ORIGINAL ARTICLE

Untargeted metabolomics of COVID-19 patient serum reveals potential prognostic markers of both severity and outcome

Ivayla Roberts[1] · Marina Wright Muelas[1] · Joseph M. Taylor[2] · Andrew S. Davison[2] · Yun Xu[1,3] · Justine M. Grixti[1] · Nigel Gotts[1,3] · Anatolii Sorokin[1] · Royston Goodacre[1,3] · Douglas B. Kell[1,4]

Abstract

Introduction The diagnosis of COVID-19 is normally based on the qualitative detection of viral nucleic acid sequences. Properties of the host response are not measured but are key in determining outcome. Although metabolic profiles are well suited to capture host state, most metabolomics studies are either underpowered, measure only a restricted subset of metabolites, compare infected individuals against uninfected control cohorts that are not suitably matched, or do not provide a compact predictive model.

Objectives Here we provide a well-powered, untargeted metabolomics assessment of 120 COVID-19 patient samples acquired at hospital admission. The study aims to predict the patient's infection severity (i.e., mild or severe) and potential outcome (i.e., discharged or deceased).

Methods High resolution untargeted UHPLC-MS/MS analysis was performed on patient serum using both positive and negative ionization modes. A subset of 20 intermediary metabolites predictive of severity or outcome were selected based on univariate statistical significance and a multiple predictor Bayesian logistic regression model was created.

Results The predictors were selected for their relevant biological function and include deoxycytidine and ureidopropionate (indirectly reflecting viral load), kynurenine (reflecting host inflammatory response), and multiple short chain acylcarnitines (energy metabolism) among others. Currently, this approach predicts outcome and severity with a Monte Carlo cross validated area under the ROC curve of 0.792 (SD 0.09) and 0.793 (SD 0.08), respectively. A blind validation study on an additional 90 patients predicted outcome and severity at ROC AUC of 0.83 (CI 0.74–0.91) and 0.76 (CI 0.67–0.86).

Conclusion Prognostic tests based on the markers discussed in this paper could allow improvement in the planning of COVID-19 patient treatment.

Keywords COVID-19 · Serum · Untargeted metabolomics · UHPLC-MS/MS · LC–MS/MS Orbitrap

---

✉ Marina Wright Muelas
m.wright-muelas@liverpool.ac.uk

✉ Douglas B. Kell
douglas.kell@liverpool.ac.uk

1   Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

2   Department of Clinical Biochemistry and Metabolic Medicine, Liverpool Clinical Laboratories, Royal Liverpool University Hospitals Trust, Liverpool, UK

3   Centre for Metabolomics Research (CMR), Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

4   Novo Nordisk Foundation Centre for Biosustainability, Technical University of Denmark, Building 220, Chemitorvet, 2000 Kgs Lyngby, Denmark

1 Introduction

The severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) outbreak started in Wuhan, China, in 2019, and quickly resulted in a worldwide pandemic, challenging healthcare systems with the need to provide intensive care to a previously inconceivable number of patients (Bennet et al., 2020). SARS-CoV-2 presents with a wide range of symptoms. These range from minor, unspecific ones, including anosmia, a dry persistent cough, fever, diarrhoea, in certain cases combined with mild pneumonia, to more severe, potentially life-threatening symptoms, such as severe pneumonia with dyspnoea, tachypnoea and disturbed gas exchange. Approximately 5% of severely infected patients develop lung dysfunction, requiring ventilation, and shock or

🖄 Springer

multiple organ failure (Marietta et al., 2020; Wu & McGoogan, 2020). In some cases, symptoms remain for an extended period ('long COVID').

The reasons behind the wide variability in individual responses to COVID-19, i.e., the illness resulting from SARS-CoV-2 infection, are still poorly understood, though some appear to involve interferon responses (Arunachalam et al., 2020; Hadjadj et al., 2020; Zhang et al., 2020). Much research, and evidence from the clinic, points towards the idea that severe complications in COVID-19 arise through a vasculopathy and coagulopathy elicited by infection rather than via the typical inflammatory responses normally observed in acute respiratory distress syndrome or cytokine release storms (Fox et al., 2020; Grobler et al., 2020; Leisman et al., 2020; Libby & Lüscher, 2020; Paranjpe et al., 2020; Pretorius et al., 2020; Zheng et al., 2020). Prognostic scores attempt to transform complex clinical pictures into tangible numerical values. However, many of these novel COVID-19 prognostic scores have been found to have a high risk of bias, possibly reflecting the fact that they have been developed in small cohorts, and many have been published without clear details of model derivation and testing (Knight et al., 2020; Wynants et al., 2020).

Understanding changes in the biochemistry of an individual who is ostensibly healthy (Dunn et al., 2011, 2015), including when they may show no overt symptoms of infection with SARS-CoV-2, remains a huge challenge (Alene et al., 2021). Similar questions apply to understanding who is likely to survive (unaided or via intervention) and who is likely to die from COVID-19 once diagnosed.

For fundamental reasons, the metabolome is a more sensitive indicator of the biochemical status of a cell or organism than is a proteome or a transcriptome (Fuhrer et al., 2017; Kell & Oliver, 2016; Oliver et al., 1998; Raamsdonk et al., 2001). Consequently, metabolomic analyses of patient samples promise to enable understanding of biochemical changes in relation to poorly understood processes (Johnson et al., 2016), for instance as recently shown in human frailty in ageing populations (Rattray et al., 2019). Metabolomics studies measure the effects on the host and not simply the presence of the infecting agent, therefore, such studies could provide a set of markers that can be of significant use for rapid tests, complementary to current polymerase chain reaction (PCR) or antibody tests, for confirmation of SARS-CoV-2 infection, disease severity, and potential outcome.

A continuously increasing number of studies have applied metabolomics to investigate COVID-19 in human patients; a selection of which is presented in Supplementary information Table S1. Most of the studies have highlighted disruption of lipid metabolism (López-Hernández et al., 2021; Overmyer et al., 2020; Thomas et al., 2020), along with tryptophan metabolism in relation to inflammation (Ansone et al., 2021; Blasco et al., 2020; López-Hernández et al.,

2021; Overmyer et al., 2020; Sindelar et al., 2021; Thomas et al., 2020) and changes in pyrimidine metabolism (Blasco et al., 2020) as metabolic features of COVID-19 patients against controls. Some of these studies were clearly limited by patient numbers and may therefore not be fully representative of the variation in human responses to COVID-19, and some lacked proper statistical tests see (Broadhurst & Kell, 2006). Moreover, many of these markers of COVID-19 infection, severity and outcome have also been described in patients with sepsis and acute respiratory failure (Migaud et al., 2020). Most importantly, many of the early studies simply compare patients with healthy controls, which, given that the disease status is in fact known (and current diagnostics very rapid), does not of itself have either diagnostic or prognostic value. However, more recent studies with larger cohorts such as (López-Hernández et al., 2021; Sindelar et al., 2021) have also considered more interesting aspects such as severity and included longitudinal follow up of the infection. Since presence of the disease is already known, our work here is focused on this more pertinent question: can the metabolome distinguish patients with COVID-19 in terms of either disease severity (mild/severe) or outcome (deceased/survived)?

To that end, we adopted an untargeted metabolomics approach using UHPLC-MS/MS to cohorts of serum samples collected at the Royal Liverpool University Hospital (RLUH). 120 patient samples were obtained at the time of admission and diagnosis with COVID-19. This study enabled us to identify prognostic biomarkers of both COVID-19 severity (severe *vs* mild) and outcome. Subsequently, the findings were validated in a blind study on 90 additional patients. Moreover, longitudinal data were acquired from 28 severe patients with divergent outcome, i.e., 13 deceased and 15 discharged to explore further the temporal evolution of the selected predictors.

Here we present the study results via a set of different models. We first validate that untargeted metabolomics provides better severity and outcome distinction *vs* using solely demographics and clinical data, via multiblock chemometric analysis. A high dimensional predictive model is then presented, based on more than 900 metabolites, showing promising predictive performance. Next, with the goal of clinical application in mind, we present results from a predictive model based of 20 compounds selected largely on the basis of our confidence in the metabolites' identity and known biological function. We discuss in detail the approaches used to make this selection via pathway enrichment analysis and manual curation. Moreover, we explore (and largely discount) the possible confounding effects of demographic factors and underlying conditions of the selected metabolic predictors. Finally, we validate the predictive power of the selected compounds in a blind study of a new patient cohort.

2 Results

2.1 Discovery study

Serum from 120 patients was acquired at the Royal Liverpool University Hospital (RLUH) on the first positive SARS-CoV-2 test. Severity scoring was based on the level of respiratory support required and overall patient outcome where severe corresponds to required fraction of inspired oxygen $(FIO_2) > 40\%$ and/or required Continuous Positive Airway Pressure (CPAP) and/or required invasive ventilation and/or did not survive. Patients were also stratified using the 4C Mortality Score (Knight et al., 2020). Patient demographics grouped by severity and outcome are presented in Table 1, where mild cases encompass mild to moderate disease severity.

Untargeted UHPLC-MS analysis was performed on the patient serum samples with a ThermoFisher Scientific Vanquish UHPLC system coupled to a ThermoFisher Scientific Q-Exactive mass spectrometer (ThermoFisher Scientific, UK). Pre-processing of UHPLC-MS data in Compound Discoverer resulted in 5234 retained metabolic features (i.e., m/z and RT) in positive electrospray ionization (ESI+) and 2465 in negative electrospray ionization (ESI-). Features were excluded based on standard filters; i.e., removal of background features (Maximum Sample/Blank parameter to 5), features with coefficient of variation (CV) > 30% in the quality control (QC) samples' and features present in less than 80% of the QCs samples as described in (Wright Muelas et al., 2020). Additionally, metabolic features detected in less than 25% of the experimental samples were also excluded. We note that this final step allowed the removal of small number of metabolic features mostly related to drugs (and their metabolites) taken by the patients for pre-existing conditions, which would otherwise simply have complicated and confounded the analysis.

2.2 Exploratory analysis

Principal components analysis (PCA) was performed on both ESI+ and ESI− data separately and no clear clustering in severity or outcome could be observed (Fig. S1), indicating that a simple, unsupervised method such as PCA would fail to capture the differences in metabolic profiles related to COVID-19 infection and potential outcome in complex data resulting from UHPLC-MS analysis. In contrast, when PCA was applied to the metadata associated with the 120 patients which include gender, age, body mass index (BMI), pulse, temperature, blood pressure, respiratory rate (full list is provided in the methods section), a clearer separation between severe and mild patients was observed (Fig. S2A). This alone is not overly interesting because the severity

assessment was drawn from a few variables in these metadata. However, when the metadata and the UHPLC-MS data were analysed together by using a multiblock PCA (Smilde et al., 2005; Xu & Goodacre, 2012) the separation improved (Figure S2B), indicating that there is relevant discriminatory information within the UHPLC-MS data which could not be revealed by PCA had been discovered in a multiblock model with presence of a meta data block. These findings are also valid in fatal outcomes where no clear clustering is observed in the metadata PCA block (Fig. S2A). Individual block figures from the multiblock analysis are available in the supplementary information (Figure S3). Clearly PCA is not a classification technique and the lack or presence of clusters does not translate directly to any predictive potential of the data. To investigate if differences in metabolic profiles are capable of predicting the severity and outcome of COVID-19 infection, the next section explores in detail predictive models based on the UHPLC-MS data.

2.3 Predictive models

Four multi-predictor models were trained: extreme gradient boosted trees (Chen & Guestrin, 2016; Friedman, 2001), Lasso regularization with elastic net (Zou & Hastie, 2005), logistic regression, and Bayesian logistic regression (Goodrich, 2020). All these models, when trained on the complete data of > 7000 metabolic features, showed clear signs of overfitting despite the use of tuning based on cross-validation and regularization. To overcome this, the set of predictors was filtered based on individual significance as determined by volcano plot analysis (p-value and fold change) as illustrated in Figure S4 for outcome and Figure S5 for severity.

Significance filtering based on volcano plot analysis reduced the number of metabolic features to 1987/ 935 (ESI± respectively) for severity and/or outcome combined. For those features, signal curation based on chromatogram and spectral quality (see UHPLC-MS signal curation section in Methods), was performed in Compound Discoverer and a total of 526/409 (ESI± respectively) features were retained. Table S2 provides a breakdown of those metabolic features along with their identification level according to the Metabolomics Standards Initiative (MSI) (Sumner et al., 2007), ranging from MSI level 1 to 4.

