Early and accurate prediction of COVID-19 hospitalization risk
and symptomatic course of SARS-CoV-2 infection

Corinna Holetschek\textsuperscript{1,2,3,*}, Mehmet Goekkaya\textsuperscript{1,2,*}, Karim Dorgham\textsuperscript{4,*}, Paul Eder\textsuperscript{1}, Daria Luschkova\textsuperscript{1}, Christophe Parizot\textsuperscript{5}, Mehmet Tekinsoy\textsuperscript{6}, Denise Rauer\textsuperscript{1,2}, Assia Samri\textsuperscript{4,5,}

Early-Opt-COVID19 study group\textsuperscript{5}, Matthias Reiger\textsuperscript{1,2}, Gertrud Hammel\textsuperscript{1,2}, Philipp Steininger\textsuperscript{7},

Stefanie Gilles\textsuperscript{1}, Ulrike Protzer\textsuperscript{6}, Christoph Römmele\textsuperscript{8}, Guy Gorochov\textsuperscript{4,5,9,}, Claudia Traidl-Hoffmann\textsuperscript{1,2,3,10,*} and Avidan U. Neumann\textsuperscript{1,2,*}

1 Environmental Medicine, Faculty of Medicine, University of Augsburg, Augsburg, Germany
2 Institute of Environmental Medicine, Helmholtz Munich, Augsburg, Germany
3 Chair of Environmental Medicine, Technical University of Munich, Augsburg, Germany
4 Centre d’Immunologie et des Maladies Infectieuses (CIMI-Paris), National Institute for Health and Medical Research (INSERM), Paris, France
5 Département d’Immunologie, Hôpital Pitié-Salpêtrière, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France
6 Institute of Virology, Technical University of Munich, Munich, Germany
7 Institute of Clinical and Molecular Virology, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
8 Internal Medicine III - Gastroenterology and Infectious Diseases, University Hospital of Augsburg, Augsburg, Germany
9 Sorbonne Université, Paris, France
10 CK-CARE, Christine Kühne Center for Allergy Research and Education, Davos, Switzerland

* Authors contributed equally as first authors.
# Authors contributed equally as last authors

Corresponding author: Avidan U. Neumann, Environmental Medicine, Faculty of Medicine, University of Augsburg, Neusaßser Str. 47, D-86156, Augsburg, Germany
E-mail: avidan.neumann@uni-a.de

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
$ Early-Opt-COVID19 study group:

Mustafa, Akaryildiz¹, Linda Bachmann¹, Anna Bayer¹, Barbara Böglmüller¹, Omaira da Mata-Jardin², Amadeo DeTomassi¹, Christin Dhillon³, Victoria Fairweather¹, Marie Freitag³, André Fuchs³, Anja Fusco¹, Carina Gülzow¹, Hans Hauner⁴,⁵, Claudia Khalfi¹, Lena Klepper¹, Elisabeth Kling⁶, Silke Kotschenreuther¹, Helmut Messmann³, Aline Metz¹, Anna Muzalyova⁷, Franziska Neitzel¹, Nicole Pochert¹, Claudia Ranieri¹, Monika Seemann¹, Magnus Wangenheim¹, Katharina Zeiser¹

¹ Environmental Medicine, Faculty of Medicine, University of Augsburg, Augsburg, Germany
² Centre d’Immunologie et des Maladies Infectieuses (CIMI-Paris), National Institute for Health and Medical Research (INSERM), Paris, France
³ Internal Medicine III - Gastroenterology and Infectious Diseases, University Hospital of Augsburg, Augsburg, Germany
⁴ Technical University of Munich, Else Kröner-Fresenius-Center for Nutritional Medicine, Chair of Nutritional Medicine, School of Life Sciences
⁵ Institute for Nutritional Medicine, School of Medicine, Technical University of Munich
⁶ Institute for Lab Medicine and Microbiology, Uniklinikum Augsburg, Augsburg, Germany
⁷ Institute for Digital Medicine, Uniklinikum Augsburg, Augsburg, Germany
Abstract

**Background:** Whilst SARS-CoV-2 infection has become endemic, COVID-19 related hospitalization and mortality are still considerably high. Both anti-viral and immune modulating therapies against COVID-19 are available, but they must be initiated early after infection and given only to patients of need. Currently, patients’ demographics and clinical pre-conditions factors are used to determine treatment eligibility. However, the latter do not provide accurate prediction and there are no useful biomarkers for early accurate prediction of COVID-19 related hospitalization risk and disease progression.

