SARS-CoV-2 RNAemia and proteomic biomarker trajectory inform prognostication in COVID-19 patients admitted to intensive care

Manuel Mayr (✉ manuel.mayr@kcl.ac.uk)
King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London https://orcid.org/0000-0002-0597-829X

Clemens Gutmann
King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London https://orcid.org/0000-0003-0675-8632

Kaloyan Takov
King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London https://orcid.org/0000-0002-8642-6306

Sean Burnap
King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London https://orcid.org/0000-0002-3408-8608

Bhawana Singh
King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London

Konstantinos Theofilatos
King's College https://orcid.org/0000-0001-6799-0553

Ella Reed
King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London

Maria Hasman
King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London

Adam Nabeebaccus
King's College Hospital NHS Foundation Trust, London

Matthew Fish
King's College London https://orcid.org/0000-0001-6462-3889

Mark McPhail
Department of Inflammation Biology, School of Immunology and Microbial Sciences, Faculty of Life Sciences and Medicine, King's College London, London

Kevin O'Gallagher

King's College Hospital NHS Foundation Trust, London

Lukas Schmidt

King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London

Christian Cassel

King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London

Marieke Rienks

King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London https://orcid.org/0000-0002-0590-9518

Xiaoke Yin

King's College London https://orcid.org/0000-0002-5172-0935

Georg Auzinger

King's College Hospital NHS Foundation Trust, London

Salvatore Napoli

Department of Inflammation Biology, School of Immunology and Microbial Sciences, Faculty of Life Sciences and Medicine, King's College London, London

Salma Mujib

Institute of Liver Studies, King's College Hospital, London

Francesca Trovato

King's College Hospital NHS Foundation Trust, London

Barnaby Sanderson

Department of Intensive Care Medicine, Guy's and St Thomas' NHS Foundation Trust, London https://orcid.org/0000-0002-9621-143X

Blair Merrick

Guy's and St Thomas' NHS Foundation Trust https://orcid.org/0000-0002-6061-6064

Umar Niazi

King's College London https://orcid.org/0000-0001-7176-8883

Mansoor Saqi

NIHR Biomedical Research Centre, Guy's and St Thomas' NHS Foundation Trust

Konstantina Dimitrakopoulou

NIHR Biomedical Research Centre, Guy's and St Thomas' NHS Foundation Trust and King's College London, London

Silke Braun

Medical Clinic I, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden

Romy Kronstein-Wiedemann

Medical Faculty Carl Gustav Carus, Technische Universität Dresden

Katie Doores

King's College London https://orcid.org/0000-0002-5507-1725

Jonathan Edgeworth
5th Floor North Wing, St Thomas' Hospital, London, UK
Ajay Shah
King's College London
Stefan Bornstein
Medical Faculty Carl Gustav Carus at the Technical University of Dresden
Torsten Tonn
Transfusion Medicine, Medical Faculty Carl Gustav Carus, Technische Universität Dresden
https://orcid.org/0000-0001-9580-2193
Adrian Hayday
King's College London https://orcid.org/0000-0002-9495-5793
Manu Shankar-Hari
King's College London https://orcid.org/0000-0002-5338-2538

Article

Keywords: COVID-19, SARS-CoV-2, RNAemia, Pentraxin-3, Biomarker, Proteomics

Posted Date: November 18th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-101592/v1

Version of Record: A version of this preprint was published at Nature Communications on June 7th, 2021.
See the published version at https://doi.org/10.1038/s41467-021-23494-1.

SARS-CoV-2 RNAemia and proteomic biomarker trajectory inform prognostication in COVID-19 patients admitted to intensive care

Clemens Gutmann[1*], Kaloyan Takov[1*], Sean A. Burnap[1*], Bhawana Singh[1*], Konstantinos Theofilatos[1], Ella Reed[1], Maria Hasman[1], Adam Nabeebaccus[1,2], Matthew Fish,[3,4] Mark J.W. McPhail[2,5,6], Kevin O'Gallagher[1,2], Lukas E. Schmidt[1], Christian Cassel[1], Marieke Rienks[1], Xiaoke Yin[1], Georg Auzinger[2], Salvatore Napoli[5], Salma F. Mujib[6], Francesca Trovato[2,5,6], Barnaby Sanderson[4], Blair Merrick[7], Umar Niazi[8], Mansoor Saqi[8], Konstantina Dimitrakopoulou[8], Silke Braun[9], Romy Kronstein-Wiedemann[10], Katie J. Doores[3], Jonathan D. Edgeworth[3,7], Ajay M. Shah[1], Stefan R. Bornstein[11,12], Torsten Tonn[10,13], Adrian C. Hayday[3,14], Manu Shankar-Hari[3,4,15], Manuel Mayr[1,11,15].

[1] King's College London British Heart Foundation Centre, School of Cardiovascular Medicine and Sciences, London, UK.

[2] King's College Hospital NHS Foundation Trust, London, UK.

[3] Department of Infectious Diseases, School of Immunology and Microbial Sciences, King's College London, London, UK.

[4] Department of Intensive Care Medicine, Guy's and St Thomas' NHS Foundation Trust, London, UK.

[5] Department of Inflammation Biology, School of Immunology and Microbial Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK.

[6] Institute of Liver Studies, King's College Hospital, London, UK

[7] Clinical Infection and Diagnostics Research group, Department of Infection, Guy's and St Thomas' NHS Foundation Trust, London, UK

[8] NIHR Biomedical Research Centre, Guy's and St Thomas' NHS Foundation Trust and King's College London, London, UK

[9] Medical Clinic I, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, Germany

[10] Experimental Transfusion Medicine, Faculty of Medicine Carl Gustav Carus, Technical University Dresden, Dresden, Germany

[11] Department of Internal Medicine III, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, Germany

[12] Department of Diabetes, School of Life Course Science and Medicine, King's College London, London, UK

[13] Institute for Transfusion Medicine, German Red Cross Blood Donation Service North East, Dresden, Germany

[14] The Francis Crick Institute, London, UK

* Equal first author contribution.

Running title: SARS-CoV-2 RNAemia and Proteomics

[15] **To whom correspondence should be addressed:**

Dr. Manu Shankar-Hari; manu.shankar-hari@kcl.ac.uk or Prof. Manuel Mayr; King's British Heart Foundation Centre, King's College London, 125 Coldharbour Lane, London, SE59NU UK; Phone: +44(0)2078485446; +44(0)2078485298; manuel.mayr@kcl.ac.uk.

Abstract

Prognostic characteristics inform risk stratification in intensive care unit (ICU) patients with coronavirus disease 2019 (COVID-19). We obtained blood samples (*n* = 474) from hospitalized COVID-19 patients (*n* = 123), non-COVID-19 ICU sepsis patients (*n* = 25) and healthy controls (*n* = 30). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA was detected in plasma or serum (RNAemia) of COVID-19 ICU patients when neutralizing antibody response was low. RNAemia was associated with higher 28-day ICU mortality (hazard ratio [HR], 1.84 [95% CI, 1.22–2.77] adjusted for age and sex). In longitudinal comparisons, COVID-19 ICU patients had a distinct proteomic trajectory associated with RNAemia and mortality. Among COVID-19-enriched proteins, galectin-3 binding protein (LGALS3BP) and proteins of the complement system were identified as interaction partners of SARS-CoV-2 spike glycoprotein. Finally, machine learning identified 'Age, RNAemia' and 'Age, pentraxin-3 (PTX3)' as the best binary signatures associated with 28-day ICU mortality.

Key Words: COVID-19 • SARS-CoV-2 • RNAemia • Pentraxin-3 • Biomarker • Proteomics

Introduction

Coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; a single-stranded RNA virus) poses an unprecedented challenge to health care systems globally. It is increasingly apparent that conventional prognostic scores for critically ill patients admitted to intensive care units (ICUs) such as the APACHE II (Acute Physiology and Chronic Health Evaluation) score[1] and SOFA (Sequential Organ Failure Assessment) score[2], are not discriminatory in COVID-19 ICU patients[3–6].

In this context, circulating SARS-CoV-2 RNA (RNAemia) has been highlighted as a promising prognostic biomarker in hospitalized COVID-19 patients, as it is associated with disease severity[7] and mortality[8–10], with an estimated prevalence of 10% (95% CI 5-18%, random effects model)[7]. Further, we hypothesized that the acute and profound alterations in the innate and adaptive immune system in COVID-19 patients[3,11–13], especially in RNAemic patients[14–18], will be accompanied by marked changes in the circulating proteome and interactome and that the proteome in COVID-19 patients will highlight mechanistically relevant signatures and trajectories, when compared to non-COVID-19 sepsis and healthy controls. Thus far, proteomics studies have focused on the determination of protein biomarkers of COVID-19 severity[19–22], but have not assessed the longitudinal relationship between proteomic changes, RNAemia and 28-day mortality.

In this study, we assessed RNAemia, antibody response against SARS-CoV-2 and proteomic profiles in serial blood samples from COVID-19 patients admitted to two ICUs. Controls included hospitalized, non-ICU COVID-19 patients as well as SARS-CoV-2-negative ICU sepsis and non-ICU patients. In the context of RNAemia, we explored the plasma protein interactions with the SARS-CoV-2 spike glycoprotein. Finally, we compared the associations of RNAemia and protein biomarkers with 28-day mortality, including established biomarkers

of acute respiratory distress syndrome (ARDS), *i.e.* receptor for advanced glycation end-products (RAGE)[23–25], and prognosis in ICU patients with sepsis, *i.e.* pentraxin-3 (PTX3)[26–29].

Results

Demographics and clinical characteristics of COVID-19 patients. 474 blood samples were available for analysis (Fig. 1, Supplementary Fig. 1): 295 longitudinal samples from ICU patients with COVID-19 admitted to two university hospitals (GSTT; $n = 62$ and KCH; $n = 16$) and samples from hospitalized, non-ICU COVID-19 patients for comparison ($n = 45$); ICU and non-ICU patients without COVID-19 served as controls ($n = 55$). The baseline clinical characteristics of all COVID-19 ICU patients are shown in Supplementary Table 1. The primary outcome measure was defined as mortality 28 days after ICU admission. As expected[30], non-survivors (23%) were older than survivors ($P = 0.0004$). COVID-19 patients admitted to ICU were predominantly males (72%). All other characteristics, including common comorbidities, the time from symptom onset to ICU admission, APACHE II score and SOFA score, were similar between ICU survivors and non-survivors. The mortality rate in COVID-19 ICU patients was twice as high as in hospitalized, non-ICU COVID-19 patients (23% *versus* 11%; Supplementary Table 2).

Frequency of SARS-CoV-2 RNAemia and association with mortality in COVID-19 ICU patients. The presence of circulating viral RNA was analyzed by RT-qPCR. Serum (GSTT; $n = 62$) and plasma (KCH; $n = 16$) samples were collected within 24 hours of admission to ICU with COVID-19 and thereafter during week 1, week 2 and again before discharge. Since all 78 COVID-19 ICU patients were administered heparin and heparin has an inhibitory effect on qPCR[31,32], RNA samples were treated with heparinase as previously described[33]. 18 of 78 (23%) COVID-19 ICU patients had detectable RNAemia within the first six days upon

admission to ICU (Supplementary Table 1). Strikingly, RNAemia within six days of admission to ICU was detectable in 56% of non-survivors but only in 13% of survivors (P = 0.0006, Supplementary Table 1). RNAemia was associated with a higher risk of 28-day mortality (hazard ratio [HR], 2.05 [95% CI, 1.38–3.04]), that was comparable to age (2.89 [1.66-5.03], Fig. 2a) and maintained after correction for age and sex (HR, 1.84 [95% CI, 1.22–2.77], Fig. 2b). In comparison, only 2 out of 45 (4%) non-ICU COVID-19 patients tested positive for RNAemia upon hospitalization (Supplementary Table 2). General demographics and baseline clinical characteristics of COVID-19 patients with and without RNAemia in the first six days of admission to ICU are presented in Supplementary Table 3. Hypertension (r = 0.33, P = 0.003), type 2 diabetes (r = 0.24, P = 0.038), bilirubin (r = 0.32, P = 0.005), respiration rate (r = 0.27, P = 0.018) and elevated potassium levels (r = 0.26, P = 0.023) were positively correlated to RNAemia, whilst monocyte counts were inversely correlated (r = -0.23, P = 0.047, Fig. 2c). A hierarchical clustering analysis of all clinical variables and RNAemia is presented in Supplementary Fig. 2. To confirm the specificity of our RT-qPCR assay, we measured SARS-CoV-2 RNAemia in 134 plasma samples from 55 non-COVID-19 patients, all of which tested negative (Supplementary Table 4 and 5).

Humoral immune response during SARS-CoV-2 RNAemia. In both COVID-19 ICU patient cohorts, IgG antibodies to the SARS-CoV-2 spike S1 glycoprotein and SARS-CoV-2 neutralizing capacity were measured by ELISA and Surrogate Virus Neutralization Test, respectively. The latter test evaluates the inhibition of binding of the receptor-binding domain (RBD) of SARS-CoV-2 spike to ACE2. For validation, neutralization potency was correlated to a HIV-1 based pseudotype neutralization assay in a subset of samples (38 samples from 16 ICU patients, r = 0.81, P <0.0001). COVID-19 ICU patients who tested positive or negative for RNAemia within the first six days in ICU showed no difference in their strong IgG response

138 to SARS-CoV-2 S1 or in their neutralization capacity (Fig. 2d). However, when individual
139 samples were compared, RNAemia positive samples had lower anti-SARS-CoV-2 spike IgG
140 levels and lower SARS-CoV-2 neutralization capacity (Fig. 2e).
141
142 **Plasma proteome alterations in COVID-19 ICU patients.** To capture the host response of
143 COVID-19 ICU patients, we interrogated their plasma proteome. Baseline plasma samples
144 from COVID-19 ICU patients (KCH cohort, n = 12) were compared to COVID-19 negative
145 sepsis ICU patients (sepsis, n = 12) and patients prior to undergoing elective cardiac surgery
146 (controls, n = 30) (Supplementary Table 4 and 5). The plasma proteome was quantified by a
147 data-independent acquisition-mass spectrometry (DIA-MS) approach, using authentic heavy
148 peptide standards representing 500 proteins[34], revealing 100 significantly altered proteins
149 across the three patient groups ($q<0.05$) (Fig. 3a). Hierarchical cluster analysis highlighted a
150 cluster of 47 plasma proteins enriched in COVID-19, including members of the complement
151 cascade, as well as proteins involved in platelet degranulation, the acute phase response and
152 coagulation (Fig. 3a, b).
153 Of the 100 circulating proteins altered across control, sepsis ICU and COVID-19 ICU
154 patients, 29 overlapped with previous proteomic reports identifying markers of COVID-19
155 severity[19,20] (Supplementary Fig. 3). However, only few were associated with 28-day mortality,
156 as determined through DIA-MS analysis of baseline serum samples obtained from a larger
157 COVID-19 ICU patient cohort (GSTT, n = 62) (Fig. 3c). Complement factor B (CFB),
158 carboxypeptidase N (CPN1) and alpha-1-antichymotrypsin (SERPINA3) were all negatively
159 associated with outcome. An independent, publicly available dataset utilizing proximity-
160 extension assays (Olink, n = 264 survivors, n = 42 non-survivors, Supplementary Table 6) also
161 confirmed the lack of outcome association for three proteins identified as markers of COVID-

19 severity in previous proteomics studies[19,20]: lipopolysaccharide binding protein (LBP), CD14, and inter-alpha-trypsin inhibitor heavy chain H3 (ITIH3) (Fig. 3c).

Protein changes that emerged as significantly associated with mortality in ICU patients but have not been previously linked to the severity of COVID-19, included an elevation of mannose binding lectin 2 (MBL2) and reductions in protein C (PROC), plasminogen (PLG) and coagulation factor 7 (F7) (Fig. 3d). The associations of PROC and F7 with 28 days mortality and the directionality of these associations were validated in the external validation cohort mentioned above (Supplementary Table 6).

