Title

Development of a Prediction Model for COVID-19 Acute Respiratory Distress Syndrome in Patients With Rheumatic Diseases: Results From the Global Rheumatology Alliance Registry

Journal

ACR Open Rheumatology, 4(10)

ISSN

2578-5745

Authors

Izadi, Zara
Gianfrancesco, Milena A
Aguirre, Alfredo
et al.

Publication Date

2022-10-01

DOI

10.1002/acr2.11481

Peer reviewed

ACR Open Rheumatology
Vol. 0, No. 0, Month 2022, pp 1–11
DOI 10.1002/acr2.11481

American College of Rheumatology
Empowering Rheumatology Professionals

Development of a Prediction Model for COVID-19 Acute Respiratory Distress Syndrome in Patients With Rheumatic Diseases: Results From the Global Rheumatology Alliance Registry

Zara Izadi,[1] Milena A. Gianfrancesco,[1] Alfredo Aguirre,[1] Anja Strangfeld,[2] Elsa F. Mateus,[3] Kimme L. Hyrich,[4] Laure Gossec,[5] Loreto Carmona,[6] Saskia Lawson-Tovey,[7] Lianne Kearsley-Fleet,[8] Martin Schaefer,[9] Andrea M. Seet,[1] Gabriela Schmajuk,[10] Lindsay Jacobsohn,[1] Patricia Katz,[1] Stephanie Rush,[1] Samar Al-Emadi,[11] Jeffrey A. Sparks,[12] Tiffany Y-T Hsu,[12] Naomi J. Patel,[13] Leanna Wise,[14] Emily Gilbert,[15] Alí Duarte-García,[16] Maria O. Valenzuela-Almada,[16] Manuel F. Ugarte-Gil,[17] Sandra Lúcia Euzébio Ribeiro,[18] Adriana de Oliveira Marinho,[19] Lilian David de Azevedo Valadares,[20] Daniela Di Giuseppe,[21] Rebecca Hasseli,[22] Jutta G. Richter,[23] Alexander Pfeil,[24] Tim Schmeiser,[25] Carolina A. Isnardi,[26] Alvaro A. Reyes Torres,[27] Gelsomina Alle,[27] Verónica Saurit,[28] Anna Zanetti,[29] Greta Carrara,[29] Julien Labreuche,[30] Thomas Barnetche,[31] Muriel Herasse,[32] Samira Plassart,[32] Maria José Santos,[33] Ana Maria Rodrigues,[34] Philip C. Robinson,[35] Pedro M. Machado,[36] Emily Sirotich,[37] Jean W. Liew,[38] Jonathan S. Hausmann,[39] Paul Sufka,[40] Rebecca Grainger,[41] Suleman Bhana,[42] Wendy Costello,[43] Zachary S. Wallace,[13] and Jinoos Yazdany,[1] on behalf of the Global Rheumatology Alliance Registry

Objective. Some patients with rheumatic diseases might be at higher risk for coronavirus disease 2019 (COVID-19) acute respiratory distress syndrome (ARDS). We aimed to develop a prediction model for COVID-19 ARDS in this population and to create a simple risk score calculator for use in clinical settings.

Methods. Data were derived from the COVID-19 Global Rheumatology Alliance Registry from March 24, 2020, to May 12, 2021. Seven machine learning classifiers were trained on ARDS outcomes using 83 variables obtained at COVID-19 diagnosis. Predictive performance was assessed in a US test set and was validated in patients from four countries with independent registries using area under the curve (AUC), accuracy, sensitivity, and specificity. A simple risk score calculator was developed using a regression model incorporating the most influential predictors from the best performing classifier.

Results. The study included 8633 patients from 74 countries, of whom 523 (6%) had ARDS. Gradient boosting had the highest mean AUC (0.78; 95% confidence interval [CI]: 0.67-0.88) and was considered the top performing classifier. Ten predictors were identified as key risk factors and were included in a regression model. The regression model that predicted ARDS with 71% (95% CI: 61%-83%) sensitivity in the test set, and with sensitivities ranging from 61% to 80% in countries with independent registries, was used to develop the risk score calculator.

Conclusion. We were able to predict ARDS with good sensitivity using information readily available at COVID-19 diagnosis. The proposed risk score calculator has the potential to guide risk stratification for treatments, such as monoclonal antibodies, that have potential to reduce COVID-19 disease progression.

INTRODUCTION

Acute respiratory distress syndrome (ARDS), affecting about 5% of patients with coronavirus disease 2019 (COVID-19) (1) and one third of hospitalized patients (2), is a life-threatening complication of the severe acute respiratory syndrome coronavirus 2 infection. ARDS in the setting of COVID-19 has a mortality rate of 26% to 62% in people admitted to a critical care setting and 66% to

The views expressed herein are those of the authors and participating members of the COVID-19 Global Rheumatology Alliance and do not necessarily represent the views of the American College of Rheumatology (ACR), the European Alliance of Associations for Rheumatology (EULAR), the UK National Health Service (NHS), the National Institute for Health Research (NIHR), or the UK Department of Health.

94% in patients who received mechanical ventilation (3). ARDS frequently causes long-lasting effects beyond hospitalization, from cognitive impairment to physical weakness (4). Given the high mortality and long-term consequences of ARDS, and the direct burden on the health care system, identification of patients at risk for this complication and use of potentially mitigating treatment strategies are important.

There is controversy regarding the existence of an increased risk of severe COVID-19 outcomes in people with rheumatic diseases (5–8). For example, reports from a Swedish nationwide study showed that the risks of COVID-19-related hospitalization and death (but not intensive care unit [ICU] admission) were increased in rheumatoid arthritis (RA), whereas for other inflammatory joint diseases, only the risk of COVID-19-related hospitalization were increased, compared with population referents. However, these risks were comparable to the increased risk of all-cause hospitalization in patients with rheumatic diseases and the increased all-cause mortality risk in patients with RA, and the increased mortality risk in 2020 in patients with RA was not different from that in 2015-2019 (5). In the United States, a multi-institutional electronic health record (EHR) study found higher risks of hospitalization, ICU admission, acute renal failure, and venous thromboembolism (but not death) in patients with rheumatic diseases compared with matched controls (9). Another study conducted at a multi-institutional health system among patients admitted to the hospital with COVID-19 showed higher odds of admission to intensive care and of mechanical ventilation in patients with rheumatic diseases compared with matched controls (10).

The risk factors most strongly associated with ARDS, the key life-threatening organ involvement in COVID-19, are not yet identified. Predicting ARDS using information available at the time of COVID-19 diagnosis has the potential to guide clinical risk stratification and the management of COVID-19 in this population. Because ARDS is a relatively rare event in people who develop COVID-19, there are special considerations in developing statistical models predicting this outcome. Prediction using traditional regression methods can lead to overfitting, limiting the number of predictors that can reliably be used in the prediction model (11). In addition, regression models typically limit the link between outcome and predictor variables to be linear and additive; as a result, regression models may fail to adequately represent complex interactions and high-dimensional relationships that may be present in patients with rheumatic diseases (12). Machine learning algorithms provide an alternative approach with the potential to improve predictive performance, in particular sensitivity, in the setting of relatively rare events, such as ARDS.

We acknowledge financial support from the ACR and EULAR. The ACR and EULAR were not involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

