ARTICLE DETAILS

TITLE (PROVISIONAL)	The Sleep Condition Indicator: a clinical screening tool to evaluate Insomnia Disorder
AUTHORS	Kyle, Simon; Espie, Colin; Hames, Peter; Gardani, Maria; Fleming, Leanne; Cape, John

VERSION 1 - REVIEW

REVIEWER	Pallesen, Staale University of Bergen, Norway
REVIEW RETURNED	31-Oct-2013

GENERAL COMMENTS	Insomnia is one of the most prevalent sleep disorders, and several instruments assessing insomnia symptoms have been developed over the years. Recently, the American Psychiatric Association released the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) where some changes in the criteria for insomnia were made – most notably the non-restorative sleep criterion has been removed and the time frame criteria has been increased from 1 to 3 months. Also new to the DSM-5 is a specific frequency criterion referring to symptom occurrence at leased 3 times per week. Due to this, development of new self-report instruments of insomnia emphasizing the new criteria is warranted. This paper presents data on such an instrument; "the Sleep Condition Indicator". Major issues: Page 3: I would prefer to see the introduction part to end with some specific research questions (or alternatively research hypotheses). The authors should describe the item construction phase somewhat more in detail. Notably the new scale does not seem to contain items pertaining to early morning awakening. It contains a question about sleep quality, however this is not mentioned in the DSM-5 criteria for insomnia disorder. These issues should be addressed. The items seem to have been generated before the release of the DSM-5 – thus the items do not seem specifically to have been based on the DSM-5 criteria which may be a limitation. Concerning the PCA – the authors should present information on how they decided on the number of factors to be retained (e.g., Kaiser, Scree plot, parallel analysis). Which criterion was used do decide on whether or not an item loaded on a factor (>.40?) The logistic regression analysis described on page 9 seems somewhat unclear. Which variables comprised the independent and which comprised the dependent variable? You present standardized B – do you mean OR? Do the R^2 refers to Nagelkerke R^2? This is

unclear.

I think it is important to note that positive score on the SCI not necessarily reflects Insomnia Disorder. Persons suffering from e.g., sleep apnea will assumably score above the cut off on the SCI although the symptoms reflect a another sleep disorder than insomnia. Thus, it should be emphasized that the scale first and foremost is a screening instrument and that other types of data (e.g, about snoring) should be included in a proper diagnostic evaluation.

In terms of future studies investigating the SCI up against polysomnographic data would be of interest.

The discussion is somewhat too general – please make it more closely/directly linked to the results of the present study – e.g. data on factor structure, internal consistency, sensitivity and specificity, convergent and discriminative validity.

Minor issues:
Abstract: Please use the total score cut-off (15) instead of 4.6.

Issues of research ethics is probably taken care of but should be mentioned in the manuscript.

Page 2: It is stated that treatment for insomnia often is poor – please be more specific concerning what is meant by "poor" – poor effect, poor treatment being offered, or few being offered treatment?

Page 2: Concerning arguments for why insomnia should be regarded as an entity of its own – see Harvey AG (Clinical Psychology Review 2001;21:1037-1059).

Page 2: Insomnia seems also to be a risk factor for sick leave, disability pension, etc. (see studies by Sivertsen et al). Also, the economical costs of insomnia on the societal level are significant and may deserve mention.

In terms of the explorative factor analysis the authors have such a large data set that half could be used for an explorative factor analysis and the other half could be used for confirmatory factor analysis (not mandatory but may be considered).

Page 8: Comparisons of scores on SCI among those who did and did not use sleep medication should be backed by statistical inferential data (e.g,, t-value and p).

Page 8: When it comes to the sensitivity and specificity of SCI I suppose the diagnosis was set by some specific clinical examination procedure (gold standard) – please describe this in more detail.

It would be of interest to see how the different items perform for different ages and across gender – thus a Rasch modeling could be warranted (can be added as a suggestion for future research).