Protein associations with SARS-CoV-2 RNAemia and clinical improvement. Nine proteins were significantly associated with RNAemia at baseline (GSTT COVID-19 ICU cohort) which included an increase in plasma protease C1 inhibitor (SERPING1) and complement C4-A (C4A); paralleled by a reduction in VE-cadherin (CDH5) and complement factor H-related protein 1 (CFHR1) (Fig. 4a). In longitudinal serum samples from the GSTT cohort (baseline, week 1 and week 2; $n = 47$), a greater increase of polymeric immunoglobulin receptor (PIGR) was observed in RNAemia positive, compared to RNAemia negative ICU patients (Fig. 4b). In contrast, kallikrein (KLKB1) levels significantly increased over time but tended to be higher in RNAemia negative ICU patients (Fig. 4b).

Hierarchical cluster analysis upon significantly changing serum proteins over the two-week period (baseline, week 1 and week 2) revealed four distinct protein clusters (Fig. 4c), which were annotated by gene ontology enrichment analysis. Alterations in PIGR correlated closely with neutrophil degranulation proteins such as S100A8 and S100A9 (Fig. 4c, Cluster 2), while KLKB1 kinetics followed members of the coagulation system (Fig. 4c, Cluster 4). A comparison of the trajectories of individual proteins between patients who survived and died is shown in Supplementary Fig. 4. The most pronounced changes were observed among

proteins constituting cluster 4, with recovery of liver-derived proteins linked to lipid metabolism and coagulation being significantly suppressed in patients who died (Fig. 4c).

LGALS3BP is enriched in COVID-19 and binds to SARS-CoV-2 spike glycoprotein. The spike glycoprotein is the largest protein in the viral envelope, responsible for cell entry and is the main target of neutralizing antibodies[35]. A magnetic affinity pull-down of a His-tagged SARS-CoV-2 spike glycoprotein mixed with plasma from COVID-19 ICU patients was coupled with proteomics to determine interaction partners. Proteomic analysis identified 32 spike-binding proteins. A large proportion were immunoglobulins (Fig. 5a) and members of the complement system, which are known to directly interact with antigen-bound antibodies (*i.e.* C1 complement complex, Fig. 5b, Supplementary Table 7). Additional interaction partners included complement component 4 binding proteins alpha and beta (C4BPA and C4BPB), CPN1 (among the proteins associated with 28-day mortality) and galectin-3-binding protein (LGALS3BP). Apart from apolipoprotein D (APOD), LGALS3BP was the only protein to be retrieved to a greater extent with spike glycoprotein from plasma of COVID-19 ICU patients compared to pre-pandemic sepsis ICU patients (Fig. 5c, Supplementary Table 8).

LGALS3BP was markedly elevated in COVID-19 patients as discovered by DIA-MS and confirmed by ELISA, but unchanged between control and sepsis patients without COVID-19 (Fig. 5d). Strikingly, LGALS3BP was among the most elevated proteins when compared to sepsis ICU patients (Fig. 5e). Of the proteins revealed to bind spike, only LGALS3BP and members of the complement cascade were also specifically elevated in COVID-19 ICU patients. LGALS3BP revealed a strong positive correlation with members of the complement cascade (C6, C9, C4BPA and C4BPB) and CPN1, but a negative correlation with adiponectin (ADIPOQ) (Fig. 5f). LGALS3BP abundance in COVID-19 patients closely correlated with

regulators of the complement cascade, platelet degranulation and the innate immune system (Fig. 5f, Supplementary Fig. 5).

SARS-CoV-2 mortality prediction using machine learning. RAGE is an established biomarker of ARDS[23–25], but remained unaffected by SARS-CoV-2 RNAemia and mortality (Supplementary Fig. 6a). PTX3, however, a protein we and others have previously highlighted as a prognostic marker in ICU patients with sepsis[26–29], positively associated with COVID-19 mortality (Supplementary Fig. 6b). Notably, PTX3 emerged as one of the best predictors for mortality among 1,526 proteins measured in the external validation cohort of hospitalized COVID-19 patients described above (n = 264 survivors; n = 42 non-survivors), outperforming all measured cytokines and chemokines (Supplementary Table 6). Thus, a machine learning-based approach was adopted to determine the best binary combination of clinical variables, RNAemia and protein biomarkers that are independently associated with 28-day COVID-19 mortality. Kaplan Meier plots highlight RNAemia (P <0.0001) as the best individual predictor (Fig. 6a-c, Supplementary Table 9), while the binary combinations 'Age, RNAemia' (P <0.0001) and 'Age, PTX3' (P <0.0001), improved sensitivity compared to single markers, and provided better survival stratification (Fig. 6d-f, Supplementary Table 9).

Discussion

To the best of our knowledge, this is the largest longitudinal assessment of RNAemia, humoral immune response against SARS-CoV-2, protein biomarkers and clinical variables in COVID-19 ICU patients to date. SARS-CoV-2 RNAemia was observed in 23% of COVID-19 ICU patients within the first six days of admission to ICU, which is more frequent than its estimated prevalence (10% [95% CI 5-18%], random effects model)[7]. Likely explanations include the fact that RNAemia is expected to be more common in ICU patients due to disease

severity[7]. Second, we optimized detection by treating isolated RNA with heparinase[33] to overcome the known inhibitory effect of heparin on qPCR[31,32]. We also performed a two-step RT-qPCR protocol rather than the one-step RT-qPCR protocol used in clinical practice and previous studies in which RNAemia has been assessed thus far. Third, RNAemia was more frequent closer to the onset of symptoms[7] and when humoral response against SARS-CoV-2 was low. The latter observation was maintained after correcting for time since onset of symptoms. Thus, this is not a mere reflection of low humoral response in early sampling points.

RNAemia within six days of ICU admission was strongly associated with 28-day mortality, which is a well-defined clinical outcome measure[36] also suitable for COVID-19 ICU patients[5]. Thus far, studies on RNAemia included predominantly non-ICU patients and associated RNAemia with disease severity[7]. Few studies also reported on the ability of RNAemia to predict mortality[8–10] but none of these studies specifically focused on ICU patients in which RNAemia is likely to be most informative. In our study, RNAemia was more frequent in ICU patients with type 2 diabetes and hypertension, two well-known risk factors for poor outcome in COVID-19. Using droplet digital PCR[15], RNAemia might become even more frequent but the clinical relevance of very low levels of RNAemia is unclear. In comparison to RNAemia as assessed in our study (HR, 1.84 [95% CI, 1.22–2.77] adjusted for age and sex), the mortality risk conferred by increased nasopharyngeal SARS-CoV-2 RNA abundance was found to be small (HR, 1.07 [95% CI, 1.03–1.11], n = 1,145)[37].

RNAemia could be a consequence of severe disease or might contribute to poor outcome. Given that the SARS-CoV-2 entry receptor ACE2 is expressed on vascular cells, including endothelial cells, smooth muscle cells and pericytes of most organs[38,39], and SARS-CoV-2 RNA was detected in lungs, pharynx, heart, liver, brain and kidneys of autopsy tissue[40], RNAemia could reflect the extent of viral dissemination. Notably, serum levels of CDH5, an endothelial specific surface protein, differed between RNAemia positive *versus* negative ICU

patients. RNAemia was also inversely associated with monocyte counts. A decrease in monocyte counts in COVID-19 patients has been attributed to extravasation and recruitment to lungs[11,41].

Strikingly, patients with RNAemia showed dysregulation in several components of the complement, the coagulation and the kinin-kallikrein system. Viral envelope glycoproteins are an important trigger of the contact activation system[42] leading to a combined activation of these pathways, a hallmark of thromboinflammation[42]. SARS-CoV spike is a ligand of MBL2[43,44] - a pattern recognition molecule that initiates the lectin complement pathway[45]. Additionally, high levels of MBL2 are known to increase lectin pathway-mediated tissue damage[46,47]. This is consistent with our observation of a higher risk of mortality in COVID-19 ICU patients with elevated MBL2 levels. Systemic complement activation has been associated with respiratory failure in hospitalized COVID-19 patients[48] and complement deficiencies have been reported to have protective effects on COVID-19-associated morbidity and mortality[49]. Besides KLKB1, PIGR showed a different trajectory in RNAemia positive ICU patients. PIGR is a receptor that transports polymeric IgA and IgM from the basolateral to the apical surface of airway and gut mucosal cells[50]. Apart from its protective role, PIGR can be used by pathogens such as *Streptococcus pneumoniae* to facilitate infection of airway epithelial cells[51] and its plasma and lung tissue levels have been associated with severity of idiopathic pulmonary fibrosis[52], and cystic fibrosis[53], respectively.

Pull-down experiments using SARS-CoV-2 spike glycoprotein returned several members of the complement system. The complement system recruits neutrophils (C3a and C5a[54]), is essential for neutrophil extracellular trap (NET) formation (C3[55] and C3aR[56]), and can trigger NET formation (C5a[57]) when neutrophils are primed by interferon alpha or gamma[57] – cytokines that we previously found elevated in severe COVID-19 patients[11]. Furthermore, binding of C1q to NETs protects NETs from degradation by DNases in the

circulation[58]. NETosis was previously shown to be promoted by SARS-CoV-2 RNAemia but the mechanism remained elusive[16]. NET formation is also a prothrombotic process[59], and thrombotic complications are highly prevalent in severe COVID-19[60]. NET formation itself is part of a positive feedback loop, leading to activation of the alternative pathway of complement[61], the contact activation system[62], kinin-kallikrein system[62] and release of neutrophil-derived proteins, including the humoral pattern recognition receptor PTX3[63]. PTX3 is important for activation (through MBL2 and C1q)[64] and regulation (through CFH and C4BPB)[65,66] of the complement system[67]. It is noteworthy that PTX3 has been validated as one of the best predictors for mortality in an independent cohort of hospitalized COVID-9 patients covering 1,526 plasma proteins (Supplementary Table 6, https://www.olink.com/mgh-covid-study/).

Besides members of the complement system, we demonstrate that LGALS3BP is a novel putative binding partner of SARS-CoV-2 spike glycoprotein. LGALS3BP is prominently expressed in the lung[68] and possesses antiviral activity[69]. The rise in circulating LGALS3BP is not observed in non-COVID-19 sepsis ICU patients, highlighting the specificity for viral over bacterial infections. LGALS3BP directly interacts with adeno-associated viruses, inducing viral particle aggregation and an impairment of transduction[70]. Similarly, LGALS3BP reduces the infectivity of human immunodeficiency virus particles[71]. It is currently unknown if LGALS3BP-spike binding also affects the infectivity of SARS-CoV-2, *i.e.* by competing with binding to ACE2 or preventing the subsequent spike cleavage, which is essential for viral entry[72]. Additionally, the direct interaction between LGALS3BP and SARS-CoV-2 spike remains to be confirmed. Pull-down assays cannot rule out indirect binding to the bait protein.

In summary, RNAemia is frequent in COVID-19 ICU patients and associated with a higher risk of mortality. To our knowledge, SARS-CoV-2 RNA is the only disease-specific biomarker that has been associated with COVID-19 severity and mortality to date. Patients

311 with RNAemia may benefit from personalized treatment options. Finally, proteomic analyses
312 of blood samples from ICU patients with COVID-19 uncovered protein trajectories that
313 associated with RNAemia status, predicted 28-day mortality and identified LGALS3BP as a
314 novel interaction partner of the SARS-CoV-2 spike glycoprotein. Further studies are required
315 to assess the role of complement activation in COVID-19 on outcomes and explore the effect
316 of LGALS3BP on the infectivity of SARS-CoV-2.
317

References

1. Knaus, W. A., Draper, E. A., Wagner, D. P. & Zimmerman, J. E. APACHE II: a severity of disease classification system. *Crit. Care Med.* **13**, 818–29 (1985).
2. Vincent, J. L. *et al.* The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* **22**, 707–10 (1996).
3. Sinha, P. *et al.* Prevalence of phenotypes of acute respiratory distress syndrome in critically ill patients with COVID-19: a prospective observational study. *Lancet. Respir. Med.* **0**, (2020).
4. Zou, X. *et al.* Acute Physiology and Chronic Health Evaluation II Score as a Predictor of Hospital Mortality in Patients of Coronavirus Disease 2019. *Crit. Care Med.* **48**, e657–e665 (2020).
5. Intensive Care National Audit And Research Centre. *ICNARC report on COVID-19 in critical care 31 July 2020*. (2020).
6. Gupta, R. K. *et al.* Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: An observational cohort study. *Eur. Respir. J.* (2020). doi:10.1183/13993003.03498-2020
7. Andersson, M. I. *et al.* SARS-CoV-2 RNA detected in blood products from patients with COVID-19 is not associated with infectious virus. *Wellcome Open Res.* **5**, 181 (2020).
8. Xu, D. *et al.* Relationship Between serum SARS-CoV-2 nucleic acid(RNAemia) and Organ Damage in COVID-19 Patients: A Cohort Study. *Clin. Infect. Dis.* (2020). doi:10.1093/cid/ciaa1085
9. Fajnzylber, J. M. *et al.* SARS-CoV-2 Viral Load is Associated with Increased Disease Severity and Mortality. *medRxiv* 2020.07.15.20131789 (2020). doi:10.1101/2020.07.15.20131789
10. Prebensen, C. *et al.* Severe Acute Respiratory Syndrome Coronavirus 2 RNA in Plasma Is Associated With Intensive Care Unit Admission and Mortality in Patients Hospitalized With Coronavirus Disease 2019. *Clin. Infect. Dis.* (2020). doi:10.1093/cid/ciaa1338
11. Laing, A. G. *et al.* A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat. Med.* 1–13 (2020). doi:10.1038/s41591-020-1038-6
12. Carter, M. J. *et al.* Peripheral immunophenotypes in children with multisystem inflammatory syndrome associated with SARS-CoV-2 infection. *Nat. Med.* 1–7 (2020). doi:10.1038/s41591-020-1054-6
13. Seow, J. *et al.* Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat. Microbiol.* 1–10 (2020). doi:10.1038/s41564-020-00813-8
14. Chen, X. *et al.* Detectable Serum Severe Acute Respiratory Syndrome Coronavirus 2 Viral Load (RNAemia) Is Closely Correlated With Drastically Elevated Interleukin 6 Level in Critically Ill Patients With Coronavirus Disease 2019. *Clin. Infect. Dis.* (2020). doi:10.1093/cid/ciaa449
15. Bermejo-Martin, J. F. *et al.* Viral RNA load in plasma is associated with critical illness and a dysregulated host response in COVID-19. *medRxiv* 2020.08.25.20154252 (2020). doi:10.1101/2020.08.25.20154252
16. Veras, F. P. *et al.* SARS-CoV-2–triggered neutrophil extracellular traps mediate COVID-19 pathology. *J. Exp. Med.* **217**, (2020).
17. Zaid, Y. *et al.* Platelets Can Associate with SARS-Cov-2 RNA and Are

