Comment on 'Prospective predictive performance comparison between clinical gestalt and validated COVID-19 mortality scores'

Héctor David Meza-Comparán

Correspondence to
Dr Héctor David Meza-Comparán, Dirección de Investigación, Instituto Nacional de Geriatría, Mexico City, Mexico; hmezacomparan@gmail.com

Accepted 6 December 2021

Dear Editor,

I read the article 'Prospective predictive performance comparison between clinical gestalt and validated COVID-19 mortality scores' with great interest.[1] The authors compared various COVID-19 mortality prediction models validated in Mexican patients — LOW-HARM, MSL-COVID-19, Nutri-CoV, and neutrophil-to-lymphocyte ratio (NLR) —, qSOFA, and NEWS2 against clinical gestalt to predict mortality among COVID-19 patients admitted to a tertiary hospital, concluding that clinical gestalt was non-inferior. I would like to comment on some issues with this article.

It is unclear what "clinical gestalt" meant in the study since no formal definition was provided by the authors other than study procedures. Others have defined clinical gestalt as "a physician's unstructured estimate"[2] or an "overall clinical impression".[3]

Additionally, it is not clear how the authors selected the prediction models to be evaluated. They mentioned that three models validated in datasets including Mexican patients were included; however, in the absence of clear inclusion criteria, other models validated in Mexican patients could have been left out. Thus, I performed a systematic search within

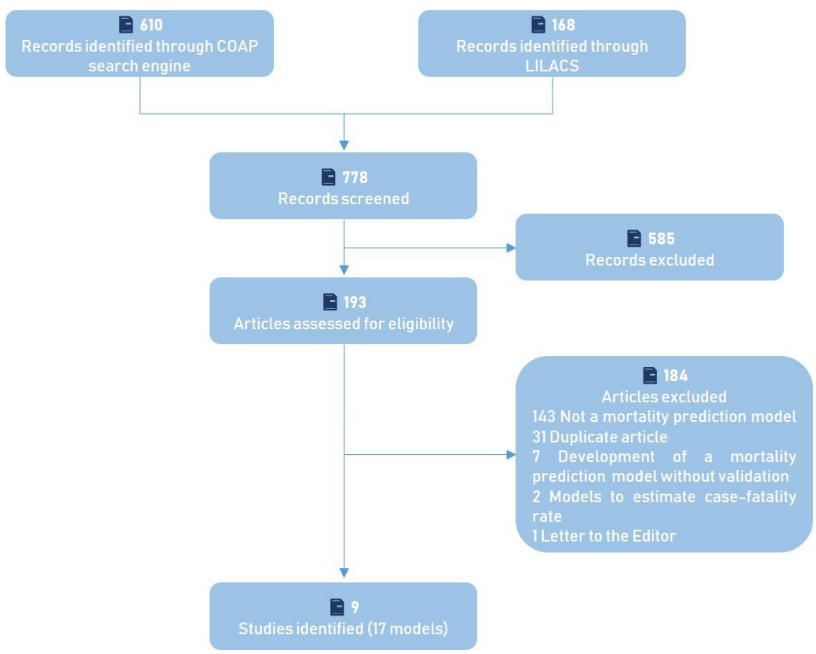

Figure 1 Systematic search flowchart of studies included and reasons for exclusion. The search within COAP[a] was performed by using the keywords and Boolean operators (mortality) AND (mexico) OR (mexican). Within LILACS[b], the keywords and Boolean operators (COVID-19) AND (mortality) AND (mexico) OR (mexican) were used; an affiliation country filter for "Mexico" was also applied in the latter case. These searches retrieved 778 records (610 and 168, respectively), of which 193 studies were retained for abstract and full-text screening. Nine studies describing 17 validated COVID-19 mortality prediction models within the Mexican population were identified. [a]COAP is a daily-updated database with SARS-CoV-2 and COVID-19 published articles from PubMed, EMBASE and PsycINFO, and preprints from medRxiv and bioRxiv (further information at https://ispmbern.github.io/covid-19/living-review/). [b]LILACS is one of the most important and comprehensive databases of scientific information in Latin America and the Caribbean with more than 880 thousand records of peer-reviewed journals, thesis and dissertations, government documents, annals of congresses and books (further information at https://lilacs.bvsalud.org/en/).

Linked

► http://dx.doi.org/10.1136/jim-2021-002037

Check for updates

To cite: Meza-Comparán HD. *J Investig Med* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/jim-2021-002243

Table 1 COVID-19 mortality prediction models validated in Mexican patients, identified through the systematic search.

