The metabolic fingerprint of COVID-19 severity

Tim Dierckx¹, Jan van Elslande², Heli Salmela³, Bram Decru², Els Wauters⁴,⁵, Jan Gunst⁶, Yannick Van Herck⁷, the CONTAGIOUS-consortium⁸, Joost Wauters⁸, Björn Stessel⁹, and Pieter Vermeersch²

1 KU Leuven, Department of Microbiology and immunology, Rega Institute for Medical Research, Laboratory of Clinical and Epidemiological Virology, Leuven, Belgium

2 University Hospitals Leuven, Clinical Department of Laboratory Medicine and National Reference Centre for Respiratory Pathogens, Leuven, Belgium

3 Nightingale Health Ltd, Helsinki, Finland

4 KU Leuven, Department of Chronic Diseases and Metabolism, Laboratory of Respiratory Diseases and Thoracic Surgery (BREATHE), Leuven, Belgium

5 University Hospitals Leuven, Department of Pneumology, Leuven, Belgium

6 KU Leuven, Department of Cellular and Molecular Medicine, Clinical Department and Laboratory of Intensive Care Medicine, Leuven, Belgium

7 KU Leuven, Department of Oncology, Laboratory of Experimental Oncology, Leuven, Belgium

8 University Hospitals Leuven, Medical Intensive Care Unit, Laboratory for Clinical Infectious and Inflammatory Diseases, Leuven, Belgium

9 Jessa Hospital, Department of Anaesthesiology and Pain Medicine, Hasselt, Belgium

* To whom correspondence should be addressed: tim.dierckx@kuleuven.be


NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Corona virus disease 2019 (COVID-19) has been associated with a wide range of divergent pathologies, and risk of severe disease is reported to be increased by a similarly broad range of co-morbidities. The present study investigated blood metabolites in order to elucidate how infection with severe acute respiratory syndrome coronavirus 2 can lead to such a variety of pathologies and what common ground they share. COVID-19 patient blood samples were taken at hospital admission in two Belgian patient cohorts, and a third cohort that included longitudinal samples was used for additional validation (total n=581). A total of 251 blood metabolite measures and ratios were assessed using nuclear magnetic resonance spectroscopy and tested for association to disease severity. In line with the varied effects of severe COVID-19, the range of severity-associated biomarkers was equally broad and included increased inflammatory markers (glycoprotein acetylation), amino acid concentrations (increased leucine and phenylalanine), increased lipoprotein particle concentrations (except those of very low density lipoprotein, VLDL), decreased cholesterol levels (except in large HDL and VLDL), increased triglyceride levels (only in IDL and LDL), fatty acid levels (decreased poly-unsaturated fatty acid, increased mono-unsaturated fatty acid) and decreased choline concentration, with association sizes comparable to those of routine clinical chemistry metrics of acute inflammation. Our results point to systemic metabolic biomarkers for COVID-19 severity that make strong targets for further fundamental research into its pathology (e.g. phenylalanine and omega-6 fatty acids).
Introduction

Severe corona virus disease 2019 (COVID-19), the disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, is associated with a wide range of pathologies including ARDS\textsuperscript{1,2}, lymphocytopenia\textsuperscript{3}, acute kidney failure\textsuperscript{4}, liver injury\textsuperscript{5} and septic shock\textsuperscript{3,4,6}. A similarly wide range of co-morbidities predisposes patients to more severe COVID-19, and includes obesity, diabetics, hypertension and cardiovascular disease (CVD)\textsuperscript{7–9}. Studies examining the effects of SARS-CoV-2 infection have implicated a wide variety of immunological processes (reviewed in\textsuperscript{10}). SARS-CoV-2 uses the angiotensin-converting enzyme 2 receptor (ACE2) as a cell entry point, which could at least partially explain the diverse effects of SARS-CoV-2 infection, as ACE2 is present on many different cell types. ACE2 has been linked to CVD, inflammation and the metabolism through the renin-angiotensin system (RAS)\textsuperscript{11}, and it has been shown to regulate amino acid homeostasis\textsuperscript{12}. The populations with increased risk of developing severe COVID-19 share low-grade inflammation as a common feature\textsuperscript{13} and the major risk factor for severe COVID-19 can be interpreted as metabolic dysregulation\textsuperscript{14}. A clear need exists for a systemic investigation into the metabolic processes underlying the diverse immunological pathways affected by SARS-CoV-2 to better understand 1) the factors associated with risk of severe infection, 2) the systemic effects of COVID-19, and 3) the cardiometabolic complications following disease resolution.

