



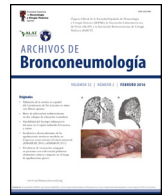
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

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



ARCHIVOS DE Bronconeumología

www.archbronconeumol.org



Original Article

Utility of Differential White Cell Count and Cell Population Data for Ruling Out COVID-19 Infection in Patients With Community-Acquired Pneumonia

Ane Uranga^{a,b,*}, Eloisa Urrechaga^{b,c}, Urko Aguirre^{d,e,f,g}, Maider Intxausti^h, Carlos Ruiz-Martinezⁱ, Maria Jose Lopez de Goicoechea^j, Cristina Ponga^{b,c}, Jose María Quintana^{d,e,f,g}, Cristina Sancho^h, Pilar Sanz^k, Pedro Pablo España^{a,b}, the LEUCOCAP STUDY GROUP[◇]

^a Department of Pneumology, University Hospital of Galdakao-Usansolo, Galdakao, Spain

^b BioCruces Bizkaia Health Research Institute, Baracaldo, Spain

^c Laboratory, University Hospital of Galdakao-Usansolo, Galdakao, Spain

^d Osakidetza Basque Health Service, Galdakao-Usansolo University Hospital, Research Unit, Galdakao, Spain

^e Health Service Research Network on Chronic Diseases (REDISSEC), Bizkaia, Spain

^f Kronikgune Institute for Health Services Research, Barakaldo, Spain

^g Network for Research on Chronicity, Primary Care, and Health Promotion (RICAPPS), Spain

^h Department of Pneumology, University Hospital of Basurto, Bilbao, Spain

ⁱ Department of Pneumology, University Hospital of San Pedro Logroño, La Rioja, Spain

^j Department of Microbiology, University Hospital of Galdakao-Usansolo, Galdakao, Spain

^k Department of Biomedical Diagnostics, University Hospital of San Pedro Logroño, La Rioja, Spain

ARTICLE INFO

Article history:

Received 18 June 2022

Accepted 12 August 2022

Available online xxx

Keywords:

COVID-19

Leucocytes

Cell population data

Pneumonia

Symex XN

Biomarkers

ABSTRACT

Introduction: The main aim of this study was to assess the utility of differential white cell count and cell population data (CPD) for the detection of COVID-19 in patients admitted for community-acquired pneumonia (CAP) of different etiologies.

Methods: This was a multicenter, observational, prospective study of adults aged ≥ 18 years admitted to three teaching hospitals in Spain from November 2019 to November 2021 with a diagnosis of CAP. At baseline, a Sysmex XN-20 analyzer was used to obtain detailed information related to the activation status and functional activity of white cells.

Results: The sample was split into derivation and validation cohorts of 1065 and 717 patients, respectively. In the derivation cohort, COVID-19 was confirmed in 791 patients and ruled out in 274 patients, with mean ages of 62.13 (14.37) and 65.42 (16.62) years, respectively ($p < 0.001$). There were significant differences in all CPD parameters except MO-Y. The multivariate prediction model showed that lower NE-X, NE-WY, LY-Z, LY-WY, MO-WX, MO-WY, and MO-Z values and neutrophil-to-lymphocyte ratio were related to COVID-19 etiology with an AUC of 0.819 (0.790, 0.846). No significant differences were found comparing this model to another including biomarkers ($p = 0.18$).

Conclusions: Abnormalities in white blood cell morphology based on a few cell population data values as well as NLR were able to accurately identify COVID-19 etiology. Moreover, systemic inflammation biomarkers currently used were unable to improve the predictive ability. We conclude that new peripheral blood biomarkers can help determine the etiology of CAP fast and inexpensively.

© 2022 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

* Corresponding author.

E-mail address: ane.uranga@osakidetza.eus (A. Uranga).

◇ Please see a list of the members of the LEUCOCAP STUDY group in [Appendix A](#).

Introduction

The coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was identified in December 2019 and the outbreak was declared a pandemic on March 11 by World Health Organization.¹ It has caused over 500 million confirmed cases and over 6 million deaths.² Early diagnosis of SARS-CoV-2 infection is crucial, not only for its long incubation period with a median of 5–7 days but also because it requires isolation to prevent the spread of the disease as well as specific management.³ During the last 2 years, reverse transcription polymerase chain reaction (RT-PCR) tests for SARS-CoV-2 have been necessary for the diagnosis leading to a significant increase in healthcare costs. Both access to these tests and the time to test results vary between regions. In addition, if the result is negative but the suspicion of COVID-19 remains high, the test tends to be repeated, possibly also requiring the collection of additional samples.⁴

Abnormalities in white blood cell morphology (size, shape and composition) have been observed and related to response to infection.⁵ Actually, in sepsis there are functional changes in certain monocytes subsets and it is reflected in a variation in the morphology of monocytes.⁶ Immune dysregulation characterizes SARS-CoV-2 infection leading to certain features in differential white blood cell count such as neutrophilia or lymphopenia and morphological abnormalities, which might be useful for early recognition of the infection.⁷ Over the past few decades, hematology analyzers have undergone rapid development due to technological advances, allowing new parameters to be reported along with the basic complete blood count (CBC). Cell population data (CPD) are background parameters providing quantitative information on the morphological and functional characteristics of blood cells. In particular, various characteristics of blood cells like granularity, volume, and nucleic acid/protein content describe their morphology.

