Heliyon 9 (2023) e14161

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Evaluation of risk factors for COVID-19 severity or death and their relationship to metabolic pathways☆

F.C. Marhuenda-Egea [a,*], J. Narro-Serrano [b]

[a] Departamento de Agroquímica y Bioquímica, Universidad de Alicante, Spain
[b] Departamento de Química Física, Universidad de Alicante, Spain

ARTICLE INFO

Keywords:
COVID-19
Metabolomics
Chemometric
Biochemistry
Prognosis

ABSTRACT

Background: Since the state of alarm was declared due to the COVID-19 pandemic, hospitals have been the main ones in charge of registering the therapeutic follow-up of affected people. The analysis of these data has allowed those different biochemical markers have been identified as predictors of the severity of the disease, but most of the published studies tend to be eminently descriptive and do not propose a biochemical hypothesis to explain the alteration of the results they are showing. Our objective is to recognize the main metabolic processes that are occurring in COVID-19 patients, as well as the identification of clinical parameters that are decisive to predict the severity of the disease.

Methods: A multivariate analysis was carried out from the clinical parameters collected in the database of the HM hospitals in Madrid, to determine the most relevant variables to predict the severity of the disease. Chemometric methods allow these variables to be obtained by applying a classification strategy with PLS-LDA.

Findings and interpretation: The variables that most contribute to separation are age in men and, in both sexes, the concentration of lactate dehydrogenase, urea and C-reactive protein.

Oxygen deficiency in the tissues, due to the loss of functionality of the lungs, could be affecting the muscle tissue with special severity. Inflammation and tissue damage is related to increased LDH and CRP. The loss of muscle mass and the increase in the concentration of urea and LDH is explained by the adaptation of muscle metabolism to this oxygen deficiency.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profits sectors.

1. Introduction

In December 2019, severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) was identified as the causal agent in a series of pneumonia cases of unknown origin in Wuhan (Hubei Province, China). It was named the 2019 coronavirus disease, or COVID-19 [1]. In March 2020, the WHO declared the status of a global pandemic for COVID-19. Since then, work has been done in a frenzy to develop vaccines and/or treatments to slow down its expansion. In parallel, a large amount of clinical data has been collected that has allowed, and will continue to allow for several years, to better understand the development of this disease and the effect it

☆ In memoriam to Joaquín Moreno Casco (1960–2020), Professor of Microbiology. Almeria University (Spain).

* Corresponding author.
E-mail address: frutos@ua.es (F.C. Marhuenda-Egea).

https://doi.org/10.1016/j.heliyon.2023.e14161
Received 28 January 2022; Received in revised form 20 February 2023; Accepted 23 February 2023
Available online 28 February 2023

causes on the body. Much remains to be explained of the side effects caused by COVID-19. Among them, one of the most evident is the great loss of muscle mass suffered by hospitalized patients and the speed with which it occurs [2–6].

Hospitals have been the main ones in charge of recording the clinical data associated with SARS-CoV-2 infection by systematically keeping the therapeutic follow-up of patients with COVID-19. The analysis of these data has allowed those different biochemical markers to have been identified as predictors of severity or prognosis that can be useful for medical decision making [7]. However, most of the studies published to date tend to be eminently descriptive and do not propose a hypothesis to explain, at the biochemical level, the alteration of the results they are showing. On the other hand, many of them have a small study population or have been carried out in Chinese populations, so the extrapolation of their conclusions to other groups can be risky [7].

Our main objective is to identify the alterations in the clinical parameters that allow predicting the severity of the disease in the Spanish population, identifying those that provide little information and those that are decisive, by analysing the data collected at the HM Hospitals in Madrid [8]. In addition, a biochemical explanation is proposed that relates the clinical alterations identified with the evident loss of muscle mass that is manifested in these patients.

2. Methods

2.1. Covid data saves lives database and study design

Data used belongs to the database of the HM Hospitals in Madrid [8]. This database has the clinical information of COVID-19 patients collected during the first months of the pandemic (2339 patients collected since 03/01/2020). The data is sorted by patient and by analysis date. The analytical data is collected according to the evolution of the patients, so that in the same day several analyses may have been carried out. Different clinical variables are evaluated, which are not always repeated in successive extractions.

2.2. Chemometric analysis

MATLAB version 2021b from MathWorks was used for the calculations. Robust principal components analysis (Robust PCA) was carried out using the *LIBRA toolbox* [9] and partial least square with linear discriminant analysis (PLS-LDA) was carried out using the *plslda toolbox* [10]. In a supervised method, such as PLS-LDA, the approach was to select several clinical variables for to make a mathematical model. This model can be used for the prediction of COVID-19 evolution.

We have used the *plslda* statistical toolbox [10] because it allows a very robust classification. This classification can be evaluated using the different statistical parameters provided by the classification algorithm (partial least squares linear discriminant analysis, PLSLDA). Within the statistical package used, we also have algorithms that allow select the most informative variables to build the classification model. We have used subwindow permutation analysis (SPA) for this purpose of variable selection. In our study, the classification model looks for the clinical variables that best help us to predict the severity of the disease.

3. Results

The multivariate analysis of the data aims to determine the variables that will make it possible to assess the severity of the infection and its subsequent evolution. In addition to helping us identify the relevant variables to predict the severity of the disease, determining which clinical variables are more important may indicate which metabolic pathways are key in the development of the disease. Chemometric methods allow these variables to be obtained when applying a classification strategy with PLS-LDA [10].

In a first step, we select the patients from whom we are going to analyze their clinical data. The main problem with the analysis of these data is that, since they are data collected from medical practice, they do not follow a standardized procedure for collecting information, but rather the analyzes required by each patient were carried out. Therefore, not all variables have been collected in all patients, nor have the same number of extractions been performed. To solve this problem, we have selected the patients who shared the highest number of variables analyzed. To do this, we have looked first at the patients who have died because they are the fewest in

Table 1
Demographics of COVID-19 patients grouped according to in-hospital mortality.