Limited reports investigating specific metabolites have already been published. These include the association to disease severity of lower levels of low density lipoprotein (LDL)\textsuperscript{15}, and cholesterol levels in high density lipoprotein (HDL) and LDL\textsuperscript{16}. However, traditional HDL, LDL and cholesterol quantification methods used in clinical settings lack the resolution to
analytically capture the lipoprotein profile. Nuclear magnetic resonance (NMR) spectroscopy can be used to address this limitation, and enables a more systemic investigation by simultaneously quantifying a large number of metrics relevant to the COVID-19 setting including energy-related metabolites, amino acid and fatty acid concentrations and glycoprotein acetylation (GlycA).

GlycA is a relatively novel serum biomarker for systemic inflammation, known to be increased in various conditions associated with a predisposition to severe COVID-19, including CVD, type-2 diabetes and elevated BMI. The GlycA NMR signal quantifies a non-specific, composite signal that arises from mobile N-acetyl methyl groups of circulating glycoproteins (predominantly alpha-1-acid glycoprotein, alpha-1-antitrypsin, haptoglobin and transferrin). These glycoproteins are acute-phase reactants, meaning their glycan structures and their concentrations change during the acute-phase response of inflammation. However, GlycA has also shown distinct advantages over traditional inflammatory biomarkers such as C-reactive protein (CRP) in its association to inflammatory conditions. Due to its more persistent and composite nature, being a read-out for many different inflammatory processes, GlycA can be interpreted as a summary measure of a subject’s ‘baseline’ chronic inflammatory burden. In light of the common theme of low-grade inflammation in populations with higher risk for severe COVID-19 and reports of increased CVD risk during and following COVID-19, GlycA is of particular relevance in the COVID-19 setting.

To elucidate how SARS-CoV-2 infection can lead to such a variety of pathologies and what common ground these pathologies share, we used NMR experiments to quantify chronic inflammatory burden, lipoprotein particle concentrations (as well as their cholesterol and
triglyceride content), amino acid and fatty acid concentrations, and energy related metabolites in three sample sets from confirmed SARS-CoV-2 infected patients in two Belgian hospitals, collected at time of hospital admission for COVID-19. We identified biomarkers associated with COVID-19 severity (mild/moderate/severe/critical) and test which of these biomarkers could be used to identify patients who will develop critical disease during their hospital stay. For a limited set of patients, we explore longitudinal data from samples taken at day 7 post admission, at hospital discharge and 30 days post hospital discharge.

**Methods**

**Ethics**

This study was approved by the Medical Ethical Committee of the University Hospitals of Leuven (MEC, UZ Leuven, s64161 and s63881) and the ethical approval committee of the Jessa hospital (20.75-20.05). Samples from UZL and Jessa cohorts were collected retrospectively, while the CONTAGIOUS cohort study is a prospective observational trial for which informed consent was obtained for all participants (clinicaltrials.gov identifier NCT04327570).

**Study populations**

This study included 581 samples from 480 patients across three different cohorts. A graphical overview of the sample sets is available in *Supplementary Figure S1*. Two of the cohorts (UZL, n=219 and JESSA, n=164) consisted of convenience samples of left-over diagnostic samples of unselected adult patients. All COVID-19 patients from the first wave in Belgium (March-June 2020) for whom a left-over sample from routine biochemistry at the time of hospital presentation was available were included. All included patients tested
positive for SARS-CoV-2 using real-time polymerase chain reaction, performed in accordance with WHO guidelines\textsuperscript{35}. Positive PCR results were obtained at time of admission sampling (median 0 days, range -6 to 2 days for the UZL cohort, range -8 to 0 days for the JESSA cohort, respectively). The UZL cohort consisted of plasma samples, while the samples assayed in the JESSA cohort were serum samples. The third sample set is a subset of plasma samples (n=198, from 97 patients) taken for the CONTAGIOUS observational clinical trial. All consecutive patients (>18 years old) admitted to the University Hospitals of Leuven between March 2020 and September 2020 with PCR-confirmed and/or CT-confirmed SARS-CoV-2 disease (confirmation a median of 2 days prior to admission, range 12 days to 1 day prior to admission) were eligible for inclusion. Samples were taken at the time of admission (within maximum 48h), at day 7, at the time of hospital discharge and 30 days after hospital discharge (if available). Patients for whom admission samples were available within the prospective CONTAGIOUS cohort were excluded from the retrospective UZL cohort.

\textit{Clinical measurements}

Severity of infection was retrospectively classified on a scale of 1-4 by a single clinician reviewing individual patient records using the following criteria\textsuperscript{36,37}: 1) Mild cases: mild clinical symptoms without manifestation of pneumonia on imaging throughout the duration of disease, 2) Moderate cases: fever, respiratory symptoms, and with radiological findings of pneumonia at any point during the disease, 3) Severe cases: meeting any one of the following criteria during the course of their disease: respiratory distress, hypoxia (SpO\textsubscript{2} \leq 93\%), or abnormal blood gas analysis (PaO\textsubscript{2} <60 mmHg, PaCO\textsubscript{2} >50 mmHg) or 4) Critical cases: meeting any one of the following criteria during the course of their disease:
respiratory failure requiring mechanical ventilation, shock, organ failure that requires intensive unit care or death of the patient due to COVID-19.

