

Developing Prediction Models for COVID-19 Outcomes: A Valuable Tool for Resource-Limited Hospitals

Irina-Maria Popescu¹, Madalin-Marius Margan², Mariana Anghel¹, Alexandra Mocanu³, Sorina Maria Denisa Laitin¹, Roxana Margan⁴, Ionut Dragos Capraru¹, Alexandra-Andreea Tene⁵, Emanuela-Georgiana Gal-Nadasan⁶, Daniela Cirnatu^{5,7}, Gratiana Nicoleta Chicin^{5,8}, Cristian Oancea⁹, Andrei Anghel¹⁰

¹Department of Infectious Diseases, Discipline of Epidemiology, "Victor Babes" University of Medicine and Pharmacy, Timisoara, Romania; ²Department of Functional Sciences, Discipline of Public Health, "Victor Babes" University of Medicine and Pharmacy, Timisoara, Romania; ³Department of Infectious Diseases, Discipline of Infectious Diseases, "Victor Babes" University of Medicine and Pharmacy, Timisoara, Romania; ⁴Department of Functional Sciences, Discipline of Physiology, "Victor Babes" University of Medicine and Pharmacy, Timisoara, Romania; ⁵Regional Center of Public Health Timisoara, Timisoara, Romania; ⁶Department of Balneology, Medical Rehabilitation and Rheumatology, Discipline of Medical Rehabilitation, "Victor Babes" University of Medicine and Pharmacy, Timisoara, Romania; ⁷Department of Medicine, "Vasile Goldis" Western University, Faculty of Medicine, Arad, Romania; ⁸Department of Epidemiology, Infectious Diseases and Preventive Medicine, "Vasile Goldis" Western University, Faculty of Medicine, Arad, Romania; ⁹Center for Research and Innovation in Precision Medicine of Respiratory Disease, "Victor Babes" University of Medicine and Pharmacy, Timisoara, Romania; ¹⁰Department of Biochemistry and Pharmacology, Discipline of Biochemistry, "Victor Babes" University of Medicine and Pharmacy, Timisoara, Romania

Correspondence: Madalin-Marius Margan, "Victor Babes" University of Medicine and Pharmacy, Eftimie Murgu Square, No. 2, Timisoara, 300041, Romania, Tel +40 726 277 354, Email marganmm@gmail.com

Purpose: Coronavirus disease is a global pandemic with millions of confirmed cases and hundreds of thousands of deaths worldwide that continues to create a significant burden on the healthcare systems. The aim of this study was to determine the patient clinical and paraclinical profiles that associate with COVID-19 unfavourable outcome and generate a prediction model that could separate between high-risk and low-risk groups.

Patients and Methods: The present study is a multivariate observational retrospective study. A total of 483 patients, residents of the municipality of Timișoara, the biggest city in the Western Region of Romania, were included in the study group that was further divided into 3 sub-groups in accordance with the disease severity form.

Results: Increased age (cOR=1.09, 95% CI: 1.06–1.11, p<0.001), cardiovascular diseases (cOR=3.37, 95% CI: 1.96–6.08, p<0.001), renal disease (cOR=4.26, 95% CI: 2.13–8.52, p<0.001), and neurological disorder (cOR=5.46, 95% CI: 2.71–11.01, p<0.001) were all independently significantly correlated with an unfavourable outcome in the study group. The severe form increases the risk of an unfavourable outcome 19.59 times (95% CI: 11.57–34.10, p<0.001), while older age remains an independent risk factor even when disease severity is included in the statistical model. An unfavourable outcome was positively associated with increased values for the following paraclinical parameters: white blood count (WBC; cOR=1.10, 95% CI: 1.05–1.15, p<0.001), absolute neutrophil count (ANC; cOR=1.15, 95% CI: 1.09–1.21, p<0.001) and C-reactive protein (CRP; cOR=1.007, 95% CI: 1.004–1.009, p<0.001). The best prediction model including age, ANC and CRP achieved a receiver operating characteristic (ROC) curve with the area under the curve (AUC) = 0.845 (95% CI: 0.813–0.877, p<0.001); cut-off value = 0.12; sensitivity = 72.3%; specificity = 83.9%.

Conclusion: This model and risk profiling may contribute to a more precise allocation of limited healthcare resources in a clinical setup and can guide the development of strategies for disease management.

Keywords: ANC, CRP, risk profiling, unfavourable outcome

Introduction

Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a global pandemic with millions of confirmed cases and hundreds of thousands of deaths worldwide.¹ While the

majority of COVID-19 cases are mild, some individuals experience severe illness requiring hospitalization and intensive care.

Predicting which individuals are at risk for severe COVID-19 can inform the allocation of limited healthcare resources and guide the development of strategies for disease prevention and management. A very important aspect in the development of such prediction models is the differentiation between science and pseudoscience, this type of prediction requires accurate data, reproducibility and specialized researchers involved in their development. Kasapçopur Ö draws attention to this aspect and the need for rigorous documentation of scientific works and the involvement of researchers who have the studied topic as an area of expertise.²

Gao YD et al highlights the importance of clinical and paraclinical factors as risk factors in the evolution of the disease, including age, underlying medical conditions and laboratory results outside the reference range, aspects also showed by Carpenter CR et al.^{3,4} It is clear that a varied range of diverse set of indicators can provide predictability regarding the evolution of the disease, including both clinical (disease like obesity, diabetes) and paraclinical factors (leucocyte formula, C-reactive protein). Understanding the role of these factors in the development and progression of COVID-19 can help to identify individuals at high risk for severe illness and guide the implementation of preventive measures.

The goal of the research was to scrutinize easy to obtain available factors, like demographic characteristics of the patient, comorbidities, and blood metrics like white blood count (WBC), absolute neutrophil count (ANC), C-reactive protein (CRP), and utilize them to generate a simple straightforward forecasting model to differentiate whether the disease progression is likely to be favourable or unfavourable. Using this kind of prediction approach empowers resource-constrained hospitals to make timely referrals of their patients to larger, more resource-equipped hospitals.

Materials and Methods

Study Design

The present study is a multivariate observational retrospective study conducted between March 2020 and December 2020 on patients hospitalized in Dr. Victor Babeş Clinical Hospital for Infectious Diseases and Pulmonology Timișoara.

