the adequacy of treatment (REF: DUHOUX, A., FOURNIER, L., MENEAR, M. Quality indicators for depression treatment in primary care: a systematic literature review. Current Psychiatry Reviews. 2011 May; 7(2):104-137). They should revise this paragraph considering a broader range of the literature on the subject.

The second paragraph of the discussion should also be revised in view of the Smolders et al. (2009) study which also used multiple data sources (case records and patient questionnaire). They found a similar result in the % of adequate care. (Smolders, M, Laurant M, Verhaak P, et al. Adherence to evidence-based guidelines for depression and anxiety disorders is associated with recording of the diagnosis. Gen Hosp Psychiatry 2009; 31(5):460-9.)

16. Response: We thank the reviewer for this important comment, and we have added all references that are based on studies with a similar setting, i.e. a population-based survey. We did not add more US references, however, because the references by Wang et al. cover the most recent survey data on the issue. We did not add references from studies conducted in primary care patients, because the setting is different. Previous Finnish studies have observed that the prevalence of depressive disorders in primary care patients is high (Vuorilehto M, Melartin T, Isometsä E. Depressive disorders in primary care: recurrent, chronic, and co-morbid. Psychol Med. 2005 May;35(5):673-82).

17. Line 261:“None of the factors chosen were related to the overall adequacy of treatment in a multivariate analysis. This suggests that services for the treatment of depression are not functioning efficiently: people with more severe or comorbid symptoms did not receive adequate treatment more often than others. However, in a bivariate analysis, the duration of depressive episode was related to treatment such that participants whose episode had lasted over a year received more often any care and guideline-concordant care in regard to all the various aspects of treatment as well as minimally adequate treatment.” The authors’ explanation is a bit too conclusive. The absence of significance can be explained by many reasons, including statistical ones. If the association is not significant in multivariate analysis, it may be explained by the low power and the small size of the sample (type 2 error). The bivariate association of duration of episode with guideline concordant care can be diluted in multivariate analysis due to lack of power.

17. Response: We agree with the reviewer in that type 2 error is possible, and we have mentioned this in the limitations.

Also the authors wrote that “people with more severe or comorbid symptoms did not receive adequate treatment more often than others.” That could be considered as a good point: even people with mild to moderate depression received adequate care.

17. Response: This is true. Our comment was related to previous survey findings that people with severe symptoms in other countries usually have received minimally adequate treatment more often than people with moderate or mild symptoms (Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, Bruffaerts R, de Girolamo G, de Graaf R, Gureje O, Haro JM, Karam EG, Kessler RC, Kovess V, Lane MC, Lee S, Levinson D, Ono Y, Petukhova M, Posada-Villa J, Seedat S, Wells JE. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. Lancet. 2007 Sep 8;370(9590):841-50). However, we could not define severity in the same way as was done in the WMH surveys, therefore we did not compare the results in detail.
18. Line 272: "This may also be one explanation for a higher estimate of treatment adequacy in our study compared to other previously mentioned studies: we investigated the most recent lifetime depressive episode whereas many other studies have looked at depressive episodes within the last 12 months. It is likely that in other studies, all participants with depressive disorder had not yet sought help though they would do so in the future." In the results section, it would be helpful to the readers if the authors could give an idea of the mean time between the last episode and the survey, and the number of participants still in their most recent episode.

18. Response: A median time between a beginning of the last depressive episode and the survey was 2.3 years, ranging from 0 to 23.6 years. 10.6% (15) of participants were still (during the last month before the interview) in their most recent episode. This information was added into a results section of a manuscript.

19. Line 278: "So it seems that physicians tend to meet the most complicated cases more often than others, as an association between comorbid substance use disorder and physician visits was also found." This sentence is the opposite of the one on lines 262-263. Those opposite conclusions are explained by the analysis considered: the conclusions are not the same if you consider bivariate or multivariate analysis. The authors should be clearer on this point (which result support which discussion point) and discuss in the limits on the divergent results (or interpretation) they found.

19. Response: We wanted to discuss the main finding, e.g. receiving minimally adequate treatment, first, and for it we did not find significant predictors. However, receiving minimally adequate treatment also depends on patient’s engagement to the treatment, not just on the health care system. The finding that physicians met patients with comorbid disorders more often suggests that referral to treatment works better, which we now mention in the manuscript.

20. Line 292: "This may indicate that psychotherapy has a beneficial effect on an individual’s psychosocial functioning ability, capacity for studying, working, and forming relationships." An alternative explanation is that being a student gives a better access to psychotherapy via the university health services. Did the authors explore this idea? Are the universities in Finland providing health services?