REVIEWER	Fernandez-Mendoza, Julio Pennsylvania State University College of Medicine, Penn State Hershey Milton S. Hershey Medical Center, United States of America
REVIEW RETURNED	04-Nov-2013

GENERAL COMMENTS	The paper by Espie et al reports on the validation of the Sleep Condition Indicator (SCI) in a large sample of volunteers drawn from a web-based survey. The results show robust internal consistency and concurrent validity of the questionnaire. However, its primary finding on criterion validity (sensitivity and specificity) is problematic, while that on sensitivity to change already published.
	1. The primary aim of this study is to validate the SCI as an instrument to evaluate DSM-5 Insomnia Disorder. The authors state that "57% screened positive for possible DSM-5 Insomnia Disorder" and a reference is provided. In their previous study, Espie et al (2012) reported the criteria and items used to establish this DSM-5 diagnosis. Criteria 1, 2a, 2b, 3, 4, 5, and 6 for DSM-5 Insomnia Disorder were established using SCI items 7, 1, 2, 4, 5, 6, 3, and 8, respectively; the only 2 diagnostic criteria that were not SCI items were early morning awakening and non-restorative sleep (which were included by exclusion). In the present study, this binary variable of DSM-5 diagnosis is the primary outcome of the study (criterion validity analyses). The major psychometric problem here is that the authors examined the sensitivity and specificity of the SCI against itself. This explains the very high sensitivity and specificity values, especially those for the 2-item score, and the "optimal" SCI cut-off scores approximating the average (e.g., 4.6 in a 0-10 range), which indicates that the authors used SCI items scores equal or greater than 2 to define "DSM-5 Insomnia disorder caseness". Therefore, unless they can demonstrate otherwise, this study cannot examine the SCI's criterion validity in terms of its sensitivity and specificity for DSM-5 Insomnia Disorder.
	2. SCI's sensitivity to change has already been published. This finding should not be included in the results of the present study.
	3. That two SCI items predicted the largest variance in SCI total score in regression analyses means that those items are most associated with higher scores in the SCI and, therefore, more severe sleep disruption but not necessarily presence or absence of an Insomnia disorder. Notice that the authors state that they used logistic regression to examine proportion of variance in SCI total score; however, the outcome was total score (linear) not a binary cut-off (logistic). Given that former DSM-IV and ICSD-2 Insomnia diagnoses have demonstrated poor reliability and validity (Edinger et al, 2011), one cannot expect 2 questions frequently asked in the clinic to have such large sensitivity and specificity if they are not being validated against themselves.
	4. The authors state that the SCI was developed as "an 8-item scale … based strictly upon DSM-5 recommendations". However, the SCI/Insomnia Disorder definition did not include: 1) "early morning awakening with inability to return to sleep" (as the authors acknowledge in the discussion), 2) "the sleep difficulty occurs despite adequate opportunity for sleep" (an important criterion for DSM-5 diagnosis), 3) "the insomnia is not better explained by and does not occur exclusively during the course of another sleep-wake disorder" (another important criterion that goes beyond comorbidity), and 4) "the insomnia is not attributable to the physiological effects of a substance". All these should be mentioned as a limitation of the SCI.
	5. The authors have the opportunity to test the concurrent validity of the SCI also against PSQI and ISI cut-offs Morin et al (2010) have provided ISI cut-off scores that define insomnia caseness.
	6. The authors may want to review the characteristics of other published insomnia assessment scales (Pittsburg, Athens, etc.).

Reviewer 1

Insomnia is one of the most prevalent sleep disorders, and several instruments assessing insomnia symptoms have been developed over the years. Recently, the American Psychiatric Association released the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) where some changes in the criteria for insomnia were made – most notably the non-restorative sleep criterion has been removed and the time frame criteria has been increased from 1 to 3 months. Also new to the DSM-5 is a specific frequency criterion referring to symptom occurrence at leased 3 times per week.   Due to this, development of new self-report instruments of insomnia emphasizing the new criteria is warranted. This paper presents data on such an instrument; "the Sleep Condition Indicator".