The inclusion criteria for the study were: age above 1 year old (to avoid juxtaposition with neonatal mortality); COVID-19 diagnosis confirmed by polymerase chain reaction (RT-PCR) testing; residency in the municipality of Timișoara, the biggest city in the Western Region of Romania, with 319,279 at the last National census (to reduce the impact of other factors like the quality of primary or emergency medicine in rural areas that could have limited access to immediate medical care prior to hospital admission); availability of data in the medical records or laboratory test results for the parameters considered in the study design. Demographic and clinical data were extracted from electronic medical records (EMR) and included sex, age, comorbidities (medical history) and stage of disease severity. Comorbidities were further categorized as following: cardiovascular diseases – CVD (that included hypertension, ischemic heart disease, arrhythmias, post myocardial infarction status, valvular diseases), chronic pulmonary disease – CPD (that included chronic obstructive bronchopulmonary disease – COPD, bronchial asthma, fibrosis, emphysema), and other comorbidities: obesity, diabetes mellitus, hepatopathy, renal diseases, neurologic disorders, immunodeficiency, autoimmune diseases, cancer.

Disease severity classification comprised 3 stages (mild, moderate and severe) defined in accordance with Protocol no. 487/23.03.2020 and appendix no. 253/27.03.2020. Paraclinical data were extracted from the Hospital's Laboratory software and included three blood test results: WBC, ANC, CRP results from the first laboratory test performed during hospitalization. The laboratory protocol was the same for patients included in the study. The blood tests were carried out on SYSMEX XN 550 by methods of flow-cytometry, impedance and colorimetry. Reference ranges of blood parameters for normal adults were the following WBC: $4.00\text{--}10.00 \times 10^3/\mu\text{L}$, ANC: $2.20\text{--}6.60 \times 10^3/\mu\text{L}$, CRP: $<5 \text{ mg/L}$. For the paediatric population, the reference values are WBC: $6.20\text{--}17 \times 10^3/\mu\text{L}$, ANC: $1.50\text{--}7.50 \times 10^3/\mu\text{L}$, CRP: $<5\text{mg/L}$.

Regarding the outcome of the disease, it was divided into two categories in accordance with the final clinical status of the patient extracted from the EMR. A favourable outcome was considered when one of the following two statuses: "discharged - clinically improved" and "discharged - disease free" was present. An unfavourable outcome was linked to the presence of "deceased" status, meaning patient death occurred during hospitalization. No other statuses besides these three were reported.

Patient data was anonymised prior to data extraction, as patients' CNPs (Personal Numeric Code – the only unique identifying number used in Romania) were replaced with an automatically generated case identification code (procedure done by the hospital staff) and all patient identifiable information (first name, last name and address) were removed.

From a number of 2693 patients admitted with confirmed COVID-19 diagnosis by PCR-testing between March 2020 and December 2020, a total of 483 patients met the inclusion criteria and were included in the present study. The study group was further divided into 3 sub-groups in accordance with the disease severity form.

None of the participants were administered the COVID-19 vaccine as its administration for the general population started in January 2021 in Romania.

The Local Committee of Ethics for Scientific Research of the “Victor Babes” University of Medicine and Pharmacy Timisoara operates under art. provisions 167 of Law no. 95/2006, art. 28, chapter VIII of order 904/2006 and with EU GCP Directives 2005/28/EC, International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), and with the Declaration of Helsinki—Recommendations Guiding Medical Doctors in Biomedical Research Involving Human Subjects. The current study protocol received ethical approval from the “Victor Babes” University of Medicine and Pharmacy Timisoara (nr. 23/2021).

Statistical Analysis

Statistical analysis was performed using R version 4.2.0 (R Foundation for Statistical Computing, Vienna, Austria).

The R Base Package was used for data manipulation, and the R Stats Package was used for performing univariate analysis (including Pearson's Chi-squared test, Fisher's Exact Test and Kruskal–Wallis *H*-test) and multivariate analysis (including Multiple Logistic Regression). The R pROC package was used for analysing and visualizing Receiver Operating Characteristic (ROC) curves, compute their confidence intervals, and perform hypothesis tests.

Continuous variables were summarized as means and standard deviations, while categorical variables were presented as counts and percentages.

For numerical variables, the study utilized non-parametric Kruskal–Wallis *H*-test to compare multiple independent samples that did not follow a normal distribution, while the chi-square test was used for categorical variables. For situations where frequencies were less than 5, Fisher's Exact Test was used to evaluate the association between two categorical variables. If these categorical variables consisted of more than two groups, the Freeman-Halton extension of Fisher's Exact Test was implemented. The severity of COVID-19 was evaluated using Spearman Correlation and univariate logistic analysis and included clinical and paraclinical collected data elements as risk factors. In addition, risk factors were further analysed using multiple logistic regression to develop diagnostic models for distinguishing between a favourable outcome (discharge) and an unfavourable outcome (death) of COVID-19 cases. The clinical and paraclinical features' ability to distinguish between the two types of COVID-19 was determined using Receiver Operating Characteristic (ROC) analysis, which found the corresponding cut-off value.

To assess the relationship between dependent variable and predictor variables, we used statistical test(s) with a significance level of 0.05.

Imputation of missing data was not used during statistical analysis.

Results

A total of 483 patients with COVID-19 met the inclusion criteria and were included in the study group for analysis.

Considering disease severity form at the moment of initial evaluation as the main predictive factor of the disease evolution, the study group was further divided into 3 subgroups: mild form subgroup (39.33%, N=190), moderate form subgroup (36.23%, N=175), severe form subgroup (24.43%, N=118). Clinical characteristics of the study group and severity forms subgroups are presented in [Table 1](#).

