A distinct metabolic profile associated with a fatal outcome in COVID-19 patients during early epidemic in Italy

Elisa Saccon1, Alessandra Bandera2,3,4, Mariarita Sciumè5, Flora Mikeloff4, Abid Ali Lashari6, Stefano Aliberti3,2, Michael Sachs6, Filippo Billi8, Francesco Blasi3,7, Erin Gabriël6, Giorgio Costantino9,10, Pasquale De Roberto5, Shuba Krishnan1, Andrea Gori2,3,4, Flora Peyvandi3,11, Luigia Scudeller12, Christian L. Lorson13,14, Luca Valenti3,15, Kamal Singh13,14, Luca Baldini5,16, Nicola Stefano Fracchiolla5* on behalf of the “COVID-19 NETWORK” working group*, Ujjwal Neogi1,13*

1Division of Clinical Microbiology, Department of Laboratory Medicine, Karolinska Institute, Stockholm, Sweden
2Infectious Diseases Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
3Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy
4Centre for Multidisciplinary Research in Health Science (MACH), University of Milan, Milan, Italy
5Hematology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
6Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
7Internal Medicine Department, Respiratory Unit and Cystic Fibrosis Adult Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
8Acute Medical Unit, Department of Medicine, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
9Department of Anesthesia, Critical Care and Emergency, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
10Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy.
11Angelo Bianchi Bonomi Hemophilia and Thrombosis Center, IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milan, Italy
12Clinical Trials Team, Scientific Direction, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milan, Italy
13Bond Life Sciences Center, University of Missouri, Columbia, USA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Keywords: COVID-19, metabolomics, fatal outcome, Italy, predictive biomarkers

Corresponding authors:

Nicola Fracchiolla (nicola.fracchiolla@policlinico.mi.it) and Ujjwal Neogi (ujjwal.neogi@ki.se)
Abstract

Leveraging the unique biological resource based upon the initial COVID-19 patients in Policlinico di Milano (Italy), our study provides the first metabolic profile associated with a fatal outcome. The identification of potential predictive biomarkers offers a vital opportunity to employ metabolomics in a clinical setting as diagnostic tool of disease prognosis upon hospital admission.

Introduction

As of February 2021, more than 100 million cases of coronavirus disease 2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, had been confirmed globally, with more than 2 million related deaths (https://covid19.who.int/). Although most individuals remain asymptomatic or display mild symptoms, 15-20% of patients exhibit severe symptoms, specifically respiratory distress, often requiring mechanical ventilation or/and intensive care (ICU) admission [1], with a mortality rate after ICU admission estimated around 40% [2]. Multiple studies have identified profound underlying conditions that demonstrate increased susceptibility to a more severe prognosis and a higher risk of fatality, including the male gender, old age [3], or certain underlying medical conditions, such as hypertension, cardiovascular diseases, diabetes or obesity [4]. Additionally, patients infected with SARS-CoV-2 present metabolic dysregulation, possibly due to immune-triggered inflammation or other changes in the host physiology, and that these alterations often reflect the disease severity [1,5,6]. For instance, levels of particular amino acids positively correlated with severe COVID-19 cases [1,7]. Moreover, perturbations in energy metabolisms, TCA and urea cycle [6] and lipid metabolism [1,8] are correlated to disease prognosis. Thus, it is essential to assemble a complete metabolic signature correlated to disease severity to identify a set of biomarkers strongly associated with the patient outcome, with the final goal of employing them for diagnostics and therapeutic purposes.

Our study retrospectively analyzes the metabolome profile of 75 COVID-19 patients with moderate and severe symptoms admitted to Policlinico di Milano (Lombardy region, Italy) following SARS-CoV-2 infection between March and April 2020. Italy was the first Western country to experience COVID-19 disease, and the Lombardy region was the epicenter of the Italian COVID-19 pandemic.
This cohort shows a higher mortality rate compared to others; therefore, it represents a unique opportunity to investigate the underlying metabolic profiles of the first COVID-19 patients in Italy and to identify potential predictive biomarkers.