In the clinical laboratory of the University Hospital Leuven (UZL patient cohort), routine laboratory parameters were measured using Roche Cobas 8000 (Roche, Basel, Switzerland) for clinical chemistry parameters and immunoassays, Sysmex XE 5000 (Sysmex, Kobe, Japan) for hematology parameters and ACL-TOP 700 (Werfen, Milan, Italy) for coagulation tests. In the clinical laboratory of the Jessa hospital, laboratory parameters were quantified using Roche Cobas 8000 (Roche, Basel, Switzerland) for CRP, electrolytes, kidney function, liver function, triglycerides and ferritin, SYMEX XN-1500 for hematology parameters, and ACL-TOP 550 (Werfen, Milan, Italy) for coagulation tests.

**NMR metabolomics**

Nightingale Health Ltd’s (Helsinki, Finland) high-throughput 1H NMR metabolomics platform was used to quantify 251 biomarkers and biomarker ratios. The experimental protocols have been described elsewhere\(^{18,20}\). The quantified biomarkers include amino acids, lipoprotein, fatty acids (FA) and triglyceride (TG), in addition to GlycA, creatinine and albumin concentration, as well as ketone bodies and glycolysis related metabolites. The full list of biomarkers quantified by NMR is available in **Supplementary Table S1**. In addition to directly measured metabolic biomarkers, we also included a composite multi-biomarker ‘Infectious Disease Score’ (IDS), calculated as posited by Julkunen et al. (under review\(^{38}\)). Briefly, this risk score was constructed to associate to the likelihood of severe pneumonia in eight years of follow-up, using NMR measurements taken in biobank samples.

**Statistics**
Differences in demographic and clinical measurements between the study populations were assessed using a T-test or Mann-Whitney U test, for normally and non-normally distributed variables respectively, as indicated by a Shapiro-Wilk test for normality. Differences between the datasets of categorical demographic variables were identified using Fisher's exact test on the contingency tables. Association to severity is reported as odds ratios in univariate cumulative logit ordinal logistic regression on standard deviation (SD) scaled variables. Where the proportional odds assumption was violated, as indicated by the Brant test, the ordinal regression was performed three times, i.e. once for each contrast in the severity scale: contrasting 1) mild versus moderate, severe and critical, 2) mild and moderate versus severe and critical, and finally 3) mild, moderate and severe versus critical. Correction for age and sex had minimal impact on the reported associations (data not shown).

Biomarkers with significant association to COVID-19 severity (Benjamini-Hochberg (BH) adjusted p-value<0.05 in both the UZL and Jessa cohorts), were tested for significant (BH adjusted p-value<0.05) differences between severe and critical cases in a third independent cohort (CONTAGIOUS) using a Mann-Whitney U test or Fischer test (for continuous and categorical variables, respectively). All statistical testing was conducted in R v4.0.2 (additional packages: ordinal, brant, ggplot).

Results

Study populations and their differences

We first tested whether there were any significant clinical differences between the two largest patient cohorts (UZL, n=219 and JESSA, n=164, Table 1). The demographics of the cohorts were largely comparable, with the exception of slightly higher prevalences of diabetes and CVD in the UZL cohort. Several clinical laboratory measurements taken at
admission were also different between the two cohorts: average white blood cell count, neutrophil count and eosinophil count, as well as CRP levels, ferritin levels and activated partial thromboplastin time were significantly higher in the Jessa cohort, whereas slightly higher sodium, potassium and calcium concentrations are noted in the UZL cohort. Some of these differences could be explained by differences in laboratory settings (e.g. UZL quantified total calcium, while Jessa quantified ionized calcium) or patient follow-up (UZL did not consistently quantify ferritin in every sample). We then aimed to identify which biomarkers consistently associated to COVID-19 severity, despite these differences between the study cohorts.

**Identifying metabolic biomarkers for COVID-19 severity**

Using ordinal regression on sd-scaled clinical and NMR measurements, 130 biomarkers were found to be significantly associated to disease severity in both the UZL and Jessa patient cohorts (full results in Supplementary Table S2). At admission, only eight of the measured demographic and clinical biomarkers were found to be significantly associated to COVID-19 severity in both patient cohorts (age, CVD status, O₂ saturation, creatinine and eGFR, CRP, and lymphocyte and eosinophil count) (Figure 1). Some demographic parameters expected to associate to severity based on previous reports (e.g. sex, BMI and diabetic status), trended towards association but fell short of statistical significance in both cohorts simultaneously.