Prediction Factors and Disease Severity Forms

Age

The mean age of the study group was 61.19 years (range: min 9 y.o, max 95 y.o., standard deviation SD=14.15). The mean age for the 3 subgroups increased as disease severity increased: for the mild group the mean age was 56.85 (range: min 9 y.o,

Table 1 Distribution of Demographic, Clinical Characteristics, and Comorbidities of Patients Across the Different Severity Forms Subgroups

Variables	Result	Mean or N (%)				p-value
		Total N=483 (100%)	Mild N=190 (39.33%)	Moderate N=175 (36.23%)	Severe N=118 (24.43%)	
Age		61.19	56.85	61.44	67.83	<0.001 ^a
Sex	Male Female	271(56.11) 212(43.89)	87(45.79) 103(54.21)	105(60) 70(40)	79(66.95) 39(33.05)	<0.001 ^b
CVD		308(63.77)	108(56.84)	110(62.86)	90(76.27)	0.002 ^b
CPD		51(10.56)	23(12.11)	14(8.00)	14(11.86)	0.385 ^b
Obesity	I	37(7.66)	9(4.74)	18(10.29)	10(8.47)	0.128 ^b
	II	30(6.21)	7(3.68)	13(7.43)	10(8.47)	0.162 ^b
	III	11(2.28)	3(1.58)	3(1.71)	5(4.24)	0.288 ^c
	Severe	11(2.28)	5(2.63)	4(2.29)	2(1.69)	0.928 ^c
	All forms	89(18.42)	24(12.63)	38(21.71)	27(22.88)	0.029 ^b
Diabetes mellitus	Type 1	35(7.25)	13(6.84)	13(7.43)	9(7.63)	0.960 ^b
	Type 2	76(15.73)	23(12.11)	31(17.71)	22(18.64)	0.206 ^b
	Both types	111(22.98)	36(18.94)	44(25.14)	31(26.27)	0.231 ^b
Hepatopathy	Hepatitis B	5(1.04)	2(1.05)	3(1.71)	0(0.00)	0.375 ^c
	Hepatitis C	8(1.66)	3(1.58)	2(1.14)	3(2.54)	0.603 ^c
	Cirrhosis	9(1.86)	4(2.11)	1(0.57)	4(3.39)	0.201 ^c
	All pathologies	22(4.55)	9(4.73)	6(3.42)	7(5.93)	0.594 ^c
Renal disease		37(7.66)	9(4.74)	8(4.57)	20(16.95)	<0.001 ^a
Neurologic disorder		37(7.66)	9(4.74)	15(8.57)	13(11.02)	0.111 ^b
Immunodeficiency		2(0.41)	1(0.53)	0(0.00)	1(0.85)	0.714 ^c
Autoimmune disease		10(2.07)	5(2.63)	4(2.29)	1(0.85)	0.658 ^c
Cancer		19(3.93)	9(4.74)	4(2.29)	6(5.08)	0.365 ^c

Notes: ^aKruskal–Wallis *H*-test *p*-value; ^bChi-squared test *p*-value; ^cFisher's exact test (the Freeman-Halton extension) *p*-value.

max 95 y.o., standard deviation SD=15.00), for the moderate group it was 61.44 (range: min 26 y.o, max 92 y.o., standard deviation SD=12.65), for the severe group it was 67.83 (range: min 36 y.o, max 92 y.o., standard deviation SD=12.19).

Sex

Males represented 56.11% (N=271) of the study group, while females represented 43.89% (N=212). Significant differences between the 3 subgroups were observed regarding sex distribution, as males were more frequently affected by moderate (sex distribution for the moderate group: male/female = 60%/40%, *p*<0.001) and severe forms (sex distribution for the severe group: male/female= 66.95/33.05%, *p*<0.001). Females were more frequent in the mild group (sex distribution for the mild group: male/female = 45.79%/54.21%).

Comorbidities

Cardiovascular pathology, obesity (all forms) and renal disease all varied with statistical significance between the three subgroups; in contrast, diabetes mellitus, hepatopathy, neurological disorder, immunodeficiency, autoimmune disease, and cancer did not vary between the subgroups.

In the severe subgroup, 76.27% (N=90) presented cardiovascular diseases, 26.27% (N=31) had diabetes mellitus and 16.95% (N=20) suffered from renal disease (Table 1).

Paraclinical Factors

All the three studied paraclinical factors (WBC, ANC, CRP) varied significantly between the subgroups (Table 2). The mean CRP value for the severe group was 114.99, almost 23 times higher than the upper limit of the normal range.

Prediction Factors and Disease Outcome

Disease outcome varied with high statistical significance between the three subgroups (Table 3). The unfavourable outcome was present in 5.27% (N=10) of the cases in the mild subgroup, 9.71% (N=17) in the moderate subgroup, and 71.79% (N=72) in the severe group ($p<0.001$).

The Spearman correlation and univariate logistic analysis of clinical risk factors for COVID-19 unfavourable outcome is presented in Table 4. The results indicate that age is strongly correlated with an unfavourable outcome ($p<0.001$). This means that there is a statistically significant difference in the mean age of individuals who died from COVID-19 in the study group compared to those who did not.

In contrast, sex does not appear to be significantly correlated with an unfavourable outcome, as indicated by a p -value of 0.741 in the Pearson's Chi-squared test. Similarly, chronic pulmonary disease (CPD), obesity, diabetes, hepatopathy, immunodeficiency, autoimmune disease and cancer do not appear to be significantly correlated with unfavourable. However, cardiovascular disease (CVD), renal diseases and neurological disorder are all significantly correlated with poor outcome, as indicated by p -values of <0.001 . Having a CVD increases the chance of death 3.37 times.

The Spearman correlation and univariate logistic analysis of paraclinical factors for COVID-19 unfavourable outcome is presented in Table 5. Increased values of WBC, ANC and CRP are all correlated with an unfavourable outcome.

In the multiple logistic regression model for clinical factors (Table 6), the unfavourable outcome was positively associated with age (aOR=1.06, 95% CI 1.03–1.09, $p<0.001$), severe form (aOR=1.81, 95% CI 1.00–34.08, $p<0.001$) and

Table 2 Comparison of Paraclinical Factors Values Across the Different Severity Forms Subgroups

Variable	Mean (Min-Max) or N (%)				p-value
	Total (N=462)	Mild (N=189)	Moderate (N=174)	Severe (N=99)	
WBC	8.2689 (1.35–48.08)	7.065 (1.73–23.55)	8.116 (1.35–26.33)	10.834 (2.11–48.08)	<0.001
Increased (Leucocytosis)	121(26,19)	33(17.46)	44(25.29)	44(44.44)	<0.001
Normal	341(73,81)	156(82.54)	130(74.71)	55(55.56)	
ANC	6.4727 (0.38–37.84)	5.1528 (0.38–21.96)	6.4624 (0.62–24.57)	9.0107 (1.43–37.84)	<0.001
Increased (Neutrophilia)	171(37.01)	46(24.33)	63(36.2)	62(62.62)	<0.001
Normal	421(62.99)	143(75.67)	111(63.8)	37(37.38)	
	Total (N=455)	Mild (N=187)	Moderate (N=171)	Severe (N=97)	
CRP value	74.7093 (0.04–368.81)	41.9270 (0.04–272.86)	87.7086 (0.45–368.81)	114.9919 (1.63–365.96)	<0.001
Increased	412(90.54)	153(81.81)	165(96.49)	94(96.9)	<0.001
Normal	43(9.46)	34(18.19)	6(3.51)	3(3.1)	