20. Response: In Finland, universities and higher education institutions provide health services of their own. In our sample, 15 participants (10.6% of the sample) had a contact via a health care system for students during their most recent depressive episode. However, our variable is a current employment meaning employment during the interview. So it is more like a consequence of treatment, which could have taken place even for several years ago. This means that someone who was a student during the interview could have received treatment as a school child. On the other hand, someone who was for example employed during the interview could have been treated previously in a setting for students. Therefore on the basis of this study, we can’t draw a conclusion on a better access to psychotherapy via the student health services.

21. Line 331: "Another significant limitation was the small size of study sample, which led to low statistical power and wide confidence intervals, particularly in the logistic regression analysis." The limits in the size of the sample are well explained. However, it is difficult to understand what is the sample representative of. The authors should explain in this section what is the representativeness of their sample. Moreover, they included a random sample of young participants with no screened symptoms. Were there any of those no symptomatic participant classified as depressive? What is the effect of the sampling strategy on the representativeness of the final sample? That discussion will help to generalize the results. For example, on the line 340, the authors wrote “A lack of adequate treatment of depressive disorders is an ongoing problem, although our results on treatments among young adults are better than in most previous studies and
encouraging in this respect.” Without an idea of the generalizability of the results, it seems difficult to draw such conclusion. Can we infer the results to all the Finland Youths?

21. Response: We have now added information on the sampling strategy to Supplementary methods. Briefly, the sample is representative of Finnish young adults living in mainland Finland. Åland was excluded from the sampling frame because the study would have been too expensive to conduct there, but the population of Åland (28900 people) represents only 0.5% of the Finnish population. We have previously reported a detailed analysis of attrition (Suvisaari et al. Psychological Medicine 2009;39:287-99). The most significant predictor of attrition was a hospital treatment for mental disorder, but we had case records for these persons and therefore were able to take this issue into account.

Another limit of the study is the time of data collection, close to 10 years ago. The quality of care could have change in a decade with improvement in treatments and dissemination of practice guidelines. The authors should discuss this point.

21. Response: The Mental Health in Early Adulthood in Finland (MEAF) study was done in 2003-2005 i.e. about ten years ago. After the years of the survey there have been changes in the health care system as well as medication in Finland, which means that figures of minimally adequate treatment may be different today. Probably, the situation has improved: based on official statistics, the use of antidepressants and the number of psychiatric outpatient visits has increased. These facts have been added to the limitations-part of the manuscript.

Conclusion:

22. Line 342: “It is alarming that dropout is related to individuals with less education, suicidality and substance use disorder, who are otherwise also persons at greatest risk of complications and social exclusion. These groups present a challenge to future health care and more efforts are needed to outreach and motivate them to receive effective treatment.” I am not comfortable with a conclusion supported by associations significant in bivariate analysis only. If the association disappears in multivariate analysis, this association maybe weak or explained by other factors. I suggest that the author write a conclusion based only on the strongest associations they found, the one significant in multivariate analysis.

22. Response: Low education and a history of suicide attempts were significant predictors of treatment dropout in multivariate analysis. We have modified the conclusion accordingly.

23. Titles of tables and figures should be more explicit with details such as sample size, source of data, country, year of data collection.

23. Response: The instructions for authors limit the number of words in the titles for tables and figures to 15. Therefore, we could not add this information to the titles.
Reviewer: Frank Jacobi

Abstract
- line 28 "...factors associated with treatment adequacy..."
Response: The sentence has been corrected.
- Line 32: please add that case records come from the same sample
Response: This information was added to the abstract.
- Line 33ff: I suggest to name or at least summarize the adequacy criteria, as well as for “dropout”
Response: The definitions of minimally adequate treatment and treatment dropout were added to the abstract.

Introduction:
- Line 43: What does “highest proportion of total burden” mean?
Response: When studying the burden (by rates of disability-adjusted life years, DALY) attributable to mental and substance use disorders as a proportion of all disease burden, globally and by region, depressive disorders had the highest proportion of total burden across all regions. This definition was taken from the Whiteford et al. article (Whiteford HA, et al. Lancet 2013, 382(9904):1575-1586), but it is true that the sentence we had was very complex (because the original article had used different methods to assess burden, and different mental disorders had been analyzed separately only in a subset of the analyses). Therefore, we decided to change the reference to the more recent Global Burden of Disease 2010 article which had specifically focused on burden related to depressive disorders. Based on this publication, we now state that “Depressive disorders are a leading cause of burden of disease worldwide. (Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, Vos T, Whiteford HA. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med. 2013 Nov;10(11):e1001547.)