	Total population n = 2339	**Survivors n = 1958 (83.7%)**	Non-survivors n = 381 (16.3%)
Age [Median (IQR)]	69 (23)	66 (22)	82 (14)
Distribution			
<30 years [n (%)]	39 (1.7)	39 (2)	0
30–39 years [n (%)]	87 (3.7)	87 (4.4)	0
40–49 years [n (%)]	218 (9.3)	217 (11.1)	1 (0.3)
50–59 years [n (%)]	339 (14.5)	331 (16.9)	8 (2.1)
60–69 years [n (%)]	490 (20.9)	446 (22.8)	44 (11.5)
70–79 years [n (%)]	536 (22.9)	435 (22.2)	101 (26.5)
≥80 years [n (%)]	630 (26.9)	403 (20.6)	227 (59.6)
Sex, distribution			
Male [n (%)]	1388 (59.3)	1137 (58.1)	251 (65.9)
Female [n (%)]	951 (40.7)	821 (41.9)	130 (34.1)

number (Table 1). Considering measured variables in non-surviving patients, we have selected surviving patients in whom the same variables were also measured.

Considering that biochemical values are usually different between men and women and that we do not have the same number of extractions in all patients, we decided to separate the data into four groups: men first extraction, which includes the data from all the tests that were carried out on the first day of admission to the hospital; men final extraction, which includes the data of all the analyzes performed on the last day in the hospital, either due to recovery or death; women initial extraction, the same as the first group but with female patients; and women final extraction, the same as the second group but with female patients. Once the groups have been selected, we screen to detect the patients who, based on their analytical values, would be outliers. We eliminate those patients using RobustPCA [9], which is an unsupervised method. The patients included in the statistical analysis are collected in Tables 2 and 3.

Subsequently, we use the algorithm subwindow permutation analysis (SPA) coupled with partial least squares linear discriminant analysis (PLSLDA), for statistical assessment of variable importance to predict the severity of the patient and the outcome of the disease [11]. The data used in the analysis have been processed as shown in the previous section and are detailed in Table 2. With the processed data (Table 2), we make a classification based on their analytical data using LDA-PLS [12]. The classification criterion was to overcome the disease or not. However, the data suggested that not all variables contained information useful for achieving an optimal classification model (Fig. 1(A–D)). The presence of informative variables is shown if there are changes in the peak prediction error between the unpermuted variable and the permuted variable, which is observed in Fig. 1(A–D). To identify the informative variables, which will allow us to obtain more accurate disease severity prediction models (Fig. 3(A–D)), we use the values obtained in Fig. 2(A and B) [13],. The higher COSS values tell us which variables could be the most informative and reflect the most severe clinical alterations. That is, the variables that have the most weight in the classification would be the ones shown with the highest values in Fig. 2(A and B),

Table 2
Demographics and clinical data in the first clinical analysis of COVID-19 patients included in the study. Patient characteristics were expressed as the median and interquartile range (IQR).

	First clinical analysis			
Sex [n (%)]	Female 135 (36.3%)		Male 237 (63.7%)	
Outcome [n (%)]	Survivors	Non-Survivors	Survivors	Non-Survivors
	82 (60.7%)	53 (39.3%)	114 (48.1%)	123 (51.9%)
Age, years	68 (24.75)	85 (15)	62 (19)	81 (13.5)
RDW, %	12.8 (2.4)	14.3 (2.475)	12.95 (1.68)	13.5 (2.1)
BAS, 10^9 cells/L	0.02 (0.02)	0.02 (0.02)	0.01 (0.02)	0.01 (0.02)
BAS%, %	0.3 (0.2)	0.2 0.2	0.2 (0.3)	0.2 (0.2)
MCHC, mg/dL	33.5 (1.6)	32.7 (1.75)	34.05 (1.5)	33.6 (2.1)
CRE, mg/dL	0.72 (0.20)	1.03 (0.86)	0.95 (0.33)	1.16 (0.73)
DD, ng/mL	603 (804)	1186 (1577)	717.5 (929)	1312 (1613.5)
EOS, 10^9 cells/L	0.01 (0.06)	0 (0.02)	0.01 (0.04)	0 (0.01)
EOS%, %	0.1 (0.8)	0 (0.15)	0.1 (0.58)	0 (0.2)
GLU, mg/dL	110.55 (26.95)	128.3 (66.7)	114.6 (32.5)	128.6 (47)
GOT, U/L	30.8 (18.3)	34 (21.3)	33.3 (24.4)	41 (31.9)
GPT, IU/L	27 (20.9)	20 (16.1)	31.15 (29.03)	25.5 (26.65)
MCH, pg/cell	29.3 (1.9)	29.45 (2.6)	30 (2.38)	30.2 (2.4)
HCT, %	39.8 (6.2)	37.45 (7.15)	42.5 (5.6)	41.8 (7.75)
RBC, million cells/μL	4.59 (0.69)	4.22 (1.09)	4.89 (0.75)	4.64 (0.77)
HGB, g/dL	13.33 (2.15)	11.9 (2.85)	14.5 (2)	14 (2.7)
INR	1.17 (0.15)	1.22 (0.38)	1.21 (0.19)	1.32 (0.26)
K, mmol/L	4.18 (0.66)	4.4 (1.4)	4.19 (0.63)	4.08 (0.67)
LDH, U/L	508 (226.6)	691.2 (489.3)	502.1 (248.4)	673.7 (299.5)
LEU, 10^9 cells/L	6.50 (3.31)	8.42 (7.52)	6.83 (3.65)	8.61 (5.86)
LYM, 10^9 cells/L	1.18 (0.93)	0.89 (0.61)	0.99 (0.64)	0.68 (0.59)
LYM%, %	19.2 (17.1)	9.3 (10.1)	15.08 (12.25)	8.1 (8.55)
MON, 10^9 cells/L	0.49 (0.35)	0.46 (0.45)	0.48 (0.39)	0.43 (0.38)
MON%, %	7.25 (5.11)	5 (4.2)	7.33 (5.1)	5.3 (4.4)
NA, mmol/L	137.4 (4.1)	137.2 (7)	136.9 (3.9)	137.9 (7.3)
NEU, 10^9 cells/L	4.71 (2.90)	7.28 (6.97)	4.92 (3.49)	7.18 (5.37)
NEU%, %	70.8 (19.1)	84.8 (14.6)	75.55 (15.35)	85.2 (10.35)
CRP, mg/L	46.20 (71.97)	132.2 (144.0)	75.88 (112.79)	143.38 (147.05)
PLAT, 10^9 cells/L	251 (114.3)	210 (111)	194 (101.5)	195.5 (100.5)
U, mg/dL	30.05 (18.5)	62.5 (48.7)	36 (18.7)	58 (43.5)
MCV, fL	87.2 (5.7)	90.2 (7.5)	87.23 (5.85)	90.3 (6.33)
MPV, fL	9.9 (1.2)	10.5 (1.4)	10.4 (1.4)	10.4 (1.25)

