Atypical lymphoid cells circulating in blood in COVID-19 infection: morphology, immunophenotype and prognosis value

Anna Merino [1,2] Alexandru Vlagea,[3] Angel Molina,[1,2] Natalia Egri,[3] Javier Laguna [1,2] Kevin Barrera,[4] Laura Boldú [1,2] Andrea Acevedo,[4] Mar Díaz-Pavón,[3] Francesc Sibina,[3] Francisca Bascón,[1] Oriol Sibila,[5] Manel Juan,[3] José Rodellar[4]

[1]Core Laboratory, Biomedical Diagnostic Center, Hospital Clinic de Barcelona, Barcelona, Spain
[2]Biochemistry and Molecular Genetics, Biomedical Diagnostic Center, Hospital Clinic de Barcelona, Barcelona, Spain
[3]Department of Immunology, Biomedical Diagnostic Center, Hospital Clinic de Barcelona, Barcelona, Spain
[4]Department of Mathematics, Universitat Politecnica de Catalunya, Barcelona, Spain
[5]Institut Clínic del tórax, Hospital Clinic de Barcelona, Barcelona, Spain

Correspondence to
Anna Merino, Department of Biochemistry and Molecular Genetics; Core Laboratory; Biomedical Diagnostic Center, Hospital Clinic de Barcelona, Barcelona, Spain; amerino@clinic.cat

AM and AV are joint first authors.

Received 4 September 2020
Revised 15 October 2020
Accepted 24 November 2020

ABSTRACT

Aims Atypical lymphocytes circulating in blood have been reported in COVID-19 patients. This study aims to (1) analyse if patients with reactive lymphocytes (COVID-19 RL) show clinical or biological characteristics related to outcome; (2) develop an automatic system to recognise them in an objective way and (3) study their immunophenotype.

Methods Clinical and laboratory findings in 36 COVID-19 patients were compared between those showing COVID-19 RL in blood (18) and those without (18). Blood samples were analysed in Advia2120i and stained with May Grünwald-Giemsa. Digital images were acquired in CellaVisionDM96. Convolutional neural networks (CNNs) were used to accurately recognise COVID-19 RL. Immunophenotypic study was performed throughflow cytometry.

Results Neutrophils, D-dimer, procalcitonin, glomerular filtration rate and total protein values were higher in patients without COVID-19 RL (p<0.05) and four of these patients died. Haemoglobin and lymphocyte counts were higher (p<0.02) and no patients died in the group showing COVID-19 RL. COVID-19 RL showed a distinct deep blue cytoplasm with nucleus mostly in eccentric position. Through two sequential CNNs, they were automatically distinguished from normal lymphocytes and classical RL with sensitivity, specificity and overall accuracy values of 90.5%, 99.4% and 98.7%, respectively. Immunophenotypic analysis revealed COVID-19 RL are mostly activated effector memory CD4 and CD8 T cells.

Conclusion We found that COVID-19 RL are related to a better evolution and prognosis. They can be detected by morphology in the smear review, being the computerised approach proposed useful to enhance a more objective recognition. Their presence suggests an abundant production of *virus-specific T cells*, thus explaining the better outcome of patients showing these cells circulating in blood.

INTRODUCTION

COVID-19 sustained by the SARS-CoV-2 has expanded in all continents.[1 2] COVID-19 includes respiratory symptoms, which may be mild in most patients, although some of them may suffer from a serious acute respiratory distress syndrome that can lead to death. Laboratory medicine plays an essential role in its early detection, diagnosis and management.[3] Several biomarkers have been described to be related to severe COVID-19, such as increased values of C-reactive protein, procalcitonin, alkaline phosphatase (AP), lactate dehydrogenase (LDH), alanine aminotransferase (ALAT), bilirubin, blood urea nitrogen and creatinine and cardiac troponin.[4]

Among haematology laboratory parameters, low lymphocyte count is frequent, which is probably related to the deficient immune response to the virus.[5] Nevertheless, some variability in lymphopoenia presentation has been associated with COVID-19,[3–6] as well as leucocytosis and neutrophilia,[3] increased neutrophil/lymphocyte ratio (NLR),[7] thrombocytopoenia,[8] atypical coagulation parameters and high values of D-dimer and fibrin/fibrinogen degradation products.[9]

Peripheral blood (PB) morphology review shows the presence of new atypical reactive lymphocytes (RL) circulating in blood[4 10–13] in some SARS-CoV-2-infected patients. In this paper, they are abbreviated as COVID-19 RL. It has been reported that these cells morphologically mimic RL of Epstein–Barr virus or cytomegalovirus infections.[10] However, COVID-19 RL show subtle morphological differences, such as more basophilic cytoplasm and occasional presence of small cytoplasmic vacuoles.[11–13] Morphological discrimination between COVID-19 RL and RL seen in other infections is a challenge. For the sake of clarity, these RL are referred as 'classical' in this paper.

PB smear review is based on visual inspection, which is time-consuming, requires well-trained personnel and is prone to subjectivity and intraobserver variability.[14] Image analysis and machine learning are technological tools increasingly used in medicine, particularly in haematopathology.[15] In a previous work, we successfully applied convolutional neural networks (CNNs) to automatically classify blood cell images.[16] Since CNNs are multilayered architectures able to extract complex and high-dimensional features from images,[17] they might be highly sensitive and specific for COVID-19 RL recognition.

