

Letter to the Editor

Interpreting Agreement in AI-Assisted American Society of Anesthesiologists (ASA) Classification

Hakan Emirkadi, MD

Department of Anesthesiology and Reanimation, Sultan 2. Abdulhamid Han Training and Research Hospital, Istanbul, Türkiye

Corresponding Author:

Hakan Emirkadi, MD
Department of Anesthesiology and Reanimation
Sultan 2. Abdulhamid Han Training and Research Hospital
Selimiye Mh, Tıbbiye Cd
Istanbul 34668
Türkiye
Phone: +90 532 260 9367
Email: hemirkadi@gmail.com

Related Articles:

Comment on: <https://medinform.jmir.org/2026/1/e83893>

Comment in: <https://medinform.jmir.org/2026/1/e112521>

JMIR Med Inform 2026;14:e111973; doi: [10.2196/111973](https://doi.org/10.2196/111973)

Keywords: AI; American Society of Anesthesiologists; ASA; ASA physical status; preoperative assessment; agreement; Cohen κ

We read with interest the study by Lin et al [1] evaluating a GPT-4o-powered system for preanesthetic assessment. The agreement between AI-generated and clinician-assigned American Society of Anesthesiologists (ASA) physical status was high. Two aspects of the analysis may help place this finding in clinical context.

First, ASA physical status is an ordinal classification. The report states that agreement was assessed using Cohen κ but does not specify whether weighting was applied. This distinction matters because an ASA III patient classified as ASA II and one classified as ASA I represent different degrees of disagreement. Weighted κ would allow these differences to be reflected in the agreement estimate [2]. The authors themselves note that comparisons with earlier interrater studies are affected by the use of weighted versus unweighted κ ; the same distinction would be useful for their own estimate.

The ASA distribution makes this particularly relevant. Of the 600 patients, 498 (83%) were ASA II, whereas only 57 were ASA III and 6 were ASA IV. The authors provide the distribution of clinician and AI classifications, but not their patient-level cross-tabulation. We therefore cannot determine whether disagreements occurred mainly between neighboring classes or included larger underclassification of higher-risk

patients. A cross-tabulation, with particular attention to ASA III and IV, together with weighted κ , would make the clinical meaning of the reported agreement clearer.

Second, the comparison with and without Taiwan National Health Insurance (NHI) data deserves further detail. The same 600 patients were evaluated under both conditions. The κ increased from 0.518 without NHI data to 0.883 with NHI data, a difference of 0.365. Our concern is not whether the two reported κ estimates differ, but how precisely the magnitude of that change was estimated. No confidence interval or other measure of uncertainty was reported for $\Delta\kappa$, and the authors state that this comparison was summarized descriptively. Because both estimates were obtained from the same patients, their dependence should be taken into account. Reporting uncertainty around $\Delta\kappa$, for example using paired resampling, would provide a more complete estimate of the incremental contribution of NHI data.

Both requests can be met with the existing dataset. Knowing where the disagreements fell, particularly among the 63 patients assigned ASA III or IV by clinicians, and how precisely the gain from NHI data was estimated, would help readers judge how far the reported agreement supports the use of this system in perioperative risk assessment.

Acknowledgments

Generative AI (ChatGPT 5.6; OpenAI) was used for text refinement and language editing. The author reviewed the final content and takes full responsibility for it.

Funding

The author declared no financial support was received for this work.

Data Availability

No datasets were generated or analyzed for this article.

Authors' Contributions

Conceptualization: HE

Writing – original draft: HE

Writing – review & editing: HE

Conflicts of Interest

None declared.

References

1. Lin BH, Lu CW, Hou IC, Lin TY, Liu CE, Khoi CS. GPT-4o-powered preanesthetic AI: development and validation study. *JMIR Med Inform*. Aug 25, 2026;14:e83893. [doi: [10.2196/83893](https://doi.org/10.2196/83893)] [Medline: [42641167](https://pubmed.ncbi.nlm.nih.gov/42641167/)]
2. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. Oct 1968;70(4):213-220. [doi: [10.1037/h0026256](https://doi.org/10.1037/h0026256)] [Medline: [19673146](https://pubmed.ncbi.nlm.nih.gov/19673146/)]

Abbreviations

ASA: American Society of Anesthesiologists

NHI: National Health Insurance

Edited by Andrew Coristine; This is a non-peer-reviewed article; submitted 14.Sep.2026; final revised version received 15.Sep.2026; accepted 02.Oct.2026; published 21.Oct.2026

Please cite as:

Emirkadı H

Interpreting Agreement in AI-Assisted American Society of Anesthesiologists (ASA) Classification

JMIR Med Inform 2026;14:e111973

URL: <https://medinform.jmir.org/2026/1/e111973>

doi: [10.2196/111973](https://doi.org/10.2196/111973)

© Hakan Emirkadı. Originally published in *JMIR Medical Informatics* (<https://medinform.jmir.org>), 21.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.