Thank you

Major issues: Page 3: I would prefer to see the introduction part to end with some specific research questions (or alternatively research hypotheses).
This is a good point. Thank you. Of course, hypotheses as such may not be so appropriate in a psychometric study, so we have inserted research questions at the end of the introduction. [page 5]

The authors should describe the item construction phase somewhat more in detail. Notably the new scale does not seem to contain items pertaining to early morning awakening. It contains a question about sleep quality, however this is not mentioned in the DSM-5 criteria for insomnia disorder. These issues should be addressed.

We acknowledge this as a limitation and have now extended our discussion section in relation to the coverage of DSM-5 criteria (page 11). We have also provided clarification about the item construction phase. [page 7]

The items seem to have been generated before the release of the DSM-5 – thus the items do not seem specifically to have been based on the DSM-5 criteria which may be a limitation.

Yes, we have clarified this in the point above. In addition, we have extensively revised the manuscript elsewhere to moderate our previous statements about association with (the now published) DSM-5. [pages 7 and 11]

Concerning the PCA – the authors should present information on how they decided on the number of factors to be retained (e.g., Kaiser, Scree plot, parallel analysis). Which criterion was used do decide on whether or not an item loaded on a factor (>.40?)

Thank you. We have added further information on these matters to the text (pages 7&8).

The logistic regression analysis described on page 9 seems somewhat unclear. Which variables comprised the independent and which comprised the dependent variable? You present standardized B – do you mean OR? Do the R2 refers to Nagelkerke R2? This is unclear.

Thank you we have clarified these matters in the revised text. We apologise that we incorrectly stated that we had used logistic regression – in fact it was linear regression, so it is standardized B, and we have now computed the Durbin-Watson statistic. [page 10]

I think it is important to note that positive score on the SCI not necessarily reflects Insomnia Disorder. Persons suffering from e.g., sleep apnea will presumably score above the cut off on the SCI although the symptoms reflect a another sleep disorder than insomnia. Thus, it should be emphasized that the scale first and foremost is a screening instrument and that other types of data (e.g, about snoring) should be included in a proper diagnostic evaluation. In terms of future studies investigating the SCI up against polysomnographic data would be of interest.

Thank you. Yes we agree with these remarks and have added text to this effect to the discussion (page 11).

The discussion is somewhat too general – please make it more closely/directly linked to the results of the present study – e.g. data on factor structure, internal consistency, sensitivity and specificity, convergent and discriminative validity.

Revised accordingly (pages 10&11)

Minor issues: Abstract: Please use the total score cut-off (15) instead of 4.6. Issues of research ethics is probably taken care of but should be mentioned in the manuscript.

Thank you. Clarified (pages 2&6)

Page 2: It is stated that treatment for insomnia often is poor – please be more specific concerning what is meant by "poor" – poor effect, poor treatment being offered, or few being offered treatment?

Thank you. Clarified (page 3)

Page 2: Concerning arguments for why insomnia should be regarded as an entity of its own – see Harvey AG (Clinical Psychology Review 2001;21:1037-1059). Page 2: Insomnia seems also to be a risk factor for sick leave, disability pension, etc. (see studies by Sivertsen et al). Also, the economical costs of insomnia on the societal level are significant and may deserve mention.

Thank you we have added some text and associated references (page 3)

In terms of the explorative factor analysis the authors have such a large data set that half could be used for an explorative factor analysis and the other half could be used for confirmatory factor analysis (not mandatory but may be considered).

This is a good point and we will look to this at some point in the future. We continue to collect data.

Page 8: Comparisons of scores on SCI among those who did and did not use sleep medication should be backed by statistical inferential data (e.g,, t-value and p).

Thank you. We have now added this information on page 6 and removed from page 8.

Page 8: When it comes to the sensitivity and specificity of SCI I suppose the diagnosis was set by some specific clinical examination procedure (gold standard) – please describe this in more detail. It would be of interest to see how the different items perform for different ages and across gender – thus a Rasch modeling could be warranted (can be added as a suggestion for future research).