Hyperactivated in COVID-19. *Circ. Res.* CIRCRESAHA.120.317703 (2020). doi:10.1161/CIRCRESAHA.120.317703
18. Zhang, S. *et al.* SARS-CoV-2 binds platelet ACE2 to enhance thrombosis in COVID-19. *J. Hematol. Oncol.* **13**, 120 (2020).
19. Shen, B. *et al.* Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59-72.e15 (2020).
20. Messner, C. B. *et al.* Ultra-High-Throughput Clinical Proteomics Reveals Classifiers of COVID-19 Infection. *Cell Syst.* **11**, 11-24.e4 (2020).
21. Overmyer, K. A. *et al.* Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Syst.* (2020). doi:10.1016/J.CELS.2020.10.003
22. Di, B. *et al.* Identification and validation of predictive factors for progression to severe COVID-19 pneumonia by proteomics. *Signal Transduct. Target. Ther.* **5**, 217 (2020).
23. Jones, T. K. *et al.* Plasma sRAGE Acts as a Genetically Regulated Causal Intermediate in Sepsis-associated Acute Respiratory Distress Syndrome. *Am. J. Respir. Crit. Care Med.* **201**, 47–56 (2020).
24. Jabaudon, M. *et al.* Plasma sRAGE is independently associated with increased mortality in ARDS: a meta-analysis of individual patient data. *Intensive Care Med.* **44**, 1388–1399 (2018).
25. Jabaudon, M. *et al.* Soluble form of the receptor for advanced glycation end products is a marker of acute lung injury but not of severe sepsis in critically ill patients. *Crit. Care Med.* **39**, 480–8 (2011).
26. Cuello, F. *et al.* Redox state of pentraxin 3 as a novel biomarker for resolution of inflammation and survival in sepsis. *Mol. Cell. Proteomics* **13**, 2545–57 (2014).
27. Mauri, T. *et al.* Persisting high levels of plasma pentraxin 3 over the first days after severe sepsis and septic shock onset are associated with mortality. *Intensive Care Med.* **36**, 621–629 (2010).
28. Muller, B. *et al.* Circulating levels of the long pentraxin PTX3 correlate with severity of infection in critically ill patients. *Crit. Care Med.* **29**, 1404–7 (2001).
29. Porte, R. *et al.* The Long Pentraxin PTX3 as a Humoral Innate Immunity Functional Player and Biomarker of Infections and Sepsis. *Front. Immunol.* **10**, 794 (2019).
30. Williamson, E. J. *et al.* Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 1–7 (2020). doi:10.1038/s41586-020-2521-4
31. Ding, M. *et al.* An optimized sensitive method for quantitation of DNA/RNA viruses in heparinized and cyropreserved plasma. *J. Virol. Methods* **176**, 1–8 (2011).
32. Kaudewitz, D. *et al.* Impact of intravenous heparin on quantification of circulating microRNAs in patients with coronary artery disease. *Thromb. Haemost.* **110**, 609–615 (2013).
33. Schulte, C. *et al.* Comparative Analysis of Circulating Noncoding RNAs Versus Protein Biomarkers in the Detection of Myocardial Injury. *Circ. Res.* **125**, 328–340 (2019).
34. Dupuis, N., Muller, S., Treiber, T. & Escher, C. Evaluation of PQ500, a 500-plasma protein blood panel in NSCLC subjects using high-throughput MRM mass spectrometry. *J. Clin. Oncol.* **37**, 110–110 (2019).
35. Ou, X. *et al.* Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **11**, 1620 (2020).
36. Vincent, J.-L. Endpoints in sepsis trials: More than just 28-day mortality? *Crit. Care Med.* **32**, S209–S213 (2004).
37. Pujadas, E. *et al.* SARS-CoV-2 viral load predicts COVID-19 mortality. *Lancet Respir. Med.* **0**, (2020).
38. Cardot-Leccia, N., Hubiche, T., Dellamonica, J., Burel-Vandenbos, F. & Passeron, T.

Pericyte alteration sheds light on micro-vasculopathy in COVID-19 infection. *Intensive Care Med.* **46**, 1777–1778 (2020).
39. Hamming, I. *et al.* Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **203**, 631 (2004).
40. Puelles, V. G. *et al.* Multiorgan and Renal Tropism of SARS-CoV-2. *N. Engl. J. Med.* **383**, 590–592 (2020).
41. Sánchez-Cerrillo, I. *et al.* COVID-19 severity associates with pulmonary redistribution of CD1c+ DC and inflammatory transitional and nonclassical monocytes. *J. Clin. Invest.* (2020). doi:10.1172/JCI140335
42. Risitano, A. M. *et al.* Complement as a target in COVID-19? *Nat. Rev. Immunol.* **20**, 343–344 (2020).
43. Zhou, Y. *et al.* A single asparagine-linked glycosylation site of the severe acute respiratory syndrome coronavirus spike glycoprotein facilitates inhibition by mannose-binding lectin through multiple mechanisms. *J. Virol.* **84**, 8753–64 (2010).
44. Ip, W. K. E. *et al.* Mannose-binding lectin in severe acute respiratory syndrome coronavirus infection. *J. Infect. Dis.* **191**, 1697–704 (2005).
45. Degn, S. E. *et al.* Complement activation by ligand-driven juxtaposition of discrete pattern recognition complexes. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13445–50 (2014).
46. Jordan, J. E., Montalto, M. C. & Stahl, G. L. Inhibition of Mannose-Binding Lectin Reduces Postischemic Myocardial Reperfusion Injury. *Circulation* **104**, 1413–1418 (2001).
47. Schafranski, M. D., Stier, A., Nisihara, R. & Messias-Reason, I. J. T. Significantly increased levels of mannose-binding lectin (MBL) in rheumatic heart disease: a beneficial role for MBL deficiency. *Clin. Exp. Immunol.* **138**, 521–5 (2004).
48. Holter, J. C. *et al.* Systemic complement activation is associated with respiratory failure in COVID-19 hospitalized patients. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 25018–25025 (2020).
49. Ramlall, V. *et al.* Immune complement and coagulation dysfunction in adverse outcomes of SARS-CoV-2 infection. *Nat. Med.* 1–7 (2020). doi:10.1038/s41591-020-1021-2
50. Li, Y. *et al.* Structural insights into immunoglobulin M. *Science* **367**, 1014–1017 (2020).
51. Zhang, J.-R. *et al.* The Polymeric Immunoglobulin Receptor Translocates Pneumococci across Human Nasopharyngeal Epithelial Cells. *Cell* **102**, 827–837 (2000).
52. Todd, J. L. *et al.* Peripheral blood proteomic profiling of idiopathic pulmonary fibrosis biomarkers in the multicentre IPF-PRO Registry. *Respir. Res.* **20**, 227 (2019).
53. Collin, A. M. *et al.* Lung immunoglobulin A immunity dysregulation in cystic fibrosis. *EBioMedicine* **60**, 102974 (2020).
54. Haas, P.-J. & van Strijp, J. Anaphylatoxins: their role in bacterial infection and inflammation. *Immunol. Res.* **37**, 161–75 (2007).
55. Yipp, B. G. *et al.* Infection-induced NETosis is a dynamic process involving neutrophil multitasking in vivo. *Nat. Med.* **18**, 1386–93 (2012).
56. Guglietta, S. *et al.* Coagulation induced by C3aR-dependent NETosis drives protumorigenic neutrophils during small intestinal tumorigenesis. *Nat. Commun.* **7**, 11037 (2016).
57. Martinelli, S. *et al.* Induction of genes mediating interferon-dependent extracellular trap formation during neutrophil differentiation. *J. Biol. Chem.* **279**, 44123–32 (2004).
58. Leffler, J. *et al.* Neutrophil extracellular traps that are not degraded in systemic lupus

erythematosus activate complement exacerbating the disease. *J. Immunol.* **188**, 3522–31 (2012).
59. Kimball, A. S., Obi, A. T., Diaz, J. A. & Henke, P. K. The Emerging Role of NETs in Venous Thrombosis and Immunothrombosis. *Front. Immunol.* **7**, 236 (2016).
60. Zhang, L. *et al.* Deep Vein Thrombosis in Hospitalized Patients With COVID-19 in Wuhan, China: Prevalence, Risk Factors, and Outcome. *Circulation* **142**, 114–128 (2020).
61. Wang, H., Wang, C., Zhao, M.-H. & Chen, M. Neutrophil extracellular traps can activate alternative complement pathways. *Clin. Exp. Immunol.* **181**, 518–27 (2015).
62. Oehmcke, S., Mörgelin, M. & Herwald, H. Activation of the human contact system on neutrophil extracellular traps. *J. Innate Immun.* **1**, 225–30 (2009).
63. Jaillon, S. *et al.* The humoral pattern recognition receptor PTX3 is stored in neutrophil granules and localizes in extracellular traps. *J. Exp. Med.* **204**, 793–804 (2007).
64. Ma, Y. J. *et al.* Heterocomplexes of mannose-binding lectin and the pentraxins PTX3 or serum amyloid P component trigger cross-activation of the complement system. *J. Biol. Chem.* **286**, 3405–17 (2011).
65. Deban, L. *et al.* Binding of the long pentraxin PTX3 to factor H: interacting domains and function in the regulation of complement activation. *J. Immunol.* **181**, 8433–40 (2008).
66. Braunschweig, A. & Józsi, M. Human Pentraxin 3 Binds to the Complement Regulator C4b-Binding Protein. *PLoS One* **6**, e23991 (2011).
67. Ma, Y. J. & Garred, P. Pentraxins in Complement Activation and Regulation. *Front. Immunol.* **9**, 3046 (2018).
68. Ullrich, A. *et al.* The secreted tumor-associated antigen 90K is a potent immune stimulator. *J. Biol. Chem.* **269**, 18401–7 (1994).
69. Loimaranta, V., Hepojoki, J., Laaksoaho, O. & Pulliainen, A. T. Galectin-3-binding protein: A multitask glycoprotein with innate immunity functions in viral and bacterial infections. *J. Leukoc. Biol.* **104**, 777–786 (2018).
70. Denard, J. *et al.* Human galectin 3 binding protein interacts with recombinant adeno-associated virus type 6. *J. Virol.* **86**, 6620–31 (2012).
71. Wang, Q., Zhang, X., Han, Y., Wang, X. & Gao, G. M2BP inhibits HIV-1 virion production in a vimentin filaments-dependent manner. *Sci. Rep.* **6**, 32736 (2016).
72. Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e8 (2020).
73. Vogels, C. B. F. *et al.* Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. *Nat. Microbiol.* 1–7 (2020). doi:10.1038/s41564-020-0761-6
74. Troyanskaya, O. *et al.* Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**, 520–525 (2001).
75. Saeys, Y., Inza, I. & Larranaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **23**, 2507–2517 (2007).
76. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
77. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
78. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–30 (2011).
79. Ojala, M. & Garrica, G. C. Permutation Tests for Studying Classifier Performance. *J. Mach. Learn. Res.* **11**, 1833–1863 (2010).
80. Kornbrot, D. Point Biserial Correlation. in *Encyclopedia of Statistics in Behavioral Science* (John Wiley & Sons, Ltd, 2005). doi:10.1002/0470013192.bsa485

Methods

Study design and recruitment. An overview of the study design is presented in Supplementary Fig. 1. *COVID-19 cohorts.* COVID-19-positive patients, as confirmed by RT-qPCR of nasopharyngeal samples, who were admitted to the ICUs of Guy's and St Thomas' NHS Foundation Trust (GSTT) and King's College Hospital (KCH) between March 12, 2020 and July 1, 2020, were recruited for an observational cohort study with serial blood sampling and analysis of clinical outcomes. The primary outcome measure was defined as mortality 28 days after ICU admission. Serial blood sampling was performed within 24 hours of admission to ICU and thereafter three measurements were taken during week 1, week 2 and again before discharge. In addition, we obtained plasma samples from COVID-19 patients upon hospitalization at GSTT (non-ICU COVID-19 cohort). *Non-COVID-19 comparator cohorts.* Plasma was collected from patients enrolled at the same time in the same KCH ICU as our COVID-19 ICU cohort but who repeatedly tested negative for nasopharyngeal SARS-CoV-2 (intra-pandemic, non-COVID-19 ICU cohort). Serial blood sampling of these samples was performed identical to our COVID-19 cohort. Additionally, pre-pandemic plasma samples from patients recruited at GSTT prior to the COVID-19 pandemic were available as controls. This included serial plasma samples from sepsis ICU patients (pre-pandemic, non-COVID-19 ICU sepsis cohort), collected upon admission and at three timepoints thereafter; as well as plasma samples from patients before elective cardiac surgery (pre-pandemic, non-COVID-19 control cohort). The study was approved by an institutional review board (REC19/NW/0750 for all patients recruited at KCH; REC19/SC/0187 for patients recruited at GSTT of the COVID-19 ICU cohort, the pre-pandemic sepsis ICU cohort, the pre-pandemic control cohort; REC19/SC/0232 for patients recruited at GSTT of the non-ICU COVID-19 cohort). Written informed consent was obtained directly from patients (if mentally competent), or from the next

of kin or professional consultee. The consent procedure was then completed with retrospective consent if the patient regained capacity.

Inactivation of serum and plasma. Plasma was collected in EDTA BD Vacutainer™ tubes (BD, 362799), whereas serum was collected in silica BD Vacutainer™ tubes (BD, 367820) and left to clot for 15 min. Plasma and serum tubes were then centrifuged at 2,000 x g for 15 min. Infectious samples were then transferred to a containment level 3 facility for safe inactivation. Samples destined for RNA extraction were inactivated by addition of 100 µL of serum or plasma to 500 µL QIAzol (Qiagen, 79306), followed by 40 s of vortexing and 5 min incubation at room temperature. Samples destined for protein analysis were inactivated by addition of 1% (v/v) Triton X-100 (Sigma, T8787) and 1% (v/v) tributyl phosphate (Sigma, 00675), followed by 15 s of vortexing and 4 h incubation at room temperature. All samples were then frozen at -80°C until further processing.

RNA extraction and heparinase treatment. Total RNA was extracted using the miRNeasy Mini kit (Qiagen, 217004) according to the manufacturer's recommendations. Total RNA was eluted in 30 µL of nuclease-free H_2O by centrifugation at 8,500 x g for 1 min at 4°C. To overcome the confounding effect of heparin on qPCR[31,32], RNA was treated with heparinase as described previously[33]. Briefly, 8 µL of RNA was added to 2 µL of heparinase 1 from Flavobacterium (Sigma, H2519), 0.4 µL RNase inhibitor (Ribo Lock 40U/µL, ThermoFisher, EO0381) and 5.6 µL of heparinase buffer (pH 7.5) and incubated at 25°C for 3 h.

Reverse transcription-quantitative polymerase chain reaction (RT-qPCR). For detection of SARS-CoV-2 RNA we performed a two-step RT-qPCR using the LunaScript® RT SuperMix Kit (NEB, E3010) and the Luna Universal Probe qPCR Master Mix (NEB, M3004)

according to the manufacturer's recommendations, apart from reducing the total qPCR reaction volume to 5 μL and loading a cDNA dilution of 1:4 instead of 1:8 when performing the qPCR reaction. Primer/probe sequences targeting the SARS-CoV-2 nucleocapsid (N) gene (N1 and N2) were predesigned by Integrated DNA Technologies (IDT, 10006821, 10006822, 10006823, 10006824, 10006825, 10006826) according to the protocol for the detection of SARS-CoV-2 of the United States Centers for Disease Control and Prevention (US CDC), using 5' FAM / ZEN™ / 3' Iowa Black™ FQ probes. The qPCR reaction concentration for probe (125 nM), forward (500 nM) and reverse primers (500 nM) were used according to the US CDC protocol. A plasmid positive control (2019-nCoV_N Positive Control plasmid, IDT, 10006625) was measured on each qPCR plate. Reactions were loaded using a Bravo Automated Liquid Handling Platform (Agilent). qPCR was performed on a ViiA7 Real-Time PCR System (Applied Biosystems). Samples were considered positive for SARS-CoV-2 if the cycle quantification (Cq) value of either N1 or N2 was below 40. Abundance of SARS-CoV-2 RNA in patients who tested positive had a mean Cq of 34.4; range: 29.8-37.6. As reported before[73], N1 primers returned lower Cq values (higher abundance) than N2 primers (Supplementary Fig. 7).