This detailed study of white blood cells and their morphological variations, which can be quantified with numerical values from the CPD, might be useful for COVID-19 screening.⁸ In this context, the aim of this study was to assess the utility of these parameters for the detection of COVID-19 in patients admitted for community-acquired pneumonia (CAP) of different etiologies.

Methods

Study design

This was a multicenter prospective observational study of patients hospitalized with CAP from November 2019 to November 2021. The study was approved by the Regional Ethics Committee (PI2019090, 17 July 2019) and registered in *ClinicalTrials.gov*: NCT 04930926. All participants provided informed consent before their inclusion and after being informed about the study and having discussed its goals, risks, and potential benefits.

Setting and study population

This study was carried out in three hospitals in the north of Spain, part of the national public health service network. Hospitalized consecutive patients diagnosed with CAP were recruited over a period of 2 years. Eligible patients were ≥ 18 years old and hospitalized with a diagnosis of CAP. Pneumonia was defined as pulmonary infiltrate on chest X-ray not seen previously, plus at least one symptom compatible with pneumonia such as cough, fever, dyspnea, and/or chest pain.⁹ Patients were excluded if they had been discharged from an acute care hospital, an onsite subacute care unit,

or a palliative care unit within the previous 14 days, or were HIV positive or chronically immunosuppressed (defined as immunosuppression for solid organ transplantation, having undergone a splenectomy, receiving ≥ 10 mg/d of prednisone or equivalent for >30 days, taking other immunosuppressive agents, or having neutropenia, i.e., neutrophil count $< 10^9/L$).

Data collection

Demographic and clinical data for each patient were collected at baseline from medical records and included comorbidities, physical examination findings, radiological presentation, and laboratory test results, as well as complications during admission. Disease severity was assessed using Confusion, Urea, Respiratory rate, Blood pressure, and age > 65 (CURB-65) and Severe Community-Acquired Pneumonia (SCAP) scores, calculated within the first 24 h after diagnosis.^{10,11} All patients were managed according to clinical guidelines that ensured the prospective and systematic collection of relevant clinical information.

Patients with COVID-19 were diagnosed following the current standards, based on positive results in real-time RT-PCR and/or antigen tests for SARS-CoV-2 in throat swab specimens. Additionally, sputum or urinary antigen tests were performed and other viruses, bacteria and atypic microorganisms were screened by standard laboratory methods.

At baseline, a Sysmex XN-20 analyzer was used to obtain detailed information related to the morphological and functional characteristics of white blood cells based on the internal cellular complexity (NE-X, LY-X, MO-X, NE-WX, LY-WX and MO-WX), cell size (NE-Z, LY-Z, MO-Z, NE-WZ, LY-WZ and MO-WZ) and quantity of nucleic acid (NE-Y, LY-Y, MO-Y, NE-WY, LY-W, and MO-WY). The membrane composition of activated cells is different from that of resting cells. The optical signals are different which allows to distinguish the morphological changes that have occurred and that are directly related to the functionality and which can be quantified with numerical values from the CPD. A detailed explanation of the cell population data is given in [Supplementary table* 1](#).

Assessment of outcomes

The primary outcome was the etiology of CAP, to distinguish between SARS-CoV-2 infection and other etiologies among patients admitted for CAP. SARS-CoV-2 infection was confirmed by real-time RT-PCR and/or antigen tests for SARS-CoV-2 in throat swab samples.

Statistical analysis

As a first step, the sample of eligible patients included in the analysis was split into two cohorts: derivation (60% of the overall sample) and validation (40%) sets.^{12,13} An exploratory data analysis was conducted of the characteristics of the sample by (1) cohort type and (2) SARS-CoV-2 positive status, calculating mean and standard deviations (or median and interquartile ranges) for continuous variables and frequencies and percentages for categorical variables. The non-parametric Wilcoxon test (for continuous variables) and the Chi-square test or Fisher's exact test as appropriate (for categorical ones) were performed to assess the homogeneity of data on demographic and clinical features.

In the derivation cohort, we first performed univariate analysis to evaluate the relationship of the demographic, clinical and laboratory CPD features with the presence of COVID-19. Again, the non-parametric Wilcoxon test (for continuous variables) and the Chi-square test or Fisher's Exact test as appropriate (for categorical ones) were used. Furthermore, the optimal cut-off points for each of the CPD that best predicted SARS-CoV-2 positive status were

Table 1
Descriptive statistics of the sociodemographic and clinical data by SARS-CoV-2 status in the derivation sample (n = 1065).