Of the 251 metabolites and metabolite ratios quantified by NMR, 122 associated with COVID-19 severity in both cohorts. In line with the clinical associations, NMR evidence of inflammation (GlycA, albumin, and the GlycA/albumin ratio) and kidney function (creatinine) showed strong associations with COVID-19 severity (Figure 2). These metrics are important
components of Julkunen et al’s ‘Infectious Disease Score’ (IDS), which can partly explain the IDS’s strong association to COVID-19 severity (odds ratios of 4 and 4.5 for the UZL and Jessa cohort, respectively). Of the 9 quantified amino acids, phenylalanine (Phe) and leucine (Leu) show consistent association to severity. Valine and total branched-chain amino acid concentrations were associated to severity in the UZL cohort, but not the Jessa cohort. Other notable associations only observed in a single cohort include glutamine (Gln), tyrosine (Tyr) and Histidine (His) (Figure 3).

Concentrations and content of analytical lipoprotein subclasses associated well with COVID-19 severity. Overall, higher lipoprotein concentrations and higher cholesterol content was associated with less severe disease, with the notable exception of VLDL particles and of the larger HDL particles. Higher triglyceride content of lipoprotein particles was associated with disease severity, but only for IDL and LDL (Figure 4). While increased absolute phospholipid content in HDL, IDL and LDL is associated with less severe disease, this contrasted to the relative phospholipid to total lipid ratios in HDL, IDL and LDL, which were associated with higher disease severity. Relative lipoprotein content ratios (including relative cholesterol, cholesteryl ester, free cholesterol, and triglyceride to total lipid ratios) otherwise showed similar association to severity as their absolute concentrations (Supplementary Table S2).

Quantification of FA showed increased poly-unsaturated FA (PUFA) content was associated with less severe disease, while increased mono-unsaturated FA (MUFA) content was associated to more severe disease (Figure 5). Linoleic acid (LA) and total Omega-6 FA showed stronger and more consistent associations with COVID-19 severity than Omega-3 FA. Both total Omega-3 FA and docosahexaenoic acid (DHA) showed opposite associations in the
two examined cohorts (OR 0.72 and 0.80 in UZL, OR 1.2 and 1.4 in Jessa, for Omega-3 FA and DHA ratios to total FA, respectively).

**Validation in CONTAGIOUS cohort**

To further validate our results, we tested which of the 130 biomarkers that were consistently associated with severity showed significant differences between severe and critical disease severity classes in admission samples (n=72) from a third, independent, patient cohort (CONTAGIOUS, clinical and demographic parameters at admission, comparisons to UZL and Jessa cohorts, **Supplementary Table S3**). In total, 72 severity-associated biomarkers (5 clinical, 68 NMR) showed statistically significant differences at admission between patients with severe and critical disease (n=48 and n=24, respectively), with strong correlations between some of the quantified moieties (**Figure 6**). Selected biomarkers are presented in **Figure 7**: albumin, GlycA/albumin ratio and IDS retained their association to disease in the CONTAGIOUS setting that does not include patients with mild or moderate disease. Phenylalanine was the only amino acid that retained its significance in this context. Omega-6 FA, PUFA and PUFA/MUFA ratio findings were consistent with our earlier observations. Again, HDL cholesterol content was significantly lower at admission in the more severe COVID-19 cases, and this difference was only evident in the small HDL sizes.

**Longitudinal associations**

For the CONTAGIOUS cohort of severe and critically afflicted patients (n=97), longitudinal samples were available taken at day 7 post admission (D7, n=38), at hospital discharge (Dis, n=41), and 30 days post discharge (Dis+30, n=47). Compared to samples taken at admission, most severity-associated biomarkers worsened at D7, and return to admission levels at time of hospital discharge (full results in **Supplementary Table S4**). At 30 days post
discharge, clinical biomarkers for severity (CRP, Ferritin) returned to healthy levels (Figure 8).
Likewise, GlycA as well as the GlycA/albumin ratio dropped below admission levels by day 30 post hospital discharge and returned to concentrations observed in healthy controls reported in demographic studies.\textsuperscript{38,42} Choline, phosphatidylcholine and phosphoglyceride concentrations, which were significantly associated to severity (Figure 5), showed no change during the first seven days of hospitalization and up until discharge, but were found to be significantly increased 30 days post discharge.

