



OPEN ACCESS

# Response to: Correspondence on 'Prospective predictive performance comparison between clinical gestalt and validated COVID-19 mortality scores' by Meza-Comparán

Adrian Soto-Mota , Braulio Alejandro Marfil-Garza

Salvador Zubiran National Institute of Medical Sciences and Nutrition, Tlalpan, Mexico

## Correspondence to

Dr Adrian Soto-Mota, Salvador Zubiran National Institute of Medical Sciences and Nutrition, Tlalpan, Mexico; [adrian.sotom@incmnsz.mx](mailto:adrian.sotom@incmnsz.mx)

Accepted 12 January 2022  
Published Online First  
27 January 2022

We would like to thank Héctor David Meza-Comparán<sup>1</sup> for their interest in our work and their insightful comments on our study published in the *Journal of Investigative Medicine*.<sup>2</sup> Here we address the issues raised in it.

As correctly pointed out, 'clinical gestalt' refers to 'a physician's unstructured estimate' or an 'overall clinical impression'. No formal definition was provided because it is a widespread term with a consistent connotation in the literature. On the other hand, as was emphasized in our introduction and in our discussion section, the main objective of our work was to make the point that the current validity of all mortality scores is likely impaired by the change in the pandemic context. Therefore, being thorough and including all Mexican COVID-19 mortality scores was beyond the purpose of our work.

Regarding the differences in years of experience between residents, we respectfully disagree. While it is true that senior residents are more likely to be confident than junior residents, this confidence applies most likely to the late clinical scenarios where senior residents have more experience (ie, nosocomial pneumonia, acute respiratory distress syndrome or pulmonary embolism) and not at hospital admission (before these complications occur). Nonetheless, because the admission process to the Internal Medicine residency program in our hospital is the most competitive in the country, the students in the top percentiles of each generation are the ones that usually conform to every resident's cohort and their graduation years tend to be quite homogenous.

With regard to the issues raised about our sample size, we would like to point out that, as was specified in our Methods section, we did not use the default input parameters of easyROC but those necessary for using Obuchowski's method which considers allocation ratio and levels of observer variability.<sup>3</sup> Arguably, and without the intention of falling into a semantic discussion about what it means to be non-inferior in terms of area under the curve (AUC), for any diagnostic

or classification tool it is reasonable to deem relevant any discrepancy beyond the original CI. Since the originally documented AUC of the LOW-HARM score was 0.96 (95% CI 0.94 to 0.98),<sup>4</sup> as detailed in our Methods section, detecting a 0.05 AUC difference with a case allocation ratio of 0.7 (because the mortality at our center is ~0.3) with a power of 0.8 and a significance cut-off level of 0.05 would require 159 patients. Since we included 166 patients and since the discrepancies with the original AUC of all scores were so large, we think it is unlikely our results are due to low statistical power.

In summary, we think that despite its inherent limitations, our work strongly suggests that the clinical utility and predictive performance of most COVID-19 mortality scores (and of many other clinical scenarios as well) demand regular reassessment. In contrast, the inherent Bayesian nature of clinical gestalt makes it continuously sensitive and quick to adapt to highly dynamic contexts (ie, hospitalization strain) while improving continuously as more information updates clinical practice (ie, novel therapies such as dexamethasone and vaccines).

**Contributors** Both authors contributed to the writing of this text and approve the current version.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study does not involve human participants.

**Provenance and peer review** Commissioned; internally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, an indication of whether changes were made, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

**ORCID iD**

Adrian Soto-Mota <http://orcid.org/0000-0002-9173-7440>



► <http://dx.doi.org/10.1136/jim-2021-002243>



© American Federation for Medical Research 2022. Re-use permitted under CC BY-NC. No commercial re-use. Published by BMJ.

**To cite:** Soto-Mota A, Marfil-Garza BA. *J Investig Med* 2022;**70**:975–976.

### REFERENCES

- 1 Meza-Comparán HD. Comment on 'Prospective predictive performance comparison between clinical gestalt and validated COVID-19 mortality scores'. *J Investig Med* 2022;70:972–4.
- 2 Soto-Mota A, Marfil-Garza BA, Castiello-de Obeso S, *et al.* Prospective predictive performance comparison between clinical gestalt and validated COVID-19 mortality scores. *J Investig Med* 2022;70:415–20.
- 3 Obuchowski NA. Sample size tables for receiver operating characteristic studies. *AJR Am J Roentgenol* 2000;175:603–8.
- 4 Soto-Mota A, Marfil-Garza BA, Martínez Rodríguez E, *et al.* The low-harm score for predicting mortality in patients diagnosed with COVID-19: a multicentric validation study. *J Am Coll Emerg Physicians Open* 2020;110.1002/emp2.12259. [Epub ahead of print: 15 Oct 2020].