Measurement of anti-SARS-CoV-2 antibodies. IgG antibodies against the SARS-CoV-2 spike S1 domain were measured by ELISA (Anti-SARS-CoV-2 IgG ELISA, Euroimmun, EI 2606-9601 G) according to the manufacturer's recommendations. Since no international reference serum for anti-SARS-CoV-2 antibodies exists, calibration was performed in ratios, giving relative antibody quantification. Neutralizing antibodies against SARS-CoV-2 were measured using a Surrogate Virus Neutralization Test (SARS-CoV-2 sVNT Kit, GenScript, L00847) according to the manufacturer's recommendations. This ELISA-based kit detects antibodies that are able to block the interaction between the SARS-CoV-2 spike receptor

592 binding domain (RBD) and the angiotensin converting enzyme (ACE2) cell receptor. For
593 validation of sVNT measurements in a subset of samples, neutralization potency was measured
594 using HIV-1 (human immunodeficiency virus-1) based virus particles, pseudotyped with
595 SARS-CoV-2 spike protein in a HeLA cell line stably expressing the ACE2 receptor, as
596 described previously[13].
597
598 **In-solution protein digestion.** 10 µL of inactivated serum or plasma were denatured by the
599 addition of urea (final concentration 7.2 M) and reduced using dithiothreitol (final
600 concentration 5 mM) for 1 h at 37 °C and 180 rpm. Reduced proteins were cooled down to
601 room temperature before being alkylated in the dark for 1 h using iodoacetamide (final
602 concentration 25 mM). An aliquot equivalent to 40 µg of alkylated protein was added to a
603 0.1 M triethylammonium bicarbonate solution (pH 8.2) and digested for 18 h at 37 °C, at
604 180 rpm using 1.6 µg of Trypsin/LysC (Promega, V5072). Digested peptide solutions were
605 acidified using trifluoroacetic acid (TFA, final concentration 1 %).
606
607 **Peptide clean-up and stable isotope-labelled standard (SIS) spike-in.** Peptide clean-up was
608 achieved using a Bravo AssayMAP Liquid Handling Platform (Agilent). After conditioning
609 and equilibration of the resin, acidified peptide solutions were loaded onto AssayMAP C18
610 Cartridges (Agilent, 5190-6532), washed using 1 % acetonitrile (ACN), 0.1 % TFA (aq) and
611 eluted using 70 % ACN, 0.1 % TFA (aq). Eluted peptides were vacuum centrifuged (Thermo
612 Scientific, Savant SPD131DDA) to dry and resuspended in 40 µL of 2 % ACN, 0.05 % TFA
613 (aq). For clinical cohort analysis, 6 µL of cleaned peptide solution was added to two injection
614 equivalents of PQ500 SIS mix (Biognosys) using a Bravo Liquid Handling Platform (Agilent).
615

Data-independent acquisition-mass spectrometry (DIA-MS) analysis. Peptides were analyzed using a high-performance liquid chromatography (HPLC)-MS assembly consisting of an UltiMate 3000 HPLC system (Thermo Scientific) which was equipped with a capillary flow selector and coupled via an EASY-Spray NG Source (Thermo Scientific) to an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific). To generate DIA data for serum samples (GSTT COVID-19 ICU cohort) and plasma samples (KCH COVID-19 ICU cohort, the pre-pandemic sepsis ICU cohort and the pre-pandemic control patients before elective cardiac surgery), peptides were injected onto a C18 trap cartridge (Thermo Scientific, 160454) at a flow rate of 25 µL / min for 1 min, using 0.1% formic acid (FA, aq). The initial capillary flow rate was reduced from 3 to 1.2 µL / min in 1 min at 1% B. Peptides were then eluted from the trap cartridge and separated on an analytical column (Thermo Scientific, ES806A, at 50 °C) using the following gradient: 1–11 min, 1–5% B; 11–32 min, 5–18% B; 32–52 min, 18–40% B; 52–52.1 min, 40–99% B; 52.1–58 min, 99% B. The flow rate was increased to 3 µL/min and the column was washed using the following gradient: 58–58.1 min, 99–1% B; 58.1–59.9 min, 1–99% B; 59.9–60 min, 99–1% B. Finally, the column was equilibrated at 1% B for 6 min. In all HPLC-DIA-MS analyses, mobile phase A was 0.1% FA (aq) and mobile phase B was 80% ACN, 0.1% FA (aq). Precursor MS1 spectra were acquired using Orbitrap detection (resolution 60000 at 200 m/z, scan range 329–1201 m/z). Quadrupole isolation was used to sequentially scan 30 precursor m/z windows of variable width (Supplementary Table 10). Per isolation window, semi-targeted Orbitrap MS2 spectra (resolution 30000 at 200 m/z) were collected following higher-energy C-trap dissociation.

MS database search for DIA-MS analysis. PQ500 SIS-spiked DIA data from all serum and plasma samples of the GSTT COVID-19 ICU cohort, the KCH COVID-19 ICU cohort, the non-COVID-19 sepsis ICU cohort and the control patients before elective cardiac surgery were

analyzed in Spectronaut v14 (Biognosys AG), using the provided PQ500 analysis plug-in. MS1 and MS2 mass tolerance strategies were set to relative at a tolerance of 20 ppm, while MS2 mass tolerance was set to dynamic. Retention time calibration was achieved using the spiked iRT peptides included in the PQ500 SIS mix. Precursor and protein Q-value cutoff was set to 0.01. Quantification was conducted at an MS2 level using peak areas and individual runs were normalized using the global strategy set to median. All peptides for reported proteins were manually checked to ensure accurate peak integration across all samples. Peptides with a Q-value of more than 0.01 or a signal to noise ratio of less than 5 were marked as missing. Peptides with more than 30% missing values across all samples were filtered out and the remaining missing values were imputed using the KNN algorithm (K = 5)[74]. Spearman correlations of peptides belonging to the same protein were computed. In case more than two peptides per protein were detected, peptides were filtered if their correlation with the remaining peptides was less than r = 0.4. In case two peptides per protein were detected, the most abundant peptide was kept even when correlation was less than r = 0.4. Final protein abundance was calculated by summing up the quantified peptide abundances. Final quantitative comparisons were conducted using the light/heavy peptide abundance ratio. For validation of our DIA-MS data, we correlated levels to clinical measurements of albumin (n = 49, r = 0.68, P <0.05) and C-reactive protein (n = 49, r = 0.83, P <0.05) as examples of high and medium-abundant proteins.

SARS-CoV-2 spike protein pull-down. His-tagged recombinant SARS-CoV-2 spike glycoprotein (RP-87680, ThermoFisher) was added to 1:2 PBS-diluted plasma from COVID-19 ICU patients (n = 8) or non-COVID-19 controls (n = 3) at 200 ng/µL and incubated overnight at 4°C with intermittent mixing. His-tagged spike was then isolated by means of metal affinity magnetic beads (Dynabeads His-Tag Isolation and Pull-down, 10103D, ThermoFisher) and eluted in imidazole-containing phosphate buffer. Proteins in the pull-down

isolates were denatured, reduced, alkylated and precipitated, as described above. Proteins interacting non-specifically with the solid phase were determined by incubating plasma samples with magnetic beads without the addition of His-tagged spike. Pull-down of His-tagged spike without addition of plasma was performed as an additional control. Spike pull-down protein digestion followed the same protocol outlined above.

Data-dependent acquisition (DDA)-MS analysis. Proteins from the spike pull-down experiments were subject to in-solution tryptic digestion and C18 cleanup as described above. Tryptic peptides were analyzed by LC-MS/MS. An UltiMate 3000 HPLC system (Thermo Scientific) with a nanoflow selector was coupled via an EASY-Spray Source (Thermo Scientific) to a Q Exactive HF mass spectrometer (Thermo Scientific). Peptides were injected onto a C18 trap cartridge (Thermo Scientific, 160454) at a flow rate of 25 µL / min for 1 min, using 0.1% FA (aq). Peptides were eluted from the trap cartridge and separated on an analytical column (EASY-Spray C18 column, 75 µm x 50 cm, Thermo Scientific, ES803A, at 45 °C) at a flow rate of 0.25 µL / min using the following gradient: 0–1 min, 1% B; 1–6 min, 1–6% B; 6−40 min, 6−18% B; 40–70 min, 18–35% B; 70–80 min, 35–45% B; 80–81 min, 45–99% B; 81–89.8 min, 99% B; 89.8–90 min, 99–1% B; 90–120 min, 1% B. Mobile phase A was 0.1% FA (aq) and mobile phase B was 80% ACN, 0.1% FA (aq). Precursor MS1 spectra were acquired using Orbitrap detection (resolution 60000 at 200 m/z, scan range 350–1600). Data-dependent MS2 spectra of the most abundant precursor ions were obtained after higher-energy C-trap dissociation and Orbitrap detection (resolution 15000 at 200 m/z) with TopN mode (loop count 15) and dynamic exclusion (duration 40 s) enabled.

MS database search for DDA-MS analysis. Proteome Discoverer software (version 2.3.0.523, Thermo Scientific) was used to search raw SARS-CoV-2 spike glycoprotein pull-

691 down data files against a human database (UniProtKB/Swiss-Prot version 2020 01, 20,365
692 protein entries) supplemented with SARS-CoV-2 spike glycoprotein (1 protein entry) using
693 Mascot (version 2.6.0, Matrix Science). The mass tolerance was set at 10 ppm for precursor
694 ions and 0.02 Da for fragment ions. Trypsin was used as the digestion enzyme with up to two
695 missed cleavages being allowed. Carbamidomethylation of cysteines and oxidation of
696 methionine residues were chosen as fixed and variable modifications, respectively.
697
698 **Machine learning**. In addition to statistical techniques, machine learning was deployed to
699 identify a prognostic classifier for COVID-19 ICU patients based on 27 clinical variables,
700 RNAemia and three ELISA measurements. The RNAemia feature was defined as a binary
701 feature which takes a true value when RNAemia was present within six days upon admission
702 to ICU. Statistical significance with P value <0.05 was used as selection criterium for singleton
703 markers. The shortlisted singleton markers were subsequently compared in binary and triplet
704 combinations with all 27 clinical variables, RNAemia and the three ELISA measurements of
705 PTX3, RAGE and LGALS3BP. In this setup, binary and triplet combinatorial feature search
706 was performed using wrapper feature selection[75] with support vector machine (SVM) classifier
707 using radial basis function (RBF) kernel. Feature combinations were evaluated using the
708 average of sensitivity, positive predicted value (PPV) and area under the receiver operating
709 characteristic curve (ROC AUC) metrics. Given the imbalanced data with positive class *i.e.*
710 non-survivors as the minority class, PPV along with sensitivity helps to balance false positives
711 and false negatives. Combined with ROC AUC, it further facilitates equilibrium between
712 sensitivity and specificity with high prediction probability. SVM uses hyperplane (decision
713 surface) leveraging only a percentage of training samples (support vectors), thus offering high
714 generalization ability attributed to its near impervious characteristic to new samples[76].
715 Combinations were restricted to a maximum of triplets to enhance ease of clinical

implementation and avoid the risk of overfitting. Additionally, 10-fold cross validation along with leave-one-out validation was used to avoid overfitting and test model generalization. The SVM Synthetic Minority Oversampling Technique (SMOTE) was used to prevent learning bias of SVM RBF towards the majority class[77]. Tuning of SVM RBF external parameter *i.e.* C was performed using grid search. The Scikit-learn default *i.e.* 'scale' was used for the SVM RBF gamma parameter[78]. A permutation test was performed to evaluate the null hypothesis that the classifier performance is by chance *i.e.* input variables and outcome labels are independent[79]. Hence, rejection of the null hypothesis implies that the classifier has found a real class structure (pattern) in the data. For technical validation of our 'Age, RNAemia' model based on SVM RBF, we employed a permutation test for statistical significance of the classifier performance; and stability of feature importance in an alternate machine learning feature ranking model *i.e.* Random forest with resampling. Age and RNAemia were ranked among the top five most important features based on mean importance across 100 resampling cycles of sensitivity analysis. A permutation test with 50 permutes *i.e.* repeating the classification procedure after random permuting of the outcome labels returned a significant *P* value (Supplementary Fig. 8). The implementation of machine learning was done using Scikit-learn 0.23.2 python package[78].

Statistical analysis. Mann-Whitney U significance test was used for continuous variables and Fisher exact test for binary variables. Spike pull-down data was analyzed by paired or unpaired Student's t-tests as appropriate. Statistical comparisons on MS data were performed using the Ebayes algorithm of the limma package correcting for age and sex. Timepoint comparisons were performed using the non-parametric Kruskal Wallis test. Correlation patterns between continuous variables were analyzed using Spearman correlation. Correlation between categorical and continuous variables was examined using point-biserial correlation[80]. Anti-

741 SARS-CoV-2 antibody data and trajectories of protein clusters were fitted using Generalized
742 Alternative Models (GAM), with *P* values reporting the effect of RNAemia or mortality in the
743 model. Survival analysis was performed using Cox regression and Kaplan-Meier plots
744 leveraging the R 'survival' package. All features were scaled to a mean of zero and a standard
745 deviation of one. Features with missing values ≥30% were dropped and not used for data
746 analysis. This resulted in two clinical variables being dropped, *i.e.* eosinophils and basophils.
747 The remaining features were imputed, as applicable, using K nearest neighbors (KNN) based
748 imputation with $K = 5$ (Supplementary Table 11)[74]. To validate DIA-MS findings a publicly
749 available proximity-extension assay proteomics-based dataset was analyzed *(Data provided by*
750 *the MGH Emergency Department COVID-19 Cohort (Filbin, Goldberg, Hacohen) with Olink*
751 *Proteomics*). Differential expression analysis of proteins in survivors and non-survivors 28-
752 days after hospitalization within the Olink dataset was achieved through the Ebayes method of
753 the limma package. Statistical analysis and associated Figures were generated with R
754 programming environment (version 4.02), Python programming environment (version 3.8.6)
755 and GraphPad software (version 8.4.3). Schematic diagrams were created with Biorender.com.
756

Acknowledgements

M.M. and A.M.S. are British Heart Foundation (BHF) Chair Holders with BHF program grant support (CH/16/3/32406, RG/16/14/32397 and CH/1999001/11735, RE/18/2/34213, respectively). C.G., E.R. and B.S. are funded by BHF PhD studentships (FS/18/60/34181, FS/17/65/33481, FS/19/58/34895). M.H. is funded by an interdisciplinary PhD studentship from King's BHF Centre of Research Excellence. M.F. is funded by a National Institute of Academic Anesthesia BJA-RCOA PhD Fellowship WKR0-2018-0047. M.J.W.M. is grateful to the Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust for support. K.O.G. is supported by a UK Medical Research Council Clinical Research Training Fellowship (MR/R017751/1). K.J.D was supported by a King's Together Rapid COVID-19 Call award, by an MRC Discovery Award (MC/PC/15068), Huo Family Foundation and the Fondation Dormeur, Vaduz. The NIHR Collaboration for Leadership in Applied Health Research and Care South London at King's College Hospital NHS Foundation Trust, awarded to J.D.E. who is also supported by a charitable donation from the Lower Green Foundation.