	SARS-CoV-2 STATUS		Total (n = 1065)	p-Value
	Negative (N = 274)	Positive (N = 791)		
Sex				0.234
Male	166 (24.52)	511 (75.48)	677 (63.57)	
Female	108 (27.84)	280 (72.16)	388 (36.43)	
Age, mean (sd)	65.42 (16.62)	62.13 (14.37)	63.07 (15.11)	<0.001
Living in residential care	9 (40.91)	13 (59.09)	22 (2.07)	0.14
Charlson Comorbidity Index				0.08
0	157 (23.82)	502 (76.18)	659 (61.88)	
1	56 (25.23)	166 (74.77)	222 (20.85)	
2	29 (32.22)	61 (67.78)	90 (8.45)	
≥3	32 (34.04)	62 (65.96)	94 (8.83)	
SCAP score, mean (sd)	7.47 (7.38)	7.91 (7.04)	7.80 (7.12)	0.64
SCAP				0.33
<2	82 (21.64)	297 (78.36)	379 (67.80)	
≥2	46 (25.56)	134 (74.44)	180 (32.20)	
CURB-65 score, mean (sd)	1.29 (1.09)	0.96 (0.91)	1.05 (0.97)	<0.001
CURB-65				<0.001
≤2	145 (22.21)	508 (77.79)	653 (93.15)	
>2	26 (54.17)	22 (45.83)	48 (6.85)	
Complications ^a				
Death within 30 days	12 (4.38)	54 (6.83)	66 (6.20)	0.19
Invasive mechanical ventilation	10 (3.65)	74 (9.36)	84 (7.89)	0.002
Use of vasopressors	7 (2.55)	51 (6.45)	58 (5.45)	0.01
Intensive care unit admission	20 (7.30)	139 (17.57)	159 (14.93)	<0.001
Noninvasive mechanical ventilation	35 (12.77)	238 (30.09)	273 (25.63)	<0.001
Readmission within 30 days	27 (9.87)	47 (5.94)	74 (6.95)	0.04

CURB65, Confusion, Urea, Respiratory rate, Blood pressure, and age > 65; SCAP, Severe Community-Acquired Pneumonia.

Values are expressed as frequencies (row percentages).

^a Values are expressed as frequencies (column percentages).

determined, and the relationship of the categorized CPD values with this status was expressed using odds ratios (ORs) and their corresponding 95% confidence intervals (CIs).

As a next step, a multivariate logistic regression model was developed to identify any independent predictors of SARS-CoV-2 status. To do so, variables with a p-value in the previous step of <0.20 were included as explanatory variables. Using a backward procedure, the final multivariate model was obtained. The ORs and 95% CIs of all selected variables are reported. Collinearity among the final explanatory variables was evaluated by means of the correlations and the variance inflation factor (VIF). VIF values > 3 will be considered as indicative of collinearity.¹⁴ The beta coefficients from this model were used to weight the relative importance of each variable for calculating the prediction score. Specifically, weights were calculated as the beta coefficient for each predictor in the model divided by the variable with the lowest beta coefficient and rounded to the nearest whole number.¹⁵ Similarly, from those CPD features more highly associated with the outcome of interest we developed a reduced multivariable version.

From the weights derived from the beta coefficient weights of the variables obtained in the multivariable model in the derivation sample we developed a continuous prediction score for each patient, and three risk categories were established (mild, moderate and severe) based on predicted versus observed SARS-CoV-2 status. We tested the continuous and categorized risk scores in the validation cohort. The same was also done for the reduced model.

Model robustness was assessed in terms of discrimination and calibration. Model discrimination was evaluated by computing the area under the receiver operating characteristic curve (AUC), considering an AUC value > 0.70 to indicate good discrimination.¹⁶ The AUC value was obtained after having performed bootstrapping with 2000 samples. The Hosmer–Lemeshow test was used to gauge the

ability of the model to match predicted and observed SARS-CoV-2 status in the logistic models.

All the statistical procedures were performed using SAS for Windows, version 9.4 (SAS Institute, Cary, NC). Figures were plotted using R version 4.4.0. All effects were considered significant at $p < 0.05$ unless stated otherwise.

Results

In this study, 1782 patients were included for analysis from three different teaching hospitals, 1065 in the training set and 717 patients in the validation set (Supplementary Fig. 1). Overall, 71.36% were positive for SARS-CoV-2 in the first set and 71.69% in the second. There were no statistically significant differences between these cohorts in baseline characteristics or main outcomes (Supplementary table* 2). Table 1 summarizes sociodemographic and clinical data by SARS-CoV-2 status in the derivation cohort. In the derivation cohort, COVID-19 was confirmed in 791 and ruled out in 274 patients, with mean ages of 62.13 (14.37) and 65.42 (16.62), respectively ($p < 0.001$). Compared to SARS-CoV-2 negative (non-COVID-19) patients, SARS-CoV-2 positive (COVID-19) patients were less severe as measured by the CURB-65 while there were no differences in severity as measured by the SCAP. The 30-day mortality rate was 6.20%, corresponding to 54 (6.83%) patients who were SARS-CoV-2 positive and 12 (4.38%) who were not ($p = 0.19$). Patients with COVID-19 were more likely to be admitted to the intensive care unit (ICU), receive mechanical ventilation, vasopressors, and/or noninvasive ventilation and be readmitted within 30 days (Table 1). Etiology distribution and coinfection data are shown in Table 2. *S. pneumoniae* was the most commonly isolated bacteria in both groups, but was more frequent among non-Covid-19 patients.