[1]Zara Izadi, MPharm, PhD, Milena A. Gianfrancesco, MPH, PhD, Alfredo Aguirre, MD, Andrea M. Seet, MPH, Lindsay Jacobsohn, BA, Patricia Katz, PhD, Stephanie Rush, BA, Jinoos Yazdany, MD, MPH: University of California, San Francisco; [2]Anja Strangfeld, MD, PhD: Deutsches Rheuma-Forschungszentrum Berlin, Berlin, Germany; [3]Elsa F. Mateus, PhD: Portuguese League Against Rheumatic Diseases, Lisbon, Portugal; [4]Kimme L. Hyrich, MD, PhD: The University of Manchester and National Institute for Health Research Manchester Biomedical Research Centre, Manchester University and NHS Foundation Trust, Manchester, UK; [5]Laure Gossec, MD, PhD: INSERM, Sorbonne Universite and Hopital Universitaire Pitie Salpetriere, AP-HP, Paris, France; [6]Loreto Carmona, MD, PhD: Instituto de Salud Musculoesquelética, Madrid, Spain; [7]Saskia Lawson-Tovey, BA: The University of Manchester and National Institute for Health Research Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust and Manchester Academic Health Science Centre, Manchester, UK; [8]Lianne Kearsley-Fleet, PhD: The University of Manchester and Manchester Academic Health Science Centre, Manchester, UK; [9]Martin Schaefer, PhD: German Rheumatism Research Center, Berlin, Germany; [10]Gabriela Schmajuk, MD, MS: University of California, San Francisco and San Francisco Department of Veterans Affairs Medical Center; [11]Samar Al-Emadi, MBBS, FRCPC, FACR: Hamad Medical Corporation, Doha, Qatar; [12]Jeffrey A. Sparks, MD, MMSc, Tiffany Y-T Hsu, MD, PhD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [13]Naomi J. Patel, MD, Zachary S. Wallace, MD, MSc: Massachusetts General Hospital and Harvard Medical School, Boston; [14]Leanna Wise, MD, MPH: University of Southern California, Los Angeles; [15]Emily Gilbert, MD, PhD: Mayo Clinic, Jacksonville, Florida; [16]Alí Duarte-García, MD, MSc, Maria O. Valenzuela-Almada, MD: Mayo Clinic, Rochester, Minnesota; [17]Manuel F. Ugarte-Gil, MD: Universidad Científica del Sur and Hospital Nacional Guillermo Almenara Irigoyen, EsSalud, Lima, Peru; [18]Sandra Lúcia Euzébio Ribeiro, MD, PhD: Universidade Federal do Amazonas, Manaus, Brazil; [19]Adriana de Oliveira Marinho, MD: Fundação Hospitalar do Acre, Rio Branco, Brazil; [20]Lilian David de Azevedo Valadares, MD: Universidade Federal de Pernambuco, Recife, Brazil; [21]Daniela Di Giuseppe, PhD: Karolinska Institutet, Stockholm, Sweden; [22]Rebecca Hasseli, MD: Justus-Liebig University Giessen, Campus Kerckhoff, Giessen, Germany; [23]Jutta G. Richter, MD: Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany; [24]Alexander Pfeil, MD: Jena University Hospital and Friedrich Schiller University Jena, Jena, Germany; [25]Tim Schmeiser, MD: Rheumatology im Veedel (Private Practice), Cologne, Germany; [26]Carolina A. Isnardi, MD: Argentine Society of Rheumatology, Buenos Aires, Argentina; [27]Alvaro A. Reyes Torres, MD, Gelsomina Alle, MD: Hospital Italiano de Buenos Aires, Buenos Aires, Argentina; [28]Verónica Saurit, MD: Hospital Privado Universitario de Córdoba, Córdoba, Argentina; [29]Anna Zanetti, MSc, Greta Carrara, MSc: Italian Society for Rheumatology and University of Milano-Bicocca, Milan, Italy; [30]Julien Labreuche, MD: Centre Hospitalier Universitaire de Lille, Lille, France; [31]Thomas Barnetche, MD: FHU ACRONIM, Centre for Autoimmune Systemic Rare Diseases, Bordeaux University Hospital, Bordeaux, France; [32]Muriel Herasse, PhD, Samira Plassart, PhD: Filière des Maladies Autoimmunes et Autoinflammatoires Rares, Hôpital Huriez, Centre Hospitalier Universitaire de Lille, Lille, France; [33]Maria José Santos, MD, PhD: Hospital Garcia de Orta, Almada, Portugal, and Instituto de Medicina Molecular Faculdade Medicina and Rheumatic Diseases Portuguese Register, Lisbon, Portugal; [34]Ana Maria Rodrigues, MD: Rheumatic Diseases Portuguese Register, Sociedade Portuguesa de Reumatologia, Nova Medical School, and Hospital dos Lusiadas, Lisbon, Portugal; [35]Philip C. Robinson, MBChB, PhD: The University of Queensland, Brisbane, Queensland, Australia, and Royal Brisbane and Women's Hospital, Metro North Hospital and Health Service, Herston, Queensland, Australia; [36]Pedro M. Machado, MD, PhD: University College London, University College London Hospitals NHS Foundation Trust and Northwick Park Hospital, London North West University Healthcare NHS Trust, London, UK; [37]Emily Sirotich, BSc: McMaster University, Hamilton, Ontario, Canada, and Canadian Arthritis Patient Alliance, Toronto, Ontario, Canada; [38]Jean W. Liew, MD, MS: Boston University School of Medicine, Boston, Massachusetts; [39]Jonathan S. Hausmann, MD: Beth Israel Deaconess Medical Center, Harvard Medical School and Boston Children's Hospital, Boston, Massachusetts; [40]Paul Sufka, MD: HealthPartners, St. Paul, Minnesota; [41]Rebecca Grainger, MBChB, BMedSci, PhD: University of Otago, Wellington, Wellington, New Zealand; [42]Suleman Bhana, MD, FACR: Pfizer Inc., New York, New York; [43]Wendy Costello: Irish Children's Arthritis Network, Tipperary, Ireland.

Author disclosures are available at https://onlinelibrary.wiley.com/action/\downloadSupplement?doi=10.1002%2Facr2.11481&file=acr211481-sup-0001-Disclosureform.pdf.

Address correspondence to Zara Izadi, MPharm, MAS, University of California, San Francisco, Department of Epidemiology and Biostatistics, 2nd floor, 550 16th Street, San Francisco, CA 94158. Email: zara.izadi@ucsf.edu.

Submitted for publication April 21, 2022; accepted May 31, 2022.

This study aimed to develop a prediction model for ARDS in individuals with COVID-19 and pre-existing rheumatic diseases using information obtained at the time of COVID-19 diagnosis and a series of machine-learning algorithms for predictor selection. An additional aim was to develop a simple and interpretable risk-score calculator for potential use in clinical settings.

MATERIALS AND METHODS

The dataset for this study, the COVID-19 Global Rheumatology Alliance (GRA) provider registry, contains only limited data; no personal identifiers, with the exception of COVID-19 diagnosis dates, are included. Because of the limited data and the noninterventional nature of the study, the GRA registry was determined to be nonhuman subjects research by the UK Health Research Authority, the University of Manchester, and the University of California, San Francisco. An institutional review board or ethics committee approval or informed consent was therefore not required.

Study design. This study used data from the COVID-19 GRA registry (13) from March 24, 2020, to May 12, 2021. Briefly, data from adults with rheumatic diseases diagnosed with COVID-19 are entered by rheumatology clinicians via one of two parallel international data entry portals: one (14) limited to European countries and a second (15) for the rest of the world. Five countries in Europe—France (8,16,17), Germany (18–20), Italy (21), Portugal (22,23), and Sweden (24)—and two countries in South America—Brazil (25,26) and Argentina (27)—host national registries supported by their respective national societies. National data from these countries are regularly transferred and merged into the GRA registry. Although GRA data largely depend on convenience sampling, rheumatology practices from two large health systems within the United States (Mass General Brigham in Massachusetts and Mayo Clinic in Minnesota and Florida) have processes in place to systematically report all symptomatic and asymptomatic COVID-19 diagnoses, irrespective of COVID-19 severity.

Patient demographics, rheumatic disease characteristics, comorbidities, COVID-19 outcomes, and complications are entered by reporting clinicians. Methods of COVID-19 diagnosis are indicated, including one or more of the following: polymerase chain reaction, antigen testing, antibody testing, metagenomic testing, computed topography scan, laboratory assay, or a presumptive diagnosis based on symptoms or close contact alone. Quality is assessed by data validation teams who remove all known or potential duplicates and address erroneous or ineligible reports. We followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis statement for prediction model development and validation (28).

Inclusion and exclusion criteria. We included patients with a reconciled status only, defined as the highest COVID-19 illness severity level being confirmed. Patients with a COVID-19 diagnosis date that preceded January 1, 2020, were excluded (n = 7). Additionally, we excluded patients with missing data on ARDS or any of the predictor variables (Supplementary Table 1). Patients reported from France, Portugal, and Germany were excluded because of the unavailability of data on ARDS or smoking status.

Outcome. ARDS was the outcome and the event being predicted in this study. A diagnosis of COVID-19–related ARDS was indicated by the reporting clinician at the point of data entry and in almost all cases reflected a diagnosis given to the patient by the inpatient team (eg, pulmonologists, critical care specialists, or internists directly caring for the patient).

Predictors. ARDS was predicted using 83 predictor variables related to patient demographics, rheumatic disease diagnoses and activity, immunomodulatory medications used for the treatment of rheumatic disease, and comorbidities (Supplementary Table 2). All variables reflect data at the time of COVID-19 diagnosis.