Evaluation of the four predictive models on this reduced set of 935 ESI± metabolic features in total from ESI+ and ESI− identified that Bayesian logistic regression had the best generalization performance; therefore, all the following results are based on this model. The mean area under the curve (AUC) was found to be at 0.836 (SD 0.069) for severity and 0.807 (SD 0.081) for outcome, demonstrating good predictive power of the patient metabolome. However, a mass spectrometry-derived model with some 900 metabolic

Table 1 Patient demographics by severity and outcome

Condition	Severity			Outcome		
	Mild	Severe	P-v	Discharged	Deceased	P-v
N	71	49	–	89	31	–
4C score	8.39 (4.75–12)	12.87 (10.00–16)	***	8.91 (5.25–12)	14.34 (13.00–16)	***
Male	38 (54%)	32 (65%)		48 (54%)	22 (71%)	
Female	33 (46%)	17 (35%)		41 (46%)	9 (29%)	
Age	66.0 (54–77)	69.2 (56–87)		63.8 (52.0–77.0)	77.1 (65.5–88.5)	***
BMI	26.4 (21.8–28.6)	27.0 (23.0–32.0)		26.7 (22.3–29.0)	26.4 (21.9–29.6)	
O_2 support	32 (45%)	49 (100%)	***	50 (56%)	31 (100%)	***
Mechanical respiratory support	1 (1%)	12 (25%)	***	3 (10%)	10 (11%)	
Max FiO_2	24.7 (21–28)	71.6 (40–100)	***	33.3 (21–35)	74.2 (36–100)	***
Required O_2 on presentation	17 (24%)	39 (80%)	***	34 (38%)	22 (71%)	**
FiO_2 (%) on presentation	22.8 (21–21)	47.0 (28–60)	***	28.2 (21.0–28)	45.5 (22.5–60)	**
NEWS	2.56 (1–4)	7.27 (5–10)	***	3.47 (1–6.0)	7.40 (6–10.8)	***
Hypertension	28 (39%)	26 (53%)		35 (39%)	19 (61%)	*
Cardiac disease	20 (28%)	17 (35%)		25 (28%)	12 (39%)	
Kidney disease	24 (34%)	19 (39%)		28 (31%)	15 (48%)	
eGFR	67.7 (49–90)	56.5 (30–90)	*	67.7 (48.0–90.0)	49.9 (26.5–68.5)	**
Liver disease	7 (9.9%)	6 (12.2%)		9 (10%)	4 (13%)	
Malignancy	13 (18%)	6 (12%)		13 (15%)	6 (19%)	
Fever	34 (48%)	30 (61%)		49 (55%)	15 (48%)	
Cough	33 (46%)	29 (59%)		43 (48%)	19 (61%)	
SOB	35 (49%)	29 (59%)		48 (54%)	16 (52%)	
Pulse (BPM)	87.9 (72.0–102)	102.8 (91.8–114)	***	90.6 (74–106)	103.8 (92–114)	**
Temp. (°C)	37.1 (36.6–37.4)	37.5 (36.7–38.1)	*	37.2 (36.6–37.5)	37.3 (36.7–37.9)	
BP (mmHg)	137 (120–151)	123 (102–138)	**	135 (120–149)	120 (102–130)	**
Respiratory rate	20.1 (17–22)	26.6 (20–30)	***	21.7 (17.0–22)	25.7 (19.2–30)	*
Hb (g/L) Male [133–168] Female [118–148]	126 (111–141)	124 (107–146)		127 (113–140)	120 (96–146)	
WBC ($\times 10^9$/L) [3.5–11.0]	8.28 (4.8–10.9)	9.98 (5.9–11.7)		8.63 (5.1–11.0)	9.97 (6.5–11.6)	
Lymphs ($\times 10^9$/L) [1.0–3.5]	1.25 (0.75–1.4)	1.06 (0.60–1.6)		1.239 (0.7–1.50)	0.994 (0.5–1.35)	
PLTS ($\times 10^9$/L) [150–400]	256 (176–306)	236 (166–292)		255 (175–305)	226 (169–259)	
HCT (%)	0.372 (0.337–0.417)	0.371 (0.329–0.436)		0.375 (0.341–0.416)	0.361 (0.287–0.443)	
ALT (U/L) Male [≤41] Female [≤33]	32.8 (13–31.8)	47.1 (16–39.0)		36.1 (15–39.0)	46.6 (13–31.5)	
Urea (mmol/L) [2.5–7.8]	8.29 (4.6–8.75)	13.05 (5.2–19.00)	**	8.53 (4.4–8.8)	15.12 (8.0–22.2)	***
Creatinine (μmol/L) Male [59–104] Female [45–84]	112 (66–104)	133 (70–168)		109 (65.0–105)	153 (84.5–184)	
CRP (mg/L) [<4]	47.7 (11–61.2)	132.9 (49–190.0)	***	61.4 (13.5–79.5)	144.3 (52.5–209.5)	***

Cohort demographics are presented as counts and percentages (%) for categorical data and means with interquartile range for continuous data. *P*-values indicating significant differences between groups follow 'star' notation i.e., '***' correspond to p-values < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1 and missing when > 0.1. P-values were not calculated for the group counts (N). This study looked at the serum samples from 120 COVID-19 patients. 49 patients developed severe symptoms and 31 patients died as a result of the infection. Severity score metrics based on the 4C Mortality score (Knight et al., 2020) are provided as group means. Some disparity can be observed in gender as women represent 13% of the severe cases and only 29% of the deceased patients. Age groups of severe and deceased patients also tend to be slightly higher. BMI is not significantly different between the groups. O_2 support indicates the number of patients that required oxygen support at any time during their

Table 1 (continued)

hospitalization. Mechanical respiratory support indicates the need of invasive support or continuous positive airway pressure (CPAP). Max FiO_2 captures the maximum fraction of inspired oxygen required by the patient during the hospitalization period, where respiratory support captures the patients requiring any support at diagnosis and FiO_2 represents the fraction of inspired oxygen required at time of sample acquisition. As expected, oxygen need and inspired fraction, are highly correlated with severity and fatal outcome. National Early Warning Score (NEWS) also showed correlation with both severity and outcome. Cardiac disease refers to multiple cardiovascular conditions, most frequently: ischemic heart disease, atrial fibrillation, and heart failure. Kidney disease is a grouping of stages G2 to G5 of chronic kidney disease as defined by the National Institute for Health and Care Excellence (NICE) (NICE, 2015). Liver disease in most cases refers to cirrhosis and hepatitis. Malignancy cases vary from lung, bladder, prostate, skin cancer to haematological. Those underlying conditions did not show significant differences in severe infections or poor outcome. Despite kidney disease classification not showing correlation, estimated Glomerular Filtration Rate (eGFR) levels were significantly different across severity and outcome classes. Fever (temperature ≥ 38 °C), cough and shortness of breath (SOB) were noted at time of sample acquisition and did not show relation to severity or outcome. Pulse, systolic blood pressure (BP) and respiratory rate also taken at samples acquisition showed correlation with severity and outcome, where higher pulse, higher respiratory rate and lower blood pressure are associated with severe cases and poor outcome. Haemoglobin levels (Hb), white blood cell count (WBC), lymphocyte count (Lymphs), platelets count (PLTs), haematocrit (HCT) and alanine aminotransferase (ALT) measured at sample acquisition did not show significant correlation with severity and outcome. Urea, creatinine, and C-reactive protein (CRP) concentrations are consistently elevated in severe and deceased patients. Hb, WBC, Lymphs, PLTs and HCT were measured on Beckman analyser. ALT, urea, creatinine and CRP were measured on a Roche analyser. Reference ranges are provided in square brackets [] when available

features is not a practical solution (Kenny et al., 2010) for an assay of general utility, and thus subsets of these metabolic features were investigated further. The sub-group selection was guided by metabolic pathway enrichment analysis and manual curation of well identified (i.e., MSI level 1 or 2) compounds with known biological relevance. A final model with 20 compounds selected in this way showed a reasonable cross-validated mean AUC of 0.793 (SD 0.080) for severity and 0.792 (0.090) for outcome. Mean balanced accuracy was calculated at 0.716 (0.088) and 0.655 (0.098) for severity and outcome, respectively. Representative receiver operating characteristic (ROC) curves for one specific train-test split are shown in Fig. 1. Lastly, age and gender adjustment

were considered in the model for severity, and these had very minor impact on AUC, did not change metabolite estimate intervals and were themselves insignificant and centred around 0.

2.4 Metabolic pathways linked to severity and outcome

Pathway enrichment analysis was performed with MUM-MICHOG (Li et al., 2013) as implemented in Metabo-Analyst (Pang et al., 2020) as a way of selecting biologically relevant sub-groups of compounds. As expected, no 'entire' pathways showed significant p-values when
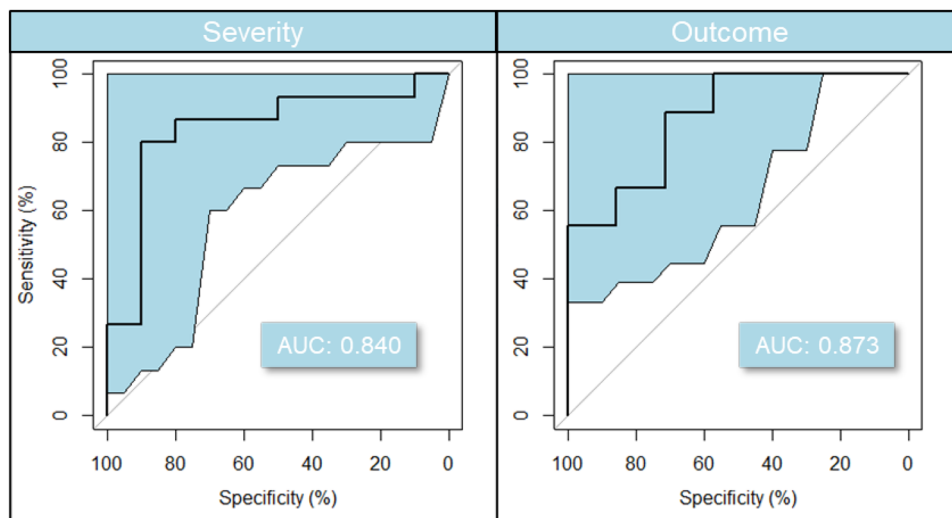

Fig. 1 Predictive model based on 20 compounds selected for their identification confidence and known biological role (see Fig. 4). Balanced accuracy and AUC for the specific train / test split are in the 0.70 s and 0.80 s, respectively. ROC 95% confidence intervals were calculated with 2000 stratified bootstrap replicates on the test data

and are presented as blue shading around the mean curve. The balanced accuracy calculated using Monte Carlo cross-validation for these models is 0.716 for severity and 0.655 for outcome. Cross-validated AUC was calculated as ~ 0.79 in both conditions

ESI + and ESI− results were grouped together (Anderson et al., 2014; Kell & Goodacre, 2014; Kell & Westerhoff, 1986). However, pathways that showed multiple significant hits were further investigated. These included pyrimidine metabolism and tryptophan metabolism. A number of these are in accordance with recently published studies comparing COVID-19 patients to healthy controls (Blasco et al., 2020; Overmyer et al., 2020; Thomas et al., 2020). The 20 compounds selected with this approach (shown with their significance in patient outcome Fig. 2) are further discussed below. Selected compound significance in the prediction of COVID-19 severity is shown in Figure S6.

2.5 Elevated serum deoxycytidine and ureidopropionate association with severity and outcome

Deoxycytidine levels were increased over twofold in patient samples at admission who went on to develop severe symptoms or subsequently died (Figs. S7A, B). Furthermore, ureidopropionate, a pyrimidine degradation product, showed the largest Fold Change (FC) increase of 2.4. Similarly, elevated cytosine levels were found in COVID-19 patients when compared to against healthy controls and 15 days post infection in (Blasco et al., 2020), but severity or outcome were not discriminated as they were here.