Table 2
Etiology distribution by SARS-CoV-2 status in the derivation sample.

	Non-COVID-19 patients (n = 449)	COVID-19 patients (n = 1333)	Total (n = 1782)	p-Value*
Type of positivity				
Viral	14 (3.12)	6 (0.45)	20 (1.12)	<0.001
Bacterial	97 (21.60)	76 (5.70)	173 (9.71)	<0.001
Virus (n = 20)*				
Adenovirus E	2 (14.29)	0 (0)	2 (10)	0.33
Bocavirus	0	0	0 (0)	-
Influenza A	5 (35.71)	0 (0)	5 (25)	0.26
Metapneumovirus	0	0	0 (0)	-
Parainfluenza	3 (21.43)	6 (100)	9 (45)	0.002
Rhinovirus	3 (21.43)	0 (0)	3 (15)	0.52
Respiratory syncytial virus	1 (7.14)	0 (0)	1 (5)	0.50
Influenza B	1 (7.14)	0 (0)	1 (5)	0.50
Bacteria (n = 173)*				
<i>S. pneumoniae</i>	42 (43.30)	19 (25)	61 (35.26)	0.02
<i>H. influenzae</i>	3 (3.09)	5 (6.58)	8 (4.62)	0.30
<i>Staphylococcus aureus</i>	4 (4.12)	10 (13.16)	14 (8.09)	0.05
<i>P. aeruginosa</i>	6 (6.19)	3 (3.95)	9 (5.20)	0.73
<i>Escherichia coli</i>	9 (9.28)	5 (6.58)	14 (8.09)	0.58
<i>K. pneumoniae</i>	5 (5.15)	1 (1.32)	6 (3.47)	0.23
<i>Serratia marcescens</i>	0 (0)	0 (0)	0 (0)	-
<i>L. pneumophila</i>	14 (14.43)	10 (13.16)	24 (13.87)	0.81
<i>M. pneumoniae</i>	3 (3.09)	11 (14.47)	14 (8.09)	0.01
<i>C. burnetti</i>	9 (9.28)	9 (11.84)	18 (10.40)	0.62
<i>C. pneumoniae</i>	11 (11.34)	13 (17.11)	24 (13.87)	0.38
<i>M. catarrhalis</i>	2 (2.06)	1 (1.32)	3 (1.73)	0.71

* Percentage calculated with respect to overall number of cases in which viruses (n = 20) or bacteria (n = 173) were detected.

Fig. 1 presents the differential white cell counts, CPD parameters and biomarker levels by SARS-CoV-2 positive status as well as cut-offs, ORs (95%CI) and p values. There were significant differences in all CPD parameters except MO-Y. In addition, COVID-19 patients presented significantly lower CRP and PCT levels, as well as a lower neutrophil/lymphocyte ratio (NLR) than non-COVID patients with a median of 5.79 (3.43–9.64) in COVID-19 patients and 9.32 (5.22–17.83) in non-COVID-19 patients (OR 3.47, [2.61, 4.61], $p < 0.001$, OR 9.41 [6.02, 14.69], $p < 0.001$, and OR 2.83 [2.13, 3.76], $p = < 0.001$, respectively). On the other hand, COVID-19 patients had a higher platelet/lymphocyte ratio (OR 1.79 [1.34, 2.37], $p < 0.001$). The white blood cell differential fluorescence (WDF) scattergram results revealed differences in the distribution of cell clusters between COVID-19 and non-COVID-19 patients, as illustrated in Fig. 2.

A weighted multivariate prediction model was developed to identify SARS-CoV-2 positive status (Table 3). Lower NE-X, NE-WY, LY-Z, LY-WY, MO-WX, MO-WY, and MO-Z values as well as NLR were related to COVID-19 etiology. The variables most strongly related to COVID-19 pneumonia were NE-WY, LY-WY, and MO-WX. This model showed good discrimination and calibration with an AUC (95% CI) of 0.819 (0.790, 0.846). No multicollinearity was present in this model (all VIF values ranged from 1.03 (NE-X) to 1.18 (NE-WY) as well as all correlation were weak). The multivariate model was compared with another one that included CRP and PCT, with an AUC value of 0.823 (0.800, 0.855) (Supplementary table* 3). Results derived from the comparison of the models indicated that differences were not statistically significant ($p = 0.18$).

In addition, a risk score based on three different groups was developed both in the derivation and validation sets, categorized as low (0–7 points), moderate (8–16 points) and high (≥ 17 points) and showing an AUC value of 0.777 (0.726, 0.825) (Table 4). Additionally, adjusted (by sex, age, Charlson Comorbidity Index and CURB-65 score) analysis was performed in the validation set.