Discussion

We performed the first in-depth characterization of circulating metabolites in SARS-CoV-2 infected patients across the spectrum of disease severity. In left-over diagnostic samples taken at time of hospital admission of two independent cohorts (n=219 and n=164), we identified a metabolic profile associated to COVID-19 severity. This metabolic fingerprint of COVID-19 is relevant to four aspects of COVID-19 related research: first, we provide further metabolic evidence of the association between COVID-19 severity and inflammation. Second, our amino acid profiling confirms observations of altered energy requirements of hyperinflammatory immune cells of COVID-19 patients\textsuperscript{43,44} and suggests an alternative explanation for the low occurrence of severe COVID-19 cases in phenylketonuria patients\textsuperscript{45}. Third, we refine existing reports on the association between lipoprotein particles, their cholesterol content, and COVID-19 severity\textsuperscript{15,16} by investigating size-specific lipoprotein subclasses and their content. Finally, we add to the literature relating COVID-19 severity to fatty acid concentrations in the blood, and observe an association with linoleic acid that is in-line with mechanistic evidence\textsuperscript{46}.
Metabolic evidence of inflammation, i.e. increased GlycA levels and reduced albumin levels, was associated with increased severity (Figure 2). Many populations reported as high-risk for developing severe COVID-19 share low-grade chronic inflammation as a common feature. Increased GlycA concentrations have been reported in populations with co-morbidities related to increased risk of severe COVID-19, such as advanced age, diabetes, obesity, and CVD. GlycA association to severity is further strengthened by calculating the GlycA to Albumin ratio (Figure 2). While routine clinical chemistry metrics of acute inflammation such as CRP and ferritin showed stronger associations to severity than GlycA (Figure 1), GlycA has higher basal healthy control levels, higher longitudinal stability compared to CRP, and is associated with inflammatory disease severity even in CRP negative patients, which could be of when investigating the longitudinal effects of COVID-19 (Figure 8).

In line with reports on reduced mitochondrial oxidative phosphorylation in hyperinflammatory immune cells of COVID-19 patients, our results confirmed higher serum pyruvate concentration to be associated with increased COVID-19 severity (Figure 2). The observed associations of amino-acid concentrations with severity could be interpreted in a similar light (Figure 3): the lower serum glutamine (GLN) concentration in severe COVID-19 could represent increased GLN catabolism in visceral tissues to accommodate the energy requirements of hyperinflammatory cells through anaplerosis. Our findings suggest that in mild cases, ongoing branched chain amino acid (BCAA) oxidation contributes to sufficient availability of GLN, while in severe cases BCAAs accumulate in serum as their oxidation is inhibited in favor of cytosolic aerobic glycolysis. Other than leucine, phenylalanine also showed a consistently significant association with COVID-19 severity (with borderline significance observed for tyrosine, Figure 3). Phenylalanine’s association to
severity is in line with reports of low occurrence of severe COVID-19 outcomes among phenylketonuria patients, which was previously ascribed to this population’s lower rates of vitamin D deficiency as a result of the protein substitutes in their diet45.

Our results confirm reports of low lipoprotein particle concentrations of LDL15, HDL, and their cholesterol content16 in severe COVID-19. However, these early reports were limited both in the scope of the considered lipoprotein particles (only examining HDL and LDL) and in the degree to which these particles are characterized (only quantifying their cholesterol content). The 1H NMR panel assayed in the present study comprehensively characterizes the range of lipoprotein particle densities and sizes, and provides an in-depth profile of each of the particle subclasses. We expand on existing observations to show that not all HDL sizes associate to severity: very large and large HDL particles, and their lipid and cholesterol content, showed no significant association to severity (Figure 4), while small HDL was among the strongest and most consistent associations to COVID-19 severity. Decreased small HDL has been reported to be associated with higher all-cause mortality risk51, though these findings contrast to associations noted in CVD studies. In the CVD context, large HDL is inversely associated to CVD risk, while small HDL typically shows no such association52.

Similarly opposite to what is reported in the CVD context, our findings show increased LDL concentration, as well as its phospholipid and cholesterol content, is associated with decreased severity. Triglyceride associations to COVID-19 severity are similar to what’s observed in CVD and diabetes53: high TG levels in IDL, LDL and VLDL and low TG levels in HDL, though this latter association does not reach significance in our cohorts (Figure 4).

While the altered cholesterol profile of severe COVID-19 can be interpreted as metabolic evidence of inflammation, it can also be construed as evidence of a genetic predisposition to
severe infection, similar to what's been shown for HDL cholesterol\textsuperscript{54}. Relevant to the hyperinflammatory state\textsuperscript{13} frequently identified as a “cytokine storm” in COVID-19\textsuperscript{55}, a meta-analysis of genome-wide association studies has shown a genetic link between cytokine activity and the serum concentration of apolipoprotein B, serum cholesterol and the cholesterol and phospholipid content of IDL, LDL and VLDL\textsuperscript{56}, all of which are shown to be lower in severe COVID-19 cases in our results. This interpretation in terms of a predisposition to severe infection finds additional support in the observed association at admission between severity and the blood biomarker score for infectious disease risk, which was constructed by Julkunen et al. from biobank samples to associate to the likelihood of severe pneumonia and severe COVID-19 during eight years of follow-up\textsuperscript{38}.