**Methods:** Non-vaccinated patients (N=185) were recruited early after the first positive SARS-CoV-2 test. Biochemistry, hematology and 8 serum cytokine levels were longitudinally measured within the first month.

**Findings:** Early levels of LDH, IL-6 or CRP, each alone or their combinations, were identified as accurate predictors for the risk of hospitalization (sensitivity=93.6-100%, specificity=93.4-96.7%, p<0.0001). Moreover, the combination of 4 cytokines (IFN-α, IFN-γ, IL-6, IL-17A) was the only accurate predictor for symptoms risk (sensitivity=97.5%, specificity=92.3%, p<0.0001). In comparison, age and BMI showed significantly lower predictive values than above biomarkers. Prediction with above biomarkers was independent of sampling time (0-11 days post symptoms onset), age, gender, BMI, clinical pre-conditions or SARS-CoV-2 variant. Furthermore, the early higher levels of LDH, CRP and inflammatory cytokines in hospitalized, as compared to non-hospitalized, patients, stayed consistently higher for at least 4 weeks.

**Interpretation:** The risk for COVID-19 hospitalization or symptoms can be accurately predicted as early as the time of the first positive SARS-CoV-2 test, with biomarkers that are feasibly measurable at point-of-testing. These findings could allow for better early personalized treatment and optimization of clinical management of COVID-19 patients.
Source of funding

This study was funded by the Bavarian State Ministry for Science and Art Program for the funding of Corona research (Early-Opt-COVID19project and research networks FOR-COVID and Bay-VOC), by grants from the Helmholtz Association's Initiative and Networking Fund (KA1-Co-02 “COVIPA” to UP; KA1-Co-06“CORAERO” to CH, MG and AUN), by the European Commission FET Open Grant VIROFIGHT (grant no. 899619), and by the Sanddorf foundation. Work in the Gorochov laboratory was supported by Institut National de la Santé et de la Recherche Médicale (INSERM), Sorbonne Université, Fondation pour la Recherche Médicale (FRM), Paris, France, program “Investissement d’Avenir” launched by the French Government and implemented by the Agence Nationale de la Recherche(ANR) (programme COFIFERON ANR-21-RHUS-08), by EU Horizon HLTH-2021-DISEASE-04UNDINE project, by Fondation pour la Recherche Médicale, Paris, France (programme Equipe FRM 2022) and by the Département Médico-Universitaire de Biologie et Génomique Médicales (DMU BioGen), APHP, Paris, France.

Acknowledgments

We thank the patients for their participation in the study. We thank the PerForM-REACT Project (funded by the Free State of Bavaria and the European Regional Development Fund - ERDF) for use of devices and consumables. The French-German collaboration was additionally supported by the BayFrance foundation.

Disclosure statement

UP received personal fees from Abbott, Abbvie, Arbutus, Gilead, GSK, Leukocare, J&J, Roche, MSD, Sanofi, Sobi and Vaccitech. UP is a co-founder and share-holder of SCG Cell Therapy. All other authors declare no conflict of interest. The study was independently designed, run, analyzed and summarized by the authors with no involvement from the funding
agencies. The manuscript was written by the authors with no involvement of the funding agencies. The funding agencies did not pay for or were involved in any way of writing the manuscript.

