



OPEN ACCESS

EDITED BY

Melanie L. Hart,
University of Freiburg Medical Center,
Germany

REVIEWED BY

Pier Maria Fornasari,
REGENHEALTHSOLUTIONS, Italy
You Zhou,
Cardiff University, United Kingdom
Shailesh Shirguppikar,
Rajarambapu Institute of Technology, India

*CORRESPONDENCE

Leon Gerard D'Cruz
✉ Leon.D'Cruz@porthosp.nhs.uk;
✉ darthcruz@gmail.com

†These authors share first authorship

RECEIVED 23 March 2023

ACCEPTED 25 July 2023

PUBLISHED 05 September 2023

CITATION

Wiffen L, D'Cruz LG, Brown T,
Higenbottam TW, Bernstein JA,
Campbell C, Moellman J, Ghosh D,
Richardson C, Weston-Davies W and
Chauhan AJ (2023) Clinical severity classes
in COVID-19 pneumonia have distinct
immunological profiles, facilitating risk
stratification by machine learning.
Front. Immunol. 14:1192765.
doi: 10.3389/fimmu.2023.1192765

COPYRIGHT

© 2023 Wiffen, D'Cruz, Brown, Higenbottam,
Bernstein, Campbell, Moellman, Ghosh,
Richardson, Weston-Davies and Chauhan.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Clinical severity classes in COVID-19 pneumonia have distinct immunological profiles, facilitating risk stratification by machine learning

Laura Wiffen^{1†}, Leon Gerard D'Cruz^{1,2†*}, Thomas Brown¹,
Tim W. Higenbottam³, Jonathan A. Bernstein⁴,
Courtney Campbell⁵, Joseph Moellman⁶, Debajyoti Ghosh⁴,
Clive Richardson³, Wynne Weston-Davies³
and Anoop J. Chauhan^{1,2}

¹Research and Innovation Department, Portsmouth Hospitals University National Health Service (NHS) Trust, Portsmouth, United Kingdom, ²School of Pharmacy & Biomedical Science, University of Portsmouth, Portsmouth, United Kingdom, ³AKARI Therapeutics PLC, London, United Kingdom, ⁴Department of Internal Medicine, Division of Rheumatology, Allergy and Immunology, University of Cincinnati College of Medicine, Cincinnati, OH, United States, ⁵Ohio State University Medical Centre, Department of Cardiovascular Medicine, Columbus, OH, United States, ⁶Department of Emergency Medicine, University of Cincinnati College of Medicine, Cincinnati, OH, United States

Objective: Clinical triage in coronavirus disease 2019 (COVID-19) places a heavy burden on senior clinicians during a pandemic situation. However, risk stratification based on serum biomarker bioprofiling could be implemented by a larger, nonspecialist workforce.

Method: Measures of Complement Activation and inflammation in patientS with CoronAvirus DisEase 2019 (CASCADE) patients ($n = 72$), (clinicaltrials.gov: NCT04453527), classified as mild, moderate, or severe (by support needed to maintain $SpO_2 > 93\%$), and healthy controls (HC, $n = 20$), were bioprofiled using 76 immunological biomarkers and compared using ANOVA. Spearman correlation analysis on biomarker pairs was visualised *via* heatmaps. Linear Discriminant Analysis (LDA) models were generated to identify patients likely to deteriorate. An X-Gradient-boost (XGB) model trained on CASCADE data to triage patients as mild, moderate, and severe was retrospectively employed to classify COROnavirus Nomacopan Emergency Treatment for covid 19 infected patients with early signs of respiratory distress (CORONET) patients ($n = 7$) treated with nomacopan.

Results: The LDA models distinctly discriminated between deteriorators, nondeteriorators, and HC, with IL-27, IP-10, MDC, ferritin, C5, and sC5b-9 among the key predictor variables during deterioration. C3a and C5 were elevated in all severity classes vs. HC ($p < 0.05$). sC5b-9 was elevated in the "moderate" and "severe" categories vs. HC ($p < 0.001$). Heatmap analysis shows a pairwise increase of negatively correlated pairs with IL-27. The XGB model indicated sC5b-9, IL-8, MCP1, and prothrombin F1 and F2 were key discriminators in nomacopan-treated patients (CORONET study).

Conclusion: Distinct immunological fingerprints from serum biomarkers exist within different severity classes of COVID-19, and harnessing them using machine learning enabled the development of clinically useful triage and prognostic tools. Complement-mediated lung injury plays a key role in COVID-19 pneumonia, and preliminary results hint at the usefulness of a C5 inhibitor in COVID-19 recovery.

KEYWORDS

COVID-19, biomarker, nomacopan, cytokine, machine learning, risk stratification, complement

Introduction

The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) identified first in Wuhan, China, from human airway epithelial cells, was responsible for the coronavirus disease 2019 (COVID-19) (1). The clinical presentation of COVID-19 is highly variable (2); severe COVID-19 pneumonia presents with acute respiratory failure, septic shock, and multiorgan failure and may result in death (3).

Initial assessment and triage of patients by senior clinicians can often accelerate treatment *via* judicious planning and appropriate use of investigations; however, this places a huge burden on senior clinicians in a pandemic situation (4). However, triage strategies incorporating serum biomarker profiling, are implementable by nonspecialist and allied healthcare staff, thus reducing the bottleneck of waiting for senior clinician reviews.

During the first (March 2020–May 2020) and second (September 2020–April 21) waves of the pandemic, we risk-stratified COVID-19 patients by their clinical status and oxygenation requirements. A variety of cytokines, serum biomarkers, blood components, and complement proteins were evaluated in the CASCADE study ($n = 72$), initially to determine if there were “immunological fingerprints”, distinct to the clinical severities and thereafter to employ these in developing two predictive algorithms: the first, to identify patients likely to deteriorate clinically, and the second, a model to risk-stratify patients as “mild”, “moderate”, and “severe”. Since these models were trained on the levels of serum immunological proteins and not on viral proteins, the future applicability of the model may not be limited to COVID-19 infections caused by viral genotypic strains of the first and second waves.

The destructive role of complement in lung injury and COVID-19 pneumonia is well documented (5–7) and was also evident from this study (8). Complement is activated by SARS-CoV-2 *via* the lectin and classical pathways (9), while the SARS-CoV-2 spike protein (subunits 1 and 2) directly activates the alternative pathway (10).

Intuitively, therefore, complement inhibitors may have a role in influencing outcomes in the pandemic, and nomacopan, a complement C5 and leukotriene B4 (LTB4) inhibitor previously

used in the treatment of paroxysmal nocturnal haemoglobinuria (PNH) (11) and bullous pemphigoid (12), was trialled in a small group of COVID-19 patients (CORONET study, $n = 7$), classified as “severe” using the model developed in the CASCADE study. The results from the CASCADE and CORONET studies are presented here, together with the analysis of biomarkers and the machine-learning algorithms developed as a consequence.

Methods

Study design, participant description, and approvals

CASCADE was an observational cohort study that enrolled 52 COVID-19 quantitative polymerase chain reaction (qPCR)-positive patients admitted to the Portsmouth Hospitals University NHS Trust, and 20 qPCR-negative volunteers (CASCADE H), without comorbidities and those with stable, chronic medical conditions including diabetes and hypertension, composed of nonclinical hospital staff (Figure 1A).

The CORONET USA study enrolled five qPCR-positive COVID-19 patients from the University of Cincinnati Medical Centre (UCMC) and two qPCR-positive patients from the Ohio State University Medical Centre (OSUMC). The same inclusion and exclusion criteria were used for the CASCADE and CORONET studies (Figure 1A) and Supplementary Table S1.

The CASCADE trial details are summarised at “clinical trials.gov” [NCT04453527 (13)] and approved by the South-Central Berkshire Research Ethics Committee (REC reference: 20/SC/0228 22nd May 2020). The CORONET-study patients were treated with nomacopan in an open-label trial under a nonemergency “investigational new drug” (IND) application approved as an “expanded access application” by the FDA (13). Data and Safety Management Committees (DMSC) at each site provided safety oversight for this study.

All patients in both studies and control subjects in the CASCADE study provided written informed consent to obtain additional blood samples if their respiratory failure worsened, and

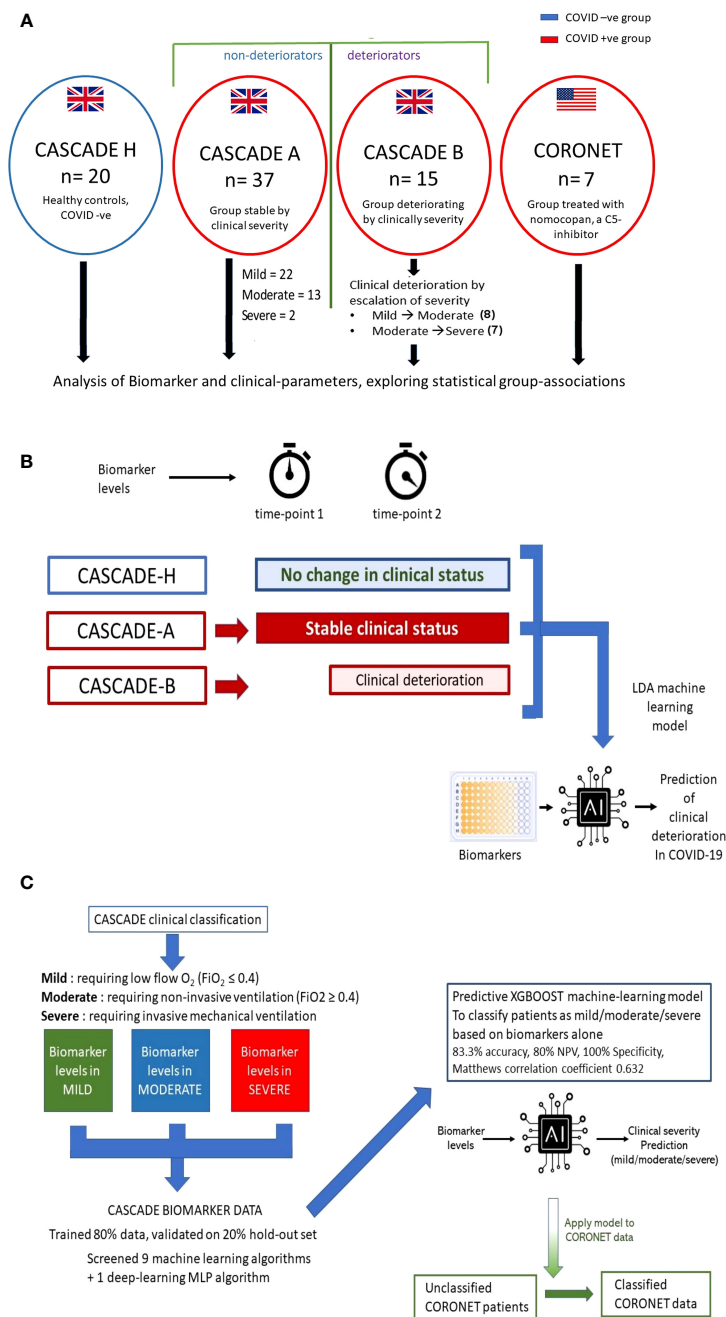


FIGURE 1

Outline of CASCADE and CORONET studies and description of predictive models generated. (A) Healthy controls (CASCADE H) were recruited from volunteers with stable, chronic medical conditions, including diabetes and hypertension, from nonclinical staff at Portsmouth Hospitals University NHS Trust (PHU). Participants in CASCADE A remained stable clinically throughout their hospital stay, whereas CASCADE B patients deteriorated clinically. In the event of deterioration with worsening respiratory failure, where patients transitioned to the next severity category (i.e., mild to moderate, moderate to severe) repeat blood samples were collected for biomarker analysis. Participants from the CORONET study were treated with nomocopan®. (B) The linear discriminant analysis (LDA) model was trained using an 80:20 split of the data, comprising CASCADE H participants, CASCADE A (patients who remained clinically stable throughout their hospital stay), and CASCADE B (patients who deteriorated at timepoint = 2). Admission biomarker data (timepoint 1) for CASCADE B patients were analysed initially together with the nondeteriorating group (CASCADE A) and healthy controls (CASCADE H) using LDA. Similarly, a second LDA classifier model was constructed using biomarker data from CASCADE B at point of clinical deterioration (time) compared to CASCADE A and CASCADE H participants. (C) Biomarker levels by clinical severity classes in the CASCADE study were used to train nine machine-learning and one deep-learning algorithm, the X-Gradient Boosting (XGBOOST) algorithm (trained and validated using an 80:20, train-test split) performed best (key performance metrics shown). The final XGBoost model was used to classify the CORONET patient bio-profile data retrospectively.

CORONET patients also provided written informed consent to receive nomacopan.