Training, test, and validation sets. Construction of the training, test, and validation sets is depicted in Figure 1. Patients reported from the United States (except those reported from Mass General Brigham in Massachusetts and Mayo Clinic in Minnesota and Florida) and all other countries that directly reported to the GRA registry were included in the training set (n = 5673). The test set comprised all patients reported from Mass General Brigham in Massachusetts and Mayo Clinic in Minnesota and Florida (n = 891). We used this approach to address any potential for provider reporting bias and to improve the generalizability of our findings by testing on a subset of data that most closely represent the underlying spectrum of COVID-19 severity among patients with pre-existing rheumatic diseases. Additionally, patients reported from these health systems had low rates of missing data (<10% of patients excluded because of incomplete data) permitting complete-case analyses. Patients reported from countries with independent registries were used as validation sets (n = 2069). We used four validation sets in total, corresponding to patients reported from Italy (n = 1060), Sweden (n = 225), Brazil (n = 201), and Argentina (n = 583). The amount of missing data varied considerably between the validation sets (Supplementary Table 1). Italy had the lowest rates of missing data (<10% of patients excluded because of incomplete data) and was therefore considered the primary validation set.

Prediction algorithms. Because ARDS is relatively rare and many predictors are potentially relevant to predicting this severe outcome, we used a machine learning approach for predictor selection, which is suited to data with high dimensionality. To identify the most important predictors of ARDS, we compared predictive performance of seven supervised machine learning

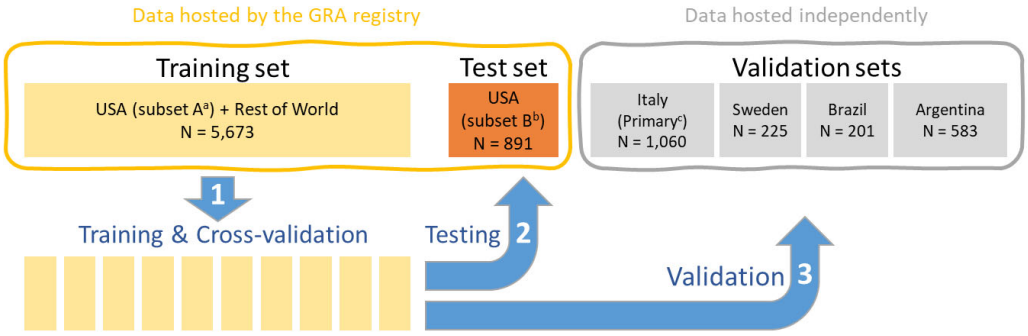


Figure 1. Data set partitioning into training, test, and validation sets. 1) Seven supervised machine learning algorithms were trained on acute respiratory distress syndrome outcomes using three repeats of 10-fold cross-validation. 2) Predictive performance was assessed in the test set. 3) Predictive performance was further assessed in the validation sets. [a]Subset A included all patients reported from the United States, except patients reported from Mass General Brigham in Massachusetts and Mayo Clinics in Minnesota and Florida. [b]Subset B included all patients reported from Mass General Brigham in Massachusetts and Mayo Clinics in Minnesota and Florida. These health systems systematically reported all coronavirus disease 2019 (COVID-19) diagnoses, irrespective of severity. [c]Italy had the lowest rates of unknown data (<7% in any variable and <10% of patients excluded because of incomplete data) among all validation sets and was therefore considered the primary validation set. GRA, Global Rheumatology Alliance.

classifiers commonly applied in the setting of rare clinical outcomes (29). The classifiers were trained on ARDS outcomes using three repeats of 10-fold cross-validation. Prediction algorithms used instance-based learning (k-nearest neighbors and support vector machines), regularization (the lasso and elastic-net regularized generalized linear models), Bayesian regression (Bayesian generalized linear models), additive models (generalized additive models), ensemble learning (gradient boosting machines [GBM]), and deep learning (neural networks). All analyses were performed in R version 3.6.1, using the Classification and Regression Training (30) package.

Model performance. Model performance was assessed using accuracy, sensitivity, specificity, and area under the curve (AUC). The prediction algorithm with the highest AUC in the test set was selected as the best performing classifier. AUC is an aggregate measure of the receiver operating characteristic curve and, unlike accuracy, does not depend on a classification threshold value. For each prediction algorithm and data set, separately, classification threshold values were selected to reduce the absolute difference between sensitivity and specificity (31). This approach was taken to maximize both metrics (Supplementary Figure 1) and to account for potential country-level differences in health system structure, health care access, and use. Mean classification thresholds, mean performance metrics, and corresponding 95% confidence intervals (CIs) were derived from 1000 samples of 500 randomly selected patients from the test set and each validation set using bootstrapping and sampling with replacement.

Risk score calculator development. The risk score calculator was derived from a multivariable logistic regression incorporating the most influential predictors from the best performing classifier (32). We used logistic regression to develop a risk score calculator that was interpretable, user friendly, and readily accessible for potential use in clinical settings across health systems. To determine the optimum number of items in the risk score calculator, a series of regressions with varying numbers of predictors (ranging from top five predictors to top n predictors, in which the importance score associated with nth predictor was >0) were trained on ARDS outcomes using 10 repeats of 10-fold cross-validation. To balance the calculator's ease of use in clinical settings (33,34) with predictive performance, our final regression model incorporated the lowest number of predictors associated with the highest mean AUC. To improve regression fit, we assessed linearity in the relationship between continuous predictors and the outcome and accounted for nonlinear relationships using interaction terms. Direction, magnitude, and statistical significance of key risk factors associated with ARDS were reported from the final regression model using adjusted odds ratios (ORs). The predictive performance of the regression model was evaluated in the test set and validation sets using the aforementioned performance metrics and methods. Additionally, we assessed calibration of the regression model by comparing the mean predicted ARDS probabilities with the mean observed probabilities within every decile of predicted risk in the test and validation sets and reported corresponding integrated calibration indices (ICIs) (35).

To aid the interpretation of predicted probabilities, risk of ARDS development was defined as "low" for predicted probabilities lower than the lowest country-specific mean classification threshold, "moderate to high" for the predicted probability region between the highest and the lowest country-specific mean classification thresholds, and "high" for predicted probabilities equal to or higher than the highest country-specific mean classification threshold.

A point-based scoring system was developed in which points were assigned to each item by multiplying each β coefficient (log OR) from the regression model by a constant arbitrary number and rounding (to the nearest integer for points 1-5 and to the nearest fifth integer for points >5) to facilitate total risk score calculation. A total risk score was assigned to each patient by summing the points for each item in the risk score calculator. Mean predicted ARDS probabilities and 95% CIs corresponding to each total risk score within the "moderate to high" category of risk are reported.