**Keywords**

SARS-CoV-2, COVID-19, Prediction, Biomarkers, Inflammatory cytokines, Type-I Interferon, LDH, CRP, IL-6

**Total Word Count main text:** 3753
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CK</td>
<td>Creatin Kinase</td>
</tr>
<tr>
<td>CoVaKo</td>
<td>Corona-Vakzin-Konsortium</td>
</tr>
<tr>
<td>COVID</td>
<td>Coronavirus Disease</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>DPSO</td>
<td>Days Post Symptom Onset</td>
</tr>
<tr>
<td>ECMO</td>
<td>Extracorporeal membran oxygenation</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>EOC</td>
<td>Early Opt Covid 19 Study</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquantile Range</td>
</tr>
<tr>
<td>J-Index</td>
<td>Youden Index</td>
</tr>
<tr>
<td>LDH</td>
<td>Lactate Dehydrogenase</td>
</tr>
<tr>
<td>MVS</td>
<td>Mechanical Ventilation</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PLT</td>
<td>Platelet</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiving Operator Curve</td>
</tr>
<tr>
<td>SARS-CoV</td>
<td>severe-acute-respiratory-syndrome-related coronavirus</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumor necrosis factor</td>
</tr>
<tr>
<td>WBC</td>
<td>White Blood Cell</td>
</tr>
</tbody>
</table>
Introduction

The SARS-CoV-2 virus started spreading in December 2019 causing a total of 6 million COVID-19 related deaths by December 31, 2021. In the last 2 years, SARS-CoV-2 became endemic, but COVID-19 mortality is still considerably high, with about 200,000 deaths in 2023, even with the less severe Omicron variant. Several Anti-viral drugs are available and more are under investigation, also for patients with mild-moderate disease. However, these need to be initiated early within 5-7 days post symptoms onset (DPSO) and targeted to patients that are of risk for hospitalization or develop considerable symptomatic disease. Also, immune modulating drugs that inhibit inflammatory cytokines elevated as part of the COVID-19 cytokine storm (e.g., IL-6 inhibitors) are available. However, currently, anti-viral therapy is underused and cytokine inhibitors are only given to hospitalized severe patients, mainly because of lack of biomarkers that allow early stratification of the appropriate target patients. Therefore, it is important to understand which SARS-CoV-2 infected individuals are at risk for hospitalization and/or to suffer from symptomatic mild to moderate disease, in order to allow early personalized treatment and optimize public health management.

COVID-19 shows a versatile range of severity, from asymptomatic, to symptomatic mild-moderate disease and up to critical severity leading to hospitalization, requiring invasive oxygen support, and potentially leading to multi-organ failure and death. COVID-19 has a mean incubation period 5.2 days (95% CI: 4.9–5.5) until symptoms onset. Common early symptoms are characterized by fever, cough, fatigue, diarrhea, and dyspnea. However, meta-analysis showed that 40-50% of confirmed SARS-CoV-2 positive cases are asymptomatic. The Kings College large app study identified the early symptoms that are most associated with subsequent hospitalization and thus allowed stratification of patients into low versus high symptomatic groups.
Patients’ characteristics, such as age and BMI, and underlying comorbidities (e.g., diabetes, asthma, cancer, cardiovascular disorder, chronic kidney disease) are known factors associated with risk for severe COVID-19 disease\textsuperscript{17} and are currently used as recommendations for early treatment with anti-viral therapy\textsuperscript{18}. However, these factors do not allow accurate prediction of hospitalization and symptomatic risk. In addition, a number of biochemistry and hematological blood biomarkers were reported to be risk factors for severe COVID-19 in hospitalized patients. C-reactive protein (CRP) levels were found to be significantly higher in severe hospitalized COVID-19 compared to moderate cases\textsuperscript{19}, and elevated blood lactate dehydrogenase (LDH) levels were reported to be associated with a worse outcome in hospitalized patients\textsuperscript{20}. Furthermore, low lymphocyte count was also associated with a more severe COVID-19 disease course\textsuperscript{21}.