Table 3 Disease Outcome in the Study Group and Severity Forms Subgroups

Outcome	N(%)				p-value
	Total (N=483)	Mild (N=190)	Moderate (N=175)	Severe (N=118)	
Unfavourable	99(20.5)	10(5.27)	17(9.71)	72(61.01)	<0.001
Favourable	384(79.5)	180(94.73)	158(90.29)	46(38.99)	

Table 4 Association Between Various Demographic Characteristics, Disease Severity Form, Comorbidities, and Unfavourable Outcome in the Study Group, Based on Spearman Correlation and Univariate Logistic Regression Analysis

Risk Factor	Unfavourable Outcome			
	Spearman Correlation		Univariate Logistic Regression	
	Rho	p-value	cOR (95% CI)	p-value
Demographic characteristics				
Age	0.3852	<0.001	1.092(1.069–1.118)	<0.001
Male sex	0.0150	0.741	1.078(0.691–1.691)	0.741
Disease severity				
Mild form	–0.3038	<0.001	0.127(0.060–0.241)	<0.001
Moderate form	–0.2013	<0.001	0.296(0.164–0.507)	<0.001
Severe form	0.5707	<0.001	19.594(11.576–34.103)	<0.001
Comorbidities				
CVD	0.2013	<0.001	3.372(1.969–6.081)	<0.001
CPD	–0.0576	0.206	0.587(0.235–1.270)	0.209
Obesity	0.0364	0.423	1.251(0.710–2.136)	0.423
Diabetes mellitus	0.1005	0.027	1.732(1.051–2.814)	0.028
Hepatopathy	0.0366	0.421	1.483(0.520–3.718)	0.423
Renal disease	0.2008	<0.001	4.269(2.133–8.527)	<0.001
Neurologic disorder	0.2394	<0.001	5.465(2.742–11.019)	<0.001
Immunodeficiency	0.0471	0.301	3.908(0.153–99.416)	0.336
Autoimmune disease	–0.0378	0.407	0.425(0.022–2.302)	0.419
Cancer	0.0027	0.951	1.035(0.290–2.930)	0.951

Abbreviations: Rho, Spearman's rank correlation coefficient; cOR, crude odds ratio for logistic univariate regression; 95% CI, 95% confidence interval.

Table 5 Correlation Between Paraclinical Factors and Unfavourable Outcome in the Study Group, Based on Spearman Correlation and Univariate Logistic Regression Analysis

Predictive Factor	Unfavourable Outcome			
	Spearman Correlation		Univariate Logistic Regression	
	Rho	p-value	cOR (95% CI)	p-value
WBC	0.2434	<0.001	1.103(1.053–1.159)	<0.001
ANC	0.2699	<0.001	1.153(1.094–1.218)	<0.001
CRP	0.2435	<0.001	1.007(1.004–1.009)	<0.001

Abbreviations: Rho, Spearman's rank correlation coefficient; cOR, crude odds ratio for logistic univariate regression; 95% CI, 95% confidence interval.

Table 6 Correlation Between Selected Clinical Risk Factors and Unfavourable Outcome in the Study Group Using Multivariate Logistic Regression

Variables	aOR (95% CI)	p-value
Age	1.067(1.037–1.099)	<0.001***
Severe form	1.811(1.002–34.080)	<0.001***
CVD	1.506(7.282–3.206)	0.275
Renal disease	1.212(4.822–3.008)	0.678
Neurological disorders	2.914(1.106–7.647)	0.029*

Notes: “***”: p-value ≤ 0.001; “**”: 0.01 < p-value ≤ 0.05; “*”: 0.1 < p-value ≤ 1 (not significant).

Abbreviations: aOR, adjusted odds ratio for logistic multivariate regression; 95% CI, 95% confidence interval.

Table 7 Correlation Between Paraclinical Risk Factors and Unfavourable Outcome in the Study Group Using Multivariate Logistic Regression

Variables	aOR (95% CI)	p-value
WBC	0.741(0.495–1.013)	0.118
ANC	1.533(1.105–2.332)	0.031*
CRP	1.004(1.001–1.008)	0.002**

Notes: “***”: 0.001 < p-value ≤ 0.01; “**”: 0.01 < p-value ≤ 0.05; “*”: 0.1 < p-value ≤ 1 (not significant).

Abbreviations: aOR, adjusted odds ratio for logistic multivariate regression; 95% CI, 95% confidence interval.

neurological disorders (aOR=2.91, 95% CI 1.10–7.64, p<0.05). The model had good fit, as indicated by low deviance residuals, the low p-value of the coefficients and the comparison of AIC (Akaike Information Criterion) value with other considered models.

In the multiple logistic regression model for paraclinical factors (Table 7), the unfavourable outcome was positively associated with ANC (aOR=1.53, 95% CI 1.10–2.33, p<0.05) and CRP (aOR=1.004, 95% CI 1.001–1.008, p<0.005). The model had good fit, as indicated by a low deviance residuals and p-value of the coefficients.

Prediction Models

Using the three numerical factors that showed high statistical correlation with unfavourable outcome in the multiple logistic regressions (age, ANC, CRP), three diagnostic binary classification models were proposed to separate high-risk and low-risk patients. The ROC curves for the three logistic models in the study group (presented in Figure 1) were Model 1: $\text{logit}(p) = -2.221925 + 0.006981 * X$ (X = CRP value); ROC curve: AUC = 0.688 (95% CI: 0.625–0.751, p<0.001); cut-off = 0.13; sensitivity = 72.3%; specificity = 58.8%. Model 2: $\text{logit}(p) = -2.988112 + 0.005666 * X1 + 0.121688 * X2$ (X1 = CRP value, X2 = ANC value); ROC curve: AUC = 0.731 (95% CI: 0.673–0.790, p<0.001); cut-off = 0.14; sensitivity = 71.0%; specificity = 65.7%. Model 3: $\text{logit}(p) = -8.647672 + 0.003347 * X1 + 0.122706 * X2 + 0.088323 * X3$ (X1 = CRP value, X2 = ANC value, X3 = Age); ROC curve: AUC = 0.824 (95% CI: 0.770–0.877, p<0.001); cut-off value = 0.14; sensitivity = 82.8%; specificity = 69.9%.