The relationship of COVID-19 RL and the evolution and prognosis of the disease has not been investigated so far. Despite the large numbers of cases and deaths, information on the immunophenotype

of SARS-CoV-2-specific cells is scarce. The objective of this work is threefold: (1) analyse if patients in which COVID-19 RL are detected show particular clinical or biological characteristics related to the evolution and prognosis of the disease; (2) develop an automatic system to characterise and recognise these lymphoid cells in an objective way and (3) analyse COVID-19 RL's immunophenotype to investigate their role in patient's outcome.

MATERIAL AND METHODS
Patients
A number of 36 COVID-19 patients were studied during their stay at Hospital Clinic of Barcelona. They were arranged in two groups: positive and negative. The positive group included 18 patients (13 males and 5 females) showing COVID-19 RL in PB. The negative group had the remaining 18 patients (11 males and 7 females) showed neither COVID-19 RL nor classical RL. Diagnoses were confirmed by positive *real-time reverse-transcription PCR*.

Laboratory parameters and statistical analysis
Blood samples were collected in EDTA and analysed in the Advia2120i. PB smears were stained with May Grünwald-Giemsa and digital cell images (363×360 pixels) were acquired by the CellaVisionDM96 (CellaVision, Lund, Sweden). All clinical and laboratory findings were compared between both groups of patients. Time (in days) between onset of symptoms and collection of samples was practically the same for both groups. A single sample was obtained from each patient.

Full blood cell parameters and counts were evaluated. Absolute numbers of NLR were calculated. Other tests included prothrombin time, D-dimer and biochemical markers, such as C-reactive protein, procalcitonin, AP, LDH, ALAT, aspartate aminotransferase, gamma glutamyl transpeptidase, bilirubin and ferritin. RL in COVID-19 patients were identified by pathologists according to their characteristic morphology. Statistical analyses were conducted using Shapiro-Wilk, Student's t parametric test and Wilcoxon non-parametric tests with R software.[18] P values<0.05 were considered statistically significant.

Development of an automatic classification system
CNNs have been successfully applied in the automatic classification of normal and abnormal leukocytes in PB.[17 19] Based on our previous works, we proposed the sequential classification scheme with two CNN models working in series, as shown in figure 1. The first CNN was trained to distinguish between normal and RL, which included both COVID-19 RL and classical RL in a single category. The second CNN was trained to discern between both classes of RL.

We used an initial set of 7555 images. A number of 187 COVID-19 RL images were collected from the 18 patients of the positive group. Images of normal lymphocytes (4928) and classical RL (2340) were collected from healthy controls and patients with other infections, respectively, which were used by the research group in previous works.[20 21] The overall set was split into two subsets: 80% was randomly selected for training the models, while the remaining 20% was saved for their assessment (1491 images). It is common to use higher proportions of images for training than for testing, typically between 70% and 80% in most practical applications. The reason is to use more information to adjust models and keep enough information for evaluating the trained models. In this work, we selected
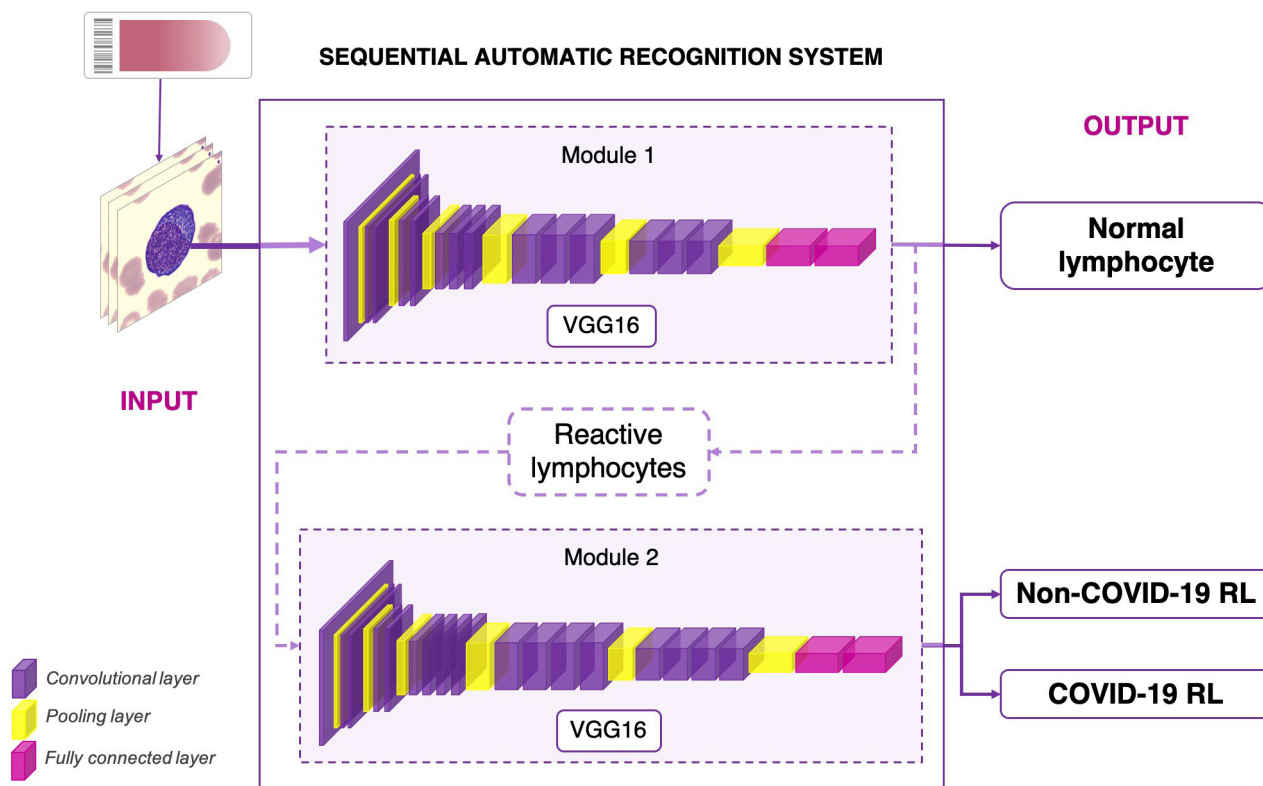

