Author’s response to reviews

Title: Characteristic distribution of the total and individual item scores on the Kessler Screening Scale for Psychological Distress (K6) in US adults

Authors:

Shinichiro Tomitaka (tomitaka.shinichiro@jp.panasonic.com; stomitaka@gmail.com)
Toshiaki Furukawa (furukawa@kuhp.kyoto-u.ac.jp)
Yohei Kawasaki (kawasaki.yohei.2r@kyoto-u.ac.jp)
Kazuki Ide (ide.kazuki.2r@kyoto-u.ac.jp)
Maiko Akutagawa (s15417@u-shizuoka-ken.ac.jp)
Yutaka Ono (yutakaon@gmail.com)
Hiroshi Yamada (hyamada@u-shizuoka-ken.ac.jp)

Version: 4 Date: 26 Jul 2017

Author’s response to reviews:

July 27, 2017

Anna Clark
Editor
BMC Psychiatry

Dear Dr. Clark:

Thank you for your comments concerning our manuscript entitled “Characteristic pattern of the sum of the depressive symptom item scores using the Kessler Psychological Distress Scale (K6) in US adults.” We have read your comments carefully and have made the required corrections, which we hope to meet your approval. The manuscript ID is BPSY-D-17-00157.
1. Authors contributions:

We sincerely appreciate your advice. In line with the ICMJE guidelines, we changed the sentences as follows.

“ST carried out the design of the study and the statistical analysis and wrote the manuscript. YK, KI, MA, contributed to the analysis of the data. HY, YO, and TAF contributed to the acquisition of data. YK, KI, MA, HY, YO, and TAF interpreted the data and wrote the manuscript. All authors read and approved the final manuscript.”

2. Text overlap:

We apologize for our manuscript containing a considering amount of text overlap with our previous publications. We have revised our manuscript in order to remove all text overlaps.

All expressions of all overlapping text have been changed as suggested. Using iThenticate, we have confirmed that the overall similarity index of our revised manuscript is under 25%, and the single match similarity index is less than 10%. However, since the manuscript has been already reviewed by two reviewers, the ideas and structure of the manuscript have not changed. The changes to the manuscript are given below.

##Abstract

1. P 3, line 10. To remove repetitive phrases, "depressive symptom measures" has been rephrased into “depression screening scales.”

2. P 3, line 12. P 5 line 53. P 19 line 7. To avoid text overlapping, "total depressive symptom scores" has been reworded into “total scores on depression screening scales”
3. P 3, line 15, P 5 line 56, P 17 line 37 and 56, P 18, line 17. To remove repetitive phrases, "except for the lower end" has been rewritten into “ with the exception of the lower end”

4. P 3, line 17, line 28, line 35. According to the title, "sum of item scores" has been changed into “total and individual item score”

5. P 3, line 28. To avoid text overlapping, "aged between 25 and 74 years old" has been changed into “between ages of 25 and 74”

6. P 3, line 30, P 10, line 23. To avoid text overlapping, "analyzed using log-normal scales" has been reworded into “investigated with histograms”

7. P3 , line 43. To remove repetitive phrases, “depending on” has been changed into “according to”

8. P3, line 51. To avoid text overlapping, the sentence “While depressive symptom of individuals is difficult to predict, the entire population may follow a certain mathematical pattern” has been rewritten as follows.

“While the degree of depressive symptoms varies from individual to individual, an entire population may show a certain mathematical distribution.”

##Introduction

1. P 5 , line 8. To avoid text overlapping, the sentence “Because depressive symptomology (number and chronicity) forms the basis of a diagnosis of clinical depression, there is considerable interest in understanding the distribution of depressive symptoms in the general population” has been changed into as follows.
“Because the degree of depressive symptoms is fundamental to a diagnosis of clinical depression, the severity distribution of depressive symptoms in the general population is a major area of interest within the field of psychiatry.”

2. P 5 , line 58. To avoid text overlapping, sentence “In an analysis of nearly 10,000 respondents to the British National Household Psychiatric Morbidity Survey, Melzer et al. reported that an exponential curve provided the best fit for total depressive and neurotic symptom scores on the Revised Clinical Interview Schedule (CIS-R), except at the lowest symptom counts [10]” has been rewritten as follows.

“An analysis of data from almost 10,000 respondents to the British National Household Psychiatric Morbidity Survey has revealed that an exponential model best fits for the total score data on the Revised Clinical Interview Schedule (CIS-R), except at the lowest symptom counts [10]”

3. P 6 , line 7. To avoid self-plagiarism, the sentence “The authors of the present study have similarly reported that the distribution of total depressive symptom scores of the Center for Epidemiologic Studies Depression Scale (CES-D) follows an exponential curve except for the lower end of the distribution, based on data from nearly 25,000 respondents to a national survey of the Japanese population [11]” has been changed as follows.

“Analyzing almost 25,000 respondents to the Japanese Active Survey of Health and Welfare, we have similarly reported that the distribution of total scores on the Center for Epidemiologic Studies Depression Scale (CES-D) approximates an exponential pattern, with the exception of the lower end of the distribution [11].”