Uridine, another pyrimidine, was found to be significantly decreased in fatal outcome cases; however, its FC was close, but not significant in severe cases. Moreover, pseudouridine, an isomer of uridine, was increased in patient samples with
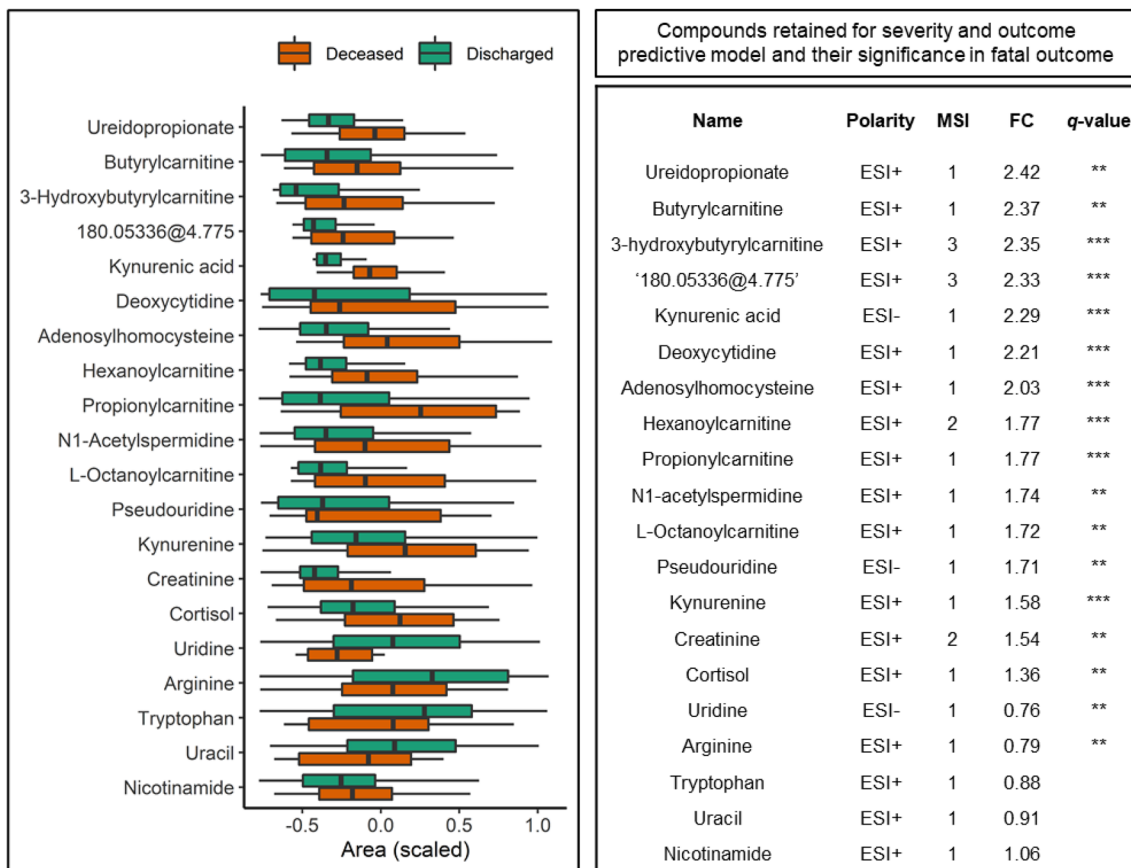

Compounds retained for severity and outcome predictive model and their significance in fatal outcome				
Name	**Polarity**	**MSI**	**FC**	**q-value**
Ureidopropionate	ESI+	1	2.42	**
Butyrylcarnitine	ESI+	1	2.37	**
3-hydroxybutyrylcarnitine	ESI+	3	2.35	***
'180.05336@4.775'	ESI+	3	2.33	***
Kynurenic acid	ESI-	1	2.29	***
Deoxycytidine	ESI+	1	2.21	***
Adenosylhomocysteine	ESI+	1	2.03	***
Hexanoylcarnitine	ESI+	2	1.77	***
Propionylcarnitine	ESI+	1	1.77	***
N1-acetylspermidine	ESI+	1	1.74	**
L-Octanoylcarnitine	ESI+	1	1.72	**
Pseudouridine	ESI-	1	1.71	**
Kynurenine	ESI+	1	1.58	***
Creatinine	ESI+	2	1.54	**
Cortisol	ESI+	1	1.36	**
Uridine	ESI-	1	0.76	**
Arginine	ESI+	1	0.79	**
Tryptophan	ESI+	1	0.88	
Uracil	ESI+	1	0.91	
Nicotinamide	ESI+	1	1.06	

Fig. 2 Compounds retained for severity and outcome predictive model. Box plot shows compound area differences between discharged and deceased patients ordered by fold change. Compound areas are standardized (mean=0, SD=1) to facilitate comparison. Boxes represent the quartiles Q1 to Q3 with Q2 (i.e., median) line in the middle. The 'whiskers' depict the upper and lower limit i.e., Q1±(Q3-Q1). For visualization simplicity the data is clipped between 10 and 90 percentiles. The table on the right side of the figure shows detailed information about the compounds including Fold Change (FC) and q-value (false discovery rate corrected p-value) following 'star' notation i.e., '***' correspond to q-values < 0.001, '**' < 0.01, and missing when > 0.1. In the case of '180.05336@4.775 ', which was initially identified as nicotinuric acid at MSI level 3, further standard based validation refuted this identification. Therefore, the compound remains unknown. Further details can be found in 'Compound identity validation' section in Methods

severe disease progression or deceased outcome. Pseudouridine (Fig. 2) is a marker of cell ribosomal RNA (rRNA) turnover (Nakano et al., 1993), for instance in heart failure (Dunn et al., 2007).

2.6 Tryptophan and kynurenine metabolism compounds associate with both severity and outcome

Kynurenine (a tryptophan degradation product) was significantly increased in both severe cases and in patients who died, with a 1.5-fold change. The difference was even more marked for kynurenic acid in outcome, with levels increasing over twofold.

In addition to changes in kynurenine and kynurenic acid, our results showed that a reduction in levels of serotonin (MSI level 3) and melatonin (MSI level 3) were associated with COVID-19 severity and death. We also observed an increase in serum levels of cortisol with q-value of 0.006 and ~ 1.3-fold in severe cases as well as those who died. The observed changes in tryptophan metabolism appear to be related to an immune response to SARS-CoV-2 since the activity of tryptophan dioxygenase, a tryptophan-degrading enzyme, is upregulated by cortisol and initiates degradation of tryptophan to kynurenine. Previous research (Thomas et al., 2020) has found a correlation between increases of interleukin 6 and kynurenine levels in COVID-19 patients compared to controls. Finally, kynurenine pathway upregulation as part of an inflammatory response is known to negatively impact serotonin levels (Hunt et al., 2020; Li et al., 2017), as was observed in our results.

Interestingly, we observed nicotinamide (MSI 1) levels to be 1.5-fold higher in severe cases, but not significantly changed with respect to outcome. It is important to note that nicotinamide is related to kynurenine/tryptophan metabolism (Murray, 2003).

2.7 Beta oxidation metabolites, acylcarnitines, increased in severe and deceased outcome patients

Multiple fatty acyl carnitines, in particular hydroxybutyrylcarnitine (> twofold), hexenoylcarnitine (1.4-fold) and hydroxyoctanoylcarnitine (~ twofold), were found to be significantly higher in both severe and poor outcome cases. Long chain fatty acyl carnitines were not reliably detected here as the UHPLC gradient used for data acquisition is optimized for more hydrophilic molecules. Therefore, fatty acyl carnitines with more than 10-carbon FA chains were excluded from the analysis. On the other hand, carnitine (MSI level 1) and its precursors, i.e., lysine (MSI 1) and methionine (MSI 1), were unchanged in both severity and outcome.

2.8 Other compounds

The final selection of compounds also includes arginine (MSI level 1) and S-adenosylhomocysteine (MSI level 1), a precursor to homocysteine. Levels of these compounds were significantly increased in patients who developed severe COVID-19 and those with a fatal outcome. Indeed, higher levels of S-adenosylhomocysteine have been found in patients with various cardiovascular disorders, metabolic syndrome, and inflammatory disorders (Catena et al., 2015; Choi et al., 2014; Schaffer et al., 2014; Soda, 2018). Moreover, the model includes N1-acetylspermidine (MSI level 1) that was reported along with spermidine by (Thomas et al., 2020) as increased in COVID-19 patient sera compared to controls. Interestingly, in contrast to N1-acetylspermidine, spermidine did not show significant changes in relation to either severity or outcome. N1-acetylspermidine was also found to be related to COVID-19 prognosis in a smaller, recent study (Danlos et al., 2021); however, in contrast to our findings lower levels were associated with an unfavourable outcome. Finally, the model includes creatinine (MSI level 2) which, in accordance with the clinically acquired data, tends to increase in patients with fatal outcomes but did not increase with disease severity.

2.9 Model compounds adjusted for demographic factors and underlying conditions

Confounders are an important issue in omics studies (Broadhurst & Kell, 2006), and several factors (age, gender, BMI, and existing inflammatory diseases) are recognised as predisposing patients infected with COVID-19 to severe disease and poor outcome. Any impact on compound levels simply from population demographics and underlying conditions was explored with logistic regression. Individual compounds' OR and 95% CI for outcome as adjusted for age, gender, BMI, diabetes, liver, kidney, and cardiac disease and all together are presented in Table 2. The severity-based table is available in the supplementary information (Table S3). It is important to note that identifying links between those factors and the compounds of interest is meaningful when evaluating their biological role; however, being 'confounded' does not invalidate their predictive power and relevance in a predictive model. Additionally, it can be observed that not all compounds (e.g. kynurenic acid, cortisol) selected for their significant q-value and fold change showed significance based on 95% CI.

An important result from this analysis is that deoxycytidine and kynurenine remained significant regardless of patient demographics or underlying conditions. By contrast, increases in abundance of short and medium fatty acylcarnitines in severity and outcome appear to be partially explained by age and BMI, as ORs tended to decrease when

adjusted; this may be correlated to frailty in ageing (Rattray et al., 2019). After adjusting for all recorded conditions in the study, only butyrylcarnitine, 3-hydroxybutyrylcarnitine and hydroxyhexanoylcarnitine showed a strong relation to disease severity. In respect to outcome, butyrylcarnitine remained the most significant. This indicates that higher levels of acylcarnitines are likely linked to metabolic differences in the patients prior to the viral infection. As such they could potentially indicate a risk group. Moreover, pseudouridine's significance was impacted by age, BMI and mildly by cardiac conditions and hypertension. This in itself is not a confounding issue since both age and BMI also contribute statistically to the outcome of COVID-19 (Hussain et al., 2020; Iaccarino et al., 2020).

Interestingly ureidopropionate, despite maintaining strong significance when adjusted for all factors, showed an increased OR when adjusted for age in fatal outcome case. In contrast, ureidopropionate's OR dropped when adjusted for age and BMI in severe (vs mild) cases. Lower ureidopropionate levels have also been associated with an increased risk of developing type 2 diabetes mellitus and coronary artery disease (Ottosson et al., 2018).

A recent publication found an association between gender and kynurenic acid levels in COVID-19 patients suggestive of sex-specific differences in immune responses and clinical outcomes (Cai et al., 2021). In our results, gender correction did not impact kynurenic acid OR in either outcome or severity. However, this simple observation does not disprove the more complex and thoughtful analysis that was performed in the paper by Cai et al. Moreover, our attempt to further consider kynurenine to kynurenic acid ratio was unsuccessful due to wide variations in our cohort resulting in numerical instabilities in the regression method when applied to those ratios.

2.10 Validation study

Validation is an important aspect in metabolomics studies not only to validate statistical results but to also ensure reproducibility (Gromski et al., 2015). The Bayesian logistic model described above was trained on the discovery cohort data and used to predict severity and outcome in a separate cohort of 90 patients in a blind manner i.e., data acquisition and data analysis (model predictions) were performed