Higher relative PUFA content and PUFA to MUFA ratio were consistently associated with lower severity, in contrast to increased MUFA levels (Figure 5). Linoleic Acid concentration and total omega-6 fatty acid content (absolute concentration and relative to total FA content) were lower in severe COVID-19 cases in our cohorts. Conflicting reports exist on PUFA association with inflammation, and omega-6 FA have been linked to increased inflammation in the airways\textsuperscript{57}. Considering a recent report showed that linoleic acid (a PU omega-6 FA) tightly binds to the spike glycoprotein which is used by SARS-CoV-2 to bind the ACE2 receptor to gain entry into host cells\textsuperscript{46}, the association between less severe COVID-19 and higher omega-6 PUFA levels could therefore be a mechanistic effect, rather than an effect mediated through inflammation.

Interestingly, one recent report showed that the precursor for phosphatidylcholine, phosphocholine, was upregulated in severe COVID-19\textsuperscript{58}, while our results show increased phosphatidylcholine is associated with less severe COVID-19, as are serum
phosphoglycerides and choline levels, in agreement with other findings\(^\text{58}\) (Figure 5). Furthermore, analysis of longitudinal follow-up samples of severe and critical patients showed no changes in these concentrations during hospitalization (Figure 8), but a marked increase of these concentrations was observed in samples taken 30 days after hospital discharge. This, combined with the convenience with which the 1H NMR spectrum can be measured and how many of the captured biomarkers are relevant to COVID-19 severity, leads us to suggest that the NMR analysis technique is well-suited to prospective longitudinal studies aiming to elucidate the long-term effects of COVID-19.

We point out several strengths and limitations of our study. The studied populations had significant differences in demographics and clinical characteristics, and differences in sample type (serum vs plasma) between our two largest cohorts discourage the merging of their experimental results, even after accounting for the batch effects introduced by differences in sample handling between the different cohorts. However, this does lend a high degree of confidence to the results that were consistently observed in all three cohorts. Availability of longitudinal follow-up samples was limited to the CONTAGIOUS cohort of patients with severe infection. It is worth noting that, to our knowledge, this is the first study using 1H NMR metabolomics in severe viral infection. As such, without comparable results from studies on other infections, we cannot attribute the results reported in this study exclusively to SARS-CoV-2 infection. We suspect that e.g. a study into the metabolic associations of severe pneumonia could yield similar results, in light of the strength of the association of severity with the infectious disease multi-biomarker score from Julkunen et al\(^\text{38}\). Many of the associations to severity noted here could be interpreted as general signs of inflammation and robust correlation network analysis on paired NMR metabolomics, cytokine
quantification and transcriptomics, while contrasting severe COVID-19 to severe (viral) pneumonia without SARS-CoV-2 infection, would allow for a more fine-grained interpretation of the results.

COVID-19 is reported to have many diverse effects and complications, many of which can be related to low-grade inflammation. Our study of systemic metabolic biomarkers reveals that the metabolic fingerprint of COVID-19 severity consists of a pervasive inflammatory signature, not unlike reported associations with CVD risk or long-term risk of pneumonia. The broad range of biomarkers reflects the wide range of COVID-19 symptoms and co-morbidities, and these biomarkers should be further investigated for their translational potential, and their precise role in the COVID-19 pathology.

Acknowledgements

The authors thank Kimberly Vanhees, Cato Jacobs, Evy Smeyers and Laurien Geebelen and for their efforts in sample selection and preparation, and the procurement of clinical data.

Funding

TD is funded by a research grant by The Research Foundation – Flanders (FWO Vlaanderen, 12Y6320N). PV is a senior clinical investigator of the FWO-Vlaanderen. JG holds a postdoctoral research fellowship supported by Clinical Research and Education Council of the University Hospitals Leuven.

Conflicts of Interest Statements

HS is an employee of Nightingale Health Ltd.
References


diagnostics and perspectives – a report of the European Academy of Allergy and Clinical Immunology (EAACI). *Allergy* **all.14462** (2020). doi:10.1111/all.14462


56. Nath, A. P. et al. Multivariate Genome-wide Association Analysis of a Cytokine Network Reveals Variants with Widespread Immune, Haematological, and


Figures

Figure 1: Odds Ratios of clinical and demographic biomarkers associations with COVID-19 severity.

Association of demographic and clinical variables with disease severity in the Jessa (n=164, circle) and UZL cohort (n=219, triangle) tested using ordinal regression on sd-scaled variables. Odds ratios of proportional odds in each cohort are presented in purple. If odds ratios differ significantly between each severity class contrast, as judged by a Brant test, the OR is visualized separately for each contrast: red: Mild (severity 1) versus Non-mild (severity 2, 3 and 4), green: Mild and Moderate (severity 1 and 2) versus Severe and Critical (severity 3 and 4), and finally, blue: non-critical (severity 1, 2 and 3) versus critical (severity 4). Multiple-comparison corrected significance of each association is indicated by the opacity of each point (opaque: q-value < 0.05, clear: q-value > 0.05). ALT: alanine

Figure 2: Odds ratios of Inflammation-related NMR biomarker associations with COVID-19 severity.