Nomacopan treatment

CORONET patients received a 45-mg subcutaneous dose of nomacopan on admission and two additional doses at 12-h intervals. After the initial three doses, nomacopan was administered at 45 mg every 24 h for a maximum of 12 days, unless they were discharged or passed away sooner. Patients were monitored for achieving normal 50% haemolytic complement (CH50) activity serum levels on day 3 or at discharge, similar to what has previously been reported for the C5 antibody, eculizumab (14, 15).

Patients received either a prophylactic beta-lactam or cephalosporin antibiotic (or an alternative agent if there was a penicillin or cephalosporin allergy) as terminal complement complex therapy has been linked to risk of meningococcal infection (16).

Clinical severity of patients ranking in CASCADE

COVID-19 patients admitted in respiratory failure and recruited to the CASCADE study were ranked as mild, moderate, or severe respiratory failure according to the level of support required to maintain arterial oxygen saturation (SpO₂) above 93%. Respiratory distress for patients was defined as mild if requiring low-flow oxygen (fraction of inspired oxygen; FiO₂ ≤ 0.4), moderate if requiring FiO₂ > 0.4, with or without noninvasive ventilation support (e.g., high-flow nasal oxygen or continuous positive airway pressure), and severe if requiring invasive mechanical ventilation.

For the CASCADE study, admitted patients who remained clinically stable entered the CASCADE A group (*n* = 37), and those who, on admission, deteriorated (moved between categories, from mild → moderate or moderate → severe or mild → severe) entered the CASCADE B group (*n* = 15), summarised in [Figure 1A](#). If CASCADE A patients demonstrated further clinical deterioration over 4 to 10 days after admission, they were reclassified to the CASCADE B group.

Biological samples

COVID-19 qPCR nasal and throat swabs were obtained from all participants and processed according to standard published protocols (17) and as described in the [Online Data Supplement](#). Blood samples obtained by venesection were handled using appropriate containment procedures (level 2). Serum and plasma samples were shipped on dry ice to external laboratories for biomarker analysis. Laboratory personnel were blinded to clinical

information. Blood samples were obtained from 20 healthy controls (CASCADE H) and the 52 patients in the CASCADE A and CASCADE B groups for biomarkers.

Statistics and machine-learning analysis

Heatmap analysis

A total of 76 biomarkers were analyzed in the CASCADE study. Spearman correlation coefficients between pairs of mean biomarker levels within each severity class of all patients and healthy controls in CASCADE were assembled into a 76 × 76 matrix. Further agglomerative hierarchical clustering and calculation of Euclidean distances between clusters were employed to build heatmaps (Seaborn API, ver 0.11.2). The colour map approaching hues of blue represents an inverse relationship between pairs of biomarkers, as one increases, the other variable decreases, moving in opposite directions while hues approaching red indicate positive (additive) correlation, where biomarker pairs change in their values together with the same sign.

Selection of machine-learning algorithms for analysis

To select the best machine-learning algorithm to address the analyses, nine machine-learning algorithms; K-nearest neighbour, ADA-Boost, Decision Trees, Random Forest, Extra Trees, Support Vector Classifier, X-Gradient Boost, Logistic Regression, and Linear Discriminant Analysis, and one multilayer perceptron (deep neural net) algorithm were screened using an iterative grid-search method in Python, using the training-dataset at the cross-validation step. The accuracy of prediction was employed as the metric. Hyperparameters for the chosen algorithm were then further optimised, further details are provided in the [Online Data Supplement](#).

Linear discriminant analyses model for clinical deterioration in COVID-19

Two Linear Discriminant Analysis (LDA) models (model 1 for biomarkers collected only at timepoint 1 and model 2, biomarkers collected at timepoint 2) were generated in the CASCADE study, coded in Python 3.8, using Scikit-learn (ver.0.16.1). Timepoint 1 was at the time of recruitment and admission to hospital, and timepoint 2 was at the point of clinical deterioration. The model was trained and validated using an 80:20 train-test split. Hyperparameters were optimised and determined following extensive experimentation and testing of models while employing a grid-search method (18). Class imbalances were addressed using the Synthetic Minority Over-sampling Technique (SMOTE) method (19).

LDA machine-learning models explored if serum biomarker levels could identify and discriminate between patients who remained stable clinically and those who were likely to deteriorate, as described schematically in [Figure 1B](#). Details are provided in the [Online Data Supplement](#).

X-Gradient Boosting model for classification of mild, moderate, and severe in COVID-19

An X-Gradient Boosting (XGBoost) model was generated and trained on biomarker levels measured on CASCADE data from patients at different clinical severity levels (Figure 1C). The data were split into 80:20, train:test ratios. Nine machine-learning algorithms and one neural-net algorithm (multi-layer perceptron) were screened using the k-means stratified cross-validation (number of splits = 5) Hyperparameters were tuned using a grid-search method, and details of these are shown in the [Online Data Supplement](#). Mean accuracy was used to evaluate the performance of the algorithms. The final validated model was used to retrospectively classify CORONET patients (Figure 1C).

Statistical analysis

All statistical analysis was carried out with SPSS version 25 (SPSS Version 25.0. Armonk, NY: IBM Corp), SciPy module (version 1.3) for Python (version 3.7.2) and R (version 3.60). Visualisation plots were created using Excel (Microsoft), Prism (GraphPad), or scripted using Python with Matplotlib (version 3.43) or Seaborn (version 0.11.2) software. Significance values in figures and tables are displayed using American Psychological Association (APA) styles (e.g., ns if $p > 0.05$; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$).

Results

Demographics and baseline features of CASCADE and CORONET participants

Figure 1A summarises patient recruitment details. Table 1 admission demographics compare patients in CASCADE A and CASCADE B COVID-19-positive groups with CASCADE H healthy controls. Patients in the CASCADE study were clinically triaged as mild, moderate, and severe. Patients under the mild category had an $\text{FiO}_2 \leq 0.4$, and those under the moderate category: $\text{FiO}_2 \geq 0.4$ and/or the use of noninvasive ventilation (NIV) or CPAP. Patients classified as severe were those who needed invasive/mechanical ventilation.

CASCADE A were patients who remained stable within their clinical triage status before either recovery and discharge from the hospital or succumbing to death. CASCADE B had patients transitioning between categories of clinical severities (the deteriorating group), as summarised schematically in Figure 1A.

Table 2 shows that age, gender, BMI, and length-of-stay were not statistically significant between the CASCADE A, CASCADE B, and CORONET groups by the Kruskal–Wallis test and, as such, were not considered confounding variables when comparing groups. Categorical variables such as sex, smoking status, comorbidities, and medication history were compared using

TABLE 1 Demographics and salient features of patients enrolled within the CASCADE (A and B) study.

	CASCADE H (<i>n</i> = 20)	CASCADE (A+B) COVID19 +ve (<i>n</i> = 52)	<i>p</i> -value	Mild	Moderate	Severe	<i>p</i> -value
Age	47.7 ± 13.5	64.4 ± 16.5	***	63.8 ± 16.2	65.5 ± 16.9	62.7 ± 16.4	ns
Gender (% male)	4 (20%)	34 (65%)	***	17 (57%)	21 (75%)	7 (78%)	ns
BMI	27.9 ± 5.1	32.4 ± 6.5	**	32.3 ± 6.6	31.9 ± 6.8	33.3 ± 6.6	ns
Smoking status							
Never (%)	6 (30%)	31 (57%)	*	20 (64.5%)	9 (45%)	2 (100%)	ns
Ex (%)	10 (50%)	22 (43%)	*	11 (35.5%)	11 (55%)	0	–
Current (%)	4 (20%)	0	–	0	0	0	–
Ethnicity							
Caucasian (%)	18 (90%)	50 (94)	ns	28 (90%)	20 (100%)	2 (100%)	ns
Asian (%)	1 (5%)	2 (4%)	ns	2 (7%)	0	0	ns
Afro-Caribbean (%)	1 (5%)	0	ns	0	0	0	ns
Other (%)	4 (20.0%)	0	ns	1 (3%)	0	0	ns
Comorbidities							
None	13 (65%)	20 (38.0%)		12 (39%)	7 (35%)	1 (50%)	–
1	6 (30%)	11 (20%)		7 (22%)	5 (25%)	0	–
≥2	1 (5%)	22 (42%)		12 (39%)	8 (40%)	1 (50%)	–
14-day outcome							
Discharged	–	–	–	24 (80%)	14 (50%)	1 (11%)	–
Inpatient	–	–	–	3 (10%)	8 (29%)	6 (67%)	**
RIP	–	–	–	3 (10%)	6 (21%)	2 (22%)	**

ns if $p > 0.05$; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

TABLE 2 Baseline demographics, co-morbidities, and relevant medication history of participants in the study.

	(CASCADE A) nondeteriorating COVID-19 patients (n = 37)	(CASCADE B) deteriorating COVID-19 patients (n = 15)	(CORONET) patients on nomacopan (n = 6)	p-value
Sex (M:F)	25 (67.6%):12 (32.4%)	10 (66.7%):5 (33.3%)	6 (100%):0 (0%)	0.249
Age	65.25 ± 16.69	63.27 ± 17.32	50.0 ± 11.3	0.120
BMI	32.6 ± 7.7	31.9 ± 5.2	32.0 ± 3.8	0.856
Average length of stay (days)	11.0 ± 7.4	13.3 ± 5.9	6.0 ± 7.0	0.137
Smoking status				
Never smoked	20 (54.1%)	10 (66.7%)	4 (66.7%)	0.645
Ex-smoker	17 (45.9%)	5 (33.3%)	2 (33.3%)	
Current smoker	0 (0%)	0 (0%)	0 (0%)	
Comorbidities				
Diabetes	8 (21.6%)	3 (20.0%)	5 (83.3%)	0.005
Hypertension	12 (32.4%)	3 (20.0%)	4 (66.7%)	0.120
Heart disease	9 (24.3)	5 (33.3%)	0 (0%)	0.272
Chronic kidney disease	5 (13.5%)	3 (20.0%)	1 (16.7%)	0.840
Previous VTE	2 (5.4%)	2 (13.3%)	1 (16.7%)	0.496
Malignancy	3 (8.1%)	1 (6.7%)	1 (16.7%)	0.749
Immunosuppression	0 (0%)	1 (6.7%)	0 (0%)	0.233
Medication history				
Long-term anticoagulation	35 (94.6%)	13 (86.7%)	1 (16.7%)	<0.001
ACE inh/ARB	8 (21.6%)	4 (26.7%)	3 (50.0%)	0.337
Statins	12 (32.4%)	5 (33.3%)	3 (50%)	<0.001
Inhaled corticosteroids	3 (8.1%)	5 (33.3%)	1 (50%)	0.740

cross-tab and Chi-square analysis, $p < 0.05$ (95% confidence interval (CI)) was regarded as statistically significant. The average BMI of all participants across the three groups was in the range clinically classified as “obese” (BMI > 30). Participants in the nomacopan[®] treatment group (CORONET) had a significantly higher proportion of diabetics (type 2 DM), compared to CASCADE A or CASCADE B. Participants in CASCADE A and CASCADE B had a significantly larger proportion of participants on long-term anticoagulation compared to the nomacopan[®] treatment group. There was a higher proportion of participants on statins in the nomacopan[®] group ($p < 0.001$) compared to the CASCADE A and CASCADE B groups.

CASCADE (mild, moderate, and severe) and CORONET patients are distinct, using ROX, NEWS2, and SOFA scores

The ROX index is a clinical assessment of the degree of a patient’s hypoxaemic respiratory failure and need for intubation.

Normal ROX indices were calculated with FiO₂ of 21%, respiratory rates between 12 and 18 breaths per minute, and SpO₂ of 95%–98%. Patients in the severe category were intubated and received oxygen-enriched air, thus having a higher than atmospheric FiO₂, this accounts for the higher mean value and the slightly elevated range of maxima and minima within the severe category compared to the moderate category.

When grouped by ROX indices, NEWS2, and mSOFA scores, participants in the clinical severity groups were significantly distinct from each other using Dunn’s multiple comparison tests (Figure 2). All seven patients in the severe category were mechanically ventilated.