Table 2 Adjusted logistic regression results by outcome

	Adjusted by:	None		Age		Gender		BMI		Liver disease		CV disease		Hypertension		Kidney disease		Diabetes		All	
	Compound	OR	Sig.	OR	Sig.	OR	Sig.	OR	Sig.	OR	Sig.	OR	Sig.	OR	Sig.	OR	Sig.	OR	Sig.	OR	Sig.
Nucleic acids metabolism	Deoxycytidine	2.1 (1.4-3.6)	**	2 (1.2-3.4)	*	2.3 (1.4-3.9)	**	2.1 (1.3-3.8)	**	2.1 (1.4-3.6)	**	2.1 (1.3-3.5)	**	2 (1.3-3.4)	**	2 (1.3-3.5)	**	2 (1.3-3.5)	**	2.5 (1.4-4.9)	**
	Uracil	0.75 (0.47-1.1)		0.84 (0.56-1.2)		0.79 (0.49-1.2)		0.83 (0.56-1.2)		0.75 (0.47-1.2)		0.72 (0.44-1.1)		0.78 (0.48-1.2)		0.79 (0.49-1.2)		0.79 (0.49-1.2)		0.81 (0.45-1.4)	
	Ureidopropionate	2.8 (1.4-6.9)	*	7.5 (2.8-24)	***	2.7 (1.4-6.5)	*	3.2 (1.4-8.4)	**	2.8 (1.4-6.9)	*	3 (1.5-7.3)	**	3.9 (1.8-10)	**	3.2 (1.5-8.3)	**	3.2 (1.5-8.3)	**	4.5 (1.8-14)	**
	Pseudouridine	1.9 (1.3-2.8)	**	1.5 (0.97-2.4)	.	1.9 (1.2-2.9)	**	1.6 (1.1-2.5)	*	1.9 (1.3-2.8)	**	1.8 (1.2-2.8)	**	1.7 (1.1-2.7)	*	2 (1.2-3.5)	*	2 (1.2-3.5)	*	1.6 (0.91-2.9)	
	Uridine	0.39 (0.21-0.67)	**	0.58 (0.37-0.88)	*	0.4 (0.22-0.7)	**	0.61 (0.39-0.91)	*	0.39 (0.21-0.67)	**	0.36 (0.19-0.63)	***	0.41 (0.22-0.7)	**	0.41 (0.22-0.71)	**	0.41 (0.22-0.71)	**	0.41 (0.19-0.82)	*
Kynurenine metabolism	Tryptophan	0.71 (0.46-1.1)		0.88 (0.59-1.3)		0.68 (0.43-1.1)	.	0.79 (0.53-1.2)		0.72 (0.46-1.1)		0.72 (0.46-1.1)		0.73 (0.47-1.1)		0.76 (0.48-1.2)		0.76 (0.48-1.2)		0.72 (0.42-1.2)	
	Kynurenine	2.7 (1.7-4.8)	***	2.3 (1.4-4.2)	**	2.8 (1.8-5.1)	***	2.3 (1.4-3.9)	**	2.7 (1.7-4.8)	***	2.9 (1.7-5.4)	***	2.6 (1.6-4.6)	***	2.6 (1.6-4.7)	***	2.6 (1.6-4.7)	***	2.6 (1.5-5)	**
	Kynurenic acid	1.5 (0.99-2.6)		1.4 (0.91-2.7)		1.5 (0.96-2.6)		1.4 (0.88-2.6)		1.5 (0.98-2.6)		1.5 (1-2.6)	.	1.4 (0.91-2.4)		1.3 (0.87-2.4)		1.3 (0.87-2.4)		1.2 (0.73-2.1)	
	Cortisol	1.4 (0.93-2.2)		1.3 (0.9-2.1)		1.4 (0.93-2.2)		1.2 (0.82-2)		1.4 (0.94-2.2)		1.4 (0.91-2.1)		1.4 (0.92-2.2)		1.4 (0.93-2.1)		1.4 (0.93-2.1)		1.4 (0.86-2.3)	
	Nicotinamide	1.1 (0.68-1.5)		1.2 (0.8-1.8)		1 (0.67-1.5)		1 (0.69-1.5)		1 (0.67-1.5)		1.1 (0.68-1.5)		1.1 (0.69-1.6)		1.1 (0.67-1.5)		1.1 (0.67-1.5)		1.2 (0.7-1.9)	
Fatty acyl carnitines	Propionylcarnitine	2.2 (1.5-3.6)	***	2 (1.3-3.2)	**	2.2 (1.4-3.5)	***	1.9 (1.2-3)	**	2.3 (1.5-3.6)	***	2.2 (1.4-3.5)	***	2.1 (1.4-3.5)	***	2.2 (1.4-3.5)	**	2.2 (1.4-3.5)	**	1.9 (1.1-3.2)	*
	Butyrylcarnitine	2.5 (1.6-4.4)	***	2.8 (1.6-5.7)	**	2.5 (1.6-4.5)	***	2.3 (1.3-4.4)	**	2.6 (1.6-4.6)	***	2.5 (1.5-4.4)	***	2.4 (1.5-4.3)	***	2.4 (1.5-4.3)	***	2.4 (1.5-4.3)	***	2.2 (1.3-4.4)	*
	3-hydroxybutyrylcarnitine	1.9 (1.3-2.9)	**	1.7 (1.1-2.8)	*	2.1 (1.4-3.3)	***	1.5 (0.98-2.4)	.	1.9 (1.3-2.9)	**	1.8 (1.2-2.9)	**	1.8 (1.2-2.9)	**	1.8 (1.2-2.8)	**	1.8 (1.2-2.8)	**	1.6 (1-2.7)	*
	Hexanoylcarnitine	1.4 (0.95-2.4)		1.2 (0.82-2.2)		1.4 (0.92-2.3)		1.3 (0.83-2.2)		1.4 (0.95-2.4)		1.4 (0.93-2.3)		1.3 (0.9-2.3)		1.3 (0.88-2.2)		1.3 (0.88-2.2)		1.1 (0.67-1.7)	
	L-Octanoylcarnitine	1.4 (0.93-2.2)		1.2 (0.78-1.9)		1.3 (0.9-2.2)		1.2 (0.79-1.9)		1.4 (0.93-2.2)		1.3 (0.91-2.2)		1.3 (0.87-2.1)		1.3 (0.84-2.1)		1.3 (0.84-2.1)		0.95 (0.56-1.5)	
Other	Creatinine	1.5 (1-2.4)	*	1.4 (0.91-2.2)		1.5 (1-2.3)	.	1.4 (0.93-2.4)		1.5 (1-2.4)	*	1.6 (1.1-2.4)	*	1.4 (0.93-2.2)		1.4 (0.91-2.3)		1.4 (0.91-2.3)		1.4 (0.82-2.4)	
	Arginine	0.53 (0.33-0.81)	**	0.66 (0.43-0.97)	*	0.52 (0.32-0.81)	**	0.69 (0.46-1)	.	0.51 (0.32-0.8)	**	0.53 (0.33-0.81)	**	0.53 (0.33-0.82)	**	0.55 (0.34-0.86)	*	0.55 (0.34-0.86)	*	0.59 (0.33-1)	.
	N1-acetylspermidine	1.9 (1.3-3)	**	1.8 (1.2-3.1)	*	1.9 (1.3-3.1)	**	1.7 (1.1-2.9)	*	2 (1.3-3.2)	**	1.9 (1.2-3)	**	1.9 (1.3-3.1)	**	1.8 (1.2-2.9)	**	1.8 (1.2-2.9)	**	1.7 (1-3)	*
	'180.05336@4.775'	1.6 (1.1-2.6)	*	1.5 (1-2.8)	.	1.6 (1.1-2.5)	*	1.6 (1-2.9)	.	1.7 (1.1-2.7)	*	1.6 (1.1-2.6)	*	1.5 (1-2.4)	*	1.6 (1-2.6)	.	1.5 (1-2.6)	.	1.5 (0.92-2.6)	
	Adenosylhomocysteine	2.3 (1.4-4.2)	**	2.2 (1.3-4.5)	*	2.3 (1.4-4.4)	**	2.3 (1.3-4.7)	*	2.3 (1.4-4.2)	**	2.2 (1.4-4.1)	**	2.2 (1.3-4.1)	**	2.3 (1.3-4.8)	*	2.3 (1.3-4.8)	*	2.1 (1.2-4.3)	*

Positive OR indicate increased levels in patients with poor outcome and are presented with OR (95% CI). Significance is presented in 'star' notation i.e., '***' correspond to p-values < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1 and missing when > 0.1. Compounds are adjusted for age, gender, BMI, liver conditions, cardiovascular diseases, hypertension, kidney disease, diabetes and all together. Details of specific diseases in each category are available in the Methods section

without knowledge of patient outcome or severity. Results from this blind prediction are as follows: the outcome AUC was 0.83 (CI 0.74–0.91) for outcome and 0.76 (CI 0.67–0.86) for severity (Fig. 3).

The validation study effort was focused on the 20 compounds identified in the discovery study (Fig. 2). UHPLC-MS/MS peak areas were extracted for these compounds in Trace Finder (see Materials and Methods) to allow for more detailed and controlled signal mapping between batches and studies. In the case of 'deoxycytidine' all three cytosine events were summed up i.e., cytosine, cytidine, deoxycytidine. This approach was employed because automatic software pre-processing in Compound Discoverer showed multiple alignment issues between batches of those cytosine events close in RT and resulting from cytosine, cytidine, and deoxycytidine source fragmentation (see Sect. 1.1). For this reason, these will be referred to as cytosine-based nucleosides in the following paragraphs.

Kynurenic acid (ESI-) showed poor acquisition signal in the validation study and was removed from the model. Uridine (ESI-) and pseudouridine (ESI-) showed significant distribution differences after normalization between the data acquired for the discovery study patients and validation study patients. Such study bias can mislead the model therefore these two compounds were also removed from the validation model.

All areas for the remaining 17 ESI + compounds were normalized by batch (using pooled QC samples for each batch) and between batches (using inter batch IQAs) as described in methods for all discovery and validation batches. This ensured relative areas of the 17 compounds were comparable between the 2 studies so the model could be trained on the discovery study and used to predict the validation study patients.

It is important to note that at the time of the validation cohort, which was later than those in the discovery phase, patients were regularly treated with dexamethasone, an anti-inflammatory corticosteroid that by this time was known to have efficacy as a treatment agent (NHS, 2020; RECOVERY, 2021). This change in patient treatment is likely a possible explanation of the drop in AUC for severity prediction as this could have influenced the patient state severity perception.

As shown in Figure S8, higher levels of *S*-adenosylhomocysteine, propionylcarnitine, ureidopropionate and kynurenine are highly influential factors in this multivariate outcome model, combined with low levels of tryptophan and uracil. Interestingly creatinine, octanoylcarnitine and hexanoylcarnitine show a negative influence once conditioned on these other compounds. In terms of severity, ureidopropionate and *S*-adenosylhomocysteine showed the strongest positive influence. However, these results need to be interpreted with caution due to conditioning e.g., on mediators.

Still, it is interesting to note that there is broad agreement on directionality between the models.

2.11 Longitudinal study

Finally, a longitudinal sample of 28 severe patients with divergent outcomes from the discovery cohort (13 deceased and 15 discharged) were analysed using UHPLC-MS in 7 batches for a total of 198 samples. The samples were acquired throughout the patients' stay on different days and with a different number of samples for each patient. Areas of the compounds previously selected in the predictive model were extracted in TraceFinder to allow for reliable comparison as alignment between batches suffered similar issues as those described in the validation study section.

Initial analysis of the longitudinal data via individual time series plots did not present clear patterns (not shown). This is most likely due to the irregular interval between samples, the different times of day of sample acquisition, medication interference, measurement variability and the general complexity of biological processes. However, when focusing the analysis on the changes of those compounds throughout the hospital stay, i.e., the difference between the first and last sample of each patient, some compounds showed significant directional changes based on logistic regression analysis. Figure 4 shows the significant compounds, ureidopropionate (A), uracil (B), arginine (C) and tryptophan(D).

It appears that levels of ureidopropionate are significantly increased in patients who subsequently died, but levels remained unchanged in discharged patients. In addition, the role of this pathway is strengthened by the significant inverse relation in uracil levels (an increase) in discharged patients. Interestingly, cytosine-based nucleotides did not show significant differences in levels between deceased and discharged patients 0.67 (0.26–1.5) OR (95% CI), with the wide CI indicating strong variability between individuals. When it comes to tryptophan levels, we can see that discharged patients experienced increases in levels, where deceased patients saw a drop in tryptophan levels. Neither kynurenine 0.66 (0.27–1.4) or kynurenic acid 1.2 (0.52–3.6) showed significant deltas probably due to the broad variation between patients or possibly suffering from precursor deficiency in some cases.

3 Discussion

In this study we employed untargeted metabolomics using UHPLC-MS/MS to detect and measure changes in the baseline serum metabolome of a cohort of 120 COVID-19-infected patients at the point of hospital admission. We used this to build a predictive model of disease severity and outcome then validated our findings in a blind study on an

additional 90 patients. Furthermore, we explored the temporal evolution of patient serum metabolome in the longitudinal data of 28 severe patients. Our aim was to find not only prognostic markers of subsequent disease severity and outcome, but to also understand what impact COVID-19 infection has on the patient's metabolome, and conversely what effect patient biochemistry and physiology might have on infection development. These results could then be used to guide patient treatment and medical attention requirements following COVID-19 diagnosis. Most early studies applying metabolomics to COVID-19 human patients have compared them against healthy controls; this is of limited predictive value, as rapid PCR and point-of-care antibody tests exist and does not provide an insight into what biochemical changes drive disease severity or outcome. Moreover, a number of these studies (Table S1), used either a small patient cohort or applied targeted metabolomics restricting the breath of possible findings; some also lacked proper statistical analysis (see Broadhurst & Kell, 2006; Trivedi et al., 2017)).

Our results showed that distinct alterations in the serum metabolome were already capable of identifying patients with higher risk of severe illness or fatal outcome at the time of diagnosis and admission. More importantly, using both univariate and a multiple predictor Bayesian logistic regression model we found a subset of 20 metabolites (16 of which could be identified to MSI Level 1) with relevant biological functions that are predictive of subsequent disease severity and patient outcome with AUC 0.792 and 0.793 respectively. These metabolite differences are centred particularly around pyrimidine, tryptophan and acylcarnitine metabolism, which can be related to viral presence/replication (and/or the virus's stimulation of host biosynthesis), host inflammation response and alterations in energy metabolism.

