



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ABC₂-SPH risk score for in-hospital mortality in COVID-19 patients: development, external validation and comparison with other available scores



Milena S. Marcolino, MD, MSc, PhD^{1,2,3,*}, Magda C. Pires, PhD^{3,4,*},
 Lucas Emanuel F. Ramos⁴, Rafael T. Silva⁴, Luana M. Oliveira, MSc^{3,5},
 Rafael L.R. Carvalho, PhD³, Rodolfo Lucas S. Mourato⁶, Adrián Sánchez-Montalvá, MD,
 PhD⁷, Berta Raventós, MSc⁷, Fernando Anschau, MD, PhD⁸, José Miguel Chatkin, MD,
 PhD^{9,10}, Matheus C.A. Nogueira, MD¹¹, Milton H. Guimarães-Júnior, MD, MSc¹²,
 Giovanna G. Vietta, MD, PhD^{13,14,15}, Helena Duani, MD, PhD^{1,16}, Daniela Ponce, MD,
 PhD^{17,18}, Patricia K. Ziegelmann, PhD^{3,19}, Luís C. Castro, PhD^{20,21},
 Karen B. Ruschel, PhD^{3,22,23}, Christiane C.R. Cimini, MD, MSc^{2,3,24,25},
 Saionara C. Francisco, MSc^{26,30}, Maiara A. Floriani, MSc²⁷, Guilherme F. Nascimento, MSc²⁸,
 Bárbara L. Farace, MD²⁹, Luanna S. Monteiro, MD^{30,31}, Maira V.R. Souza-Silva, MD³⁰,
 Thais L.S. Sales, MSc⁶, Karina Paula M.P. Martins, MD, MSc^{3,16,29}, Israel J. Borges do
 Nascimento¹, Tatiani O. Fereguetti, MD³², Daniel T.M.O. Ferrara³³,

* Corresponding author

E-mail addresses: milenamarc@ufmg.br (M.S. Marcolino), magda@est.ufmg.br (M.C. Pires), luckermos19@gmail.com (L.E.F. Ramos), rafaelsilva@posteo.net (R.T. Silva), luanalmo19.09@gmail.com (L.M. Oliveira), rafaelsjdr@hotmail.com (R.L.R. Carvalho), rodolfo_use@hotmail.com (R.L.S. Mourato), adrian.sanchez.montalva@gmail.com (A. Sánchez-Montalvá), berta.raventos@vhir.org (B. Raventós), afernando@ghc.com.br (F. Anschau), jmchatkin@puhrs.br (J.M. Chatkin), mathnogueira42@gmail.com (M.C.A. Nogueira), miltonhenriques@yahoo.com.br (M.H. Guimarães-Júnior), ggvietta@gmail.com (G.G. Vietta), hduani@yahoo.com.br (H. Duani), daniela.ponce@unesp.br (D. Ponce), patriciakz99@gmail.com (P.K. Ziegelmann), pharmlucamsc@gmail.com (L.C. Castro), karenbruschel@gmail.com (K.B. Ruschel), christiane.cimini@gmail.com (C.C.R. Cimini), saionaracf@gmail.com (S.C. Francisco), maiara.floriani@hmv.org.br (M.A. Floriani), guilhermefagundesn@hotmail.com (G.F. Nascimento), barbarafarace@gmail.com (B.L. Farace), luannasmonteiro@gmail.com (L.S. Monteiro), mairavsouza@gmail.com (M.V.R. Souza-Silva), thaislorennass30@yahoo.com.br (T.L.S. Sales), kkpmp Prado2@gmail.com (K.P.M.P. Martins), israeljbn@ufmg.br (I.J. Borges do Nascimento), tatianifereguetti@gmail.com (T.O. Fereguetti), daniel@taiar.com.br (D.T.M.O. Ferrara), fbotoni@medicina.ufmg.br (F.A. Botoni), anabsetges@gmail.com (A.P.B.S. Etges), alexvspoa@gmail.com (A.V. Schwarzbald), amandaoliveira.maurilio@gmail.com (A.O. Maurílio), analuiza.bahia@yahoo.com.br (A.L.B.A. Scotton), andrepweber@gmail.com (A.P. Weber), andresmc@gmail.com (A.S.M. Costa), andressa.glaeser@hmv.org.br (A.B. Glaeser), angelicaacmadureira@gmail.com (A.A.C. Madureira), lindabhering@gmail.com (A.R. Bhering), brunocastro1199@gmail.com (B.M. Castro), carlatcas@gmail.com (C.T.C.A. Silva), carol.marques@live.com (C.M. Ramos), carolinegomes.pesquisaclinica@gmail.com (C.D. Gomes), carvalho.cintiaa@gmail.com (C.A. de Carvalho), danielvez@gmail.com (D.V. Silveira), crfsae3_hu@centrodepesquisaclinica.com.br (E. Cezar), elaynepp@yahoo.com.br (E.C. Pereira), manu.kroger@gmail.com (E.M.S. Kroger), vallt@medufmg.com (F.B. Vallt), fb_lucas@yahoo.com.br (F.B. Lucas), fgaranha@icloud.com (F.G. Aranha), fredlazzi@hotmail.com (F. Bartolazzi), gabrielapecry@gmail.com (G.P. Crestani), gisele.nader@hmv.org.br (G.A.N. Bastos), glicia.cristina.castro@gmail.com (G.C.d.C. Madeira), helenacnoal@gmail.com (H.C. Noal), hrianna@hotmail.com (H.R. Vianna), hcerqueirag@gmail.com (H.C. Guimarães), isabela_moraes1@yahoo.com.br (I.M. Gomes), srael.molina@fiocruz.br (I. Molina), joannalyra@gmail.com (J.d.L. Batista), joice-alvarenga@hotmail.com (J.C. Alvarenga), jusabatino@gmail.com (J.D.S.S. Guimarães), juliadrumond@hotmail.com (J.D.P. de Moraes), jr.machado@unesp.br (J.M. Rugolo), karen_lua3@hotmail.com (K.C.J.R. Pontes), kauane.ams@gmail.com (K.A.M. Santos), seixasleo@yahoo.com.br (L.S. Oliveira), lilian.pinheiro98@hotmail.com (L.S. Pinheiro), lilianespacheco@gmail.com (L.S. Pacheco), ld-sousa@hotmail.com (L.D. Sousa), lucianasiuves@gmail.com (L.S.F. Couto), kluciane@ghc.com.br (L. Kopitke), luis.moura@tacchini.com.br (L.C.S. Moura), luisa_elem@hotmail.com (L.E.A. Santos), madersonalvares@hotmail.com (M.A.S. Cabral), mairadiassouza@gmail.com (M.D. Souza), mgttofani@gmail.com (M.G.T. Tofani), marceloc@unisc.br (M. Carneiro), mpiferreira@hcpa.edu.br (M.A.P. Ferreira), macbicalho@gmail.com (M.A.C. Bicalho), mariacларapontellob@gmail.com (M.C.P.B. Lima), mfdgodoy@gmail.com (M.F. Godoy), marilia.cardoso@unesp.br (M.M.A. Cardoso), mpfen@yahoo.com.br (M.P. Figueiredo), natsamster@gmail.com (N.C.S. Sampaio), netx_11@hotmail.com (N.L. Rangel), nat_trifiletti@hotmail.com (N.T. Crespo), neimyramos@gmail.com (N.R. Oliveira), pedro.ledic@hmdcc.com.br (P.L. Assaf), petroniocarla@uol.com.br (P.J. Martelli), rafacharao@gmail.com (R.S.C. Almeida), rapahaelcm@gmail.com (R.C. Martins), raquellutk83@gmail.com (R. Lutkmeier), ravalacio@hotmail.com (R.A. Valacio), renanfinger@yahoo.com.br (R.G. Finger), ricardocardoso@gmail.com (R.B. Cardoso), robertapozza@tacchini.com.br (R. Pozza), xaviercampos.roberta@gmail.com (R.X. Campos), rochelemenezes@unisc.br (R.M. Menezes), roger.abreu@hmdcc.com.br (R.M. Abreu), rufino@ufsj.edu.br (R.F. Silva), smangeon@gmail.com (S.M.M. Guimarães), silviaferreiragastro@gmail.com (S.F. Araújo), susany2808@gmail.com (S.A. Pereira), talitafischeroliveira@gmail.com (T.F. Oliveira), kurtz@unisc.br (T. Kurtz), thainarastaehler@hotmail.com (T.C. Oliveira), thaizaalbino16@gmail.com (T.S.M.A. Araújo), thuliodinizmed@gmail.com (T.H.O. Diniz), veridbaldon@hotmail.com (V.B. Santos), vgvirginigomes@gmail.com (V.M.R. Gomes), vitoravale@gmail.com (V.A.L. Vale), yuri.ramires@gmail.com (Y.C. Ramires), hboersma@erasmusmc.nl (E. Boersma), carisi.anne@gmail.com (C.A. Polanczyk).

<https://doi.org/10.1016/j.ijid.2021.07.049>

1201-9712/© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Fernando A. Botoni, MD^{1,34}, Ana Paula B.S. Etges, PhD^{3,9,19}, Alexandre V. Schwarzbald, MD, PhD³⁵, Amanda O. Maurílio, MD^{36,37}, Ana Luiza B.A. Scotton, MD³⁸, André P. Weber²⁰, André S.M. Costa, MD¹¹, Andressa B. Glaeser²⁷, Angélica Aparecida C. Madureira, MD³⁹, Angelinda R. Bhering, MD²⁸, Bruno Mateus de Castro⁴⁰, Carla Thais C.A. da Silva⁴¹, Carolina M. Ramos, MD³⁴, Caroline D. Gomes²², Cíntia A. de Carvalho⁴², Daniel V. Silveira, MD, MSc²⁸, Edilson Cezar²³, Elayne C. Pereira, PhD^{13,15}, Emanuele Marianne S. Kroger, MD³⁴, Felipe B. Vallt¹, Fernanda B. Lucas, MD⁴³, Fernando G. Aranha, MD¹⁵, Frederico Bartolazzi, MD, MSc⁴³, Gabriela P. Crestani²², Gisele A.N. Bastos, MD, ScD²⁷, Glícia Cristina de C. Madeira, MD³⁹, Helena Carolina Noal, MSc³⁵, Heloisa R. Vianna, MD, MSc^{11,39,44}, Henrique C. Guimarães, MD, PhD²⁹, Isabela M. Gomes¹, Israel Molina, MD, PhD^{45,46}, Joanna d'Arc L. Batista, MD, PhD^{3,47,48}, Joice C. de Alvarenga, MD⁴², Júlia D.S.S. Guimarães, MD⁴⁹, Júlia D.P. de Moraes, MD^{39,44}, Juliana M. Rugolo¹⁸, Karen Cristina J.R. Pontes⁴⁸, Kauane Aline M. dos Santos⁴⁰, Leonardo S. de Oliveira²⁵, Lílian S. Pinheiro²⁴, Liliane S. Pacheco, MD, MSc³⁵, Lucas de D. Sousa, MD³⁸, Luciana S.F. Couto³⁵, Luciane Kopittke, PhD⁸, Luis Cesar S. de Moura, PhD⁵⁰, Luisa Elem A. Santos⁵¹, Máderson A.S. Cabral, MD^{1,16,26}, Maíra D. Souza, MD³¹, Marcela G.T. Tofani, MD⁵², Marcelo Carneiro, MD, PhD^{53,54}, Maria Angélica P. Ferreira, PhD⁴⁰, Maria Aparecida C. Bicalho, MD, PhD^{1,55}, Maria Clara P.B. Lima⁵⁶, Mariana F. Godoy, MD¹⁰, Marília M.A. Cardoso, PhD¹⁸, Meire P. Figueiredo⁴³, Natália C.S. Sampaio³², Natália L. Rangel, MD¹², Natália T. Crespo, MD¹², Neimy R. de Oliveira, MD³², Pedro L. Assaf, MD²⁶, Petrônio José de L. Martelli, MD, PhD^{57,58}, Rafaela S.C. Almeida⁸, Raphael C. Martins^{19,50}, Raquel Lutkmeier⁸, Reginaldo Aparecido Valacio³¹, Renan G. Finger, MD⁴⁸, Ricardo B. Cardoso, MSc^{3,19}, Roberta Pozza, MD, PhD⁵⁰, Roberta X. Campos, MD⁴⁹, Rochele M. Menezes, MD^{53,54}, Roger M. de Abreu, MD²⁶, Rufino de F. Silva, MD^{6,36}, Silvana M.M. Guimarães, MD⁵², Silvia F. Araújo, MD⁵², Susany Anastácia Pereira¹, Talita F. Oliveira, MD³¹, Tatiana Kurtz, MD, PhD^{53,54}, Thainara C. de Oliveira²³, Thaíza Simônia M.A. Araújo¹, Thulio Henrique O. Diniz, MD³⁶, Veridiana B. dos Santos⁸, Virginia Mara R. Gomes⁵⁹, Vitor Augusto L. do Vale⁴⁹, Yuri C. Ramires, MD²⁰, Eric Boersma, MD, PhD⁶⁰, Carisi A. Polanczyk, MD, PhD^{3,19,27}

¹ Department of Internal Medicine, Medical School, Universidade Federal de Minas Gerais. Belo Horizonte, Brazil

² Telehealth Center, University Hospital, Universidade Federal de Minas Gerais. Belo Horizonte, Brazil

³ Institute for Health Technology Assessment (IATS/ CNPq). Rua Ramiro Barcelos, 2359. Prédio 21 | Sala 507, Porto Alegre, Brazil

⁴ Department of Statistics, Universidade Federal de Minas Gerais. Belo Horizonte, Brazil

⁵ Center for Research and Graduate Studies in Business Administration, Universidade Federal de Minas Gerais. Belo Horizonte, Brazil

⁶ Universidade Federal de São João del-Rei. Divinópolis, Brazil

⁷ Infectious Diseases Department, Vall d'Hebron University Hospita, Universitat Autònoma de Barcelona, PROSICS, Barcelona, Spain

⁸ Hospital Nossa Senhora da Conceição and Hospital Cristo Redentor, Grupo Hospitalar Conceição. Porto Alegre, Brazil

⁹ Pontifícia Universidade Católica do Rio Grande do Sul (Sul RGS), Porto Alegre, Brazil

¹⁰ Hospital São Lucas PUCRS, Porto Alegre, Brazil

¹¹ Rede Mater Dei de Saúde. Belo Horizonte, Brazil

¹² Hospital Márcio Cunha. Ipatinga, Brasil

¹³ Universidade do Sul de Santa Catarina (UNISUL), Palhoça, Brazil

¹⁴ Dissertare Scientific Advice, Florianópolis, Brazil

¹⁵ SOS Cardio Hospital, Florianópolis, Brazil

¹⁶ Internal Medicine Department, University Hospital, Universidade Federal de Minas Gerais. Belo Horizonte, Brazil

¹⁷ Faculdade de Medicina de Botucatu, Universidade Estadual Paulista "Júlio de Mesquita Filho". Botucatu, Brazil

¹⁸ Hospital das Clínicas da Faculdade de Medicina de Botucatu. Botucatu, Brazil

¹⁹ Universidade Federal do Rio Grande do Sul. Porto Alegre, Brazil

²⁰ Hospital Bruno Born. Lajeado, Brazil

²¹ Research Center of Vale do Taquari. Estrela, Brazil

²² Hospital Mãe de Deus. Porto Alegre, Brazil

²³ Hospital Universitário de Canoas. Canoas, Brazil

²⁴ Mucuri Medical School, FAMMUC, Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM. Teófilo Otoni, Brazil