RDW: Red cell Distribution Width (Anisocytosis Coefficient); BAS: Basophils; BAS%: Basophils%; MCHC: Mean Corpuscular Hemoglobin Concentration; CRE: Creatinine; DD: D-dimer; EOS: Eosinophils; EOS%: Eosinophils%; GLU: Glucose; GOT: Glutamic oxaloacetic transaminase (Aspartate transaminase); GPT: Glutamic-pyruvic transaminase (Alanine transaminase); MCH: Mean Corpuscular Hemoglobin; HCT: Hematocrit; RBC: Red Blood Cells; HGB: Hemoglobin; INR: International Normalized Ratio; K: Potassium; LDH: Lactate dehydrogenase; LEU: Leukocytes; LYM: Lymphocytes; LYM%: Lymphocytes%; MON: Monocytes; MON%: Monocytes%; NA: Sodium; NEU: Neutrophils; NEU%: Neutrophils%; CRP: C-reactive protein; PLAT: Platelet count; U: Urea; MCV: Mean Corpuscular Volume; MPV: Mean Platelet Volume. Clinical data are expressed as [Median (IQR)].

Table 3
Demographics and clinical data in the last clinical analysis of COVID-19 patients included in the study. Patient characteristics were expressed as the median [interquartile range (IQR)].

	Last clinical analysis			
Sex [n (%)]	Female		Male	
	139 (36.4%)		243 (63.3%)	
Outcome [n (%)]	Survivors	Non-Survivors	Survivors	Non-Survivors
	86 (61.9%)	53 (38.1%)	124 (51.0%)	119 (49.0%)
Age, years	66 (17)	85 (21.5)	60.5 (12)	83 (17.5)
RDW, %	12.5 (2.9)	14.4 (2.1)	12.5 (1.8)	13.9 (3.15)
BAS, 10^9 cells/L	0.02 (0.03)	0.02 (0.03)	0.02 (0.03)	0.02 (0.02)
BAS%, %	0.4 (0.4)	0.2 (0.2)	0.3 (0.5)	0.2 (0.1)
MCHC, mg/dL	33.2 (1.6)	32.5 (2.1)	33.7 (1.5)	32.7 (1.85)
CRE, mg/dL	0.65 (0.19)	1.13 (1.16)	0.845 (0.25)	1.18 (0.94)
DD, ng/mL	736.5 (1108)	1918 (3463)	730.5 (828.5)	2548 (6147.5)
EOS, 10^9 cells/L	0.07 (0.13)	0 (0.01)	0.1 (0.16)	0 (0.045)
EOS%, %	1.2 (2.45)	0 (0.1)	1.6 (2.65)	0 (0.4)
MCH, pg/cell	29.5 (2.1)	29.3 (3.1)	29.3 (2.13)	29.9 (2.55)
HCT, %	37.6 (5.5)	38.5 (7.7)	40.9 (6)	39.5 (8.2)
RBC, million cells/μL	4.3 (0.63)	4.07 (0.99)	4.66 (0.68)	4.45 (0.89)
HGB, g/dL	12.3 (1.8)	12.3 (2.4)	13.8 (2.3)	13.2 (2.8)
LDH, U/L	384 (143.7)	891 (703.9)	403.3 (184.3)	905.2 (631.75)
LEU, 10^9 cells/L	6.01 (3.07)	12.19 (12.36)	6.78 (3.86)	11.53 (6.94)
LYM, 10^9 cells/L	1.46 (0.93)	0.75 (0.74)	1.52 (0.89)	0.68 (0.56)
LYM%, %	25.1 (13.7)	7.8 (8.95)	23.65 (16.55)	5.8 (6.2)
MON, 10^9 cells/L	0.61 (0.31)	0.46 (0.46)	0.655 (0.355)	0.46 (0.44)
MON%, %	10 (4.9)	3.9 (3.5)	9.3 (4.7)	4.1 (4.4)
NA, mmol/L	138.8 (3.65)	140.6 (8.5)	138 (3.6)	141.1 (9.65)
NEU, 10^9 cells/L	3.68 (2.10)	9.88 (10.93)	4.34 (3.26)	10.25 (5.92)
NEU%, %	61.6 (16.45)	87.2 (11)	64.1 (20.6)	89.3 (10.75)
CRP, mg/L	9.12 (19.07)	101.31 (118.81)	11.37 (29.57)	102.8 (128.51)
PLAT, 10^9 cells/L	321 (154.5)	220 (128)	306.5 (130.5)	200 (140.5)
U, mg/dL	32.2 (19)	73 8 (44.3)	37.8 (21.55)	83 (63.8)
MCV, fL	88.7 (4.55)	90.3 (9.6)	87.65 (4.73)	91 (7.5)
MPV, fL	9.7 (1.15)	10.8 (1.5)	10.1 (1.23)	10.7 (1.6

RDW: Red cell Distribution Width (Anisocytosis Coefficient); BAS: Basophils; BAS%: Basophils%; MCHC: Mean Corpuscular Hemoglobin Concentration; CRE: Creatinine; DD: D-dimer; EOS: Eosinophils; EOS%: Eosinophils%; GLU: Glucose; GOT: Glutamic oxaloacetic transaminase (Aspartate transaminase); GPT: Glutamic-pyruvic transaminase (Alanine transaminase); MCH: Mean Corpuscular Hemoglobin; HCT: Hematocrit; RBC: Red Blood Cells; HGB: Hemoglobin; INR: International Normalized Ratio; K: Potassium; LDH: Lactate dehydrogenase; LEU: Leukocytes; LYM: Lymphocytes; LYM%: Lymphocytes%; MON: Monocytes; MON%: Monocytes%; NA: Sodium; NEU: Neutrophils; NEU%: Neutrophils%; CRP: C-reactive protein; PLAT: Platelet count; U: Urea; MCV: Mean Corpuscular Volume; MPV: Mean Platelet Volume. Clinical data are expressed as [Median (IQR)].

depending on the COSS value (COnditional Synergetic Score) (Fig. 2(A and B)).