Figure 1 Block diagram that illustrates the sequential classification of normal lymphocyte and reactive lymphocyte (RL) images in non-COVID-19 and COVID-19 infection based on morphology. Using the first CNN model, normal lymphocytes are excluded from the RL group. In the following step, the second CNN discerns between COVID-19 RL and non-COVID-19 RL. CNN, convolutional neuralnetworks.

80%–20% for training and test subsets since, after some preliminary trials, we obtained the best performance measures.

A CNN has a modular structure that can be explained in two parts. The first part combines the following elements: (1) the input layer, which reads the pixels contained in the images; (2) a number of convolutional layers able to detect specific patterns and extract quantitative features of the images (feature maps); (3) pooling layers, which reduce the size of feature maps, while preserving relevant information and eliminating irrelevant details. Through subsequent passing of the input image along the different layers, the final result is a set of relevant quantitative features that represent the image. The second part is formed by a number of fully connected layers as those used in a regular neural network. This part is trained to learn how to combine the obtained features to perform the final classification of the input image. This is done by assigning a probability to each possible class and predict the class with the highest score.

In general, training of CNN models requires a balanced availability of images from all classes. To cope with the unbalanced proportions of COVID-19 RL images, data augmentation was performed. It consists on randomly applying transformations to the original images, such as vertical and horizontal flips and rotations.[22] With this up-sampling, we finally arranged a data set with 5000 images of normal lymphocytes and 5000 of RL, from which 2500 images were non-COVID-19 RL and 2500 were COVID-19 RL.

Training is an iterative process, where in each iteration all the images of the training set are processed forward by the network. The classification outputs are compared with the ground truth assigned by the clinical pathologists and used to calculate a loss function to quantify the error. Cross-entropy was the loss function used in this work. In a second step, the error is propagated backwards to update the parameters (weights) involved in the network using the gradient descent algorithm to minimise the loss function. Using the updated weights, in the end of each iteration, the images are passed through the network. The objective is to check the performance of the model using the loss function and also the accuracy obtained in the classification of the validation images (proportion of images correctly classified). In this work, we used the one cycle learning rate policy to obtain optimal classification results with fewer iterations. Following the same learning scheme, we obtained an accuracy of 99% for each CNN classifier.

We analysed several CNN architectures already pretrained with the ImageNet database.[23–25] VGG16 architecture was selected for both CNN (see figure 1) according to the following criteria: (1) they showed the best accuracy, which is the proportion of images correctly classified; and (2) this architecture is simpler compared with the other CNN models and had the best classification speed, which is an advantage for a potential real-time implementation.

After the development stage, the system was assessed with the testing data set (see Results section).

Immunophenotypic study

We selected the population of large lymphocytes to perform the immunophenotypic study, since COVID-19 RL cells are morphologically large and complex lymphocytes. For the characterisation of these large lymphocytes, we used flow cytometry

Table 1 Clinical characteristics, therapies and the number of patients who needed intensive care unit and mechanical ventilation in patients with and without COVID-19 atypical lymphocytes (COVID-19 LY) circulating in peripheral blood (n=36)

			Patients with COVID-19 LY n=18	Patients without COVID-19 LY n=18
Age (years)			53 (30–88)	74 (46–90)
Sex	Men		13 (72%)	11 (61%)
	Women		5 (28%)	7 (39%)
Symptoms	Fever		18 (100%)	16 (89%)
	Cough		13 (72%)	14 (78%)
	Dyspnoea		10 (56%)	9 (50%)
	Nausea		0 (0%)	2 (11%)
	Vomiting		0 (0%)	2 (11%)
	Diarrhoea		1 (6%)	1 (6%)
	Myalgia		3 (17%)	2 (11%)
	Dysgeusia		4 (22%)	0 (0%)
	Anosmia		4 (22%)	0 (0%)
Death			0 (0%)	4 (22%)
Drugs	Antibiotic therapy		17 (94%)	18 (100%)
	Antiviral therapy		16 (89%)	16 (89%)
	Antiparasitic therapy		17 (94%)	17 (94%)
	Antifungal therapy		0 (0%)	1 (6%)
	Immunosuppression	*Tocilizumab*	3 (17%)	7 (39%)
		Tocilizumab+Anakinra	2 (11%)	2 (11%)
		Tocilizumab+Dexamethasone		1 (5%)
		Siltuximab		1 (5%)
		Anakinra		1 (5%)
Intensive care unit	Yes		1 (6%)	9 (50%)
	No		17 (94%)	9 (50%)
Mechanical ventilation	Yes		1 (6%)	8 (44%)
	No		17 (94%)	10 (56%)

using PB samples from 13 patients of the positive group, 4 of the negative group and 5 samples of healthy controls. After compensation, data acquisition was performed with a BD FACSCanto II flow cytometer. For analysis, BD FACSDiva (Becton Dickinson, Franklin Lakes, NJ) software was used. A minimum of 100 000 events or 30 000 T cells were acquired.