**Methods**

Plasma samples from 75 COVID-19 patients were collected at the moment of admission to Policlinico di Milano (Italy) from March to April 2020 and SARS-CoV-2 positivity was confirmed by PCR (Ethical Clearance No. 462-2020-bis). Patients were initially classified based on the type of ventilation received: Nasal cannula (n=25) and VentMask (n=25) (moderate cases) or CPAP (n=25) (severe cases).

Plasma untargeted metabolomics was performed by Global Metabolomics (HD4) in Metabolon, (USA) as previously described [9]. The metabolomics method is ISO 9001:2015 certified and the lab is accredited by the College of American Pathologist (CAP), USA. Clinical data was collected from the registries.

Logistic regression was used to model associations of each biomarker with COVID-19-related in-hospital mortality adjusting for age, gender, and body mass index (BMI). As BMI was missing for many patients, multivariate imputation by chained equations (MICE) was used (via the MICE R-package) to impute BMI. Imputation was done in two stages, first including no biomarkers, and then including those biomarkers that were found to be significant under the original imputation. The resulting pooled estimates and inference were obtained using Rubin's rules. P-values were not adjusted in this analysis.

Metabolomics data were log2 transformed and plotted using histograms with normal distribution superimposed. R package LIMMA was applied for differential abundance analysis between different mask types (Nasal cannula/VentMask/CPAP), outcome (survivors/non-survivors), and severity (moderate/severe). Adjustment for multiple testing was assessed using false discovery rate (FDR) <0.05. Heatmap was built using the R package ComplexHeatmap. Uniform Manifold Approximation and Projection (UMAP) representations were done using the R package UMAP. Metabolites with variance equal to zero were removed and positive significant pairwise correlations after Bonferroni.
correction (Spearman, adjusted $p<0.00001$) were used for association analysis. The strength of the connections was evaluated by plotting distribution of correlation coefficients in a graphical network using igraph (https://igraph.org/python/). The network was compared to a random network with similar dimensions to validate the structure of the network to be not due to chance. Community detection was performed using Leiden algorithm (https://leidenalg.readthedocs.io/en/stable/index.html). For each community large enough (n>30), metabolite set enrichment analysis (MSEA) with KEGG and Metabolon terms via the Python module gseapy was performed. The average degree and clustering coefficient were calculated for each community. The final network was build using Cytoscape and biomarkers that were significantly associated with death were highlighted.

Results and discussion

Among the 75 patients enrolled in this study, 24% (18/75) succumbed to death in the hospital (Supplementary Table 1) and a significantly higher rate of fatal outcomes was observed for severe cases compared to moderate cases ($p=0.0218$, Chi-square test). Among patients who died, comorbidities were nearly universally observed (94%, 17/18). Underlying conditions, like hypertension, were recorded for 46% (26/57) of the survivors and 50% (9/18) of the non-survivors, while cardiovascular diseases were significantly more common in non-survivors ($p=0.0320$, Fisher’s exact test) compared to survivors. Among the serological markers, lactate was significantly higher ($p<0.0001$, Mann-Whitney test) in non-survivors compared to survivors, confirming its tight link to disease severity [10].

In a logistic regression analysis (adjusted to age, gender, and BMI), 35 metabolites, among the >1000 tested, were significantly associated with COVID-19 mortality ($p<0.05$, unadjusted), among which 10 biomarkers were significant at the unadjusted 0.025 level (Figure 1A). Interestingly, cyclic adenosine monophosphate (cAMP) is significantly increased in non-survivors compared to survivors (OR: 7.4 95% CI 1.5 - 37). cAMP is a well-known intracellular messenger that functions as a regulator of various cellular activities, including cell growth and differentiation, gene transcription, protein expression, and is intimately involved in mitochondrial dynamics [11]. As cAMP plays a role in
SARS-CoV-2 endocytosis in the initial phases of the infection [12], its involvement in disease progression is worthy of further investigations as a potential biomarker. Using UMAP, we observed that the distribution of the patients in the biomarker enrichment showed a separation between survivors and non-survivors (Figure 1B). No specific pattern or clustering was observed for comorbidities (Fig 1C), the number of comorbidities (Fig 1D), or diabetes (Fig 1E), indicating that these metabolite sets only differentiate related to the clinical outcome.

