Author's response to reviews

Title: Reliability and validity of ADHD diagnostic criteria in the Assessment System for Individuals with ADHD (ASIA), a Japanese semi-structured diagnostic interview

Authors:

Toshinobu Takeda (t-tak@umin.ac.jp)
Yui Tsuji (pyhlg_scooy@yahoo.co.jp)
Teruhisa Uwatoko (uwatoko.teruhisa.8s@kyoto-u.ac.jp)
Hiroshi Kurita (hkurita@mvf.biglobe.ne.jp)

Version: 3
Date: 24 April 2015

Author's response to reviews: see over
April 23, 2015

Dear Dr. Teicher and Dr. Ballegooijen;

On behalf of all the authors, first of all I would like to thanks to all of you for careful, thoughtful and instructive suggestions to our manuscript and ask you again to consider our manuscript entitled “Reliability and validity of ADHD diagnostic criteria in the Assessment System for Individuals with ADHD (ASIA), a Japanese semi-structured diagnostic interview” for publication in BMC psychiatry as a research article.

We provided our answers to all of your suggestions below. In addition to the answers to the suggestions from you and the two reviewers, we made minimum amendments in the manuscript solely for the clarity of the statements.

All study participants provided informed consent, and the study design was approved by an ethics review board.

This manuscript has not been published and is not under consideration for publication elsewhere. All the authors have read the revised manuscript and have approved this submission.

The authors declare no conflict of interest in this paper.

Sincerely,

Toshinobu Takeda, M.D., Ph.D.

Department of Clinical Psychology, Ryukoku University, 125-1 Daiku-cho, Oomiya-higashihairu, Shichijyo-dori, Shimogyo-ku, Kyoto 600-8268, Japan.

t-tak@umin.ac.jp, +81-75-343-3074
An answer to the comment from Dr. Levy

Comment: The construction of a DSM-based structured interview for adolescent and adult ADHD based on DSM5 is timely. The present MS is obviously of great interest to Japanese investigators but the MS would benefit by publishing one or two of the questions in the supplementary material as examples (one hyperactivity and one inattention) in the main document, perhaps in reduced form as examples.

Answer: Thank you for this valuable comment. We added two example questions for one symptom from hyperactivity-impulsivity to our manuscript. Thus, now it has example questions for hyperactivity-impulsivity as well as inattention as shown in Appendix.


Answers to the comments from Dr. Kristiansen

Comments

Major Compulsory Revisions

1. Limitations:
The paper concludes upon the validity of ASIA compared to the consensus diagnosis. As it is already stated by the authors in the limitations-section the second author is blinded to the consensus diagnosis – however, whether the first author is blinded or not does not appear in the paper. As the second authors ratings are based on a videotaped interview performed by the first author, this person is [as it is mentioned in the paper] likely to be affected of any bias from the first author. If the first author is biased due to not being blinded this is a limitation and this should be explained in the paper.

Answer: Thank you for this very important comment. Our answer to this comment consists of two groups of sentences placed in Procedure and Discussion as follows:

The first one (from line 221 on page 11 to line 224 on page 12 in Procedure) concerns the exact state of blindness of the first author about participants’ information: “The first author administered the ASIA to the participants and was blind to participant information, including the J-CAARS-S scores and the additional records used in the consensus diagnoses. This author knew only that participants suspected that they suffered from or were interested in ADHD or other psychiatric disorders.”

The second one (from line 362 on page 18 to line 369 on page 19 in limitations in Discussion) mentions how to address the limitation in our methodology to study inter-rater reliability and the merit to report the findings concerning inter-rater reliability despite the limitation this time as follows: “However, as the second author rated the interview while watching the videotape of the first author’s interviews, she may have been influenced by the first author’s biases, even though all possible measures were taken to make the first author blind to any information about participants except for the ASIA ratings. This shortcoming needs to be addressed by a further study employing a more rigorous methodology (i.e., two interviewers independently conducting ASIA interviews). Even with the inherent methodological
problems, it was important to examine interrater reliability in the present study, as this semi-structured diagnostic interview can be administered by a variety of clinicians.”

2. Discussion/conclusion:
   The paper concludes that “ASIA ADHD criteria showed satisfactory internal consistency reliability, inter-rater reliability, and concurrent, discriminant, and content validity.”
   In order to support this statement, the discussion should compare the findings to the existing literature regarding what is considered to be satisfactory levels of these measures. An example of where this could be of great value is the following statement from the discussion: “Overall, these results showed concurrent validity of ASIA ADHD criterion A.”

Answer: Thank you for this important comment. We used a more conservative term “acceptable” instead of “satisfactory” to evaluate the magnitudes of variables concerning reliability and validity throughout the paper. Our answers to this comment are as follows:

(1) Internal consistency reliability ($\alpha$)
   Except for $\alpha$ 0.64 in one symptom “h. difficulty in waiting turn in adulthood,” values of $\alpha$ (ranged from 0.73 to 0.92 as shown in Tables 1 and 2) in the remaining 35 symptoms can be considered acceptable (acceptable $\alpha$ range is from 0.70 to 0.95 [Ref. 22]), though not satisfactory (as noted from line 283 on page 15 to line 288 on page 15 in Discussion).