We can see that the most important variable in men is age, both for the first and for the last analyses. In the model, that allows us to predict the outcome of the disease. In women, other variables are more important, such as LDH, urea, or CRP. With these variables we construct the LDA-PLS models that allow us to predict the development of the disease and its severity (Fig. 3(A–D)). Once the most important variables in the classification of patients have been selected, the statistic clearly improves, confirming the selection (Table 4 [13],). The variables that most contribute to separation are age in men and, in both sexes, the concentration of lactate dehydrogenase (LDH), urea and C-reactive protein (CRP).

4. Discussion

The fundamental problem of COVID-19 patients is bilateral pneumonia [1,14–18]. This pneumonia causes an oxygen deficit in the tissues. This oxygen deficit would be affecting muscle tissue with special severity. The increase in the concentration of urea and LDH is explained by the adaptation of muscle metabolism to this oxygen deficiency.

Under optimal oxygenation conditions, glucose is oxidized in muscle via glycolysis to pyruvate, generating ATP and NADH. Pyruvate oxidizes to Acetyl-CoA. Acetyl-CoA is oxidized to CO_2 in the Krebs cycle producing GTP, NADH and $FDAH_2$. NADH and $FDAH_2$ are the electron transporters. Electrons enter the mitochondrial electron transport chain to deplete O_2 and produce massive amounts of ATP through oxidative phosphorylation [19].

For this to happen, and for enough ATP to be generated for cellular metabolism, O_2 is needed as the final electron acceptor. As there is an O_2 deficit caused by pneumonia, the muscle tissue adapts to this stress situation by breaking down muscle proteins to generate energy (Fig. 4). Proteins are made up of amino acids, which have a carbon skeleton that will be oxidized in the Krebs cycle, but they also have an amino group that must be eliminated, as it is toxic. To remove ammonium, the muscle cell must transport it to the liver, where it is transformed into urea, which is excreted through the urine.

Because ammonia cannot pass directly into the blood, the body uses alanine to transport it to the liver [19]. Therefore, a very

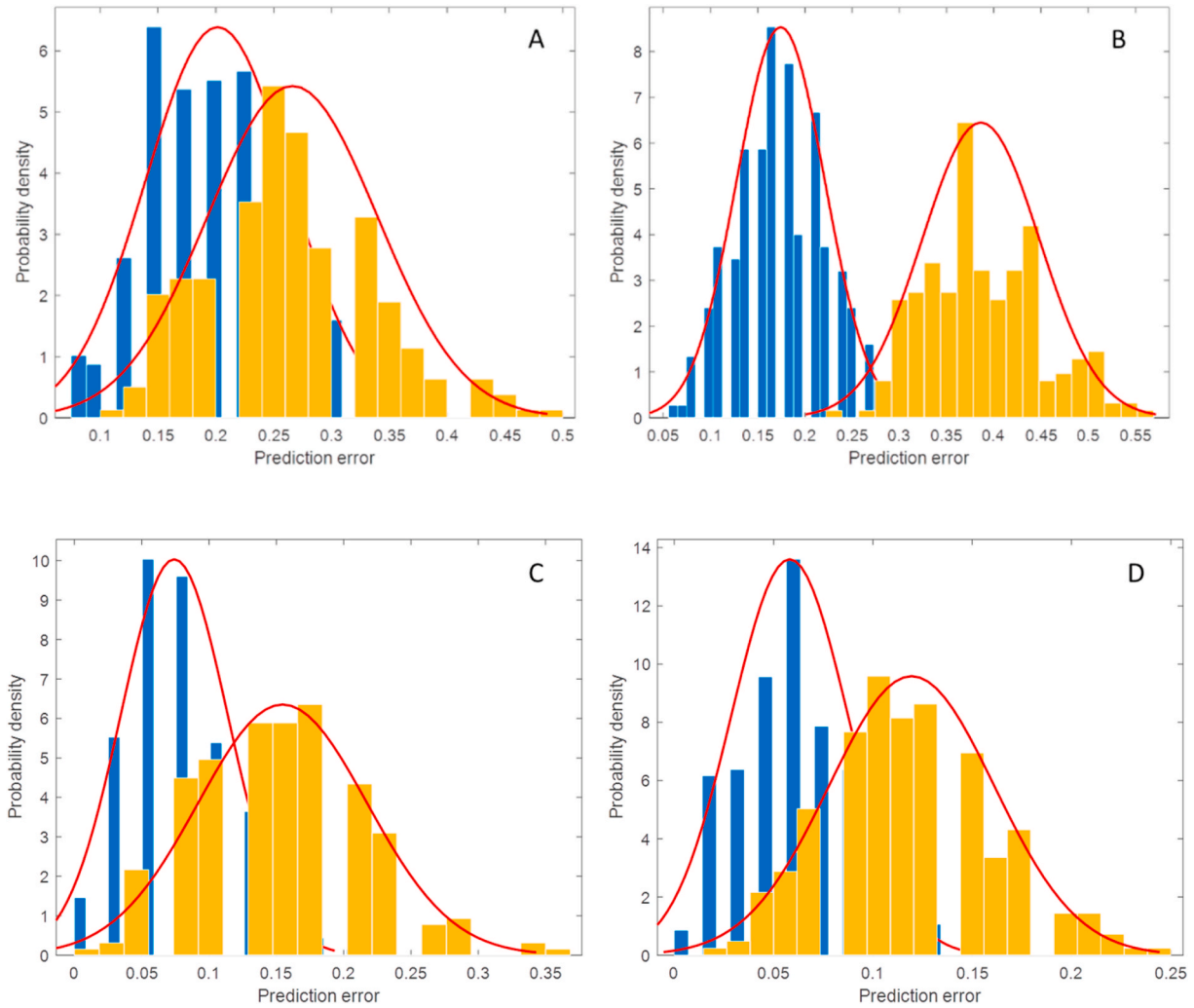

Fig. 1. A) Male initial, B) female initial, C) male final and D) female final. The plots show the distributions of normal prediction errors (blue bars) and permuted prediction errors (orange bars). The peak of the curve generated by the blue bars corresponds to the prediction error distribution of a non-permuted informative variable, while the peak generated by the orange bars corresponds to the same permuted informative variable. The prediction error was obtained with the subwindow permutation analysis (SPA) algorithm of the *plslda* toolbox (13). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

important step in this process is the generation of alanine. Alanine is generated by transamination: the different amino acids from protein degradation give up their amino group to pyruvate, through transaminases. Pyruvate is produced in the last step of glycolysis. Alanine passes into the blood and reaches the liver. In the liver, by transamination, the amino group is transferred to the oxaloacetate to produce glutamate and initiate the urea cycle. Urea passes into the blood to be excreted in the urine. Pyruvate is used for the synthesis of glucose through gluconeogenesis. Glucose also passes into the blood and reaches the muscle and other extrahepatic tissues, to be oxidized in glycolysis and produce pyruvate, in what is known as the glucose-alanine cycle, or Cahill cycle [19].