Table 1. Demographic and clinical characteristics of the study population

			Validation sets			
			Primary	Other		
	Training set n = 5673	Test set n = 891	Italy n = 1060	Sweden n = 225	Brazil n = 201	Argentina n = 583
Age, years, mean (SD)	53.2 (15.2)	58.0 (17.1)	56.6 (14.6)	53.5 (14.7)	47.8 (14.1)	49.2 (14.2)
Sex, n (%)						
Male	1585 (27.9)	236 (26.5)	311 (29.3)	88 (39.1)	57 (28.4)	126 (21.6)
Female	4088 (72.1)	655 (73.5)	749 (70.7)	137 (60.9)	144 (71.6)	457 (78.4)
Smoking status, n (%)						
Never smoker	4212 (74.2)	543 (60.9)	752 (70.9)	127 (56.4)	188 (93.5)	395 (67.8)
Former smoker	1152 (20.3)	307 (34.5)	199 (18.8)	88 (39.1)	5 (2.5)	154 (26.4)
Current smoker	309 (5.4)	41 (4.6)	109 (10.3)	10 (4.4)	8 (4)	34 (5.8)
Most common diagnoses, n (%)						
Rheumatoid arthritis	2472 (43.6)	322 (36.1)	360 (34)	100 (44.4)	60 (29.9)	299 (51.3)
Psoriatic arthritis	569 (10)	81 (9.1)	220 (20.8)	46 (20.4)	23 (11.4)	47 (8.1)
Spondyloarthritis	554 (9.8)	45 (5.1)	108 (10.2)	40 (17.8)	54 (26.9)	48 (8.2)
Other inflammatory arthritis	145 (2.6)	63 (7.1)	12 (1.1)	12 (5.3)	0 (0)	0 (0)
Systemic lupus erythematosus	689 (12.1)	99 (11.1)	80 (7.5)	5 (2.2)	25 (12.4)	110 (18.9)
Vasculitis	171 (3)	49 (5.5)	40 (3.8)	8 (3.6)	1 (.5)	23 (3.9)
Sjogren syndrome	195 (3.4)	34 (3.8)	29 (2.7)	0 (0)	9 (4.5)	31 (5.3)
Polymyalgia rheumatica	102 (1.8)	47 (5.3)	25 (2.4)	0 (0)	0 (0)	3 (.5)
Systemic sclerosis	165 (2.9)	23 (2.6)	63 (5.9)	1 (.4)	11 (5.5)	20 (3.4)
Disease activity, n (%)						
Remission or low	4554 (80.3)	695 (78)	894 (84.3)	194 (86.2)	166 (82.6)	457 (78.4)
Moderate or high	1119 (19.7)	196 (22)	166 (15.7)	31 (13.8)	35 (17.4)	126 (21.6)
Most common comorbidities, n (%)						
None	2040 (36)	182 (20.4)	315 (29.7)	102 (45.3)	81 (40.3)	317 (54.4)
At least one comorbidity	3633 (64)	709 (79.6)	745 (70.3)	123 (54.7)	120 (59.7)	266 (45.6)
Interstitial lung disease	288 (5.1)	42 (4.7)	70 (6.6)	5 (2.2)	6 (3)	33 (5.7)
Obstructive lung disease	433 (7.6)	145 (16.3)	69 (6.5)	28 (12.4)	6 (3)	9 (1.5)
Obesity	926 (16.3)	273 (30.6)	131 (12.4)	16 (7.1)	26 (12.9)	93 (16)
Diabetes	786 (13.9)	167 (18.7)	102 (9.6)	15 (6.7)	20 (10)	52 (8.9)
Hypertension	1921 (33.9)	412 (46.2)	365 (34.4)	56 (24.9)	67 (33.3)	161 (27.6)
Cardiovascular disease	463 (8.2)	129 (14.5)	169 (15.9)	21 (9.3)	13 (6.5)	19 (3.3)
Chronic kidney disease	274 (4.8)	114 (12.8)	66 (6.2)	3 (1.3)	8 (4)	17 (2.9)
Cancer	191 (3.4)	70 (7.9)	64 (6)	4 (1.8)	4 (2)	12 (2.1)
Liver disease	156 (2.7)	24 (2.7)	66 (6.2)	1 (.4)	0 (0)	8 (1.4)
Neurological or neuromuscular disease	77 (1.4)	40 (4.5)	53 (5)	6 (2.7)	0 (0)	5 (.9)
Psychiatric disease	91 (1.6)	44 (4.9)	27 (2.5)	2 (.9)	2 (1)	22 (3.8)
Psoriasis	291 (5.1)	54 (6.1)	184 (17.4)	13 (5.8)	6 (3)	28 (4.8)
Medications, n (%)						
No DMARDs	939 (16.6)	265 (29.7)	175 (16.5)	13 (5.8)	25 (12.4)	5 (.9)
csDMARDs alone	2501 (44.1)	338 (37.9)	396 (37.4)	77 (34.2)	80 (39.8)	405 (69.5)
b/tsDMARDs alone	1196 (21.1)	193 (21.7)	278 (26.2)	85 (37.8)	64 (31.8)	91 (15.6)
csDMARDs + b/tsDMARDs	1037 (18.3)	95 (10.7)	211 (19.9)	50 (22.2)	32 (15.9)	82 (14.1)
GC use						
No use, n (%)	3942 (69.5)	635 (71.3)	659 (62.2)	172 (76.4)	180 (89.6)	335 (57.5)
GC user, n (%)	1731 (30.5)	256 (28.7)	401 (37.8)	53 (23.6)	21 (10.4)	248 (42.5)
GC dose,[a] mg, median (IQR)	5 (5)	5 (5)	5 (0)	5 (2.5)	10 (5)	5 (5)
ARDS, n (%)						
Yes	355 (6.3)	35 (3.9)	57 (5.4)	12 (5.3)	17 (8.5)	47 (8.1)
No	5318 (93.7)	856 (96.1)	1003 (94.6)	213 (94.7)	184 (91.5)	536 (91.9)

Abbreviations: ARDS, acute respiratory distress syndrome; b/tsDMARDs, biologic or targeted synthetic DMARDs; csDMARDs, conventional systemic DMARDs; DMARD, disease-modifying antirheumatic drug; GC, glucocorticoid; IQR, interquartile range.
[a]Average daily prednisone-equivalent dose among GC users.

RESULTS

Characteristics of the study population. A total of 8633 patients reported from 74 countries were included in the study. Of these, 5673 were partitioned into the training set and 891 and 2069 into the test set and validation sets, respectively, as described in Methods. Among patients composing the training set, the mean (SD) age was 53.2 (15.2) years, 4088 (72.1%) were female, and 4212 (74.2%) were nonsmokers. RA, reported among 2472 (43.6%) individuals, was the most common diagnosis, followed by systemic lupus erythematosus (12.1%) and psoriatic arthritis (10.0%). Treatment with conventional synthetic disease-modifying antirheumatic drugs alone was the most common treatment modality (44.1%). A majority of individuals were in remission or had low disease activity (80.3%; Table 1). ARDS was reported among 355 (6.3%) patients in the training set, 35 (3.9%) patients in the test set, and 57 (5.4%) patients in the primary validation set (Italy). In the other validation sets, the prevalence of ARDS ranged from 3.3% (Sweden) to 8.5% (Brazil).

Predictive performance of machine learning algorithms. Among the seven machine learning classifiers, GBM had the highest AUC in the test set (mean: 0.78; 95% CI: 0.67-0.88) and was considered the top performing model (Supplementary Table 3). In the test set, at the optimum classification threshold, GBM had a mean accuracy, sensitivity, and specificity of 0.70. In the primary validation set, GBM had a mean AUC of 0.79 (95% CI: 0.70-0.87) and a mean accuracy, sensitivity, and specificity of 0.73 at the optimum classification threshold (Supplementary Table 4). In other validation sets, GBM's mean AUC ranged from 0.74 to 0.85, with mean sensitivity and mean specificity ranging from 0.65 to 0.78 and 0.66 to 0.78, respectively. In order of predictor importance, age, average daily prednisone-equivalent glucocorticoid dose, and pulmonary hypertension were the most influential predictors identified by GBM (Supplementary Figure 2).

Important risk factors and risk score calculator. The risk score calculator was derived from a multivariable logistic regression model incorporating the 10 most influential predictors from GBM because 10 was the smallest number of predictors that corresponded to the highest mean AUC (0.77) in cross-validation (Supplementary Materials). Average daily prednisone-equivalent glucocorticoid doses greater than 60 mg were considered clinically high doses. We fitted an interaction term to account for the potential effect modification in dose response in patients receiving glucocorticoid doses greater than 60 mg. The resulting regression was equivalent to a simpler regression that winsorized glucocorticoid