- Line 59: How was “treatment need” defined?
Response: In the article by Haarasilta et al, treatment sought, perceived need for psychiatric help and at least moderate psychosocial impairment were considered as indicators of treatment need. We did not add this information to the manuscript, however, because we wanted to keep the Introduction concise. (Haarasilta L, Marttunen M, Kaprio J, Aro H: Major depressive episode and health care use among adolescents and young adults. Soc Psychiatry Psychiatr Epidemiol 2003, 38(7):366-372.)

- Line 62: How was “minimally adequate treatment” defined?
Response: Eisenberg and Chung defined minimally adequate treatment as antidepressant use for at least 2 months and discussions about medication with a provider at least 3 times in the previous year or at least 7 visits of psychotherapy. As in our study, their criteria was based on the criteria in the study of Wang, where 4 provider visits in addition to medication for 2 months or eight psychotherapy sessions were needed for minimally adequate treatment. However, they had to use a little more liberal threshold because of the categories that had been used in their answer choices (for example, for psychotherapy visits the answer choices had been 1–3, 4–6, 7–9 and 10+).

- Line 68: “…to define comorbid psychiatric disorders…” -> “…to identify comorbid mental disorders…”
Response: The expression has been corrected.

- Treatment rates etc. highly depend on the health care systems and structures of a country or region. It would be helpful if you added a few sentences (or perhaps references to European comparisons) on the Finnish mental health care situation
Response: In Finland, mental-health services are mainly provided by local authorities (eg, municipalities), as part of public-health and social-care services, and public mental health outpatient services are free of charge for users. Young adults may also use student health services (http://www.yths.fi/en), or occupational health services that often are run by private sector. Psychotherapy is often given in private sector, but our social insurance institution provides reimbursement for it (http://www.kela.fi/web/en/rehabilitative-psychotherapy). Third sector also provides low-threshold services for young people.
According to the most recent OECD statistics, the availability of mental health professionals in Finland is better than the OECD average. For example, the number of psychiatrists per 100000 is 20 (average 16), and the corresponding numbers for psychologists are 58 (OECD average 26) and for mental health nurses 50 (OECD average 50). Psychiatric care bed number is close to the OECD average (71 in Finland, OECD average 68). There are models for collaboration between psychiatric specialists and primary care, but these are not as well developed as in some other OECD countries. (OECD (2014), Making Mental Health Count: The Social and Economic Costs of Neglecting Mental Health Care,OECD Health Policy Studies, OECD Publishing. doi: 10.1787/9789264208445-en)
Overall, it is difficult to compare the Finnish treatment system with other countries because of lack of relevant and valid comparable data on mental health systems (Wahlbeck K. European comparisons between mental health services. Epidemiol Psychiatr Sci. 2011 Mar;20(1):15-8). We have now added a sentence about this in the Discussion.

Methods:
In general please point out more specific from the beginning what information was used for which criterion. Especially try to clarify the use of “case records” in this study (because in other countries it might be quite unusual to be in a situation linking complete health records to epidemiological studies). E.g., how sure can you be that no positive case record information is “true” non utilization? What was done when self-report was discordant to case records etc.?
Response: To clarify the use of case records:
Information on outpatient health care contacts was based on the interview, where we asked about treatment received for mental health problems. We located the case records based on this information to validate the information given in the interview. All information on outpatient visits was collected this way – the participants gave information on where they had been treated and gave permission to access the case records. If the participant had reported a treatment episode but we were unable to locate the case records, we used the information given in the interview. This was sometimes the case in treatment provided in the private sector.
When we obtained the case records, we checked their concordance to what the participant had reported and in case of discrepancies relied on the case records, where e.g. the visits to a physician were all recorded, as was treatment dropout (scheduled visits which the person did not attend). Obviously, people tend to have difficulties
in remembering the details of the treatment they had received, and the case records were a valuable source of additional information. Case records in Finnish mental health care are quite detailed, including information on symptoms and treatment plan, and instances when scheduled visits are not used are also reported and also the reasons for ending the treatment episode (planned vs. patient’s dropping out from the treatment).

For hospital care, we were also able to use the information from the Finnish Hospital Discharge Register and were able to access all case records this way. However, most of the treatment episodes were outpatient treatments.

- Were data used from the former Health2000 study?
  Response: We did not use data on mental health from the Health 2000 baseline study. All data on mental health and socioeconomic variables is from the MEAF questionnaire and interview.