This has very important biochemical consequences. On the one hand, it allows the muscle to obtain energy from muscle proteins continuously, oxidizing amino acids. On the other hand, it leaves the entire energy load of glucose production (gluconeogenesis) and urea elimination (urea cycle) to the liver [19]. That is, the muscle does not expend unnecessary energy in any process other than supplying ATP to its own metabolic processes.

In this situation of limitation in the supply of O_2 by pneumonia, all the processes that need energy, gluconeogenesis and the urea cycle, occur in the liver. The liver obtains its energy from the oxidation of lipids, through β-oxidation, generating in turn ketone bodies. An increase in ketone bodies and plasma lipoproteins has been observed in COVID-19 patients, indicating a greater mobilization of this energy source from adipose tissue [20,21].

The enormous loss of muscle mass [4–6,22,23] would be related to the increase in urea production and its importance as a key variable to assess the severity of COVID-19. It must also be increasing the oxidation of NADH to NAD^+, by LDH, which, under anaerobic conditions, is reducing pyruvate to lactate. This effect worsens with the development of the disease, as the deficiency in O_2 becomes

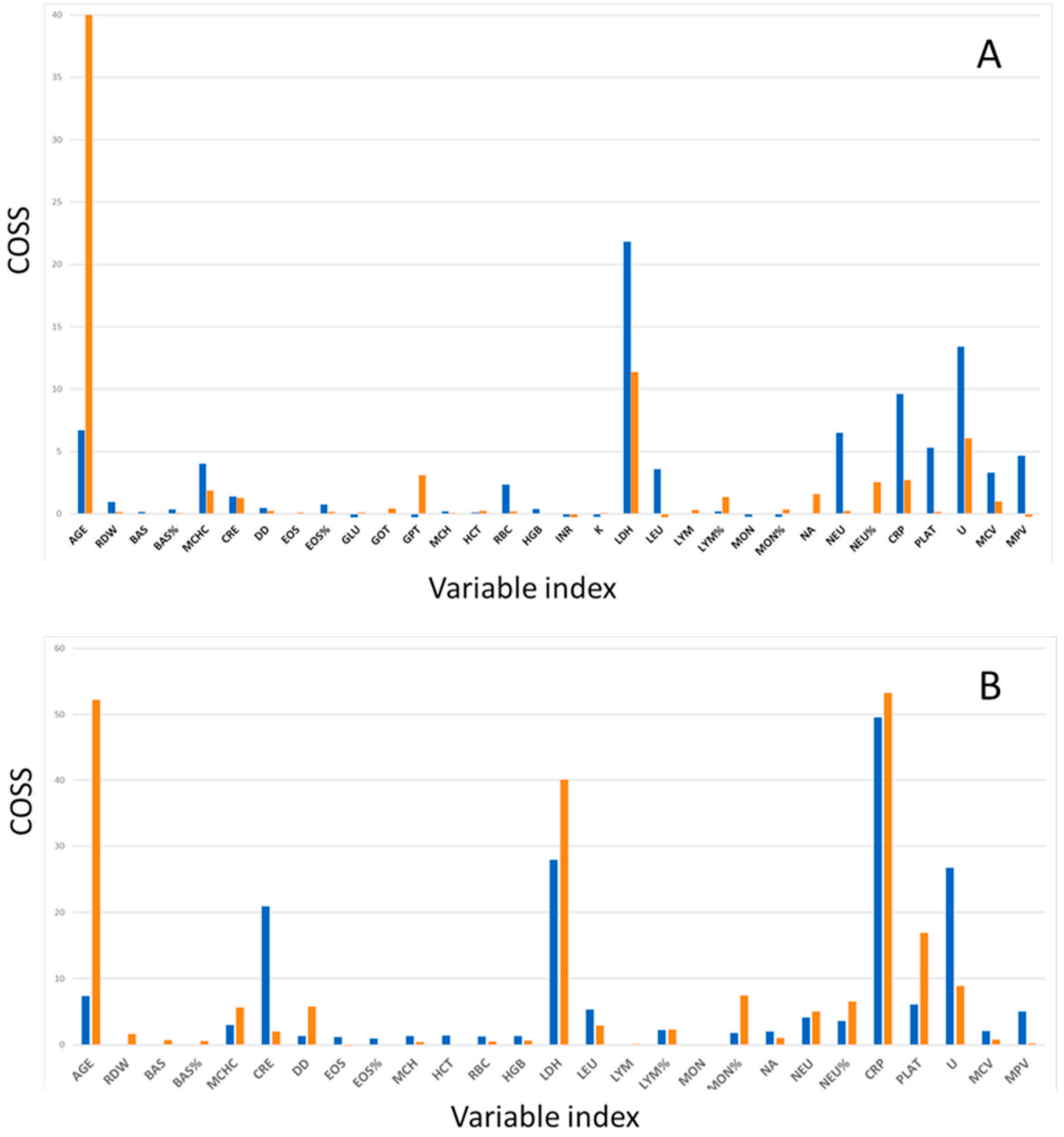

Fig. 2. COSS (Conditional Synergetic Score) values are the output of subwindow permutation analysis (SPA) algorithm [13], and it was used to rank variables. The top figure (A) was obtained with the first analytical data and the bottom figure (B) was obtained the last analytical data. The column blue was for females and the column orange for males. We selected the most informative variables based on the highest COSS values. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

more pronounced. In the later stage of the disease, these factors become even more important, since muscle protein reserves have already been depleted and, probably, O_2 deficiency has worsened, as shown by the importance of an increase in the concentration of LDH, both men and women.