RESULTS
Clinical and biological data of patients
Median values and SD of age (years) were 53 ± 16 in patients with COVID-19 RL (positive group), and 74 ± 13 in patients of the negative group, p<0.00009. Most frequent initial clinical symptoms included fever (94%), cough (75%), dyspnoea (53%), myalgia (14%), anosmia (11%), dysgeusia (11%), diarrhoea (6%), nausea and vomiting (6%). Myalgia, anosmia and dysgeusia were present exclusively in the negative group (table 1).

Positive patients showed lower absolute neutrophil counts ($\mu=2.9\times10^9$/L) and higher absolute lymphocyte counts ($\mu=1.6\times10^9$/L) than negative patients ($\mu=8.1\times10^9$/L and $\mu=0.8\times10^9$/L), p=0.04 and p=0.01, respectively. NLR showed significant increased values in negative patients ($\mu=19.2$) as compared with positive ones ($\mu=2.2$), p=0.0002 (table 2). Large unstained cells greater than 5% or atypical lymphocyte flags on the Advia2120i were found in the positive group.

Table 2 Haematological and coagulation and biochemistry parameters in patients with and without COVID-19 reactive lymphocytes (COVID-19 RL)

		Patients with COVID-19 RL		Patients without COVID-19 RL		
	Parameter	Mean	SD	Mean	SD	P
Haematological	RBC (10^{12}/L)	4.54	0.73	3.27	0.82	<0.001*
	Haemoglobin (g/L)	136	22	101	25	<0.001*
	Haematocrit (L/L)	0.405	0.060	0.318	0.080	0.001*
	MCV (fL)	89.6	4.9	97.2	5.4	<0.001*
	MCH (pg)	30.1	1.5	31.0	1.4	0.058
	MCHC (g/L)	335.7	8.1	319.4	9.2	<0.001*
	Platelets (10^9/L)	268	148	202	121	0.090
	MPV (fL)	8.3	1.0	9.9	2.1	0.007*
	MPC (gL)	25.0	2.3	23.9	2.0	0.149
	MPM (pg)	2.75	3.36	2.09	0.29	0.326
	WBC (10^9/L)	5.3	4.4	9.8	8.8	0.235
	Neutrophils (10^9/L)	2.9	3.4	8.1	7.9	0.044*
	Lymphocytes (10^9/L)	1.6	1.5	0.8	0.7	0.012*
	NLR	2.2	1.6	19.5	32.2	<0.001*
	Monocytes (10^9/L)	0.39	0.25	0.54	0.53	0.620
	Eosinophils (10^9/L)	0.20	0.43	0.07	0.10	0.127
	Basophils (10^9/L)	0.02	0.04	0.02	0.05	0.961
	LUC (10^9/L)	0.16	0.18	0.18	0.24	0.702
	PMN (%)	55.5	13.2	76.5	13.1	<0.001*
	MN (%)	43.4	13.1	22.7	13.0	<0.001*
	PMN/MN ratio	1.5	0.8	4.7	2.8	<0.001*
	RL (10^9/L)	0.21	0.36	0.00	0.00	<0.001*
Coagulation and biochemistry	PT (s)	12.8	1.3	14.2	3.3	0.255
	D-dimer (µg/L)	856	572	2900	1744	<0.001*
	CRP (mg/L)	24	25	45	50	0.242
	ASAT (U/L)	58	51	44	36	0.349
	ALAT (U/L)	77	71	46	34	0.272
	GGT (U/L)	115	209	111	97	0.221
	ALP (U/L)	90	81	108	53	0.021*
	Total bilirubin (mmol/L)	0.011	0.006	0.016	0.023	1.000
	Direct bilirubin (mmol/L)	0.004	0.003	0.009	0.017	0.322
	LDH (U/L)	277	97	383	172	0.076
	Ferritin (nmol/L)	2.4	1.5	2.3	2.7	0.346
	Procalcitonin (µg/L)	0.06	0.03	0.58	1.13	0.023*
	BUN (mmol/L)	6.9	2.7	18.7	11	0.032*
	Creatinine (mmol/L)	0.08	0.02	0.15	0.11	0.067
	GFR (mL/min/1.73 m^2)	85.11	12.4	50.92	31.69	0.001*
	Total protein (g/L)	662	71	541	65	<0.001*
	Albumin (g/L)	390	48	315	39	<0.001*

*Significant differences.

ALAT, alanine aminotransferase; ALP, alkaline phosphatase; ASAT, aspartate aminotransferase; BUN, blood urea nitrogen; CRP, C-reactive protein; GFR, glomerular filtration rate; GGT, gamma-glutamyl transferase; LDH, lactate dehydrogenase; LUC, large unstained cells; MCH, mean corpuscular haemoglobin; MCHC, mean corpuscular haemoglobin concentration; MCV, mean corpuscular volume; MN, mononuclear leucocytes; MPC, mean platelet component; MPM, mean platelet mass; MPV, mean platelet volume; NLR, neutrophil/lymphocyte ratio; PMN, polymorphonuclear leucocytes; PT, prothrombin time; RBC, red blood cells; WBC, white blood cells.