4. P7 , line 33. To avoid self-plagiarism, the sentence “Furthermore, we observed that (1) the estimated rate parameters of the exponential model were similar across groups with the same number of items, and increased with an increasing number of chosen items; and (2) the ratio of
“none” to “some” for each item response contributed to the non-exponential pattern of total scores at the lowest scores [20]” has been changed as follows.

“Furthermore, we observed that (1) the total score distributions with the same number of chosen items exhibited similar estimated rate parameters; (2) the estimated rate parameters increased in proportion to the number of chosen items; and (3) the ratio of “none” to “some” of item responses contributed to the non-exponential pattern of total score distributions at the lowest scores [20].”

5. P6 , line 21. To remove repetitive phrases, “To confirm the reproducibility of such findings” has been rephrased into “To verify the findings.”

6. P7, line 41. To avoid text overlapping, the sentence “These findings suggest that the item response of such items is key to predicting the distribution of the sum of the item scores.” has been rewritten into “These findings suggest that the distributional patterns of the total scores are dependent on the item responses of such items.”

##Method

1. P 8, line 34. To avoid self-plagiarism, the first paragraph of the Dataset has been rewritten as follows.

“Data were drawn from the MIDUS 1 [21,22]. The MacArthur Foundation Research Network on Successful Midlife Development conducted the MIDUS 1 between 1995 and 1996. The average age of the sample was 46.4 years (SD = 13). Non-English speaking and institutionalized individuals were excluded. The MIDUS 1 sample consisted of 7,108 participants (male: n = 3,395) including four subsamples: the main sample from a national random-digit-dialing (RDD)
procedure, oversamples from five US areas, siblings of participants from the RDD sample, and a national sample of twin pairs.”

2. P 8, line 56. The section started “The main RDD sample was selected from the working telephone banks …” to “The MIDUS 1 researchers also carried out oversampling of five metropolitan areas.” has been removed because it is presented a lot in detail elsewhere (reference 22).

3. P 9, line 16. To avoid self-plagiarism, the paragraph of the Ethics statement has been rewritten as follows.

This study used de-identified data available to the public. The author’s institutional review board did not consider the secondary analysis of publicly available data as human subjects research.

4. P 10, line 4 and 9. To avoid text overlapping, the phrase “compared to those predicted from the exponential pattern of total scores” has been rephrased into “than predicted from the exponential regression curve.”

5. P 10, line 26. P 13, line 9. P 14, line 27. P15, line 17. P 16, line 9 and 58. P 20, line 33. To avoid repetitive phrases, the phrase “a log-normal scale” has been rephrased into “a semi-logarithmic scale”
## Results

1. P 11, line 6. To avoid self-plagiarism, the sentence “A ranking in ascending order was allocated according to the degree of the ratio of “a little” to “none.” has been rewritten into “All 6 items were ranked according to the ratio of “a little” to “none” in ascending order.”

2. P12, line 48 and 50. To remove repetitive phrases, “item response rates” has been rephrased into “item response frequencies.”

3. P 13, line 24 and 53. To avoid text overlapping, “independent variable” and “dependent variables” have been rephrased into “predictor variable” and “predicted variable,” respectively.

4. P13, line 43. To avoid self-plagiarism, the sentence “In support of the small sample sizes being a cause of the fluctuations, the frequencies at the higher scores were very low (less than 10)” has been reworded as follows.

“The possibility that the small sample sizes caused the fluctuations is supported by the fact that the frequencies at the higher scores were very low (less than 10).”


6. P 15, line 21 and 23. P16, line 14 and 16. P 17 line 7 and 9. To avoid repetitive phrases, “compared to those predicted from the exponential patterns” have been rephrased into “than expected from the linear patterns.” “linear patterns” is correct because it is on a semi-logarithmic scale.

7. P15, line 24. P16, line 17. To avoid self-plagiarism, the sentence “The divergence of the actual data from the predicted exponential pattern at the lower end of the distribution was not evident in the middle ratio group (Figure 4E)” has been reworded as follows.
“An apparent gap between the frequencies at the lowest end of the total score and the linear pattern was not detected in the middle ratio group.”

##Discussion

1. P 17, line 34. To avoid self-plagiarism, the first paragraph “The main findings of this study were as follows: (1) regardless of the choice of items, the sum of item scores of K6 approximated an exponential pattern, except for the lower end of the distribution; (2) the estimated rate parameters of the exponential model were similar across groups with the same number of items and increased as the number of chosen items increases; and (3) at the lower end of the distribution, the distributional pattern of the sum of the item scores varied depending on the ratio of “none” to “a little” of the chosen items.” has been rewritten as follows.

“The present studies yielded three main findings. First, irrespective of the choice of items, the sum of item scores of K6 exhibited an exponential pattern, with the exception of the lower end of the distribution. Second, the estimated rate parameters of regression analysis were similar among groups with the same number of items, and increased in proportion to the number of chosen items. Lastly, the ratio of “none” to “a little” of the chosen items contributed to the distributional pattern of the total scores at the lower end of the distribution.”