M.M.'s research was made possible through the support of the BIRAX Ageing Initiative and funding from the EU Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 813716 (TRAIN-HEART), the Leducq Foundation (18CVD02), the excellence initiative VASCage (Centre for Promoting Vascular Health in the Ageing Community, project number 868624) of the Austrian Research Promotion Agency FFG (COMET program–Competence Centers for Excellent Technologies) funded by the Austrian Ministry for Transport, Innovation and Technology; the Austrian Ministry for Digital and Economic Affairs; and the federal states Tyrol (via Standortagentur), Salzburg, and Vienna (via Vienna Business Agency), two BHF project grant supports (PG/17/48/32956, SP/17/10/33219) and the BHF Centre for Vascular Regeneration with Edinburgh/Bristol (RM/17/3/33381). M.M. and S.R.B. acknowledge support as visiting professors as part of the

Transcampus TU Dresden King’s College London Initiative. The work of A.C.H. is supported by a Cancer ImmunoTherapy Accelerator award from CRUK; the Wellcome Trust (106292/Z/14/Z); the Rosetrees Trust; King’s Together Seed Fund; The John Black Charitable Foundation; Royal Society Grant IES\R3\170319 and the Francis Crick Institute, which receives core funding from Cancer Research UK (FC001093), the MRC (FC001093) and the Wellcome Trust (FC001093). M.S.H. is supported by the National Institute for Health Research Clinician Scientist Award (CS-2016-16-011). The research was funded/supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Fig. 1

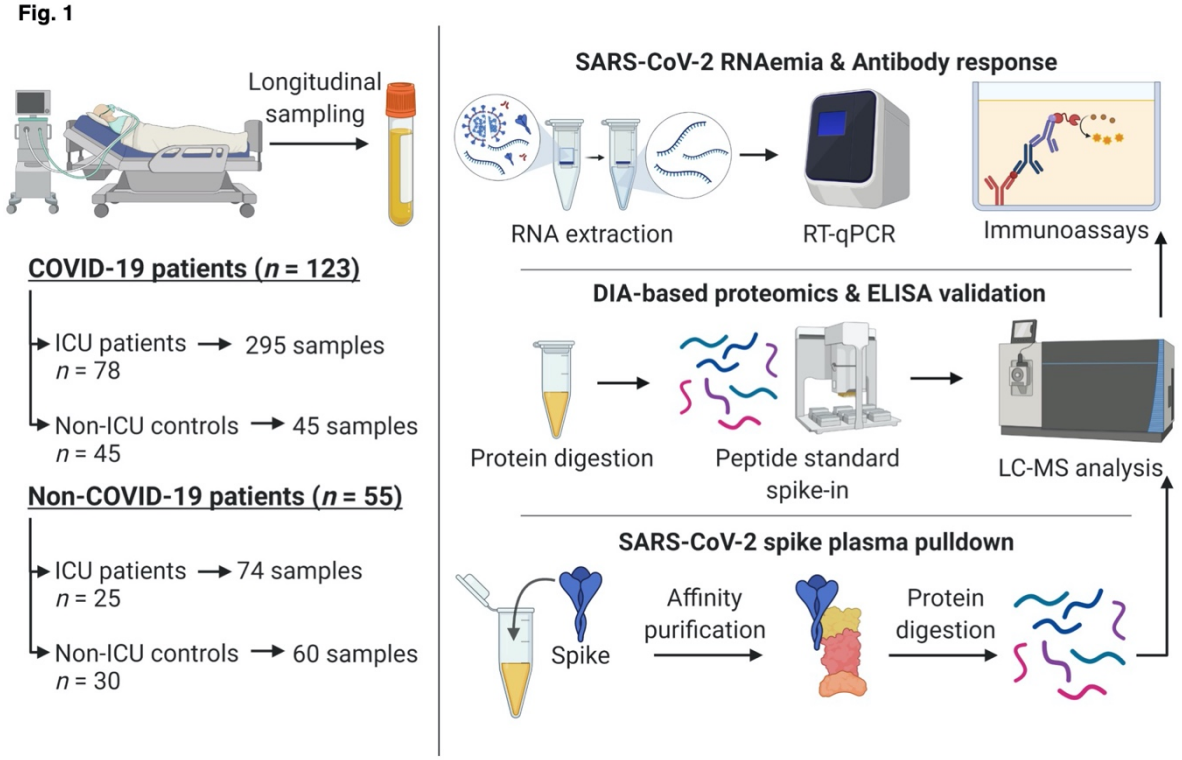
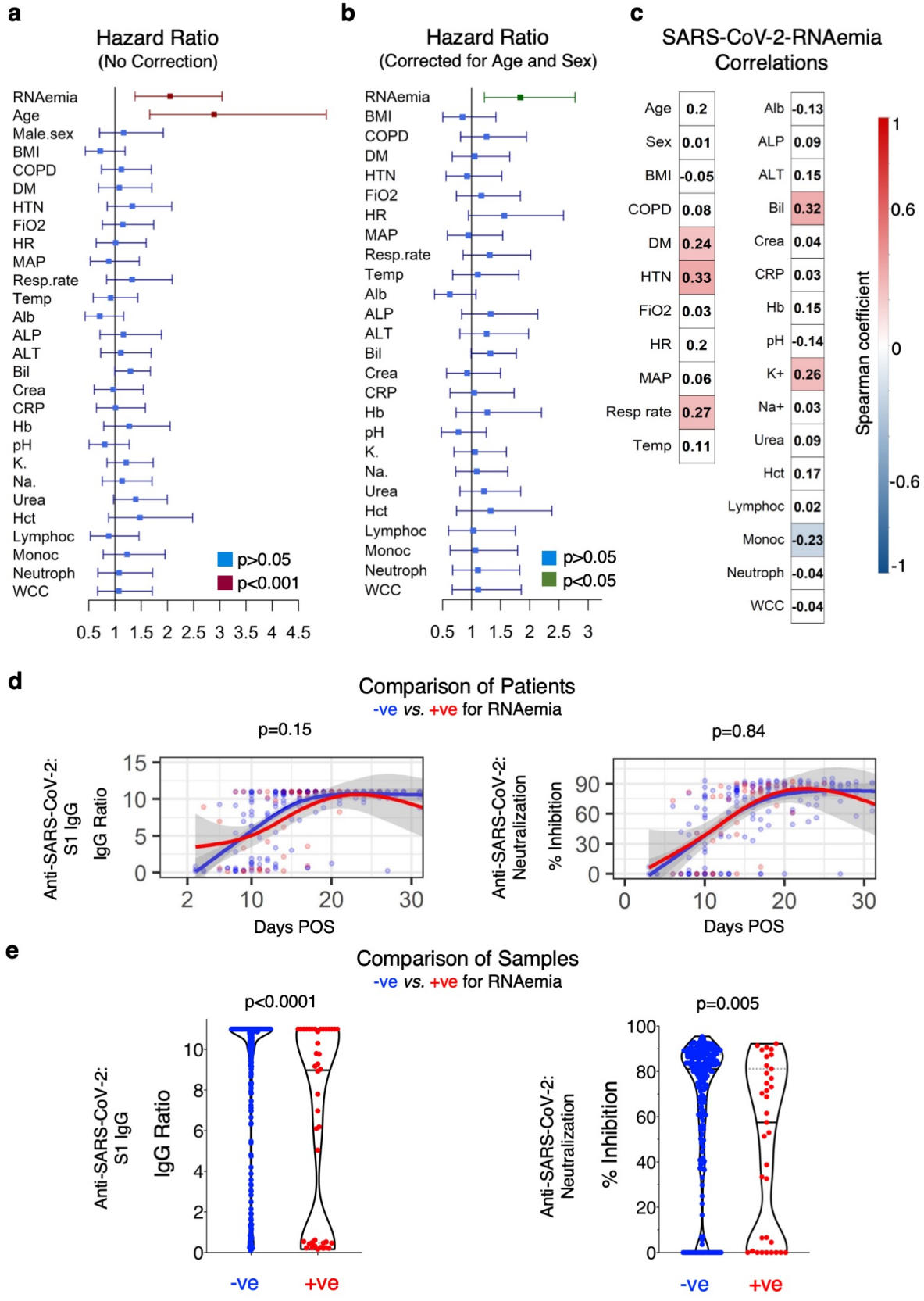


Fig. 1. Schematic of study design. Plasma and serum samples were obtained from multiple patient cohorts across two UK-based university hospitals, including 123 COVID-19 patients: 78 SARS-CoV-2 positive patients in ICU were sampled at multiple time points over a 2-week period and compared to hospitalized non-ICU SARS-CoV-2 positive patients (n = 45). We used non-COVID-19 ICU patients (n = 25) and patients before and after undergoing elective cardiac surgery (n = 30) as controls. Patient samples were assessed for SARS-CoV-2 RNAemia, antibody responses and protein changes in the circulation. Finally, plasma protein interactions with SARS-CoV-2 spike glycoprotein were determined using a pull-down assay followed by mass spectrometry analysis.

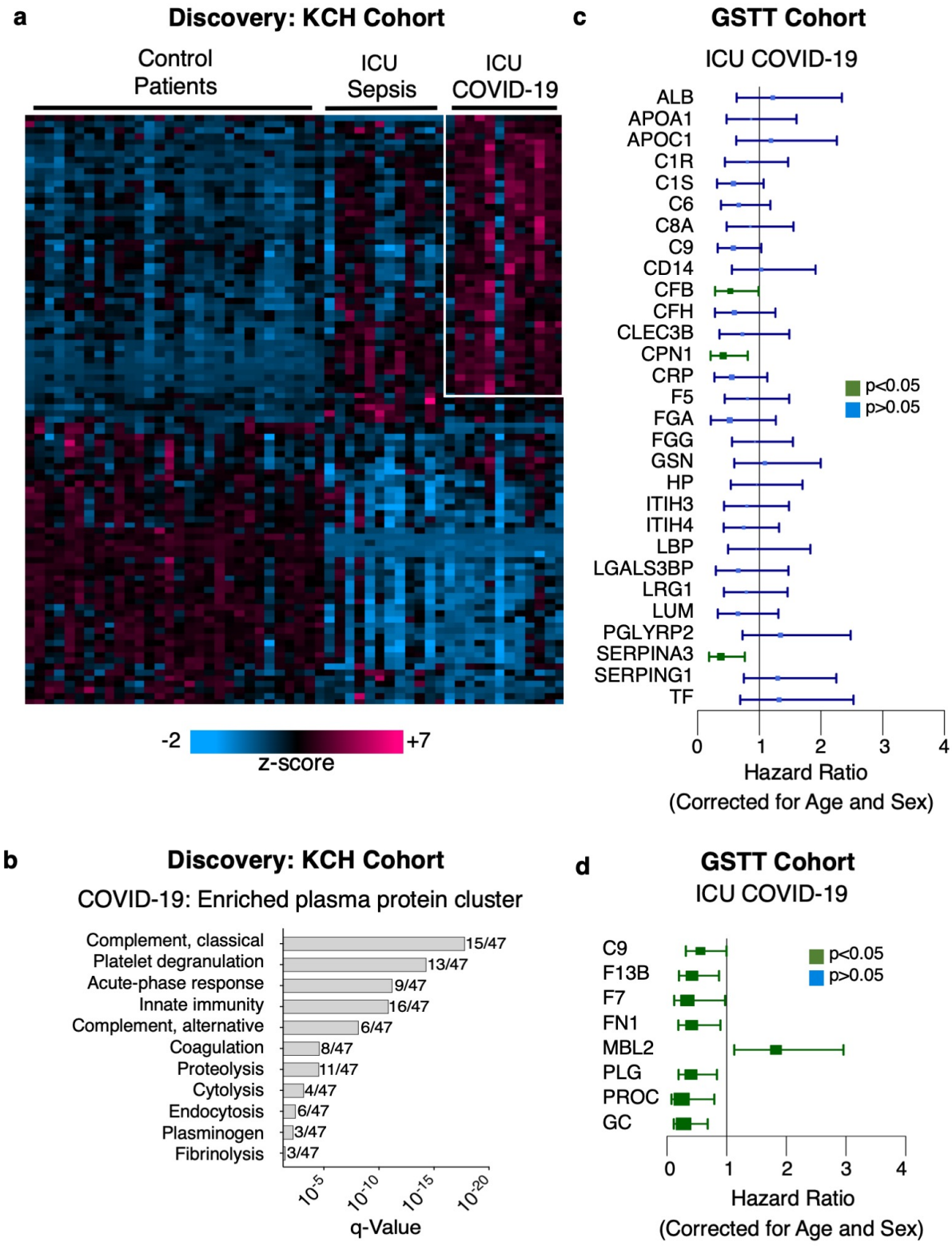
Fig. 2


a
Hazard Ratio
(No Correction)
RNAemia
Age
Male.sex
BMI
COPD
DM
HTN
FiO2
HR
MAP
Resp.rate
Temp
Alb
ALP
ALT
Bil
Crea
CRP
Hb
pH
K.
Na.
Urea
Hct
Lymphoc
Monoc
Neutroph
WCC
p>0.05
p<0.001
0.5 1 1.5 2 2.5 3 3.5 4 4.5
b
Hazard Ratio
(Corrected for Age and Sex)
RNAemia
BMI
COPD
DM
HTN
FiO2
HR
MAP
Resp.rate
Temp
Alb
ALP
ALT
Bil
Crea
CRP
Hb
pH
K.
Na.
Urea
Hct
Lymphoc
Monoc
Neutroph
WCC
p>0.05
p<0.05
0.5 1 1.5 2 2.5 3
c
SARS-CoV-2-RNAemia
Correlations
Age 0.2
Sex 0.01
BMI -0.05
COPD 0.08
DM 0.24
HTN 0.33
FiO2 0.03
HR 0.2
MAP 0.06
Resp rate 0.27
Temp 0.11
Alb -0.13
ALP 0.09
ALT 0.15
Bil 0.32
Crea 0.04
CRP 0.03
Hb 0.15
pH -0.14
K+ 0.26
Na+ 0.03
Urea 0.09
Hct 0.17
Lymphoc 0.02
Monoc -0.23
Neutroph -0.04
WCC -0.04
Spearman coefficient
1
0.6
0
-0.6
-1
d
Comparison of Patients
-ve vs. +ve for RNAemia
p=0.15
Anti-SARS-CoV-2: S1 IgG
IgG Ratio
15
10
5
0
2 10 20 30
Days POS
p=0.84
Anti-SARS-CoV-2: Neutralization
% Inhibition
90
60
30
0
0 10 20 30
Days POS
e
Comparison of Samples
-ve vs. +ve for RNAemia
p<0.0001
Anti-SARS-CoV-2: S1 IgG
IgG Ratio
10
8
6
4
2
0
-ve
+ve
p=0.005
Anti-SARS-CoV-2: Neutralization
% Inhibition
100
80
60
40
20
0
-ve
+ve

803

Fig. 2. SARS-CoV-2 RNAemia and the humoral immune response. a, Unadjusted hazard ratios based on two ICU patient cohorts (KCH and GSTT, 60 survivors and 18 non-survivors). Green indicates *P* value <0.05, maroon indicates *P* value <0.001 and blue indicates *P* value >0.05. **b**, Hazard ratios after adjustment for age and sex. **c**, Association of SARS-CoV-2 RNAemia with binary variables (Spearman correlation) and continuous variables (point-biserial correlation). Red indicates positive and blue negative correlation with *P* value <0.05. Abbreviations: Alb: albumin, ALP: alkaline phosphatase, ALT: alanine aminotransferase, Bil: bilirubin, COPD: Chronic obstructive pulmonary disease, Crea: Creatinine, CRP: C-reactive protein, DM: Diabetes, Hct: Hematocrit, Hb: Hemoglobin, HR: Heart rate, HTN: Hypertension, Lymphoc: Lymphocytes, MAP: Mean arterial pressure, Monoc: Monocytes, Neutroph: Neutrophils, K^+: Potassium, Resp. rate: Respiratory rate, Na^+: Sodium, Temp: Body temperature, WCC: White cell count. **d,** Anti-SARS-CoV-2 spike IgG and (**d**) anti-SARS-CoV-2 neutralization response (**d**) based on days post onset of symptoms (POS) in patients who tested positive (red) or negative (blue) for plasma/serum SARS-CoV-2 RNA within the first six ICU days. Lines show fitted Generalized Additive Models (GAM) with grey bands indicating the 95% interval of trust, correcting for age and sex. **e**, Anti-SARS-CoV-2 spike IgG levels (**e**) and anti-SARS-CoV-2 neutralization capacity (**e**) in individual samples positive or negative for SARS-CoV-2 RNA. Significance was determined through the Mann-Whitney U test. *P* values are corrected for age, sex and days POS.