In the Jessa (n=164, circle) and UZL cohort (n=219, triangle), associations of NMR biomarkers were tested using ordinal regression on sd-scaled variables. Odds ratios of proportional odds in each cohort are presented in purple. If odds ratios differ significantly between each severity class contrast, as judged by a Brant test, the OR is visualized separately for each contrast: red: Mild (severity 1) versus Non-mild (severity 2, 3 and 4), green: Mild and Moderate (severity 1 and 2) versus Severe and Critical (severity 3 and 4), and finally, blue: non-critical (severity 1, 2 and 3) versus critical (severity 4). Multiple-comparison corrected significance of each association is indicated by the opacity of each point (opaque: q-value <0.05, clear: q-value >0.05). GlycA: Glycoprotein Acetylation.
Figure 3: Odds ratios of amino-acid concentration associations with COVID-19 severity. In the Jessa (n=164, circle) and UZL cohort (n=219, triangle), associations of NMR biomarkers were tested using ordinal regression on sd-scaled variables. Odds ratios of proportional odds in each cohort are presented in purple. If odds ratios differ significantly between each severity class contrast, as judged by a Brant test, the OR is visualized separately for each contrast: red: Mild (severity 1) versus Non-mild (severity 2, 3 and 4), green: Mild and Moderate (severity 1 and 2) versus Severe and Critical (severity 3 and 4), and finally, blue: non-critical (severity 1, 2 and 3) versus critical (severity 4). Multiple-comparison corrected significance of each association is indicated by the opacity of each point (opaque: q-value < 0.05, clear: q-value > 0.05). BCAA: branched-chain amino acids.
Figure 4 Odds ratios of lipoprotein particle and lipoprotein particle content associations with COVID-19 severity. In the Jessa (n=164, circle) and UZL cohort (n=219, triangle), associations of NMR biomarkers were tested using ordinal regression on sd-scaled variables. Odds ratios of proportional odds in each cohort are presented in purple. If odds ratios differ significantly between each severity class contrast, as judged by a Brant test, the OR is visualized separately for each contrast: red: Mild (severity 1) versus Non-mild (severity 2, 3 and 4), green: Mild and Moderate (severity 1 and 2) versus Severe and Critical (severity 3 and 4), and finally, blue: non-critical (severity 1, 2 and 3) versus critical (severity 4). Multiple-comparison corrected significance of each association is indicated by the opacity of each point (opaque: q-value < 0.05, clear: q-value > 0.05). VLDL: very-low-density lipoprotein, IDL: intermediate-density lipoprotein, LDL: low-density lipoprotein, HDL: high-density lipoprotein. Lipoprotein subclass sizes are listed in Supplementary Table S1.
Figure 5: Odds ratios of COVID-19 severity associations of fatty acids and lipid-associated metabolites. In the Jessa (n=164, circle) and UZL cohort (n=219, triangle), associations of NMR biomarkers were tested using ordinal regression on sd-scaled variables. Odds ratios of proportional odds in each cohort are presented in purple. If odds ratios differ significantly between each severity class contrast, as judged by a Brant test, the OR is visualized separately for each contrast: red: Mild (severity 1) versus Non-mild (severity 2, 3 and 4), green: Mild and Moderate (severity 1 and 2) versus Severe and Critical (severity 3 and 4), and finally, blue: non-critical (severity 1, 2 and 3) versus critical (severity 4). Multiple-comparison corrected significance of each association is indicated by the opacity of each point (opaque: q-value < 0.05, clear: q-value > 0.05). FA: Fatty Acid. MUFA: mono-unsaturated FA, PUFA: poly-unsaturated FA.
Figure 6: Hierarchical clustering of Spearman correlations of COVID-19 severity associated biomarkers. All depicted biomarkers are significantly associated to severity in both the UZL and Jessa cohorts and show significant differences between severe and critical cases of the CONTAGIOUS patient cohort. Hierarchical clustering was achieved by average linkage of the absolute value of Spearman’s rho. C: cholesterol, CE cholesteryl ester, FC, free cholesterol, TG: triglyceride, FA: fatty acid, L: total lipid concentration, P: lipoprotein particle concentration, PL: phospholipid concentration, PUFA poly-unsaturated FA, MUFA: mono-unsaturated FA, VLDL: very-low-density lipoprotein, IDL: intermediate-density lipoprotein, LDL: low-density lipoprotein, HDL: high-density lipoprotein. Lipoprotein subclass sizes are listed in Supplementary Table S1.
Figure 7: Selected biomarkers in severe and critical cases of the CONTAGIOUS cohort. Severity 3 and 4, severe and critical, respectively. Where depicted, the y-axis on the right side of the graph normalizes the data using the mean and standard deviations observed in Finnish demographic cohorts\textsuperscript{42}. n.s. not significant, * p<0.05, ** p<0.01. CRP: c-reactive protein, FA: fatty acid, HDL: high-density lipoprotein, LDL, low-density lipoprotein.
Figure 8: Longitudinal analysis of severe and critical COVID patients of the CONTAGIOUS cohort.