We have addressed this concern below because it was also raised by the second reviewer (pages 9&11).

Reviewer Name Julio Fernandez-Mendoza   Institution and Country Pennsylvania State University College of Medicine, Penn State Hershey Milton S. Hershey Medical Center, United States of America   Please state any competing interests or state 'None declared': None declared

The paper by Espie et al reports on the validation of the Sleep Condition Indicator (SCI) in a large sample of volunteers drawn from a web-based survey. The results show robust internal consistency and concurrent validity of the questionnaire. However, its primary finding on criterion validity (sensitivity and specificity) is problematic, while that on sensitivity to change already published.

1. The primary aim of this study is to validate the SCI as an instrument to evaluate DSM-5 Insomnia Disorder. The authors state that "57% screened positive for possible DSM-5 Insomnia Disorder" and a reference is provided. In their previous study, Espie et al (2012) reported the criteria and items used to establish this DSM-5 diagnosis. Criteria 1, 2a, 2b, 3, 4, 5, and 6 for DSM-5 Insomnia Disorder were established using SCI items 7, 1, 2, 4, 5, 6, 3, and 8, respectively; the only 2 diagnostic criteria that were not SCI items were early morning awakening and non-restorative sleep (which were included by exclusion). In the present study, this binary variable of DSM-5 diagnosis is the primary outcome of the study (criterion validity analyses). The major psychometric problem here is that the authors examined the sensitivity and specificity of the SCI against itself. This explains the very high sensitivity and specificity values, especially those for the 2-item score, and the "optimal" SCI cut-off scores approximating the average (e.g., 4.6 in a 0-10 range), which indicates that the authors used SCI items scores equal or greater than 2 to define "DSM-5 Insomnia disorder caseness". Therefore, unless they can demonstrate otherwise, this study cannot examine the SCI's criterion validity in terms of its sensitivity and specificity for DSM-5 Insomnia Disorder.

We thank the reviewer for this careful articulation and acknowledge that this is a problem in the way we have presented the data and our argument for the scale. We did not have an independent clinical evaluation of disorder status for the great majority of participants (open access surveys). We have extensively revised the manuscript to qualify reference to caseness per se, and to present the SCI more as a symptom screening measure that reflects DSM-5 criteria. We have also removed our sensitivity and specificity analysis, but based on the reviewer's helpful suggestion below (point 5), we now investigate how SCI cut-offs perform against ISI cut-offs (pages 9&11). Please note these changes are extensive from title of the paper, through abstract, methodology, results and discussion.

2. SCI's sensitivity to change has already been published. This finding should not be included in the results of the present study.

Agreed, we now emphasise reference to the existing publication for these data (page 10)

3. That two SCI items predicted the largest variance in SCI total score in regression analyses means that those items are most associated with higher scores in the SCI and, therefore, more severe sleep disruption but not necessarily presence or absence of an Insomnia disorder. Notice that the authors state that they used logistic regression to examine proportion of variance in SCI total score; however, the outcome was total score (linear) not a binary cut-off (logistic). Given that former DSM-IV and ICSD-2 Insomnia diagnoses have demonstrated poor reliability and validity (Edinger et al, 2011), one cannot expect 2 questions frequently asked in the clinic to have such large sensitivity and specificity if they are not being validated against themselves.

Thank you this was an error in reporting. See comments and response to first reviewer. We agree

with the reviewer that the nature of the dataset would artificially inflate sensitivity and specificity values and have therefore removed this analysis from the manuscript.

4. The authors state that the SCI was developed as "an 8-item scale … based strictly upon DSM-5 recommendations". However, the SCI/Insomnia Disorder definition did not include: 1) "early morning awakening with inability to return to sleep" (as the authors acknowledge in the discussion), 2) "the sleep difficulty occurs despite adequate opportunity for sleep" (an important criterion for DSM-5 diagnosis), 3) "the insomnia is not better explained by and does not occur exclusively during the course of another sleep-wake disorder" (another important criterion that goes beyond comorbidity), and 4) "the insomnia is not attributable to the physiological effects of a substance". All these should be mentioned as a limitation of the SCI.