3.1 Alterations in pyrimidine metabolites predictive of severity and outcome

Viral infections induce characteristic changes in host cell metabolism to enable effective viral replication (Thaker et al., 2019). Moreover, the resulting metabolic impact and cellular reprogramming varies between viruses (even within the same family) and host cell type. In the case of SARS-CoV-2, cytosine has been described as pivotal in the virus' evolution (Danchin & Marlière, 2020) where avoidance of host defence mechanisms have favoured a reduced cytosine proportion in viral RNA, estimated at 17.6%. When compared to typical human RNA, the cytosine proportion is significantly lower. In contrast to cytosine, the proportion of uracil in SARS-CoV-2 is estimated at around 32.4% (Danchin & Marlière, 2020) and it is significantly higher than that in human RNA. This difference between host nucleotides anabolic processes and viral RNA needs

could result in significant levels of cytosine-based ribonucleotides and deoxyribonucleotides being accumulated in infected cells.

CTP synthetase, the enzyme converting UTP to CTP, is allosterically inhibited by CTP, therefore allowing the virus to thrive despite its RNA composition differences. However, the reduction of the ribonucleotides to deoxyribonucleotides by ribonucleotide reductase is nonspecific, and its activity is not modulated by CDP or CTP (Hofer et al., 2012). This could allow the small differences in CTP to UTP ratios to be amplified in their reduced form, consistent with the principles of metabolic control analysis (Kell & Mendes, 2012). The accumulation of CDP and CTP in infected cells could result in breakdown and potentially deoxycytidine being excreted from the cells as waste product or released at cell death. It is important to note that SARS-CoV-2 has been found to use lysosomal trafficking to exit the cell (Ghosh et al., 2020), therefore leaving an open question on how the deoxycytidine is being released into the systemic circulation.

The high levels of ureidopropionate, a product of pyrimidine catabolism, observed in patients with severe disease and poor outcome also show activation of salvage pathways, most likely driven by an excess of nucleosides. It is important to note that the first step of cytidine breakdown is the conversion to uridine by cytidine aminohydrolase therefore, reducing the gap between cytidine / uridine levels. High levels of ureidopropionate have been linked to neuropathology and showed to act as endogenous neurotoxin (Kölker et al., 2001).

Cytosine has been reported previously to discriminate between COVID-19 patients and uninfected controls (Blasco et al., 2020), but has not been assessed regarding severity or outcome. Here, our analyses showed that deoxycitidine levels were predictive of subsequent disease severity and outcome. Serum deoxycitidine levels were increased over twofold in serum from patient samples at admission who went on to develop severe symptoms or subsequently died (Figure S7 A and B). Given the reduced incorpration of cytosine into SARS-CoV-2, these results may indicate a higher viral load and replication, subsequently leading to the development of severe symptoms, in some cases resulting in death.

Increased viral reproduction could be due to an initial higher viral load, a host environment favourable to viral reproduction, or the effective stage of the infection. It is thus not possible to draw mechanistic conclusions from our results; nevertheless, the deoxycitidine OR remained stable when adjusted for demographic factors and known underlying conditions indicating that those factors are not significant with respect to viral replication rates. This would be more consistent with the fact that the initial load before admission is the most important parameter affecting both disease severity and outcome.

Given these results, measurement of deoxycitidine and ureidopropionate, could potentially allow tracking of viral activity and predict recovery or aggravation. However, as highlighted by (Migaud et al., 2020) the levels of certain nucleobases are also increased in patients who die from sepsis and acute respiratory failure. Further investigation would be required to tease out the contributions of SARS-CoV-2 viral replication to the levels of these pyrimidines against the secondary effects of the virus on the host (human).

Pseudouridine, was also noted to be increased in severe cases and poor outcome. As mentioned previously, pseudouridine (Fig. 2) is a marker of cell (rRNA) turnover (Nakano et al., 1993), for instance in heart failure (Dunn et al., 2007). When adjusting for cardiovascular disease this compound remained significant but showed slight correlation with hypertension and age (Table 2). Finally, when adjusted for all factors pseudouridine 95% CI lost significance indicating correlation with more than one factor. Interestingly, growing evidence points toward complications in COVID-19 arising through a vasculopathy and coagulopathy elicited by the infection (Pretorius et al., 2020) and thus, increased pseudouridine levels may be indicative of this process.

3.2 Tryptophan—kynurenine degradation

The degradation of tryptophan to kynurenine is often associated with increase in inflammatory processes (Schröcksnadel et al., 2006). In the context of this study the non-significant decrease of tryptophan in severe cases could be explained either by a change in dietary habits, possible weight loss, sarcopenia, or more likely by the higher stimulation of tryptophan to kynurenine degradation indicated by the significant increase in levels of kynurenine and kynurenic acid. Upregulation of this process was further confirmed by the higher levels of cortisol, which stimulates tryptophan degradation, especially in severe cases (Table 2). An increase in patient kynurenine levels has also been reported in COVID-19 patients compared to controls (Thomas et al., 2020), associated with severity (Overmyer et al., 2020), and also linked to fatal sepsis development (Migaud et al., 2020). This provides strong evidence of higher levels of immune response in severe cases and those with a fatal outcome. More recently, upregulation of tryptophan metabolites including kynurenine has been found to play a protective role in radiation injury during cancer radiotherapy (Guo et al., 2020) indicating that these relationships are more complicated than previously thought, and also involve the gut microbiome.

3.3 Fatty acid beta oxidation

The levels of short and medium chain acylcarnitines have been previously reported by (Thomas et al., 2020) as being reduced in COVID-19 patients *versus* controls irrespective of Interleukin 6 (IL6) levels. In this study we identified multiple short chain acyl carnitines as significantly increased in severe and fatal cases compared to mild cases and discharged patients. Changes in serum acylcarnitines have been previously associated with cardiovascular disease, diabetes and inflammation (Anderson et al., 2014). Moreover, increased levels of octanoyl-l-carnitine have been previously associate with arterial stiffness (Kim et al., 2015) and dysregulation of the carnitine shuttle. Long chain acyl carnitines, not detected in this study, have been additionally reported in relation to frailty in the elderly (Rattray et al., 2019). Moreover, accumulation of fatty acid oxidation products was found linked to mitochondrial disfunction (Wajner & Amaral, 2016). Finally, fatty acid β-oxidation has been associated with T-cell development and differentiation (Lochner et al., 2015).

Indeed, adjusted logistic regression results (Table 2) for these compounds in our study show that some of their significance can be explained by BMI levels. This could possibly indicate differences in energy metabolism in response to viral infection linked to pre-existing phenotype; i.e., BMI and therefore represent a high-risk group.

3.4 Compounds requiring further investigation

Our study highlighted several other compounds (not discussed here) that changed significantly in severity and outcome. However, as is common in metabolomic studies (Blaženović et al., 2018; Salek et al., 2013; Shrivastava et al., 2021), these require further rigorous identification following the Metabolomics Standards Initiative (MSI) for reporting metabolite identification (Sumner et al., 2007) and so were not included in the predictive model at this stage. Examples of such compounds include pentahomomethionine (MSI 3) and trihomomethionine (MSI 3), both sulphur-containing amino acids. These were both increased in severe cases and were especially high in patients with a fatal outcome. Other sulphur-containing compounds such as cysteine and taurine have been found to be reduced in COVID-19 cases compared to controls, and in COVID-19 patients with moderate-high IL-6 levels (Thomas et al., 2020). Homomethionines such as penta- and tri-homomethionine identified in our results are formed by transamination of oxo-acids that are themselves formed during fatty acid breakdown, possibly indicating a pro-catabolic phenotype as is common in inflammation (Underwood et al., 2006) and possibly suggestive of sarcopenia.

Another compound worthy of further attention is ergothioneine. Ergothioneine is a potent exogenous antioxidant, usually acquired via the consumption of mushrooms (Borodina et al., 2019; Cheah & Halliwell, 2012). The potential protective value in SARS-CoV-2 infections of ergothioneine was recently reviewed by (Cheah & Halliwell, 2020). The

Fig. 3 Predictive model performance based on 17 ESI + compounds previously selected in the discovery study. ROC 95% confidence intervals were calculated with 2000 stratified bootstrap replicates on the test data and are presented as blue shading around the mean curve. A Monte Carlo cross-validation results estimate the model balanced accuracy at 0.63 for severity and outcome
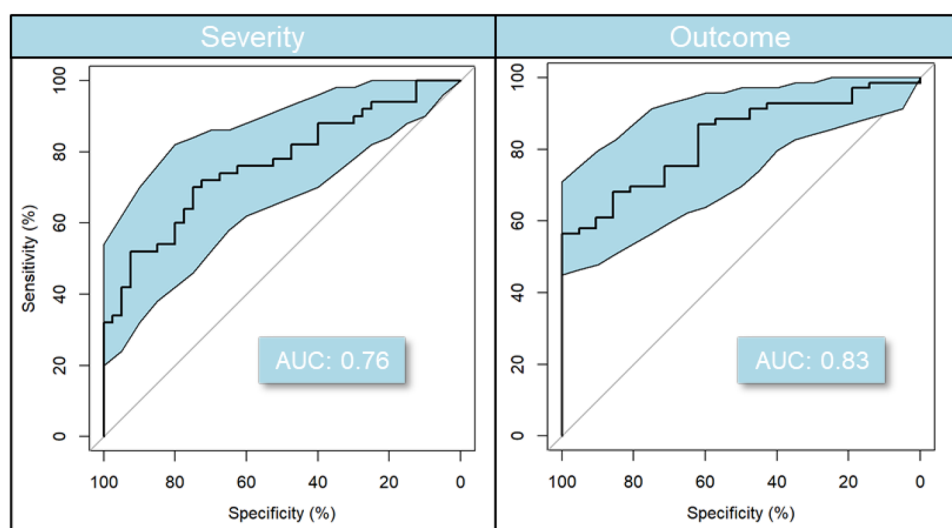
results of our study found ergothioneine levels to follow the expected trend, i.e., lower in severe cases and poor outcome in accordance with findings by (Wu et al., 2020); however its *q*-values (in a population whose mushroom consumption was neither monitored nor controlled) fell just slightly short of significance (see Fig. S12A, B). Moreover, piperine (MSI 2) (representative of black pepper consumption) was interestingly found significantly decreased in severe cases and poor outcome (see Fig. S12C, D). This most likely reflects dietary changes in the patients experiencing severe symptoms; however, the potential impact of piperine to the host organism is not fully understood.

3.5 Limitations and future work

Whilst untargeted UHPLC-MS analysis can detect large number of compounds of diverse chemical classes, limitations in the compound coverage can result from sample preparation methods, LC solvents and gradient elution profiles. In this study the serum extraction and LC gradient were targeted at hydrophilic compounds, and hence numerous lipids were not reliably measured, e.g., long chain acyl carnitines. Despite this limitation several lipids that exhibit amphiphilic properties e.g., phospholipids and fatty acids were detected and showed significant changes between the patient groups studied. However, confirmation of their identity will require further work before being integrated into a predictive model.

Additional limitations in the presented work come from the preselection of metabolic features that were individually significant following volcano plot analysis. While this simple method of feature selection narrows the list of compounds it also prevents us from identifying more complex interactions in the case of disjoint populations.

Despite these limitations, multiple significant biological processes were identified as key in discriminating between disease severity and outcome. Future work will aim to quantitate those changes.

3.5.1 Conclusions

We have here performed a well-powered, untargeted metabolomics analysis of serum of COVID-19 patients with the aim of finding prognostic markers of disease severity and outcome. Using both univariate and a multivariable Bayesian logistic regression model we found a subset of 20 metabolites with relevant biological functions that are predictive of subsequent disease severity and patient outcome. Although, no individual metabolite appeared be *strongly* discriminative on its own, a combined model based on viral activity, host immune response and underlying metabolic differences showed promising predictive results with AUC 0.792 and 0.793, for outcome and severity, respectively. Furthermore, we validated those findings in a blind study on 90 additional patients with AUC of 0.83 and 0.76. Longitudinal exploration of some of these potential markers showed strong relation to poor outcome opening opportunities to differentiate between recovery and aggravation states. These markers hold promise to improve patient care upon COVID-19 infection and diagnosis.