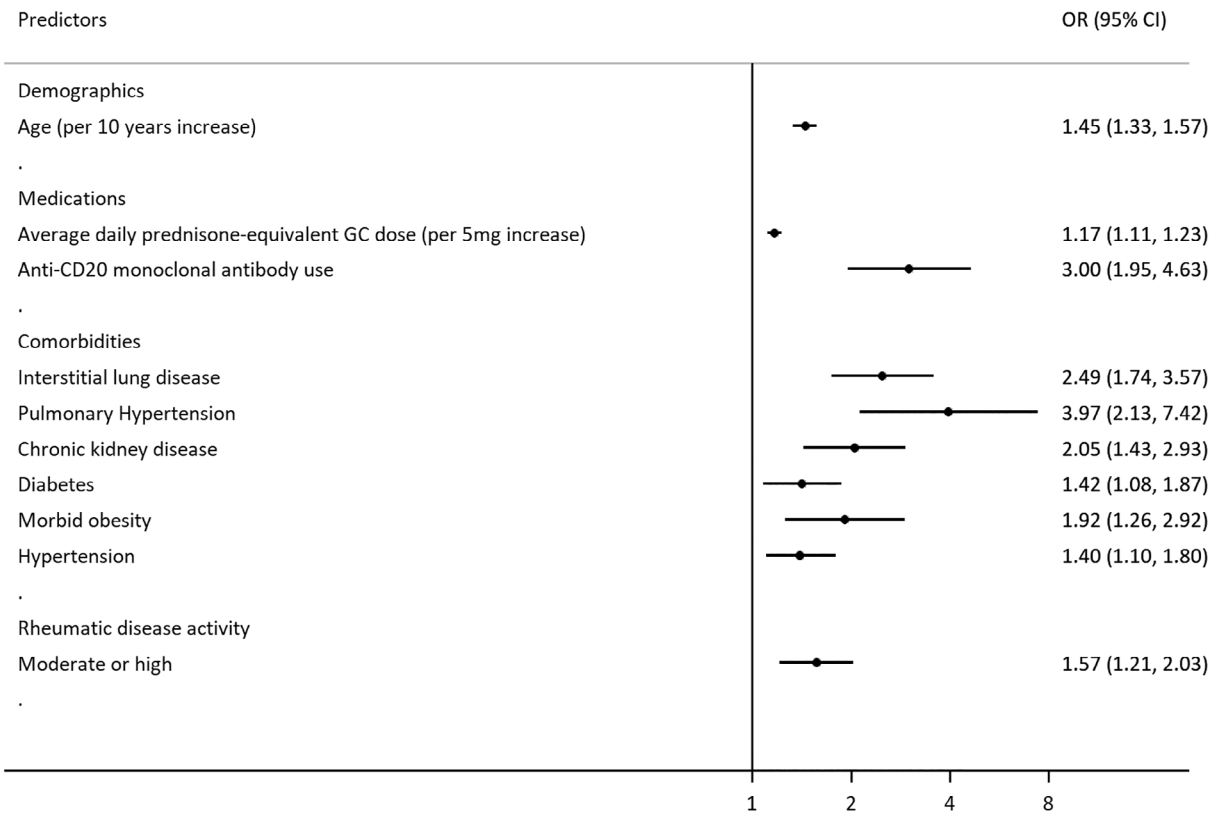


Figure 2. Adjusted odds ratios (ORs) obtained from the multivariable logistic regression model. Top 10 most influential predictors identified by the gradient boosting machine. CI, confidence interval; GC, glucocorticoid.

Table 2. Predictive performance of the multivariable logistic regression model in the test set and across validation sets

	Classification threshold, mean (95% CI)	Percentile of predicted risk[a]	Performance metrics, mean (95% CI)			
			AUC	Accuracy	Sensitivity	Specificity
Test set	0.096 (0.074-0.128)	71.1 (58.8-80.4)	0.79 (0.68-0.88)	0.71 (0.62-0.82)	0.71 (0.61-0.83)	0.71 (0.62-0.82)
Primary validation set						
Italy	0.069 (0.056-0.085)	70.7 (61.9-77.8)	0.77 (0.68-0.86)	0.73 (0.65-0.80)	0.73 (0.64-0.81)	0.73 (0.65-0.80)
Other validation sets						
Sweden	0.058 (0.044-0.081)	70.7 (60.9-81.8)	0.82 (0.72-0.92)	0.74 (0.62-0.84)	0.74 (0.59-0.85)	0.74 (0.62-0.84)
Brazil	0.036 (0.033-0.041)	60.2 (54.2-67.7)	0.71 (0.63-0.78)	0.61 (0.55-0.70)	0.61 (0.52-0.71)	0.61 (0.55-0.70)
Argentina	0.060 (0.053-0.069)	75.5 (70.5-79.9)	0.85 (0.79-0.91)	0.80 (0.75-0.85)	0.80 (0.74-0.86)	0.80 (0.75-0.85)

Abbreviations: AUC, area under curve; CI, confidence interval.
[a]Percentiles of predicted risk correspond to the mean (95% CI) classification thresholds. Classification thresholds, performance metrics, and corresponding 95% CIs were derived from 1000 random samples of 500 patients from each data set using bootstrapping and sampling with replacement.

doses greater than 60 to 60 mg (calibration slope: 0.99 [1.00 indicating perfect calibration]; calibration intercept: 0.00; correlation coefficient: 0.99; $P < 0.0001$). We therefore opted for the simpler regression model in creating the risk score calculator. All 10 predictors were independently and statistically significantly associated with the development of ARDS (Figure 2): older age (OR 1.45; 95% CI: 1.33-1.57, per decade increase in age), higher average daily prednisone-equivalent glucocorticoid doses (1.17; 95% CI: 1.11-1.23, per 5-mg increase in dose), pulmonary hypertension (3.97; 95% CI: 2.13-7.42), interstitial lung disease (2.49; 95% CI: 1.74-3.57), chronic renal insufficiency or end-stage renal disease (2.05; 95% CI: 1.43-2.93), anti-CD20 monoclonal antibody use (3.00; 95% CI: 1.95-4.63), diabetes (1.42; 95% CI: 1.08-1.87), hypertension (1.40; 95% CI: 1.10-1.80), moderate or high rheumatic disease activity (1.57; 95% CI: 1.21-2.03), and morbid obesity (1.92; 95% CI: 1.26-2.92).

Predictive performance of the final regression model was assessed in the test set and each validation set from countries

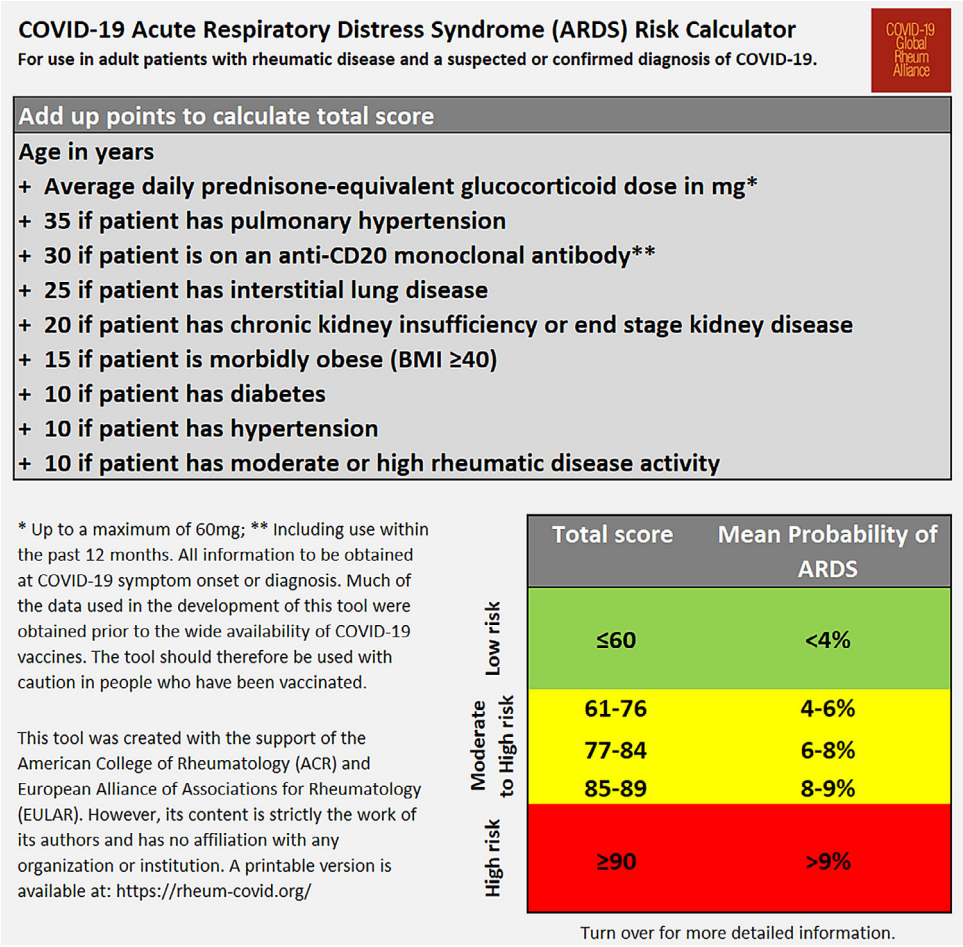
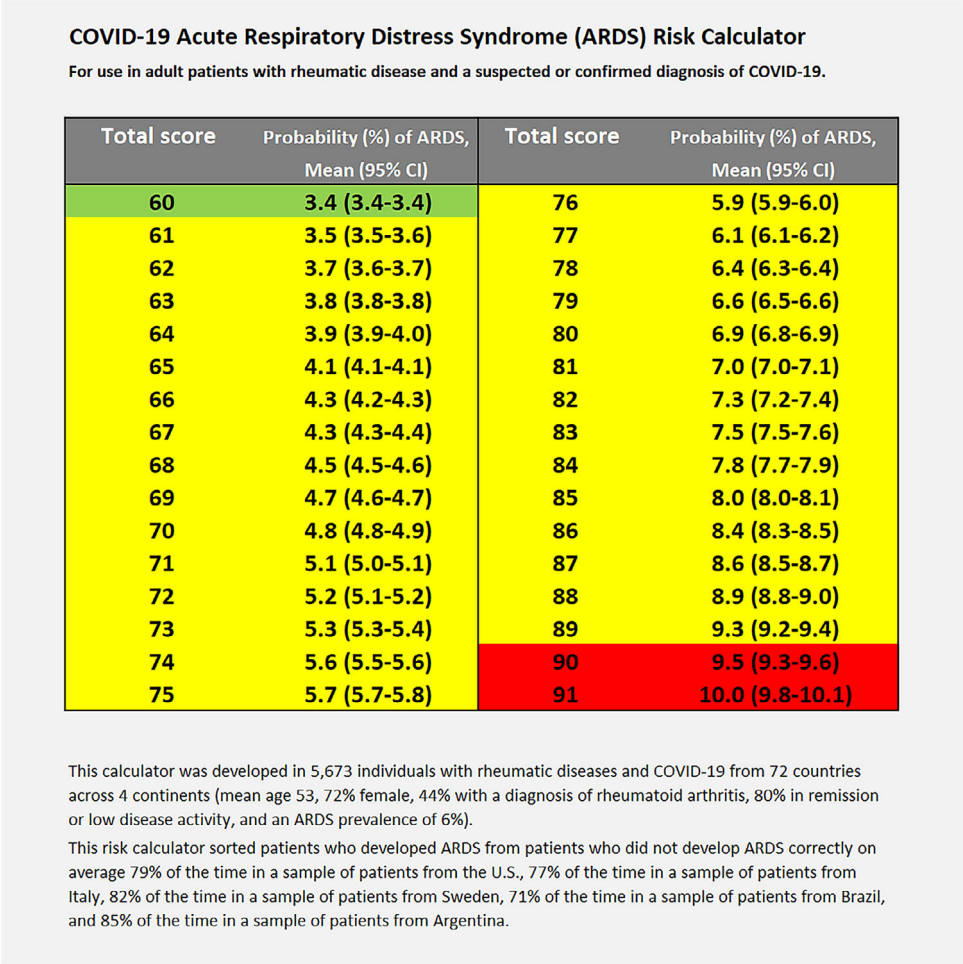