We found higher values of haemoglobin and platelet count in positive patients (136 ± 22 g/L and $268\pm148\times10^9$/L) than in negative patients (101 ± 25 g/L and $202\pm121\times10^9$/L), p=0.00007 and (p=0.09) respectively (table 2). Four patients showed platelet counts lower than 100×10^9/L in the negative group and none in the positive one.

D-dimer values were higher in the negative group (2900 ± 1744 ng/mL) than in the positive one (856 ± 572), p=0.0004. In addition, we found significantly increased values of procalcitonin in the negative group (0.58 ± 1.13 ng/mL, normal values: <0.50 ng/mL) than in the positive group (0.06 ± 0.03 ng/mL), p=0.02. Significant abnormal values of blood urea nitrogen, total protein, albumin and glomerular filtration rate were found in the negative group (see table 2).

There were no differences between both groups in the antibiotic, antiviral or hydroxychloroquine treatments. Nevertheless, 65% of negative patients received immunosuppression (dexamethasone in one patient, as it is shown in table 1), while only 28% of positive patients received it.

Comparing both groups, we found significant differences in: (1) number of days hospitalised, which was longer for negative patients (28 ± 13 days) than for the positive ones (13 ± 8) (p=0.0005); (2) period between the onset of symptoms and discharge, which was longer for negative patients (35 ± 12 days) than for positive ones (21 ± 9) (p=0.0007); (3) patients that required admission to the intensive care unit (ICU), which were 50% in the negative group and 6% in the positive group and (4) mechanical ventilation was necessary in 44% of negative patients, while in only one positive patient (6%). Finally, four negative patients (22%) died and none from the positive group.

Morphological description of atypical lymphocytes in COVID-19

In the positive group, the atypical lymphocyte count reached values between 1% and 15% in PB ($\mu=0.21\times10^9$/L). Figure 2 shows COVID-19 RL images in PB. They showed a large-medium
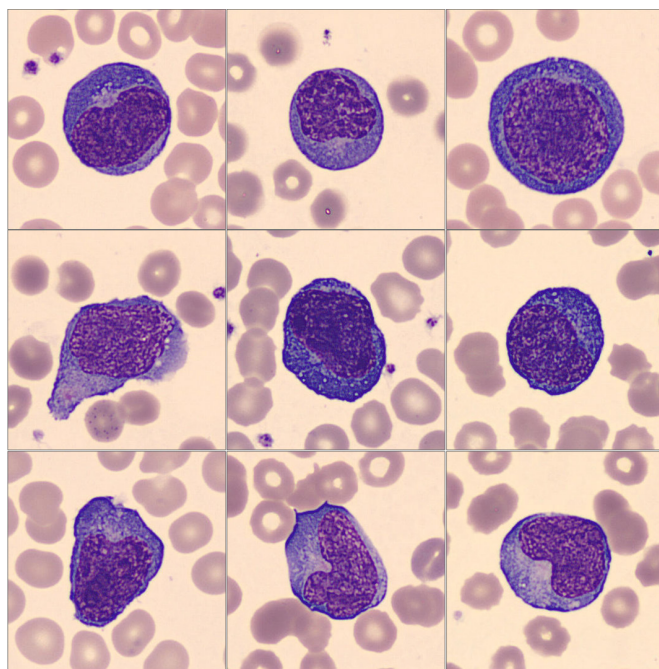
Figure 2 Images obtained in the CellaVisionDM96 showing atypical lymphocytes circulating in peripheral blood in patients with COVID-19 infection. May Grünwald-Giemsa staining (x1000).

		Predicted class		
		COVID-19 RL	**N**	**Non COVID-19 RL**
True class	**COVID-19 RL**	**38 (90.5%)**	1 (2.4%)	3 (7.1%)
	N	0 (0%)	**960 (99.4%)**	6 (0.6%)
	Non COVID-19 RL	8 (1.7%)	1 (0.2%)	**474 (98.1%)**

Figure 3 Confusion matrix of the automatic recognition results of the lymphoid cells of the testing set. Absolute image numbers and percentages (in brackets). COVID-19 RL, COVID-19 reactive lymphocytes, N, normal lymphocytes; non-COVID-19 RL, non-COVID-19 reactive lymphocytes.

size, moderate nucleus-cytoplasmic ratio, regular or kidney-shaped nucleus with a spongy chromatin pattern, usually with one nucleolus, and a distinct deep blue cytoplasm with occasional presence of small vacuoles. In some of them, nucleus showed an eccentric position.