As the three included parameters (age, ANC, CRP) were available at the moment of diagnosis in a total of 1186 patients with COVID-19, models and ROC curve were reiterated for this extended group.

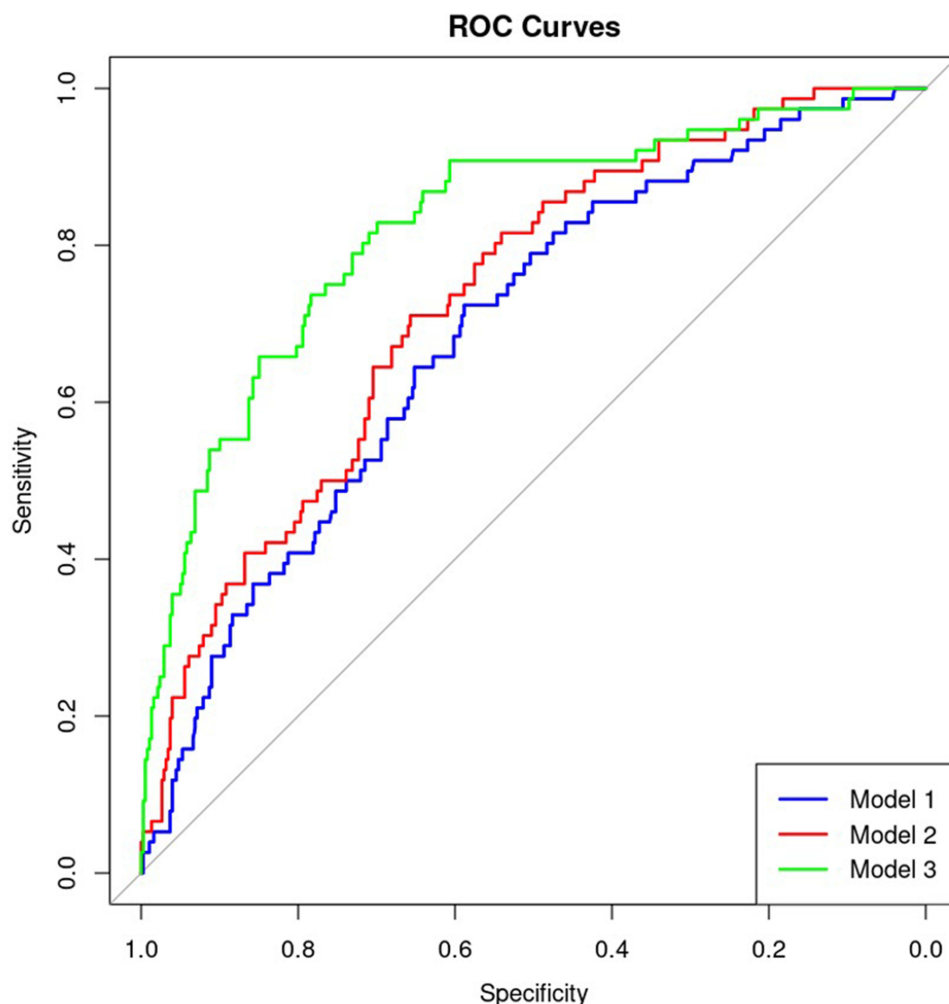


Figure 1 Receiver operating characteristic (ROC) curves illustrating the performance of proposed predictive models in distinguishing high-risk and low-risk patients within the study group dataset.

The ROC curves for the three logistic models in the extended group (presented in Figure 2) were as follows: Model 1: $\text{logit}(p) = -2.5072060 + 0.0076854 * X$ ($X = \text{CRP value}$); ROC curve: $\text{AUC} = 0.737$ (95% CI: 0.698–0.776, $p < 0.001$); cut-off = 0.11; sensitivity = 72.9%; specificity = 64.9%. Model 2: $\text{logit}(P) = -3.393918 + 0.005682 * X1 + 0.146949 * X2$ ($X1 = \text{CRP values}$, $X2 = \text{ANC value}$); ROC curve: $\text{AUC} = 0.780$ (95% CI: 0.743–0.818, $p < 0.001$); cut-off = 0.10; sensitivity = 64.7%; specificity = 79.5%. Model 3: $\text{logit}(P) = -8.295374 + 0.003499 * X1 + 0.132701 * X2 + 0.079695 * X3$ ($X1 = \text{CRP value}$, $X2 = \text{ANC value}$, $X3 = \text{Age}$); ROC curve: $\text{AUC} = 0.845$ (95% CI: 0.813–0.877, $p < 0.001$); cut-off value = 0.12; sensitivity = 72.3%; specificity = 83.9%.

When the extended group was used in the models' creation, the accuracy of the results slightly increased. This stands as proof that the usability of the proposed model brings value in a clinical setup at initial presentation and independently of the disease severity form (a classification that was not available for the extended group and was not used in the model).

Discussion

In respect of age as a risk factor, the mean age was 61.19 years for the study group and 68.83 for the severely diseased group. In contrast, a study from the same medical centre published by Laza et al in 2022 reported a mean age of 48.76 years in the general group.⁵ This important difference could have been caused by the different residency of included patients, as the Laza et al study group also included hospitalized patients from all neighbouring counties, from both rural and urban areas. Because all cases from these administrative divisions were sent to a tertiary centre like the one where the study was performed, patients