Additionally, as severe cases of COVID-19 suffer from a "cytokine storm"\textsuperscript{22,23} which is an abnormal regulation and excessive release of different pro-inflammatory cytokines as a response to the infection, various cytokines were identified as biomarkers predictive of severe COVID-19. IL-6 is known to play a role as a prognostic marker in COVID 19 since most patients show elevated IL-6 levels during a SARS-CoV-2 infection\textsuperscript{23-25}. Several studies have shown high IL-10 levels in severe COVID-19\textsuperscript{26,27}. Similarly, IL-17A\textsuperscript{28} and tumor necrosis factor alpha (TNF-\(\alpha\))\textsuperscript{25,29} were reported to be associated with disease severity and progression. As to type-I interferon response, it was reported that hospitalized patients with increased IFN-\(\alpha\) levels showed improvement in the COVID-19 disease course\textsuperscript{30}. In contrast, high IFN-\(\gamma\) levels in hospitalized cases were shown to be correlated with a worse diagnosis\textsuperscript{31}. We had previously reported that the ratio between type-I interferon response and inflammatory cytokines, at the day of hospitalization, shows the highest accuracy for predicting COVID-19 severity and mortality\textsuperscript{32}.
However, all the above studies only tested biomarkers predicting severity in patients that were already hospitalized. It would be rather highly beneficial to discover biomarkers for early predictors of severity at the earliest time possible (i.e., as soon as possible after the first positive SARS-CoV-2 diagnosis and within 5-7 DPSO), which are feasible to measure at the point-of-testing. Thus, here we are combining biochemistry, hematological and cytokine biomarkers to predict as early as possible the risk of hospitalization, and the symptomatic versus asymptomatic course of infection.

Methods

Patients and study cohorts

The Early-opt-COVID-19 study, a longitudinal study was performed at the University Hospital Augsburg and approved by the local ethics committee of the Technical University of Munich (internal code 799/20 S). Adult study participants were enrolled after written informed consent. Patients were included at the day of the SARS-CoV-2 PCR test at the test center or within 1-5 days after the PCR test through passive recruitment via flyers. The main inclusion criteria were either the presence of COVID-19-related symptoms or contact with a SARS-CoV-2 infected person within the last five days. Exclusion criteria encompassed individuals with chronic virus infections, individuals taking immunosuppressants, and individuals with immune deficiency or pregnancy. Only non-vaccinated and pre-Omicron infected patients were included in the analysis, in order to avoid the effect of a COVID-19 vaccination or infection with less severe variants such as omicron (B.1.1.529). Patients were divided into two main groups, non-hospitalized and hospitalized patients.

Visit 1 at the day of PCR test, for patients recruited at the test center, was available for 36 patients, while the other patients (n=56) started at visit 2 (1-5 days after PCR test), with following visit 3 (2-3 weeks after) and visit 4 (4-5 weeks after). Only patients with their first
sample available at DPSO≤11 were included in the analysis, since at later time points there was a significant decline in the various biomarker levels. The patients completed a symptoms questionnaire at each visit to classify the disease course.

From 1.2.2021 until 14.9.2021, n=115 patients were recruited in Augsburg, of which 83 non-hospitalized patients were recruited within the Early-Opt-COVID-19 study, n=9 non-hospitalized patients were included within the COVID-19 Vaccine Consortium (CoVaKo) study and 23 hospitalized patients were included at the University Hospital Augsburg. In addition, we used results from our previous study for n=70 patients hospitalized at the Pitié-Salpêtrière hospital in Paris. Additionally, n=14 control subjects with negative by SARS-CoV-2 PCR test were enrolled. Only patients with the first visit at most 11 days post symptoms onset (DPSO average 6.2, IQR: 4-9) were included.