²⁵ Hospital Santa Rosália. Teófilo Otoni, Brazil

²⁶ Hospital Metropolitanou Doutor Célio de Castro. Belo Horizonte, Brazil

²⁷ Hospital Moinhos de Vento. Porto Alegre, Brazil

²⁸ Hospital Unimed BH. Belo Horizonte, Brazil

²⁹ Hospital Risoleta Tolentino Neves. Belo Horizonte, Brazil

³⁰ Post-graduation Center, Medical School, Universidade Federal de Minas Gerais. Belo Horizonte, Brazil

³¹ Hospital Metropolitanou Odilon Behrens. Belo Horizonte, Brazil

³² Hospital Eduardo de Menezes. Belo Horizonte. Brazil

- ³³ Universidade FUMEC. Belo Horizonte, Brazil
³⁴ Hospital Julia Kubitschek. Belo Horizonte, Brazil
³⁵ Hospital Universitário de Santa Maria. Santa Maria, Brazil
³⁶ Departamento de Medicina Interna, Universidade Federal de Santa Maria. Santa Maria, Brazil
³⁷ Hospital São João de Deus. Divinópolis, Brazil
³⁸ Hospital Regional Antônio Dias. Patos de Minas, Brazil
³⁹ Faculdade Ciências Médicas de Minas Gerais. Belo Horizonte, Brazil
⁴⁰ Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul. Porto Alegre, Brazil
⁴¹ Faculdade de Ciências Humanas de Curvelo. Curvelo, Brazil
⁴² Hospital João XXIII. Av. Professor Alfredo Balena, 400, Belo Horizonte, Brazil
⁴³ Hospital Santo Antônio. Curvelo, Brazil
⁴⁴ Hospital Universitário Ciências Médicas. Belo Horizonte, Brazil
⁴⁵ Vall d'Hebron University Hospital, PROSICS Barcelona. Barcelona, Spain
⁴⁶ Instituto René Rachou-FIOCRUZ Minas. Belo Horizonte, Brazil
⁴⁷ Universidade Federal da Fronteira Sul, Chapecó, Brazil
⁴⁸ Hospital Regional do Oeste. Chapecó, Brazil
⁴⁹ Pontifícia Universidade Católica de Minas Gerais. Belo Horizonte, Brazil
⁵⁰ Hospital Tacchini. Bento Gonçalves, Brazil
⁵¹ Centro Universitário de Patos de Minas. RPatos de Minas, Brazil
⁵² Hospital Semper. Belo Horizonte, Brazil
⁵³ Hospital Santa Cruz. Santa Cruz do Sul, Brazil
⁵⁴ Universidade de Santa Cruz - Santa Cruz do Sul
⁵⁵ Fundação Hospitalar do Estado de Minas Gerais – FHEMIG. Belo Horizonte, Brazil
⁵⁶ Universidade Federal de Ouro Preto. Ouro Preto, Brazil
⁵⁷ Hospital das Clínicas da Universidade Federal de Pernambuco. Recife, Brazil
⁵⁸ Universidade Federal de Pernambuco, Centro de Ciências Médicas. Recife, Brazil
⁵⁹ Centro Universitário de Belo Horizonte (UniBH). Belo Horizonte, Brazil
⁶⁰ Erasmus MC, University Medical Center Rotterdam, Department of Cardiology, Rotterdam, The Netherlands

ARTICLE INFO

Article history:

Received 11 June 2021
 Revised 18 July 2021
 Accepted 20 July 2021

Keywords:

COVID-19
 SARS-CoV-2
 Mortality
 Prognosis
 Risk factors
 Hospitalizations
 Score

ABSTRACT

Objectives: The majority of available scores to assess mortality risk of coronavirus disease 2019 (COVID-19) patients in the emergency department have high risk of bias. Therefore, this cohort aimed to develop and validate a score at hospital admission for predicting in-hospital mortality in COVID-19 patients and to compare this score with other existing ones.

Methods: Consecutive patients (≥ 18 years) with confirmed COVID-19 admitted to the participating hospitals were included. Logistic regression analysis was performed to develop a prediction model for in-hospital mortality, based on the 3978 patients admitted between March–July, 2020. The model was validated in the 1054 patients admitted during August–September, as well as in an external cohort of 474 Spanish patients.

Results: Median (25–75th percentile) age of the model-derivation cohort was 60 (48–72) years, and in-hospital mortality was 20.3%. The validation cohorts had similar age distribution and in-hospital mortality. Seven significant variables were included in the risk score: age, blood urea nitrogen, number of comorbidities, C-reactive protein, SpO₂/FiO₂ ratio, platelet count, and heart rate. The model had high discriminatory value (AUROC 0.844, 95% CI 0.829–0.859), which was confirmed in the Brazilian (0.859 [95% CI 0.833–0.885]) and Spanish (0.894 [95% CI 0.870–0.919]) validation cohorts, and displayed better discrimination ability than other existing scores. It is implemented in a freely available online risk calculator (<https://abc2sph.com/>).

Conclusions: An easy-to-use rapid scoring system based on characteristics of COVID-19 patients commonly available at hospital presentation was designed and validated for early stratification of in-hospital mortality risk of patients with COVID-19.

© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Introduction

Coronavirus disease 2019 (COVID-19) is still the main global health, social and economic challenge, overwhelming healthcare systems in many countries and heavily burdening others (WHO 2020, Emanuel et al., 2020). Case rates continue to rise, and some hospitals are nearly at their full capacity of intensive care unit (ICU) beds. The emergence of new variants of SARS-CoV-2 in the UK, South Africa, Brazil and India is currently a cause of huge concern – with very high viral growth, being more transmissible, less detectable with the RT-PCR technique, or deadlier than the wild-type SARS-CoV-2, and with evidence of lower vaccine efficacy (Conti et al., 2021, Zhang, 2021, Faria et al., 2021, Rubin, 2021).

To empower early identification and intervention of patients at higher risk of poor outcome, fast and efficient assessment of prognosis of the disease is needed to optimize the allocation of healthcare and human resources. A proper assessment tool will guide decision-making to develop an appropriate plan of care for each patient (Zhang et al., 2020). In this context, rapid scoring systems, which combine different variables to estimate the risk of a poor outcome, may be extremely helpful for quick and effective assessment of those patients in the emergency department (Leeuwenberg and Schuit, 2020).

Although different scores have been proposed to assess prognosis in COVID-19 patients, the majority of them lack benefit to clinical decision-making, and there is a lack of reliable prognostic prediction models (Fumagalli et al., 2020, Gupta et al., 2020). Most scores were developed from small cohorts at high risk of bias, with selected study samples and relatively few outcome events, without

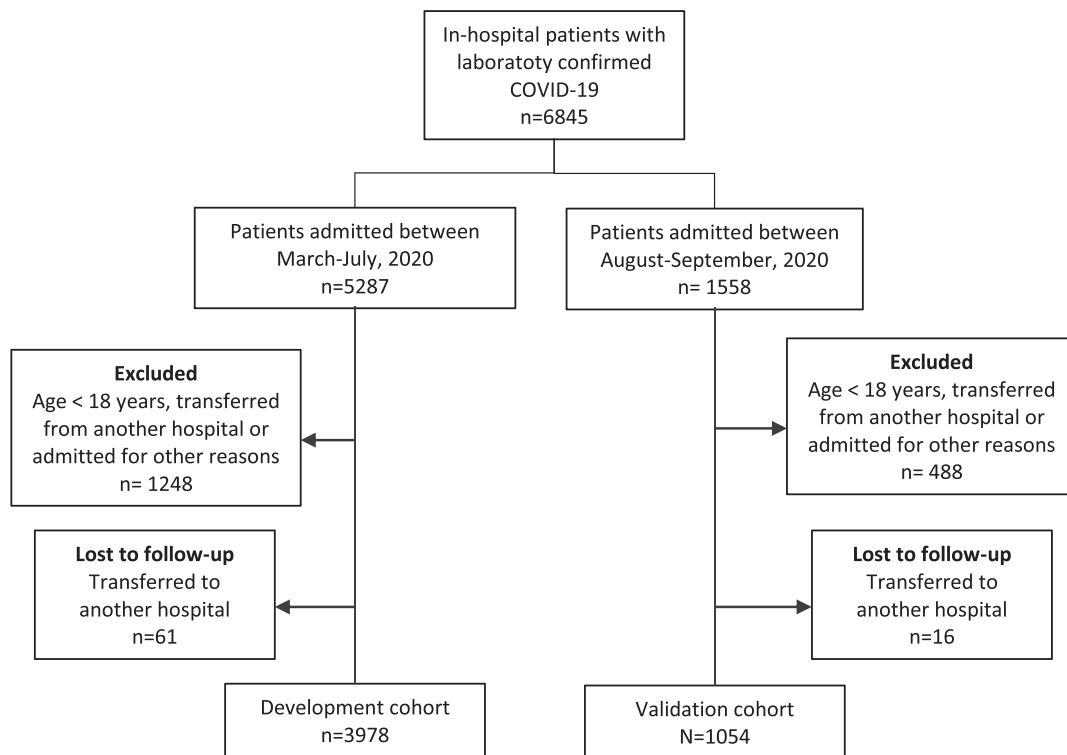


Figure 1. Flowchart of COVID-19 patients included in the study.

clear details of model derivation and validation, and unclear reporting on intended use (Wynants et al., 2020, Wang et al., 2020, Allenbach et al., 2020, Kim et al., 2020, Zhou et al., 2020). These issues have led to a high risk of model overfitting, thus their predictive performance when used in clinical practice may be different to that reported (Gupta et al., 2020, Wynants et al., 2020) and external validation has rarely been performed (Goel et al., 2020, Gupta et al., 2020).

In this context, this cohort aimed to develop and validate an easily applicable score that employs routinely available clinical and laboratory data at hospital presentation to predict in-hospital mortality in patients with COVID-19, and able to discriminate high-risk vs non-high-risk patients. Additionally, it aimed to compare this score with other existing ones.

Methods

This study is part of the Brazilian COVID-19 Registry, an ongoing multicenter observational study described elsewhere (Marcolino et al., 2021), and a collaboration with Vall d'Hebron University Hospital, in Barcelona, Spain, for independent external validation. The Brazilian COVID-19 Registry is being conducted according to a predefined protocol, in 36 Brazilian hospitals, located in 17 cities, from five Brazilian states. Model development, validation and reporting followed guidance from the Transparent Reporting of a Multivariable Prediction Model for Individual Prediction or Diagnosis (TRIPOD) checklist and Prediction model Risk Of Bias Assessment Tool (PROBAST) (Supplementary Material Tables S6 and S7) (Moons et al., 2015, Wolff et al., 2019).

Study subjects

Consecutive adult patients (aged ≥ 18 years) with laboratory-confirmed COVID-19 (Organization, 2020) admitted to the participating hospitals from 01 March 1 to 30 September 2020 were enrolled. Patients who were transferred between hospitals, and ad-

mission data from the first hospital (as it was aimed to develop a score to be used at the first assessment) or the last hospital that were unavailable were excluded (Table S1), as well those who were admitted for other reasons and developed COVID-19 symptoms during their stay (Figure 1).

Measurement and data quality assessment

Demographic information, clinical characteristics, laboratory and outcome data were collected from the medical records by using a prespecified case report form applying Research Electronic Data Capture (REDCap) tools (Harris et al., 2009, Harris et al., 2019). Data were collected by trained hospital staff or interns. A detailed data management plan (DMP) was developed and provided to all participating centers, and online DMP training was mandatory before local research personnel were allowed to start collecting study data (Gregory and Radovinsky, 2012). Comprehensive data quality checks were undertaken to ensure high quality. A code was developed in R software to identify values likely related to data entry errors for vital signs and laboratory variables, based on expert-guided rules. Data were sent to each center for checking and correction.

Potential predictors for in-hospital mortality

All variables used to calculate the risk score were obtained at hospital admission. A set of potential predictor variables for in-hospital mortality was selected a priori, as recommended (Wolff et al., 2019), taking into account the evidence in literature of association with worse prognosis in patients with COVID-19 or pneumonia, and availability of predictor measurement at the time the model would be used (i.e., hospital admission) (Wynants et al., 2020). All laboratory tests were performed at the discretion of the treating physician. Imaging test results were not included as X-rays and CT scans are not always performed at patient admission and

their interpretation involves subjective judgement. Candidate predictor variables that were unavailable for at least two-thirds of patients within the derivation cohort (more than one-third of missing data) were excluded.

Data analysis

Continuous variables were summarized using medians and interquartile ranges (IQR), whereas counts and percentages were used for categorical variables. This study reported 95% confidence intervals, and a p -value < 0.05 was considered statistically significant. Statistical analysis was performed with R software (version 4.0.2) with the *mgcv*, *finalfit*, *mice*, *glmnet*, *pROC*, *rms*, *rmda*, and *psfmi* packages.

Missing data

Considering missing at random after analyzing missing data patterns, multiple imputation with chained equations (MICE) was used to handle missing values on candidate variables (outcomes were not imputed). Mortality outcome was used as a predictor in MICE in the derivation dataset, but not in the validation dataset. The predictive mean matching (PMM) method was used for continuous predictors and polytomous regression for categorical variables (two or more unordered levels). The results of 10 imputed datasets, each with 10 iterations, were combined following Rubin's rules (Rubin, 2004).

Development of the risk score model

Patients who were admitted before 31 July were included in the development cohort (Figure S1). Predictor selection was conducted based on clinical reasoning and literature review before modeling. Generalized additive models (GAM) were used to examine the relationships between in-hospital mortality and continuous (through penalized thin plate splines) and categorical (as linear components) predictors. During this stage, variable selection was based on D1-statistic (multivariate Wald test) and D2-statistic (pools test statistics from repeated analyses). Subsequently, for easier application of the risk score model at the bedside, continuous variables were categorized based on widely accepted cut points, current evidence and/or categories defined in established scores for pneumonia and sepsis. Lastly, least absolute shrinkage and selection operator (LASSO) logistic regression were used to derive the mortality score by scaling the (L1 penalized) shrunk coefficients. The penalty parameter λ in LASSO was chosen using 10-fold cross-validation methods based on mean squared error criterion. Risk groups were proposed based on predicted probabilities: low risk ($< 6.0\%$), intermediate risk (6.0 – 14.9%), high risk (15.0 – 49.9%), and very high risk ($\geq 50.0\%$), as recommended by TRIPOD Guidelines (Moons et al., 2015). Risk strata was based on other scores (Knight et al., 2020).

External validation

An external (temporal) validation analysis was performed using patients who were admitted from 01 August to 30 September 2020. Independent external validation was also performed in a cohort of patients from Vall d'Hebron University Hospital, in Barcelona, Spain, admitted from 01 March to 31 May 2020 (Montalva et al., 2020). Inclusion and exclusion criteria were the same as the previously mentioned ones. All included patients were followed for at least 28 days. The aforementioned methods for data imputation were used.

Performance measures

Overall performance was evaluated using the Brier score (Rufibach, 2010). Calibration was graphically assessed by plotting the predicted mortality probabilities against the observed mortality, testing intercept equals zero and slope equals one. The area under the receiver operating characteristic curve (AUROC) described the model's discrimination. Confidence intervals (95% CI) for AUROC were obtained through 2000 bootstrap samples. Positive and negative predictive values of the derived risk groups were calculated.

Model comparisons

The developed model was compared within the validation cohort with existing scores. These scores were identified through a literature search of Medline, medRxiv and BioRxiv, with no language or date restrictions, using the search terms "COVID-19", "COVID", "SARS-CoV-2", "coronavirus" combined with "score" and "mortality". The last search was performed on 19 November 2020. Two authors independently performed article selection and data extraction. Additionally, established scores for pneumonia and sepsis were included (Lim et al., 2003, Liu et al., 2016, Fine et al., 1997, Cag et al., 2018, Olsson et al., 2004). From the set of identified scores, those which with predictors were available within the database and had accessible methods for calculation were selected. Model comparisons were performed using AUROC and decision curve analysis.