Figure 3. The risk score calculator pocket care side 1. BMI, body mass index; COVID-19, coronavirus disease 2019.



COVID-19 Acute Respiratory Distress Syndrome (ARDS) Risk Calculator

For use in adult patients with rheumatic disease and a suspected or confirmed diagnosis of COVID-19.

Total score	Probability (%) of ARDS, Mean (95% CI)	Total score	Probability (%) of ARDS, Mean (95% CI)
60	3.4 (3.4-3.4)	76	5.9 (5.9-6.0)
61	3.5 (3.5-3.6)	77	6.1 (6.1-6.2)
62	3.7 (3.6-3.7)	78	6.4 (6.3-6.4)
63	3.8 (3.8-3.8)	79	6.6 (6.5-6.6)
64	3.9 (3.9-4.0)	80	6.9 (6.8-6.9)
65	4.1 (4.1-4.1)	81	7.0 (7.0-7.1)
66	4.3 (4.2-4.3)	82	7.3 (7.2-7.4)
67	4.3 (4.3-4.4)	83	7.5 (7.5-7.6)
68	4.5 (4.5-4.6)	84	7.8 (7.7-7.9)
69	4.7 (4.6-4.7)	85	8.0 (8.0-8.1)
70	4.8 (4.8-4.9)	86	8.4 (8.3-8.5)
71	5.1 (5.0-5.1)	87	8.6 (8.5-8.7)
72	5.2 (5.1-5.2)	88	8.9 (8.8-9.0)
73	5.3 (5.3-5.4)	89	9.3 (9.2-9.4)
74	5.6 (5.5-5.6)	90	9.5 (9.3-9.6)
75	5.7 (5.7-5.8)	91	10.0 (9.8-10.1)

This calculator was developed in 5,673 individuals with rheumatic diseases and COVID-19 from 72 countries across 4 continents (mean age 53, 72% female, 44% with a diagnosis of rheumatoid arthritis, 80% in remission or low disease activity, and an ARDS prevalence of 6%).

This risk calculator sorted patients who developed ARDS from patients who did not develop ARDS correctly on average 79% of the time in a sample of patients from the U.S., 77% of the time in a sample of patients from Italy, 82% of the time in a sample of patients from Sweden, 71% of the time in a sample of patients from Brazil, and 85% of the time in a sample of patients from Argentina.

Figure 4. The risk score calculator pocket care side 2. CI, confidence interval; COVID-19, coronavirus disease 2019.

with independent registries. In the test set, the model had a mean AUC of 0.79 (95% CI: 0.68-0.88) and a mean accuracy, sensitivity, and specificity of 0.71 at the optimum classification threshold (Table 2). In the primary validation set, the model had a mean AUC of 0.77 (95% CI: 0.68-0.86) and a mean accuracy, sensitivity, and specificity of 0.73 at the optimum classification threshold. In other validation sets, mean AUCs ranged from 0.71 to 0.85, with both mean sensitivity and mean specificity ranging from 0.61 to 0.80. The calibration plot showed relatively poor agreement between the observed and predicted ARDS risk in the test set (calibration slope: 0.43; intercept: 0.00; ICI: 0.056) and good agreement in the primary validation set (calibration slope: 0.80; intercept: 0.00; ICI: 0.024). The model had relatively poor to moderate calibration in other validation sets, with calibration slopes, intercepts, and ICIs ranging from 1.38 to 1.91, –0.03 to 0.01, and 0.029 to 0.049, respectively (Supplementary Figure 3).

Figures 3 and 4 provide details of the ARDS risk score calculator developed from the multivariable regression model. Predicted ARDS probabilities less than 4% (corresponding to total scores ≤60) were defined as "low" risk, predicted ARDS probabilities between 4% and 9% (corresponding to total scores 61-89) were defined as "moderate to high" risk, and predicted ARDS probabilities greater than 9% were defined as "high" risk. As described in methods, these thresholds were not quantitatively derived but instead reflect probabilities that were felt to be clinically meaningful.

DISCUSSION

In this global sample of patients with rheumatic diseases, we developed a simple ARDS risk score calculator that has the potential for risk stratification and to guide management of COVID-19 among individuals with rheumatic diseases in routine clinical settings. A machine learning classifier, GBM, predicted the onset of ARDS with 70% sensitivity in the test set and with 73% sensitivity in the primary validation set using information obtained at COVID-19 diagnosis. A multivariable regression model using the 10 most influential predictors from GBM predicted ARDS with 71% sensitivity in the test set and with 73% sensitivity in the primary validation set. Rheumatic disease characteristics and medications identified as important risk factors in predicting COVID-19 ARDS align with previously reported factors

associated with COVID-19 hospitalization or death in patients with immune-mediated inflammatory diseases (36–41). Other risk factors, including older age, obesity, chronic lung disease, and chronic kidney disease, were also consistent with risk factors identified by a recently published prognostic model for adverse COVID-19 outcomes using information obtained at diagnosis in a general population-based cohort from Iceland (42).

Our study findings help identify patients with underlying rheumatic diseases who may be at a higher risk for ARDS from COVID-19. Use of baseline information at COVID-19 symptom onset or at COVID-19 exposure or diagnosis in asymptomatic patients facilitates early triage of high-risk patients for monitoring, prophylaxis, or treatment interventions. For example, with the recent Food and Drug Administration Emergency Use Authorizations (43,44) for the use of monoclonal antibodies for the treatment of ambulatory patients with COVID-19 or as postexposure prophylaxis for high-risk individuals exposed to the virus, a risk calculator coupled with clinical judgment can prioritize which patients are most likely to derive benefit from this therapy. Our findings also identify potentially modifiable risk factors that rheumatologists can consider when making patient care decisions to minimize the risk of adverse COVID-19 outcomes, namely, glucocorticoid dose, rituximab use (45), and rheumatic disease activity.