- Fig 1: In my view, one stage/differentiation is missing between “MEAF questionnaire returned” and “Invited to participate”: please insert screen positive vs. screen negatives
  Response: In response to this question and questions raised by the other reviewer, we have written supplementary methods text which includes also this information as follows:

Based on the MEAF questionnaire screen, 768 participants were screen-positive in at least one of the screening questionnaires/questions, and they were all invited in the interview. Of screen-negative participants, 161 (29.4%) out of 548 were invited in the MEAF interview, and 88 (54.6%) participated. A detailed analysis of attrition is presented in the supplementary material of Suvisaari et al. (2009).

- Fig 1: N=432 refusals could be commented a little bit more in the manuscript
  Response: We have previously reported the results concerning attrition in detail (Suvisaari et al. Psychol Med. 2009 Feb;39(2):287-99, supplementary tables and text.). In brief, attrition was higher in men, in people with low baseline education, and in people with a history of hospital treatment for mental disorders. However, none of the scores in any of the screens we used for the mental health interview differed between interview participants and nonparticipants. In fact, self-reported mental disorder diagnosed by a physician was more common in screen-positive participants (22.8%) than in screen-positive non-participants (17.3%), although the difference was not statistically significant. Since most of the participants with depressive disorders had not received hospital treatment, we assume that attrition did not have a major effect on the results.

- Fig 1: In my view, there should be also information about the case records in the flow chart; but if this should be too complicated perhaps create an own one with participants vs. non-participants with regard to having no case report (e.g. when no consent was given), incomplete, and complete case records.
  Response: We have added this information in the manuscript.

- Table 1: right column: please add total number of screened persons and/or add a % column
  Total numbers of screened persons and percentages have been added to the table 1.

- From the start I felt a bit confused about the operationalization of constructs such as “adequate treatment” and corresponding data sources and suggest to add (at least as supplement) another table where it is stated which information from interview and/or “case records” were used determining the six characteristics in columns of tables 3ff
Response: We used all available information, i.e. both the interview and case records, if available, to determine minimally adequate treatment and its components. We counted the visits and evaluated the duration of medication based on the information in the case records, if they were available. If case records were missing or incomplete (for example in the case of treatment in the private sector) we could often estimate the visits and medication use based on information given in the interview. For example a participant could have told that he had visited a private psychotherapist once a week for a year or used an antidepressant for about half a year. This has been added to the supplement where also the definition of minimally adequate treatment is now explained more detailed. Note: Finnish medical records are usually detailed. Each visit is recorded separately, including the date, the professional, things that were discussed during the visit and all decisions regarding medications. Prescriptions are recorded separately. Also missed visits are recorded, as are drop-outs from planned care. Therefore, information related to different quality indicators of care was relatively easy to retrieve from the records. Although these issues were also asked about in the interview, sometimes the participants had difficulties in remembering details of the care they had received.

Table 2: add footnote to “Major Depression” about time frame (lifetime?) and what depressive disorders were included in the “No” row
Response: These specifications were made in the table 2.

“spesific” -> “specific” (several times in manuscript and tables)
Response: Misspellings have been corrected.

Line 150: specify “treatment strategy was assessed to be adequate according to the case records but the patient discontinued the visits “
Response: Treatment dropout was rated if the treatment strategy was assessed to be adequate according to the case records but the patient discontinued the visits by his own decision. Typically a missed appointment was recorded in the case records as “did not attend, nor cancelled the appointment” and afterwards there could be an epicrisis which also told that the treatment had terminated prematurely because the patient had dropped out from the treatment. It could also be that visits just seemed to stop without an explanation although there had been a plan to continue them. This information has been added to the supplement.

Line 159: please specify “high school” (>10 years in school?)
Response: High school in Finland means twelve years of education usually followed by a matriculation examination. This specification has been made to the manuscript.

Line 178: I suggest to skip the topic of dropout as an “independent variable”: since dropout was defined as discontinuing adequate treatment it makes no sense to determine “how dropping out affects the adequacy of treatment”
Response: We excluded a dropout as an independent variable because the weaknesses of it.

Line 181: please make always clear (also in tables) when associations are bivariate and multivariate/adjusted (to determine “independent” associations)
Response: We have added specifications of the type of associations into the manuscript. The information of the type of analysis was also added in the tables as a footnote.

Results:
Response: Any hospital treatment was received by 9.2% of participants, of 77% spent more than four days in the hospital. In the analysis we took account those 7.0% of participants with more than four days of hospitalization among all participants in our sample. This information has been written in more detailed to the manuscript.