ABC₂-SPH risk score calculator

The risk score calculator was developed in Javascript, using the Svelte framework, while the website was developed in R language (blogdown package).

Results

The derivation cohort included 3978 patients from 267 cities in Brazil (Figure 2). Median age was 60 [IQR, 48–72] years, 53.8% were male, 70.1% had at least one comorbidity, and 20.3% died during hospitalization. Table 1 shows demographic, clinical characteristics and laboratory findings for the derivation and validation datasets.

Development of the risk score model

Thirty-six potential predictor variables were identified (Table S2). The number of comorbidities was created as a composite of 10 individual comorbidities shown to have prognostic impact in COVID-9 – hypertension, diabetes mellitus, obesity, coronary artery disease, heart failure, atrial fibrillation or flutter, cirrhosis, chronic obstructive pulmonary disease, cancer, and previous stroke (Harrison et al., 2020, Harrison et al., 2020) – as in other scores (Knight et al., 2020, Tuty Kuswardhani et al., 2020). Twelve variables were excluded due to the excessive number of missing values, two for high collinearity, and one was not recorded within the database. Inotrope use was combined with blood pressure. Therefore, 20 variables were tested.

Through a generalized additive model (GAM), a combination of seven variables was selected as the best predictor of in-hospital mortality (Table S3). For an easier application to the risk score model at bedside, continuous selected predictors were categorized for LASSO logistic regression. All categories were defined a priori,

Table 1
Demographic and clinical characteristics for derivation and validation cohorts of patients admitted to hospital with COVID-19.

Characteristics	Derivation cohort (n=3978)		Brazilian validation cohort (n=1054)		Spanish validation cohort (n=856)	
	Frequency (%) or median (IQR)	Non-missing cases (%)	Frequency (%) or median (IQR)	Non-missing cases (%)	Frequency (%) or median (IQR)	Non-missing cases (%)
In-hospital mortality	806 (20.3%)	3978 (100%)	208 (19.7%)	1054 (100%)	172 (20.1%)	856 (100%)
Age (years)	60.0 (48.0, 72.0)	3978 (100%)	62.0 (48.2, 73.0)	1054 (100%)	62.0 (49.0, 74.0)	856 (100%)
Sex at birth		3976 (99.9%)		1054 (100%)		856 (100%)
Male	2138 (53.8%)		582 (55.2%)		498 (58.2%)	
Comorbidities						
Hypertension	2147 (54.0%)	3978 (100%)	563 (53.4%)	1054 (100%)	377 (44.0%)	856 (100%)
Coronary artery disease	215 (5.4%)	3978 (100%)	60 (5.7%)	1054 (100%)	65 (7.6%)	856 (100%)
Heart failure	269 (6.8%)	3978 (100%)	58 (5.5%)	1054 (100%)	46 (5.4%)	856 (100%)
Atrial fibrillation or flutter	139 (3.5%)	3978 (100%)	27 (2.6%)	1054 (100%)	89 (10.4%)	856 (100%)
Stroke	146 (3.7%)	3978 (100%)	43 (4.1%)	1054 (100%)	42 (4.9%)	856 (100%)
COPD	253 (6.4%)	3978 (100%)	60 (5.7%)	1054 (100%)	48 (5.6%)	856 (100%)
Diabetes mellitus	1151 (28.9%)	3978 (100%)	297 (28.2%)	1054 (100%)	151 (17.6%)	856 (100%)
Obesity (BMI > 30 kg/m ²)	696 (17.5%)	3978 (100%)	181 (17.2%)	1054 (100%)	202 (23.6%)	856 (100%)
Cirrhosis	25 (0.6%)	3978 (100%)	9 (0.9%)	1054 (100%)	5 (0.6%)	856 (100%)
Cancer	194 (4.9%)	3978 (100%)	65 (6.2%)	1054 (100%)	35 (4.1%)	856 (100%)
Number of comorbidities		3978 (100%)		1054 (100%)		856 (100%)
0	1189 (29.9%)		309 (29.3%)		325 (38.0%)	
1	1173 (29.5%)		328 (31.1%)		222 (25.9%)	
2	1013 (25.5%)		269 (25.5%)		167 (19.5%)	
3	429 (10.8%)		106 (10.1%)		91 (10.6%)	
4	131 (3.3%)		33 (3.1%)		34 (4.0%)	
≥ 5	43 (1.1%)		9 (0.9%)		17 (2.0%)	
Clinical assessment at admission						
SF ratio	428.6 (332.1, 452.4)	3845 (96.7%)	433.3 (339.3, 452.4)	1034 (98.1%)	457.1 (423.8, 466.5)	842 (98.4%)
Respiratory rate (irpm)	20 (18, 24)	3236 (81.3%)	20 (18, 24)	870 (82.5%)	20 (18, 26)	452 (92.3%)
Heart rate (bpm)	88 (78, 100)	3787 (95.2%)	88 (77, 100)	1020 (96.8%)	93 (80, 105)	842 (98.4%)
Glasgow coma score	15 (15, 15)	3695 (92.9%)	15 (15, 15)	982 (93.2%)	15 (15, 15)	838 (97.9%)
Systolic blood pressure		3762 (94.6%)		1014 (96.2%)		843 (98.5%)
≥ 90 (mm Hg)	3076 (81.8%)		825 (81.4%)		829 (98.3%)	
< 90 (mm Hg)	510 (13.6%)		146 (14.4%)		14 (1.7%)	
Inotrope requirement	176 (4.7%)		43 (4.2%)		0	
Diastolic blood pressure		3776 (94.9%)		1022 (97.0%)		842 (98.4%)
> 60 (mm Hg)	3541 (93.8%)		962 (94.1%)		712 (84.6%)	
≤ 60 (mm Hg)	59 (1.6%)		17 (1.7%)		130 (15.4%)	
Inotrope requirement	176 (4.7%)		43 (4.2%)		0	
Laboratory parameters						
Hemoglobin (g/L)	13.3 (12.1, 14.4)	3871 (97.3%)	13.3 (11.9, 14.5)	1021 (96.9%)	13.4 (12.2, 14.6)	851 (99.4%)
Platelet count (10 ⁹ /L)	196.0 (154.0, 257.0)	3824 (96.1%)	203.0 (154.0, 260.2)	1016 (96.4%)	188.0 (149.0, 243.0)	851 (99.4%)
NLR	4.7 (2.8, 7.8)	3759 (94.5%)	4.9 (3.0, 8.4)	989 (93.8%)	NA	NA
Lactate (mmol/L)	1.4 (1.1, 1.9)	2,742 (68.9%)	1.5 (1.2, 2.1)	720 (68.3%)	NA	NA
C-reactive protein (mg/L)	77.0 (38.0, 143.0)	3487 (87.7%)	74.1 (33.8, 143.0)	881 (83.6%)	98.4 (41.8, 186.3)	581 (67.9%)
BUN (mg/dL)	16.3 (11.5, 24.3)	3636 (91.4%)	17.3 (12.9, 25.2)	942 (89.4%)	16.8 (12.1, 25.2)	653 (76.3%)
Creatinine (mg/dL)	0.9 (0.8, 1.2)	3765 (94.6%)	1.0 (0.8, 1.3)	967 (91.7%)	0.8 (0.7, 1.1)	847 (98.9%)
Sodium (mmol/L)	137.0 (135.0, 140.0)	3550 (89.2%)	137.0 (134.3, 140.0)	930 (88.2%)	136.1 (134.0, 138.0)	846 (98.8%)
Bicarbonate (mEq/L)	23.0 (21.0, 25.0)	3222 (81.0%)	23.0 (20.6, 25.0)	807 (76.6%)	NA	NA
pH	7.4 (7.4, 7.5)	3232 (81.2%)	7.4 (7.4, 7.5)	808 (76.7%)	NA	NA
pO ₂ (mmHg)	75.0 (63.0, 96.0)	3183 (80.0%)	73.4 (63.0, 94.6)	800 (75.9%)	NA	NA
pCO ₂ (mmHg)	35.0 (31.3, 39.0)	3194 (80.3%)	34.0 (30.0, 38.0)	801 (76.0%)	NA	NA

BMI: body mass index; BUN: blood urea nitrogen; COPD: chronic obstructive pulmonary disease; NA: not available; NLR: neutrophils-to-lymphocytes ratio; SF ratio: SpO₂/FiO₂ ratio

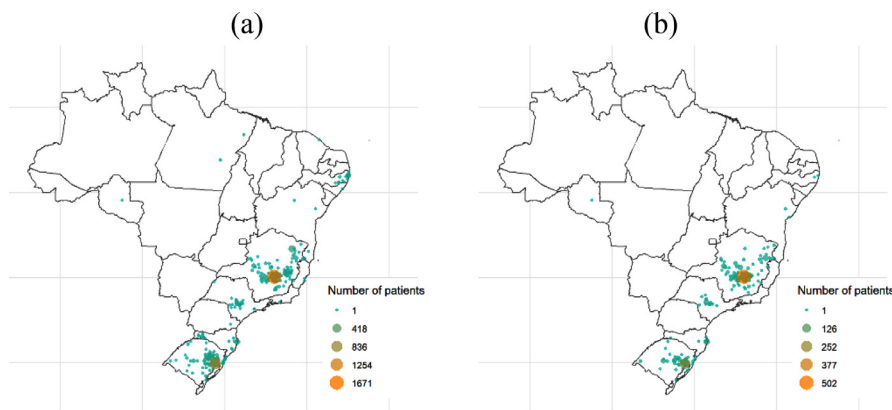


Figure 2. City of residence of patients within (a) development and (b) validation cohorts.

Table 2
ABC₂-SPH score for in-hospital mortality in patients with COVID-19.

Variable	ABC ₂ -SPH score
A Age (years)	
< 60	0
60 – 69	1
70 – 79	3
≥ 80	5
B Blood urea nitrogen (mg/dL)*	
< 42	0
≥ 42	3
C₂ Comorbidities	
0 – 1	0
≥ 2	1
C -reactive protein (mg/L)	
< 100	0
≥ 100	1
S SF ratio (%)	
> 315	0
> 235 – 315	1
> 150 – 235	3
≤ 150	6
P Platelet count (x10 ⁹ /L)	
> 150	0
100 – 150	1
< 100	2
H Heart rate (bpm)	
≤ 90	0
91 – 130	1
≥ 131	2

* When converted to urea, the cut-off is 90 mg/dL

based on widely accepted cut points, current evidence and/or categories defined in established rapid scoring systems from pneumonia and sepsis (Wolff et al., 2019).

All variables were statistically significant predictors for in-hospital mortality (Table S4 and Figure S2). Shrunken coefficients were scaled to provide a prognostic index, which was denoted as the ABC₂-SPH risk score (Table 2). The sum of the prediction scores ranged 0–20, with a high score indicating higher risk of in-hospital mortality. Risk groups were proposed based on predicted probabilities (Table 3): low risk (0–1 score, observed in hospital mortality 2.0%), intermediate risk (2–4 score, 11.4%), high risk (5–8 score, 32.0%), and very high risk (≥ 9 score, 69.4%). Subject-specific risks can be assessed using the developed ABC₂-SPH risk score web-based calculator (<https://abc2sph.com/>), which is freely available to the public, and it can also be assessed through infographics (Figure S3). As well as GAM and LASSO, the ABC₂-SPH risk score showed good overall performance (Brier score: 0.114) and good discrimination (AUROC equal 0.842; 95% CI 0.840–0.843) within the derivation cohort (Table 4).

External validation – Brazilian cohort

A total of 1054 patients were included in the validation cohort. Median age was 62 (IQR 48–73) years, 55.2% were male, 70.7% had at least one comorbidity, and 19.7% died during hospitalization. The distribution of patients across the range of ABC₂-SPH scores in derivation and validation cohorts is presented in Figure 3. It was observed that the ABC₂-SPH risk score had an AUROC of 0.859 (95% CI 0.833 to 0.885), good overall performance (Brier = 0.108) and calibration (slope = 1.138, intercept = 0.114, p-value = 0.184; Figure S4a) under the validation cohort (Figure 4). A good performance was also demonstrated in sensitivity analyses using complete case data (Table S5). Low-risk, intermediate-risk and high-risk groups showed good negative predictive values (99.7%, 88.1% and 71.0%, respectively). A positive predictive value of 73.7% was observed in patients classified as very high mortality risk.

External validation – Spanish cohort

A second external (geographic) validation was performed within a Spanish cohort with 856 patients and 172 (20.1%) in-hospital mortalities. The demographic and clinical characteristics at admission are listed in Table 1. The median follow-up time was 21 (IQR, 7–40) days. The ABC₂-SPH score showed AUROC = 0.894 (95% CI 0.870 to 0.919; Figure 4) and good overall performance (Brier = 0.093). Calibration is shown in Figure S4b.

Literature review

The literature search identified 39 scores to predict mortality in COVID-19 patients (Table 5). In 36% the derivation cohort was from China, 21% from the United States and none from South America.

Comparison with other scores

Based on a complete case validation cohort, the ABC₂-SPH score achieved better discrimination (Table 6, Figure 5a) than other prediction scoring systems for COVID-19, pneumonia and sepsis (0.85; 95% CI 0.82 – 0.88). Xie’s and Zhang’s score (Zhang et al., 2020, Xie et al., 2020, Zhang et al., 2020) showed good discrimination, but the number of complete cases and deaths were relatively small. Considering clinical utility (Figure 5b), the ABC₂-SPH score showed a better performance compared to the three most discriminating scores for in-hospital mortality that were tested in more than 600 patients (4C Mortality Score, A-DROP and CURB-65) (Knight et al., 2020, Liu et al., 2016).

Table 3
Predicted mortality and mortality rates for ABC₂-SPH score risk groups.

Risk group	Predicted mortality	Derivation cohort		Validation cohort	
		No. of patients	No. of deaths (%)	No. of patients	No. of deaths (%)
Low (0–1)	< 6%	1133	23 (2.0%)	290	1 (0.3%)
Intermediate (2–4)	6 – 14.9%	1470	168 (11.4%)	394	47 (11.9%)
High (5–8)	15 – 49.9%	907	290 (32.0%)	252	73 (29.0%)
Very high (≥ 9)	≥ 50%	468	325 (69.4%)	118	87 (73.7%)
Overall	–	3978	806 (20.3%)	1054	208 (19.7%)

Table 4
Discrimination and model overall performance in derivation and validation cohorts.

Model	Derivation cohort		Brazilian validation cohort	
	AUROC (95% CI)	Brier score	AUROC (95% CI)	Brier score
GAM	0.884 (0.879; 0.888)	0.101	0.871 (0.862; 0.879)	0.102
LASSO	0.844 (0.842; 0.846)	0.115	0.859 (0.855; 0.862)	0.110
ABC ₂ -SPH	0.842 (0.840; 0.843)	0.114	0.857 (0.854; 0.860)	0.108

GAM: generalized additive models; LASSO: least absolute shrinkage and selection operator logistic regression

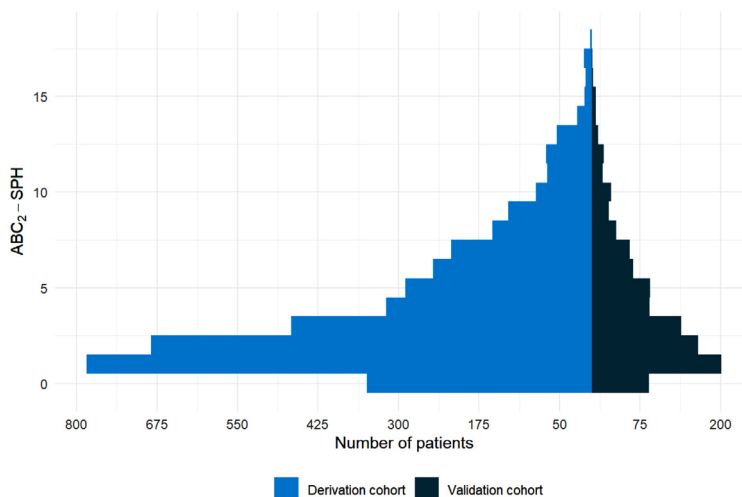


Figure 3. ABC₂-SPH score in derivation and validation cohorts.