Analysis of markers of inflammation, coagulation, and complement

Markers of hyperinflammation, degree of lymphopaenia, neutrophil levels, TNF- α , INR levels, and proinflammatory cytokines (interleukin (IL)-6 and IL-8) were significantly

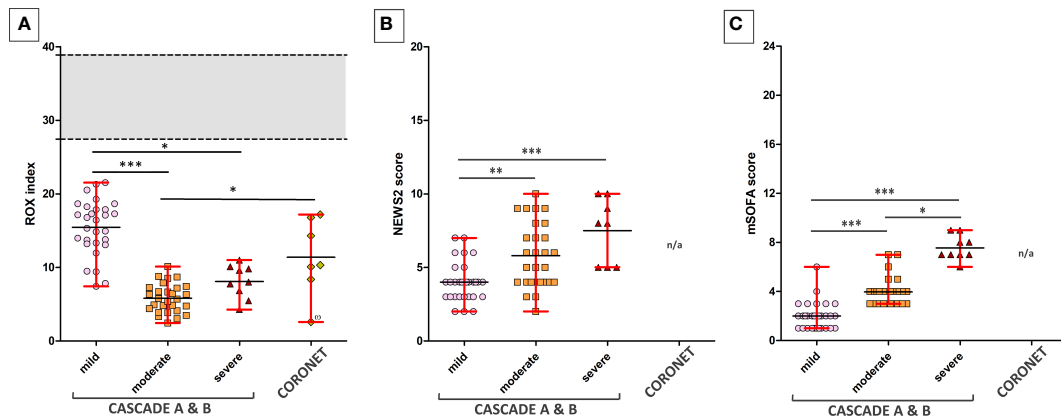


FIGURE 2

ROX indices of patients across various clinical severities in the CASCADE and CORONET studies. (A) The ROX index for normal, healthy individuals is shown as the grey area. The upper and lower limits shown in red bars represent the maximum and minimum data within that group. Black horizontal lines represent the average ROX index. CASCADE B comprised patients who deteriorated while in hospital, progressing either from mild to moderate or moderate to severe categories, while CASCADE A patients remained stable within that category. CORONET patients treated with nomacopan[®] also had ROX indices below that of normal, healthy individuals. One patient in the CORONET study (data point marked with the number sign) had late commencement of treatment due to delays in reaching the treatment centre in Ohio, required invasive mechanical ventilation, and unfortunately died 13 days following admission due to COVID-19-related complications. (B) Mild, moderate, and severe patients in CASCADE A & B grouped by NEWS2 scores (C) and grouped by mSOFA scores. Data for NEWS2 and mSOFA were not available (n/a) for CORONET patients. ns if $p > 0.05$; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

increased in COVID-19 patients compared to healthy controls ($p < 0.05$, 95% CI) (Figure 3). In addition, D-dimer levels were significantly elevated in COVID-19 patients in the severe category ($p < 0.05$, 95% CI) above levels for normal-healthy individuals (normal reference ranges: <250 ng/ml).

IL-5 levels were significantly raised in the mild and moderate groups when compared to the healthy controls (Figure 3); IL-13 levels, however, did not show any difference between any of the groups (Figure 3).

Figure 3 shows that patients infected with COVID-19 have deranged levels of IL-6, IL-8, IL-10, IL-12p40 (a component of IL-12 and IL-23 and a chemoattractant for macrophages to sites of inflammation), MIP-1b (responsible for macrophage chemotactic action), monocyte chemoattractant protein-1 (MCP-1, a cytokine related to thrombosis), chemokine (C-X-C motif) ligand 9 (CXCL9, a chemoattractant for activated T cells) and TNF- α (a proinflammatory cytokine, a target for immune biologics such as infliximab, adalimumab, or etanercept). INR levels in COVID-19 patients were significantly raised (normal healthy levels = 1.1). D-dimer levels were significantly raised (Figure 3), implying high circulating levels of fibrin degradation products, and concomitantly raised fibrinogen and prothrombin levels suggest a pro-coagulative state in COVID-19.

IL-27 levels appear low in healthy controls but are raised in all severity classes (Figure 3). Interestingly, the average IL-27 level appears to be lower in the severe class, compared to that in the mild and moderate classes, although the decrease does not appear to be significant ($p > 0.05$) (Figure 3). Interferon gamma-induced protein 10 (IP-10) appears to be raised significantly above the levels of healthy controls in all severity classes.

The complement components, C3a, C5a, and C5, were significantly elevated in all severities of COVID-19-infected

patients (mild, moderate, and severe) with pneumonia compared to healthy controls (Figure 4). Of note, the C5a ($p < 0.01$) and sC5b-9 ($p < 0.001$) levels were significantly different between the moderate and severe categories compared to healthy controls.

Levels of the terminal complement complex, sC5b-9 (the membrane-attack complex (MAC)) were significantly raised in COVID-19 patients (Figure 4). Levels of C5a and C5 were raised in COVID-19 patients, thereby offering a rationale for C5 inhibitor therapy in COVID-19; this is reported in the results of the CORONET study. Levels of C3a (an anaphylatoxin and cleavage product in the formation of the C3 convertase) were significantly raised, strongly confirming complement cascade activation in COVID-19 (Figure 4). This is also reflected in the significant rise of factor Bb levels (Figure 4), the fragment together with C3b which forms the C3 convertase, a key step in the activation of the complement cascade.

Levels of leukotriene B4 (LTB4), a member of the eicosanoid family of lipid mediators, are also shown here (Figure 4). Although not part of the complement cascade, nomacopan[®], the C5-inhibitor drug used in the CORONET study, binds tightly to LTB4. Limited samples were obtained for LTB4, and although these levels were higher in severe COVID-19 patients compared to stable COVID-19 patients and healthy controls, they were not statistically significant.

Heatmap and correlation of biomarkers by disease severity

The heatmaps in Figures 5–8 demonstrate a correlation between pairs of biomarkers, with red pixels indicating a positive correlation and blue pixels indicating a negative correlation. Together, the biomarker pairs show an additive effect when they highlight a

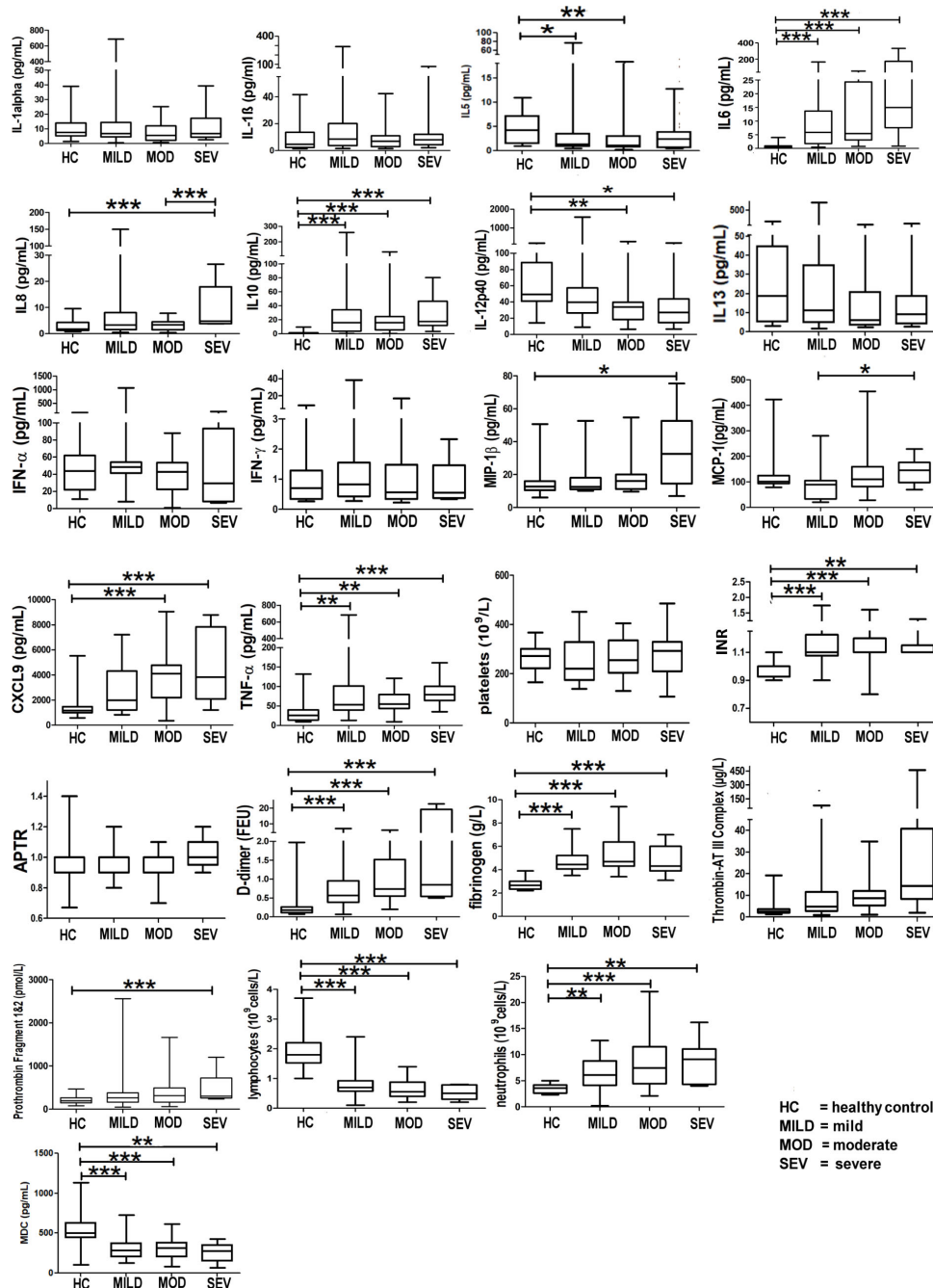


FIGURE 3

Markers of hyperinflammation and coagulation in healthy vs. COVID-19-infected patients. Mean values are shown as a horizontal line within the boxes; the whiskers indicate min and max ranges. (HC, healthy controls; MOD, moderate; MILD, mild; SEV, severe). ns if $p > 0.05$; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

positive correlation. An additional clustering step, applied to the heatmap analysis, redistributes the biomarker pairs into clusters, where those positively correlated are clustered together while the negatively correlated are grouped together. Cluster heatmaps have an advantage over unordered heatmaps, as they reorder the matrix based on the hierarchical clustering step, thus they display and condense large amounts of rank-ordered data into a compact space. The hierarchical structure is displayed as a dendrogram on the top

edge of Figures 5–8. This analysis demonstrated that the patient bio-profiles and their biomarkers changed with the severity of respiratory failure. For example, the number of blue pixels increased from Figures 5 to 8, and Figure 8 contained the greatest number of negatively correlated biomarkers (blue).

Neutrophil to lymphocyte ratio (NLR), a known severity predictor in COVID-19 (20), is positively correlated with Complement C5a and sC5b-9 and Complement C5 appears also to

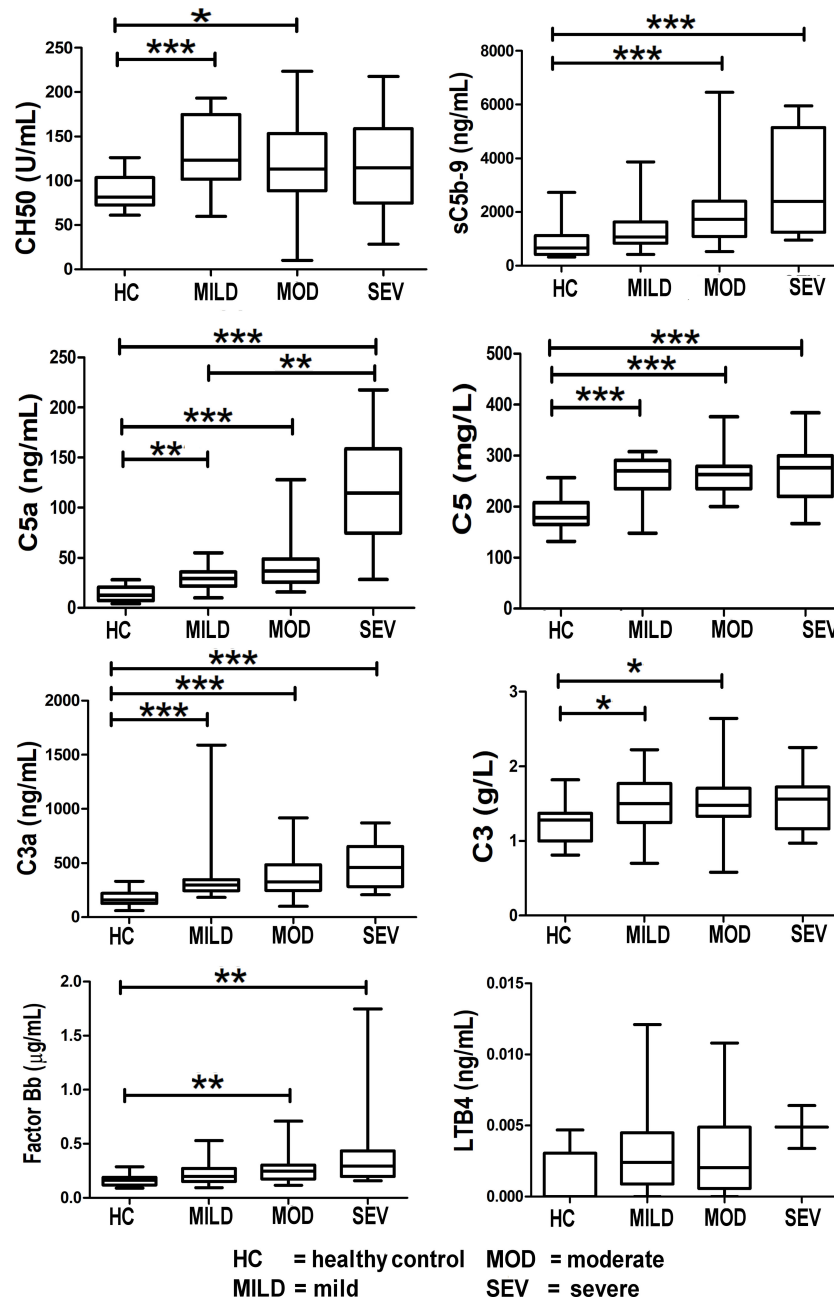


FIGURE 4

Levels of complement components and LTB4 in healthy controls compared to COVID-19-infected patients. Mean values are shown as a horizontal line within the boxes; whiskers indicate min and max ranges. Data were available for only two patients in the "severe" category for LTB4. (HC, healthy controls; MOD, moderate; MILD, mild; SEV, severe). ns if $p > 0.05$; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

be positively correlated with levels of D-dimer (Figure 8). NLR was positively correlated with CRP levels in the moderate group (Figure 7) while the relationship between NLR and CRP is inversely correlated in the "severe" class (Figure 8). Interestingly, complement C5 levels are correlated with levels of troponin in the severe class (Figure 8).