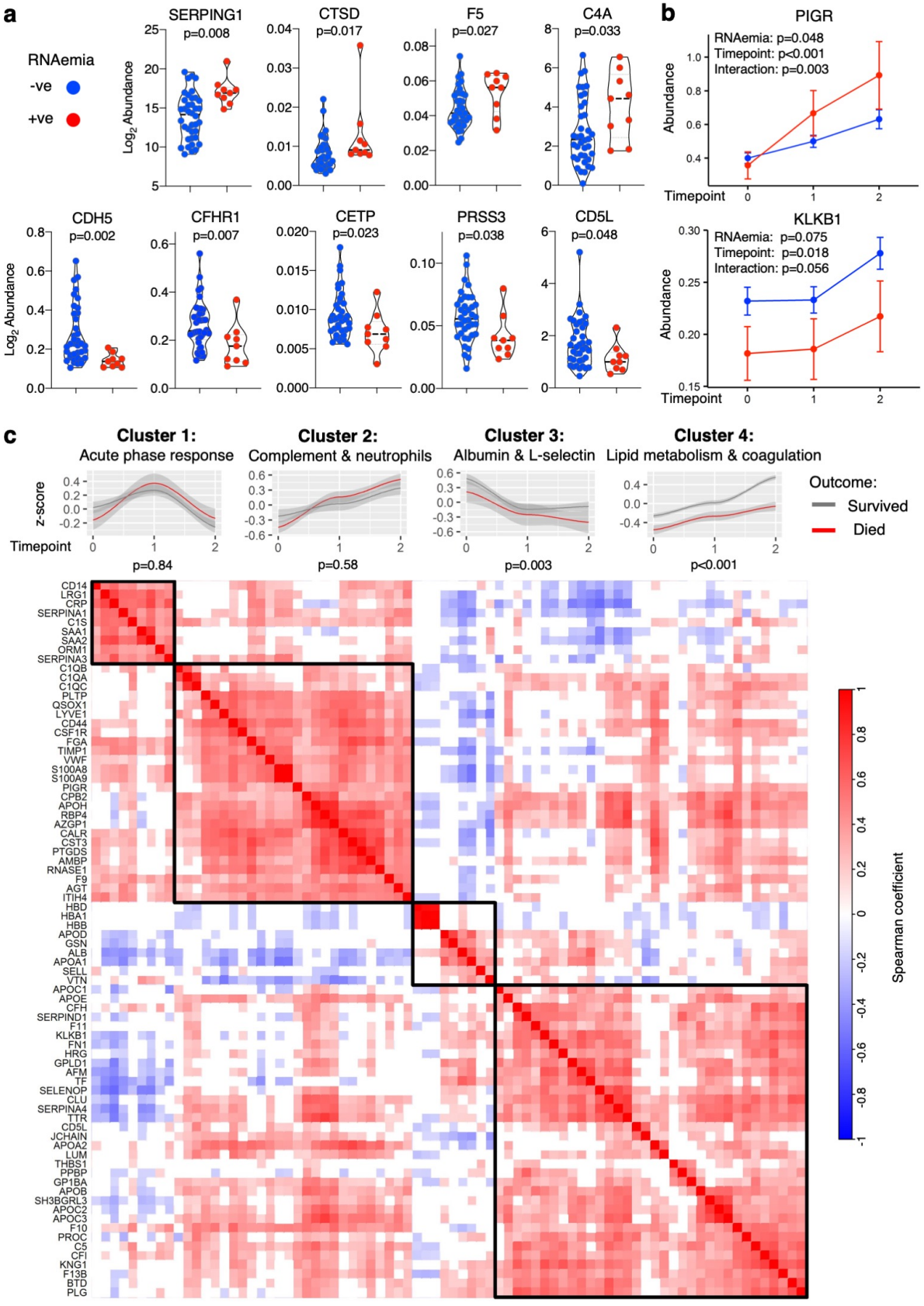
Fig. 3


a
Discovery: KCH Cohort
Control Patients
ICU Sepsis
ICU COVID-19
-2
+7
z-score
c
GSTT Cohort
ICU COVID-19
ALB
APOA1
APOC1
C1R
C1S
C6
C8A
C9
CD14
CFB
CFH
CLEC3B
CPN1
CRP
F5
FGA
FGG
GSN
HP
ITIH3
ITIH4
LBP
LGALS3BP
LRG1
LUM
PGLYRP2
SERPINA3
SERPING1
TF
p<0.05
p>0.05
0
1
2
3
4
Hazard Ratio
(Corrected for Age and Sex)
b
Discovery: KCH Cohort
COVID-19: Enriched plasma protein cluster
Complement, classical
Platelet degranulation
Acute-phase response
Innate immunity
Complement, alternative
Coagulation
Proteolysis
Cytolysis
Endocytosis
Plasminogen
Fibrinolysis
15/47
13/47
9/47
16/47
6/47
8/47
11/47
4/47
6/47
3/47
3/47
10-5
10-10
10-15
10-20
q-Value
d
GSTT Cohort
ICU COVID-19
C9
F13B
F7
FN1
MBL2
PLG
PROC
GC
p<0.05
p>0.05
0
1
2
3
4
Hazard Ratio
(Corrected for Age and Sex)

823

Fig. 3. COVID-19 circulating proteome signature and associations with 28-day mortality. a, Plasma proteome profiling was conducted using a data-independent acquisition-mass spectrometry (DIA-MS) approach with spiked standards for 500 proteins. Hierarchical cluster analysis was conducted upon significantly changing plasma proteins across control patients before elective cardiac surgery (n = 30), ICU patients with sepsis (n = 12) and ICU patients with COVID-19 (n = 12, KCH). The heatmap highlights 47 proteins enriched in COVID-19. Kruskal-Wallis, BH correction $q<0.05$. **b**, Gene ontology enrichment analysis was conducted upon these 47 proteins and significantly enriched pathways are represented. **c**, 29 common proteins cross-referenced against two published proteomic studies, exploring protein biomarkers of COVID-19 severity. The ability of these 29 proteins to predict 28-day mortality was explored in an independent ICU patient cohort (GSTT) by DIA-MS, and hazard ratio plots are shown. **d**, Proteomic analysis by DIA-MS conducted upon the serum samples of the GSTT COVID-19 ICU cohort returned additional biomarker candidates that predict 28-day mortality. Significance was determined through the Kruskal-Wallis test with Benjamini and Hochberg's FDR correction.

Fig. 4


a
RNAemia
-ve
+ve
SERPING1
p=0.008
CTSD
p=0.017
F5
p=0.027
C4A
p=0.033
CDH5
p=0.002
CFHR1
p=0.007
CETP
p=0.023
PRSS3
p=0.038
CD5L
p=0.048
Log2 Abundance
b
PIGR
RNAemia: p=0.048
Timepoint: p<0.001
Interaction: p=0.003
KLKB1
RNAemia: p=0.075
Timepoint: p=0.018
Interaction: p=0.056
Abundance
Timepoint
c
Cluster 1:
Acute phase response
Cluster 2:
Complement & neutrophils
Cluster 3:
Albumin & L-selectin
Cluster 4:
Lipid metabolism & coagulation
z-score
Timepoint
p=0.84
p=0.58
p=0.003
p<0.001
Outcome:
Survived
Died
CD14
LRG1
CRP
SERPINA1
C1S
SAA1
SAA2
ORM1
SERPINA3
C1QB
C1QA
C1QC
PLTP
QSOX1
LYVE1
CD44
CSF1R
FGA
TIMP1
VWF
S100A8
S100A9
PIGR
CPB2
APOH
RBP4
AZGP1
CALR
CST3
PTGDS
AMBP
RNASE1
F9
AGT
ITIH4
HBD
HBA1
HBB
APOD
GSN
ALB
APOA1
SELL
VTN
APOC1
APOE
CFH
SERPIND1
F11
KLKB1
FN1
HRG
GPLD1
AFM
TF
SELENOP
CLU
SERPINA4
TTR
CD5L
JCHAIN
APOA2
LUM
THBS1
PPBP
GP1BA
APOB
SH3BGRL3
APOC2
APOC3
F10
PROC
C5
CFI
KNG1
F13B
BTD
PLG
Spearman coefficient

838

Fig. 4. Circulating protein changes associated with SARS-CoV-2 RNAemia status over time. a, DIA-MS analysis upon serum samples from the GSTT COVID-19 ICU cohort was used to determine proteins that associate with the presence of SARS-CoV-2 RNAemia. Proteins that were significantly associated with RNAemia at baseline are individually represented as violin plots. Significance was determined through the Limma linear model analysis using Benjamini and Hochberg's FDR correction. Abbreviations: CDH5, cadherin-5 or VE-cadherin; CFHR1, complement factor H-related protein 1; SERPING1, plasma protease C1 inhibitor; CTSD, cathepsin D; CETP, cholesteryl ester transfer protein; F5, coagulation factor 5; C4A, complement factor 4a; PRSS3, trypsin-3; CD5a, CD5 antigen-like. **b**, Proteins with significantly different trajectories over time (baseline, week 1 – time point 1, week 2 – time point 2) between RNAemia positive and negative patients. PIGR, polymeric immunoglobulin receptor; KLKB1, kallikrein B1. **c,** Serial serum samples from COVID-19 ICU patients (GSTT, baseline, week 1 and week 2) were analyzed by DIA-MS to determine protein changes over time in ICU. The heat map represents a hierarchical cluster analysis conducted upon a Spearman correlation network of significantly changing proteins over time in ICU. Comparison of the trajectories of protein clusters in COVID-19 ICU patients based on 28-day mortality. Gene ontology enrichment analysis was used to determine functional pathways associated with the distinct protein clusters identified. Listed are the protein clusters that show a significant change between 28-day survivors (grey) and non-survivors (red) - and having a significant interaction with time points (baseline, week 1 – time point 1, week 2 – time point 2). Lines show fitted Generalized Additive Models (GAM) with grey bands indicating the 95% interval of trust. *P* values represent the significance of the outcome term in a fitted GAM model when correcting for age and sex.

Fig. 5

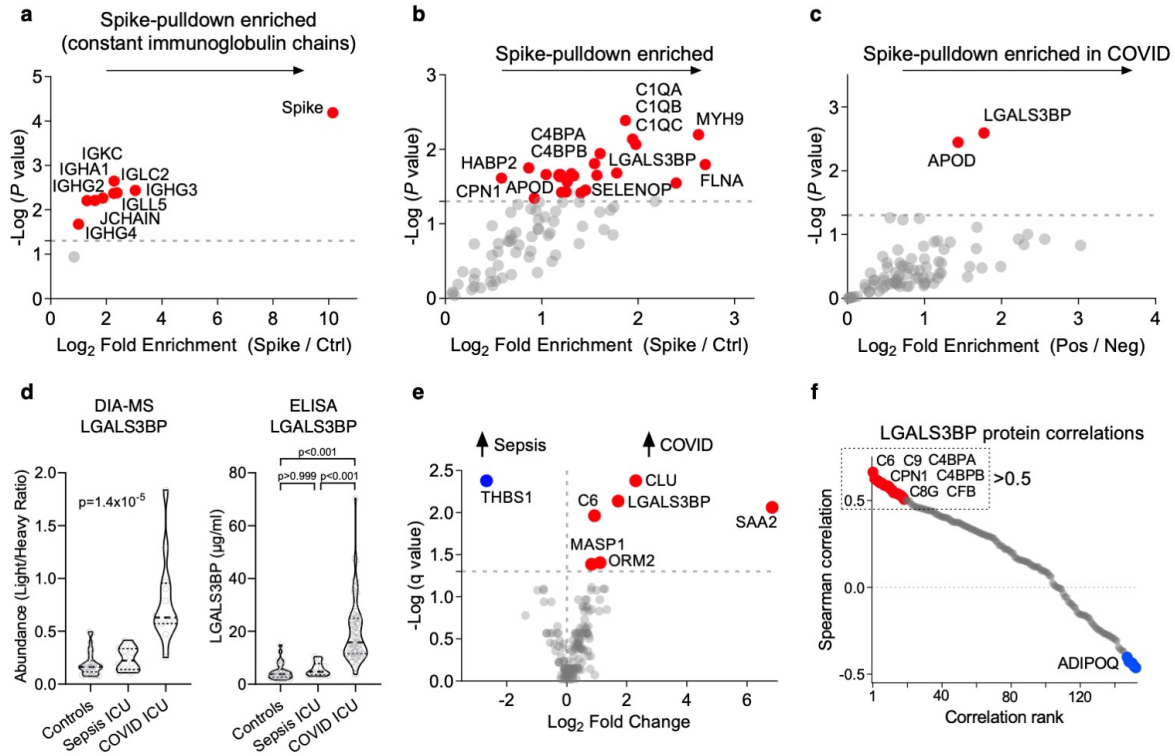
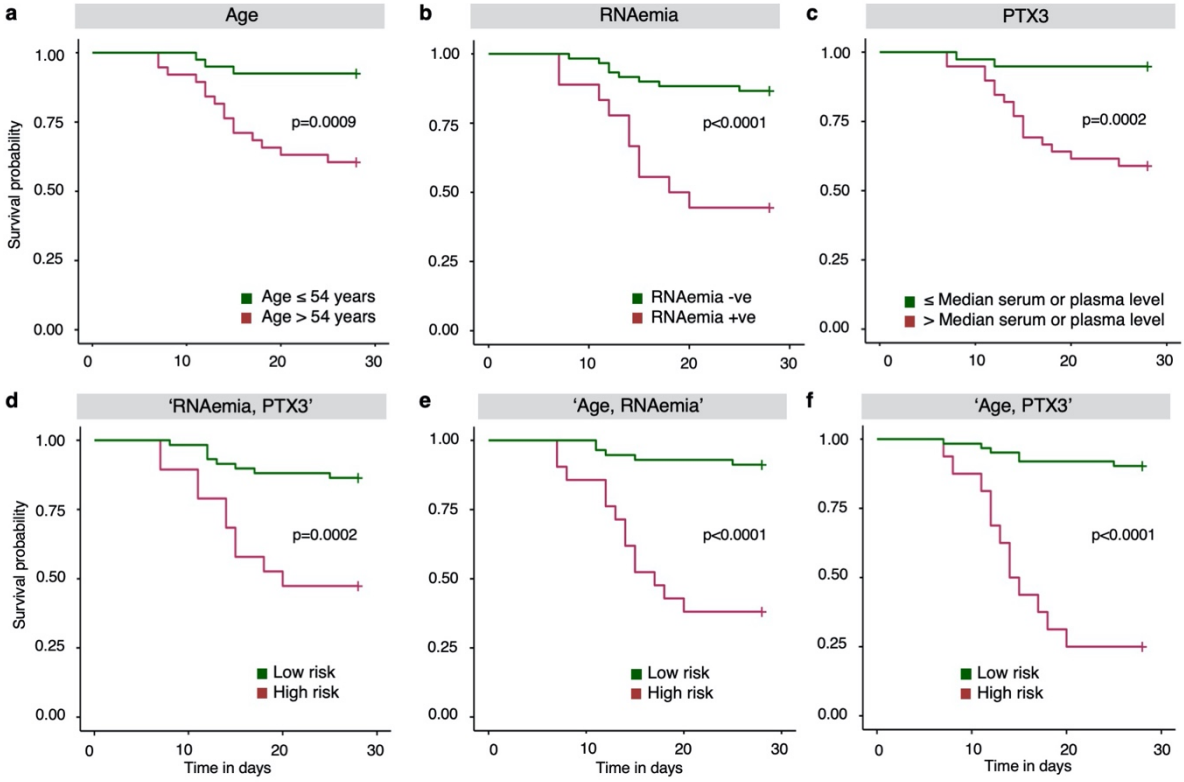


Fig. 5. LGALS3BP interacts with SARS-CoV-2 spike glycoprotein. a, Magnetic bead-based affinity isolation of binding partners using His-tagged SARS-CoV-2 spike glycoprotein as a bait for proteins in SARS-CoV-2-positive patient plasma. **b**, Volcano plot with significantly enriched proteins. **c**, Comparison of SARS-CoV-2 spike glycoprotein pull-down using plasma from COVID-19 ICU patients and non-COVID-19 sepsis ICU patients. Significance was determined by Student's t-test. **d,** LGALS3BP levels across three patient cohorts as determined by DIA-MS or ELISA: control patients before undergoing elective cardiac surgery (n = 30), pre-pandemic sepsis ICU patients (n = 12) and COVID-19 ICU patients (n = 74). Kruskal-Wallis and Dunn's multiple comparisons test were used to determine statistical significance. **e**, Volcano plot representing protein changes between baseline plasma samples from patients in ICU with either sepsis or COVID-19. Significance was determined through the Mann-Whitney U test with Benjamini and Hochberg's FDR correction. **f**, Plasma proteins correlating to LGALS3BP after age and sex corrections are highlighted by a Spearman correlation matrix across the proteomic dataset. Proteins with a Spearman correlation coefficient greater than 0.5 were used for Gene ontology pathway enrichment analysis (Supplementary Fig. 5).