Thank you. See also our response to the first reviewer, and our explanation about the point in time when we were working on the SCI. We have addressed these limitations further in the discussion section (page 11).

5. The authors have the opportunity to test the concurrent validity of the SCI also against PSQI and ISI cut-offs Morin et al (2010) have provided ISI cut-off scores that define insomnia caseness.

This is a great help thank you and we have incorporated a further analysis in relation to ISI cut-offs (page 9)

6. The authors may want to review the characteristics of other published insomnia assessment scales (Pittsburg, Athens, etc.).

Thank you we have added a further sentence in reference to the ISI and Athens scale when discussing EMA. (page 11)

VERSION 2 – REVIEW

REVIEWER	Pallesen, Staale University of Bergen, Norway
REVIEW RETURNED	08-Jan-2014

GENERAL COMMENTS	I think the authors still needs to comment on the fact that symptoms of other sleep disorders (e.g. delayed sleep phase syndrome, sleep apnea) might lead to high scores on the SCI. It should also be mentioned as limitaton that sensitivity and specificty was investigated against another self-report measures of insomnia (ISI) and not against clinical interview data (gold standard). Beside these minor points, I think the authors have done a good job revising the ms.

REVIEWER	Fernandez-Mendoza, Julio Pennsylvania State University College of Medicine, Penn State Milton S. Hershey Medical Center, United States of America
REVIEW RETURNED	08-Jan-2014

GENERAL COMMENTS	Espie and colleagues have addressed most of the reviewers' concerns and revised their paper satisfactorily. There are only some

minor issues that need to be addressed.

1. The sentences in the Methods section starting with "Around 57% screened positive for Insomnia Disorder" should be moved to the Limitations section and phrased in a more specific way "Around 57% scored 16 or lower in the SCI…". Interestingly, in their Science Center sample the frequency of insomnia cases (i.e., "moderate to severe insomnia") was about 10%, which is consistent with the prevalence of insomnia (as a chronic complaint or syndrome) in general population samples; the latter provides further support to the concurrent validity findings.

2. Please move the Internal consistency findings before concurrent validity, given that a test's reliability always places a ceiling on its validity.

3. Because the authors examined Concurrent validity, set a theoretical/pre-defined threshold for the SCI, and did not use methods of criterion validity (statistically, ROC curves and clinically, a gold standard), they may want to refrain from using the terms discriminant ability and discriminant validity. In fact, the data reported refers solely to concurrent agreement between SCI and ISI in detecting insomnia cases.

4. Please refer to ISI categories in a consistent manner (i.e., either as "no insomnia disorder" vs. "probable insomnia disorder" or "absence or sub-threshold insomnia" vs. "moderate or severe insomnia" or "no insomnia" vs. "probable insomnia"). In regards to the cut-off used, it appears that the authors have used the more stringent criterion (ISI score 15 or higher), while Morin et al (2011) suggested that a cutoff score of 10 and 11 appear to be the best compromise to define/detect insomnia cases in population-based samples and clinical trials, respectively. The authors may want to provide concurrent validity data also using this cut-off or Discuss this issue in relation to the characteristics of the Science Center sample (see point 1 above).

5. SCI's sensitivity to change has already been published. This finding should not be included in the Results section or in the Abstract as part of the aims of the present study. Being a single sentence it can be included in the Discussion.

6. Please identify what method was used in the linear regression analyses to ascertain the best predictors of SCI total score (e.g., backward, forward, stepwise, force entry). In fact, a Table with the output of this regression model would be very helpful here, given that multicollinearity was not an issue.

7. It is unclear whether it was item 7 (severity) or item 8 (duration) that was the best single predictor of SCI total score together with item 3. Please double-check the regression model and the Results and Discussion sections before publication.