To further understand the patterns of metabolic changes related to COVID-19, we performed a weighted correlation network analysis on the metabolomics dataset, using significant positive correlations (Spearman, adjusted \( p<0.00001 \)). We identified 10 metabolites communities highly connected in a network of 916 nodes (metabolites) and 11453 edges (Figure 1F). Six predicted biomarkers (out of the 10 previously identified with \( p<0.025 \), unadjusted) were highly correlated and were present in the network and belong to the lipid pathways. These metabolites are also known to be associated with peroxisomal fatty acid oxidation disorders (3-hydroxysebacate) [13] or insulin resistance (5-dodecenate (12:1n7), tetradecadienoate (14:2)* and myristoleate (14:1n5)) [14].

Finally, to further discriminate the subset of metabolites significantly associated with a fatal clinical outcome, we selected all the biomarkers and their first neighbors in the network analysis previously described (238 metabolites). Based on these data, we found a clear clustering of non-survivors in opposition to survivors (Fig 1G). No clustering according to mask type was observed, indicating that the metabolic signature associated with mortality appears to be independent of the oxygen demand at the moment of hospitalization, providing the first identified correlation between a metabolite profile and disease severity in COVID-19 patients.

It is important to consider that when this study was planned (March-April 2020), there was little knowledge about the COVID-19. Therefore, several clinical data were missing. Despite that, this is the first set of biomarkers identified from high throughput metabolomics data that are associated with mortality and are not confounded by other preexisting conditions.

**Conclusions**
Our analysis has identified metabolic biomarkers that in our data differentiate between COVID-19 survivors and non-survivors and that may be predictive of death from COVID-19, from the early stage of the epidemic, independently from oxygen demand at the moment of diagnosis. Our results on high throughput metabolomics contribute to a better understanding of COVID-19-related metabolic disruption and may represent a useful starting point for the identification of independent prognostic factors to be employed in the therapeutic practice.

Authors' contribution


Funding

The study was supported by Swedish Research Council Grants (2017-01330) and Karolinska Institute Stiftelser och Fonder (2020-01554) to U.N and by Office of Research, University of Missouri, Columbia, USA to K.S. E.S. was partially supported by VR (2020-05836).

Declaration of competing interest

C.L.L is co-founder and CSO of Shift Pharmaceuticals. All other authors declare no competing conflicts of interest.

Acknowledgments

The authors would also like to thank all patients involved in this study, as well as the dedicated medical and research staff who fight against SARS-CoV-2. We also like to acknowledge the work of the “COVID-19 NETWORK” working group:

Responsible person: Prof. Silvano Bosari
Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Milan, Italy

Scientific Direction: Silvano Bosari, Luigia Scudeller, Giuliana Fusetti, Laura Rusconi, Silvia Dell’Orto. Transfusion Medicine (Biobank): Daniele Prati, Luca Valenti, Silvia Giovannelli.

Infectious Diseases Unit: Andrea Gori, Alessandra Bandera, Antonio Muscatello, Davide Mangioni, Laura Alagna, Giorgio Bozzi, Andrea Lombardi, Riccardo Ungaro, Teresa Itri, Valentina Ferroni, Valeria Pastore, Roberta Massafra, Ilaria Rondolini, Angelo Bianchi Bonomi. Internal Medicine,

