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: 1 Date: 09 Jun 2017

Author’s response to reviews:

June 9, 2017

René Ernst Nielsen
Associate Editor
BMC Psychiatry

Dear Dr. Nielsen:

Thank you for your letter and for the reviewer 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 the comments of the two 
reviewers carefully and have made the required corrections, which we hope meet your approval. The manuscript ID is BPSY-D-17-00157.

Our responses to the reviewers’ comments are given below. We thank the reviewers for carefully reading our manuscript and for giving critical input. The reviewers’ comments were most helpful and gave us useful suggestions to improve our paper. 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

Reviewer 1 (Mirko Manchia):

## Major concerns

1. Comment: “However, this manuscript misses to translate these results in clinical terms. In other words, the authors should make an effort to highlight the clinical importance of these findings, and importantly this should reflect in the conclusions of the Abstract as well. Why should we care about a specific distribution (normal, exponential, log-normal) of severity of depressive symptom in the general population? Have these findings any relevance on predicting clinically relevant depressive symptoms? This is marginally discussed by the authors in lines 14-31 page 16. However, the exponential distribution of unidimensional latent trait of depression might not necessarily translate into a clinically relevant measure.”
Response: We agree with the reviewer’s comment. We have added sentences to translate the results in clinical terms as follows.

P 3, Line 40. The present results potentially enable estimation of how depressive symptoms distribute in the general population. While depressive symptoms of individuals are difficult to predict, the entire population may follow a certain mathematical pattern.

P17, Line 47. A score of 13 has been frequently suggested as a cut-off point for serious mental illness. Our results indicate that the exponential pattern covers a wide range of K6 scores beyond the cut-off point. These results suggest that depressive symptoms are better conceptualized as a continuously distributed syndrome rather than a discrete diagnostic entity. 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 [26,27].

P18, Line 26. The finding that the total scores followed an exponential distribution potentially enables further insight into the mechanism of depressive symptomatology. 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 [30]. With respect to individual variability and total stability, mood of individuals often changes depending on the situation, and the distribution of the depressive symptom scores in a general population is stable from 30–69 years [11,12]. Further studies are needed to clarify the mechanism of depressive symptomatology.

P21, Line 23 (Conclusion). Based on the distributional patterns observed in this study, it is necessary to reconsider the statistical model of normally distributed depressive symptom scores, which is often presupposed in population studies on depressive symptomatology. The present results potentially enable further insight into the mechanism of depressive symptomatology, and estimation on how depressive symptoms distribute in the general population.

2. Comment: “Another major point is whether this instrument is appropriate to measure the distribution of depressive symptoms in the general population.”
Response: The K6 is a reliable and valid tool to screen for depressive symptoms and mood disorders in epidemiological survey research. An explanation has been added as suggested.

P 5, Line 34. The K6 is a reliable and valid tool to screen for depressive symptoms and mood disorders in epidemiological survey research [15,16].

3. Comment: “A related point concerns the validity of K6 in measuring depressive symptoms. K6 is a self-rated instrument. It would be of importance to know if there is concordance with clinician-rated measures of depression (Hamilton?) and if these objective measures have distributional properties similar to those found by the authors in K6 (if such data do exist).

Response: In general, moderate-to-strong correlations are reported to exist between clinician-rated scales and self-report questionnaires in the measurement of depression severity. However, to the best of our knowledge, the distributions of total scores or item responses using clinician-rated scales have not been reported. The sentence has been changed as suggested.

P 21Line 15. The degree to which these findings can be generalized to clinician-rated scales is unclear but warrants examination

4. Comment: The following statement is convoluted: "We assume that the latent trait of depressive symptoms follows an exponential distribution, because this assumption enables us to explain the mathematical patterns of the sum of item scores and item responses."

Response: We agree with the reviewer’s comment. The sentence has been changed as suggested.

P 18, Line 9. Based on the mathematical patterns of the total and individual item scores, we assume that the latent trait of depressive symptoms follows an exponential distribution.
## Minor concerns

1. Comment: “There are typos in the title as well as in the text (please see line 26 page 14).”

Response: Thank you for your advice. The typos have been corrected as suggested (title and P 15, Line 34).

Reviewer 2 (Rob Saunders):

We really appreciate your careful reading of our manuscript and your critical input. We have tried to answer your questions to the best of our ability.