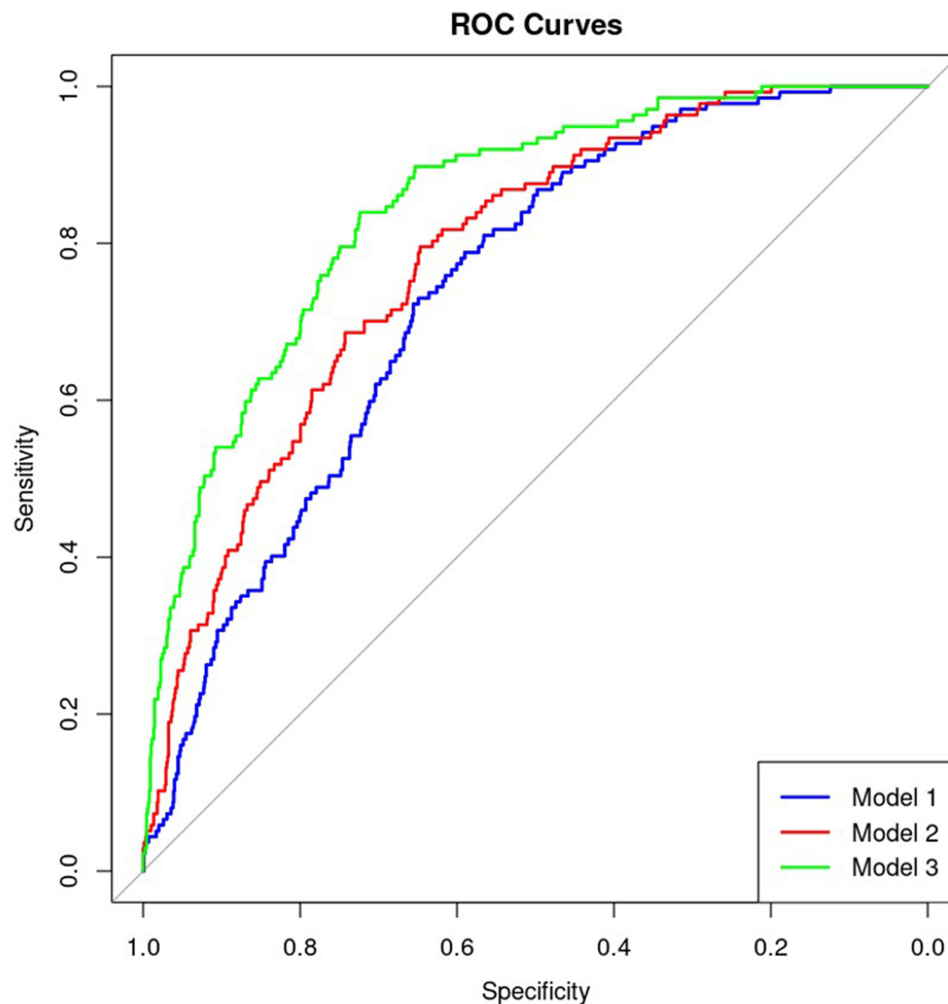


Figure 2 Receiver operating characteristic (ROC) curves illustrating the performance of proposed predictive models in distinguishing high-risk and low-risk patients within the extended group dataset.

were admitted independently of age and severity, thus generating a lower mean age. Our results indicate a strong correlation between age and unfavourable outcome, establishing age as an independent risk factor for COVID 19 poor outcome. Laza et al objectivated the same result.⁵ Ho FK et al, 2020 came to the same conclusion, proposing an attenuated immune response among the elderly as the possible explanation.⁶ A systematic review by Romero et al, 2021 that included studies from variate geographic regions (Europe, USA) found the same correlation between increased disease severity and higher age.⁷ Flaherty et al, in a critical literature review from 2020 also reported age as a risk factor, patients over 66 years of age developed more frequent severe forms requiring admission to ICU.⁸

Regarding comorbidities, CVD, renal disease, and neurological disorder were all significantly correlated with an unfavourable outcome in the study group. Our findings show that having a neurologic disorder increases the risk of an unfavourable outcome 5.46 times, a renal disease increases the risk 4.26 times, and CVD 3.37 times.

Garcia et al, in a cohort study from 2020 including 576 patients, found that the 105 patients (18.3% of the total) with neurological disorders prior to COVID-19 infection had a survival rate that was 30 days lower compared with the group of patients without neurological pathology. Patients with chronic neurological disorders developed severe forms of the disease and having a neurological disorder prior to COVID-19 infection led to an unfavourable outcome.⁹ According to Alberti et al, (2021) patients diagnosed with a cerebrovascular disease prior to COVID-19 infection presented a greater vulnerability to secondary events and to an unfavourable outcome.¹⁰

Our study showed that the presence of a kidney disease prior to confirmation of COVID-19 infection increased the risk of an unfavourable outcome 4 times. The same conclusion was reported by Pecky et al in a meta-analysis from 2021 including 73 studies that aimed to evaluate the association between multi-organ dysfunction and the evolution of the disease in patients with COVID-19.¹¹ Pakhchanian et al, in a multicentre cohort study from 2021, that included a total number of 152,463 patients and 8810 patients presenting renal pathology prior to COVID-19 infection, also concluded that renal disease is an independent risk factor for COVID-19 unfavourable outcome.¹²

Li et al, in a study carried out in Wuhan China on a study group containing 596 patients hospitalized between January and March 2020, from which 215 (36.1%) presented CVD prior to COVID-19 infection, also highlighted CVD as an independent risk factor for an unfavourable outcome.¹³ Cordero et al, in a meta-analysis from 2020 that included a number of 307,596 patients with COVID-19 from which 46,321 (15.1% of the total) had CVD, demonstrated that CVD clearly augments the severity of COVID-19, increasing 4 times the risk of an unfavourable outcome.¹⁴

Concerning the contribution of paraclinical parameters, our results show that the unfavourable outcome was positively associated with increased values for ANC and CRP. Bouayed et al, in a monocentric descriptive retrospective study conducted in 2022, correlated the existence of elevated values for CRP in patients with COVID-19 with the severity of the disease and proved that increased values of this marker are a significant risk factor for an unfavourable outcome of the disease.¹⁵ The same conclusion was reported by Lentner et al in 2021 in a retrospective cohort study performed on 541 patients, mentions the presence of high values for CRP at admission as a predictive factor for unfavourable disease evolution.¹⁶

Mosquera-Sulbaran et al, 2021 studied the role of increased CRP value as a prediction risk factor for COVID-19 unfavourable outcome and showed association with severe lung damage.¹⁷ Mocanu et al, in a retrospective assessment from 2022 with the aim of determining the characteristics of Roma patients with severe forms of the disease, also established the increased level of CRP as a risk factor for the severity of the disease.¹⁸

Huang et al, 2020, in a meta-analysis that included 1289 cases of COVID-19, reported the presence of leucocytosis in patients with severe forms of disease and an unfavourable outcome.¹⁹ This finding was also reported by Yamada et al in a systematic review from 2020, which concluded that leucocytosis and elevated CRP values represented predictive factors for an unfavourable outcome.²⁰ A retrospective study conducted by Zhu et al in 2021 on 163 patients reported a significant association between the presence of leucocytosis at admission and severe forms of the disease that had an unfavourable outcome.²¹

In our study, the increased level of ANC was correlated with a severe form of the disease and an unfavourable outcome. McKenna et al, in a review article from 2022, showed that an increased number of neutrophils in patients with COVID-19 indicated an exacerbation of the disease and was correlated with an unfavourable outcome.²² Moreover, in a 2021 study conducted in Delhi, on a group of 203 patients diagnosed with COVID-19, Agarwal et al reported neutrophilia as a risk factor for a severe evolution of the disease and an unfavourable outcome.²³

Our intention was to establish a model that could be useful to select different management patterns for patients at admission based on outcome prediction: favourable or unfavourable. In our multivariate logistic model, when all three paraclinical factors were considered, just ANC and CRP proved to be independent risk factors. Therefore, after adding age (also being a numerical independent risk factor), we have created several predicting models that use these 3 easy-to-collect parameters: age, ANC and CRP. The best model obtained an AUC of 0.845, sensitivity of 72.3% and specificity of 83.9%, which are on par with other models in the literature.