VERSION 2 – AUTHOR RESPONSE

Reviewer Name Ståle Pallesen
Institution and Country University of Bergen, Norway
Please state any competing interests or state 'None declared': Non declared

I think the authors still needs to comment on the fact that symptoms of other sleep disorders (e.g. delayed sleep phase syndrome, sleep apnea) might lead to high scores on the SCI.

= Completed (page 11-12).

It should also be mentioned as limitaton that sensitivity and specificty was investigated against

another self-report measures of insomnia (ISI) and not against clinical interview data (gold standard).

= In line with reviewer 2's comment, we have removed all mention of sensitivity/specificity and discriminant validity. We make reference to the need for assessment against clinical interview data. (page 11)

Beside these minor points, I think the authors have done a good job revising the ms.

= Thank you.

Reviewer Name Julio Fernandez-Mendoza, Ph.D.
Institution and Country Pennsylvania State University College of Medicine, Penn State Milton S. Hershey Medical Center, United States of America
Please state any competing interests or state 'None declared': Non declared

Espie and colleagues have addressed most of the reviewers' concerns and revised their paper satisfactorily. There are only some minor issues that need to be addressed.

1. The sentences in the Methods section starting with "Around 57% screened positive for Insomnia Disorder" should be moved to the Limitations section and phrased in a more specific way "Around 57% scored 16 or lower in the SCI…". Interestingly, in their Science Center sample the frequency of insomnia cases (i.e., "moderate to severe insomnia") was about 10%, which is consistent with the prevalence of insomnia (as a chronic complaint or syndrome) in general population samples; the latter provides further support to the concurrent validity findings.

= We have now moved these sentences to the discussion section. We thank the reviewer for pointing out the 10% figure and we now make reference to this in the discussion section, consistent with point number 4.

2. Please move the Internal consistency findings before concurrent validity, given that a test's reliability always places a ceiling on its validity.

= We don't feel this change is appropriate. One could equally or perhaps more strongly argue that validity takes precedence over reliability because if a test is not valid then it is hardly worth even exploring reliability; and validity comes first in terms of construction.

3. Because the authors examined Concurrent validity, set a theoretical/pre-defined threshold for the SCI, and did not use methods of criterion validity (statistically, ROC curves and clinically, a gold standard), they may want to refrain from using the terms discriminant ability and discriminant validity. In fact, the data reported refers solely to concurrent agreement between SCI and ISI in detecting insomnia cases.

= Completed.

4. Please refer to ISI categories in a consistent manner (i.e., either as "no insomnia disorder" vs. "probable insomnia disorder" or "absence or sub-threshold insomnia" vs. "moderate or severe insomnia" or "no insomnia" vs. "probable insomnia"). In regards to the cut-off used, it appears that the authors have used the more stringent criterion (ISI score 15 or higher), while Morin et al (2011) suggested that a cutoff score of 10 and 11 appear to be the best compromise to define/detect insomnia cases in population-based samples and clinical trials, respectively. The authors may want to provide concurrent validity data also using this cut-off or Discuss this issue in relation to the characteristics of the Science Center sample (see point 1 above).

= We have standardised the terminology and made reference to the science centre data in the discussion section.

5. SCI's sensitivity to change has already been published. This finding should not be included in the Results section or in the Abstract as part of the aims of the present study. Being a single sentence it can be included in the Discussion.

= We would prefer to leave it in the results section but have removed it from the abstract. The rationale here is that sensitivity to change is a central psychometric function and we think it helps the reader to see the range of psychometric qualities. It is very clear from the statement that it is not a new result.

6. Please identify what method was used in the linear regression analyses to ascertain the best predictors of SCI total score (e.g., backward, forward, stepwise, force entry). In fact, a Table with the output of this regression model would be very helpful here, given that multicollinearity was not an issue.

= We used stepwise linear regression. We have incorporated clarification as requested in the text but have elected not to take up the reviewer's suggestion of possibly including a table.

7. It is unclear whether it was item 7 (severity) or item 8 (duration) that was the best single predictor of SCI total score together with item 3. Please double-check the regression model and the Results and Discussion sections before publication.

= Thank you. It is item 7, we have amended this spelling error.