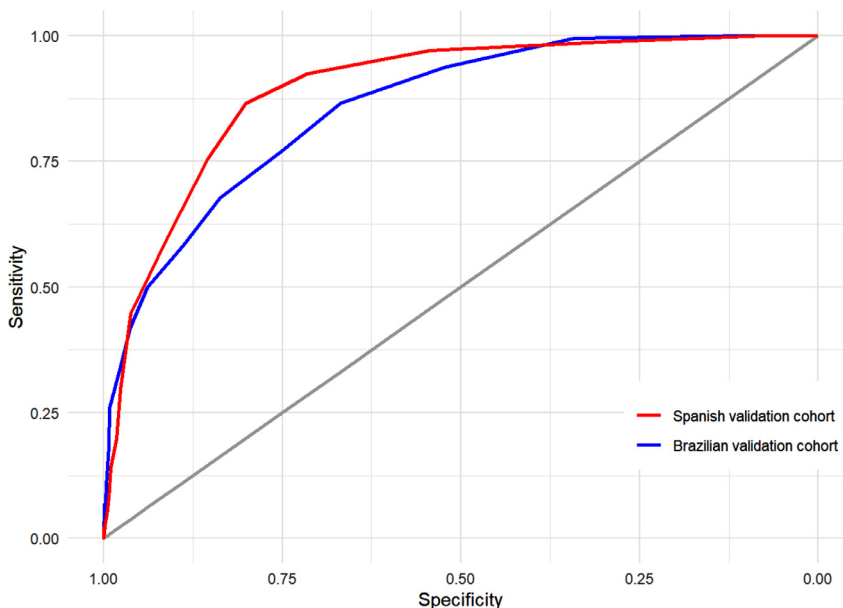


Figure 4. Discrimination of ABC₂-SPH score in external validation cohorts.

Table 5
Main characteristics of the studies.

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Halalau (Halalau et al., 2021)	Retrospective cohort	March 1, 2020 to April 1, 2020	USA	USA	2025	Not clear	1290	Not clear	Confirmed SARS-CoV-2 patients who required hospital admission at 8 hospitals in Beamount, excluding patients who remained hospitalized beyond May 12, 2020
Fumagalli (Fumagalli et al., 2020)	Retrospective cohort	February 22, 2020 to April 10, 2020	Italy	Italy	516	516	NA	Consecutive adult patients with COVID-19 from 2 Italian tertiary hospitals	
Knight (Knight et al., 2020)	Prospective cohort	May 21, 2020 to June, 29 2020	England, Scotland, and Wales	England, Scotland, and Wales	57824	35463	22361	Consecutive adult patients with COVID-19 from 260 hospitals, admitted up to May 20, 2020	The same as the development population, admitted after May 20, 2020
Liang (Liang et al., 2020)	Retrospective cohort	November 21, 2019 to January 31, 2020	China	China	2300	1590	710	Patients with COVID-19 from 575 hospitals in 31 provincial administrative regions	Data from hospitals not included in the development cohort
Nicholson (Nicholson et al., 2020)	Retrospective cohort	First patient to May 19, 2020	USA	USA	1042	578	464	Consecutive adult patients with laboratory-confirmed COVID-19 patients from Mass General Brigham hospitals	
Garibaldi (Garibaldi et al., 2021)	Retrospective cohort	March 4, 2020 to April 24, 2020, with follow-up through June 27, 2020	USA	USA	832	832	NA	Consecutive confirmed COVID-19 patients from 5 hospitals (John Hopkins Medicine)	
Sourij (Sourij et al., 2020)	Prospective and retrospective cohort	April 15, 2020 to June 30, 2020	Austria	NA	238	238	NA	Adult patients with confirmed COVID-19 and diabetes or pre-diabetes	NA
Gavelli (Gavelli et al., 2020)	Retrospective single-center cohort	March 16, 2020 to April 22, 2020	Italy	Italy	480	Apparently, it was developed by expert consensus	480	NA	Adult patients with confirmed COVID-19 patients admitted to one university hospital
Kazemi (Kazemi et al., 2020)	Retrospective cohort	February 25, 2020 to April 25, 2020	Iran	NA	91	91	NA	Adult patients with confirmed COVID-19 who had undergone CT scan < 8 days from the beginning of symptoms, excluding the ones with RT-PCR > 7 days from CT. CT score developed not based on the data. Authors tested CT score and clinical variables in a model	NA

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Núñez-Gil (Nunez-Gil et al., 2020)	Retrospective cohort	February 8, 2020 to April 1, 2020	Spain and Italy	NA	908	908	NA	Patients with confirmed COVID-19 from centers in Italy (n=88) and Spain (n=820)	
Allenbach (Allenbach et al., 2020)	Prospective single-center cohort	March 16, 2020 to April 4, 2020	France	France	152	152	131	Adult patients with confirmed COVID-19 from one tertiary care university hospital	Not described
Kim (Kim et al., 2020)	Retrospective single-center cohort	February 19, 2020 to March 15, 2020	Korea	NA	38	38	NA	Adult patients with confirmed COVID-19 admitted to a tertiary university hospital	NA
Altschul (Altschul et al., 2020)	Retrospective single-center cohort	March 1, 2020 to April 16, 2020	USA	USA	4711	2355	2356	Patients with confirmed COVID-19 from an academic hospital	The same as the development population (split 50/50%, apparently by admission date)
Hajifathalian (Hajifathalian et al., 2020)	Retrospective cohort	March 4, 2020 to April 9, 2020	USA	USA	929	664	265	Adult patients with confirmed COVID-19 patients presenting to emergency department of 2 hospitals in Manhattan (did not exclude patients who were discharged within 24 hours)	Adult patients with confirmed COVID-19 patients presenting to emergency department of 9 hospitals in Massachusetts (did not exclude patients who were discharged within 24 hours)
Wang (Wang et al., 2020)	Retrospective single-center cohort	January 28, 2020 to March 4, 2020	China	China	243	199	44	Adult patients with confirmed COVID-19 from one university hospital	The same as the development population (the criteria used to divide patients in training and testing sets was not clear)
Zhou (Zhou et al., 2020)	Retrospective single-center cohort	January 12, 2020 to February 26, 2020	China	NA	118	118	NA	Elderly patients (> 60 years) with "clinically diagnosed" COVID-19 (RT-PCR or chest CT) from one university hospital	NA
Goméz (Gomez et al., 2020)	Retrospective single-center cohort	February 24, 2020 to March 16, 2020	Spain	NA	163	163	NA	Adult patients with suspected COVID-19 admitted to one university hospital	NA
Galloway (Galloway et al., 2020)	Retrospective cohort	March 24, 2020 to April 17, 2020	England	NA	1157	1157	NA	Patients with confirmed COVID-19 from 2 academic hospitals	NA
Bello-Chavolla (Bello-Chavolla et al., 2020)	Registry data from an open source database from the Mexican Ministry of Health	First patient up to May 18, 2020	Mexico	Mexico	51633	41307	10326	Patients with confirmed COVID-19 from the open source Mexican Ministry of Health database (inpatients and outpatients)	The same as the development population (split by random sampling stratified by mortality status)

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Weng (Weng et al., 2020)	Retrospective cohort	January 1, 2020 to February 15, 2020	China	China	301	176	125	Adult patients with laboratory-confirmed COVID-19 from 2 hospitals	The same as the development population (the criteria used to divide patients in training and testing sets was not clear)
Ko (Weng et al., 2020)	Retrospective cohort	Development cohort: January 10, 2020 to February 24, 2020; Validation cohort: February to July 2020	China	China	467	361	106	Patients with COVID-19 (not clear if laboratory-confirmed) from one hospital, excluding 14 patients without a blood test within 1 day after the hospital admission	Patients with COVID-19 (not clear if (laboratory-confirmed) from 3 hospitals
Xie (Xie et al., 2020)	Retrospective cohort	January and February 2020	China	China	444	299	145	Patients with confirmed COVID-19 from one hospital in Wuhan who had been discharged or died	Patients with confirmed COVID-19 from another hospital in Wuhan, excluding 6 patients who died quickly
Yoo (Yoo et al., 2020)	Retrospective cohort	March 1, 2020 to April 28, 2020	USA	USA	4.840	1.613	1.614	Adult patients with confirmed COVID-19 from 5 hospitals, up to 99 years-old. The sample was randomly split in 3 datasets, the second one was used for development	The same as the development population: randomly split in 3 datasets, the third one was used for validation
Zhang (Zhang et al., 2020)	Retrospective cohort	Not reported	China	United Kingdom	1001	775	226	Adult patients with confirmed COVID-19 from one hospital	Adult patients with confirmed COVID-19 from another hospital
Yadaw (Yadaw et al., 2020)	Retrospective and prospective cohort	March 9, 2020 to April 7, 2020	USA	USA	5051	3841	961	Inpatients and outpatients (including those attended by telehealth) with confirmed COVID-19 from the Mont Sinai Health System (8 hospitals and over 400 ambulatory practices) until April 6, 2020	The same as the development population (randomly split 80/20%) and patients admitted to Mont Sinai Hospitals who were included in the database (with the outcome) on April 7, 2020
Shang (Shang et al., 2020)	Retrospective Cohort	January 1, 2020 to March 27, 2020	China	China	452	113	339	Consecutive patients with confirmed COVID-19 from 2 hospitals in Wuhan, who had severe or critical illness	The same definition as the development population, but from a third hospital in Wuhan
Faisal (Faisal et al., 2020)	Registry data	March 11, 2020 to June 13, 2020	United Kingdom	United Kingdom	6444	3924	2520	Consecutive adult non-elective or emergency medical admissions (COVID-19 and non-COVID-19 patients) from one hospital, who were discharged over a course of three months and had electronic NEWS2 recorded	Consecutive adult non-elective or emergency medical admissions (COVID-19 and non-COVID-19 patients) from another hospital, who were discharged over a course of three months and had electronic NEWS2 recorded

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Mei (Mei et al., 2020)	Retrospective cohort	January 21, 2020 to February 27, 2020	China	China	492	237	Validation 1 = 120 2 = 135	Adult patients with confirmed COVID-19, diagnosed with pneumonia by CT scan, from one hospital in Wuhan. Patients who died within the first 24 hours, with not clinical outcome available or who refused to participate were excluded	The same as the development population, from other 3 hospitals
Zhang (Zhang et al., 2020)	Retrospective cohort	January 12, 2020 to February 9, 2020	China	China	828	516	312	Adult patients with confirmed COVID-19 from one hospital	Adult patients with confirmed COVID-19 from the same hospital in a different time span (February 8–9, 2020) and from another hospital
Lu (Lu et al., 2020)	Retrospective single-center cohort	January 21, 2020 to February 5, 2020	China	NA	577	577	NA	Patients with confirmed or suspected COVID-19 from one hospital	NA
Soto-Mota (Soto-Mota et al., 2020)	Retrospective Cohort	April 30, 2020 to May 20, 2020	Mexico	NA	400	Score developed by consensus	400	NA	Consecutive patients with confirmed COVID-19 from 12 hospitals, with complete clinical information and outcome
Yan (Yan et al., 2020)	Retrospective cohort	Development cohort: January 10, 2020 to February 18, 2020; Validation cohort: February 19–24, 2020	China	China	485	375	110	Adult patients with COVID-19 (not clear if patients had laboratory-confirmed disease), from one hospital, excluding patients with > 20% missing values and breast-feeding women	The same as the development population, admitted after February 18, 2020
Williams (Williams et al., 2020)	Retrospective cohort	Development cohort: any time prior to 2020; validation cohort: January 1st 2020 to April 20, 2020	USA, South Korea, Spain, Australia, Japan, Netherlands	South Korea, Spain, USA	2,126,784	2,082,277	44,507	Healthcare database of 6 countries, in which adult patients with GP, EP or OP visit with influenza or flu-like symptoms, at least 365 days of prior observation, and no symptoms in the preceding 60 days	Adult patients with confirmed with COVID-19, presenting at an initial healthcare provider interaction in a GP, ER or OP visit, and who had no diagnosis of influenzae or pneumonia and no flu-like symptoms in the preceding 60 days
Gue (Gue et al., 2020)	Retrospective single-center cohort	March 10, 2020 to May 30, 2020	United Kingdom	NA	316	316	NA	Consecutive patients with confirmed COVID-19 from a general hospital, who had clinical symptoms at admission	NA

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Das (Das et al., 2020)	Retrospective cohort	January 20, 2020 to May 30, 2020.	South Korea	South Korea	3.524	3.524	NA	Data shared by Korea Centers for Disease Control and Prevention, from 17 provinces. Patients with confirmed COVID-19, with availability of demographic, exposure and diagnosis confirmation features along with the outcome	NA
Levy (Levy et al., MedRxiv)	Retrospective and prospective cohort	March 1, 2020 to May 12, 2020	USA	USA	8391	6162	2229	Adult patients with confirmed COVID-19 from 11 acute care hospitals in New York, from March 1, 2020 to April 23, 7 2020. Patients were excluded if they were still in the hospital at the study end point with a length of stay less than 7 days; if they were transferred to a hospital outside of the health system and their outcomes were unknown; or if they expired but were not marked as discharged in the EH	The same as the development cohort from another hospital in New York from March 1, 2020 to May, 7 2020, and all 12 hospitals from April 24, 2020 to May 6, 2020.
Chen (Chen et al., 2020)	Retrospective cohort	The first patient to January 31, 2020	China	China	1590	1590	NA	Patients with confirmed COVID-19 from 575 hospitals throughout China, excluding cases with incomplete medical records (20.8%)	NA
Sarkar (Sarkar and Chakrabarti, 2020)	Registry data	13th January, 2020 to 28th February, 2020	22 countries in Asia, Australia, Europe and North America	NA	115	115	NA	Open source databased of COVID-19 patients (inclusion criteria is not clear)	NA

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Hu (Hu et al., 2020)	Retrospective cohort	28 January 2020 and 11 March 2020	China	China	247	183	64	Patients with severe confirmed COVID-19 infection admitted to one hospital in Wuhan. patients who had >10% missing values, stayed in the hospital <7 days, were afflicted by a severe disease before admission (e.g. cancer, aplastic anemia or uremia), were unconscious at admission or were directly admitted to the intensive care unit (ICU) were excluded	The same as the development population, admitted at another hospital
Halalau (Halalau et al., 2021)	Hospital admission and in-hospital mortality	In-hospital	Multivariate logistic regression	No	No	Yes	Not clear	No	Not clear
Fumagalli (Fumagalli et al., 2020)	Mortality	In-hospital	Cox regression analysis	No	No	Yes	20	Yes	120
Knight (Knight et al., 2020)	Mortality	In-hospital	LASSO logistic regression	Yes. Multiple imputation with chained equations	Yes. ML	Yes (4C mortality score)	21	No	11426
Liang (Liang et al., 2020)	Composite of ICU admission, need of invasive mechanical ventilation or death	In-hospital	LASSO logistic regression	Yes (if <20%). Predictive mean matching to impute numeric features, logistic regression to impute binary variables, and Bayesian polytomous regression to impute factor features	No	Yes (COVID-GRAM)	72	No	51 (3.2%)
Nicholson (Nicholson et al., 2020)	Need of mechanical ventilation and in-hospital mortality	In-hospital	Multivariate logistic regression	No	No	Yes: one to predict ventilation need (VICE score) and another one for death (DICE score)	49	Yes	Not reported
Garibaldi (Garibaldi et al., 2021)	In-hospital mortality and a composite of disease severity (WHO scale) or in-hospital mortality	In-hospital	Cox regression analysis	Yes. Imputed missing values by chained equations (MICE) with predictive mean matching	Yes. NLP was used to identify presenting symptoms	Yes: COVID-19 Inpatient Risk Calculator (CIRC)	24	No	131