In a study by Hu et al from 2020 a severe COVID-19 risk model was also built using multivariate logistic regression analysis adjusted by age, lobular involvement score and lymphocyte cell count, showing that an objective score based on medical imaging could also be accustomed.²⁴ Huang et al, who used the same parameters to create the prediction model, reached a higher AUC of 0.903 and a higher sensitivity rate of 90.9%.²⁵ Even using CT score alone can generate a sensitivity of 80.0% and a specificity of 82.8% in the ROC curve interpretation, as shown by Li et al.²⁶ Nevertheless, computer tomography is an investigation that might not be available in remote clinical setups.

A systematic review published by Miller et al in 2022 attempted to assess existing research on prediction models for serious COVID-19 disease that can be used in the emergency room or public health scenarios.²⁷ Knight et al created a prediction model that might be used to divide patients admitted with COVID-19 into various care groups. The

recommended easy-to-use 4C Mortality Score included 8 variables that could be measured at initial hospital presentation, one of them being CRP. This calculated score proved to be superior to the existing models, being valuable when including patients in various risk groups for an unfavourable outcome and establishing the treatment behaviour appropriate for the case.²⁸ The PAWNN score proposed by Liu et al only uses age and some complete blood count (CBC) results to precisely forecast mortality for COVID-19 patients throughout their entire hospital. It achieved an AUROC of 0.97 in a validation cohort of 2949 patients with only 1 CBC record, demonstrating its role of helping medical professionals prioritize treatment for COVID-19 patients.²⁹ Lopez et al found that distinctly adding the rate of changes of both hemogram ratios (neutrophil-to-platelet-ratio and neutrophil-to-lymphocyte-ratio) during the first week after admission slightly improves AUC and prediction power of the proposed models.³⁰

Nevertheless, the majority of these studies solely investigated in-hospital mortality as the outcome of interest, a common limitation also presented in our study.^{27–30}

This research has several constraints. Firstly, the study was limited by the relatively small number of participants. Secondly, the study was based on data from a single center, limiting the generalizability of the findings. Furthermore, the study did not incorporate frequently used parameters in ICU settings, such as oxygen saturation levels, clotting cascade activation, and supportive drug treatments, which could have provided further insights into patient outcomes and disease progression.

Other studies have also achieved similar findings, underscoring the need for a straightforward and swift predictive model that could be implemented to differentiate between patient's outcomes. This would enable timely and appropriate treatment for those predicted to have an unfavourable outcome, ideally in specialized medical units equipped with resources tailored to their needs.

Conclusion

The present study showed that advanced age, the presence of comorbidities (CVD, renal disease, neurological pathology), and increased values for WBC, ANC and CRP represent important risk factors for an unfavourable COVID-19 outcome. Thus, a composite profile that includes older age, certain comorbidities, leucocytosis, neutrophilia, and elevated CRP values might suggest an unfavourable prognosis. Conversely, a younger age, absence of comorbidities, and normal CBC and CRP values might signify a more favourable outcome.

To aid in predicting outcomes, we proposed a prediction model for an unfavourable outcome based on very accessible parameters such as age, ANC and CRP value. The model characterizes COVID-19 in the capital city of the Western Region of Romania, Timișoara.

Building on the results of our study, this model could be particularly beneficial for decision-making in healthcare, such as whether to admit patients in smaller hospitals or to directly transfer those potentially facing an unfavourable outcome to larger hospitals equipped with more extensive care options. In this way, the model could assist healthcare practitioners in optimizing patient management and treatment, ultimately improving COVID-19 outcomes.

Abbreviations

WBC, white blood count; ANC, absolute neutrophil count; CRP, C-reactive protein; ROC, receiver-operating characteristic; AUC, area under the curve; COVID-19, coronavirus disease 2019; SARS-CoV-2, severe acute respiratory syndrome coronavirus; RT-PCR, polymerase chain reaction; EMR, electronic medical records; CVD, cardiovascular diseases; CPD, chronic pulmonary disease; COPD, chronic obstructive bronchopulmonary disease; CNP, personal numeric code; ICH, international conference on harmonization of technical requirements for registration of pharmaceuticals for human use; CBC, complete blood count.

Data Sharing Statement

The data presented in this study are available on request from the corresponding author.

Ethics Approval

The study was conducted according to the guidelines of the Declaration of Helsinki, EC Directive 86/609/EEC for animal experiments, and Uniform Requirements for manuscripts submitted to Biomedical journals. The study protocol was approved by the Ethics Committee of the University of Medicine and Pharmacy “Victor Babes” Timisoara (nr. 23/2021).