## General concerns

1. Comment: The phrase "lower end of the distribution" is used nearly 30 times in the manuscript which is a bit too much. Please change the wording periodically to aid readability as it seem repetitive at the moment - mixing in phrases like 'lowest symptom counts', 'individuals with low depression scores'.

Response: We agree with the comment. The sentences have been changed as suggested.

P 3, line 31; P3, line 33; "lower end of the distribution" has been changed to “lower scores.”

P 6 line 43; P17, line 25, and P20 line 7. "lower end of the distribution" has been changed to “lowest scores.”

P 5, line 7; P 5 line 55, and P17 line 25. "lower end of the distribution" has been changed to “at the lowest symptom counts.”
P 9, line 24. "lower end of the distribution" has been changed to “at the lowest end of the scores.”

2. Comment: “I feel a bit more balance is needed, as there is a danger readers might think you are ignoring the higher end and focusing on your lower-end theory instead.”

Response: As you pointed out, the distribution certainly fluctuated at the highest scores. We think that the fluctuation at the highest scores reflects the small sample sizes. In support of the small sample sizes being a cause of the fluctuations, the total scores fluctuated more with increases in total scores and when frequencies at higher scores were very low (less than 10). In previous papers, we have repeatedly discussed fluctuations at the higher scores. Thus, in the present paper, we did not focus on the “higher end.” We are sorry for our lack of explanation. To explain fluctuations at the highest scores, we added sentences as follows.

P 12, line 39. Using a log-normal scale (Figure 2B), the total scores of the 6 items showed a linear pattern for almost the entire range of total scores. Of note, the total scores fluctuated more with the increase of total scores, consistent with previous studies [11,17]. It is likely that the small sample sizes for the highest scores caused the fluctuation. 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).

P 13, line 34. Consistent with the total scores of the 6 items (Figure 2B), the distribution fluctuated at the highest scores (Figure 3C and 3D).

P14, line 34. The distribution fluctuated at the highest scores.
1. Comment: “Typo - please correct (characteristic pattern)”

Response: The typos have been corrected.

2. Comment: “I also recommend it is changed slightly, as currently it isn't that clear. The phrase 'sum of the depressive symptom item scores' doesn't seem right - wouldn't changing to 'distribution of total and individual item scores' be more clear considering the contents of the paper?”

Response: We agree your advice. The title has been changed as suggested (P 1, line 4)

## Abstract

1. Comment: “Start with "The distributional pattern…""

Response: The words have been changed as suggested (P 3, line 8).

2. Comment: “Replace ‘lower end of the distribution’ with 'lower scorers', or similar.”

Response: We agree with this comment.

P 3, line 31; P3, line 33; "lower end of the distribution" has been changed to “lower scores.”
## Introduction

1. Comment: “The very first sentence is copied directly from your 2015 and 2016 papers, please amend.”

Response: We agree with the comment. The sentence has been changed as suggested.

P 4, line 8. Depression is a common psychological disorder that affects about 350 million people worldwide.

2. Comment: “Explain why understanding the distribution of depressive symptoms in the general population is important? Your description is limited but it is the justification of this study.”

Response: We agree with this comment. A sentence has been added as suggested.

P 4, line 18. Understanding the distribution of depressive symptoms in the general population is important because statistical hypothesis tests and statistical estimators are derived from statistical models, which are assumed to adequately approximate the empirical distribution [5].

3. Comment: “Suggest re-writing first sentence, second paragraph. Increasing sample size reduces uncertainly that findings reflect the population. There's an assumption a larger sample more closely approximates the population distribution, but it isn't guaranteed.”

Response: We deleted this sentence because it is uncertain.
4. Comment: "exponential pattern, except for the lower end of the distribution" - this needs to be defined for the general reader of this journal, who may not be as comfortable with the concepts used in this paper. This concept is fundamental to the interpretation of the findings later.

Response: We think that "exponential pattern, except for the lower end of the distribution" is not a concept, but a concrete pattern. As is often the case with patterns, it is difficult to define the pattern more clearly with words. From our point of view, to have general readers understand the aim of the study, it may be better to explain the importance of visualization with histograms. The point of our study is that “graphs force us to note the unexpected; nothing could be more important (John Tukey).” Thus, we added explanations as follows.