198 References


## Table 1: Characteristics of the study population

<table>
<thead>
<tr>
<th></th>
<th>Survivors</th>
<th>Non-survivors</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number</strong></td>
<td>57</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td>0.1169</td>
</tr>
<tr>
<td>- Male (%)</td>
<td>17 (30.4%)</td>
<td>9 (50%)</td>
<td></td>
</tr>
<tr>
<td>- Female (%)</td>
<td>40 (69.4%)</td>
<td>9 (50%)</td>
<td></td>
</tr>
<tr>
<td><strong>Age in years, median (IQR)</strong></td>
<td>62 (52-73.5)</td>
<td>76.5 (68.5-82)</td>
<td>0.0013</td>
</tr>
<tr>
<td><strong>BMI, median (IQR)</strong>*</td>
<td>26.3 (24.4-29.5)</td>
<td>29.6 (27-32.2)</td>
<td>0.1697</td>
</tr>
<tr>
<td><strong>SpO2, median (IQR)</strong>*</td>
<td>97 (94.25-98)</td>
<td>95 (89.25-98)</td>
<td>0.0768</td>
</tr>
<tr>
<td><strong>pH, median (IQR)</strong>*</td>
<td>7.48 (7.45-7.5)</td>
<td>7.49 (7.45-7.53)</td>
<td>0.6672</td>
</tr>
<tr>
<td><strong>Lactate, median (IQR)</strong>*</td>
<td>0.9 (0.7-1.1)</td>
<td>1.5 (1.1-1.8)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Comorbidities (%)</strong></td>
<td></td>
<td></td>
<td>0.0297</td>
</tr>
<tr>
<td>- Hypertension</td>
<td>26 (46.4%)</td>
<td>9 (50%)</td>
<td>0.7913</td>
</tr>
<tr>
<td>- Diabetes</td>
<td>9 (16%)</td>
<td>3 (16.7%)</td>
<td>NS</td>
</tr>
<tr>
<td>- Lung disease</td>
<td>4 (7.1%)</td>
<td>1 (5.6%)</td>
<td>NS</td>
</tr>
<tr>
<td>- COPD</td>
<td>4 (7.1%)</td>
<td>2 (11.1%)</td>
<td>0.6257</td>
</tr>
<tr>
<td>- Obesity</td>
<td>8 (14.3%)</td>
<td>6 (33%)</td>
<td>0.0867</td>
</tr>
<tr>
<td>- Renal disease</td>
<td>4 (7.1%)</td>
<td>3 (16.7%)</td>
<td>0.3481</td>
</tr>
<tr>
<td>- Liver disease</td>
<td>1 (1.8%)</td>
<td>1 (5.6%)</td>
<td>0.4249</td>
</tr>
<tr>
<td>- Cardiovascular disease</td>
<td>7 (12.5%)</td>
<td>7 (38.9%)</td>
<td>0.0320</td>
</tr>
<tr>
<td>- Cerebrovascular disease</td>
<td>2 (3.6%)</td>
<td>0</td>
<td>NS</td>
</tr>
<tr>
<td>- Others**</td>
<td>14 (24.6%)</td>
<td>11 (61.1%)</td>
<td></td>
</tr>
<tr>
<td><strong>Number of comorbidities</strong></td>
<td></td>
<td></td>
<td>0.0047</td>
</tr>
<tr>
<td>(listed above):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- None</td>
<td>19 (33.4%)</td>
<td>1 (5.6%)</td>
<td></td>
</tr>
<tr>
<td>- One</td>
<td>16 (28%)</td>
<td>3 (16.7%)</td>
<td></td>
</tr>
<tr>
<td>- Two</td>
<td>9 (15.8%)</td>
<td>6 (33%)</td>
<td></td>
</tr>
<tr>
<td>- Three or more</td>
<td>13 (22.8%)</td>
<td>8 (44.4%)</td>
<td></td>
</tr>
<tr>
<td><strong>Disease severity$$:</strong></td>
<td></td>
<td></td>
<td>0.0218</td>
</tr>
<tr>
<td>- Moderate</td>
<td>42 (73.7%)</td>
<td>8 (44.4%)</td>
<td></td>
</tr>
<tr>
<td>- Severe</td>
<td>15 (26.3%)</td>
<td>10 (55.6%)</td>
<td></td>
</tr>
</tbody>
</table>

*Calculated from the available data
**Gastrointestinal reflux, dyslipidemia, neoplasia, dementia
§ Based on the masks
Figure legends:

Figure 1: 10 metabolites are significantly associated with COVID-19 related in-hospital mortality.

A. All biomarkers that were significant at the 0.05 level, after adjustment of age, gender, and BMI, were included in the plot and order by the value of the odds ratio. Color coding for lower p-values is presented in the legend.

B-E. UMAP visualization of patients’ data after selecting the 10 death biomarkers. Patients were labeled for the B. outcome, C. presence of comorbidities, D. amount of comorbidities and E. diabetes.

F. Global weighted network after communities’ detection. Biomarkers with COVID-19 related mortality belong to communities 4 and 5 (labeled). The size of the bubble represents the connectivity of each metabolite.

G. Heatmap of potential biomarkers and first neighbors in patients. Outcome (survivors/non-survivors) and mask type is indicated. Data were log-transformed.