(2) Inter-rater reliability
   Values of $\kappa$ in 144 questions ranging from 0.98 to 1.00 are high ($\kappa$ between 0.81 and 1.00 is interpreted as almost perfect [Ref. 23]) and raw agreement rates in $\kappa$ incariculable criteria B, C, D and E are also high (0.97-1.00) indicate acceptable inter-rater reliability despite the limitation of the methodology (the second author viewed the first author’s videotape interview) as mentioned in the first paragraph in Discussion (from line 283 on page 15 to line 294 on page 15). The need of addressing this limitation is also described from line 362 on page 19 to line 369 on page 19 in limitations in Discussion as the latter half of the answer to the first comment.
(3) Concurrent validity
Concurrent validity of ASIA ADHD criterion A with the consensus diagnoses as criteria seems acceptable, because the consensus diagnoses agreed with the ASIA ADHD diagnoses (ADHD vs. non-ADHD and subtype diagnoses) in 59 of the 60 cases. In one case, the consensus diagnosis was other specified ADHD, whereas the ASIA diagnosis was non-ADHD. This case met the other specified ASIA ADHD diagnosis except for criteria A and D by showing three positive inattention symptoms and four positive hyperactivity–impulsivity symptoms in childhood, and three positive inattention symptoms and three positive hyperactivity–impulsivity symptoms in adulthood (as mentioned in detail from line 295 on page 15 to line 307 on page 16 in Discussion).

In addition, the correlations between ASIA ADHD criterion A subscale scores and J-CAARS-S subscale scores were moderate to high ($r = 0.44$ to 0.857) and the correlations for the corresponding subscale scores were higher ($r = 0.66$ to 0.857). Whereas, in a similar comparison, correlations between corresponding items on the CAADID DSM-IV ADHD symptom counts and CAARS-S (n = 30) ranged from 0.20 to 0.52 [26], lower than the correlations in our study. These findings also indicate acceptable concurrent validity for ASIA ADHD criterion A (as noted in lines 308 to 316 on page 16 in Discussion and in Table 3).

(4) Discriminant validity
Our findings that compared to the non-ADHD group, the ADHD group showed significantly higher scores in 125 of the 144 questions and non-significantly higher scores in the remaining 19 questions seem to show acceptable, though not satisfactory, discriminant validity (as noted in line 317 on page 16 to line 329 on page 17).

(5) Content validity
We think that the exhausting efforts of developing the 144 questions, key components of the ASIA ADHD criterion A, by experienced clinicians in ADHD introduced in the latter half of Part 1 in Instrument in Methods (in lines 127 to 141 on page 7) ensure one aspect of content validity of the ASIA ADHD criterion A (as also noted in Discussion (in line 328 and 329 on page 17).

3. Inter-rater-reliability
The inter-rater-reliability is being questioned by the author due to limitations (inter-rater-bias), and an argument as to why these results are still considered important, would be useful, for example by refereeing to similar studies.
Answer: Thank you for this comment. The answer to this comment is already presented in inter-rater reliability in the answer (2) to comment 2 above.

Minor Essential Revisions

1. English: This paper is well written. However, it would benefit from some minor editing as there are some grammatical errors throughout the paper.

Answer: Thank you this valuable comment. We asked a professional editing service, Edanz which BioMed Central recommends to improve the English in our manuscript.

Discretionary Revisions

Comments

1. Methods: This paragraph thoroughly describes ASIA and the development of this. The paragraph is long, and I would suggest that you use subheadings in order to make it more easily read. A suggestion would be to separate the development part from the description of ASIA, and further a subheading for each of parts One and Two of ASIA, as well as a third for “Assessment” or “Application of the interview”.

Answer: Thank you for this instructive comment. We provided the subheadings, Part One, Part Two, and Parts Three and Four in Instrument in Methods according to your suggestion although they are not exactly as suggested.

2. Discussion:

2.1: Could you support the following with a reference? “This disagreement typically shows that the milder ADHD symptoms become, the more difficult to determine whether there is disorder [impairment] or not”

Answer: Thank you for this comment. To address this comment, we revised sentences by quoting Diller’s statement [25] “The disagreement in diagnoses for this case supports Diller’s point that “although severe and even moderate ADHD might not be
too difficult to diagnose, mild ADHD, especially the inattentive type, is wide open to interpretation as to what behavior crosses the line from a variation of normal to the deviancy or disorder” (in line 303 on page 16 to 307 on page 16).

2.2: Could you support the following statement with a reference? “Even in adults with ADHD, these symptoms could be suppressed through a long period of discipline.”

**Answer:** Thank you for this comment. We found this statement has no support from a literature. Rather the statement in DSM-5 [4] “excessive running/climbing in adolescent or adults, may be limited to feeling restless is far much plausible. Accordingly, we deleted the sentence “Even in adults with ADHD, these symptoms could be suppressed through a long period of discipline” and inserted the following sentence instead in line 322 on page 16 to line 325 on page 17.

“The small number of ADHD combined presentation cases might partly be a result of the attenuating nature of hyperactivity-impulsivity symptoms in adulthood, as described in the DSM-5 ADHD criteria for hyperactivity symptom c, which indicates that “excessive running/climbing may be limited to feeling restless in adolescents or adults.”