## PEER REVIEW HISTORY



## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | What do measures of agreement (κ) tell us about quality of exposure assessment? Theoretical analysis and numerical simulation. |
| **AUTHORS** | Burstyn, Igor; de Vocht, Frank; Gustafson, Paul |

## VERSION 1 - REVIEW

| | |
|---|---|
| **REVIEWER** | Anneli Ojajärvi<br>Finnish Institute of Occupational Health, Finland |
| **REVIEW RETURNED** | 11-Oct-2013 |

| | |
|---|---|
| **GENERAL COMMENTS** | Method: page 5, second paragraph: Explain the purposes of notes [b] and [c].<br><br>Results: Also what is step [b]? |

| | |
|---|---|
| **REVIEWER** | Mary McHugh, PhD<br>National University<br>United States of America |
| **REVIEW RETURNED** | 21-Oct-2013 |

| | |
|---|---|
| **GENERAL COMMENTS** | The paper is well written, addresses an important topic. The author makes a good case for the simulation approach to testing the hypothesis that the kappa and prevalence of exposure can be used to calculate the sensitivity and specificity of an instrument when a direct measure of the validity of the binary classifier.<br><br>Two minor editorial corrections are recommended as follows:<br><br>Page 5, line 13: There seems to be a problem on this line. Should the sentence stated: “[b] selecting candidate set (SN, SP) from values uniformly distributed between lower bounds, (SNl, SPl)” have stated: “[b] selecting candidate set (SN, SP) from values uniformly distributed between lower bounds, (SNl, SPl) and the upper bound (1).”?<br><br>Page 5 of 10, line 38: The word, “our” is misspelled as “out”. Please correct.<br><br>This is an extremely interesting approach to validity estimation. Simulations of this nature are what were used to originally discover the Central Limit Theorem, and to test sample size needed for Factor Analysis. Thus, the simulation approach has a strong precedent and it is good to see it expanded to other uses. |

**VERSION 1 – AUTHOR RESPONSE**

*Comment 1 of Anneli Ojajärvi: Method: page 5, second paragraph: Explain the purposes of notes [b] and [c].

**Response to comment 1 of Anneli Ojajärvi: We added explanation of the purpose of steps [a], [b], and [c]: “The purpose of step [a] in the procedure is to calculate lower bounds on sensitivity and specificity. The purpose of step [b] is to sample candidate values of sensitivity and specificity that lie between their respective theoretical lower and upper boundaries. The purpose of step [c] is to limit values of sensitivity and specificity to only those that, first, are congruent with the theoretical model that relates validity to reliability (Eq. 1), and, second, satisfy the assumption that classification of exposure is better than random (Eq. 2).”

*Comment 2 of Anneli Ojajärvi: Results: Also what is step [b]?

**Response to comment 2 of Anneli Ojajärvi: The original description contained an omission that is now corrected and clarified in explanation of the purpose of step [b].

*Comment 1 of Mary McHugh: The paper is well written, addresses an important topic. The author makes a good case for the simulation approach to testing the hypothesis that the kappa and prevalence of exposure can be used to calculate the sensitivity and specificity of an instrument when a direct measure of the validity of the binary classifier. Two minor editorial corrections are recommended as follows:

**Response to comment 1 of Mary McHugh: Thank you very much for your kind words.

*Comment 2 of Mary McHugh: Page 5, line 13: There seems to be a problem on this line. Should the sentence stated: “[b] selecting candidate set (SN, SP) from values uniformly distributed between lower bounds, (SNl, SPl)” have stated: “[b] selecting candidate set (SN, SP) from values uniformly distributed between lower bounds, (SNl, SPl) and the upper bound (1).”?

**Response to comment 2 of Mary McHugh: many thanks for catching the omission. We corrected it and are glad that the rest of the text was sufficiently helpful to allow the reviewer to identify the correction that was needed.

*Comment 3 of Mary McHugh: Page 5 of 10, line 38: The word, “our” is misspelled as “out”. Please correct.

**Response to comment 3 of Mary McHugh: Correction made.

*Comment 4 of Mary McHugh: This is an extremely interesting approach to validity estimation. Simulations of this nature are what were used to originally discover the Central Limit Theorem, and to test sample size needed for Factor Analysis. Thus, the simulation approach has a strong precedent and it is good to see it expanded to other uses.

**Response to comment 4 of Mary McHugh: We are delighted that our approach is seen as valuable and are encouraged to pursue this line of work in the future.