Finally, a reduced model was developed including NE-WY, LY-WY and MO-WY with an AUC (95%) of 0.758 (0.725, 0.790) (Supplementary table* 4). Risk groups of the reduced model is

shown in Supplementary table* 5. Original model showed significantly higher predictive accuracy (AUC (95%) 0.819 (0.790, 0.846) vs 0.758 (0.725, 0.790), respectively, $p < 0.001$)

Discussion

The present study highlights the utility of commonly used white blood cell analysis to distinguish between pneumonia due to COVID-19 from pneumonia of other etiologies. Moreover, NLR and abnormalities in white blood cell morphology assessed based on a few cell population data values were able to accurately identify COVID-19 etiology in these patients over a long study period including different waves of the pandemic. In addition, current systemic inflammation biomarkers were not able to improve the predictive ability.

Complete blood count and differential white cell count have been widely studied among patients with SARS-CoV-2 infection.¹⁷ Both lymphopenia and neutrophilia have been reported in patients with COVID-19.^{18,19} In relation to this, several laboratory predictors have been explored for assessing disease severity. A recent meta-analysis identified high white blood cell count and low lymphocyte and platelet counts as markers for progression to critical illness.²⁰ Likewise, high NLR has been positively associated with higher mortality in these patients.^{21–23} In our results, NLR was associated with the presence of SARS-CoV-2 infection.

Alterations in white blood cell morphology like granularity, volume, and nucleic acid/protein content (size, shape and composition) have been observed in response to stimuli, such as infections, and may offer valuable information on the cells' state of activation and functional activity.²⁴ Activated cells can be distinguished morphologically from resting ones by various features, i.e., nucleic acid content and volume. Atypical lymphocytes circulating in blood have been reported in COVID-19 patients.²⁵

Little is known about the utility of differential white cell count and cell population data to distinguish between COVID-19 and other forms of CAP. Our data suggest that the straightforward evaluation of neutrophil activation can identify COVID-19 in

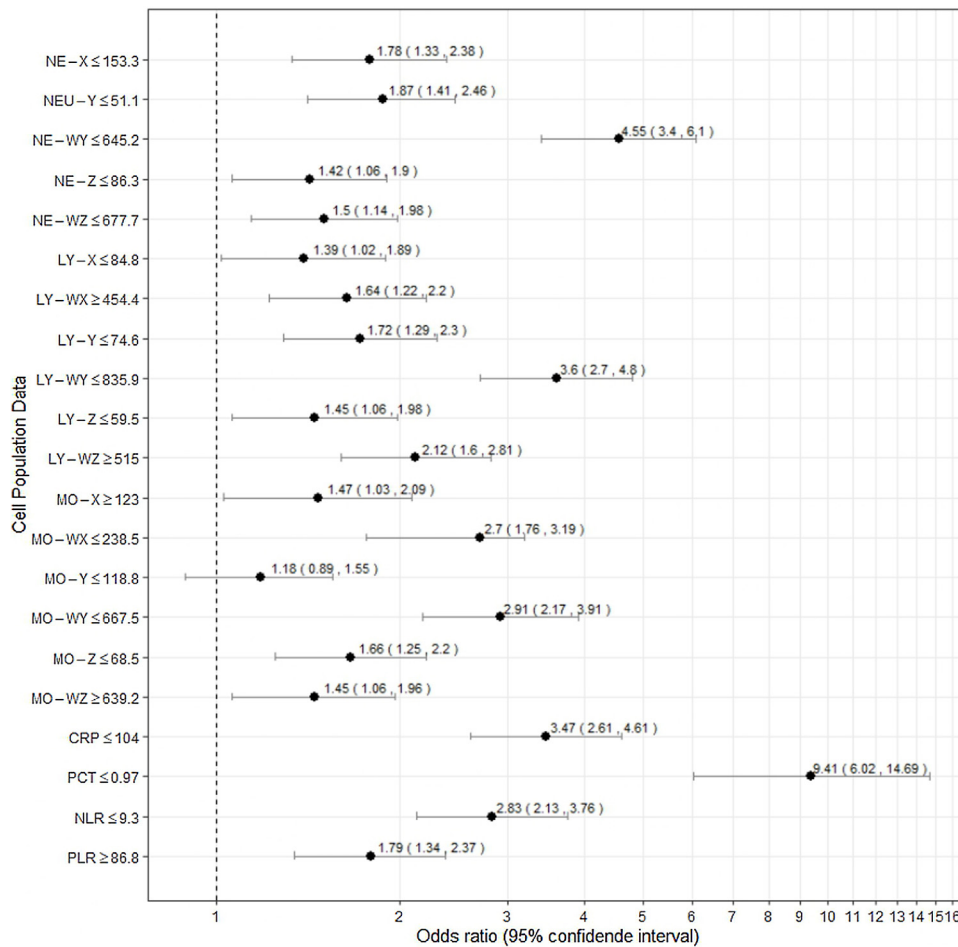


Fig. 1. Summary statistics of the laboratory tests by SARS-CoV-2 positive status in the derivation sample (n = 1065).