With only 10 items, the proposed calculator is simple to use and can be easily implemented in clinical settings. Additionally, information required for the scoring system is available in both outpatient and inpatient settings or even remotely without the need for close contact, which is not the case with existing ARDS prediction models that require physical examination, laboratory measurements, and imaging data (46–48). In classification, there is typically an inverse relationship between sensitivity and specificity. In this study, we selected classification thresholds that maximized both sensitivity and specificity by minimizing the absolute difference between them. This choice is somewhat arbitrary; in practice, the trade-off between specificity and sensitivity must account for the underlying population risk for ARDS, health gains from available treatment or monitoring interventions, and the regional health system structure that governs the availability and access to health resources.

With the exception of Brazil, both GBM and the GBM-based regression model performed slightly better in validation sets than the test set. This may be explained by the fact that the training set was more similar in nature to the validation sets than the test set, such that provider reporting bias affecting the training set was of a similar magnitude of the bias affecting the validation sets. It is plausible that rheumatology practices that systematically reported all COVID-19 diagnoses and composed the test set also captured information on important risk factors, such as comorbidities, more comprehensively than practices that composed the training set and validation sets. Calibration plots support this hypothesis: predicted probabilities of ARDS were higher than the observed risk in the test set, whereas they were largely comparable in Italy, Sweden, and Argentina and lower than the observed risk in Brazil. Without processes in place to systematically report all COVID-19 diagnoses and capture complete information on baseline characteristics, it is possible that provider reporting patterns were influenced by COVID-19 severity, provider perceptions of factors related to COVID-19 severity, and availability of information through direct interactions between the patient and their rheumatologist during the pandemic. Other factors that may have impacted calibration include nature of the institutions included in the test versus training sets, capabilities of care teams (eg, presence of dedicated COVID-19 care teams and units), and characteristics of the study populations beyond those that we were able to account for. Furthermore, patients may underreport important social and behavioral factors, such as smoking. This social desirability bias can vary across countries and cultures (49) and may additionally explain the discordances observed in predictive performance.

This study has important strengths. First, to our knowledge, this is the first study predicting COVID-19 ARDS among individuals with rheumatic diseases. Second, the prediction models were trained on a global sample of individuals with rheumatic disease, thus increasing the heterogeneity and likely generalizability of patient characteristics. This has the potential to improve prediction accuracy by increasing the number of potential predictors and accounting for complex high-dimensional relationships between them. Importantly, active rheumatic disease status was captured as a predictor. The registry is unique among other data sources in rheumatic diseases in being able to capture data on disease activity that are not typically available in administrative data or EHRs. Furthermore, reporting occurred directly via rheumatology clinicians, which likely increased the accuracy of the information. Third, we tested the performance of prediction models in a subset of practices that had processes in place to minimize potential provider reporting bias. Maximizing the heterogeneity of COVID-19 outcomes in the test set improves the generalizability of our findings to the target population of individuals with pre-existing rheumatic diseases with COVID-19. Fourth, the external validity of our prediction models was assessed using external data sets from Europe and Latin America.

Limitations of this work include potential provider reporting bias and missing data in the training set and external validation sets; the tool should therefore be used with caution outside the United States. Assessments of calibration showed relatively poor agreement between observed and predicted probabilities of ARDS in the test set and in external validation sets; we therefore recommend that the tool be used as a guide for COVID-19 prognosis and in conjunction with clinical judgement. Although we attempted to account for country-level differences in health system structure, health care access, and use through optimizing ARDS classification thresholds at the regional level, a residual impact by these factors may remain. Additionally, we were unable

to account for other important clinical, sociodemographic, or environmental factors, such as the continuation or withholding of rheumatic disease treatments at the time of COVID-19 diagnosis, race and ethnicity, alcohol consumption, occupation, poverty, housing conditions, or air pollution, all of which may influence the outcomes of COVID-19, including the development of ARDS. Much of the data were obtained prior to the wide availability of COVID-19 vaccines, which may lower risk of developing severe COVID-19 outcomes, such as ARDS. However, vaccinated patients with COVID-19 with rheumatic diseases have been reported to experience breakthrough infection possibly because of inadequate humoral vaccine immune response associated with some immunosuppressors (50).

In summary, a GBM-based regression model predicted COVID-19 ARDS with good sensitivity and specificity in patients with pre-existing rheumatic diseases using demographics and basic clinical characteristics that can be easily obtained at COVID-19 exposure or onset. Prediction accuracies were largely comparable or better in external data sets from four countries that hosted independent COVID-19 registries. Age, daily glucocorticoid dose, pulmonary hypertension, interstitial lung disease, chronic kidney disease, anti-CD20 monoclonal antibody use, diabetes, hypertension, active rheumatic disease, and morbid obesity were the most influential factors in predicting the onset of ARDS. Further studies including vaccinated individuals and more recent COVID-19 variants (such as omicron) are needed to prospectively evaluate the clinical utility of the proposed risk score calculator for its potential to guide risk stratification, prophylaxis with monoclonal antibodies, and treatment of COVID-19 in high-risk patients with rheumatic diseases.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Izadi had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Izadi.
Acquisition of data. All authors.
Analysis and interpretation of data. Izadi, Gianfrancesco, Yazdany.

REFERENCES

1. World Health Organization. Clinical management of COVID-19: interim guidance, 27 May 2020. Geneva: World Health Organization; 2020.
2. Tzotzos SJ, Fischer B, Fischer H, Zeitlinger M. Incidence of ARDS and outcomes in hospitalized patients with COVID-19: a global literature survey. Crit Care 2020;24:516.
3. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med 2020;180:934–43.
4. Mart MF, Ware LB. The long-lasting effects of the acute respiratory distress syndrome. Expert Rev Respir Med 2020;14:577–86.
5. Bower H, Frisell T, Di Giuseppe D, Delcoigne B, Ahlenius GM, Baecklund E, et al. Impact of the COVID-19 pandemic on morbidity and mortality in patients with inflammatory joint diseases and in the general population: a nationwide Swedish cohort study. Ann Rheum Dis 2021;80:1086–93.
6. Cordtz R, Lindhardsen J, Soussi BG, Vela J, Uhrenholt L, Westermann R, et al. Incidence and severeness of COVID-19 hospitalisation in patients with inflammatory rheumatic disease: a nationwide cohort study from Denmark. Rheumatology (Oxford) 2021;60: SI59–67.
7. Peach E, Rutter M, Lanyon P, Grainge MJ, Hubbard R, Aston J, et al. Risk of death among people with rare autoimmune diseases compared with the general population in England during the 2020 COVID-19 pandemic. Rheumatology (Oxford). 2021;60:1902–9.
8. FAIR/SFR/SNFMI/SOFREMIP/CRI/IMIDIATE consortium and contributors. Severity of COVID-19 and survival in patients with rheumatic and inflammatory diseases: data from the French RMD COVID-19 cohort of 694 patients. Ann Rheum Dis 2021;80:527–38.
9. D'Silva KM, Jorge A, Cohen A, McCormick N, Zhang Y, Wallace ZS, et al. COVID-19 outcomes in patients with systemic autoimmune rheumatic diseases compared to the general population: a US multicenter, comparative cohort study. Arthritis Rheumatol 2021;73: 914–20.
10. Hsu TY, D'Silva KM, Patel NJ, Wang J, Mueller AA, Fu X, et al. Laboratory trends, hyperinflammation, and clinical outcomes for patients with a systemic rheumatic disease admitted to hospital for COVID-19: a retrospective, comparative cohort study. Lancet Rheumatol 2021;3:e638–47.
11. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.
12. Zhang Z, Zhao Y, Canes A, Steinberg D, Lyashevska O, written on behalf of AME Big-Data Clinical Trial Collaborative Group. Predictive analytics with gradient boosting in clinical medicine. Ann Transl Med 2019;7:152.
13. Gianfrancesco MA, Hyrich KL, Gossec L, Strangfeld A, Carmona L, Mateus EF, et al. Rheumatic disease and COVID-19: initial data from the COVID-19 Global Rheumatology Alliance provider registries. Lancet Rheumatol 2020;2:e250–3.
14. EULAR COVID-19 Registry. 2020. URL: https://www.eular.org/eular_covid_19_registry.cfm.
15. COVID-19 Global Rheumatology Alliance. 2020. URL: https://rheum-covid.org/.
16. Avouac J, Drumez E, Hachulla E, Seror R, Georgin-Lavialle S, El Mahou S, et al. COVID-19 outcomes in patients with inflammatory rheumatic and musculoskeletal diseases treated with rituximab: a cohort study. Lancet Rheumatol 2021;3:e419–26.
17. Trefond L, Drumez E, Andre M, Costedoat-Chalumeau N, Seror R, Devaux M, et al. Impact of hydroxychloroquine used as DMARD on SARS-CoV-2 tests and infection evolution in a population of 871 patients with inflammatory rheumatic and musculoskeletal diseases. Joint Bone Spine 2021;88:105226.
18. Hasseli R, Mueller-Ladner U, Schmeiser T, Hoyer BF, Krause A, Lorenz HM, et al. National registry for patients with inflammatory rheumatic diseases (IRD) infected with SARS-CoV-2 in Germany (ReCoVery): a valuable mean to gain rapid and reliable knowledge of the clinical course of SARS-CoV-2 infections in patients with IRD. RMD Open 2020;6:e001332.
19. Hasseli R, Mueller-Ladner U, Hoyer BF, Krause A, Lorenz HM, Pfeil A, et al. Older age, comorbidity, glucocorticoid use and disease activity are risk factors for COVID-19 hospitalisation in patients with inflammatory rheumatic and musculoskeletal diseases. RMD Open 2021;7: e001464.