P 4, line 26. Population studies of depressive symptoms have been conducted using parametric statistics, factor analysis, and item response theory. These methods are called confirmatory data analysis (CDA), which presupposes a statistical model, tests a hypothesis, and estimates parameters [6]. On the other hand, exploratory data analysis (EDA) is an inductive approach designed to reveal characteristics and patterns in the data, often with visual methods [6,7]. Both inductive and deductive approaches are necessary for data analysis. However, whereas CDA has been performed by a great many researchers, exploratory data analysis is still not well developed in the field of psychiatry. For example, little work has been done to visualize the distributions of depressive symptom scores in the general population to understand their characteristics patterns.

Visualizations are central to EDA because the rich information they provide is unrivaled in its ability to detect data patterns [8,9].

5. Comment: “I would suggest that you change the depression of the K6 items as saying "anxiety" is a 'depressive symptom' might cause confusion (and clinical disagreement). Changing to 'nervousness' or 'worry' might be safer.”

Response: We agree with this comment. The word “anxiety” has been replaced with “worry” as suggested. (P 5, line 34)
6. Comment: ‘Rephrase the sentence "Furthermore, we investigated the total scores of the PHQ" to make it clear that it's an upcoming study and place the reference at the end of the first sentence rather than end of the paragraph. No issue with describing finding in this part though, as it's relevant.

Response: This study is now “in submission.” As you suggested, we placed the reference at the end of the first sentence (P 5, line 46).

7. Comment: With the following paragraph ("taken together,…"), I think this is an important part of the introduction but it isn't strong enough. It links to a part about the assumption of normal distributions and psychometric methods (i.e. item response theory) you make in the discussion. This justifies the value of this study.

Response: We agree with this comment. A sentence has been added as suggested.

P6, line 14. Furthermore, if the pattern of the empirical distribution follows a non-normal distribution, the statistical model of normal distributed depressive symptom scores, which is presupposed in parametric statistics and item response theory, requires reconsideration.

8. Comment: “page 6 - phrase "reproducibility of the aforementioned findings" is repeated in lines 14 and 17”

Response: As you suggested, we changed the words as follows: “to confirm the preceding findings” (P 7, line 7).
# Method

1. Comment: I wonder why the section starting "Of the RDD respondents who reported…" to "….solicit their participation in the survey" is of value to this study? I think you could drop and spend the words elsewhere. I think you could cut the dataset section even further as surely it's presented in a lot of detail elsewhere.

Response: We agree with this comment. The sentences have been deleted as suggested. (P 8, line 15)

2. Comment: “Analysis procedure - please provide percentage of individuals excluded from the sample.”

Response: We agree with this comment. The percentage of individuals excluded from the sample has been added as suggested (P 9, line 12).

3. Comment: “As so much analysis has been conducted as part of this study, the 'analysis procedure' section needs to be clear, and would benefit from a bit more detail. The part about "non-exponential distribution at the lowest end of the scores" could be more clearly defined, and why calculating ratios helps this.”

Response: We agree with this comment. The explanation of the “analysis procedure” has been added as suggested. We added an explanation as to why calculating ratios is necessary to understand the "non-exponential distribution at the lowest end of the scores."

P 9, line 22. First, item response rates were calculated for all 6 items. As the previous study demonstrated, the ratio of “a little” to “none” contributes to the non-exponential distribution at the lowest end of the scores [20]. According to the previous study, the total scores of items with
high values of the ratio of “a little” to “none” are expected to exhibit lower frequencies compared to those predicted from the exponential pattern of total scores, whereas the total scores of items with low values of the ratio of “a little” to “none” are expected to exhibit higher frequencies compared to those predicted from the exponential pattern of total scores. On the other hand, the ratios of “some” to “a little” are expected to be similar among all 6 items [20]. Therefore, the ratios of “a little” to “none,” and “some” to “a little” were calculated for all 6 items.

4. Comment: “Linked to above, why was a log-normal scale and regression analysis used? This may not be obvious to all readers, who may not be knowledgeable about methods but it is important to understand why you have done this for the interpretation of the results.”

Response: We agree with this comment. An explanation as to why a log-normal scale and regression analysis was used has been added as suggested.

P 9, line 51. As the graph of an exponential distribution follows a linear pattern with a log-normal scale, a log-normal scale allows us to identify the range of an exponential pattern. Regression analysis was used to estimate the relationship between the sum of the K6 and the frequencies.

##Results

1. Comment: “Figure 1a/b - I couldn't see item 4 in this graph at all. It might be hiding behind another line but at present it appears you've missed it. Please comment on it at least, as looks like an omission.”