Fig. 6



876

877 **Fig. 6. SARS-CoV-2 mortality prediction using machine learning.** **a**, Kaplan-Meier plot for
878 age (using median age of 54 years). **b**, Kaplan-Meier plot for SARS-COV-2 RNAemia. As a single
879 predictor, RNAemia provides the best stratification for survival. **c**, Kaplan-Meier plot for PTX3 using
880 median of serum or plasma. **d, e, f**, Kaplan-Meier plots for 'RNAemia, PTX3', 'Age, RNAemia' and
881 'Age, PTX3' combined in SVM RBF machine learning model. The machine learning model selected
882 binary combinations of 'Age, RNAemia' and 'Age, PTX3' as the best predictors.

883

Figures

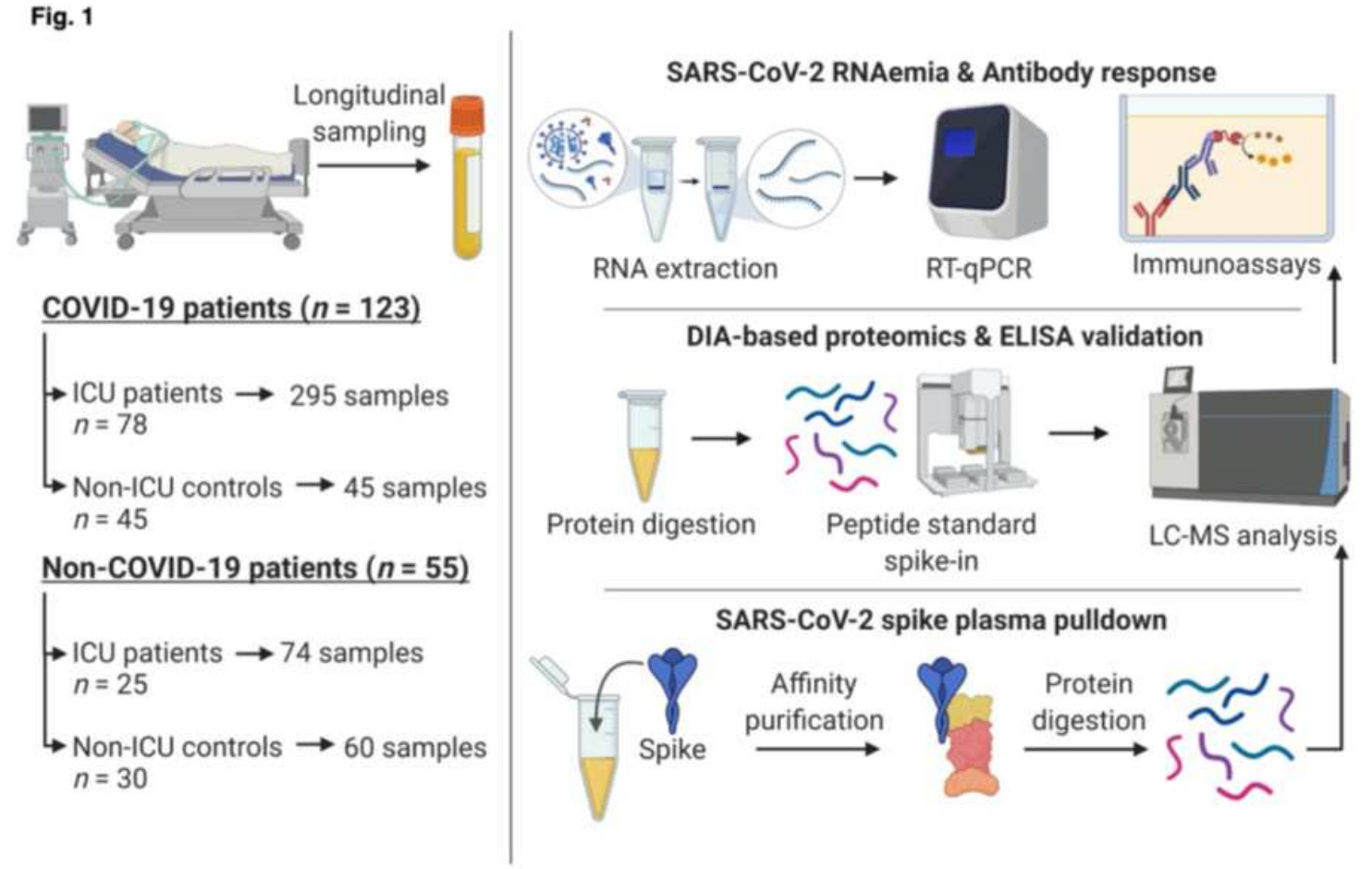


Figure 1

Schematic of study design. Plasma and serum samples were obtained from multiple patient cohorts across two UK-based university hospitals, including 123 COVID-19 patients: 78 SARS-CoV-2 positive patients in ICU were sampled at multiple time points over a 2-week period and compared to hospitalized non-ICU SARS-CoV-2 positive patients (n = 45). We used non-COVID-19 ICU patients (n = 25) and patients before and after undergoing elective cardiac surgery (n = 30) as controls. Patient samples were assessed for SARS-CoV-2 RNAemia, antibody responses and protein changes in the circulation. Finally, plasma protein interactions with SARS-CoV-2 spike glycoprotein were determined using a pull-down assay followed by mass spectrometry analysis.

Fig. 2

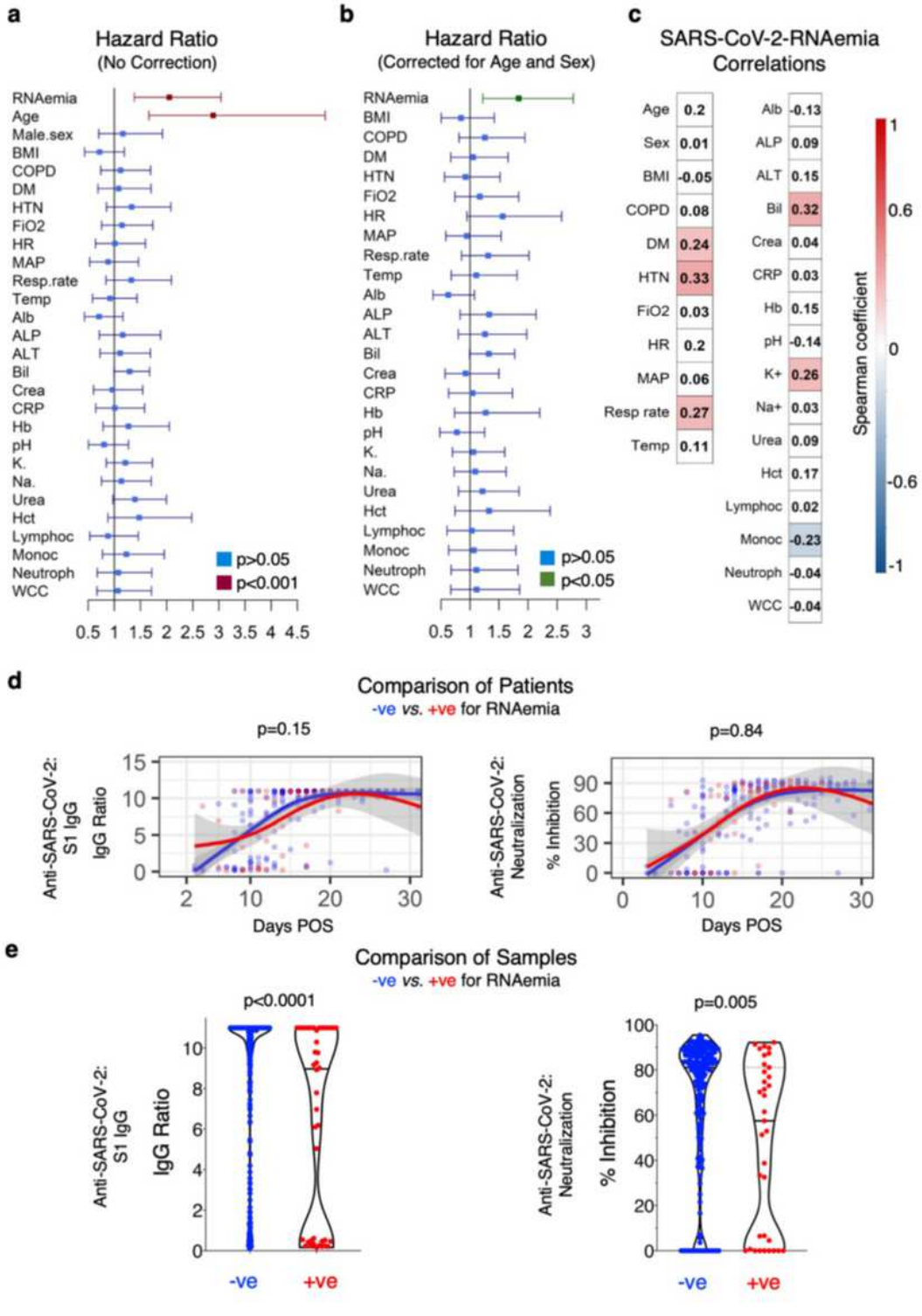


Figure 2

SARS-CoV-2 RNAemia and the humoral immune response. a, Unadjusted hazard ratios based on two ICU patient cohorts (KCH and GSTT, 60 survivors and 18 non-survivors). Green indicates P value <0.05, maroon indicates P value <0.001 and blue indicates P value >0.05. b, Hazard ratios after adjustment for age and sex. c, Association of SARS-CoV-2 RNAemia with binary variables (Spearman correlation) and continuous variables (point-biserial correlation). Red indicates positive and blue negative correlation with

P value <0.05. Abbreviations: Alb: albumin, ALP: alkaline phosphatase, ALT: alanine aminotransferase, Bil: bilirubin, COPD: Chronic obstructive pulmonary disease, Crea: Creatinine, CRP: C-reactive protein, DM: Diabetes, Hct: Hematocrit, Hb: Hemoglobin, HR: Heart rate, HTN: Hypertension, Lymphoc: Lymphocytes, MAP: Mean arterial pressure, Monoc: Monocytes, Neutroph: Neutrophils, K+: Potassium, Resp. rate: Respiratory rate, Na+: Sodium, Temp: Body temperature, WCC: White cell count. d, Anti-SARS-CoV-2 spike IgG and (d) anti-SARS-CoV-2 neutralization response (d) based on days post onset of symptoms (POS) in patients who tested positive (red) or negative (blue) for plasma/serum SARS-CoV-2 RNA within the first six ICU days. Lines show fitted Generalized Additive Models (GAM) with grey bands indicating the 95% interval of trust, correcting for age and sex. e, Anti-SARS-CoV-2 spike IgG levels (e) and anti-SARS-CoV-2 neutralization capacity (e) in individual samples positive or negative for SARS-CoV-2 RNA. Significance was determined through the Mann-Whitney U test. P values are corrected for age, sex and days POS.

Fig. 3

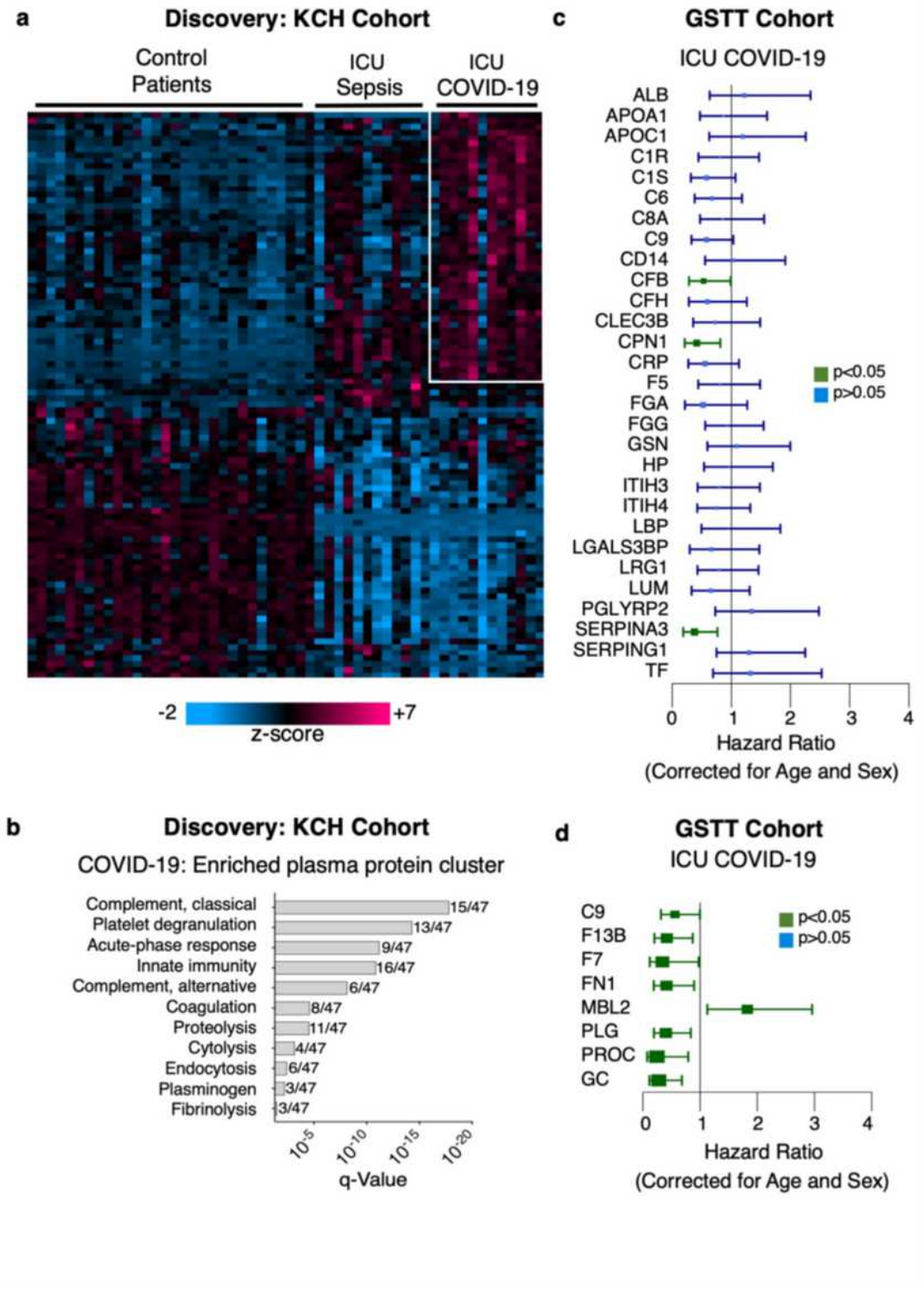


Figure 3

COVID-19 circulating proteome signature and associations with 28-day mortality. a, Plasma proteome profiling was conducted using a data-independent acquisition-mass spectrometry (DIA-MS) approach with spiked standards for 500 proteins. Hierarchical cluster analysis was conducted upon significantly changing plasma proteins across control patients before elective cardiac surgery (n = 30), ICU patients with sepsis (n = 12) and ICU patients with COVID-19 (n = 12, KCH). The heatmap highlights 47 proteins

enriched in COVID-19. Kruskal-Wallis, BH correction $q<0.05$. b, Gene ontology enrichment analysis was conducted upon these 47 proteins and significantly enriched pathways are represented. c, 29 common proteins cross-referenced against two published proteomic studies, exploring protein biomarkers of COVID-19 severity. The ability of these 29 proteins to predict 28-day mortality was explored in an independent ICU patient cohort (GSTT) by DIA-MS, and hazard ratio plots are shown. d, Proteomic analysis by DIA-MS conducted upon the serum samples of the GSTT COVID-19 ICU cohort returned additional biomarker candidates that predict 28-day mortality. Significance was determined through the Kruskal-Wallis test with Benjamini and Hochberg's FDR correction.

