

Letter to the Editor

Authors' Reply to: Minimizing Selection and Classification Biases Comment on "Clinical Characteristics and Prognostic Factors for Intensive Care Unit Admission of Patients With COVID-19: Retrospective Study Using Machine Learning and Natural Language Processing"

Jose Luis Izquierdo¹, MD; Joan B Soriano², MD

¹Universidad de Alcalá, Madrid, Spain

²Hospital Universitario de La Princesa, Madrid, Spain

Corresponding Author:

Joan B Soriano, MD

Hospital Universitario de La Princesa

Diego de Leon 62

Servicio de Neumología

Madrid, 28006

Spain

Email: jbsoriano2@gmail.com

Related Articles:

Comment on: <https://www.jmir.org/2020/10/e21801/>

Comment on: <https://www.jmir.org/2021/5/e27142>

(*J Med Internet Res* 2021;23(5):e29405) doi: [10.2196/29405](https://doi.org/10.2196/29405)

KEYWORDS

artificial intelligence; big data; COVID-19; electronic health records; tachypnea; SARS-CoV-2; predictive model; prognosis; classification bias; critical care

We acknowledge the letter by Martos Pérez et al [1] and take this opportunity to clarify related issues from our publication [2]. Of our 10,504 patients with COVID-19, 2737 (26.5%) were tested with PCR (polymerase chain reaction). Within the 5 provinces of Castilla-La Mancha, the province that tested the most was Toledo (28.9%), while the least was Guadalajara (21.2%). Those patients in whom PCR was performed were 6.5 years older (63.0 vs 56.5 years). All these differences were highly statistically significant.

You must take into account that our study period was from March 1 to 29, 2020, and including only microbiologically confirmed cases or prolonging the period of inclusion would have resulted in a biased assessment. From March 30, 2020, onwards, most intensive care units (ICUs) at our hospitals collapsed and ICU admissions were highly distorted due to a lack of beds. As we commented in the *Discussion* section, the ICU capacity in Castilla-La Mancha during the study period had not yet been compromised, which protects against possible bias in our training data (all patients requiring critical care were indeed admitted to the ICU). Therefore, it is unlikely that the

absence of a confirmed diagnosis with PCR during the first weeks of the pandemic influenced our results. This was a generalized situation throughout Spain and in most European countries early in 2020. At that time, when a patient was hospitalized, a wide battery of viruses was considered for which there were reagents before performing PCR for coronaviruses. Patients seen during the month of March, in the midst of an avalanche of COVID-19 cases in our region, with negative tests for other viruses and clinical, radiologic, and blood tests highly compatible, did not raise doubts about their diagnosis of COVID-19, and the probability of error was considered negligible [3-5]. For all these reasons, bias in our AI (artificial intelligence) algorithms is highly unlikely. We, however, agree that admission to the ICU can be related to many factors. One strength of our study is that it analyzes the usual clinical practice in the whole population cared for in an entire health care region of Spain during a period when the lack of beds was not a limiting factor. It was not a sample—it was the entire population. Finally, our study objective was not mortality. In other studies, when we addressed mortality, the study period was extended to reliably collect this variable [6,7].

Conflicts of Interest

None declared.

References

1. Martos Pérez F, Gomez Huelgas R, Martín Escalante MD, Casas Rojo JM. Minimizing Selection and Classification Biases. Comment on "Clinical Characteristics and Prognostic Factors for Intensive Care Unit Admission of Patients With COVID-19: Retrospective Study Using Machine Learning and Natural Language Processing". J Med Internet Res 2021 May;23(5):e27142. [doi: [10.2196/27142](https://doi.org/10.2196/27142)]
2. Izquierdo JL, Ancochea J, Savana COVID-19 Research Group, Soriano JB. Clinical Characteristics and Prognostic Factors for Intensive Care Unit Admission of Patients With COVID-19: Retrospective Study Using Machine Learning and Natural Language Processing. J Med Internet Res 2020 Oct 28;22(10):e21801 [FREE Full text] [doi: [10.2196/21801](https://doi.org/10.2196/21801)] [Medline: [33090964](https://pubmed.ncbi.nlm.nih.gov/33090964/)]
3. Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? Eur J Radiol 2020 May;126:108961. [doi: [10.1016/j.ejrad.2020.108961](https://doi.org/10.1016/j.ejrad.2020.108961)] [Medline: [32229322](https://pubmed.ncbi.nlm.nih.gov/32229322/)]
4. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology 2020 Aug;296(2):E32-E40 [FREE Full text] [doi: [10.1148/radiol.2020200642](https://doi.org/10.1148/radiol.2020200642)] [Medline: [32101510](https://pubmed.ncbi.nlm.nih.gov/32101510/)]
5. Xu J, Wu R, Huang H, Zheng W, Ren X, Wu N, et al. Computed Tomographic Imaging of 3 Patients With Coronavirus Disease 2019 Pneumonia With Negative Virus Real-time Reverse-Transcription Polymerase Chain Reaction Test. Clin Infect Dis 2020 Jul 28;71(15):850-852 [FREE Full text] [doi: [10.1093/cid/ciaa207](https://doi.org/10.1093/cid/ciaa207)] [Medline: [32232429](https://pubmed.ncbi.nlm.nih.gov/32232429/)]
6. Graziani D, Soriano JB, Del Rio-Bermudez C, Morena D, Díaz T, Castillo M, et al. Characteristics and Prognosis of COVID-19 in Patients with COPD. J Clin Med 2020 Oct 12;9(10):E3259 [FREE Full text] [doi: [10.3390/jcm9103259](https://doi.org/10.3390/jcm9103259)] [Medline: [33053774](https://pubmed.ncbi.nlm.nih.gov/33053774/)]
7. Izquierdo JL, Almonacid C, González Y, Del Rio-Bermudez C, Ancochea J, Cárdenas R, et al. The impact of COVID-19 on patients with asthma. Eur Respir J 2021 Mar;57(3):57 [FREE Full text] [doi: [10.1183/13993003.03142-2020](https://doi.org/10.1183/13993003.03142-2020)] [Medline: [33154029](https://pubmed.ncbi.nlm.nih.gov/33154029/)]

Abbreviations

AI: artificial intelligence

ICU: intensive care unit

PCR: polymerase chain reaction

Edited by T Derrick; this is a non-peer-reviewed article. Submitted 06.04.21; accepted 13.05.21; published 26.05.21.

Please cite as:

Izquierdo JL, Soriano JB

Authors' Reply to: Minimizing Selection and Classification Biases Comment on "Clinical Characteristics and Prognostic Factors for Intensive Care Unit Admission of Patients With COVID-19: Retrospective Study Using Machine Learning and Natural Language Processing"

J Med Internet Res 2021;23(5):e29405

URL: <https://www.jmir.org/2021/5/e29405>

doi: [10.2196/29405](https://doi.org/10.2196/29405)

PMID: [33989164](https://pubmed.ncbi.nlm.nih.gov/33989164/)

©Jose Luis Izquierdo, Joan B Soriano. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 26.05.2021. 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 the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.