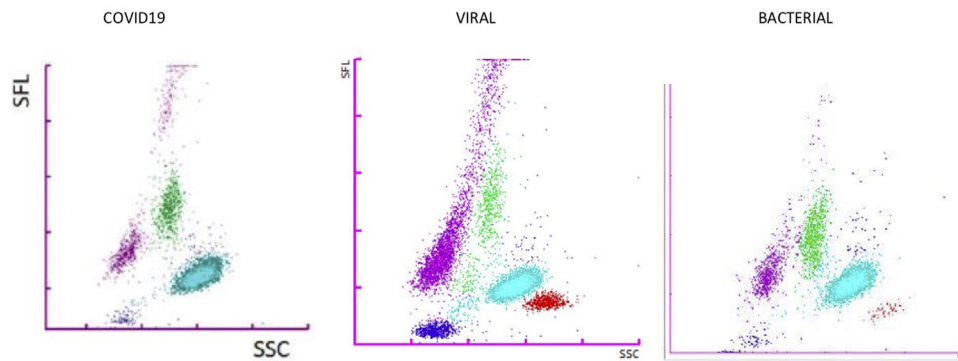


Fig. 2. The white blood cell differential fluorescence scattergram results showing differences between etiologies in cell cluster distribution. Side scatter (SSC) is shown on the X-axes and side fluorescence (SFL) on the Y-axes.

patients with CAP in a way that is both easy and inexpensive. In a retrospective study, Dai et al. constructed a machine learning model that was able to promptly distinguish COVID-19 from other forms of CAP.²⁶ Similarly, an Italian group reported basophil and eosinophil optimal cut-off values that could rule out COVID-19 in patients with CAP.²⁷ These authors focused on the ability of high values to exclude SARS-CoV-2 infection, but below proposed cut-offs, the performance to detect the disease decreased considerably.

The COVID-19 pandemic has exposed weaknesses in many health systems, even those previously considered resilient, and exacted a serious economic toll. In this context, it is crucial to ensure rational use of tests both now and in possible future waves.

Advantages of the approach we propose include the fact that CPD and differential white cell analysis are easy to perform avoiding unnecessary costs. The drawback to the use of CPD is that these are research parameters which are complex to explain to clinicians and this is why we have developed a score rather than proposing the use of numerical values from CPD. Actually, the model that included CRP and PCT, which would be more complex and expensive to obtain, did not show better accuracy. Additionally, we developed a short risk score model with just three CPD features, easier to use but with less predictive ability than the longer one.

Our study has certain strengths: first, its prospective and multicenter design; second, the large sample size including cases from

Table 3
Multivariate prediction model for identifying SARS-CoV-2 positive status in the internal derivation sample (n = 1065).

	Model 1			
	Beta (s.e.)	OR (95% CI)	p-Value	Weight
NE-X ≤ 153	0.64 (0.17)	1.90 (1.35, 2.67)	<0.001	3
NE-WY ≤ 645.5	1.27 (0.18)	3.55 (2.49, 5.06)	<0.001	6
LY-Z ≤ 59.5	0.65 (0.20)	1.91 (1.30, 2.81)	0.001	3
LY-WY ≤ 835.9	1.23 (0.17)	3.41 (2.45, 4.73)	<0.001	5
MO-WX ≤ 246	0.72 (0.18)	2.05 (1.45, 2.92)	<0.001	3
MO-WY ≤ 667.5	1.13 (0.17)	3.08 (2.20, 4.32)	<0.001	5
MO-Z ≤ 68.5	0.73 (0.18)	2.08 (1.45, 2.97)	<0.001	3
NLR ≤ 9	0.47 (0.17)	1.61 (1.14, 2.26)	0.007	2
AUC (95% CI)*		0.819 (0.790, 0.846)	0.56	
Hosmer-Lemeshow test				

AUC, area under the receiver operating characteristic curve; NLR, neutrophil-to-lymphocyte ratio.

* Value obtained after bootstrapping with 2000 samples.

Table 4
Risk groups in the derivation and validation samples (n = 1065 and n = 717 respectively).

	Derivation sample			Validation sample				
	No of events/total	OR (95% CI)	p-Value	No of events/total	Unadjusted analysis		Adjusted analysis**	
					OR (95% CI)	p-Value	OR (95% CI)	p-Value
Continuous risk score**	-	1.25 (1.21, 1.29)	<0.001	-	1.22 (1.18, 1.27)	<0.001	1.23 (1.17, 1.29)	<0.001
AUC (95% CI)		0.814 (0.784, 0.843)			0.783 (0.742, 0.823)		0.806 (0.757, 0.854)	
Hosmer-Lemeshow test		0.83			0.19		0.65	
Risk groups								
0-7	17/83 (20.48)	Reference		11/47 (23.40)	Reference		Reference	
8-16	169/297 (56.90)	5.13 (2.87, 9.16)	<0.001	121/204 (59.31)	4.77 (2.30, 9.91)	<0.001	4.87 (2.22, 10.69)	<0.001
≥17	605/685 (88.32)	29.36 (16.41, 52.54)	<0.001	410/466 (87.98)	23.96 (11.54, 49.76)	<0.001	25.04 (10.75, 58.32)	<0.001
AUC (95% CI)*		0.757 (0.725, 0.789)			0.727 (0.688, 0.764)		0.777 (0.726, 0.826)	
Hosmer-Lemeshow test		0.99			0.99		0.73	

AUC, area under the receiver operating characteristic curve.