Response: The responses of Items 4 and 6 are very similar (Table 1). As you suggested, the line of item 4 is hiding behind that of item 6 (Fig 1). An explanation has been added as follows:

P 33, line 19 (Figure Legends). The line of item 4 is hiding behind that of item 6. The responses of Item 4 and 6 are very similar (Table 1).
2. Comment: “Page 10 - "…after which they show a decreasing pattern" - I'd argue they were all decreasing before the crossover anyway, so please say a little more about this. It's an interesting finding.”

Response: We agree with this comment. The words have been changed to “after which they decreased regularly and converged at one point.” (P 11, line 53)

3. Comment: “Page 11 start - change 'inevitably' to 'as expected'.”

Response: The word has been changed as suggested (P 11, line 7).

4. Comment: "total scores of the 6 items showed linear patterns…” - there is also quite a bit of fluctuation as well, especially at the lower end which deserves comment. Appreciate it's likely due to lower numbers at this level of symptoms. Fluctuation needs a bit a comment in the discussion as well.

Response: We agree with this comment. Sentences have been added as suggested. Since the comment on the cause of the fluctuations is short, we added it in the results section.

P 12, line 39. Using a log-normal scale (Figure 2B), the total scores of the 6 items showed a linear pattern for almost the entire range of total scores. Of note, the total scores fluctuated more with the increase of total scores, consistent with previous studies [11,17]. It is likely that the small sample sizes for the highest scores caused the fluctuation. 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).
5. Comment: “Typo - Ration instead is ratio.”

Response: Thank you for your advice. The typo has been corrected as suggested (P 15, line 34).

## Discussion

1. Comment: “Major point with the discussion and first paragraph - although you have focused on the 'lower end of the distribution' throughout the paper, the higher end of the distribution is equally interesting but hasn't much mention in the discussion at all. In not every set of graphs you present is your statement correct and therefore a bit more balance to your argument is needed. There's a danger that the reader might interpret this as you ignoring the higher end of the distribution just to fit with your 'lower end' hypothesis.”

Response: As you pointed out, the distribution certainly fluctuated at the highest scores. We think that the fluctuation at the highest scores reflects the small sample sizes. 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). Compared to the “lower end of the distribution,” this fluctuation was probably due to the small sample size, is not very important scientifically, and we have repeatedly discussed it in previous papers (11,12,17). Thus, we did not focus on the “higher end” in the present paper. We are sorry for our lack of explanation. To explain the fluctuation at the highest scores, we added sentences as follows

P 12, line 39. Using a log-normal scale (Figure 2B), the total scores of the 6 items showed linear patterns for almost the entire range of total scores. Of note, the total scores fluctuated more with the increase of total scores, consistent with previous studies [11,17]. It is likely that the small sample sizes for the highest scores caused the fluctuation. 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).

P 13, line 34. Consistent with the total scores of the 6 items (Figure 2B), the distribution fluctuated at the highest scores (Figure 3C and 3D).
P 14, line 34. The distribution fluctuated at the highest scores.

2. Comment: “whereas the rest of the discussion is really strong, the conclusion looks rather weak at the moment in comparison which is a shame as this is the take home message. I suggest it's re-written to improve. The first sentence is the same statement repeated. The next sentence doesn't have the desired effect as it's not clear why this information in the general population impacts clinical practice, and this isn't mentioned elsewhere adequately in this manuscript - but I suggest it is to improve reach.”

Response: We agree with this comment. The conclusion has been re-written as suggested.

Conclusion (P 20, line 24).
The findings from this study support the hypothesis that total scores on the self-rating depression scales follow an exponential pattern, except for the lower end of the distribution. Based on the distributional patterns observed in this study, it is necessary to reconsider the statistical model of normally distributed depressive symptom scores, which is often presupposed in population studies on depressive symptomatology. The present results potentially enable further insight into the mechanism of depressive symptomatology, and estimation on how depressive symptoms distribute in the general population.

##Figure legends

1. Comment: “You have written 'rate' rather than 'ratio' for all legends/graphs which appears wrong.”

Response: We agree with this comment. “Rate” has been changed to “ratio” as suggested (legends and graphs)
2. Comment: “I recommend in the legends you specify which items are in the low/medium/high groups for each set of graphs.”

Response: We agree with this comment. We have specified which items are in the low/medium/high groups for each set of graphs (legends).