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Sourij (Sourij et al., 2020)	Mortality	In-hospital	Multivariate logistic regression	No	No	Yes	Not clear	Yes	58
Gavelli (Gavelli et al., 2020)	In-hospital mortality and in-hospital clinical stability	In-hospital	Multivariable logistic regression and Cox Regression Hazard models	No	No	Yes (NOVARA score)	NA	No	NA (consensus)
Kazemi (Kazemi et al., 2020)	Mortality	In-hospital	Multivariate logistic regression	No	No	Yes (authors created a CT score not based on the data)	Not available	No	11
Núñez-Gil (Nunez-Gil et al., 2020)	Mortality	In-hospital	Multivariate logistic regression	No	No	Yes	Not clear	Yes	311
Allenbach (Allenbach et al., 2020)	Composite of ICU admission or death	14 days	Multivariate logistic regression	No	No	Yes	42	Yes	32
Kim (Kim et al., 2020)	Mortality	In-hospital	Consensus	No	No	Yes	3	No	7
Altschul (Altschul et al., 2020)	Mortality	In-hospital	Multivariate logistic regression	No	No	Yes	Not clear	Yes	621
Hajifathalian (Hajifathalian et al., 2020)	Mortality	7 days and 14 days	Multivariable logistic regression	Yes. Imputation by chained equations	No	Yes (COVID-AID)	38	Yes	93
Wang (Wang et al., 2020)	Mortality	28 days	Multivariable logistic regression	No	No	Yes (FAD-85)	41	No	24
Zhou (Zhou et al., 2020)	Mortality	In-hospital	Multivariable logistic regression	No	No	Yes (NLAUD)	37	No	51
Goméz (Gomez et al., 2020)	Mortality	30 days	Multivariable logistic regression	No	No	Yes (COVEB)	20	No	33
Galloway (Galloway et al., 2020)	Composite of transfer to ICU or death	In-hospital	LASSO logistic regression	No	No	Yes	19	No	244
Bello-Chavolla (Bello-Chavolla et al., 2020)	Mortality	30 days	Cox proportional risk regression analysis	No	No	Yes	12	No	4276
Weng (Weng et al., 2020)	Mortality	In-hospital	LASSO logistic regression	Yes, for variables with <10% missing values (>10% were excluded from model development). RF.	No	Yes (ANDC)	24	No	21

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Ko (Weng et al., 2020)	Mortality	In-hospital	Machine learning techniques	Yes, imputed with mean values for development and training datasets	Yes, DLN and RF model	Yes (EDRnet)	73	Yes	212 (58.7%)
Xie (Xie et al., 2020)	Mortality	In-hospital	Multivariate logistic regression	No	No	Yes	28	No	155
Yoo (Yoo et al., 2020)	Mortality	In-hospital	Gray's K-sample tests, DeLong's test	No	No	Yes	48	Yes	Not reported
Zhang (Zhang et al., 2020)	Death and poor outcome (developing ARDS, receiving intubation or ECMO treatment, ICU admission or death)	In-hospital	LASSO logistic regression	No	No	Yes (DCS, DC SL, DL)	19	No	33 (4.3%)
Yadaw (Yadaw et al., 2020)	Mortality	In-hospital	Artificial intelligence techniques	Yes, using means	Yes. Recursive feature elimination method for feature selection, and logistic regression, SVM, RF model, and XGBoost algorithms for prediction	Yes (17F and 3F models)	17	No	313 (8.15%)
Shang (Shang et al., 2020)	Mortality	In-hospital	LASSO logistic regression	Yes, multiple imputation methods for variables with <10% missing values	No	Yes (CSS score)	52	No	49
Faisal (Faisal et al., 2020)	Mortality	In-hospital	Multivariable logistic regression	No	No	Yes (CARMc19_N and CARMc19_NB)	Not clear	No	323
Mei (Mei et al., 2020)	Mortality	In-hospital	LASSO logistic regression	No	No	Yes	43	No	105
Zhang (Zhang et al., 2020)	Mortality	14 days and 28 days	Cox regression analyses	Yes. Multiple imputations (method not reported)	No	Yes	30	No	96

(continued on next page)

Table 5 (continued)

Study	Study design	Patient time span	Country of derivation	Country of validation	Sample size (n)	Development sample (n) (for mortality)	Validation sample (n) (for mortality)	Development population	Validation population
Lu (Lu et al., 2020)	Mortality	12 days	Cox regression analysis	No	No	Yes	Not clear	Yes	39
Soto-Mota (Soto-Mota et al., 2020)	Mortality	In-hospital	Consensus	No	No	Yes (LOW-HARM)	NA	No	200 (50%)
Yan (Yan et al., 2020)	Mortality	In-hospital	Machine learning techniques	No	Yes, XGBoost machine learning algorithm	Yes	75	No	174
Williams (Williams et al., 2020)	Hospitalization with pneumonia, hospitalization with pneumonia requiring intensive services or death and death in the 30 days after index date	In-hospital and 30 days after index rate	LASSO logistic regression	No	Yes, ML (train-test-split)	Yes, 3 scores (COVER-F for death)	31.917	No	11407
Gue (Gue et al., 2020)	Mortality	30 days	Multivariable logistic regression	No	No	Yes (COVID-19 Mortality Score)	15	No	145
Das (Das et al., 2020)	Mortality	In-hospital	Logistic regression and machine learning techniques	No	Yes. SVM, K nearest neighbor, RFM and gradient boosting	Yes (CoCoMoRP)	4	No	74
Levy (Levy et al., MedRxiv)	Mortality	7 days	LASSO logistic regression	Yes, imputation of means. Variables with > 50% missing values were excluded.	No	Yes (NOCOS Calculator)	42	No	Not clear
Chen (Chen et al., 2020)	Mortality	14, 21 and 28 days	Multivariate Cox regression analysis	No	No	Yes (nomogram)	37	No	50
Sarkar (Sarkar and Chakrabarti, 2020)	Mortality	In-hospital	Machine learning techniques	No	Yes, RF classification algorithm	Yes	6	No	37
Hu (Hu et al., 2020)	Mortality	In-hospital	LASSO logistic regression	Yes, using bagging tree. Variables with > 30% missing values were excluded	Yes. Logistic regression, PLS regression, EN model, random forest and bagged flexible discriminant analysis (FDA).	Yes	51	No	68

Table 5
Continued.

Study	Variables included in the final model (for mortality)	External validation	How are predictors combined?	AUC in derivation cohort	AUC in validation cohort	Limitations
Halalau (Halalau et al., 2021)	Age, male sex, congestive heart failure, end-stage renal disease, chronic pulmonary disease, DM, hypertension, obesity, nursing home residence, immunocompromised status, congenital heart disease, coronary artery disease, end-stage liver disease and pregnancy	Yes	Points-based score	Not available	0.75 (0.71 – 0.78)	Selection bias: Excluded patients who were hospitalized beyond May 12, 2020. Data on how the score was developed not reported. Absence of an initial validation cohort. Uniform scoring weights of different risk factors. Complete case analysis.
Fumagalli (Fumagalli et al., 2020)	Age, number of comorbidities (CV disease, hypertension, DM, depression, dementia and cancer), respiratory rate, PaO ₂ /FiO ₂ , serum creatinine and platelet count obtained on admission	No	Points-based score	0.90 (0.87 – 0.93)	NA	Modest sample size. No external validation. Variables were selected by univariate analysis. Complete case analysis.
Knight (Knight et al., 2020)	Age, sex, number of comorbidities (chronic cardiac disease, chronic respiratory disease excluding asthma, chronic renal disease defined as estimated glomerular filtration rate \leq 30, mild to severe liver disease, dementia, chronic neurological conditions, connective tissue disease, DM, HIV or AIDS, and malignancy), respiratory rate, SpO ₂ , level of consciousness, urea and CPR obtained on admission	Yes	Points-based score	0.786 (0.781 – 0.790)	0.767 (0.760 – 0.773)	Several potentially relevant comorbidities, such as hypertension, previous myocardial infarction, and stroke, were not included in data collection. The authors considered that inclusion of these comorbidities might have impacted upon or improved the performance and generalizability of the 4C Mortality Score. Secondly, a proportion of recruited patients (3.3%) had incomplete episodes, so there is a possibility of selection bias, if patients with incomplete episodes, such as those with prolonged hospital admission, had a differential mortality risk to those with completed episodes.
Liang (Liang et al., 2020)	Chest radiographic abnormality, age, hemoptysis, dyspnea, unconsciousness, number of comorbidities (COPD, hypertension, DM, coronary heart disease, chronic kidney disease, cancer, cerebrovascular disease, hepatitis B, immunodeficiency), cancer history, neutrophil-to-lymphocyte ratio, lactate dehydrogenase and direct bilirubin obtained on admission	Yes	Logistic Regression	0.88 (0.85 – 0.91)	0.88 (0.84 – 0.93)	Modest sample size for score development and a relatively small sample for validation. The data for score development and validation are entirely from China, which could potentially limit the generalizability of the risk score in other areas of the world. Mortality was quite low (3.2%). Apparently, patients with cancer should gain points for both cancer history and number of comorbidities, not clear.
Nicholson (Nicholson et al., 2020)	Age, sex, diabetes mellitus, chronic statin use, albumin, C-reactive protein, neutrophil-lymphocyte ratio, mean corpuscular volume, platelet count, and procalcitonin obtained on admission	Yes	Logistic Regression	0.87 (0.83 – 0.91)	0.80 (0.75 – 0.85)	Modest sample sizes in both our derivation and validation cohorts. The number of events on the derivation and validation cohort separately was not informed (211 in total). Variables were selected by univariate analysis. Complete case analysis.
Garibaldi (Garibaldi et al., 2021)	Age, nursing home residence, sex, BMI, Charlson Comorbidity Index, SaO ₂ /FiO ₂ ratio obtained on admission	No	Cox regression analysis	Not available	Not available	Modest sample size. No external validation. Too many variables tested in the model for the number of events (24/131). To try to overcome that, authors tested variables "in blocks"

(continued on next page)

Table 5 (continued)

Study	Variables included in the final model (for mortality)	External validation	How are predictors combined?	AUC in derivation cohort	AUC in validation cohort	Limitations
Sourij (Sourij et al., 2020)	Age, arterial occlusive disease, CRP, estimated GFR and aspartate AST levels obtained on admission	No	Nomogram	0.889 (0.837 – 0.941)	NA	Small sample size and number of events. Number of variables tested not clear. Complete case analysis, and predictors with > 20% missing values were excluded. No external validation
Gavelli (Gavelli et al., 2020)	Presence of comorbidity (any disease on active therapy), SpO2 and respiratory rate after a trial of 15 minutes with oxygen at a FiO2 0.5	No	Points-based score	NA	Not reported	Score developed by consensus. Modest sample size. Number of events is not clear. Single-center study. No external validation. AUC and accuracy not presented.
Kazemi (Kazemi et al., 2020)	Age, sex, comorbidity (cardiovascular and pulmonary), diffused distribution of CT abnormality, total CT-score and dyspnea at admission	No	Logistic Regression	0.73 (95% CI not reported)	NA	Small sample size and number of events. Too many variables tested for the low number of events. Comorbidities were not well defined, percentage of involvement included in CT score is subjective and peripheral involvement is not well defined. Complete case analysis. High risk of selection bias: All 3 hospitals were referral centers for COVID-19 patients, so it is possible that the overall CT- score of the patients in this study would not be representative of the general population
Núñez-Gil (Nunez-Gil et al., 2020)	Age, hypertension, obesity, renal insufficiency, any immunosuppressive condition, SpO2, CRP obtained on admission	No	Points-based score	0.88 (0.85 – 0.91)	NA	No external validation. Variables were selected by univariate analysis. Complete case analysis. Variables included in the model not clearly defined. Authors reported that some incident events in the participating centers may not have been diagnosed and/or not been reported. The data analysis and modeling focused on only two countries (Italy and Spain) of the four initially considered, since as previously mentioned heterogeneity among countries with regard to clinical features and death-risk assessment could limit the representative nature of the sampling.
Allenbach (Allenbach et al., 2020)	Age, WHO clinical scale, CRP and lymphocytes count obtained on admission	No	Points-based score (but AUC presented based on the logistic regression model)	0.786 for the composite outcome and 0.803 for death (after correction for over-optimism; IC95% not reported)	0.787 for the composite outcome and 0.827 for death (after correction for over-optimism; IC95% not reported)	Small sample size of both development and validation samples. Too many predictors tested for a small number of events. Complete case analysis. External validation sample not described. The external sample consisted of patients from a regional non-university hospital, which could explain the differences on catchment area and patient recruitment. In the acute context of the first SARS-CoV-2 epidemic wave in France, we relied on a sample prospectively defined by consecutive eligible patients in the study center.
Kim (Kim et al., 2020)	Myocardial damage marker (creatinine kinase-MB [CK-MB] or troponin-I > the 99th percentile upper reference limit) + Heart failure marker (NT-proBNP ≥ 125 pg/mL) + Electrical abnormality marker (first detected or newly developed supraventricular tachycardia, ventricular tachycardia, ventricular fibrillation, atrial fibrillation, bundle branch block, ST-segment elevation/depression, T-wave flattening/inversion, and QT interval prolongation on ECG)	No	Points-based score	Not reported	NA	Score developed by consensus. Small sample size and small number of events. Accuracy not assessed. The protocol for the evaluation of cardiac injury was not controlled. The attending physician decided each category of the test according to the patient's condition at the time of the management. When the test was not performed, it is assumed as a negative result because the physician considered it as an unnecessary test or the result might be negative.