IL-27 levels appear to have fewer blue pixels in the mild (Figure 6) (7 blue pixels out of 30 horizontal pixel pairs) and

moderate (Figure 7) (6 blue pixels out of 30 horizontal pixel pairs), compared to 12 blue pixels out of 30 in the severe category (Figure 8). The negative correlation is observed since IL-27 levels appear to decrease with increasing severity; thus, in the pairwise relationship visualised by heatmap analysis, this is displayed as an increase in blue pixels between mild (Figure 6) and moderate (Figure 7) compared to the severe class (Figure 8).

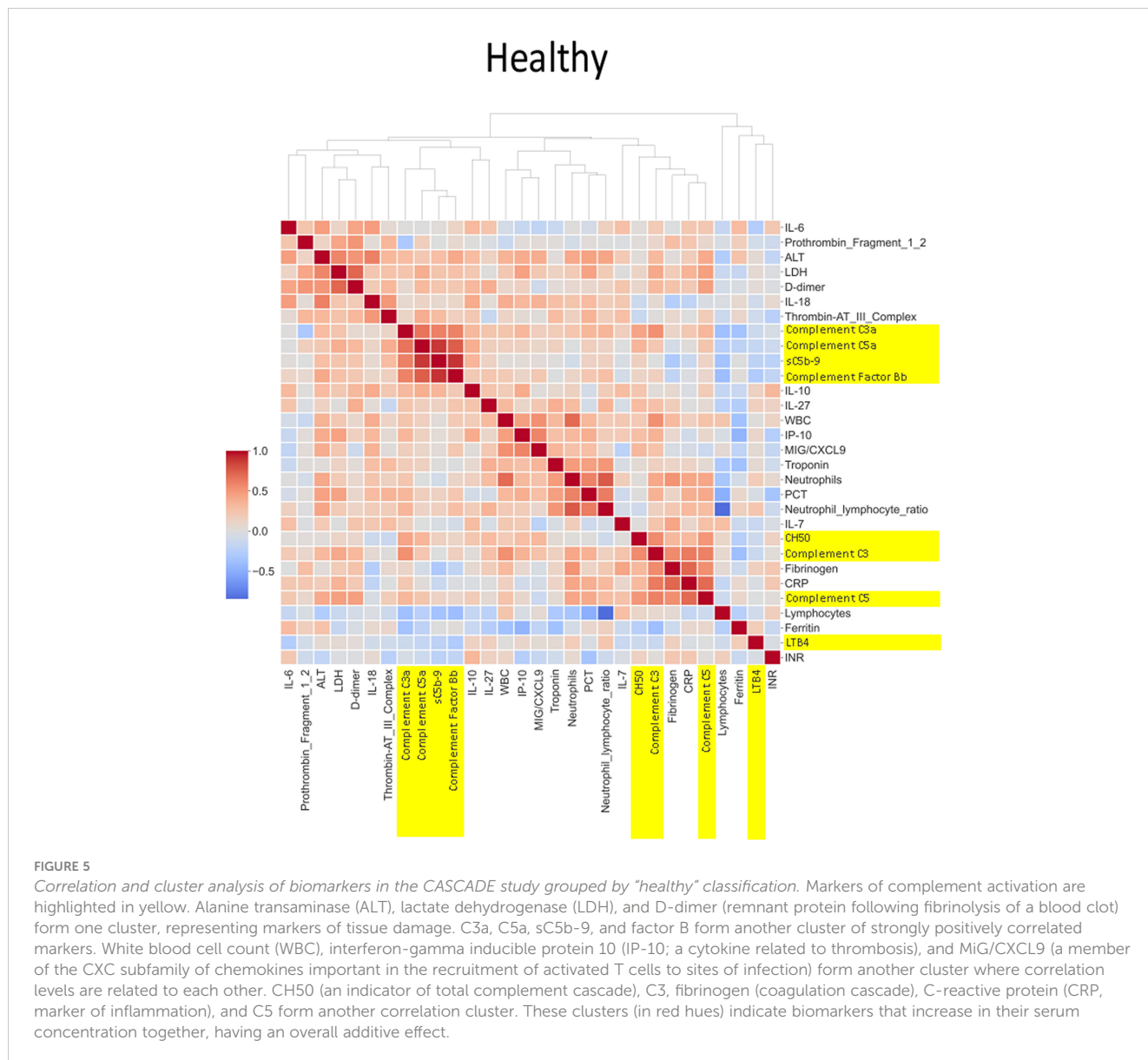


FIGURE 5
 Correlation and cluster analysis of biomarkers in the CASCADE study grouped by “healthy” classification. Markers of complement activation are highlighted in yellow. Alanine transaminase (ALT), lactate dehydrogenase (LDH), and D-dimer (remnant protein following fibrinolysis of a blood clot) form one cluster, representing markers of tissue damage. C3a, C5a, sC5b-9, and factor B form another cluster of strongly positively correlated markers. White blood cell count (WBC), interferon-gamma inducible protein 10 (IP-10; a cytokine related to thrombosis), and MIG/CXCL9 (a member of the CXC subfamily of chemokines important in the recruitment of activated T cells to sites of infection) form another cluster where correlation levels are related to each other. CH50 (an indicator of total complement cascade), C3, fibrinogen (coagulation cascade), C-reactive protein (CRP, marker of inflammation), and C5 form another correlation cluster. These clusters (in red hues) indicate biomarkers that increase in their serum concentration together, having an overall additive effect.

Linear discriminant analysis and discrimination between non-deteriorators (CASCADE A), deteriorators (CASCADE B), and healthy volunteers (CASCADE H)

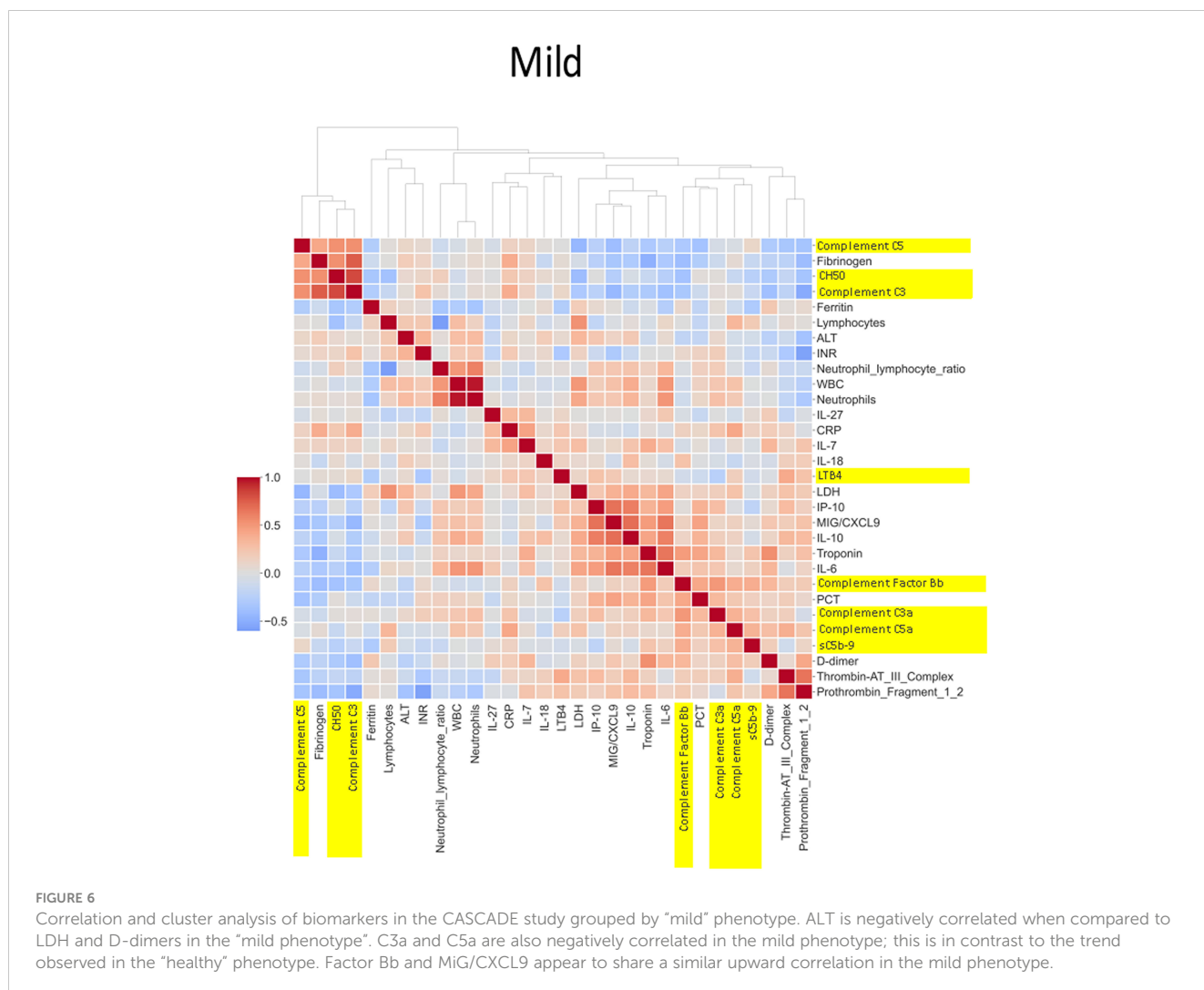
Two separate LDA models are summarised in Figure 9. One reflects the immunological landscape of COVID-19 patient predeterioration, and the other is captured from biomarker levels during deterioration. The biomarkers of importance vary between timepoint 1 and timepoint 2; however, some biomarkers appear in both timepoints, although their ranking in order of importance varies between the two (IL-27, macrophage-derived chemokine (MDC), PDGFAA, ferritin, and IP-10). The differences in the order and type of biomarkers provide a discriminative potential between the status of patients at these two time points. IL-27 emerged as the top predictor of clinical deterioration (Figure 9D). Complement C5 was among the top five biomarkers of predictive importance (for clinical deterioration) to the model (Figure 9D).

The metrics for both the LDA models showed 73% accuracy, 77% specificity, and 90.9% negative-predictive value (NPV).

Biomarker thresholds derived from the LDA models

Table 3 shows critical threshold values achieving the maximum sensitivity and specificity (associated with the Youden’s index) in both the LDA models, differentiating between the clinically deteriorating subcohort from the nondeteriorating subcohort. The Youden Index is the greatest potential effectivity of a biomarker, a common measure of the ROC curve (21, 22). The threshold values identify a point when the true-positive rate (TPR) is high and the false-positive rate (FPR) is low.

The level of circulating terminal complement complex sC5b-9 was between 8 and 11 times higher than that reported in normal, healthy donors (23). Threshold values help discriminate between



the deteriorating class of patients (CASCADE B) from those who were clinically stable (CASCADE A).

Comparison of clinical screening tools

The performance of LDA-model 2 against currently available surrogates of clinical deterioration (ROX index, mSOFA score, and NEWS2 score) was compared using ROC-AUC analysis (Table 4) using the “training data” of CASCADE A, CASCADE B, and CASCADE H. The performance of the LDA model is represented on the hold-out set as shown in Table 4 (ROC-AUC analysis using the holdout or test data (20%).

The mSOFA scores performed slightly better than NEWS2 scores, while LDA model 2 gave an overall superior performance (Table 4).

Machine-learning model for classification of clinical severity by biomarker levels

As explained in Figure 1C, the XGBoost model was trained on the CASCADE immune-biological data, to help triage patients by

their clinical severity. The model was validated on the hold-out set of data (20%) and gave a prediction accuracy of 83.33%, a positive predictive value of 1.0, a negative predictive value of 0.8, and a Matthew’s correlation coefficient of 0.632.