2. To avoid repetitive phrases, the subheadings in the Discussion has been removed.

3. P 18, line 1. To avoid self-plagiarism, the sentence “Our theory comprises the following three conditions: (1) depressive symptom items are manifest variables influenced by a unidimensional latent trait of depressive symptoms; (2) the latent trait of depressive symptom items follows an exponential distribution; and (3) the threshold of item scoring forms a distribution in accordance with the unidimensional latent trait [11,24,25]” has been rewritten as follows.
“Our theory consists of three main parts. First, all item scores are manifest variables, which correspond to a single latent variable. Second, the latent variable of all item scores follows an exponential distribution. Finally, the threshold of item scoring differs from individual to individual, and forms a frequency distribution according to the single latent variable [11,24,25].”

4. P 18, line 28. To avoid text overlapping, the sentence “Our findings are in line with the results from taxometric analysis, which is a statistical technique specifically designed to determine whether a given construct is best conceptualized as discrete latent subcategories or as one continuous latent dimension [27,28]” has been reworded as follows.

“Our findings concur with the results from taxometric analysis, which is a statistical method specially designed to identify whether given constructs are categorical or dimensional [27,28].”

3. P 18, line 55. To avoid text overlapping, the sentence, “In general, an exponential distribution is observed where individual variability and total stability are organized together, such as the Boltzmann–Gibbs law and income distribution [31] “ has been rewritten as follows.

“Typically, an exponential distribution occurs when individual variability coexists with total stability, for examples, energy levels within atoms and individual income [31].”

5. P 19, line 1. To avoid repetitive phrases, the phrase “With respect to individual variability and total stability” has been changed into “Regarding individual variability and total stability.”

6. P 19, line 55. P20, 4. To avoid repetitive phrases, the phrase, “compared to those predicted from the exponential pattern” have been rephrased into “than expected from the fitted exponential curve.”
7. P 20, line 14. To avoid text overlapping, the sentence “Analyzing the data of the British National Household Psychiatric Morbidity Survey, several studies have shown that, although an exponential curve provided the best fit for total CIS-R scores, the actual frequencies at the lower end of the total CIS-R scores were higher than those predicted from the fitted exponential curve [10,32,33]” has been rewritten as follows.

“Several analyses of CIS-R data from the British National Household Psychiatric Morbidity Survey have revealed that, although exponential regression analysis fitted total CIS-R scores, the actual frequencies at the lower end of the total CIS-R scores were higher compared to those expected from the regression analysis [10,33,34].”

8. P 20, line 41. To avoid text overlapping, the paragraph “The present study has some limitations. Although we evaluated whether the sum of the item scores and item response approximated an exponential distribution, we did not perform an analysis based on other mathematical models. In general, the most important part of model evaluation is testing whether the model fits empirical data better than other models. However, to the best of our knowledge, no other mathematical models for the sum of item scores have yet been reported. Thus, using graphical analysis and regression analysis, we performed the analysis limited to an exponential model. Future studies might seek to evaluate the comparative fit of other models to the data from the Interuniversity Consortium for Political and Social Research (ICPSR) repository” has been rewritten as follows.

“The findings of this research need interpretation in the light of some limitations. Although we demonstrated that an exponential curve fitted the total and individual item score data, we did not fit other functions to the data. Generally, the most important part of model selection is to identify patterns in data. However, there is little published data on the patterns of total and individual item scores. Thus, using histograms and regression analysis, we identified the patterns. More research is needed to assess the comparative fit of other patterns to the data from the Interuniversity Consortium for Political and Social Research (ICPSR) repository”
9. P 21, line 21. To avoid self-plagiarism, the sentence, “The degree to which these findings can be generalized to clinician-rated scales is unclear but warrants examination” has been rewritten as follows.

“It would be interesting to determine whether these findings can be generalized to clinician-rated depression scales.”

##References


##Figure legends

1. P 34, line 14, 26, 39,51, and 56. P 35, line 4, 7, 17, and 20. P 36, line 10 and 13. To avoid repetitive phrases, the phrase “a log-normal scale” has been rephrased into “a semi-logarithmic scale”

2. P 34, line 15 and 19. P 35 line 1 and14. P36, line 7. According to the title, "sum " has been changed into “total scores”

3. P 34, line 9 and 11. P 35, line 2 and 4. P 36, line 16 and 18. To avoid repetitive phrases, the phrase “frequencies compared to those predicted from the exponential pattern” has been changed into “frequencies than predicted from the linear pattern.”
We believe that our manuscript has been improved satisfactorily and hope it will be accepted for publication in BMC Psychiatry.

Yours Sincerely,

Shinichiro Tomitaka M.D.

Department of Mental Health, Panasonic Health Center
Landic building 3F, Nishishinbashi 3-8-3, Minato-ku, Tokyo 105-0003, Japan
E : tomitaka.shinichiro@jp.panasonic.com