* Value obtained after bootstrapping with 2000 samples.

** Adjusted for patient gender and age, Charlson Comorbidity Index and Confusion, Urea, Respiratory rate, Blood pressure, and age > 65 total score.

different waves of the COVID-19 pandemic; third, the analysis of coinfection; and finally, the performance of comparisons with current systemic biomarkers. On the other hand, this study has some limitations. In particular, a Sysmex XN analyzer is required to perform the analysis; however, this method is used frequently, Sysmex being the leader in the market for this type of laboratory equipment in Europe. Further research is required focusing on children, and less severe and asymptomatic patients as well as validating these data in an external cohort.

In conclusion, new peripheral blood biomarkers can help determine the etiology of CAP in a fast and inexpensive way. In a future lower incidence scenario, this could provide a useful tool, reducing the number of PCR tests performed, and providing a rapid and inexpensive diagnostic approach which could lead to faster initiation of treatment and instigation of preventive measures to reduce the risk of transmission.

Authors' contributions

AU and PPE contributed to the study design, data collection and data interpretation. AU wrote the manuscript. AU, PPE and EU took responsibility for study design, data integrity review, editing and approval the final version. UA carried out data analysis and review the manuscript. JMQ contributed to the study design, review and approval the final version. MI, CRM, CP, CS, MJLG and PS contributed substantially to acquisition and interpretation of data. All authors reviewed critically the manuscript. LEUCOCAP STUDY GROUP contributed to the data collection, data interpretation, review, editing and approval the final version.

Access to data

Please contact the first author for data requests.

Funding

This work was supported in part by grant 2018111033 from the Basque Country Health Department and grant 2019/968 from the Spanish Pulmonology and Thoracic Surgery Society.

Conflict of interest

All authors have nothing to declare.

Acknowledgments

Authors would like to thank Ideas Need Communicating Language Services for improving the use of English in the manuscript.

Appendix A. LEUCOCAP STUDY GROUP

Ane Uranga, Amaia Artaraz, Aitor Ballaz, Sandra Dorado, Silvia Pascual (Department of Pneumology, University Hospital of Galdakao-Usansolo, Galdakao, Spain), Urko Aguirre, Jose María Quintana, Ane Villanueva (Osakidetza Basque Health Service, Galdakao-Usansolo University Hospital, Research Unit, Galdakao, Spain), María Jose Lopez de Goicoechea (Department of Microbiology, University Hospital of Galdakao-Usansolo, Galdakao, Spain), Carmen Mar, Eloisa Urrechaga, Cristina Ponga (Laboratory, University Hospital of Galdakao-Usansolo, Galdakao, Spain) Ignacio Arriaga, Maider Intxausti, (Department of Pneumology,

University Hospital of Basurto, Bilbao, Spain), Domingo Fernandez (Department of Microbiology, University Hospital of Basurto, Bilbao, Spain), Itziar Benito (Laboratory, University Hospital of Basurto, Bilbao, Spain), Carlos Ruiz-Martinez, Javier Ugeda (Department of Pneumology, University Hospital of San Pedro Logroño, La Rioja, Spain), Pilar Sanz, Ivan Bernardo (Department of Biomedical Diagnostics, University Hospital of San Pedro Logroño, La Rioja, Spain), Pedro Pablo España (Department of Pneumology, University Hospital of Galdakao-Usansolo, Galdakao, Spain).

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.arbres.2022.08.011.