4 Methods

All UHPLC-MS data acquired is freely available in mzML format in the MetaboLights repository (Haug et al., 2019) with study identifier MTBLS2997 (www.ebi.ac.uk/metabolights/).
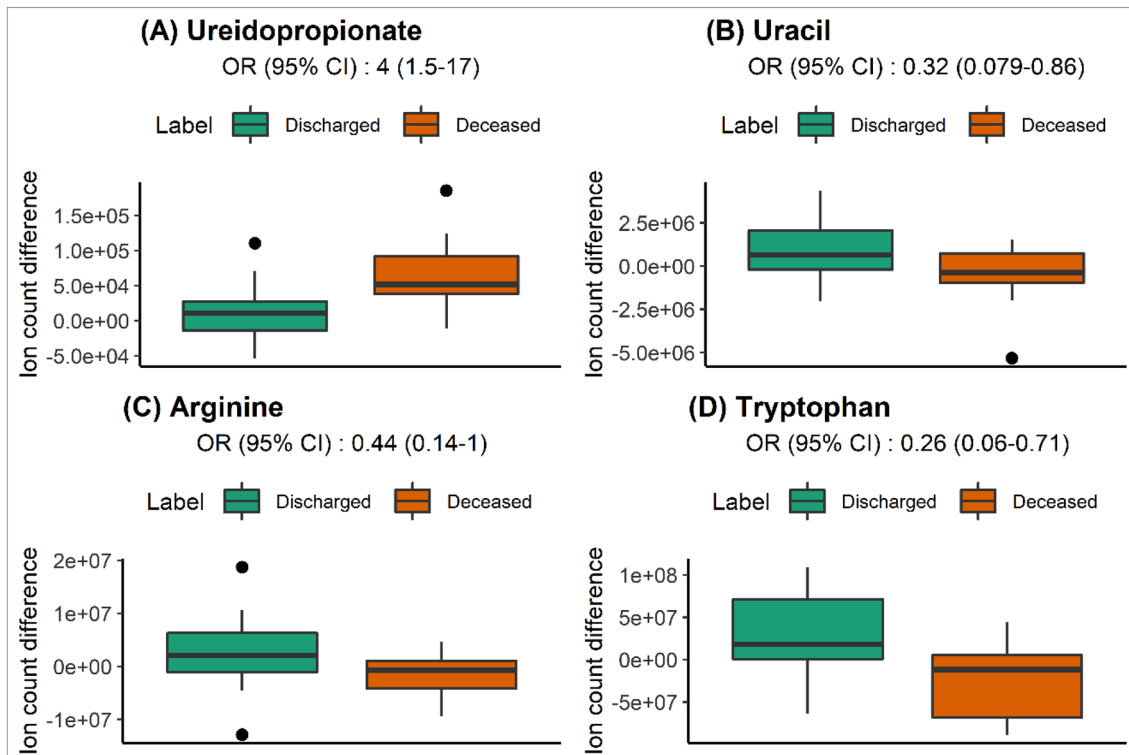
Fig. 4 Level of selected compounds for significant change in level between first and last patient sample. Ureidopropionate (**A**), uracil (**B**), arginine (**C**) and tryptophan (**D**) levels showed significantly different evolution measured by logistic regression OR and 95%CI. Y axes reflect the difference in ion counts or peak areas difference between the first sample and last sample per patient. Boxplot markings follow the same standard as Fig. 4

4.1 Sample acquisition

All samples were acquired at the Royal Liverpool University Hospital (RLUH) on the first positive SARS- CoV-2 test (not fasted and different times of the day). Surplus serum was saved after routine diagnostic testing on patients admitted to the hospital who subsequently tested positive for SARS-CoV-2 via PCR. Blood was initially collected into VACUETTE Clot Activator tubes (Greiner, Germany) within approximately 48 h of presentation and centrifuged at 1500×g for 10 min within 60 min of collection. Surplus serum was stored at − 80 °C prior to processing and analysis. Ethical approval for the use of serum samples and associated metadata in this study was obtained from North West – Haydock research ethics committee (REC ref: 20/NW/0332).

4.2 Sample preparation for metabolomics analysis

Patient serum samples were thawed at room temperature and maintained on ice throughout the sample preparation process. Samples were prepared by addition of 100 μL of sample to a 2 mL Eppendorf containing 350 μL methanol (LC–MS grade) previously cooled at − 80 °C and maintained on dry ice. The mixture of serum and methanol was vortexed vigorously followed by centrifugation at 18,000×g for 15 min at 4 °C to pellet proteins. Multiple 75 μL aliquots (for extraction replicates) of the resulting supernatant were then dried in a vacuum centrifuge (ScanVac MaxiVac Beta Vacuum Concentrator system, LaboGene ApS, Denmark) with no temperature application and stored at − 80 °C until required for UHPLC-MS/MS analysis.

Batch quality controls (QC) and conditioning QC samples were also prepared in this way by first pooling together equal amounts from each patient sample within a batch. Inter-batch quality assurance (IQA) samples were prepared with the same protocol using pooled serum samples of non-COVID-19 patients provided by RLUH. Therefore, those samples are not representative of the study samples, but representative of the general hospital population and sample acquisition and storage practices. All IQA samples were prepared at the same time at the beginning of the study to minimize variation between batches resulting from sample preparation.

Additional inter-batch quality assurance samples were prepared using commercial pooled human serum (BioIVT, Lot BRH1413770, Cat: HMSRM, mixed gender 0.1 um filtered) spiked with internal standards. Here, 100 μL sample

were added to a 2 mL Eppendorf containing 330 μL methanol (LC–MS grade) and 20 μL of an internal standard mixture (ISTDs) as described in (Wright Muelas et al., 2020). The mixture of methanol and internal standards (ISTDs) was previously cooled at − 80 °C and maintained on dry ice when adding serum. A template run order and description of samples is provided in the Supplementary information (Table S4).

Extraction blanks were prepared in the same way as serum samples replacing serum with 100 μL of water (LC–MS grade).

Prior to analysis, samples were resuspended in 40 μL water (LC–MS), centrifuged at 17,000×g for 15 min at 4 °C to remove any particulates and transferred to glass sample vials.

4.3 UHPLC-MS/MS analysis of patient serum samples

Untargeted UHPLC-MS/MS data acquisition was performed as described in (Wright Muelas et al., 2020) and using published methodologies and guidelines (Broadhurst & Kell, 2006; Broadhurst et al., 2018; Brown et al., 2005; Dunn et al., 2011; Mullard et al., 2015; Raad et al., 2016). Data were acquired using a ThermoFisher Scientific Vanquish UHPLC system coupled to a ThermoFisher Scientific Q-Exactive mass spectrometer (ThermoFisher Scientific, UK). Samples for the longitudinal aspect of this study were analysed in the same way but using a ThermoFisher Scientific ID-X Tribrid mass spectrometer (ThermoFisher Scientific, UK). Mass spectrometer operation and details are provided in Supplementary Information.

Samples were analysed following guidelines set out in (Dunn et al., 2011) and (Broadhurst et al., 2018) and in the order per batch as described in Table S4. Briefly, blank extraction samples were injected at the beginning and end of each batch to assess carry over and lack of contamination. QC samples, prepared by pooling equal aliquots of analytical samples in each batch, were applied to condition the analytical platform, assess reproducibility and to correct for systematic errors within batches. Quality Assurance (QA) samples were also incorporated in every batch at regular intervals. Two pools of hospital patient serum (referred as Inter-batch quality assurance (IQA)) and commercial serum (referred as SQA) were prepared at the beginning of the study and used in every batch allowing batch alignment in the data processing step. Samples with isotopically labelled internal standards spiked in (SQA, see Sample preparation for metabolomics analysis) were used to monitor mass accuracy. IQA samples were subsequently used to correct peak areas across all batches. Four randomly selected samples from each analytical batch were run in duplicate replicates to further assess reproducibility.

4.4 UHPLC-MS/MS data pre-processing and analysis

Untargeted compound pre-processing was used for the discovery study and a targeted area extraction was employed in the validation study. However, data acquisition in both studies was performed with the same untargeted UHPLC-MS method.

4.5 Discovery study data pre-processing

Raw instrument data from all batches in.RAW file format were exported to Thermo Fisher scientific Compound Discoverer 3.1 (CD3.1) for deconvolution, alignment and annotation (full workflow and settings as described in (Wright Muelas et al., 2020)) based on IQA samples.

Compound grouping was performed in CD3.1 where 6971 metabolic features in ESI + and 3122 metabolic features in ESI- were retained. Retained compounds are selected based on presence in more than 80% of the IQA, CVs less than 30% and signal 5 times higher than the blank injections. From those metabolic features 267 in ESI + and 142 in ESI- were identified against ThermoFisher Scientific 'mzCloud' spectral library with score higher than 70% (MSI 2) or against in-house spectral library with score higher than 75% (MSI 1) and full match on proposed molecular formula from CD 3.1. For all data acquired, annotation and identification criteria were according to (Sumner et al., 2007).

4.6 Validation and longitudinal study data pre-processing

Areas for the 20 compounds previously selected for the model were extracted in Thermo Fisher Scientific Trace Finder 5.1 software from the.RAW files. A compound fragmentation database was obtained from Compound Discoverer from previous runs. The quan master method was employed to allow summing up of all events in cases of metabolic features eluting across multiple peaks such as cytosine and cytosine-containing compounds where source fragmentation was observed, or butylcarnitine where multiple elution peaks were observed. The ICIS detection algorithm was employed for all metabolic features; however, other detection algorithm settings (e.g., peak detection strategy, peak threshold type) and retention times settings (i.e., detection type and RT window) were tuned individually to each metabolic feature allowing for controlled selection of the peak area in every sample. Exported areas were further processed in R for normalization and statistical analysis.

4.7 Area normalization and batch correction

A custom 2-step normalization was used to correct for small within-batch runtime drift and larger between-batch

variations in both studies. To perform this, non-normalized peak areas from CD3.1 were exported as a.csv file. QC-based area correction was performed in R (version 4.0.2) as discussed in (Dunn et al., 2011) using the *fANCOVA* package. The QC correction was performed on each batch independently to remove runtime drift intensity variations. Subsequently, variation between batches was corrected using correction to the mean based on IQA injections for each metabolite separately. PCA plots comparing the results for different normalization approaches are included in the supplementary information (Figure S13). This approach was taken as no available tool (Misra, 2021) allowed to normalize taking into account inter-batch time drift and between batch variation.

4.8 Metadata and multiblock analysis

The following metadata variables were used for multivariate analysis: gender, age, BMI, Glasgow coma scale (GCS), NEWS, pulse, temperature, blood pressure, respiratory rate, Hb, WBC, Lymphs, PLTs, HCT, ALT, total bilirubin, urea, creatinine, eGFR, CRP, FiO_2 (%) and O_2 saturation (%). The description of the biochemical tests is presented in Table 1. There were 1.44% missing values in this meta data set; these were imputed by using *K*-NN imputation algorithm (Troyanskaya et al., 2001).

The multiblock analysis was performed on three data blocks: UHPLC-MS ESI + /ESI− data and metadata. For each block, the data matrix was auto-scaled so that each variable has a mean of 0 and a standard deviation of 1. In addition, a block scaling factor which is the inverse of the square root of the number of variables of this block was applied to compensate differences in variance due the difference in the number of variables. A multiblock PCA (MB-PCA) model called consensus PCA (Smilde et al., 2003; Xu & Goodacre, 2012) was applied to the three blocks of data. The results of this MB-PCA model are consisted of one super scores matrix which represents the common trend of all the blocks and three block scores matrix which represents the pattern of each block under perspective of the common trend.

4.9 Selection of significant metabolic features

Background features and metabolic features detected in less than 25% of the samples were excluded from further analysis. Metabolic features were further filtered based on False Discovery Rate (FDR) corrected *p*-value (i.e., *q*-value) significance < 0.05 and $\log_2$ fold change > 0.5 in severity (i.e., severe vs mild cases) and outcome (i.e., diseased vs discharged patients). *p*-values were calculated using the Mann–Whitney as the data did not satisfy normality assumption required for a T-test. The usually adopted log transformation approach to satisfy T-test requirements also tended to overemphasize low abundance metabolic features.

Significant differences (q-value < 0.05 and absolute $\log_2$ fold change > 0.5) were found for 1143/601 metabolic features in ESI ± when comparing severe against mild cases, with 1650/680 metabolic features (ESI ±) when comparing between outcomes. For simplicity, due to the overlap in patients across the two comparisons performed (severity and outcome), the union of significant features across these two comparisons was taken forward for further analysis, with a total of 1987/973 ESI ± metabolic features.

4.10 UHPLC-MS signal curation

Significant metabolic features were manually curated in CD3.1 based on the UHPLC signal quality and MS spectra. UHPLC quality was assessed based on between-batch retention time (RT) overlap and clear peak separation with smooth peak appearance. MS spectra were evaluated on the detection of preferred ion i.e., $[M+H]^+$ and $[M-H]^-$ with at least 2 isotopes. Additionally, metabolic features where the signal was filled by CD gap fill option for more than 20% of the QC samples and 90% of the samples were also excluded.

Where metabolic features of interest were detected in both ESI+ and ESI−, the clearest signal was retained for further analysis. When necessary, standards were run to confirm MS/MS, RT and signal intensity in each ESI polarity. Specifically, in the case of uridine and pseudouridine best separation and signal intensity detection of standards was achieved in ESI-, therefore negative polarity data was used for those compounds.