20. Hasseli R, Pfeil A, Hoyer BF, Krause A, Lorenz HM, Richter JG, et al. Do patients with rheumatoid arthritis show a different course of COVID-19 compared to patients with spondyloarthritis? [Full paper]. Clin Exp Rheumatol 2021;39:639–47.
21. Scirè CA, Carrara G, Zanetti A, Landolfi G, Chighizola C, Alunno A, et al. COVID-19 in rheumatic diseases in Italy: first results from the Italian registry of the Italian Society for Rheumatology (CONTROL-19). Clin Exp Rheumatol 2020;38:748–53.
22. Canhão H, Faustino A, Martins F, Fonseca JE, Rheumatic Diseases Portuguese Register Board Coordination, Portuguese Society of Rheumatology. Reuma.pt: the rheumatic diseases Portuguese register. Acta Rheumatol Port 2011;36:45–56.
23. Santos M, Canhão H, Mourão A, Ramos FO, Ponte C, Duarte C, et al. Reuma.pt contribution to the knowledge of immune-mediated systemic rheumatic diseases. Acta Rheomatol Port 2017;42:232–9.
24. The Swedish Rheumatology Quality Register. URL: https://srq.nu/en/welcome-health-professional/.
25. Marques CD, Kakehasi AM, Pinheiro MM, Mota LM, Albuquereque CP, Silva CR, et al. High levels of immunosuppression are related to unfavourable outcomes in hospitalised patients with rheumatic diseases and COVID-19: first results of ReumaCoV Brasil registry. RMD Open 2021;7:e001461.
26. Marques C, Kakehasi AM, Gomides AP, Paiva ED, Neto ET, Pileggi GC, et al. A Brazilian cohort of patients with immuno-mediated chronic inflammatory diseases infected by SARS-CoV-2 (ReumaCoV-Brasil Registry): protocol for a prospective, observational study. JMIR Res Protoc 2020;9:e24357.
27. Isnardi CA, Gómez G, Quintana R, Roberts K, Berbotto G, Báez RM, et al. Características epidemiológicas y desenlaces de la infección por SARS-CoV-2 en pacientes con patologías reumáticas: primer reporte del registro argentino SAR-COVID. Rev Arg Reumatol 2021;32:7–15.
28. Moons KG, Altman DG, Reitsma JB, Loannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–73.
29. Schaefer J, Lehne M, Schepers J, Prasser F, Thun S. The use of machine learning in rare diseases: a scoping review. Orphanet J Rare Dis 2020;15:145.
30. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008;28:1–26.
31. Unal I. Defining an optimal cut-point value in ROC analysis: an alternative approach. Comput Math Methods Med 2017;2017:3762651.
32. Shipe ME, Deppen SA, Farjah F, Grogan EL. Developing prediction models for clinical use using logistic regression: an overview. J Thorac Dis 2019;11 Suppl 4:S574–84.
33. Richter AN, Khoshgoftaar TM. A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. Artif Intell Med 2018;90:1–14.
34. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess 2004;8:iii-iv, ix-xi, 1–158.
35. Austin PC, Steyerberg EW. The Integrated Calibration Index (ICI) and related metrics for quantifying the calibration of logistic regression models. Stat Med 2019;38:4051–65.
36. Gianfrancesco M, Hyrich KL, Al-Adely S, Carmona L, Danila MI, Gossec L, et al. Characteristics associated with hospitalisation for COVID-19 in people with rheumatic disease: data from the COVID-19 Global Rheumatology Alliance physician-reported registry. Ann Rheum Dis 2020;79:859–66.
37. Brenner EJ, Ungaro RC, Gearry RB, Kaplan GG, Kissous-Hunt M, Lewis JD, et al. Corticosteroids, but not TNF antagonists, are associated with adverse COVID-19 outcomes in patients with inflammatory bowel diseases: results from an international registry. Gastroenterology 2020;159:481–91.
38. Mahil SK, Dand N, Mason KJ, Yiu ZZ, Tsakok T, Meynell F, et al. Factors associated with adverse COVID-19 outcomes in patients with psoriasis: insights from a global registry-based study. J Allergy Clin Immunol 2021;147:60–71.
39. Strangfeld A, Schäfer M, Gianfrancesco MA, Lawson-Tovey S, Liew JW, Ljung L, et al. Factors associated with COVID-19-related death in people with rheumatic diseases: results from the COVID-19 Global Rheumatology Alliance physician-reported registry. Ann Rheum Dis 2021;80:930–42.
40. Bachiller-Corral J, Boteanu A, Garcia-Villanueva MJ, de la Puente C, Revegna M, Diaz-Miguel MC, et al. Risk of severe COVID-19 infection in patients with inflammatory rheumatic diseases. J Rheumatol 2021;48:1098–102.
41. Pablos JL, Abasolo L, Alvaro-Gracia JM, Blanco FJ, Blanco R, Castrejón, et al. Prevalence of hospital PCR-confirmed COVID-19 cases in patients with chronic inflammatory and autoimmune rheumatic diseases. Ann Rheum Dis 2020;79:1170–3.
42. Eythorsson E, Bjarnadottir V, Runolfsdottir HL, Helgason D, Ingvarsson RF, Bjornsson HK, et al. Development and validation of a prognostic model for COVID-19: a population-based cohort study in Iceland [preprint]. medRxiv 2021.
43. US Food and Drug Administration. FDA authorizes REGEN-COV monoclonal antibody therapy for post-exposure prophylaxis (prevention) for COVID-19. 2021. URL: https://www.fda.gov/drugs/drug-safety-and-availability/fda-authorizes-regen-cov-monoclonal-antibody-therapy-post-exposure-prophylaxis-prevention-covid-19.
44. US Food and Drug Administration. RE: emergency use authorization 091. 2022. URL: https://www.fda.gov/media/145610/download.
45. Sparks JA, Wallace ZS, Seet AM, Gianfrancesco MA, Izadi Z, Hyrich KL, et al. Associations of baseline use of biologic or targeted synthetic DMARDs with COVID-19 severity in rheumatoid arthritis: results from the COVID-19 Global Rheumatology Alliance physician registry. Ann Rheum Dis 2021;80:1137–46.
46. Zeiberg D, Prahlad T, Nallamothu BK, Iwashyna TJ, Wiens J, Sjoding MW. Machine learning for patient risk stratification for acute respiratory distress syndrome. PLoS One 2019;14:e0214465.
47. Ding XF, Li JB, Liang HY, Wang ZY, Jiao TT, Liu Z, et al. Predictive model for acute respiratory distress syndrome events in ICU patients in China using machine learning algorithms: a secondary analysis of a cohort study. J Transl Med 2019;17:326.
48. Xu W, Sun NN, Gao HN, Chen ZY, Yang Y, Ju B, et al. Risk factors analysis of COVID-19 patients with ARDS and prediction based on machine learning. Sci Rep 2021;11:2933.
49. Middleton KL, Jones JL. Socially desirable response sets: the impact of country culture. Psychol Mark 2000;17:149–63.
50. Cook C, Patel NJ, D'Silva KM, Hsu T, Dilorio M, Prisco L, et al. Clinical characteristics and outcomes of COVID-19 breakthrough infections among vaccinated patients with systemic autoimmune rheumatic diseases. Ann Rheum Dis 2022;81:289–91.