Response: As mentioned, we excluded a dropout as a determining independent variable because of weaknesses you mentioned. Dropout is also defined more detailed in the supplement now. Dropout was defined when we could see in case records that there was an adequate treatment plan (for example related to medication use or psychotherapy) but the patient had discontinued the visits by his own decision. “Dropouts were more likely to have had all included treatments” means that dropout was associated with having pharmacotherapy, having at least 1 session of psychotherapy and at least 1 and as well as 4 visits with a physician. So, it seems that there had been a plan of intensive treatment, but although these people had been in treatment long enough to e.g. visit a physician four times, they had then discontinued their visits. This reflects the clinical challenges of engaging young people to long-term mental health treatment.

Response: We added a supplement table 3 about logistic regression models of variables associated with treatments received and dropout during the most intensively treated depressive episode into the supplement. Short descriptions of differences between multivariate analysis of most recent and most intensively treated depressive episodes have been made in the result-part of the manuscript.

Response: The line of “treatment dropout” was dropped.

Response: Footnote h was added into a table 4.

Response: We added to the footnotes of table 5 the information that all the variables were entered simultaneously into a logistic regression model.

Discussion:

Response: Health 2000 adult (30+ years) and young adult (18-29 years) study samples were drawn at the same time, using a two-stage stratified cluster sampling design covering the total adult population in mainland Finland. However, while the adult sample was administered a structured psychiatric interview (M-CIDI), this was not done for the young adult sample, and therefore the current mental health substudy was conducted. In terms of sampling strategy and questions used to evaluate the treatment received for mental health problems, the
adult sample and the young adult sample correspond well with each other. However, we have now modified the Discussion and have taken away the text describing the detailed results from the adult survey.

- Line 277ff: Is this conclusion (most complicated cases more often at physicians) a) really supported by the data and b) isn’t it normal that physicians meet complicated cases more often than “the others” (who?)?
  Response: People with comorbid substance use disorders had visited a doctor more often, and people with MDD has more often visited a doctor at least 4 times, which would support our conclusion. It was also possible that a person had visited e.g. a psychologist without ever meeting a physician, for example in student health care services.

Our main result that none of the socioeconomic or clinical factors we studied was associated with receiving minimally adequate treatment gives an impression that our service system operates in a random way and is not able to identify patients with more treatment need and treat them adequately. Therefore, we wanted to point out that at least these more complicated cases more often visit physicians.

- Line 290/295: the gender effect topic should be discussed at the beginning of this paragraph and not just be added as supplemental information in the end (or skip the whole paragraph concerning “other”)
  Response: We decided to skip the whole paragraph concerning current employment, because it was employment during the interview. Therefore it was not a predictor but rather a consequence of treatment.

- Line 301: an increase from 14% to 47% in less than 10 years must have also methodological reasons
  Our finding actually fits with register-based information on e.g. medication use and outpatient clinic visits in young adults. Unfortunately, the national SOTKAnet Statistics and Indicator Bank (see http://uusi.sotkanet.fi/portal/page/portal/etusivu) only reports antidepressant medication use in the age group 18-24 years separately from all adults, but this is how it looks like since 1996:

Reimbursements for depression medicines, recipients aged 18-24, as % of total population of the same age (id: 2355 [info])

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole</td>
<td>1.3</td>
<td>1.6</td>
<td>1.7</td>
<td>2.0</td>
<td>2.5</td>
<td>2.9</td>
<td>3.2</td>
<td>3.5</td>
<td>3.9</td>
<td>4.2</td>
<td>4.7</td>
<td>4.7</td>
<td>5.2</td>
<td>5.5</td>
<td>5.6</td>
<td>5.7</td>
<td>5.7</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Outpatient mental health visits are only recorded for the whole adult population, and their trend shows an increase between 1996 and 2003-2005 as our results also suggest:

Mental health outpatient visits of adults per 1000 persons aged 18 and over (id: 3075 [info])

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole</td>
<td>366</td>
<td>365</td>
<td>377</td>
<td>395</td>
<td>362</td>
<td>403</td>
<td>448</td>
<td>446</td>
<td>441</td>
<td>439</td>
<td>432</td>
<td>442</td>
<td>440</td>
<td>422</td>
<td>460</td>
<td>439</td>
<td>456</td>
<td>446</td>
</tr>
</tbody>
</table>

We have now added the following sentence in the Discussion:
The Mental Health in Early Adulthood in Finland (MEAF) study was done in 2003-2005 i.e. about ten years ago. After the years of the survey there have been changes in the health care system as well as medication in
Finland, which means that figures of minimally adequate treatment may be different today. Probably, the situation has improved: based on official statistics, the use of antidepressants and the number of psychiatric outpatient visits has increased.

- Line 306ff: see above: I suggest to skip the topic of association between dropout and treatment adequacy. We skipped the topics of associations between dropout and treatment received.