References

- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395:497–506. [http://dx.doi.org/10.1016/S0140-6736\(20\)30183-5](http://dx.doi.org/10.1016/S0140-6736(20)30183-5).
- World Health Organization. COVID-19 weekly epidemiological update. Edition 88, published 20 April 2022. Available from: <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19-20-april-2022>.
- Sharma A, Ahmad Farouk I, Lal SK. COVID-19: a review on the novel coronavirus disease evolution, transmission, detection, control and prevention. *Viruses*. 2021;13:202. <http://dx.doi.org/10.3390/v13020202>. PMID: 33572857; PMCID: PMC7911532.
- Lai CKC, Lam W. Laboratory testing for the diagnosis of COVID-19. *Biochem Biophys Res Commun*. 2021;538:226–30. <http://dx.doi.org/10.1016/j.bbrc.2020.10.069>.
- Urrechaga E, Aguirre U, España PP, García de Gadiana L. Complete blood counts and cell population data from Sysmex XN analyser in the detection of SARS-CoV-2 infection. *Clin Chem Lab Med*. 2020;59:e57–60. <http://dx.doi.org/10.1515/ccim-2020-1416>. PMID: 33079699.
- Yona S, Jung S. Monocytes: subsets, origins, fates and functions. *Curr Opin Hematol*. 2010;17:53–9.
- Jamal M, Bangash HI, Habiba M, Lei Y, Xie T, Sun J, et al. Immune dysregulation and system pathology in COVID-19. *Virulence*. 2021;12:918–36. <http://dx.doi.org/10.1080/21505594.2021.1898790>. PMID: 33757410; PMCID: PMC7993139.
- Osman J, Lambert J, Templé M, Devaux F, Favre R, Flaujac C, et al. Rapid screening of COVID-19 patients using white blood cell scattergrams, a study on 381 patients. *Br J Haematol*. 2020;190:718–22. <http://dx.doi.org/10.1111/bjh.16943>.
- Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med*. 1997;336:243–50. <http://dx.doi.org/10.1056/NEJM199701233360402>. PMID: 8995086.
- Lim WS, Baudouin SV, George RC, Hill AT, Jamieson C, Le Jeune I, et al. Pneumonia Guidelines Committee of the BTS Standards of Care Committee BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax*. 2009;64:1–55.
- Espana PP, Capelastegui A, Gorordo I, Esteban C, Mikel Oribe M, Ortega M, et al. Development and validation of a clinical prediction rule for severe community-acquired pneumonia. *Am J Respir Crit Care Med*. 2006;174:1249–56.
- Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–31. <http://dx.doi.org/10.1093/eurheartj/ehu207>.
- Kuhn M, Johnson K. *Applied predictive modeling*. New York: Springer; 2018.
- Kock N, Lynn G. Lateral collinearity and misleading results in variance-based sem: an illustration and recommendations. *J Assoc Inf Syst*. 2012;13:doi:10.17705/1jais.00302.
- Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med*. 2004;23:1631–60. <http://dx.doi.org/10.1002/sim.1742>. PMID: 15122742.
- Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York, NY: John Wiley and Sons; 2000. p. 160–4 [chapter 5].
- Zini G, bellesi S, Ramundo F, d'Onofrio G. Morphological anomalies of circulating blood cells in COVID-19. *Am J Hematol*. 2020;95:870–2.
- Frater JL, Zini G, d'Onofrio G, Rogers HJ. COVID-19 and the clinical hematology laboratory. *Int J Lab Hematol*. 2020;42 Suppl. 1:11–8. <http://dx.doi.org/10.1111/ijlh.13229>. PMID: 32311826; PMCID: PMC7264622.
- Fan BE, Chong VCL, Chan SSW, Lim GH, Lim KGE, Tan GB, et al. Hematologic parameters in patients with COVID-19 infection. *Am J Hematol*. 2020;95:E131–4. <http://dx.doi.org/10.1002/ajh.25774>. Epub 2020 Mar 19; Erratum in: *Am J Hematol*. 2020 Nov;95(11):1442. PMID: 32129508.
- Henry BM, de Oliveira MHS, Benoit S, Plebani M, Lippi G. Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (COVID-19): a meta-analysis. *Clin Chem Lab Med*. 2020;58:1021–8. <http://dx.doi.org/10.1515/ccim-2020-0369>. PMID: 32286245.
- Yang AP, Liu JP, Tao WQ, Li HM. The diagnostic and predictive role of NLR, d-NLR and PLR in COVID-19 patients. *Int Immunopharmacol*. 2020;84:106504. <http://dx.doi.org/10.1016/j.intimp.2020.106504>. Epub 2020 Apr 13; PMID: 32304994; PMCID: PMC7152924.
- Lagunas-Rangel FA. Neutrophil-to-lymphocyte ratio and lymphocyte-to-C-reactive protein ratio in patients with severe coronavirus disease 2019 (COVID-19): a meta-analysis. *J Med Virol*. 2020;92:1733–4.
- Sun S, Cai X, Wang H, He G, Lin Y, Lu B, et al. Abnormalities of peripheral blood system in patients with COVID-19 in Wenzhou, China. *Clin Chim Acta*. 2020;507:174–80. <http://dx.doi.org/10.1016/j.cca.2020.04.024>. Epub 2020 Apr 24; PMID: 32339487; PMCID: PMC7194694.
- Urrechaga E. Reviewing the value of leukocytes cell population data (CPD) in the management of sepsis. *Ann Transl Med*. 2020;8:953. <http://dx.doi.org/10.21037/atm-19-3173>. PMID: 32953753; PMCID: PMC7475430.
- Pakos IS, Lo KB, Salacup G, Pelayo J, Bhargav R, Peterson E, et al. Characteristics of peripheral blood differential counts in hospitalized patients with COVID-19. *Eur J Haematol*. 2020;105:773–8. <http://dx.doi.org/10.1111/ejh.13509>. Epub 2020 Sep 20; PMID: 32794205; PMCID: PMC7436857.
- Dai W, Ke P, Li Z, Zhuang Q, Huang W, Wang Y, et al. Establishing classifiers with clinical laboratory indicators to distinguish COVID-19 from community-acquired pneumonia: retrospective cohort study. *J Med Internet Res*. 2021;23:e23390. <http://dx.doi.org/10.2196/23390>. <https://www.jmir.org/2021/2/e23390>
- Biolo G, Fiotti N, Cominotto F, Di Girolamo FG, Panizon E, Altamura N, et al. Ruling out coronavirus disease 2019 in patients with pneumonia: the role of blood cell count and lung ultrasound. *J Clin Med*. 2021;10:3481. <http://dx.doi.org/10.3390/jcm10163481>. PMID: 34441777; PMCID: PMC8397060.