Model or authors	Predictors	Reference of validation study
CALL score	Comorbidities, age, lymphocyte count, LDH	4
Charlson Comorbidity Index (CCI)	Age, MI, CHF, PVD, CVA or TIA, dementia, COPD, connective tissue disease, PUD, liver disease, DM, hemiplegia, moderate to severe CKD, solid tumor, leukemia, lymphoma, AIDS	4
HScore	Immunosuppression, body temperature, organomegaly, cytopenias, ferritin, triglycerides, fibrinogen, AST, features of hemophagocytosis in bone marrow aspirate	4
Inflammation-based risk scoring system	Albumin, hs-CRP, WBC	4
Karaismailoglu *et al*	Age, pneumonia, CKD, COPD, DM	8
Kimura-Sandoval *et al*	Percentage of total opacity >51% in non-contrast chest CT, LDH	10
LOW-HARM score	Lymphopenia, SpO_2, WBC, HTN, age, renal injury, myocardial injury	4
NLR	Absolute neutrophil count divided by absolute lymphocyte count	4
Nutri-CoV score	MSL-COVID-19 score, SpO_2, RR	7
Obesity and diabetes score (MSL-COVID-19)	Pneumonia, DM, DM and age <40 years, age ≥65 years, age <40 years, CKD, immunosuppression, COPD, obesity	4
ODL-COVID	$CD8^+$ T lymphocyte count, D-dimer, LDH, CRP, HTN, DM	5
PH-Covid19 score	Age, sex, DM, COPD, immunosuppression, HTN, obesity, CKD	4, 9
PhenoAge components	Albumin, creatinine, CRP, CA	12
PhenoAgeAccel+CA	PhenoAgeAccel value, CA	12
Quiroz-Juárez *et al*	DM, COPD, immunosuppressive drugs, HTN, CKD, CVD, obesity, presence of other chronic illnesses, sex, state of birth (Mexico), state of residence (Mexico), age, units of viral respiratory diseases (USMER) designation, sector (medical facility), state of treatment (Mexico), days symptoms-treatment, COVID-19 status, COVID-19-related pneumonia, hospitalization status, intubation, ICU	6
Wollenstein-Betech *et al*	Age, sex, immunosuppression, CKD, obesity, DM	11
Wollenstein-Betech *et al* (extended model)	Age, sex, immunosuppression, CKD, obesity, DM, hospitalization, pneumonia, need for ICU or ventilator	11

AIDS, acquired immune deficiency syndrome; AST, aspartate aminotransferase; CA, chronological age; CHF, congestive heart failure; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CRP, C-reactive protein; CT, computed tomography; CVA, cerebrovascular accident; CVD, cardiovascular diseases; DM, diabetes mellitus; hs-CRP, high-sensitivity C-reactive protein; HTN, hypertension; ICU, intensive care unit; LDH, lactate dehydrogenase; MI, myocardial infarction; NLR, neutrophil-to-lymphocyte ratio; PUD, peptic ulcer disease; PVD, peripheral vascular disease; RR, respiratory rate; SpO_2, peripheral oxygen saturation; TIA, transient ischemic attack; WBC, white blood cell count.

the COAP search engine and LILACS of studies published to November 5, 2021 (figure 1). Nine studies describing 17 validated COVID-19 mortality prediction models within the Mexican population were identified (table 1),[4–12] four of which were evaluated by Soto-Mota and colleagues (LOW-HARM, MSL-COVID-19, Nutri-CoV, and NLR).[4 7] Therefore, the authors did not evaluate a number of the important prediction models validated in Mexican patients to predict mortality.

Although the authors mentioned the median years of hospital experience (which could include medical internship and social service in Mexico) in medical residents who performed predictions, disclosing their corresponding postgraduate year (PGY) would have been important, since confidence of predictions was generally low in this study — only ~35% had >80% confidence of prediction. While they argued that "with the COVID-19 pandemic, clinicians of all levels of training started their learning curve at the same time", senior residents are less likely to be underconfident compared with junior residents.[13]

Furthermore, the statement "no score was significantly better than clinical gestalt predictions" might be questionable, due to concerns regarding sample size. An inadequate sample size could have led to the inability to detect differences, especially since the authors used *easyROC* — an open web calculator that estimates, among others, sample sizes for non-inferior ROC comparisons — to estimate the sample size for their study. Of note, *easyROC* requires an input for the "smallest difference" between tests' AUC, not the "maximal AUC difference" as the authors report. Most important is the fact that *easyROC* was not developed to estimate sample sizes to evaluate non-inferiority between prognostic predictive models; instead, it was developed to compare diagnostic test models.[14]

Finally, it is worthwhile mentioning that while in younger patients obesity is the strongest risk factor for short-term mortality,[15] chronological age remains the single most important predictor of in-hospital COVID-19 mortality.[9]

Twitter Héctor David Meza-Comparán @HectorMezaMD

Acknowledgements I would like to thank Javier Mancilla-Galindo and Ashuin Kammar-García for their invaluable comments and recommendations regarding the manuscript, as well as Vianey Fragoso-Saavedra for her support.