Assessment of the automatic classification system

The 1491 images of the testing set were analysed with the classification system (see figure 1). Results are summarised in the confusion matrix shown in figure 3. Rows are the true values and columns are the predicted ones. The principal diagonal contains the true positive rates (TPRs) for each class. The overall accuracy is the percentage of images correctly classified over the 1491 images, which was 98.7%. Since this is a multiclass classification, we considered a 'one versus all' approach, where the performance metrics were calculated for each class. Focussing only on COVID-19 RL as the positive class, we calculated the sensitivity or TPR, specificity or true negative rate (TNR) and precision or positive predictive value (PPV) as follows:

$$Sensitivity\ (TPR) = \frac{TP}{TP+FN} = \frac{38}{38+4} = 0.905$$

$$Specificity\ (TNR) = \frac{TN}{TN+FP} = \frac{1441}{1441+8} = 0.994$$

$$Precision\ (PPV) = \frac{TP}{TP+FP} = \frac{38}{38+8} = 0.826$$

Immunophenotypic analysis of the large lymphocyte population

The large lymphocyte population studied by high forward scatter/side scatter contained less B cells ($\mu=4.9\%$) than NK ($\mu=18.9\%$) and T ($\mu=71.2\%$) cells (see table 3). T cells showed a CD4+ predominance (CD4/CD8 ratio >1).

Once we found that these large lymphocytes were mostly T cells, CD45RA, CCR7 and HLA−DR+ cell markers were employed to further analyse the following T cell subpopulations: naïve (CD45RA+CCR7+), central memory (CD45RA−CCR7+), effector memory (CD45RA−CCR7−), effector/TEMRA (Effector memory T cells re-expressing CD45RA) (CD45RA+CCR7−). The performed analysis revealed a significant enrichment of CD4 and CD8 effector memory (CD45RA−CCR7−) T cells in the positive group in comparison to four negative patients (p<0.05). In addition, large lymphocytes in positive patients were particularly rich in activated T cells (HLA−DR+) when compared with healthy controls (see

Table 3 Percentages of B, NK and T cells in the subset of large lymphocytes obtained throughflow cytometry (high FSC/SSC)

	Large lymphocytes Mean (range) (%)
B cells (CD19+)	4.9 (1.1–15.2)
NK cells (CD16/56+)	18.9 (8.9–31.8)
T cells (CD3+)	71.2 (40.4–88.1)
CD4/CD8 ratio	1.6 (0.1–5.3)

It is also shown the CD4/CD8 ratio to see the predominance of CD4 or CD8 in these patients.
FSC/SSC, forward scatter/side scatter.

figure 4). The remaining subpopulations did not show significant differences between both groups of patients.

DISCUSSION

The discussion section is organised in the three lines along which our study has progressed: (1) clinical and biological characteristics related to the evolution and prognosis, (2) morphological classification and (3) immunophenotype findings.

Clinical and biological characteristics related to evolution and prognosis

Clinical symptomatology in COVID-19 is variable. Indeed, patients may be asymptomatic or show a severe acute respiratory syndrome. Clinical, laboratory data and treatments have been described in recent studies,[3 5 26] in which certain haematological and biochemical parameters have been related to the severity of the disease.[27] Nevertheless, the possible role of the presence of RL in PB in the evolution and prognosis of the COVID-19 infection has not been reported previously. In this work, we observed that those patients with RL circulating in blood showed significant differences in some clinical symptoms, biological markers, hospitalisation time and recovery, with respect to those who did not present them.

Lymphopoenia is common in COVID-19[5] and it has been related to a defective immune response to the virus.[26] Nevertheless, our study revealed that patients with atypical lymphocytes had significantly higher lymphocyte numbers and, in consequence, lower NLR than patients without them. Increase in NLR values in patients with severe disease has been reported in the literature.[7 28] Therefore, our findings support a better outcome related to the presence of RL in COVID-19 patients, which might be associated with a better regulation of the immune response. Moreover, thrombocytopoenia has been considered an important indicator of severe disease in this infection.[8] It is important to mention that low platelet counts were found in our work exclusively in patients in which RL in blood were not observed.

Most of severe cases previously published showed elevated levels of infection-related biomarkers and inflammatory cytokines.[28] Our results show that indicators of disease severity, such as D-dimer and procalcitonin,[3] reached significant high values in those patients in which RL were absent in PB. High number of these patients showed critical illness and required immunosuppression drugs, as it was shown in table 1. In addition, considering the group without RL in blood, the number of days in the hospital was significantly longer, as well as the period between onset of symptoms and discharge. Moreover, the number of patients who required mechanical ventilation or died because of severe acute respiratory syndrome were also higher in this group.

The results of this study support that patients with the presence of RL in blood have a more effective immune response against the virus infection, with a better evolution and prognosis. Considering these findings, the presence of atypical lymphocytes in PB smear review might be helpful in the early screening of critical illness.