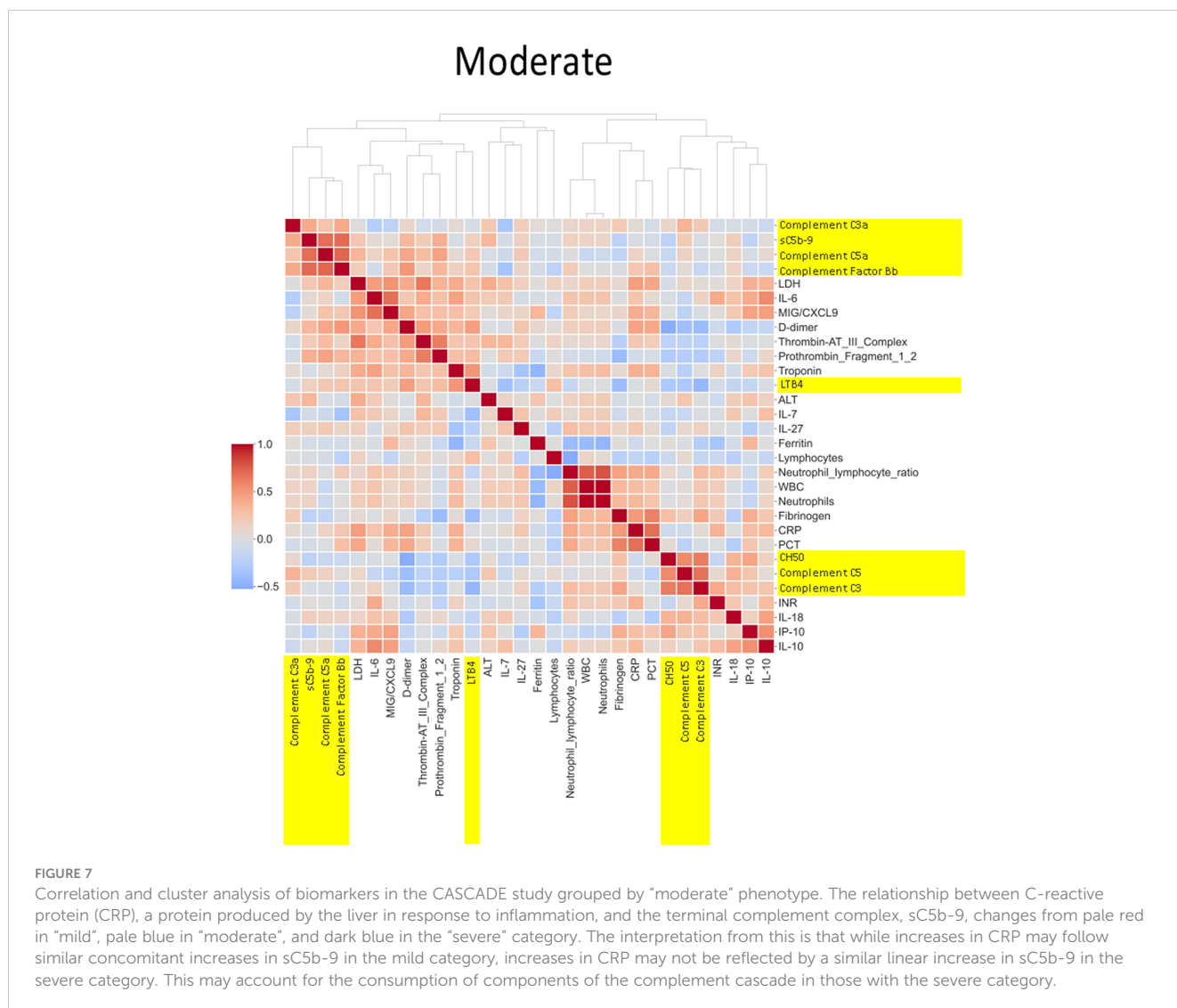
The XGBoost model was retrospectively used to determine the admission severity of patients in the CORONET study (Figure 1C). All patients receiving nomacopan in the CORONET study were classified as “severe” by the model trained on the CASCADE classification.

The terminal complement component, sC5b9, was the leading biomarker of importance in the classification of clinical severity by biomarker levels in the CASCADE study, as shown by the plot of mean SHAP values (Figure 10).

IL-8, a well-studied marker with associations with ARDS and lung inflammation, ranked second highest in its predictive properties and impact on the performance of the XGBoost model (Figure 10).

Nomacopan treatment in the *CORONET study

All nomacopan patients had ROX indices (SpO₂ <93%) below normal healthy individuals from CASCADE H (Figure 2). All were



admitted to the ICU, and two patients required early invasive mechanical ventilation. Six out of seven patients survived and were discharged 3 to 22 days after admission with normal CH50 measurements (Table 5). One female patient who died had delayed treatment with nomacopan by 3 days.

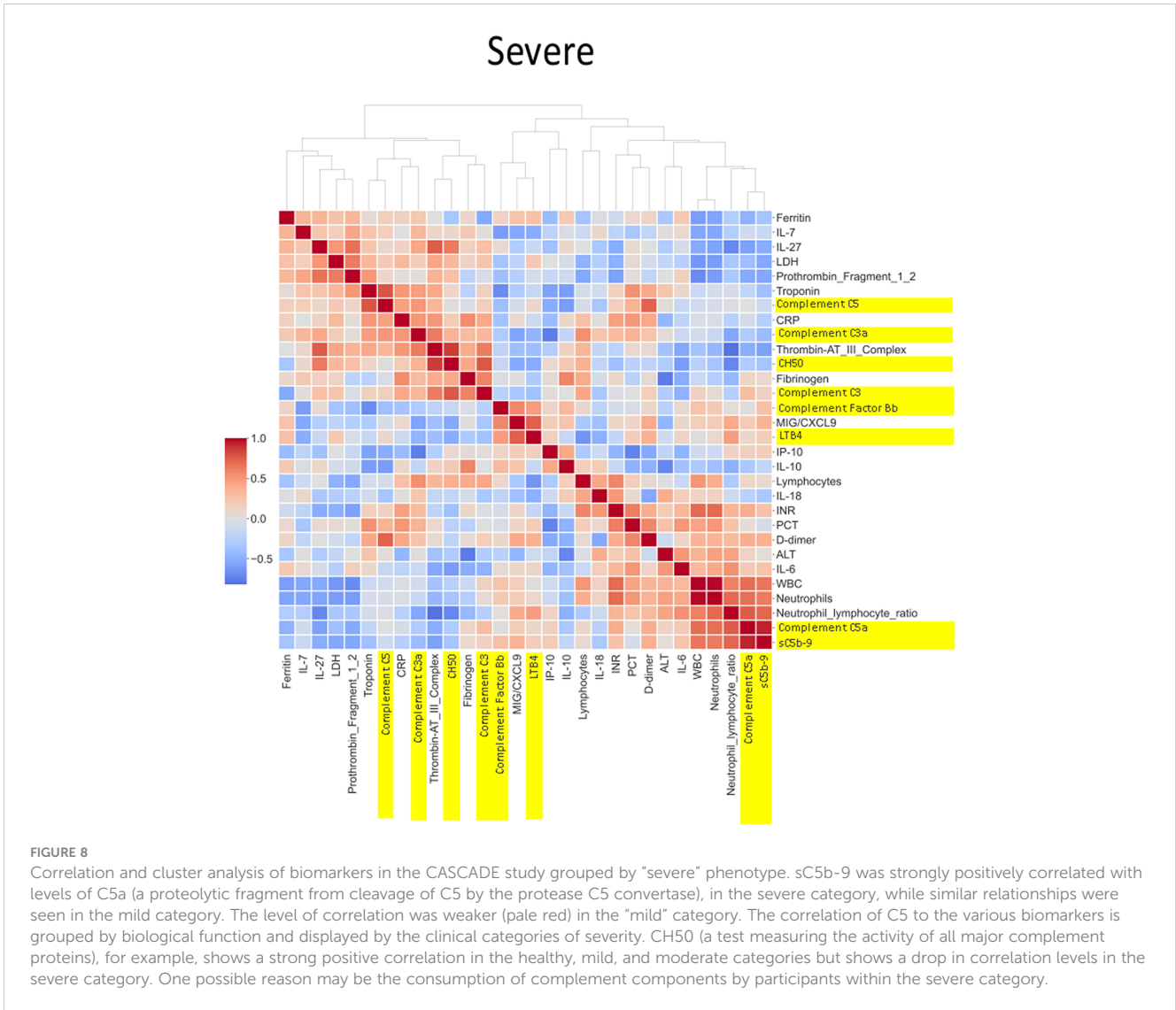
Discussion

The CASCADE study carried out in the UK was designed to determine if the separate classes (“healthy”, mild, moderate, and severe) of COVID-19 patients determined by clinical judgement and risk assessment would reveal distinct “immunological fingerprints”, by biomarker levels. Our findings have shown such fingerprints exist, and when the classification is automated *via* machine-learning algorithms, it would serve as a rapid risk assessment and triage tool. Secondly, distinct immunological fingerprints in COVID-19 patients were determined, allowing early identification of patients who might have the propensity to deteriorate clinically.

The CORONET study carried out in the USA applied a similar but smaller panel of biomarker assessment, but crucially, this study recruited seven patients with severe symptoms who were treated with an off-label complement C5 inhibitor, nomacopan, administered under specially procured permissions under the “compassionate-use” regulations from the FDA.

The two studies intended to determine the “immunological fingerprint” of patients categorised by their symptoms, unique to the clinical severity demonstrated during their hospital admission for COVID-19. Clinical-triage efforts during a pandemic can be a huge responsibility and burden on senior clinical physicians; however, if the risk assessment were based on bioprofiling using serum biomarker levels and aided by an artificial-intelligence-based model, this could be implemented by a wider, less specialist workforce.

The correlation coefficients displayed in the heatmaps describe the direction of the relationship between pairs of variables, where a positive correlation means that the pairs of variables are either both high or both low at the time of biomarker measurement. This is represented as red hues in the heatmap analysis. A negative



correlation occurs when one variable is high and the other is low at the time of biomarker measurement; these are represented as blue hues in the heatmap analysis.

The heatmap analysis displays a “bird’s eye view” of an immunological fingerprint by comparing pairs of variables within each severity classification. Clusters with red pixels indicate variable pairs that are increasing or decreasing in their values together. The blue pixels indicate values in opposite directions to each other. The paired relationships shown here are primarily a statistical representation of the direction of the relationship in the biomarker levels; they may not necessarily reflect a biological relationship where the level of one biomarker affects the direction of the other.

An example of where the levels of one biomarker might impinge on the levels of another is illustrated in the relationship of the neutrophil-lymphocyte ratio (NLR) to CRP levels. Our analysis shows that NLR is positively correlated to CRP levels in the moderate group (Figure 7). The importance of both these biomarkers has been shown in a separate study, where the utility of these markers are excellent diagnostic predictors of COVID-19 in

a binary logistic regression model (24). However, our heatmap analysis adds a very important angle to the immunological biomarker study since Figure 8 shows the NLR ratio is inversely correlated to CRP in the “severe” class. This conflicting observation (between that in the moderate vs. severe) could be explained as follows: neutrophil concentrations when high, (as in the “severe” class), restrict viral replication *via* degranulation, phagocytosis, and the release of neutrophil traps and subsequently reduce viral titres with a concomitant lowering of CRP values as have been reported previously (25). Thus, the “compartmentalisation” of the immunological fingerprints in COVID-19 through the mild, moderate, and severe severity classes are captured uniquely in our use of heatmaps as a visual representation. The association of specific levels or ranges of cytokines with disease severity is not a new concept and has been described before in the literature (26, 27).