Another key variable is the concentration of C-reactive protein (CRP). CRP is considered a systemic inflammatory marker and the elevation of its plasma levels is associated with the worsening of different pathologies such as cardiovascular problems [24,25], pneumonia [26], different types of cancer [27,28], or COVID-19 [29,30]. CRP plays a fundamental role in the body's response to infectious and inflammatory processes. Its main function is to promote complement activation in order to facilitate neutralization and phagocytosis by macrophages, which have specific receptors for CRP. The highest levels are found in the first 24–48 h after infection

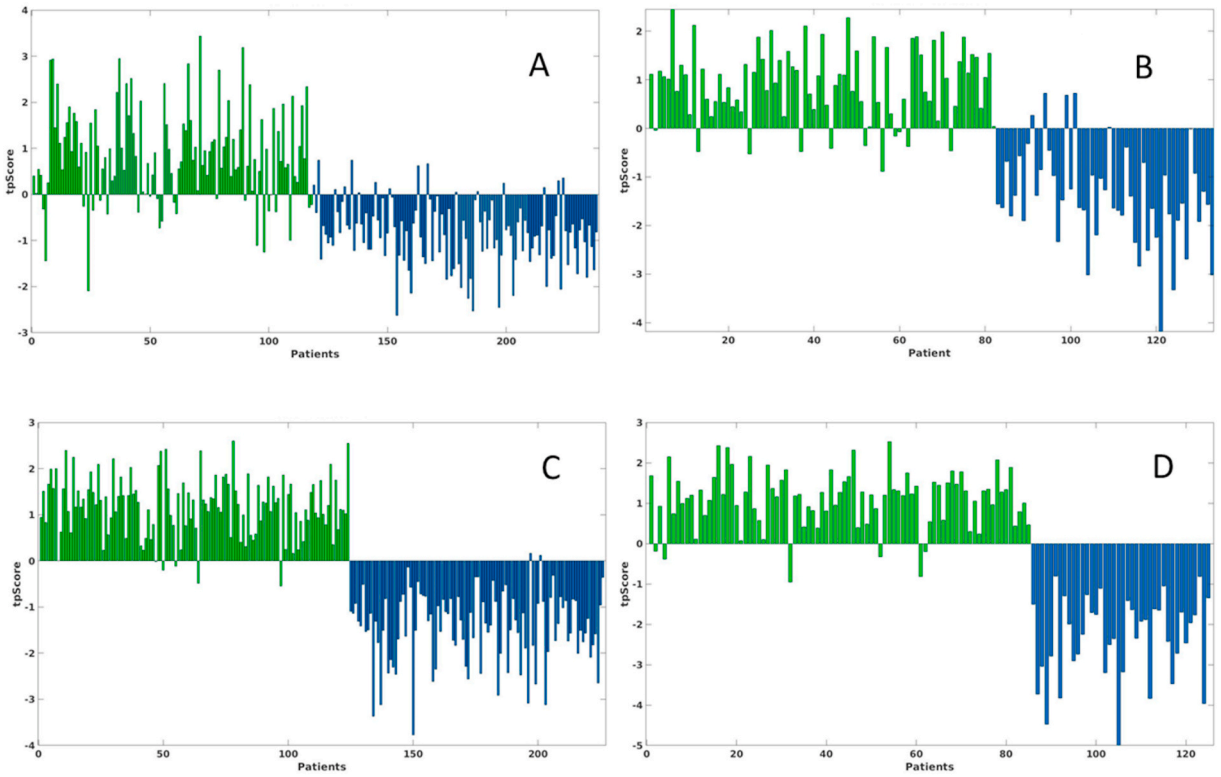

Fig. 3. A) Male initial, B) female initial, C) male final and D) female final. Scores on the target-projected (tpScore) component for survivor patients (blue) and for non-survivor patients (green). PLS-LDA models with 3 components transformed to produce the best of possible single-component predictive model. Positive values of tpScores for survivor patients indicate misclassification and negative values of tpScores for non-survivor patients also indicate misclassification of the PLS-LDA model. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Prediction ability of LDA-PLS before and after variable selection using COSS (Conditional Synergetic Score) [13].

Groups	Variables	R2Y	R2X	Error (%)	Sensitivity (%)	Specificity (%)	AUC (%)
Females (initial data)	All	51.35	15.27	12.03	87.80	88.24	94.24
	12[a]	54.12	27.28	12.03	89.02	86.27	93.14
Males (initial data)	All	39.41	15.95	10.97	91.23	86.99	93.67
	7[b]	40.34	39.69	16.81	79.66	86.67	91.19
Females (final data)	All	74.36	24.30	7.20	89.41	100.00	97.44
	11[c]	76.13	38.29	4.80	92.94	100.00	97.41
Males (final data)	All	70.60	25.01	2.65	97.58	97.06	98.83
	8[d]	72.84	35.72	3.54	95.16	98.04	98.78

[a] Age; MCHC: Mean Corpuscular Hemoglobin Concentration; CRE: Creatinine; RBC: Red Blood Cells; HGB: Hemoglobin; LDH: Lactate dehydrogenase; LEU: Leukocytes; NEU: Neutrophils; CRP: C-reactive protein; PLAT: Platelet count; U: Urea; MCV: Mean Corpuscular Volume.
[b] Age; CRE: Creatinine; LDH: Lactate dehydrogenase; LYM%: Lymphocytes%; NEU: Neutrophils; CRP: C-reactive protein; U: Urea.
[c] Age; MCHC: Mean Corpuscular Hemoglobin Concentration; CRE: Creatinine; LDH: Lactate dehydrogenase; LEU: Leukocytes; CRP: C-reactive protein; NA: Sodium; NEU: Neutrophils; NEU%: Neutrophils%; PLAT: Platelet count; U: Urea.
[d] Age; MCHC: Mean Corpuscular Hemoglobin Concentration; CRE: Creatinine; DD: D-dimer; LDH: Lactate dehydrogenase; CRP: C-reactive protein; PLAT: Platelet count; U: Urea.

and it remains high until the infectious/inflammatory process is resolved [31].

One of the main alterations that were associated with the severity of the disease in COVID-19 patients is the dysregulation of inflammatory cytokines and chemokines, commonly called cytokine storm or inflammatory storm [32]. Considering the relationship of CRP with the activation of the immune response in response to certain inflammatory mediators, such as cytokines, it seems reasonable that their levels correlate positively with severity or, in our case, with a poor prognosis.