(continued on next page)

Table 5 (continued)

Study	Variables included in the final model (for mortality)	External validation	How are predictors combined?	AUC in derivation cohort	AUC in validation cohort	Limitations
Altschul (Altschul et al., 2020)	Age, sex, SpO ₂ , MAP, INR, creatinine, BUN, interleukin-6 (IL-6), CRP and procalcitonin obtained on admission	Yes	Points-based score	0.824 (0.814 to 0.851)	0.798 (0.789 to 0.818)	Complete case analyses, variables selected by univariate analyses
Hajifathalian (Hajifathalian et al., 2020)	Age, mean arterial pressure, serum creatinine and severity of hypoxia at hospital presentation.	Yes	Multivariate logistic regression	7 days: 0.877 (95%CI 0.831–0.923); 14 days: 0.847 (95%CI 0.806–0.888)	7 day (0.851 [0.781 to 0.921]); 14 day (0.825 [0.764 to 0.887])	Modest sample size for development and validation, less than 100 events both in the development and validation cohorts, short follow-up time
Wang (Wang et al., 2020)	Age, ferritin and D-dimer obtained on admission	Yes	Logistic regression and nomogram	0.871 (based on its optimal cut-off value = 85)	Not available (link for supplemental material does not work)	Single-center study, with small sample for development and validation, less than 100 events both in the development and validation cohorts. Complete-case analysis. D-dimer assay not described. AUC for external validation not available to the readers.
Zhou (Zhou et al., 2020)	Lactate dehydrogenase, albumin, BUN, NLR and D-dimer obtained on admission	No	Nomogram	0.955 (95% CI not provided)	NA	Single-center study, with small sample size, including cases not confirmed by RT-PCR, and less than 100 events. Complete-case analysis and tests too many variables for the number of events. D-dimer assay not described.
Goméz (Gomez et al., 2020)	Age, creatinine, glucose and white blood cells obtained on admission	No	Not clear	0.874 (0.816–0.933)	NA	Single-center study, with small sample size, including cases not confirmed by RT-PCR, and less than 100 events. Complete-case analysis and tests too many variables for the number of events.
Galloway (Galloway et al., 2020)	Age, sex, ethnicity, DM, hypertension, chronic lung disease, SpO ₂ , radiographic severity score, neutrophil count, respiratory rate, CRP, albumin, creatinine obtained on admission	No	Points-based score	0.697 (0.652,0.741)	NA	Modest sample size. No external validation. Complete case analysis. AUC < 0.70
Bello-Chavolla (Bello-Chavolla et al., 2020)	Age, diabetes, obesity, CKD, COPD, hypertension, immunosuppression and COVID-19 pneumonia	Yes	Points-based score	0.823 (95% CI not reported)	0.830 (95% CI not reported)	The use of data collected from a sentinel surveillance system model, what raises concern about data quality. The same score for inpatient and outpatients and sensitivity analysis was not performed to assess accuracy for patients who were hospitalized. Apparently, complete case analysis.
Weng (Weng et al., 2020)	Age, neutrophil-to-lymphocyte ratio, D-dimer and C-reactive protein obtained on admission	Yes	Nomogram and logistic regression	0.921 (0.835–0.968)	0.975 (0.947–1.0)	Small sample size for development and validation, with < 100 events in both cohorts. Variables with > 10% missing values were excluded. D-dimer assay was not reported.
Ko (Weng et al., 2020)	Lymphocytes, neutrophils, albumin, LDH, neutrophil count (?), CRP, prothrombin activity, calcium, urea, estimated GFR, monocytes, globulin, eosinophils, glucose, RDW, bicarbonate, RDW standard deviation, platelet count, mean platelet volume, platelet large-cell ratio, prothrombin time, total protein, platelet distribution width, aspartate aminotransferase, thrombocytocrit, eosinophil count, alkaline phosphatase, INR	Yes	AI model	Not reported	Not reported	Small sample size for development and validation, too many variables tested for the limited number of events, high mortality rate, with possibility of selection bias. Not clear if included laboratory-confirmed COVID-19 patients only. The number of predictors make it difficult to be applicable at bedside.

(continued on next page)

Table 5 (continued)

Study	Variables included in the final model (for mortality)	External validation	How are predictors combined?	AUC in derivation cohort	AUC in validation cohort	Limitations
Xie (Xie et al., 2020)	Age, lymphocyte count, lactate dehydrogenase and SpO2 obtained on admission	Yes	Logistic regression and nomogram	0.880 (95% CI not reported)	0.980 (0.958–1.00)	High risk of selection bias: the cohort was conducted early in the pandemic, there was a high mortality rate (51.8% in development cohort and 47.6% in the validation cohort), and it may not accurately represent patients with mild or asymptomatic COVID-19 (as they were not being tested). Small sample size for development and validation, less than 100 events both. Complete case analysis.
Yoo (Yoo et al., 2020)	Glasgow coma scale, oxygen support level, BUN, age, lymphocyte percentage, troponin	Yes	Points-based score	Not reported, as AUC was used to define the variables for the score.	At admission 0.81; maximum through admission 0.91; mean through admission 0.92	The authors reported that documentation of all kinds was inconsistent during the first wave of Covid-19, and the environments at different hospitals varied substantially. While it is unlikely that a laboratory result or medication administration was missed, inconsistencies in flowsheet documentation during this period could mean that the timings of different modes of oxygen administration were not always accurately capture. The statistical test used to produce the score is not adequate according to the TRIPOD and may lead to over optimism.
Zhang (Zhang et al., 2020)	DCS (demographic, comorbidities and symptoms): age, sex, chronic lung disease, DM, hypertension, immunosuppression, cancer, CKD, heart disease, cough, dyspnea, diarrhea; DCSL (demographic, comorbidities, symptoms and laboratory tests): age, sex, chronic lung disease, DM, cancer, cough, dyspnea, CRP, creatinine, platelets, neutrophils and lymphocytes counts; DL (demographic and laboratory tests): age, sex, CRP, creatinine, platelets, neutrophils and lymphocytes counts (around admission)	Yes	Logistic regression	DCS: 0.79; DCS: 0.89; DL: 0.91 (95% CI not reported)	DL: 0.74 (95% CI not reported)	Authors reported that clinical datasets were collected when healthcare services were under severe strain. Data extraction sought to ensure consistency and accuracy, but there is missing data in both datasets, and the analysis was complete case based. Sample sizes for development and validation were small, with < 100 events. Clinical assessments at admission such as SpO2 were not available in either dataset. The external validation dataset has very different case-mix, and only had follow-up to a fixed date (6–39 days). Although the Wuhan cohort includes many people with less severe disease, in the validation cohort most admitted patients are likely to have severe disease. Although the authors reported all variables were included in the model, for most of the included ones the 95% CI of the OR included 1.0
Yadaw (Yadaw et al., 2020)	17F: age, sex, ethnicity, encounter type, temperature, diastolic blood pressure, oxygen saturation at presentation, minimum oxygen saturation, smoking, asthma, COPD, obesity, DM, HIV, cancer; 3F: age, minimum oxygen saturation, and type of patient encounter, obtained the day of admission	Yes	Artificial intelligence (XGBoost)	0.91 (95% CI not provided)	0.91 (95% CI not provided)	As it includes inpatients and outpatients, important laboratory parameters were not tested. The authors reported that the clinical features available were limited to those routinely collected during hospital encounters, and they pointed out that development of even better prediction models should be possible using a richer set of features.
Shang (Shang et al., 2020)	Age, coronary heart disease, % of lymphocytes, procalcitonin, D-dimer	Yes	Points-based score	0.919 (95% CI 0.870–0.970)	0.938 (95% CI 0.902–0.973)	Small sample size in development (113 participants) and validation cohorts, with < 100 events in the development one. Too many variables tested for the number of events.
Faisal (Faisal et al., 2020)	CARMc19_N: 10 [age, sex, COVID-19 (yes/no), NEWS2 score and subcomponents] and CARMc19_NB: 18. All variables from CARMc19_N + 7 blood test results + AKI score	Yes	Points-based score	CARMc19_NB = 0.87 (95% CI 0.85–0.89) vs CARMc19_N 0.86 (95% CI 0.84–0.87)	CARMc19_NB = 0.88 vs CARMc19_N = 0.86	Not exclusively for COVID-19 patients. COVID-19 was identified by ICD-10 code which depends on clinical judgment. Risk of selection bias, as only patients with NEWS2 recorded were included. Complete case analysis.

(continued on next page)

Table 5 (continued)

Study	Variables included in the final model (for mortality)	External validation	How are predictors combined?	AUC in derivation cohort	AUC in validation cohort	Limitations
Mei (Mei et al., 2020)	Age, NLR, admission body temperature, AST, total protein	Yes	Points-based score	0.912 (95% CI 0.878–0.947)	VC1 = 0.928 (95% CI 0.884–0.971) and VC2 = 0.883 (0.815–0.952)	Risk of selection bias due to inclusion/exclusion criteria, included only patients from Wuhan. Small sample size for development and validation. Complete case analysis.
Zhang (Zhang et al., 2020)	Age, LDH, NLR and direct bilirubin obtained on admission	Yes	Nomogram	0.886 (95% CI 0.873–0.899)	0.879 (95% CI, 0.856–0.900) and 0.839 (95% CI [0.798–0.880] for each one of the hospitals	Small sample size for development and validation, < 100 events for both cohorts. The amount of missing data differed between the survivor and non-survivor groups. The study included a high population of patients who were severely ill, the authors pointed out there may be a selection bias when identifying the risk factors of mortality
Lu (Lu et al., 2020)	Age, CPR	No	Cox regression analysis, decision tree	Not reported	NA	Included both patients with confirmed and not confirmed disease, small sample size with < 100 events, number of potential predictors tested was not clear. No external validation.
Soto-Mota (Soto-Mota et al., 2020)	Age, hypertension, white blood cell count, lymphocyte count, myocardial necrosis marker, creatinine, SpO2 (not clear in which moment)	No	Logistic regression	NA	Provided by different cut-offs, ranging from 0.61 to 0.90 (95% ranges from 0.59 to 0.93), with best AUC for 25 points (0.90 [95% CI 0.87–0.93])	Score developed by consensus. Not clear the moment it is meant to be used. Risk of selection bias, high mortality in the cohort (50%)
Yan (Yan et al., 2020)	LDH, lymphocytes and CRP obtained at hospital admission	Yes	Multi-tree XGBoost model	0.978 (IC 95% not provided)	0.951 (CI 95% not provided)	Single-center study, with small sample for development and validation, less than 100 events in the validation cohort. Apparently, complete-case analysis.
Williams (Williams et al., 2020)	Age, sex, history of cancer, COPD, diabetes, heart disease, hypertension, hyperlipidemia and kidney disease.	Yes	Points-based score	0.896 (95% CI 0.72 – 0.90)	CUIMC database 0.820 (95% CI 0.796–0.840); HIRA database 0.898 (95% CI 0.857–0.940); SIDIAP 0.895 (95% CI 0.881–0.910); VA 0.717 (0.642–0.791)	The authors reported they were unable to develop a model on COVID-19 patient data due the scarcity of databases that contains this information in sufficient numbers. Based on secondary data, with possibility of misclassifications of predictors (diseases is incorrectly recorded in a patient's history, incorrect recording of influenza or COVID-19, and authors were unable to include some suspected diseases predictors such as BMI/obesity in the analysis due to the inconsistency with which these measures are collected and reported across the databases included in the study. Patients may day after 30 days, and this will be recorded as a non-event. Apparently, complete case analysis.

(continued on next page)

Table 5 (continued)

Study	Variables included in the final model (for mortality)	External validation	How are predictors combined?	AUC in derivation cohort	AUC in validation cohort	Limitations
Gue (Gue et al., 2020)	Age, sex, hypertension, coronary artery disease, heart failure, atrial fibrillation, oral anticoagulants, modified sepsis-induced coagulopathy (mSIC) score (INR, platelet count, qSOFA score)	No	Points-based score	0.793 (95% CI 0.745–0.841)	NA	Small sample size from a single center, no external validation. Complete case analysis. Authors pointed out that patients at the highest risk may be deemed too sick for maximal intervention and may be denied ICU treatment; predictors and their assigned weights in the final model.
Das (Das et al., 2020)	Age, sex, province (in South Korea) and exposure (nursing home, hospital, religious gathering, call center, community center, shelter and apartment, gym facility, overseas inflow, contact with patients and others)	No	Logistic regression (SMOTE)	0.830 (95% CI not reported)	NA	Risk of selection bias (only patients with complete data were included), unavailability of crucial clinical information on symptoms, risk factors and clinical parameters. Less than 100 events. No external validation
Levy (Levy et al., MedRxiv)	Age, length of stay, SpO ₂ , neutrophil, RDW, sodium urea (on admission and every 2 days)	Yes	Logistic regression	0.86 (95% CI not reported)	0.82 (95% CI not reported)	Data were imputed for variables with up to 50% missing values. Follow up was too short (7 days), what causes a high risk of bias, as a significant proportion of patients may die after 7 days. Authors did not show how to calculate the score.
Chen (Chen et al., 2020)	Age, coronary heart disease, cerebrovascular disease, dyspnea, procalcitonin, aspartate aminotransferase, total bilirubin upon admission	No	Nomogram	0.91 (95% CI, 0.85–0.97)	NA	High risk of selection bias (20.8% patients with incomplete data were excluded), modest sample size, with < 100 events. No external validation. Complete case analysis. Authors did not show how to calculate the score.
Sarkar (Sarkar and Chakrabarti, 2020)	Age, sex, from Wuhan, visit to Wuhan, days from symptom onset to hospitalization	No	RF classification algorithm	0.97 (95% CI not reported)	NA	Small sample size, with < 100 events. High risk of selection bias: from 1085 patients, 652 (60.1%) were excluded due to missing values, and the model was developed using one 115 patients(10.6%). Data quality is questionable, as the study is based in open source database.
Hu (Hu et al., 2020)	Age, CRP, D-dimer, lymphocyte count at admission	Yes	Points-based score	0.895 (95% CI not reported)	0.881 (95% CI not reported)	Small sample size of both development and validation samples, with < 100 events. Too many predictors tested for a small number of events. The authors did not exclude patients transferred from other hospitals (so the assessment was not the first hospital assessment in all patients). Single center study, patients from both derivation and validation sets were from Tongji Hospital, which is one of the hospitals with a high level of medical care in China (the authors reported that some critically ill patients who recovered there might die in other hospitals with suboptimal or typical levels of medical care).

AUC: area under the curve; BMI: body mass index; CI: confidence interval; CPOD: chronic obstructive pulmonary disease; CPR: C-reactive protein; CT: computed tomography; DLN: deep learning networks; DM: diabetes mellitus; GFR: glomerular filtration rate; ICU: intensive care unit; LASSO: least absolute shrinkage and selection operator logistic regression; NA: not applicable; RDW: red blood cell distribution width; PLS: partial least squares RF: Random Forest; SF ratio: SpO₂/FiO₂ ratio; SVM: support-vector machine; XGBoost: eXtreme Gradient Boosting; WHO: World Health Organization.

Table 6
Discrimination of risk scores within validation cohort (complete case)

Score	Number of patients	Number of deaths (%)	AUROC (95%CI)
4C Mortality Score	625	110 (17.6%)	0.790 (0.750–0.830)
A-DROP	704	148 (21%)	0.780 (0.740–0.820)
ABC ₂ SPH	779	148 (19%)	0.853 (0.822–0.885)
AID-14	929	187 (20.1%)	0.752 (0.714–0.790)
AID-7	929	187 (20.1%)	0.751 (0.713–0.789)
CURB65	770	165 (21.4%)	0.748 (0.709–0.786)
E-CURB65	146	33 (22.6%)	0.768 (0.682–0.853)
NEWS-FAST	578	112 (19.4%)	0.739 (0.692–0.786)
NEWS2	425	90 (21.2%)	0.746 (0.687–0.804)
NOVARA	865	176 (20.3%)	0.656 (0.613–0.699)
qSOFA	850	172 (20.2%)	0.653 (0.609–0.697)
REMS	780	145 (18.6%)	0.753 (0.712–0.793)
SOFA	288	59 (20.5%)	0.778 (0.712–0.843)
Xie	475	93 (19.6%)	0.816 (0.768–0.863)
Yan	431	81 (18.8%)	0.650 (0.603–0.697)
Zhang	279	67 (24%)	0.810 (0.751–0.869)

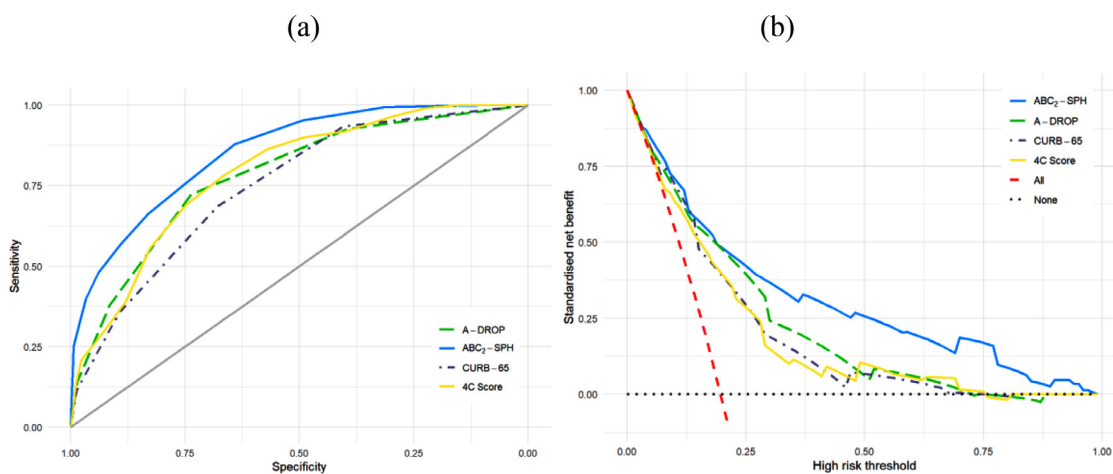


Figure 5. ROC curves (a) and decision curves (b) for best performing scores.