The release of neutrophil elastase from highly activated neutrophils mediates lung injury in sepsis, including COVID-19 (28), and inhibitors to neutrophil elastase (sivelestat) are protective in inflammatory lung injury (29), and have been proposed to be beneficial in COVID-19-related lung injury (30). Increased levels of

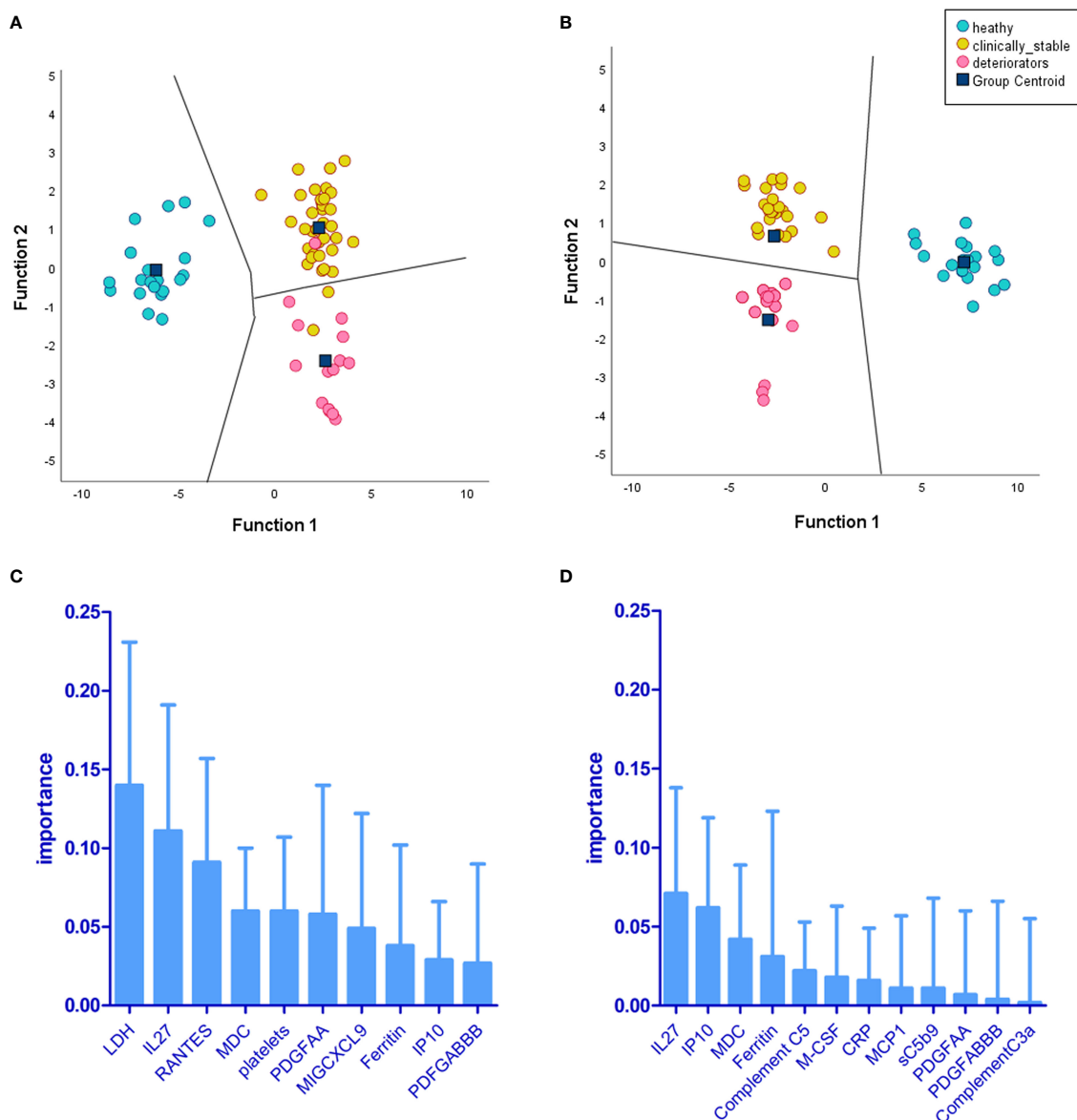


FIGURE 9

Linear discriminant analysis (LDA) demonstrates levels of biomarkers and cell counts, providing separation of deteriorating patients (CASCADE B) from nondeteriorating (CASCADE A) and healthy participants (CASCADE H). Two LDA models were constructed: (A) the territorial map with canonical discriminant functions incorporating timepoint 1 (at admission) of markers for patients in CASCADE A, CASCADE B, and CASCADE H and (B) the territorial map with canonical discriminant functions at timepoint 2 of markers for patients in CASCADE A, CASCADE B, and CASCADE H (bloods assayed for markers on day 1 of clinical deterioration of CASCADE B patients). The territorial maps (A, B) show separation between the groups based on the ranking of biomarkers (C) at timepoint 1 and (D) at timepoint 2. There was an increased separation of deteriorating patients (CASCADE B) from clinically stable patients (CASCADE A) and healthy controls (CASCADE H) in LDA model 2 (B) without overlap of points between groups in the territorial map areas.

circulating neutrophil elastase have been shown to correlate with complement activation (31). Both neutrophils and C5a (a surrogate marker of complement activation) are raised in the severe phenotype in the CASCADE study (Figures 3, 4), and they show a positive correlation in the heatmaps (Figure 8); therefore, the severity of inflammation in COVID-19 patients due to complement activation might be attenuated by the release of neutrophil elastase in COVID-19.

Significant elevations of cardiac troponin-T, correlating with an increased risk of cardiac damage, have been reported in COVID-19

(32, 33). In mouse models of myocardial ischaemia/reperfusion injury, in the absence of viral infection, C5 levels were elevated along with increases in serum troponin levels (34). Our heatmap analysis shows elevated serum complement C5 is positively correlated with raised troponin levels. This finding could be explained by observations from other studies where the high inflammatory burden from the spectrum of cytokines and complement proteins released in the disease is thought to induce myocardial injury (35). Thus, the heatmap analysis provides a unique “aerial perspective” of the “immunological landscape”,

TABLE 3 Threshold levels of biomarkers pertinent in the LDA models.

Biomarker	Threshold value
Deteriorator timepoint 1	
LDH	460 (Units/L)
IL-27	551.15 (pg/ml)
RANTES	338.13 (pg/ml)
MDC	235.44 (pg/ml)
Platelets	200 ($\times 10^9/L$)
PDGFAA	62.92 (pg/ml)
MIGCXCL9	4,453.95 (pg/ml)
Ferritin	80 (mg/L)
IP-10	781.86 (pg/ml)
PDFGABBB	6,154.8 (pg/ml)
Deteriorator timepoint 2	
IL-27	547.27 (pg/ml)
IP-10	599.45 (pg/ml)
MDC	201.52 (pg/ml)
Ferritin	80 (mg/L)
Complement C5	247 (mg/L)
MCSF	7.55 (pg/ml)
CRP	9 (mg/L)
MCP1	24.21 (pg/ml)
sC5b9	1,434.18 (ng/ml)
PDGFAA	72.06 (pg/ml)
PDFGABBB	3,867.24 (pg/ml)
Complement C3a	238.85 (ng/ml)

within the boundaries of clinical severity types, to help guide the clinician and investigative scientists, of the potential trend various markers are adopting during hospitalisation and treatment of a patient with COVID-19.

IL-33, a member of the IL-1 cytokine family, has been suggested to play an important role in severe COVID-19 (36), upregulating Th-2 cytokines such as IL-5 and IL-13 (37–39). The cytokine panel used in our study did not contain assays for IL-33; however, as IL-5 and IL-13 were surrogates of IL-33 activation, there was a significant increase in IL-5 levels above normal levels, although no concomitant increase of IL-13 was observed.

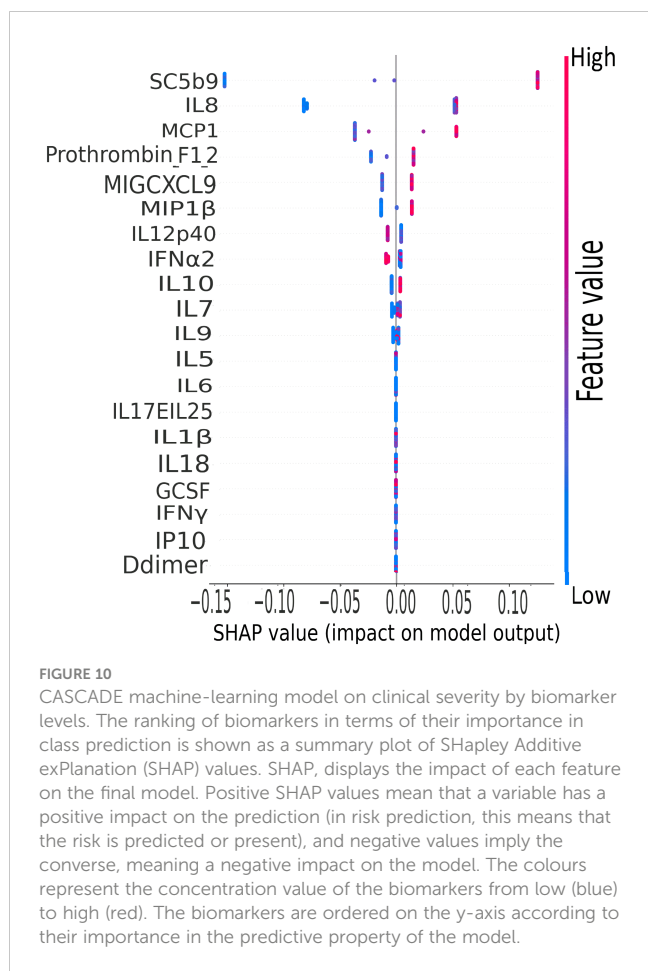
LDA model 2 (Figure 9D) identified IL-27 and IP-10 as the top two predictors of clinical deterioration in COVID-19. A decrease in the levels of IL-27 has been shown in other studies as a reliable predictor of adverse clinical outcomes in COVID-19 (40). Our heatmap analysis captures the trend of IL-27, with fewer negatively correlated pairs with IL-27 in the mild and moderate categories compared to the severe category ($12_{\text{blue pixels}}/30_{\text{total horizontal pixels}}$). IP-10 has been demonstrated in a previous study as a marker associated with clinical severity in COVID-19 and correlated with disease progression (41), and it is interesting that our model identified this as a key marker in the recognition of a patient with the propensity to deteriorate clinically.

MDC was identified as third in order of importance in our LDA model (Figure 9D). MDC has been shown to inhibit the replication of CCR5-dependent HIV in macrophages (42) and play a protective role over CD4⁺ T cells from infection by HIV (43). Our findings show that the levels of MDC are significantly lower than that found in normal, healthy individuals (Figure 3), and this is in agreement with a previous study that reported reduced levels of the chemokine in COVID-19 patients (44).

Raised ferritin levels have been observed in a range of inflammatory diseases (45), and intracellular ferritin is thought to leak into serum from damaged intracellular stores (46). Ferritin was identified as fourth, in order of importance, in our LDA model in

TABLE 4 Comparison of clinical screening tools with the LDA model.

Model or screening tool	AUC \pm SE	95% confidence interval
ROC-AUC analysis on the training data (80%)		
LDA model	1.000 \pm 0.000	1.000–1.000
ROX index	0.433 \pm 0.106	0.225–0.641
mSOFA score	0.667 \pm 0.101	0.469–0.865
NEWS2 score	0.633 \pm 0.103	0.431–0.836
ROC-AUC analysis using the holdout or test data (20%)		
LDA model	0.940 \pm 0.051	0.862–1.000
ROX index	0.192 \pm 0.109	0.000–0.406
mSOFA score	0.827 \pm 0.107	0.617–1.000
NEWS2 score	0.654 \pm 0.176	0.309–0.999



timepoint 2 (Figure 9D). The presence of excess iron in the internal milieu is known to favour the growth of numerous viruses (47); this might explain why the order of importance of ferritin in timepoint 2 (point of clinical deterioration) is moved up by rank compared to that in timepoint 1. The role of ferritin as an important biomarker in the progression of the disease in COVID-19 has been reviewed previously (48).

COVID-19 patient levels for CH50, sC5b-9, C5a, C5, C3a, C3, and factor B were all higher than levels in healthy individuals (Figure 4). Complement C5 and CH50 levels correlate with CRP levels in all severity classes, as shown in our heatmap analysis (Figures 6–8). Complement activation has been documented to positively correlate with CRP levels since CRP activates complement *via* the classical pathway, activating C1q, which then activates the rest of the complement cascade (49).

The XGBoost model highlighted that IL-8 ranked second highest as a predictive variable (Figure 10). IL-8 is a well-studied neutrophil chemotactic factor that plays a key role in numerous pathological conditions. IL-8 is expressed in neutrophils, epithelial cells, hepatocytes, fibroblasts, endothelial cells, and alveolar macrophages (50, 51). It has been proposed that the presence of IL-8 in the bronchoalveolar lavage fluid (BALF) is a useful prognostic variable in ARDS patients (50). Furthermore, IL-8 has been implicated in the recruitment of neutrophils to the lungs in acute inflammation of the lung (52). IL-8 has been associated with the development of respiratory failure following the reduction of PaO₂/FiO₂ (53), and the reduction in lung oxygenation levels across the severity classes has been demonstrated in this study by our analysis of ROX scores (Figure 2).

TABLE 5 Admission and treatment details of patients in the CORONET study treated with the anti-C5 inhibitor, nomacopan.