Morphologic detection and classification

In recent years, approaches have been proposed for the automatic recognition of different blood cells by combining image
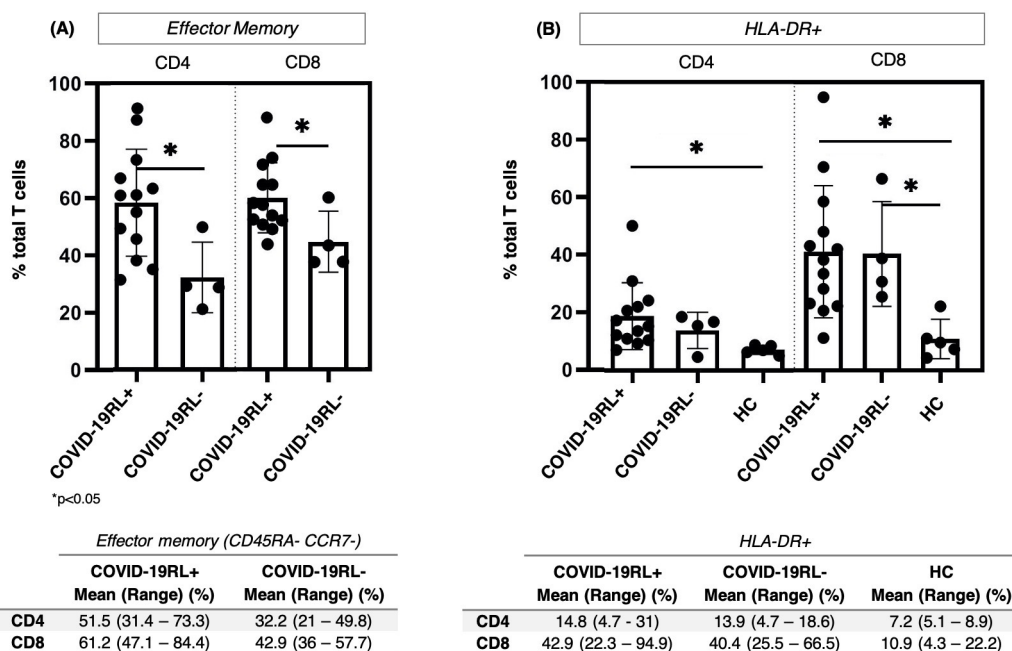

*p<0.05

	Effector memory (CD45RA- CCR7-)		
	COVID-19RL+ Mean (Range) (%)	COVID-19RL- Mean (Range) (%)	
CD4	51.5 (31.4 – 73.3)	32.2 (21 – 49.8)	
CD8	61.2 (47.1 – 84.4)	42.9 (36 – 57.7)	

	HLA-DR+		
	COVID-19RL+ Mean (Range) (%)	COVID-19RL- Mean (Range) (%)	HC Mean (Range) (%)
CD4	14.8 (4.7 – 31)	13.9 (4.7 – 18.6)	7.2 (5.1 – 8.9)
CD8	42.9 (22.3 – 94.9)	40.4 (25.5 – 66.5)	10.9 (4.3 – 22.2)

Figure 4 Immunophenotyping analysis of COVID-19 reactive lymphocytes (RL) showing the tendency of (A) effector memory and (B) activated (HLA−DR+) T cells in graphs. Mean and range values (in percentage) for each subpopulation are shown in each respective table.

analysis and artificial intelligence within a computational haematopathology framework.[29] Since morphological review requires high skills and may be prone to subjectivity, computerised methods are designed to add objectivity through quantitative features. Examples are the classification of abnormal lymphocytes and blasts associated with lymphomas and leukaemia, respectively.[21 30]

Two main difficulties have been faced in this work to develop an automatic image classifier using CNNs: (1) the similarity between COVID-19 RL and RL detected in other infections;[4 10–13] and (2) the availability of a reduced number of images of COVID-19 RL. We believe that the sequential structure of the proposed classification scheme has been successful to cope with this problem. The first CNN model was designed for a first discrimination of normal lymphocytes, while the second model was specialised in detecting COVID-19 RL, reducing the system to a couple of binary classifiers showing high accuracies. To the best of the authors' knowledge, this is the first time that this strategy is used to classify these new lymphocytes in an objective way. The system is not computationally complex and could be implemented as a rapid diagnostic tool on a simple computer alongside the pathologists. Sensitivity and specificity, considering COVID-19 RL as the positive class, reached very high values (90.5% and 99.4%, respectively).

In this work, the scarcity of COVID-19 RL images was compensated using image augmentation. Applicability and validation of data augmentation techniques in medical image classification problems have been reported,[22] in particular, in histopathological images. We believe that, although 90.5% sensitivity is satisfactory, this score may be improved when using a larger set of atypical lymphocytes from more patients.

Immunophenotype findings

In a first insight, immunophenotype results in our study show that COVID-19 RL in PB are mostly T cells enriched in activated effector memory CD4 and CD8 T cells. In a further insight, our results support that these COVID-19 RL are activated effector memory T cells (CD3+CCR7−CD45RA−TCRαβ+HLA−DR+). In addition, integrating our results with a previous work,[31] we propose that COVID-19 RL are in fact SARS-CoV-2-specific T cells.

Previous publications showed that the presence of SARS-CoV-2-specific CD4 and CD8 T cells is associated with less severe disease.[32] In accordance with this, our work has shown that patients showing COVID-19 RL have a clearly better clinic outcome. Morphological assessment of the smear is important in these patients since the visualisation of the presence of these atypical lymphocytes may be an indicator of the production of abundant virus-specific T cells.

CONCLUSIONS

In summary, this paper has three main contributions:
1. We found that RL circulating in blood in COVID-19 patients are related to a better evolution and prognosis.
2. We demonstrated that these atypical reactive lymphoid cells can be detected by morphology in the smear review, being the computerised approaches proposed herein useful to enhance a more objective recognition.
3. We found that the presence of RL in COVID-19 patients suggests an abundant production of *virus-specific T cells*, thus explaining the better outcome of patients showing these cells circulating in blood.