Other authors have observed a positive correlation between CRP and LDH levels with PaO_2/FiO_2 values, which has been used classically to define acute lung injury (ALI) and acute respiratory distress syndrome (ARDS) [33]. As noted in the article, this may be

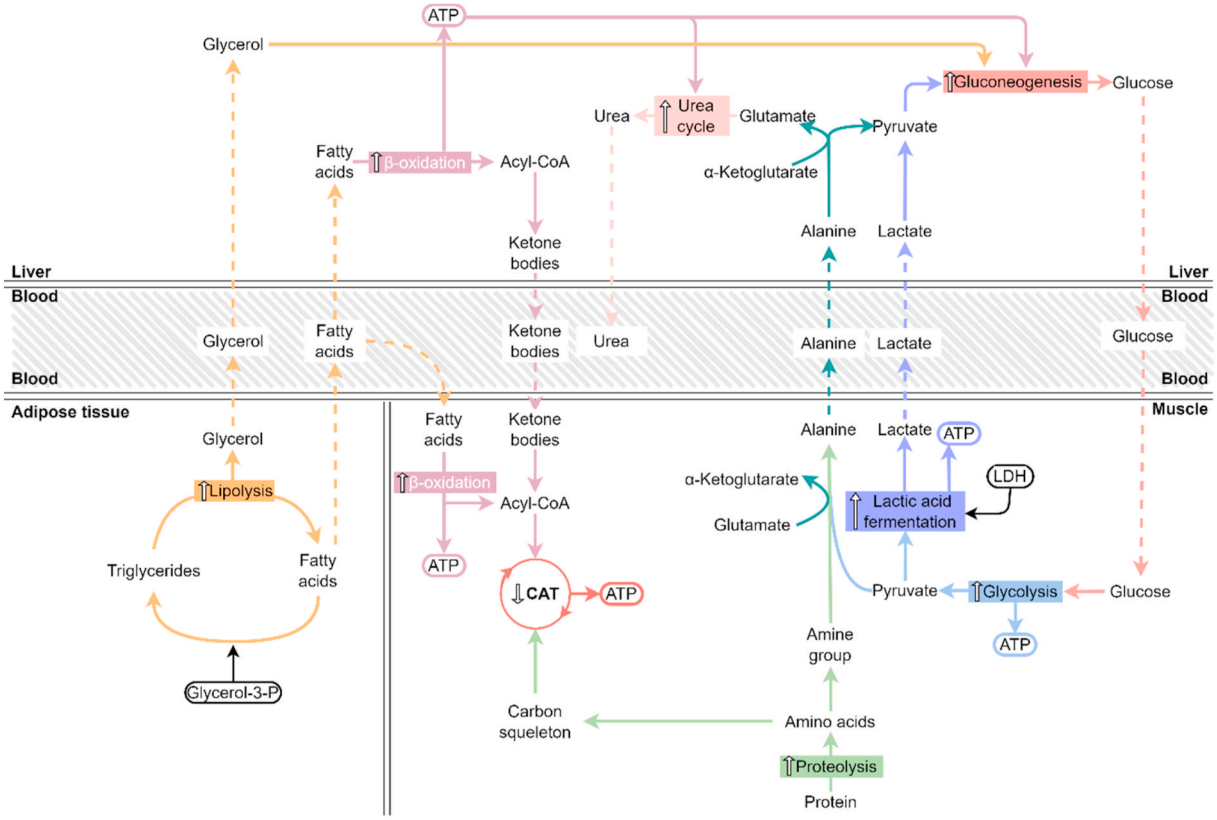

Fig. 4. Changes in metabolic pathways that correspond to the alterations found in patients with a worse prognosis. White upward arrows indicate increased metabolic pathway flux. The black downward arrow indicates that the CAT would be slowing down.

suggesting a possible relationship between tissue damage and the level of infection. The inflammatory response to the virus induces apoptosis of the epithelial and endothelial cells of the lung. Inflammatory cytokine-induced infiltration of inflammatory cells causes apoptosis of the airways and alveolar epithelial cells [34]. Apoptosis of endothelial cells and epithelial cells damages lung microvascular barriers and causes vascular leakage and alveolar edema, leading to ARDS.

LDH is an enzyme related to energy metabolism that is present in most tissues. Its levels are related to tissue damage, both acute and chronic [35]. This relationship is because, when tissue deterioration occurs, the enzymes that are generally contained within the cells are discharged into the bloodstream, therefore, the greater the cellular damage, the greater the concentration of LDH in the blood. Another factor to consider is that the amount of LDH that migrates into the bloodstream due to tissue damage will also be related to the amount of enzyme present in the cell before lysis occurs. Therefore, the increase in LDH concentration observed in COVID-19 patients can, on the one hand, be an indicator of tissue damage and, on the other hand, indicate an increase in the synthesis of the enzyme LDH as response to metabolic adaptation discussed in this article.

We consider that the main strength of our results is based on the enormous amount of data analyzed, as well as the large number of patients considered in the study. This is thanks to the initiative of HM Hospitals, which has made its databases available for this type of study. However, this is also where the main difficulty lies, as we have not been able to carry out any type of planning to select the most appropriate type of clinical tests to assess the risks and consequences of the disease. In conclusion, we have identified a few routine laboratory markers as predictors of disease severity and risk of death. These predictors of the risk of death have a high predictive capacity even with the values of the first clinical analysis at the time of admission to the hospital. The factors with the highest risk of severity or death would be those indicated in the last paragraph of the Results section: age for men, while for both sexes, are the concentration of lactate dehydrogenase (LDH), urea and C-reactive protein (CRP). We have also tried to make a biochemical interpretation of the metabolic changes caused by infection.

5. Research in context

5.1. Evidence before this study

During the current pandemic, one of the problems has been not being able to predict the evolution of the disease. Having a clear vision of the development of the disease in each person is essential to decide on the most effective treatments. If considering some key

values we can predict the evolution of the disease, it will be possible to intervene to try to avoid a fatal outcome.

5.2. Added value of this study

Here, we show that certain clinical variables can discriminate between survival and not survival COVID-19 patients. We can identify, with a high probability and according to the values of certain variables of the first clinical analysis, patients who may not survive. That is, by entering the hospital we can predict the survival rate of each patient.

5.3. Implications of all the available evidence

All the evidence suggests that we could intervene in each patient to decide how to treat it to improve the survival rate.

Author contribution statement

Frutos Marhuenda-Egea: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.
Jennifer Narro-Serrano: Analyzed and interpreted the data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data associated with this study has been deposited at HMHospitales. Covid Data Save Lives. Available online: https://www.hmhospitales.com/coronavirus/covid-data-save-lives/english-version;2020.

Declaration of interest's statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank HM Hospitals for allowing us access and availability of their Data Set.