Fig. 4

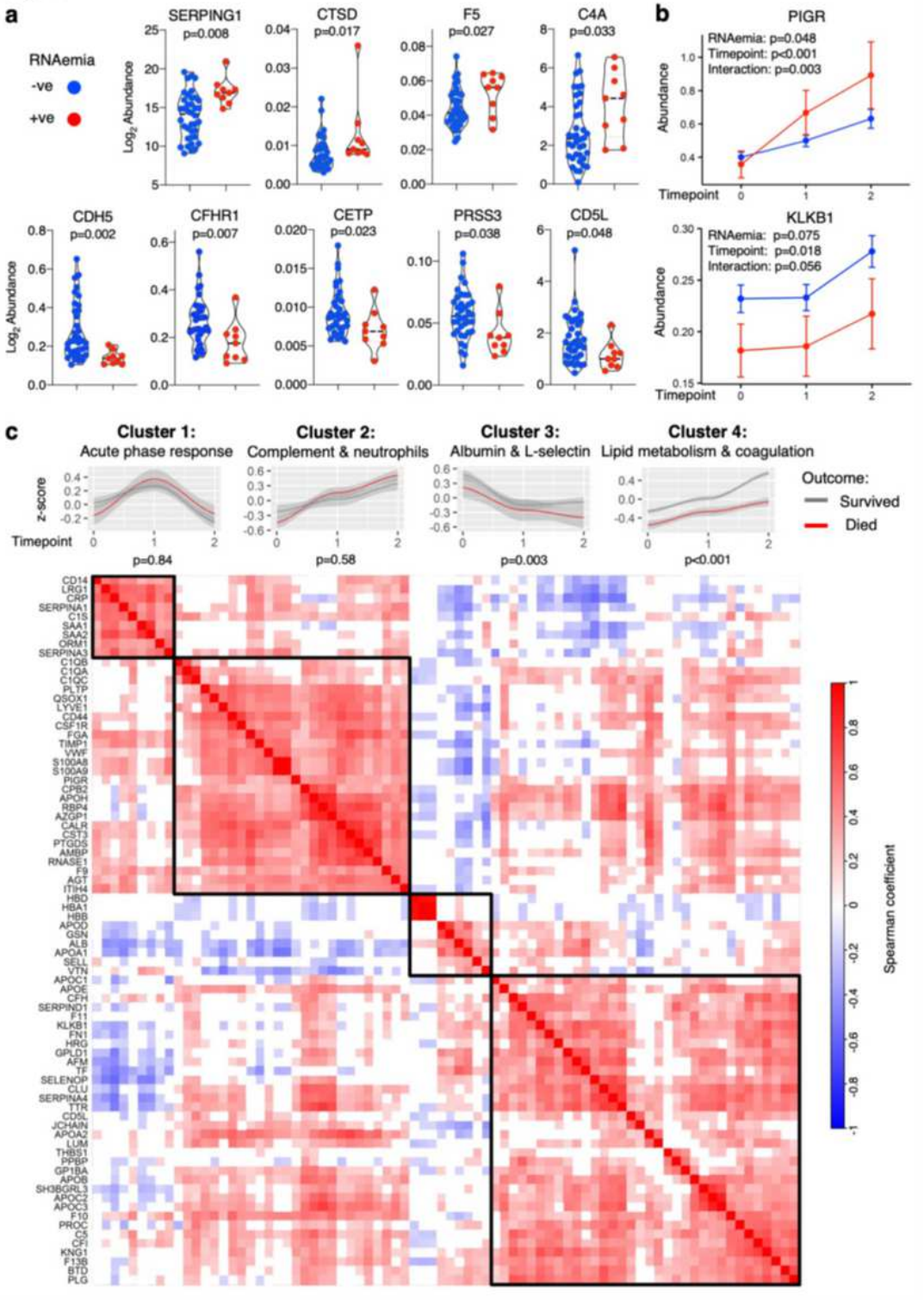


Figure 4

Circulating protein changes associated with SARS-CoV-2 RNAemia status over time. a, DIA-MS analysis upon serum samples from the GSTT COVID-19 ICU cohort was used to determine proteins that associate with the presence of SARS-CoV-2 RNAemia. Proteins that were significantly associated with RNAemia at baseline are individually represented as violin plots. Significance was determined through the Limma linear model analysis using Benjamini and Hochberg's FDR correction. Abbreviations: CDH5, cadherin-5 or

VE-cadherin; CFHR1, complement factor H-related protein 1; SERPING1, plasma protease C1 inhibitor; CTSD, cathepsin D; CETP, cholesteryl ester transfer protein; F5, coagulation factor 5; C4A, complement factor 4a; PRSS3, trypsin-3; CD5a, CD5 antigen-like. b, Proteins with significantly different trajectories over time (baseline, week 1 – time point 1, week 2 – time point 2) between RNAemia positive and negative patients. PIGR, polymeric immunoglobulin receptor; KLKB1, kallikrein B1. c, Serial serum samples from COVID-19 ICU patients (GSTT, baseline, week 1 and week 2) were analyzed by DIA-MS to determine protein changes over time in ICU. The heat map represents a hierarchical cluster analysis conducted upon a Spearman correlation network of significantly changing proteins over time in ICU. Comparison of the trajectories of protein clusters in COVID-19 ICU patients based on 28-day mortality. Gene ontology enrichment analysis was used to determine functional pathways associated with the distinct protein clusters identified. Listed are the protein clusters that show a significant change between 28-day survivors (grey) and non-survivors (red) - and having a significant interaction with time points (baseline, week 1 – time point 1, week 2 – time point 2). Lines show fitted Generalized Additive Models (GAM) with grey bands indicating the 95% interval of trust. P values represent the significance of the outcome term in a fitted GAM model when correcting for age and sex.

Fig. 5

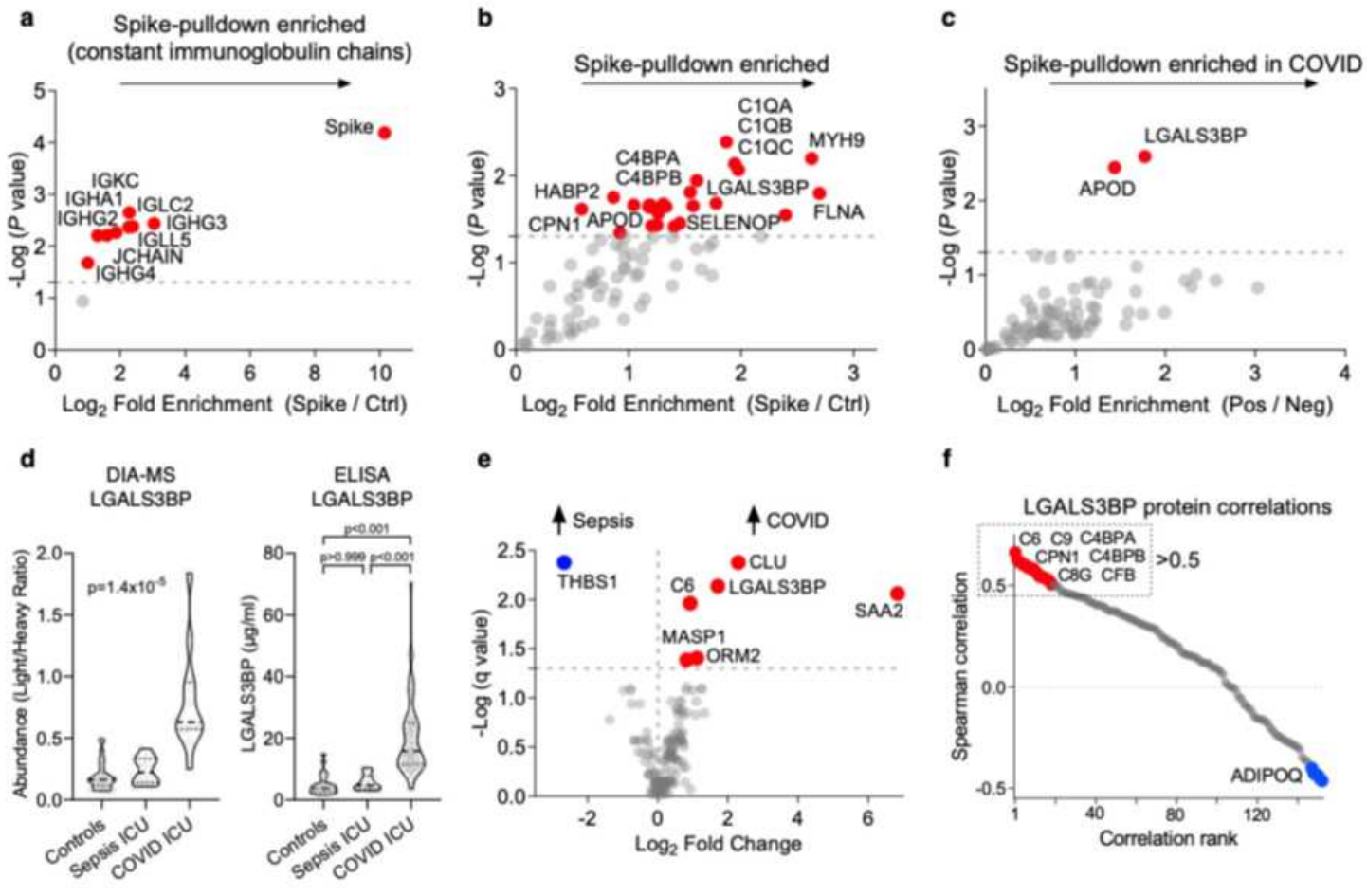


Figure 5

LGALS3BP interacts with SARS-CoV-2 spike glycoprotein. a, Magnetic bead-based affinity isolation of binding partners using His-tagged SARS-CoV-2 spike glycoprotein as a bait for proteins in SARS-CoV-2-positive patient plasma. b, Volcano plot with significantly enriched proteins. c, Comparison of SARS-CoV-2 spike glycoprotein pull-down using plasma from COVID-19 ICU patients and non-COVID-19 sepsis ICU patients. Significance was determined by Student's t-test. d, LGALS3BP levels across three patient cohorts as determined by DIA-MS or ELISA: control patients before undergoing elective cardiac surgery (n = 30), pre-pandemic sepsis ICU patients (n = 12) and COVID-19 ICU patients (n = 74). Kruskal-Wallis and Dunn's multiple comparisons test were used to determine statistical significance. e, Volcano plot representing protein changes between baseline plasma samples from patients in ICU with either sepsis or COVID-19. Significance was determined through the Mann-Whitney U test with Benjamini and Hochberg's FDR correction. f, Plasma proteins correlating to LGALS3BP after age and sex corrections are highlighted by a Spearman correlation matrix across the proteomic dataset. Proteins with a Spearman correlation coefficient greater than 0.5 were used for Gene ontology pathway enrichment analysis (Supplementary Fig. 5).

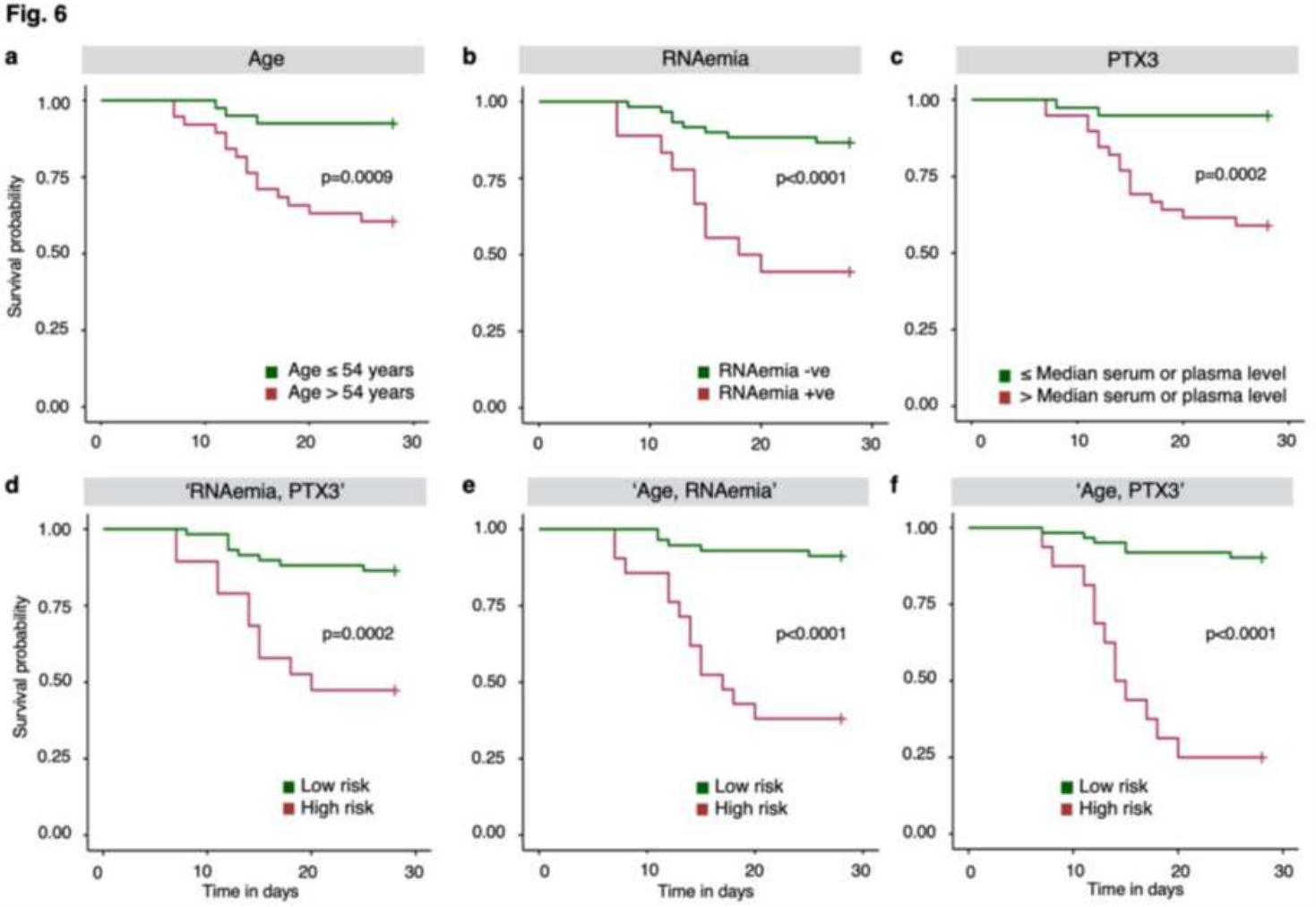


Figure 6

SARS-CoV-2 mortality prediction using machine learning. a, Kaplan-Meier plot for age (using median age of 54 years). b, Kaplan-Meier plot for SARS-COV-2 RNAemia. As a single predictor, RNAemia provides the

best stratification for survival. c, Kaplan-Meier plot for PTX3 using median of serum or plasma. d, e, f, Kaplan-Meier plots for ‘RNAemia, PTX3’, ‘Age, RNAemia’ and ‘Age, PTX3’ combined in SVM RBF machine learning model. The machine learning model selected binary combinations of ‘Age, RNAemia’ and ‘Age, PTX3’ as the best predictors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationSARSCoV2RNAemiaandProteomics.pdf