4.11 Pathway enrichment analysis

Pathway enrichment analysis was performed in MUMMIC-HOG (Li et al., 2013) version 2 incorporated in Metabo-Analyst (Pang et al., 2020). To match the MUMMICHOG requested *m/z* feature format, resolved masses retrieved from CD analysis were altered to achieve $[M+H]^+$ adduct for ESI + and ESI- results. The MUMMICHOG adduct option was set to recognize exclusively $[M+H]^+$ adducts. This allowed processing of ESI+ and ESI− results together and minimize false hits due to multiple adduct matches. Pathways with a high number of significant hits were further manually investigated and hits subsequently verified by MS, RT, MS/MS and match against standards when available.

4.12 Compound identity validation

The identity of the compounds retained for the model was validated against standards whenever possible. Standard validation was performed against RT and MS/MS profile of 1 μM and 10 μM standards in water or appropriate solution.

In some cases, e.g., cytosine derivatives, asparagine and ureidopropionate the standards were also spiked in to pooled study serum at 5 µM to validate the RT in the matrix.

MS/MS profile exploration was performed in the case of ureidopropionate due to a poor MS/MS capture in the untargeted data acquisition. A targeted Parallel Reaction Monitoring (PRM) method focused on the 133.06077 m/*z* ion at a limited scan range of 66.7 to 200 m/z was performed. The quadrupole isolation window was set to 1.2 m/z reflecting the original method value. This setup allowed for the acquisition of the MS/MS profile of this compound at any UHPLC gradient stage for all ions in a 1.2 m/z window range around 133.06077 m/z. This approach provided best results in capturing low intensity signal in areas of the UHPLC gradient with high event load as was the case with ureidopropionate.

In the case of the unknown metabolic feature '180.05336@4.775 ', which was initially identified as nicotinuric acid at MSI level 3, further validation demonstrated that despite significant fragmentation profile overlap with nicotinuric acid the compound of interest structure is undeniably different based on UHPLC gradient elution. Therefore, its identity remains unknown. More details related to compound identity validation and area calculation choices are available in the Supplementary information Sect. 1.1.

4.13 Multiple predictor models

Reported results for multiple predictor models were produced with a Bayesian logistic regression model implemented in R with *rstanarm* R package (Goodrich, 2020). Comparative analysis was performed with the extreme gradient boosting *xgboost* R package (Chen & Guestrin, 2016; Friedman, 2001), Logistic regression with Generalized Linear Models (GLM) *glm* R package and GLM with Elastic net regularization *glmnet* R package (Zou & Hastie, 2005). Data were separated into training and test groups (80:20) with balanced label ratios. Conservative regularization parameters were used to reduce overfitting. Bayesian GLM approach was set to increased regularization with prior scale = 1, for glmnet *alpha* was set to 0.1 and xgboost *eta* to 0.01 with *max_depth* = 1. Due to the large group disparities in outcome, weights were incorporated in the model to handle class imbalance. Sparsity inducing parameters such as L1 regularization in glmnet and lasso or hierarchical shrinkage prior in Bayesian GLM were avoided despite better results in some cases. This allowed to perform subgroup selection accounting for compound identity confidence level and known biological role. From the four compared methods Bayesian logistic regression consistently showed better generalization and was therefore retained for results reporting.

Mean and standard deviation for accuracy and AUC were estimated using cross-validation with 100 iterations. These cross-validation results were used to describe the model sensitivity to the data and not for hyperparameters optimization. ROC plots with 95% confidence intervals were obtained from a randomly selected train/test partition using pROC R package with 2000 stratified bootstrap replicates on the test data.

4.14 Adjusted compounds significance

Individual OR and 95% CI for selected compounds were evaluated with univariate logistic regression with Generalized Linear Models (*glm*) implementation in R *stats* package. OR (95% CI) and *p*-values for significance are presented in the reported tables. Compounds are adjusted for age, gender, BMI, liver disease, cardiovascular disease, hypertension, kidney disease (i.e., chronic kidney disease stages 2 to 5), diabetes mellitus (type 1 or 2) and all together. Liver conditions include cirrhosis, hepatitis, alcoholic hepititis, autoimmune hepatitis, ascites, transplant, fatty liver disease. Cardiovascular conditons include ischaemic heart disease (IHD), atrial fibrillation (AF), cardiomegaly, cardiomyopathy, left ventricular systolic dysfunction, left ventricular hypertrophy, left ventricular failure, congestive heart failure, drug induced myocarditis, heart failure, angina and Takotsubo cardiomyopathy with IHD and AF being most frequent.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

Data availability UHPLC-MS data for the discovery and validation studies in mzML format is freely available at MetaboLights repository (Haug et al., 2019) with study identifier MTBLS2997 (www.ebi.ac.uk/metabolights/). The code covering data normalization, predictive models and basic visualizations is available at https://github.com/dbkgr

oup/COVID. Multiblock PCA analysis can be found at https://github.com/Biospec/cluster-toolbox-v2.0.

References

Alene, M., Yismaw, L., Assemie, M. A., Ketema, D. B., Mengist, B., Kassie, B., & Birhan, T. Y. (2021). Magnitude of asymptomatic COVID-19 cases throughout the course of infection: A systematic review and meta-analysis. *PLoS ONE, 16*, e0249090.

Anderson, S. G., Dunn, W. B., Banerjee, M., Brown, M., Broadhurst, D. I., Goodacre, R., Cooper, G. J. S., Kell, D. B., & Cruickshank, J. K. (2014). Evidence that multiple defects in lipid regulation occur before hyperglycemia during the prodrome of type-2 diabetes. *PLoS ONE, 9*, e103217.

Ansone, L., Ustinova, M., Terentjeva, A., Perkons, I., Birznice, L., Rovite, V., Rozentale, B., Viksna, L., Kolesova, O., Klavins, K., et al. (2021). Tryptophan and arginine metabolism is significantly altered at the time of admission in hospital for severe COVID-19 patients: Findings from longitudinal targeted metabolomics analysis. medRxiv, 2021.2003.2031.21254699.

Arunachalam, P. S., Wimmers, F., Mok, C. K. P., Perera, R. A. P. M., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tak-Yin Tsang, O., et al. (2020). Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science, 369*, 1210.

Bennet, B.M., Wolf, J., Laureano, R., & Sellers, R.S. (2020). Review of current vaccine development strategies to prevent coronavirus disease 2019 (COVID-19). Toxicol Pathol 192623320959090.

Blasco, H., Bessy, C., Plantier, L., Lefevre, A., Piver, E., Bernard, L., Marlet, J., Stefic, K., Benz-de Bretagne, I., Cannet, P., et al. (2020). The specific metabolome profiling of patients infected by SARS-COV-2 supports the key role of tryptophan-nicotinamide pathway and cytosine metabolism. *Scientific Reports, 10*, 16824.

Blaženović, I., Kind, T., Ji, J., & Fiehn, O. (2018). Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. *Metabolites, 8*, 31.

Borodina, I., Kenny, L.C., McCarthy, C.M., Paramasivan, K., Pretorius, E., Roberts, T.J., van der Hoek, S.A., & Kell, D.B. (2019). The biology of ergothioneine, an antioxidant nutraceutical. Nutrition Research Reviews, 1–28.

Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., Lewis, M. R., & Dunn, W. B. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics, 14*, 72.

Broadhurst, D. I., & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics, 2*, 171–196.

Brown, M., Dunn, W. B., Ellis, D. I., Goodacre, R., Handl, J., Knowles, J. D., O'Hagan, S., Spasić, I., & Kell, D. B. (2005). A metabolome pipeline: From concept to data to knowledge. *Metabolomics, 1*, 39–51.

Cai, Y., Kim, D. J., Takahashi, T., Broadhurst, D. I., Yan, H., Ma, S., Rattray, N. J. W., Casanovas-Massana, A., Israelow, B., Klein, J., et al. (2021). Kynurenic acid may underlie sex-specific immune responses to COVID-19. *Science Signaling, 14*, eabf8483.

Catena, C., Colussi, G., Nait, F., Capobianco, F., & Sechi, L. A. (2015). Elevated homocysteine levels are associated with the metabolic syndrome and cardiovascular events in hypertensive patients. *American Journal of Hypertension, 28*, 943–950.

Cheah, I. K., & Halliwell, B. (2012). Ergothioneine; antioxidant potential, physiological function and role in disease. *Biochimica Et Biophysica Acta Molecular Basis of Disease, 1822*, 784–793.

Cheah, I. K., & Halliwell, B. (2020). Could ergothioneine aid in the treatment of coronavirus patients? *Antioxidants, 9*, 595.

Chen, T., and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Paper presented at Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.*

Choi, J. W., Lee, M. H., Fujii, T., Fujii, N., & Moon, Y. (2014). Association of the urine homocysteine/creatinine ratio to proinflammatory cytokine, natural anticoagulant, and nitric oxide levels in cerebrovascular disease. *Annals of Clinical & Laboratory Science, 44*, 461–465.

Danchin, A., & Marlière, P. (2020). Cytosine drives evolution of SARS-CoV-2. *Environmental Microbiology, 22*, 1977–1985.

Danlos, F.-X., Grajeda-Iglesias, C., Durand, S., Sauvat, A., Roumier, M., Cantin, D., Colomba, E., Rohmer, J., Pommeret, F., Baciarello, G., et al. (2021). Metabolomic analyses of COVID-19 patients unravel stage-dependent and prognostic biomarkers. *Cell Death & Disease, 12*, 258.

Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J. D., Halsall, A., Haselden, J. N., et al. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols, 6*, 1060–1083.

Dunn, W. B., Broadhurst, D. I., Deepak, S. M., Buch, M. H., McDowell, G., Spasic, I., Ellis, D. I., Brooks, N., Kell, D. B., & Neyses, L. (2007). Serum metabolomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate. *Metabolomics, 3*, 413–426.

Dunn, W. B., Lin, W., Broadhurst, D., Begley, P., Brown, M., Zelena, E., Vaughan, A. A., Halsall, A., Harding, N., & Knowles, J. D. (2015). Molecular phenotyping of a UK population: Defining the human serum metabolome. *Metabolomics, 11*, 9–26.

Fox, S. E., Akmatbekov, A., Harbert, J. L., Li, G., Quincy Brown, J., & Vander Heide, R. S. (2020). Pulmonary and cardiac pathology in African American patients with COVID-19: An autopsy series from New Orleans. *The Lancet Respiratory Medicine, 8*, 681–686.

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 1189–1232.

Fuhrer, T., Zampieri, M., Sévin, D. C., Sauer, U., & Zamboni, N. (2017). Genomewide landscape of gene–metabolome associations in *Escherichia coli*. *Molecular Systems Biology, 13*, 907.

Ghosh, S., Dellibovi-Ragheb, T. A., Kerviel, A., Pak, E., Qiu, Q., Fisher, M., Takvorian, P. M., Bleck, C., Hsu, V. W., Fehr, A. R., et al. (2020). β-Coronaviruses use lysosomes for egress instead of the biosynthetic secretory pathway. *Cell, 183*, 1520-1535.e1514.

Goodrich, B. (2020). rstanarm: Bayesian applied regression modeling via Stan., I.A. Jonah Gabry, Sam Brilleman, ed.

Grobler, C., Maphumulo, S. C., Grobbelaar, L. M., Bredenkamp, J. C., Laubscher, G. J., Lourens, P. J., Steenkamp, J., Kell, D. B., & Pretorius, E. (2020). Covid-19: The rollercoaster of fibrin (Ogen), D-dimer, Von Willebrand factor, P-Selectin and their interactions

with endothelial cells, platelets and erythrocytes. *International Journal of Molecular Sciences, 21*, 5168.

Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., & Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta, 879*, 10–23.

Guo, H., Chou, W.-C., Lai, Y., Liang, K., Tam, J.W., Brickey, W.J., Chen, L., Montgomery, N.D., Li, X., & Bohannon, L.M. (2020). Multi-omics analyses of radiation survivors identify radioprotective microbes and metabolites. *Science, 370*.

Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., Péré, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., et al. (2020). Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science, 369*, 718.

Haug, K., Cochrane, K., Nainala, V. C., Williams, M., Chang, J., Jayaseelan, K. V., & O'Donovan, C. (2019). MetaboLights: A resource evolving in response to the needs of its scientific community. *Nucleic Acids Research, 48*, D440–D444.

Hofer, A., Crona, M., Logan, D. T., & Sjöberg, B.-M. (2012). DNA building blocks: Keeping control of manufacture. *Critical Reviews in Biochemistry and Molecular Biology, 47*, 50–63.

Hunt, B.C., e Cordeiro, T.M., Robert, S., de Dios, C., Leal, V.A.C., Soares, J.C., Robert, D., Antonio, T., and Sudhakar, S.M. (2020). Effect of mmune activation on the kynurenine pathway and depression symptoms–A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*.

Hussain, A., Mahawar, K., Xia, Z., Yang, W., and Shamsi, E.-H. (2020). Obesity and mortality of COVID-19. Meta-analysis. *Obesity Research & Clinical Practice*.