Patient ID.	Age bin	Sex	qPCR (+ve/-ve)	Admission (month/year)	O ₂ flow rate/min for SpO ₂ >93% [mode of delivery]	Outcome (alive/dead)	Number of days to recovery	CH ₅₀ at day 3 (U/ml)	Highest C5 (U/ml)	Highest CRP (mg/ml)	Status at discharge
								UL = 40 U/ml	UL = 10 mg/ml	UL = 10 mg/ml	
1	5	Male	+ve	September/2020	8L [NC]	Alive	11	<17.0	39.0	40.5	Normoxia on air
2	4	Male	+ve	November/2020	2L [NC]	Alive	3	22.0	40.0	199.0	Normoxia on air
3	3	Male	+ve	January/2021	6L [NC]	Alive	5	21.0	26.0	116.0	Normoxia on air
4	3	Male	+ve	February/2021	6L [IMV 4 days, then NC]	Alive	22	20.0	54.0	193.0	4L O ₂ at rest, 8–10 L on exertion
5	3	Male	+ve	April/2021	4L [NC]	Alive	5	15.0	36.0	109.0	Normoxia on air
6	5	Male	+ve	July/2020	4L [NC]	Alive	6	<10.0	22.2	154.18	Normoxia on air
7	1	Female (delayed Tx nomacopan for 3 days)	+ve	July/2020	[IMV 7 days postadmission]	Dead 14 days postadmission	N/A	>60.0 prior to nomacopan Tx	24.3	274.65	Dead 14 days postadmission

IMV, intubated and mechanically ventilated; UL, upper limit; NC, nasal cannulae; N/A, not applicable; qPCR, quantitative polymerase chain reaction; Tx, treatment. Age bins: 20–29 = 1, 30–39 = 2, 40–49 = 3, 50–59 = 4, 60–69 = 5, 70–79 = 6. The primary endpoint was full respiratory recovery, defined as nondependence on mechanical ventilation and SpO₂ of >93% by pulse oximetry, breathing air without the need for supplementary oxygen. Secondary endpoints included treatment-related adverse events, time to SpO₂ >93% with no oxygen supplementation, ventilator-free days between Day 0 and discharge, and duration from hospital admission to discharge and recovery. Treatment lasted for a maximum of 14 days.

MCP-1 was ranked third in the XGBoost model (Figure 10); this is also in agreement with the findings of other groups where MCP-1 expression levels were found to be higher in patients with COVID-19, especially those admitted to intensive care units (54). Increased levels of MCP-1 have been isolated in lung tissue of COVID-19 patients (55). It has been suggested from other studies that monitoring MCP-1 levels during hospitalisation could help prevent COVID-19 from progressing from a mild to a severe presentation (56); this finding is coherent with our XGB model of stratification of risk classes and the prominent ranking of this marker in the predictive order of importance.

The CORONET study was a “compassionate use” study, allowing off-label nomacopan treatment; nomacopan is currently in phase III studies for a coagulopathic disease, thrombotic microangiopathy, in children treated with human stem cell transplants (57). COVID-19 coagulopathy has been well studied (58), and a marker associated with coagulation (prothrombin F12) has been identified as key in our XGBoost risk stratification model (Figure 10).

Six patients received nomacopan on admission, and one patient’s treatment was delayed (for nonmedical reasons) for 3 days and died from COVID-19-related complications. The remaining six CORONET patients at admission presented with ROX scores consistent with severe hypoxaemic respiratory failure (Figure 2). All surviving patients were discharged home between 3 and 22 days after admission with high circulating levels of sC5b-9 but normal CH50 values.

The terminal complement complex exists in two forms: soluble sC5b-9 and the MAC. C3 convertase cleave C5 to C5a and C5b. C5b initiates the activation of the MAC that includes C6, C7, C8, and C9 complement components important for allowing the MAC complex to penetrate the cell membrane (59). The bound membrane attack complex forms a cytolytic pore. The complement regulators, CD46, CD55, and CD59, are important for controlling complement activation and prevention of the MAC pore assembly in the cell membrane (60).

Although pre-nomacopan treatment levels of CH50 of patients in the CORONET study were not available for comparison to the day-3 levels of CH50, the pre-nomacopan treatment C5 levels (all above upper limits) hint that complement components were high in the systemic circulation at admission (Table 3). CH50, in the CORONET study, is a useful index for identifying people at risk of deteriorating respiratory failure in COVID-19 pneumonia. Hospital-discharged nomacopan-treated CORONET patients appear to be healthy and stable without any long-term or residual sequelae. While these are preliminary results on a small-sized “compassionate-use” cohort, it warrants a randomised controlled trial of nomacopan for minimising the progression of severe lung injury in COVID-19.

One drawback of clinical risk stratification in a “resource-poor” and extremely busy pandemic setting is the dependence on senior clinical reviews, whereas automated algorithms trained on serum biomarker evaluation, could be implemented by a larger portion of the nursing workforce (61), particularly in precision nursing, with less pressures on senior clinical staff (61, 62).

Thus, the LDA model, based on biomarker profiling of COVID-19 patients, could identify patients with the propensity of deteriorating clinically. The dimensionality-reduction feature of LDA allowed a minimal set of biomarkers to be used as a prognostic guide; this is both advantageous clinically (less laboratory resources) and economically, as fewer markers mean lower costs. The XGBoost model is also useful to triage patients according to the mild, moderate, and severe categories based on biomarker profiles; such models could be easily implemented by nonspecialist staff during a pandemic.

These models were generated based on biomarker profiles from patients of the first and second COVID-19 waves, and while these bioprofiles look at the downstream translational or proteomic signature, we do not know if the proteomic signature might be affected by future variants of the SARS-CoV-2 virus.

Nevertheless, there is a need to improve the management of resources during a pandemic, and several studies have proposed exploring the use of digital health solutions and artificial intelligence in management and triage situations (63). Lung diseases carry significant mortality and morbidity worldwide, and COVID-19 is a sentinel example of the need to expand the utility of biomarker diagnosis in diagnosis and triage situations (64).

Limitations of the study

The biological lability of complement components and leukotriene B4 in blood were confounding issues for assessing patient outcomes. The average age of participants in CASCADE A and CASCADE B was in their sixth decade, and it was very difficult to obtain age-matched healthy controls in their sixth decade without significant comorbidities, particularly during a pandemic period where a large proportion of the elderly were self-isolating. We recognise the effect of age and gender on the components of the immune system and acknowledge this limitation. Our studies are underrepresented by minority ethnic groups due to the population demographics of our local hospitals. Future studies in the validation of the models and trials of nomacopan should address these issues. Future studies should be undertaken with significantly larger sample sizes.

Data availability statement

Requests for raw data pertaining to the study can be communicated to the corresponding author, they will be reviewed following agreements.

Ethics statement

The studies involving human participants were reviewed and approved by South-Central Berkshire Research Ethics Committee (REC reference: 20/SC/0228 May 22nd, 2020) and Investigational New Drug’ (IND) application approved as an “Expanded Access Application” by the FDA. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AC, LW, TB, LGD, TH, JB, CC, JM, DG, CR, WW-D: conceived the study and delineated the hypothesis and designed the study. AC, LW, TB, JB, CC, JM: made final decisions regarding case selection and inclusions. LW, LGD, TB, DG, JB, CC, JM: collected samples from patients and facilitated downstream biological processing. LGD: implemented the machine-learning and neural net aspects of the study and conducted statistical analysis of the study. LGD is a fellow of the Royal Society of Statistics. All authors read and edited the drafted manuscript for intellectual content and discussed the data interpretation. All authors approved the final version of the manuscript.

Funding

This study received funding from AKARI Ltd. TH, CR, and WW-D revised, reviewed the manuscript critically for important intellectual content, provided approval for publication of the content, confirmed on the accuracy of the interpretation, and vouched for the integrity of the final work.

Acknowledgments

The selfless and sterling efforts of the research office team, notably Elaine Baddeley, research nurses at PHU, especially John Winter and Sharon Macready, and the efforts of numerous research fellows, consultants, and care staff at PHU, UCMC, and OSUMC are acknowledged. We thank Milan Chauhan, our database manager, for his support and oversight of the data for CASCADE, Mr. David Muckelt, and Mr. James Hayward of our finance team. Our patient representative group and Sharon Court are particularly acknowledged for their support and feedback during the ethics application process and in drafting the patient information leaflets. Sharon Glaysher (PHU Laboratory Manager)

References

- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *New Engl J Med* (2020) 382:727–33. doi: 10.1056/NEJM0A2001017/SUPPL_FILE/NEJM0A2001017_DISCLOSURES.PDF
- Gattinoni L, Gattarello S, Steinberg I, Busana M, Palermo P, Lazzari S, et al. COVID-19 pneumonia: pathophysiology and management. *Eur Respir Rev* (2021) 30 (162):210138. doi: 10.1183/16000617.0138-2021
- Chauhan AJ, Wiffen LJ, Brown TP. COVID-19: a collision of complement, coagulation and inflammatory pathways. *J Thromb Haemostasis* (2020) 18:2110–7. doi: 10.1111/JTH.14981
- Iacrossi L, Fauci AJ, Napoletano A, D'angelo D, Salomone K, Latina R, et al. Triage protocol for allocation of critical health resources during the COVID-19 health emergency. *Acta Bio Medica: Atenei Parmensis* (2020) 91:1–11. doi: 10.23750/ABM.V91I4.10393
- Siggins MK, Davies K, Fellows R, Thwaites RS, Baillie JK, Semple MG, et al. Alternative pathway dysregulation in tissues drives sustained complement activation and predicts outcome across the disease course in COVID-19. *Immunology* (2023) 168 (3):473–92. doi: 10.1111/IMM.13585
- Risitano AM, Mastellos DC, Huber-Lang M, Yancopoulou D, Garlanda C, Ciceri F, et al. Complement as a target in COVID-19? *Nat Rev Immunol* (2020) 20:343–4. doi: 10.1038/S41577-020-0320-7
- Java A, Apicelli AJ, Kathryn Liszewski M, Coler-Reilly A, Atkinson JP, Kim AHJ, et al. The complement system in COVID-19: friend and foe? *JCI Insight* (2020) 5(15): e140711. doi: 10.1172/JCI.INSIGHT.140711
- Malaquias MAS, Gadotti AC, Motta-Junior J da S, Martins APC, Azevedo MLV, Benevides APK, et al. The role of the lectin pathway of the complement system in SARS-CoV-2 lung injury. *Transl Res* (2021) 231:55–63. doi: 10.1016/J.TRSL.2020.11.008
- Ali YM, Ferrari M, Lynch NJ, Yaseen S, Dudler T, Gragerov S, et al. Lectin pathway mediates complement activation by SARS-coV-2 proteins. *Front Immunol* (2021) 12:714511. doi: 10.3389/FIMMU.2021.714511
- Yu J, Yuan X, Chen H, Chaturvedi S, Braunstein EM, Brodsky RA. Direct activation of the alternative complement pathway by SARS-CoV-2 spike proteins is blocked by factor D inhibition. *Blood* (2020) 136:2080–9. doi: 10.1182/BLOOD.2020008248

and her team assisted in collecting and storing samples for downstream analysis. We acknowledge bioinformatics advice and statistical audit from Dr. Samuel Robson (Portsmouth University), a fellow of the Royal Society of Statistics. We are indebted to AKARI PLC for supplying the drug nomacopan for use in the CORONET trial. The clinical and laboratory staff at UCMC and OSUMC are acknowledged for their kind support and professionalism. A research grant from Akari Therapeutics made the collection of patient data, clinical samples, and biological analyses possible. The sponsors of the CASCADE project (Portsmouth Hospitals University NHS Trust) supplied the machine learning and statistical analysis expertise for both studies. Finally, we thank our patients for making both studies possible.