Discussion

The ABC₂-SPH score is simple, objective, easily available at hospital admission, and easily calculated, employing seven well-defined and routinely recordable variables. It has been shown to be a reliable tool in estimating in-hospital mortality in COVID-19 patients. Model performance comparison surpassed other existing scores. The majority of developed scores are limited by methodological bias in development cohorts. Robust models require large sample sizes, which produce more reliable and accurate results (Moons et al., 2015). Among the models analyzed for comparison, 30.8% used a sample with > 1000 patients, 41.0% used a sample with < 500 patients, and 41.0% were developed and validated in a sample with < 100 events. Most of the models that were analyzed for comparison (69.2%) did not perform or did not describe whether imputation methods were used for the missing data; therefore, there is a high risk of bias related to the treatment of missing data. The approach of excluding the missing data and performing the analysis with the complete cases can lead to biased results, since the complete cases may not adequately represent the entire original study sample, generating a selection bias (Moons et al., 2015). External validation of the developed model was not performed in 43.6% of the analyzed studies. As the accuracy of a prediction model is always high, whether the model is validated on the development cohort used to derive the model only, the assessment

of accuracy in those studies may be overoptimistic (Moons et al., 2015). Previous studies have observed the variables included in the ABC₂-SPH score as risk factors for severe COVID-19, which shows that the current results are in line with the available evidence. Age and number of comorbidities were reported in several publications as independent risk factors for developing severe COVID-19 and mortality (Fumagalli et al., 2020, Knight et al., 2020, Liang et al., 2020). The strong age gradient per decade after 60 years of age is in line with other series (Fumagalli et al., 2020, Allenbach et al., 2020). Aging is associated with a well-known decline in adaptive and innate immunity, which plays a major role in the increased susceptibility of infections (Fuentes et al., 2017). It is also related to an increased severity in pro-inflammatory response and increased cytokine production, which is believed to increase patient vulnerability to the unregulated inflammatory response in COVID-19 (Sherwani and Khan, 2020). The number of comorbidities indicates the importance of pre-existing conditions to the severity of COVID-19. Even though comorbidities are age-dependent factors, the number of comorbidities remained an independent risk factor in the final model. As it was aimed to use variables easily available at emergency department admission at any institution, it was opted to evaluate the SpO₂/FiO₂ ratio (SF ratio) instead of the ratio of arterial oxygen partial pressure over the fraction of inspired oxygen (PaO₂/FiO₂ ratio), like the COVID-AID score (Hajifathalian et al., 2020). Arterial

blood gas puncture and analysis is an invasive and complex procedure, which may be time-consuming for the team. The SF ratio was already validated as a substitute for the PaO₂/FiO₂ ratio in assessing the oxygenation criterion of patients with acute lung injury and acute respiratory distress syndrome (Rice et al., 2007).

COVID-19-associated hyperinflammation and coagulopathy are correlated with a wide deviation in various inflammatory markers and hemostasis parameters, and thus these are potential prognostic markers of increased mortality in COVID-19 (Gungor et al., 2021, IJBD et al., 2020). Consistent with prior studies (Nicholson et al., 2020, Weng et al., 2020), this study also observed utility of C-reactive protein (CRP), an acute phase reactor with established prognostic prediction roles in ICU septic and non-septic patients (Qu et al., 2020, Koozi et al., 2020), and thrombocytopenia. The prognostic value of thrombocytopenia in COVID-19 patients was shown in a recent meta-analysis (Bashash et al., 2020), and it has also been included in other scores (Fumagalli et al., 2020, Nicholson et al., 2020). The exact explanation is still unknown, and it is probably multifactorial, related to direct infection of bone marrow cells by the virus, resulting in abnormal hematopoiesis; platelet destruction by the immune system; endothelial damage triggering platelet activation, aggregation and microthrombi in the lungs; and abnormal platelet defragmentation in the lungs (Bashash et al., 2020).

A recent meta-analysis (16 studies, n = 3480) showed significantly higher levels of D-dimer on admission in patients who died compared with those who were discharged (Gungor et al., 2021). This exam was included as a predictor in different scores (Wang et al., 2020, Zhou et al., 2020, Weng et al., 2020, Shang et al., 2020, Hu et al., 2020). However, the value has to be determined with the same methodology, preferably from the same manufacturer, and this information was not available in any of the studies. A recent publication highlighted confusion and potential for misinformation in reporting D-dimer data in COVID-19 (Favaloro and Thachil, 2020). The authors provided examples of serious errors in the reported values and/or units, as reported in the literature related to COVID-19, even in high impact journals.

Most studies have not reported how they dealt with cases who were transferred between hospitals. As the score is intended to be used at hospital admission and to avoid the possibility of patients already having critical disease, the current study opted to exclude patients who were transferred between hospitals and if admission data from the first hospital were not available.

Blood urea nitrogen elevation was a strong predictor for mortality, which is in line with other scores (Knight et al., 2020, Ko et al., 2020, Levy et al., MedRxiv). Although autopsy studies did not find conclusive evidence of SARS-CoV2 infection in the kidney, some authors have hypothesized that the damage may be mediated by direct cytopathic effects of SARS-Cov2 on the kidney tissue, immune-mediated damage due to virus-induced immune complexes, and the effects of the inflammatory response, hypoxia and shock (Huang et al., 2020, Lemos et al., 2003, Yao et al., 2020). A major strength of the ABC₂-SPH score is its simplicity, with the use of objective parameters, which helps to reduce inter-user variability, easily available at the emergency department presentation, even in under-resourced settings. The present study followed strict methodological criteria, as recommended by the TRIPOD (Moons et al., 2015) checklist and PROBAST (Wolff et al., 2019), and was based on robust samples of patients with confirmed SARS-CoV-2 infection from 36 hospitals in four Brazilian states, to ensure diversity of the studied population and representativeness of the intended target population. The majority of published scores were developed in China or the USA (56.4%) and Europe (25.6%); this is the first study in the Latin American population. Data were obtained by detailed medical chart reviews, and comprehensive data were able to be collected from a large number

of patients; 98.5% of the patients were followed from admission to discharge or death. Decisions about which predictors to retain in the final model did not rely on potentially biased univariable selection of predictors. They were based on clinical reasoning, previous evidence from other cohorts and systematic reviews on prognostic factors for COVID-19 patients and availability of predictor measurement at hospital admission (Moons et al., 2015).

With regards to study limitations, due to the pragmatic study design, laboratory tests were performed at the discretion of the treating physician, and a full dataset on all laboratory parameters of interest was not available. Some laboratory parameters, which proved to be of prognostic relevance in other studies, were unavailable for at least two-thirds of patients in this sample. Therefore, it cannot be ruled out that variables with a higher proportion of missing data would have had a significant impact on mortality prediction. Additionally, it was unable to assess the predictive ability of some scores, as some required variables that were unavailable. In the Spanish validation cohort, true mortality risk was somewhat underestimated in very high-risk patients. The majority of those patients came from the beginning of the pandemic, and at that time doctors used to admit patients earlier in the course of the disease. Therefore, patients were admitted in a less severe condition than Brazilian ones, which can be inferred from the difference in SF ratio. This is probably the explanation for the underestimation in very high-risk patients among the Spanish cohort. Even with this difference, the score was able to show very good discrimination to differentiate patients with high risk vs. non-high risk, which is the main goal in using such scores.

Implications for clinicians and policymakers

The ABC₂-SPH score may be very useful in a real-world setting to provide healthcare practitioners with the support that is needed to help them better identify and prioritize the care of patients who have a higher risk of death. Its development and validation followed strict methodological criteria, and the score fulfils the characteristics of an ideal score (Oprita and Aignatoaie, 2014). It can be used in all emergency departments, regardless of the level of resource settings. The results represent the experience of 36 hospitals in 17 cities in Brazil, and one hospital in Spain, and they are highly relevant to the current pandemic. It can be easily calculated at the bedside or could be easily integrated into electronic medical records for automatic computation. It may help clinicians to identify high-risk patients from the triage phase, as well as to identify those most appropriate to be enrolled into therapeutic trials, and may make it possible to expand inclusion criteria through the early identification of patients who may benefit from therapy (Haimovich et al., 2020). It might also be useful to help guide recommendations for early palliative consultation (Altschul et al., 2020, Sheehan et al., 2020).

Different from what has mistakenly been suggested (Knight et al., 2020, Gavelli et al., 2020, Halalau et al., 2021, Galloway et al., 2020), the results from this study do not suggest that patients from the low-risk group be discharged for home treatment. No score so far has specifically tested this hypothesis. A recent editorial highlighted the importance of taking into account the “treatment paradox”: patients identified for the low-risk group were at low risk due to the interventions received in hospital (Sperrin and McMillan, 2020). It must not be interpreted as the risk to a patient if no actions are taken. Sperrin and McMillan used counterfactual prediction modelling as a potential solution to minimize bias from treatment paradox (Sperrin and McMillan, 2020). More importantly, due to the treatment paradox, scoring systems developed and validated in in-hospital settings cannot be used in outpatient settings without further validation, as has mistakenly been suggested (Zhang et al., 2020).

Unanswered questions and future research

It is believed that the ABC₂-SPH score may hold potential generalizability for other countries. However, prediction models are population-specific and may produce different results in different populations (KGM et al., 2015). Considering that thresholds for admission may vary, hospitalized COVID-19 populations may be different, the outcome events may be different and patient management may be different; therefore, further validation (and recalibration) in different healthcare settings is recommended. In particular, the current model might underestimate mortality in high-risk individuals.

As it was opted to develop the score focusing on information available at admission, as this would make it more useful for clinicians, other important factors during hospitalization that may impact prognosis were not included. Further analysis involving these factors is required.

The ABC₂-SPH score may help clinicians to make a prompt and reasonable decision to optimize patient management and potentially reduce mortality. However, further prospective studies are needed to investigate whether the use of the score in the emergency department may trigger actions that result in reduced complications and hospital mortality. Additionally, due to the rapidly changing nature of COVID-19 and disease management, model performance should be closely monitored over time and space (Sperrin and McMillan, 2020).

Future studies may also investigate risk factors for mortality among patients who develop COVID-19 symptoms during hospital admissions due to other conditions.

Conclusion

In conclusion, the ABC₂-SPH rapid scoring system and a web-based risk calculator have been developed and validated. This score – based on age, number of comorbidities, blood urea nitrogen, C-reactive protein, platelet count, peripheral oxygen saturation, and oxygen support at admission – is an inexpensive tool that has been shown to objectively and accurately predict in-hospital mortality in COVID-19 patients. It may be used at the bedside for earlier identification of in-hospital mortality risk and, thus, inform clinical decisions and the assignment to the appropriate level of care and treatment for COVID-19 patients.

Declarations

Ethics approval and consent to participate: The study protocol was approved by the Brazilian National Commission for Research Ethics (CAAE 30350820.5.1001.0008). Individual informed consent was waived due to the severity of the situation and the use of deidentified data, based on medical chart review only. For the independent external validation cohort, it was approved by the and Vall d'Hebron University Hospital Research Ethics Committee (PR(AG)183/2020).

Conflicts of interest: The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement: Data are available upon reasonable request.

Data Access, Responsibility, and Analysis: The lead authors (MSM and MCP) had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Transparency declaration: The lead authors (MSM and MCP) affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of

the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

Funding: This study was supported in part by Minas Gerais State Agency for Research and Development (*Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG*) [grant number APQ-00208-20], National Institute of Science and Technology for Health Technology Assessment (*Instituto de Avaliação de Tecnologias em Saúde – IATS*)/ National Council for Scientific and Technological Development (*Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq*) [grant number 465518/2014-1], and CAPES Foundation (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*) [grant number 88887.507149/2020-00]. AS was supported by a postdoctoral grant “Juan Rodés” (JE18/00022) from Instituto de Salud Carlos III through the Ministry of Economy and Competitiveness, Spain.

Role of the funder/sponsor: The sponsors had no role in the study design; data collection, management, analysis, and interpretation; writing the manuscript; and decision to submit it for publication. MSM and MP had full access to all the data in the study and had responsibility for the decision to submit for publication.

Author contributions: Substantial contributions to the conception or design of the work: MMS, LSM, KPMPM, IJBN, CAP, LMO. Substantial contributions to the acquisition, analysis, or interpretation of data for the work: MMS, LSM, KPMPM, MVRSS, IJBN, TLSS, RLRC, MCP, RTS, LMO, AVS, AOM, ALBAS, APW, ABG, AACM, ARB, BLF, BMC, CTCAS, CMR, CDG, CCRC, CAC, DVS, DP, DHV, EC, ECP, EMSK, FBL, FA, FAB, FGA, FB, GPC, GGV, GANB, GCCM, GFN, HCN, HD, HRV, HCG, IMG, JDLB, JCA, JMC, JDSSG, JDP, JMR, KBR, KCJRP, KAMS, LSO, LSP, LDS, LSFC, LK, LCC, LCSM, LEAS, MASC, MAF, MDS, MGTT, MC, MAPF, MCPBL, MFG, MMAC, MCAN, MPF, MHGJ, NCSS, NLR, NTC, NRO, PKZ, PLA, PJLM, RSCA, RCM, RL, RAV, RGF, RBC, RP, RXC, RMM, RLMS, RMA, RFS, SCF, SMMG, SFA, SAP, TFO, TK, TCO, TSMMA, THOD, VBS, VMRG, VALV, YCR, DTMOF, AS, BR. Drafted the work: MMS, MCP, LMO. Revised the manuscript critically for important intellectual content: all authors. Final approval of the version to be published: all authors. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: MSM and MCP.

Acknowledgments

We would like to thank the hospitals, which are part of this collaboration, for supporting this project: Hospital Bruno Born; Hospital Cristo Redentor; Hospital das Clínicas da Faculdade de Medicina de Botucatu; Hospital das Clínicas da UFMG; Hospital das Clínicas da Universidade Federal de Pernambuco; Hospital de Clínicas de Porto Alegre; Hospital Santo Antônio; Hospital Eduardo de Menezes; Hospital João XXIII; Hospital Julia Kubitschek; Hospital Mãe de Deus; Hospital Márcio Cunha; Hospital Mater Dei Betim-Contagem; Hospital Mater Dei Contorno; Hospital Mater Dei Santo Agostinho; Hospital Metropolitan Dr. Célio de Castro; Hospital Metropolitan Odilon Behrens; Hospital Moinhos de Vento; Hospital Nossa Senhora da Conceição; Hospital Regional Antônio Dias; Hospital Regional de Barbacena Dr. José Américo; Hospital Regional do Oeste; Hospital Risoleta Tolentino Neves; Hospital Santa Cruz; Hospital Santa Rosália; Hospital São João de Deus; Hospital São Lucas da PUCRS; Hospital Semper; Hospital SOS Córdio; Hospital Tachini; Hospital Unimed-BH; Hospital Universitário Canoas; Hospital Universitário Ciências Médicas; Hospital Universitário de Santa Maria. We also thank all the clinical staff at those hospitals, who cared for the patients, and all undergraduate students who helped with data collection.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijid.2021.07.049.