References

[1] N. Zhu, et al., A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. 382 (8) (2020) 727–733.
[2] W. Ahmed, et al., First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community, Sci. Total Environ. 728 (2020) 138764.
[3] L. Alvarez-Belon, A. Sarnowski, L.G. Forni, COVID-19 infection and the kidney, Br. J. Hosp. Med. 81 (10) (2020).
[4] M.S. Anker, et al., Weight loss, malnutrition, and cachexia in COVID-19: facts and numbers, J. Cachexia Sarcopenia Muscle 12 (1) (2021) 9–13.
[5] I.P.A. Virgens, et al., Can COVID-19 be a risk for cachexia for patients during intensive care? Narrative review and nutritional recommendations, Br. J. Nutr. (2020) 1–9.
[6] P.-Y. Wang, Y. Li, Q. Wang, Sarcopenia: an underlying treatment target during the COVID-19 pandemic, Nutrition (Burbank, Los Angeles County, Calif.) 84 (2020) 111104.
[7] L. García de Guadiana-Romualdo, et al., Characteristics and laboratory findings on admission to the emergency department among 2873 hospitalized patients with COVID-19: the impact of adjusted laboratory tests in multicenter studies. A multicenter study in Spain (BIOCOVID-Spain study), Scand. J. Clin. Lab. Invest. (2021) 1–7.
[8] HMHospitales, Covid Data Save Lives, Available online: https://www.hmhospitales.com/coronavirus/covid-data-save-lives/english-version, 2020.
[9] M. Hubert, S. Verboven, A robust PCR method for high-dimensional regressors, J. Chemometr. 17 (2003) 438–452.
[10] H. Li, et al., Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration, Anal. Chim. Acta 648 (1) (2009) 77–84.
[11] H.-D. Li, et al., Recipe for revealing informative metabolites based on model population analysis, Metabolomics 6 (3) (2010) 353–361.
[12] S. Rezzi, et al., Classification of Gilthead Sea Bream (Sparus aurata) from 1H NMR lipid profiling combined with principal component and linear discriminant analysis, J. Agric. Food Chem. 55 (24) (2007) 9963–9968.
[13] H.D. Li, Q.S. Xu, Y.Z. Liang, libPLS: an integrated library for partial least squares regression and linear discriminant analysis, Chemometr. Intell. Lab. Syst. 176 (2018) 34–43.
[14] S. Akcay, T. Ozlu, A. Yilmaz, Radiological approaches to COVID-19 pneumonia, Turk. J. Med. Sci. 50 (2020) 604–610.
[15] A. Fumagalli, et al., Pulmonary function in patients surviving to COVID-19 pneumonia, Infection 49 (1) (2021) 153–157.
[16] R.M.A. Garcia, Clinical case serious pneumonia from covid-19, Rev. Rol Enfermeria 44 (5) (2021) 42–46.
[17] J.J. Qu, et al., Clinical characteristics of COVID-19 and its comparison with influenza pneumonia, Acta Clin. Belg. 75 (5) (2020) 348–356.
[18] M. Solomon, L. Ganti, Hospital course of a man with viral pneumonia caused by COVID-19, Cureus 12 (7) (2020).
[19] D.L. Nelson, M.M. Cox, in: Lehninger: Principles of Biochemistry, sixth ed., W. H. Freeman & Co, New York, 2014, p. 1119 (plus 17 pp glossary).
[20] C. Bruzzone, et al., SARS-CoV-2 infection dysregulates the metabolomic and lipidomic profiles of serum, iScience 23 (10) (2020).
[21] P. Panjawatanan, et al., A case of concomitant COVID-19 infection-induced acute respiratory distress syndrome and diabetic ketoacidosis: another challenge in fluid management, Cureus 12 (11) (2020).
[22] J.E. Morley, K. Kalantar-Zadeh, S.D. Anker, COVID-19: a major cause of cachexia and sarcopenia? J. Cachexia Sarcopenia Muscle 11 (4) (2020) 863–865.

[23] E. Pleguezuelos, et al., Severe loss of mechanical efficiency in COVID-19 patients, J. Cachexia Sarcopenia Muscle 12 (4) (2021) 1056–1063.
[24] T. Oikawa, et al., Association between temporal changes in C-reactive protein levels and prognosis in patients with previous myocardial infarction - a report from the CHART-2 Study, Int. J. Cardiol. 293 (2019) 17–24.
[25] T. Xu, et al., Combined effect of Helicobacter pylori infection and elevated C-reactive protein on 3-month prognosis of ischemic stroke, Atherosclerosis (2019) 287.
[26] S. Pothal, et al., C- reactive protein kinetic in the prognosis of ventilator-associated pneumonia, Eur. Respir. J. 54 (2019) 2264.
[27] L. DW, et al., C-Reactive protein is a predictor of prognosis of prostate cancer: a systematic review and meta-analysis, Ann. Clin. Lab. Sci. 50 (2) (2020) 161–171.
[28] S. Konishi, et al., C-reactive protein/albumin ratio is a predictive factor for prognosis in patients with metastatic renal cell carcinoma, Int. J. Urol. 26 (2019) 992–998.
[29] B.R. Sahua, et al., C-reactive protein: a promising biomarker for poor prognosis in COVID-19 infection, Clin. Chim. Acta 509 (2020) 91–94.
[30] T. Yamada, 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 509 (2020) 235–243.
[31] N.R. Sproston, J.J. Ashworth, Role of C-reactive protein at sites of inflammation and infection, Front. Immunol. 9 (2018).
[32] Q. Ye, B. Wang, J. Mao, The pathogenesis and treatment of the 'Cytokine Storm' in COVID-19, J. Infect. 80 (6) (2020) 607–613.
[33] E. Poggiali, et al., Lactate dehydrogenase and C-reactive protein as predictors of respiratory failure in CoVID-19 patients, Clin. Chim. Acta 509 (2020) 135–138.
[34] K. Högner, et al., Macrophage-expressed IFN-β contributes to apoptotic alveolar epithelial cell injury in severe influenza virus pneumonia, PLoS Pathog. 12 (6) (2021).
[35] J.L. Sepulveda, Challenges in routine clinical chemistry analysis: proteins and enzymes, in: A. Dasgupta, J.L. Sepulveda (Eds.), Accurate Results in the Clinical Laboratory, Elsevier, 2019, pp. 141–163.