Take home messages

► One of the contributions of this paper is that reactive lymphocytes circulating in blood in COVID-19 patients are related to a better evolution and prognosis. This finding may have clinical relevance since it may allow a better selection of patients who will require a more intensive treatment.
► We demonstrated that these atypical reactive lymphoid cells can be detected by morphology in the smear review, being the computerised approaches proposed herein useful to enhance a more objective recognition.
► We found that the presence of reactive lymphocytes in COVID-19 patients suggests an abundant production of *virus-specific T cells*, thus explaining the better outcome of patients showing these cells circulating in blood.

ORCID iDs
Anna Merino http://orcid.org/0000-0002-1889-8889
Javier Laguna http://orcid.org/0000-0003-0777-6998
Laura Boldú http://orcid.org/0000-0002-5162-3182

REFERENCES

1 Lu H, Stratton CW, Tang Y-W. Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle. *J Med Virol* 2020;92:401–2.
2 World Health Organization (WHO). Coronavirus disease (COVID-19) pandemic [Internet], 2020. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [Accessed cited Jun 17, 2020].
3 Lippi G, Plebani M. Laboratory abnormalities in patients with COVID-2019 infection. *Clin Chem Lab Med* 2020;58:1131–4.
4 Frater JL, Zini G, d'Onofrio G, et al. COVID-19 and the clinical laboratory. *Int J Lab Hematol* 2020.
5 Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497–506.
6 Fan BE, Chong VCL, Chan SSW, et al. Hematologic parameters in patients with COVID-19 infection. *Am J Haematol* 2020;95:E131–4.
7 Qin C, Zhou L, Hu Z, et al. Dysregulation of immune response in patients with coronavirus 2019 (COVID-19) in Wuhan, China. *Clin Infect Dis* 2020;71:762–8.
8 Lippi G, Plebani M, Henry BM. Thrombocytopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: a meta-analysis. *Clin Chim Acta* 2020;506:145–8.
9 Perlman S. Another decade, another coronavirus. *N Engl J Med* 2019;382:760–2.
10 Gérard D, Henry S, Thomas B. SARS-CoV-2: a new aetiology for atypical lymphocytes. *Br J Haematol* 2020;189:845.
11 Zini G, Bellesi S, Ramundo F, et al. Morphological anomalies of circulating blood cells in COVID-19. *Am J Hematol* 2020;95:870–2.
12 Weinberg SE, Behdad A, Ji P. Atypical lymphocytes in peripheral blood of patients with COVID-19. *Br J Haematol* 2020;190:36–9.
13 Jones JR, Ireland R. Morphological changes in a case of SARS-CoV-2 infection. *Blood* 2020;135:2324.
14 Merino A, Boldú L, Ermens A. Acute myeloid leukaemia: how to combine multiple tools. *Int J Lab Hematol* 2018;40:109–19.
15 El Achi H, Khoury JD. Artificial intelligence and digital microscopy applications in diagnostic hematopathology. *Cancers* 2020;12:797.

16 Acevedo A, Alférez S, Merino A, *et al*. Recognition of peripheral blood cell images using convolutional neural networks. *Comput Methods Programs Biomed* 2019;180:105020.

17 Kumar A, Kim J, Lyndon D, *et al*. An ensemble of fine-tuned Convolutional neural networks for medical image classification. *IEEE J Biomed Health Inform* 2017;21:31–40.

18 Arriaza Gómez AJ, Fernández Palacín F, López Sánchez MA, *et al*. Estadística básica Con R Y R-Commander. *Univ Cádiz* 2008:1–160.

19 Loey M, Naman M, Zayed H. Deep transfer learning in diagnosing leukemia in blood cells. *Computers* 2020;9:29.

20 Acevedo A, Merino A, Alférez S, *et al*. A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. *Data Brief* 2020;30:105474.

21 Boldú L, Merino A, Alférez S, *et al*. Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis. *J Clin Pathol* 2019;72:755–61.

22 Mikolajczyk A, Grochowski M. Data augmentation for improving deep learning in image classification problem. In: *2018 international interdisciplinary PHD workshop (IIPhDW)*, 2018: 117.

23 He K, Zhang X, Ren S, *et al*. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016: 770–8.

24 Huang G, Liu Z, Van Der Maaten L, *et al*. Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017: 4700–8.

25 Hu J, Shen L, Albanie S, *et al*. *Squeeze-and-Excitation networks*, 2017.

26 Lippi G, Plebani M. The critical role of laboratory medicine during coronavirus disease 2019 (COVID-19) and other viral outbreaks. *Clin Chem Lab Med* 2020;58:1063–9.

27 Tjendra Y, Al Mana AF, Espejo AP, *et al*. Predicting disease severity and outcome in COVID-19 patients: a review of multiple biomarkers. *Arch Pathol&Lab Med* 2020.

28 Ding X, Yu Y, Lu B, *et al*. Dynamic profile and clinical implications of hematological parameters in hospitalized patients with coronavirus disease 2019. *Clin Chem Lab Med* 2020;58:1365–71.

29 Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies. *Lancet Haematol* 2020;7:e541–50.

30 Rodellar J, Alférez S, Acevedo A, *et al*. Image processing and machine learning in the morphological analysis of blood cells. *Int J Lab Hematol* 2018;40:46–53.

31 Weiskopf D, Schmitz KS, Raadsen MP, *et al*. Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci Immunol* 2020;5:eabd2071.

32 Moderbacher CR, Ramirez SI, Dan JM, *et al*. Antigen-Specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* 2020.