Consent for Publication

Informed consent was obtained from all subjects involved in the study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. WHO Coronavirus (COVID-19) Dashboard - Overview. 2023. Available from: <https://covid19.who.int/>. Accessed May 18, 2023.
2. Kasapçopur Ö. Science and pseudoscience during the COVID-19 pandemic. *Turk Pediatri Ars.* 2020;55(4):335–336. doi:10.14744/TurkPediatriArs.2020.35902
3. Gao YD, Ding M, Dong X, et al. Risk factors for severe and critically ill COVID-19 patients: a review. *Allergy.* 2021;76(2):428–455. doi:10.1111/all.14657
4. Carpenter CR, Mudd PA, West CP, Wilber E, Wilber ST. Diagnosing COVID-19 in the Emergency Department: a Scoping Review of Clinical Examinations, Laboratory Tests, Imaging Accuracy, and Biases. *Acad Emerg Med.* 2020;27(8):653–670. doi:10.1111/acem.14048
5. Laza R, Lazureanu VE, Musta VF, et al. COVID-19 Independent Risk Factors for Unfavorable Disease Progression: a Cross-Sectional Study from Romania. *Int J Gen Med.* 2022;15:2025–2036. doi:10.2147/IJGM.S350920
6. Petermann-Rocha FK, Gray F, Jani SR, et al. Is older age associated with COVID-19 mortality in the absence of other risk factors? General population cohort study of 470,034 participants. *PLoS One.* 2020;15(11):e0241824. doi:10.1371/journal.pone.0241824
7. Romero Starke K, Reissig D, Peterleit-Haack G, Schmauder S, Nienhaus A, Seidler A. The isolated effect of age on the risk of COVID-19 severe outcomes: a systematic review with meta-analysis. *BMJ Global Health.* 2021;6:e006434. doi:10.1136/bmjgh-2021-006434
8. Flaherty GT, Hession P, Liew CH, et al. COVID-19 in adult patients with pre-existing chronic cardiac, respiratory and metabolic disease: a critical literature review with clinical recommendations. *Trop Dis Travel Med Vaccines.* 2020;6:16. doi:10.1186/s40794-020-00118-y
9. Garcia-Azorin D, Martínez-Pias E, Trigo J, et al. Neurological Comorbidity Is a Predictor of Death in Covid-19 Disease: a Cohort Study on 576 Patients. *Front Neurol.* 2020;11:781. doi:10.3389/fneur.2020.00781
10. Alberti P, Beretta S, Brighina L, Isella V, Ferrarese C. Impact of COVID-19 on pre-existing neurological diseases. In: Priori A, editor. *NEUROLOGY of COVID-19 [Internet]*. Milano, Italy: Milano University Press; 2021: 243–264
11. Pecly IMD, Azevedo RB, Muxfeldt ES, et al. COVID-19 and chronic kidney disease: a comprehensive review. *J Bras Nefrol.* 2021;43(3):383–399. doi:10.1590/2175-8239-JBN-2020-0203
12. Pakhchanian H, Raiker R, Mukherjee A, et al. Outcomes of COVID-19 in CKD Patients: a Multicenter Electronic Medical Record Cohort Study. *Clin J Am Soc Nephrol.* 2021;16(5):785–786. doi:10.2215/CJN.13820820
13. Li J, Guo T, Dong D, et al. Defining heart disease risk for death in COVID-19 infection. *QJM.* 2020;113(12):876–882. doi:10.1093/qjmed/hcaa246
14. Cordero A, García-Gallego SC, Bertomeu-González V, et al. Mortality associated with cardiovascular disease in patients with COVID-19. *Rec Cardioclinics.* 2021;56(1):30–38. doi:10.1016/j.rccl.2020.10.005
15. Bouayed MZ, Laaribi I, Chatar CEM, et al. C-Reactive Protein (CRP): a poor prognostic biomarker in COVID-19. *Front Immunol.* 2022;13:1040024. doi:10.3389/fimmu.2022.1040024
16. Lentner J, Adams T, Knutson V, et al. C-reactive protein levels associated with COVID-19 outcomes in the United States. *J Osteopathic Med.* 2021;121(12):869–873. doi:10.1515/jom-2021-0103
17. Mosquera-Sulbaran JA, Pedrañeñez A, Carrero Y, Callejas D. C-reactive protein as an effector molecule in Covid-19 pathogenesis. *Rev Med Virol.* 2021;31(6):e2221. doi:10.1002/rmv.2221
18. Mocanu A, Lazureanu VE, Marinescu AR, et al. A Retrospective Assessment of Laboratory Findings and Cytokine Markers in Severe SARS-CoV-2 Infection among Patients of Roma Population. *J Clin Med.* 2022;11:6777. doi:10.3390/jcm11226777
19. Huang G, Kovalic AJ, Graber CJ. Prognostic Value of Leukocytosis and Lymphopenia for Coronavirus Disease Severity. *Emerg Infect Dis.* 2020;26(8):1839–1841. doi:10.3201/eid2608.201160
20. Yamada T, Wakabayashi M, Yamaji T, et al. Value of leukocytosis and elevated C-reactive protein in predicting severe coronavirus 2019 (COVID-19): a systematic review and meta-analysis. *Clin Chim Acta.* 2020;509:235–243. doi:10.1016/j.cca.2020.06.008
21. Zhu B, Feng X, Jiang C, et al. Correlation between white blood cell count at admission and mortality in COVID-19 patients: a retrospective study. *BMC Infect Dis.* 2021;21(1):574. doi:10.1186/s12879-021-06277-3

22. McKenna E, Wubben R, Isaza-Correa JM, et al. Neutrophils in COVID-19: not Innocent Bystanders. *Front Immunol.* 2022;13:864387. doi:10.3389/fimmu.2022.864387
23. Agarwal N, Dua D, Sud R, Yadav M, Agarwal A, Vijayan V. COVID-19 mortality prediction model, 3C-M, built for use in resource limited settings - understanding the relevance of neutrophilic leukocytosis in predicting disease severity and mortality. *medRxiv.* 2021;2021:21261565. doi:10.1101/2021.08.05.21261565
24. Hu H, Du H, Li J, et al. Early prediction and identification for severe patients during the pandemic of COVID-19: a severe COVID-19 risk model constructed by multivariate logistic regression analysis. *J Glob Health.* 2020;10(2):020510. doi:10.7189/jogh.10.020510
25. Huang J, Gao J, Zhu W, et al. Indicators and prediction models for the severity of Covid-19. *Int J Clin Pract.* 2021;75(10):e14571. doi:10.1111/ijcp.14571
26. Li K, Wu J, Wu F, et al. The clinical and chest CT features associated with severe and critical COVID-19 pneumonia. *Invest Radiol.* 2020;55:327–331. doi:10.1097/RLI.0000000000000672
27. Miller JL, Tada M, Goto M, et al. Prediction models for severe manifestations and mortality due to COVID-19: a systematic review. *Acad Emerg Med.* 2022;29(2):206–216. doi:10.1111/acem.14447
28. 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. doi:10.1136/bmj.m3339
29. Liu H, Chen J, Yang Q, et al. Development and validation of a risk score using complete blood count to predict in-hospital mortality in COVID-19 patients. *Medicine.* 2021;2(4):435–447.e4. doi:10.1016/j.medj.2020.12.013
30. López A, Madurga R, Castellano JM, et al. Risk Score for Predicting In-Hospital Mortality in COVID-19 (RIM Score). *Diagnostics.* 2021;11:596. doi:10.3390/diagnostics11040596

International Journal of General Medicine

Dovepress

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>