Conflict of interest

Authors TH, CR, and WW-D are employed by AKARI Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1192765/full#supplementary-material>

11. Schols S, Nunn MA, Mackie I, Weston-Davies W, ichi NJ, Kanakura Y, et al. Successful treatment of a PNH patient non-responsive to eculizumab with the novel complement C5 inhibitor coversin (nomacopan). *Br J Haematol* (2020) 188:334–7. doi: 10.1111/BJH.16305
12. Sadik CD, Rashid H, Hammers CM, Diercks GFH, Weidinger A, Beissert S, et al. Evaluation of nomacopan for treatment of bullous pemphigoid: A phase 2a nonrandomized controlled trial. *JAMA Dermatol* (2022) 158:641–9. doi: 10.1001/JAMADERMATOL.2022.1156
13. The CASCADE study - Measures of Complement Activation and Inflammation in Patients With Covid-19 - Full Text View - ClinicalTrials.gov. Available at: <https://clinicaltrials.gov/ct2/show/NCT04453527> (Accessed December 16, 2021).
14. Jodele S, Dandoy CE, Danziger-Isakov L, Myers KC, El-Bietar J, Nelson A, et al. Terminal Complement Blockade after Hematopoietic Stem Cell Transplantation Is Safe without Meningococcal Vaccination. *Biol Blood Marrow Transplant* (2016) 22:1337. doi: 10.1016/j.bbmt.2016.03.032
15. Costabile M. Measuring the 50% haemolytic complement () activity of serum. *J Vis Exp* (2010) 37:1923. doi: 10.3791/1923
16. For Physicians: A Guide to Non-emergency Single Patient Expanded Access Submissions | FDA. Available at: <https://www.fda.gov/drugs/investigational-new-drug-ind-application/physicians-guide-non-emergency-single-patient-expanded-access-submissions> (Accessed October 20, 2021).
17. LabCorp COVID-19 RT-PCR test EUA Summary ACCELERATED EMERGENCY USE AUTHORIZATION (EUA) SUMMARY COVID-19 RT-PCR TEST (LABORATORY CORPORATION OF AMERICA). Available at: <https://www.fda.gov/media/136151/download> (Accessed March 23, 2023).
18. 3.2. Tuning the hyper-parameters of an estimator — scikit-learn 1.2.2 documentation. Available at: https://scikit-learn.org/stable/modules/grid_search.html (Accessed March 22, 2023).
19. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* (2002) 16:321–57. doi: 10.1613/JAIR.953
20. Huang Z, Fu Z, Huang W, Huang K. Prognostic value of neutrophil-to-lymphocyte ratio in sepsis: A meta-analysis. *Am J Emerg Med* (2020) 38:641–7. doi: 10.1016/j.ajem.2019.10.023
21. Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biom J* (2008) 50:419. doi: 10.1002/BIMJ.200710415
22. Youden WJ. Index for rating diagnostic tests. *Cancer* (1950) 3(1):32–5. doi: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3
23. Barnum SR, Bubeck D, Schein TN. Soluble membrane attack complex: biochemistry and immunobiology. *Front Immunol* (2020) 11:585108. doi: 10.3389/FIMMU.2020.585108
24. Yufei Y, Mingli L, Xuejiao L, Xuemei D, Yiming J, Qin Q, et al. Utility of the neutrophil-to-lymphocyte ratio and C-reactive protein level for coronavirus disease 2019 (COVID-19). *Scand J Clin Lab Invest* (2020) 80:1. doi: 10.1080/00365513.2020.1803587
25. Ma Y, Zhang Y, Zhu L. Role of neutrophils in acute viral infection. *Immun Inflammation Dis* (2021) 9:1186. doi: 10.1002/IID3.500
26. Antunes G, Evans SA, Lordan JL, Frew AJ. Systemic cytokine levels in community-acquired pneumonia and their association with disease severity. *Eur Respir J* (2002) 20:990–5. doi: 10.1183/09031936.02.00295102
27. Guo J, Wang S, Xia H, Shi D, Chen Y, Zheng S, et al. Cytokine signature associated with disease severity in COVID-19. *Front Immunol* (2021) 12:681516/FULL. doi: 10.3389/FIMMU.2021.681516/FULL
28. Tsujimoto H, Ono S, Majima T, Kawarabayashi N, Takayama E, Kinoshita M, et al. Neutrophil elastase, MIP-2, and TLR-4 expression during human and experimental sepsis. *Shock* (2005) 23:39–44. doi: 10.1097/01.SHK.0000145936.31967.D7
29. Pan T, Tuoxun T, Chen X, Yang CJ, Jiang CY, Zhu YF, et al. The neutrophil elastase inhibitor, sivelestat, attenuates acute lung injury in patients with cardiopulmonary bypass. *Front Immunol* (2023) 14:1082830. doi: 10.3389/FIMMU.2023.1082830
30. Sahebnaasagh A, Saghafi F, Safdari M, Khataminia M, Sadremomtaz A, Talaei Z, et al. Neutrophil elastase inhibitor (sivelestat) may be a promising therapeutic option for management of acute lung injury/acute respiratory distress syndrome or disseminated intravascular coagulation in COVID-19. *J Clin Pharm Ther* (2020) 45:1515–9. doi: 10.1111/JCPT.13251
31. Gardinali M, Padalino P, Vesconi S, Calcagno A, Ciappellano S, Conciato L, et al. Complement activation and polymorphonuclear neutrophil leukocyte elastase in sepsis. Correlation with severity of disease. *Arch Surg* (1992) 127:1219–24. doi: 10.1001/ARCHSURG.1992.01420100077014
32. Katsoularis I, Fonseca-Rodriguez O, Farrington P, Lindmark K, Fors Connolly AM. Risk of acute myocardial infarction and ischaemic stroke following COVID-19 in Sweden: a self-controlled case series and matched cohort study. *Lancet* (2021) 398:599–607. doi: 10.1016/S0140-6736(21)00896-5
33. Wei JF, Huang FY, Xiong TY, Liu Q, Chen H, Wang H, et al. Acute myocardial injury is common in patients with COVID-19 and impairs their prognosis. *Heart* (2020) 106:1154–9. doi: 10.1136/HEARTJNL-2020-317007
34. Busche MN, Stahl GL. Role of the complement components C5 and C3a in a mouse model of myocardial ischemia and reperfusion injury. *GMS German Med Sci* (2010) 8:8. doi: 10.3205/000109
35. Chang WT, Toh HS, Liao C, Yu WL. Cardiac involvement of COVID-19: A comprehensive review. *Am J Med Sci* (2021) 361:14–22. doi: 10.1016/J.AMJMS.2020.10.002
36. Furci F, Murdaca G, Allegra A, Gammeri L, Senna G, Gangemi S. IL-33 and the cytokine storm in COVID-19: from a potential immunological relationship towards precision medicine. *Int J Mol Sci* (2022) 23(23):14532. doi: 10.3390/IJMS232314532
37. Smith DE. IL-33: a tissue derived cytokine pathway involved in allergic inflammation and asthma. *Clin Exp Allergy* (2010) 40:200–8. doi: 10.1111/J.1365-2222.2009.03384.X
38. van de Veen W, Akdis M. The use of biologics for immune modulation in allergic disease. *J Clin Invest* (2019) 129:1452. doi: 10.1172/JCI124607
39. Byers DE, Alexander-Brett J, Patel AC, Agapov E, Dang-Vu G, Jin X, et al. Long-term IL-33-producing epithelial progenitor cells in chronic obstructive lung disease. *J Clin Invest* (2013) 123:3967. doi: 10.1172/JCI65570
40. Tamayo-Velasco Á, Martínez-Paz P, Peñarubia-Ponce MJ, de la Fuente I, Pérez-González S, Fernández I, et al. HGF, IL-1 α , and IL-27 are robust biomarkers in early severity stratification of COVID-19 patients. *J Clin Med* (2021) 10(9):2017. doi: 10.3390/JCM10092017/S1
41. Yang Y, Shen C, Li J, Yuan J, Wei J, Huang F, et al. Plasma IP-10 and MCP-3 levels are highly associated with disease severity and predict the progression of COVID-19. *J Allergy Clin Immunol* (2020) 146:119. doi: 10.1016/J.JACI.2020.04.027
42. Cota M, Mengozzi M, Vicenzi E, Panina-Bordignon P, Sinigaglia F, Transidico P, et al. Selective inhibition of HIV replication in primary macrophages but not T lymphocytes by macrophage-derived chemokine. *Proc Natl Acad Sci U.S.A.* (2000) 97:9162–7. doi: 10.1073/PNAS.160359197
43. Cocchi F, DeVico AL, Lu W, Popovic M, Latinovic O, Sajadi MM, et al. Soluble factors from T cells inhibiting X4 strains of HIV are a mixture of β chemokines and RNases. *Proc Natl Acad Sci U.S.A.* (2012) 109:5411–6. doi: 10.1073/PNAS.1202240109/-DCSUPPLEMENTAL
44. Tufa A, Gebremariam TH, Manyazewal T, Getinet T, Webb DL, Hellström PM, et al. Inflammatory mediators profile in patients hospitalized with COVID-19: A comparative study. *Front Immunol* (2022) 13:964179. doi: 10.3389/FIMMU.2022.964179
45. Kernan KF, Carcillo JA. Hyperferritinemia and inflammation. *Int Immunol* (2017) 29:401–9. doi: 10.1093/INTIMM/DXX031
46. Edeas M, Saleh J, Peyssonnaud C. Iron: Innocent bystander or vicious culprit in COVID-19 pathogenesis? *Int J Infect Dis* (2020) 97:303–5. doi: 10.1016/J.IJID.2020.05.110
47. Drakesmith N, Prentice A. Viral infection and iron metabolism. *Nat Rev Microbiol* (2008) 6:541–52. doi: 10.1038/NRMICRO1930
48. Kaushal K, Kaur H, Sarma P, Bhattacharyya A, Sharma DJ, Prajapat M, et al. Serum ferritin as a predictive biomarker in COVID-19: A systematic review, meta-analysis and meta-regression analysis. *J Crit Care* (2022) 67:172. doi: 10.1016/J.JCRC.2021.09.023
49. Mosquera-Sulbaran JA, Pedrañez A, Carrero Y, Callejas D. C-reactive protein as an effector molecule in Covid-19 pathogenesis. *Rev Med Virol* (2021) 31(6):e2221. doi: 10.1002/RMV.2221
50. Cesta MC, Zippoli M, Marsiglia C, Gavioli EM, Mantelli F, Allegretti M, et al. The role of interleukin-8 in lung inflammation and injury: implications for the management of COVID-19 and hyperinflammatory acute respiratory distress syndrome. *Front Pharmacol* (2022) 12:808797/BIBTEX. doi: 10.3389/FPHAR.2021.808797/BIBTEX
51. Qazi BS, Tang K, Qazi A. Recent advances in underlying pathologies provide insight into interleukin-8 expression-mediated inflammation and angiogenesis. *Int J Inflamm* (2011) 2011:1–13. doi: 10.4061/2011/908468
52. Ronit A, Berg RMG, Bay JT, Haugaard AK, Ahlström MG, Burgdorf KS, et al. Compartmental immunophenotyping in COVID-19 ARDS: A case series. *J Allergy Clin Immunol* (2021) 147:81–91. doi: 10.1016/J.JACI.2020.09.009
53. Bülow Anderberg S, Luther T, Berglund M, Larsson R, Rubertsson S, Lipscey M, et al. Increased levels of plasma cytokines and correlations to organ failure and 30-day mortality in critically ill Covid-19 patients. *Cytokine* (2021) 138:155389. doi: 10.1016/J.CYTO.2020.155389
54. 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. doi: 10.1016/S0140-6736(20)30183-5
55. Chu H, Chan JFW, Wang Y, Yuen TTT, Chai Y, Hou Y, et al. Comparative replication and immune activation profiles of SARS-coV-2 and SARS-coV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19. *Clin Infect Dis* (2020) 71:1400–9. doi: 10.1093/CID/CIAA410
56. Xi X, Guo Y, Zhu M, Wei Y, Li G, Du B, et al. Higher expression of monocyte chemoattractant protein 1 in mild COVID-19 patients might be correlated with inhibition of Type I IFN signaling. *Viral J* (2021) 18:12. doi: 10.1186/S12985-020-01478-9
57. Fettiplace J. Multicentre Study of nomacopan in Paediatric Haematopoietic Stem-Cell Transplant Associated Thrombotic Microangiopathy. Clinical Trials Register, EudraCT number: 2020-000086-17. Akari Therapeutics PLC. *Clin Trials Register* (2020). Available at: <https://classic.clinicaltrials.gov/ct2/show/NCT04784455>. (Accessed August 24, 2023).
58. Conway EM, Mackman N, Warren RQ, Wolberg AS, Mosnier LO, Campbell RA, et al. Understanding COVID-19-associated coagulopathy. *Nat Rev Immunol* 2022 22:10 (2022) 22:639–49. doi: 10.1038/s41577-022-00762-9

59. Menny A, Lukassen M v., Couves EC, Franc V, Heck AJR, Bubeck D. Structural basis of soluble membrane attack complex packaging for clearance. *Nat Commun* (2021) 12(1):6086. doi: 10.1038/S41467-021-26366-W
60. Xie CB, Jane-Wit D, Pober JS. Complement membrane attack complex: new roles, mechanisms of action, and therapeutic targets. *Am J Pathol* (2020) 190:1138–50. doi: 10.1016/J.AJPAT.2020.02.006
61. Ielapi N, Andreucci M, Licastro N, Faga T, Grande R, Buffone G, et al. Precision medicine and precision nursing: the era of biomarkers and precision health. *Int J Gen Med* (2020) 13:1705. doi: 10.2147/ijgm.S285262
62. Corwin EJ, Ferranti EP. Integration of biomarkers to advance precision nursing interventions for family research across the life span. *Nurs Outlook* (2016) 64:292–8. doi: 10.1016/J.outlook.2016.04.007
63. Gunasekeran DV, Tseng RMWW, Tham YC, Wong TY. Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. *NPJ Digit Med* (2021) 4(1):40. doi: 10.1038/S41746-021-00412-9
64. Viswanathan VS, Toro P, Corredor G, Mukhopadhyay S, Madabhushi A. The state of the art for artificial intelligence in lung digital pathology. *J Pathol* (2022) 257:413–29. doi: 10.1002/path.5966