Iaccarino, G., Grassi, G., Borghi, C., Ferri, C., Salvetti, M., & Volpe, M. (2020). Age and multimorbidity predict death among COVID-19 patients: Results of the SARS-RAS study of the Italian Society of hypertension. *Hypertension, 76*, 366–372.

Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: Beyond biomarkers and towards mechanisms. *Nature Reviews Molecular Cell Biology, 17*, 451–459.

Kell, D. B., & Goodacre, R. (2014). Metabolomics and systems pharmacology: Why and how to model the human metabolic network for drug discovery. *Drug Discovery Today, 19*, 171–182.

Kell, D.B., and Mendes, P. (2012). Metabolic control analysis and biotechnology in the post-genomic era. *Technological and Medical Implications of Metabolic Control Analysis, 74*.

Kell, D. B., & Oliver, S. G. (2016). The metabolome 18 years on: A concept comes of age. *Metabolomics, 12*, 148.

Kell, D. B., & Westerhoff, H. V. (1986). Metabolic control theory: Its role in microbiology and biotechnology. *FEMS Microbiology Letters, 39*, 305–320.

Kenny, L. C., Broadhurst, D. I., Dunn, W., Brown, M., North, R. A., McCowan, L., Roberts, C., Cooper, G. J., Kell, D. B., & Baker, P. N. (2010). Robust early pregnancy prediction of later preeclampsia using metabolomic biomarkers. *Hypertension, 56*, 741–749.

Kim, M., Jung, S., Lee, S.-H., & Lee, J. H. (2015). Association between Arterial Stiffness and Serum L-Octanoylcarnitine and Lactosylceramide in Overweight Middle-Aged Subjects: 3-Year Follow-Up Study. *PLoS ONE, 10*, e0119519.

Knight, S. R., Ho, A., Pius, R., Buchan, I., Carson, G., Drake, T. M., Dunning, J., Fairfield, C. J., Gamble, C., Green, C. A., et al. (2020). Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Development and validation of the 4C Mortality Score. *BMJ, 370*, m3339.

Kölker, S., Okun, J. G., Hörster, F., Assmann, B., Ahlemeyer, B., Kohlmüller, D., Exner-Camps, S., Mayatepek, E., Krieglstein, J., & Hoffmann, G. F. (2001). 3-Ureidopropionate contributes to the neuropathology of 3-ureidopropionase deficiency and severe propionic aciduria: A hypothesis. *Journal of Neuroscience Research, 66*, 666–673.

Leisman, D. E., Deutschman, C. S., & Legrand, M. (2020). Facing COVID-19 in the ICU: Vascular dysfunction, thrombosis, and dysregulated inflammation. *Intensive Care Medicine, 46*, 1105–1108.

Li, S., Park, Y., Duraisingham, S., Strobel, F. H., Khan, N., Soltow, Q. A., Jones, D. P., & Pulendran, B. (2013). Predicting network activity from high throughput metabolomics. *PLoS Computational Biology, 9*, e1003123.

Li, Y., Hu, N., Yang, D., Oxenkrug, G., & Yang, Q. (2017). Regulating the balance between the kynurenine and serotonin pathways of tryptophan metabolism. *The FEBS Journal, 284*, 948–966.

Libby, P., & Lüscher, T. (2020). COVID-19 is, in the end, an endothelial disease. *European Heart Journal, 41*, 3038–3044.

Lochner, M., Berod, L., & Sparwasser, T. (2015). Fatty acid metabolism in the regulation of T cell function. *Trends in Immunology, 36*, 81–91.

López-Hernández, Y., Monárrez-Espino, J., Herrera-van Oostdam, A.-S., Delgado, J.E.C., Zhang, L., Zheng, J., Valdéz, J.O., Mandal, R., González, F.O., & Borrego, J.C. (2021). Targeted metabolomics identifies high performing diagnostic and prognostic biomarkers for COVID-19.

Marietta, M., Ageno, W., Artoni, A., De Candia, E., Gresele, P., Marchetti, M., Marcucci, R., & Tripodi, A. (2020). COVID-19 and haemostasis: A position paper from Italian Society on Thrombosis and Haemostasis (SISET). *Blood Transfusion, 18*, 167–169.

Migaud, M., Gandotra, S., Chand, H. S., Gillespie, M. N., Thannickal, V. J., & Langley, R. J. (2020). Metabolomics to predict antiviral drug efficacy in COVID-19. *American Journal of Respiratory Cell and Molecular Biology, 63*, 396–398.

Misra, B. B. (2021). New software tools, databases, and resources in metabolomics: Updates from 2020. *Metabolomics, 17*, 49.

Mullard, G., Allwood, J. W., Weber, R., Brown, M., Begley, P., Hollywood, K. A., Jones, M., Unwin, R. D., Bishop, P. N., Cooper, G. J. S., et al. (2015). A new strategy for MS/MS data acquisition applying multiple data dependent experiments on Orbitrap mass spectrometers in non-targeted metabolomic applications. *Metabolomics, 11*, 1068–1080.

Murray, M. F. (2003). Nicotinamide: An oral antimicrobial agent with activity against both mycobacterium tuberculosis and human immunodeficiency virus. *Clinical Infectious Diseases, 36*, 453–460.

Nakano, K., Nakao, T., Schram, K. H., Hammargren, W. M., McClure, T. D., Katz, M., & Petersen, E. (1993). Urinary excretion of modified nucleosides as biological marker of RNA turnover in patients with cancer and AIDS. *Clinica Chimica Acta, 218*, 169–183.

NHS. (2020). COVID-19 therapy: corticosteroids including dexamethasone and hydrocortisone. In NHS.

NICE. (2015). Chronic kidney disease in adults: assessment and management. NICE *Clinical guideline [CG182]*.

Oliver, S. G., Winson, M. K., Kell, D. B., & Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology, 16*, 373–378.

Ottosson, F., Smith, E., Melander, O., & Fernandez, C. (2018). Altered asparagine and glutamate homeostasis precede coronary artery disease and type 2 diabetes. *The Journal of Clinical Endocrinology & Metabolism, 103*, 3060–3069.

Overmyer, K.A., Shishkova, E., Miller, I.J., Balnis, J., Bernstein, M.N., Peters-Clarke, T.M., Meyer, J.G., Quan, Q., Muehlbauer, L.K., and Trujillo, E.A. (2020). Large-scale multi-omic analysis of COVID-19 severity. Cell systems.

Pang, Z., Chong, J., Li, S., & Xia, J. (2020). MetaboAnalystR 3.0: toward an optimized workflow for global metabolomics. *Metabolites, 10*, 186.

Paranjpe, I., Fuster, V., Lala, A., Russak, A. J., Glicksberg, B. S., Levin, M. A., Charney, A. W., Narula, J., Fayad, Z. A., Bagiella, E., et al. (2020). Association of treatment dose anticoagulation with in-hospital survival among hospitalized patients with COVID-19. *Journal of the American College of Cardiology, 76*, 122–124.

Pretorius, E., Venter, C., Laubscher, G.J., Lourens, P.J., Steenkamp, J., and Kell, D.B. (2020). Prevalence of amyloid blood clots in COVID-19 plasma. medRxiv, 2020.2007.2028.20163543.

Raad, M.d., Fischer, C.R., and Northen, T.R. (2016). High-throughput platforms for metabolomics. *Current Opinion in Chemical Biology, 30*, 7–13.

Raamsdonk, L. M., Teusink, B., Broadhurst, D., Zhang, N., Hayes, A., Walsh, M. C., Berden, J. A., Brindle, K. M., Kell, D. B., & Rowland, J. J. (2001). A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology, 19*, 45–50.

Rattray, N. J. W., Trivedi, D. K., Xu, Y., Chandola, T., Johnson, C. H., Marshall, A. D., Mekli, K., Rattray, Z., Tampubolon, G., Vanhoutte, B., et al. (2019). Metabolic dysregulation in vitamin E and carnitine shuttle energy mechanisms associate with human frailty. *Nature Communications, 10*, 5027.

RECOVERY, C.G. (2021). Dexamethasone in hospitalized patients with Covid-19. *New England Journal of Medicine, 384*, 693–704.

Salek, R.M., Steinbeck, C., Viant, M.R., Goodacre, R., & Dunn, W.B. (2013). The role of reporting standards for metabolite annotation and identification in metabolomic studies. GigaScience *2*, 2047–2217X-2042–2013.

Schaffer, A., Verdoia, M., Cassetti, E., Marino, P., Suryapranata, H., De Luca, G., and Group, N.A.S. (2014). Relationship between homocysteine and coronary artery disease. Results from a large prospective cohort study. *Thrombosis Research, 134*, 288–293.

Schröcksnadel, K., Wirleitner, B., Winkler, C., & Fuchs, D. (2006). Monitoring tryptophan metabolism in chronic immune activation. *Clinica Chimica Acta, 364*, 82–90.

Shrivastava, A. D., Swainston, N., Samanta, S., Roberts, I., Wright Muelas, M., & Kell, D. B. (2021). MassGenie: A transformer-based deep learning method for identifying small molecules from their mass spectra. *Biomolecules, 11*, 1793. https://doi.org/10.3390/biom11121793

Sindelar, M., Stancliffe, E., Schwaiger-Haber, M., Anbukumar, D.S., Albrecht, R.A., Liu, W.-C., Travis, K.A., García-Sastre, A., Shriver, L.P., and Patti, G.J. (2021). Longitudinal metabolomics of human plasma reveals robust prognostic markers of COVID-19 disease severity. medRxiv, 2021.2002.2005.21251173.

Smilde, A. K., van der Werf, M. J., Bijlsma, S., & van der Werff-van der Vat, B.J., and Jellema, R.H. (2005). Fusion of mass spectrometry-based metabolomics data. *Analytical Chemistry, 77*, 6729–6736.

Smilde, A. K., Westerhuis, J. A., & de Jong, S. (2003). A framework for sequential multiblock component methods. *Journal of Chemometrics: A Journal of the Chemometrics Society, 17*, 323–337.

Soda, K. (2018). Polyamine metabolism and gene methylation in conjunction with one-carbon metabolism. *International Journal of Molecular Sciences, 19*, 3106.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W. M., Fiehn, O., Goodacre, R., Griffin, J. L., et al. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics, 3*, 211–221.

Thaker, S. K., & Ch'ng, J., and Christofk, H.R. (2019). Viral hijacking of cellular metabolism. *BMC Biology, 17*, 59.

Thomas, T., Stefanoni, D., Reisz, J.A., Nemkov, T., Bertolone, L., Francis, R.O., Hudson, K.E., Zimring, J.C., Hansen, K.C., Hod, E.A., et al. (2020). COVID-19 infection alters kynurenine and fatty acid metabolism, correlating with IL-6 levels and renal status. JCI Insight, *5*.

Trivedi, D. K., Hollywood, K. A., & Goodacre, R. (2017). Metabolomics for the masses: The future of metabolomics in a personalized world. *New Horizons in Translational Medicine, 3*, 294–305.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., & Altman, R. B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics, 17*, 520–525.

Underwood, B. R., Broadhurst, D., Dunn, W. B., Ellis, D. I., Michell, A. W., Vacher, C., Mosedale, D. E., Kell, D. B., Barker, R. A., & Grainger, D. J. (2006). Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. *Brain, 129*, 877–886.

Wajner, M., and Amaral, Alexandre U. (2016). Mitochondrial dysfunction in fatty acid oxidation disorders: insights from human and animal studies. *Bioscience Reports,* 36.

Wright Muelas, M., Roberts, I., Mughal, F., O'Hagan, S., Day, P. J., & Kell, D. B. (2020). An untargeted metabolomics strategy to measure differences in metabolite uptake and excretion by mammalian cell lines. *Metabolomics, 16*, 1–12.

Wu, D., Shu, T., Yang, X., Song, J.-X., Zhang, M., Yao, C., Liu, W., Huang, M., Yu, Y., Yang, Q., et al. (2020). Plasma metabolomic and lipidomic alterations associated with COVID-19. *National Science Review, 7*, 1157–1168.

Wu, Z., & McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) Outbreak in China: Summary of a report of 72314 cases From the Chinese Center for Disease Control and Prevention. *JAMA, 323*, 1239–1242.

Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., Bonten, M. M. J., Dahly, D. L., Damen, J. A. A., Debray, T. P. A., et al. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ, 369*, m1328.

Xu, Y., & Goodacre, R. (2012). Multiblock principal component analysis: An efficient tool for analyzing metabolomics data which contain two influential factors. *Metabolomics, 8*, 37–51.

Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., Ogishi, M., Sabli, I. K. D., Hodeib, S., Korol, C., et al. (2020). Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science, 370*, eabd4570.

Zheng, Y.-Y., Ma, Y.-T., Zhang, J.-Y., & Xie, X. (2020). COVID-19 and the cardiovascular system. *Nature Reviews Cardiology, 17*, 259–260.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (statistical Methodology), 67*, 301–320.