

MATTERS ARISING

Open Access



Matters arising: methodological issues on evaluating agreement between medical students' attitudes towards drugs/alcohol use during pregnancy by Cohen's kappa analysis

Tianfei Yu^{1*}, Lei Yang², Xinjie Jiang¹, Shuli Shao¹, Wei Sha¹ and Ming Li³

Abstract

Background The purpose of this article is to discuss the statistical methods for agreement analysis used in Richelle's article (*BMC Med Educ* 22:335, 2022). The authors investigated the attitudes of final-year medical students regarding substance use during pregnancy and identified the factors that influence these attitudes.

Methods We found that Cohen's kappa value for measuring the agreement between these medical students' attitudes towards drugs/alcohol use during pregnancy was questionable. In addition, we recommend using weighted kappa instead of Cohen's kappa for agreement analysis at the presence of three categories.

Results The agreement improved from "good" (Cohen's kappa) to "very good" (weighted kappa) for medical students' attitudes towards drugs/alcohol use during pregnancy.

Conclusion To conclude, we recognize that this does not significantly alter the conclusions of the Richelle et al. paper, but it is necessary to ensure that the appropriate statistical tools are used.

Keywords Substance use, Pregnancy, Medical students, Attitudes, Agreement, Cohen's kappa, Weighted kappa

Background

We read with interest the article entitled: "Factors influencing medical students' attitudes towards substance use during pregnancy" which was published in *BMC Medical Education* on 2 May 2022 [1]. The authors investigated the attitudes of final-year medical students regarding substance use during pregnancy and identified the factors that influence these attitudes. They focused on two items,

including drugs and alcohol, regarding the punishment of substance use during pregnancy. Nonetheless, we found that Cohen's kappa value for measuring the agreement between these medical students' attitudes towards drugs/alcohol use during pregnancy was questionable. We recommend using weighted kappa instead of Cohen's kappa for agreement analysis at the presence of three categories. The agreement improved from "good" (Cohen's kappa) to "very good" (weighted kappa) for medical students' attitudes towards drugs/alcohol use during pregnancy. To conclude, we recognize that this does not significantly alter the conclusions of the Richelle et al. paper, but it is necessary to ensure that the appropriate statistical tools are used.

*Correspondence:

Tianfei Yu
yutianfei2001@163.com

¹ College of Life Science and Agriculture Forestry, Qiqihar University, Qiqihar 161006, China

² College of Foreign Languages, Qiqihar University, Qiqihar 161006, China

³ College of Computer and Control Engineering, Qiqihar University, Qiqihar 161006, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Table 1 Punishment for pregnant women using drugs or alcohol

Drugs	Alcohol		
	Disagree	Undecided	Agree
Disagree	200	10	6
Undecided	12	57	11
Agree	13	1	50

k_C : 0.775 ($p < 0.001$, 95% [confidence interval] CI = 0.714–0.837)
 k_{lw} : 0.804 ($p < 0.001$, 95% [confidence interval] CI = 0.746–0.863)
 k_{qw} : 0.831 ($p < 0.001$, 95% [confidence interval] CI = 0.770–0.892)

The data has been cited from the article published by Richelle et al. [1] and undergone modification. k_C : Cohen's kappa, k_{lw} : linear weighted kappa, k_{qw} : quadratic weighted kappa

Main text

Cohen's kappa statistic is generally suitable for evaluating two raters [2]. Especially in Cohen's kappa statistic, the weighted kappa statistic should be used to estimate the inter-rater reliability in the presence of more than two categories [3]. In contrast to Cohen's kappa statistic, weighted kappa statistic relies more on predefined cell weights that reflect agreement or disagreement.

Cohen's kappa is calculated as follows:

$$k_C = \frac{\sum_{j=1}^n u_{jj}(i'i') - \sum_{j=1}^n p_{ij}p_{i'j}}{1 - \sum_{j=1}^n p_{ij}p_{i'j}} \tag{1}$$

Weighted kappa is calculated as follows:

$$k_w = 1 - \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij}p_{ij}}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}p_i q_j} \tag{2}$$

The value of $u_{jj}(i'i')$ is the proportion of objects put in the same category j by both raters i and i' . The value of p_{ij} is the proportion of objects that rater i assigned to category j , and k is the number of raters. Cohen [4] suggested the k value should be interpreted as follows: < 0.20 as poor agreement, $0.20-0.40$ as fair agreement, $0.41-0.60$ as moderate agreement, $0.61-0.80$ as good agreement, and > 0.80 as very good agreement.

In the authors' Table 1, according to the authors' calculation, inter-rater reliability was good for medical students' attitudes towards drugs/alcohol use during pregnancy (Cohen's kappa = 0.775, 95% [confidence interval] CI = 0.714–0.837). But, in our opinion, weighted kappa was more applicable than Cohen's kappa for the presence of three categories. Consequently, we performed the weighted kappa statistics (linear and quadratic) for evaluating the agreement according to the authors' data. The linear weighted kappa value was 0.804 (95% [confidence interval] CI = 0.746–0.863) as very good agreement. The quadratic weighted kappa value

was 0.831 (95% [confidence interval] CI = 0.770–0.892) as very good agreement, too. The greater the difference between the ratings of the same data, the stronger the hint of inconsistency. For example, the penalty for category disagree into category agree should be significantly greater than that for predicting category disagree into category undecided. If using Cohen's kappa, there is no difference between the former and the later. If using linear weights, the penalty for the former is equal to 2 times the later. If using quadratic weights, the penalty for the former equal to 4 times the later. Therefore, we recommend that quadratic weighted kappa should be used to evaluate the agreement for magnifying the degree of inconsistency in the judgment of large level distance.

In conclusion, the authors underestimated the agreement between medical students' attitudes towards drugs/alcohol use during pregnancy. The reasonable agreement was "very good" for medical students' attitudes towards drugs/alcohol use during pregnancy. Anyway, we recognize that this does not significantly alter the conclusions of the Richelle et al. paper, but it is necessary to ensure that the appropriate statistical tools are used. We highlight that the rigor and use of the correct statistical approach is crucial for any scientific publication. Applying appropriate statistical methods can enhance the scientific accuracy of research results.

Abbreviation

CI Confidence interval

Acknowledgements

Not applicable.

Authors' contributions

TY wrote the original draft of the manuscript. LY, XJ and SS were involved in the analysis and interpretation of the data. WS was a major contributor in revising the manuscript. ML contributed to the conception and design of the study. All authors read and approved the final manuscript.

Funding

This work was supported by the Heilongjiang Province Higher Education Teaching Reform Project (SJGY20200799), Fundamental Research Funds in Heilongjiang Provincial Universities (135509160) and Qiqihar University Degree and Postgraduate Education and Teaching Reform Research Project (JGXM_QUG_Z2019003, JGXM_QUG_Z2019002).

The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 4 August 2022 Accepted: 30 January 2023

Published online: 17 February 2023

References

1. Richelle L, Dramaix-Wilmet M, Roland M, Kacenenbogen N. Factors influencing medical students' attitudes towards substance use during pregnancy. *BMC Med Educ.* 2022;22(1):335. <https://doi.org/10.1186/s12909-022-03394-8>.
2. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb).* 2012;22(3):276–82. <https://doi.org/10.11613/BM.2012.031>.
3. Marasini D, Quatto P, Ripamonti E. Assessing the inter-rater agreement for ordinal data through weighted indexes. *Stat Methods Med Res.* 2016;25(6):2611–33. <https://doi.org/10.1177/0962280214529560>.
4. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70(4):213–20. <https://doi.org/10.1037/h0026256>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