Patients of the CONTAGIOUS cohort were sampled at COVID admission (A, red), 7 days post admission (Day7, green), time of hospital discharge (Dis, blue) and 30 days post hospital discharge.
Indicated are p-values of Mann-Whitney U tests comparing the indicated timepoints. Where depicted, the y-axis on the right side of the graph normalizes the data using the mean and standard deviations observed in Finnish demographic cohorts. Full comparison results for all variables associated to severity are available in supplementary table S4. ApoA1: Apolipoprotein A1, ApoB: Apolipoprotein B, CRP: c-reactive protein, GlycA: glycoprotein acetylation.
Tables

<table>
<thead>
<tr>
<th>Severity: total samples</th>
<th>UZL</th>
<th>Jessa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>219</td>
<td>67</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>154</td>
<td>26.5</td>
</tr>
<tr>
<td>Sex (% female)</td>
<td>154</td>
<td>48%</td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>219</td>
<td>27%</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>219</td>
<td>35%</td>
</tr>
<tr>
<td>CVD (%)</td>
<td>219</td>
<td>35%</td>
</tr>
<tr>
<td>Chronic Pulmonary Disease (%)</td>
<td>219</td>
<td>16%</td>
</tr>
<tr>
<td>Hemoglobin (g/dL)</td>
<td>217</td>
<td>13.3</td>
</tr>
<tr>
<td>Neutrophil (10⁹/L)</td>
<td>207</td>
<td>4.1</td>
</tr>
<tr>
<td>Eosinophil (10⁹/L)</td>
<td>207</td>
<td>0</td>
</tr>
<tr>
<td>Lymphocyte (10⁹/L)</td>
<td>208</td>
<td>1</td>
</tr>
<tr>
<td>Platelets (10⁹/L)</td>
<td>216</td>
<td>195</td>
</tr>
<tr>
<td>APTT (s)</td>
<td>145</td>
<td>27.8</td>
</tr>
<tr>
<td>Fibrinogen (g/L)</td>
<td>38</td>
<td>4.02</td>
</tr>
<tr>
<td>D-dimer (µg/L)</td>
<td>70</td>
<td>624.5</td>
</tr>
<tr>
<td>CRP (mg/L)</td>
<td>213</td>
<td>40.8</td>
</tr>
<tr>
<td>Sodium (mmol/L)</td>
<td>213</td>
<td>138.5</td>
</tr>
<tr>
<td>Potassium (mmol/L)</td>
<td>214</td>
<td>4.01</td>
</tr>
<tr>
<td>Creatinine (mg/dL)</td>
<td>215</td>
<td>0.55</td>
</tr>
<tr>
<td>eGFR (mL/min/1.73m²)</td>
<td>213</td>
<td>74</td>
</tr>
<tr>
<td>Calcium (mmol/L)</td>
<td>186</td>
<td>2.29</td>
</tr>
<tr>
<td>Glucose (mg/dL)</td>
<td>193</td>
<td>115</td>
</tr>
<tr>
<td>HbA1c (% ratio)</td>
<td>116</td>
<td>6.05</td>
</tr>
<tr>
<td>Total Bilirubin (mg/dL)</td>
<td>198</td>
<td>0.465</td>
</tr>
<tr>
<td>Alkaline Phosphatase (U/L)</td>
<td>210</td>
<td>70</td>
</tr>
<tr>
<td>Gamma GT (U/L)</td>
<td>211</td>
<td>40</td>
</tr>
<tr>
<td>AST (U/L)</td>
<td>213</td>
<td>32</td>
</tr>
<tr>
<td>ALT (U/L)</td>
<td>214</td>
<td>24</td>
</tr>
<tr>
<td>Creatine Kinase (U/L)</td>
<td>133</td>
<td>86</td>
</tr>
<tr>
<td>Troponin (µg/L)</td>
<td>141</td>
<td>0.016</td>
</tr>
<tr>
<td>Ferritin (µg/L)</td>
<td>66</td>
<td>327.5</td>
</tr>
</tbody>
</table>

Table 1: Demographic and clinical comparison of study populations. Variables with significant differences between the study populations are indicated in bold. CVD: cardiovascular disease, WBC: white blood cell count, aPTT: activated partial thromboplastin time, CRP: c-reactive protein, eGFR: estimated glomerular filtration rate, HBA1c: glycated hemoglobin, GammaGT: gamma-glutamyltransferase, AST: aspartate transaminase, ALT: alanine transaminase.