References

- Allenbach Y, Saadoun D, Maalouf G, et al. Development of a multivariate prediction model of intensive care unit transfer or death: A French prospective cohort study of hospitalized COVID-19 patients. *PLoS One* 2020;15(10).
- Altschul DJ, Unda SR, Benton J, et al. A novel severity score to predict inpatient mortality in COVID-19 patients. *Scientific Reports* 2020;10(1):16726, 1–8.
- Bashash D, Hosseini-Baharanchi FS, Rezaie-Tavirani M, et al. The Prognostic Value of Thrombocytopenia in COVID-19 Patients; a Systematic Review and Meta-Analysis. *Archives of Academic Emergency Medicine* 2020;8(1):e75.
- Bello-Chavolla OY, Bahena-López JP, Antonio-Villa NE, et al. Predicting mortality due to SARS-CoV-2: A mechanistic score relating obesity and diabetes to COVID-19 outcomes in Mexico. *The Journal of Clinical Endocrinology & Metabolism* 2020;105(8):2752–61.
- Cag Y, Karabay O, Sipahi OR, et al. Development and validation of a modified quick SOFA scale for risk assessment in sepsis syndrome. *PLoS One* 2018;13(9).
- Chen R, Liang W, Jiang M, et al. Risk Factors of Fatal Outcome in Hospitalized Subjects With Coronavirus Disease 2019 From a Nationwide Analysis in China. *Chest* 2020;158(1):97–105.
- Conti P, Caraffa A, Gallenga C, et al. The British variant of the new coronavirus-19 (Sars-Cov-2) should not create a vaccine problem. *Journal of Biological Regulators Homeostatic Agents* 2021;35(1).
- Das AK, Mishra S, Gopalan SS. Predicting CoVID-19 community mortality risk using machine learning and development of an online prognostic tool. *Peer J* 2020;8:e10083.
- Emanuel EJ, Persad G, Upshur R, et al. Fair Allocation of Scarce Medical Resources in the Time of Covid-19. *New England Journal of Medicine* 2020;382(21):2049–55.
- Faisal M, Mohammed MA, Richardson D, Fiori M, Beatson K. Development and validation of automated computer aided-risk score for predicting the risk of in-hospital mortality using first electronically recorded blood test results and vital signs for COVID-19 hospital admissions: a retrospective development and validation study. *MedRxiv* 2020.
- Faria NR, Clar IM, Candido D, et al. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. 2021. <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586> (accessed 4 June 2021).
- Favaloro EJ, Thachil J. Reporting of D-dimer data in COVID-19: some confusion and potential for misinformation. *Clinical Chemistry and Laboratory Medicine* 2020;58(8):1191–9.
- Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997;336(4):243–50.
- Fuentes E, Fuentes M, Alarcon M, Palomo I. Immune system dysfunction in the elderly. *Anais da Academia Brasileira de Ciências* 2017;89(1):285–99.
- Fumagalli C, Rozzini R, Vannini M, et al. Clinical risk score to predict in-hospital mortality in COVID-19 patients: a retrospective cohort study. *BMJ Open* 2020;10(9).
- Galloway JB, Norton S, Barker RD, et al. A clinical risk score to identify patients with COVID-19 at high risk of critical care admission or death: an observational cohort study. *Journal of Infection* 2020;81(2):282–8.
- Garibaldi BT, Fiksel J, Muschelli J, et al. Patient Trajectories Among Persons Hospitalized for COVID-19: A Cohort Study. *Annals of Internal Medicine* 2021;174(1):33–41.
- Gavelli F, Castello LM, Bellan M, et al. Clinical stability and in-hospital mortality prediction in COVID-19 patients presenting to the Emergency Department. *Minerva Medica* 2020;1–18.
- Goel S, Jain T, Hooda A, et al. Clinical Characteristics and In-Hospital Mortality for COVID-19 Across The Globe. *Cardiology and Therapy* 2020;9(2):553–9.
- Gomez NFP, Lobo IM, Cremades IG, et al. [Potential biomarkers predictors of mortality in COVID-19 patients in the Emergency Department]. *Rev Esp Quimioter* 2020;33(4):267–73.
- Gregory KE, Radovinsky L. Research strategies that result in optimal data collection from the patient medical record. *Appl Nurs Res* 2012;25(2):108–16.
- Gue YX, Tennyson M, Gao J, Ren S, Kanji R, Gorog DA. Development of a novel risk score to predict mortality in patients admitted to hospital with COVID-19. *Sci Rep* 2020;10(1):21379.
- Gungor B, Atici A, Baycan OF, et al. Elevated D-dimer levels on admission are associated with severity and increased risk of mortality in COVID-19: A systematic review and meta-analysis. *The American Journal of Emergency Medicine* 2021;39:173–9.
- Gupta RK, Marks M, Samuels THA, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: An observational cohort study. *European Respiratory Journal* 2020;56(6).
- Haimovich AD, Ravindra NG, Stoytchev S, et al. Development and validation of the quick COVID-19 severity index: a prognostic tool for early clinical decompensation. *Annals of Emergency Medicine* 2020;76(4):442–53.
- Hajifathalian K, Sharaiha RZ, Kumar S, et al. Development and external validation of a prediction risk model for short-term mortality among hospitalized U.S. COVID-19 patients: A proposal for the COVID-AID risk tool. *PLoS One* 2020;15(9).
- Halalau A, Imam Z, Karabon P, et al. External validation of a clinical risk score to predict hospital admission and in-hospital mortality in COVID-19 patients. *Annals of Medicine* 2021;53(1):78–86.
- Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics* 2019;95.
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics* 2009;42(2):377–81.
- Harrison SL, Fazio-Eynullayeva E, Lane DA, Underhill P, Lip GY. Atrial fibrillation and the risk of 30-day incident thromboembolic events, and mortality in adults \geq 50 years with COVID-19. *Journal of Arrhythmia* 2020(00):1–7.
- Harrison SL, Fazio-Eynullayeva E, Lane DA, Underhill P, Lip GY. Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: A federated electronic medical record analysis. *PLoS Med* 2020;17(9).
- Hu C, Liu Z, Jiang Y, et al. Early prediction of mortality risk among severe COVID-19 patients using machine learning. *MedRxiv* 2020.
- Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* 2020;395(10223):497–506.
- IjBd Nascimento, Tcv Grootte, Mathúna DPO, et al. Clinical, laboratory and radiological characteristics and outcomes of novel coronavirus (SARS-CoV-2) infection in humans: A systematic review and series of meta-analyses. *PLoS One* 2020;15(9).
- Kazemi MA, Ghanaati H, Moradi B, et al. Prognostic factors of chest CT findings for ICU admission and mortality in patients with COVID-19 pneumonia. *MedRxiv* 2020.
- KGM M, DG A, JB R, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of internal medicine* 2015;162(1):W1–W73.
- Kim I-C, Song JE, Lee HJ, et al. The Implication of Cardiac Injury Score on In-hospital Mortality of Coronavirus Disease 2019. *Journal of Korean Medical Science* 2020;35(39):e349.
- Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* 2020;370:m3339.
- Ko H, Chung H, Kang WS, et al. An Artificial Intelligence Model to Predict the Mortality of COVID-19 Patients at Hospital Admission Time Using Routine Blood Samples: Development and Validation of an Ensemble Model. *Journal of Medical Internet Research* 2020;22(12):e25442.
- Koozi H, Lengquist M, Frigyesi A. C-reactive protein as a prognostic factor in intensive care admissions for sepsis: A Swedish multicenter study. *Journal of Critical Care* 2020;56:73–9.
- Leeuwenberg AM, Schuit E. Prediction models for COVID-19 clinical decision making. *The Lancet Digital Health* 2020;2(10):e496–e4e7.
- Lemos JA, McGuire DK, Drazner MH. B-type natriuretic peptide in cardiovascular disease. *The Lancet* 2003;362(9380):316–22.
- Levy TJ, Richardson S, Coppa K, et al. Development and validation of a survival calculator for hospitalized patients with COVID-19. *MedRxiv* 2020.
- Liang W, Liang H, Ou L, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Internal Medicine* 2020;180(8):1081–9.
- Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003;58(5):377–82.
- Liu JL, Xu F, Zhou H, et al. Expanded CURB-65: a new score system predicts severity of community-acquired pneumonia with superior efficiency. *Scientific Reports* 2016;6:22911.
- Lu J, Hu S, Fan R, et al. ACP risk grade: a simple mortality index for patients with confirmed or suspected severe acute respiratory syndrome coronavirus 2 disease (COVID-19) during the early stage of outbreak in Wuhan, China. *Preprints with The Lancet* 2020.
- Marcolino MS, Ziegelmann PK, Souza-Silva MV, et al. Clinical characteristics and outcomes of patients hospitalized with COVID-19 in Brazil: results from the Brazilian COVID-19 Registry. *International Journal of Infectious Diseases* 2021.
- Mei Q, Wang AY, Bryant A, et al. Development and validation of prognostic model for predicting mortality of COVID-19 patients in Wuhan, China. *Sci Rep* 2020;10(1):22451.
- Montalva AS, Nadal JS, Pereira JE, et al. Early outcomes of tocilizumab in adults hospitalized with severe COVID-19. An initial report from the Vall d'Hebron COVID-19 prospective cohort study. *MedRxiv* 2020.
- Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162(1):W1–73.
- Nicholson CJ, Wooster L, Sigurðslid HH, et al. Estimating Risk of Mechanical Ventilation and Mortality Among Adult COVID-19 patients Admitted to Mass General Brigham: The VICE and DICE Scores. *Medrxiv* 2020;17:1–33.
- Nunez-Gil JJ, Fernandez-Perez C, Estrada V, et al. Mortality risk assessment in Spain and Italy, insights of the HOPE COVID-19 registry. *Intern Emerg Med* 2020.
- Olsson T, Terént A, Lind L. Rapid Emergency Medicine Score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. *Journal of Internal Medicine* 2004;255(5):579–87.
- Oprita B, Aignatoaie B, Gabor-Postole D. Scores and scales used in emergency medicine. *Practicability in toxicology. Journal of Medicine and Life* 2014;7(Spec Iss 3):4–7.
- Organization WH. Diagnostic testing for SARS-CoV-2: interim guidance, 11 September 2020. *World Health Organization*; 2020.

- Qu R, Hu L, Ling Y, et al. C-reactive protein concentration as a risk predictor of mortality in intensive care unit: a multicenter, prospective, observational study. *BMC Anesthesiology* 2020;20(292):1–9.
- Rice TW, Wheeler AP, Bernard GR, et al. Comparison of the SpO₂/FIO₂ ratio and the PaO₂/FIO₂ ratio in patients with acute lung injury or ARDS. *Chest* 2007;132(2):410–17.
- Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons; 2004.
- Rubin R. COVID-19 Vaccines vs Variants—Determining How Much Immunity Is Enough. *JAMA* 2021;325(13):1241–3.
- Rufibach K. Use of Brier score to assess binary predictions. *Journal of Clinical Epidemiology* 2010;63(8):938–9.
- Sarkar J, Chakrabarti P. A machine learning model reveals older age and delayed hospitalization as predictors of mortality in patients with COVID-19. *MedRxiv* 2020.
- Shang Y, Liu T, Wei Y, et al. Scoring systems for predicting mortality for severe patients with COVID-19. *E Clinical Medicine* 2020;24.
- Sheehan J, Ho KS, Poon J, Sarosky K, Fung JY. Palliative care in critically ill COVID-19 patients: the early New York City experience. *BMJ Supportive & Palliative Care* 2020;0:1–5.
- Sherwani S, Khan MWA. Cytokine Response in SARS-CoV-2 Infection in the Elderly. *Journal of Inflammation Research* 2020;13:737–47.
- Soto-Mota A, Marfil-Garza BA, Rodriguez EM, et al. The low-harm score for predicting mortality in patients diagnosed with COVID-19: A multicentric validation study. *Am Coll Emerg Physicians Open* 2020;1(6):1436–43.
- Sourij H, Aziz F, Bräuer A, et al. COVID-19 fatality prediction in people with diabetes and prediabetes using a simple score upon hospital admission. *Diabetes, Obesity and Metabolism* 2020;23:589–98.
- Sperrin M, McMillan B. Prediction models for covid-19 outcomes. *The BMJ* 2020;371:m3777.
- Tuty Kuswardhani RA, Henrina J, Pranata R, Anthonius Lim M, Lawrensia S, Suastika K. Charlson comorbidity index and a composite of poor outcomes in COVID-19 patients: A systematic review and meta-analysis. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* 2020;14(6):2103–9.
- Wang J, Zhang H, Qiao R, et al. Thrombo-inflammatory features predicting mortality in patients with COVID-19: The FAD-85 score. *Journal of International Medical Research* 2020;48(9):1–14.
- Weng Z, Chen Q, Li S, et al. ANDC: an early warning score to predict mortality risk for patients with Coronavirus Disease 2019. *Journal of Translational Medicine* 2020;18(328):1–10.
- WHO. Coronavirus disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update. 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/> (accessed 4 June 2021).
- Williams RD, Markus AF, Yang C, et al. Seek COVER: Development and validation of a personalized risk calculator for COVID-19 outcomes in an international network. *MedRxiv* 2020.
- Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Annals of Internal Medicine* 2019;170(1):51–8.
- Wynants L, Calster BV, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *British Medical Journal Publishing Group* 2020;369:m1328.
- Xie J, Hungerford D, Chen H, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. *MedRxiv* 2020.
- Yadav AS, Li YC, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *The Lancet Digit Health* 2020;2(10):e516–ee25.
- Yan L, Zhang H-T, Goncalves J, et al. An interpretable mortality prediction model for COVID-19 patients. *Nat Mach Intell* 2020;2:283–8.
- Yao XH, Li TY, He ZC, et al. [A pathological report of three COVID-19 cases by minimal invasive autopsies]. *Zhonghua Bing Li Xue Za Zhi* 2020;49(5):411–17.
- Yoo E, Percha B, Tomlinson M, et al. Development and calibration of a simple mortality risk score for hospitalized COVID-19 adults. *MedRxiv* 2020.
- Zhang C, Qin L, Li K, et al. A Novel Scoring System for Prediction of Disease Severity in COVID-19. *Front Cell Infect Microbiol* 2020;10:318.
- Zhang H, Shi T, Wu X, et al. Risk prediction for poor outcome and death in hospital in-patients with COVID-19: derivation in Wuhan, China and external validation in London, UK. Preprints with The Lancet 2020.
- Zhang S, Guo M, Duan L, et al. Development and validation of a risk factor-based system to predict short-term survival in adult hospitalized patients with COVID-19: a multicenter, retrospective, cohort study. *Critical Care* 2020;24(1):1–13.
- Zhang S. A Troubling New Pattern Among the Coronavirus Variants. 2021. <https://www.theatlantic.com/health/archive/2021/01/coronavirus-evolving-same-mutations-around-world/617721/> (accessed 22 May 2021).
- Zhou J, Huang L, Chen J, et al. Clinical features predicting mortality risk in older patients with COVID-19. *Current Medical Research and Opinion* 2020;36(11):1753–9.