Contributors HDM-C assumes sole responsibility for the drafting, writing and revising of the manuscript.

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 Not commissioned; externally peer reviewed.

ORCID iD
Héctor David Meza-Comparán http://orcid.org/0000-0002-1712-8401

REFERENCES

1 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* 2021. doi:10.1136/jim-2021-002037. [Epub ahead of print: 07 Oct 2021].
2 Lucassen W, Geersing G-J, Erkens PMG, *et al*. Clinical decision rules for excluding pulmonary embolism: a meta-analysis. *Ann Intern Med* 2011;155:448.
3 Dale AP, Marchello C, Ebell MH. Clinical gestalt to diagnose pneumonia, sinusitis, and pharyngitis: a meta-analysis. *Br J Gen Pract* 2019;69:e444–53.
4 González-Flores J, García-Ávila C, Springall R, *et al*. Usefulness of easy-to-use risk scoring systems rated in the emergency department to predict major adverse outcomes in hospitalized COVID-19 patients. *J Clin Med* 2021;10:3657.
5 Elghamrawy SM, Hassanien AE, Vasilakos AV. Genetic-based adaptive momentum estimation for predicting mortality risk factors for COVID-19 patients using deep learning. *Int J Imaging Syst Technol* 2021. doi:10.1002/ima.22644. [Epub ahead of print: 13 08 2021].
6 Quiroz-Juárez MA, Torres-Gómez A, Hoyo-Ulloa I, *et al*. Identification of high-risk COVID-19 patients using machine learning. *PLoS One* 2021;16:e0257234.
7 Bello-Chavolla OY, Antonio-Villa NE, Ortiz-Brizuela E, *et al*. Validation and repurposing of the MSL-COVID-19 score for prediction of severe COVID-19 using simple clinical predictors in a triage setting: the Nutri-CoV score. *PLoS One* 2020;15:e0244051.
8 Karaismailoglu E, Karaismailoglu S. Two novel nomograms for predicting the risk of hospitalization or mortality due to COVID-19 by the naïve Bayesian classifier method. *J Med Virol* 2021;93:3194–201.
9 Mancilla-Galindo J, Vera-Zertuche JM, Navarro-Cruz AR, *et al*. Development and validation of the patient history COVID-19 (PH-Covid19) scoring system: a multivariable prediction model of death in Mexican patients with COVID-19. *Epidemiol Infect* 2020;148:e286.
10 Kimura-Sandoval Y, Arévalo-Molina ME, Cristancho-Rojas CN, *et al*. Validation of chest computed tomography artificial intelligence to determine the requirement for mechanical ventilation and risk of mortality in hospitalized coronavirus disease-19 patients in a tertiary care center in Mexico City. *Rev Invest Clin* 2021;73:2.
11 Wollenstein-Betech S, Cassandras CG, Paschalidis IC. Personalized predictive models for symptomatic COVID-19 patients using basic preconditions: hospitalizations, mortality, and the need for an ICU or ventilator. *Int J Med Inform* 2020;142:104258.
12 Márquez-Salinas A, Fermín-Martínez CA, Antonio-Villa NE, *et al*. Adaptive metabolic and inflammatory responses identified using accelerated aging metrics are linked to adverse outcomes in severe SARS-CoV-2 infection. *J Gerontol A Biol Sci Med Sci* 2021;76:e117–26.
13 Fleming M, Vautour D, McMullen M, *et al*. Examining the accuracy of residents' self-assessments and faculty assessment behaviours in anesthesiology. *Can Med Educ J* 2021;12:17–26.
14 Goksuluk D, Korkmaz S, Zararsiz G, *et al*. EasyROC: an interactive web-tool for ROC curve analysis using R language environment. *R J* 2016;8:213–30.
15 Vera-Zertuche JM, Mancilla-Galindo J, Tlalpa-Prisco M, *et al*. Obesity is a strong risk factor for short-term mortality and adverse outcomes in Mexican patients with COVID-19: a national observational study. *Epidemiol Infect* 2021;149:e109.