Demographic and clinical patient characteristics are summarized in Table 1. Hospitalized patients were divided into three groups according to the oxygen support needed. Patients ventilated with nasal cannula, oxygen mask or without oxygen support were grouped as No-MVS (n=36). Mechanical ventilatory support (MVS, n=43) patients required invasive mechanical ventilation and ECMO patients (n=14) patients required extracorporeal membrane oxygenation. The non-hospitalized patients were divided into 3 symptom groups (Table 1 and Supplementary Figure 1): no symptoms (asymptomatic), mild and few flu-like symptoms (low-symptomatic), and at least 2 out of 5 severe (confusion, shortness of breath, fever, fatigue, diarrhea) symptoms (high-symptomatic), according to the Kings College. Low-symptomatic patients exhibited on average 3.75 symptoms, significantly (p<0.001) lower than 7.6 symptoms in high-symptomatic patients. More details in supplementary methods.

**Biological sampling**

On visit 1, only one serum Gel S-Monovette (SARSTEDT) was drawn from the patients. On visits 2-4, four serum Gel S-, a citrate, and an EDTA Monovettes of blood were drawn. On
visits 2–4 an oropharyngeal swab was taken and stored in Guanidine Thiocyanate (PrimeStore MTM) medium for RNA stabilization. More details in supplementary methods.

**Biochemistry, hematology, immunoassays and SARS-CoV-2 variant**

From two serum Gel S, a citrate and an EDTA Monovette, biochemistry, hematology and differential blood count measurements for: CRP, LDH, creatine kinase (CK), platelet (PLT), ferritin, percent of lymphocytes count and neutrophils count, were performed at the University Hospital Augsburg by standard clinical procedures.

Serum Gel S-Monovette was used to assess cytokines, chemokines and interferons levels. IFN-γ, IL-6, IL-8, IL-10, IL-22, and TNF-α were measured using multiplex planar array immunoassay (CorPlex Human Cytokines 10-Pley Panel on SP-X Platform Quanterix). IFN-α and IL-17A were measured by ultra-sensitive single-plex bead-based Simoa assay (HD-1 Analyser, Quanterix).

SARS-CoV-2 variant detection was performed by cDNA synthesis followed by ARTIC-PCR, Nextera XT (Illumina) library preparation and sequencing on Illumina NextSeq 1000. More details in supplementary methods.

**Bioinformatics and statistical analysis**

Descriptive and statistical analyses were performed using R and Python. Batch correction was applied to mitigate batch effect bias. Principal component analysis (PCA) was used to cluster patients and compute principal component weights for each parameter. Predictive value of the various biomarkers was evaluated by the Youden J-index (Informedness, Specificity + Sensitivity -1), since accuracy is to be taken only indicatively due to the ratios between number of patients in the different groups does not represent real world ratios. The threshold used for prediction was selected as the point with the highest J-index on the ROC curve for each biomarker. Longitudinal data was analyzed using piecewise linear regression and mean...
negative control data was used as a baseline. Statistical significance between groups was evaluated using the nonparametric Mann-Whitney-U test for continuous variables and the Fisher-Exact test for discrete variables. Correlations were evaluated using the nonparametric Spearman test. A two-sided p-value lower than 0.05 was considered statistically significant. Since 19 biomarkers and 5 biomarker combinations were tested, by the Bonferroni rule a p-value of lower than 0.003 was considered significant after multiple testing correction. More details in supplementary methods.

Results

Hospitalization risk prediction

Of all blood biochemistry and hematology biomarkers measured at the first sample per patient (0-11 days post symptoms onset), LDH and CRP showed the most significant (p<0.001) elevation in hospitalized versus non-hospitalized patients (Fig 1A-B), and were the only ones to show accurate prediction of hospitalization risk with a Youden J-index larger than 0.8 (LDH: J-index=91.8%, sensitivity=96.2% and specificity=95.6%; CRP: J-index=87.8%, sensitivity=94.4% and specificity=93.4%; Fig 1E, Table 2). D-dimers, ferritin, creatinine kinase and neutrophils percentage were also significantly elevated, while lymphocytes percentage was significantly lower, in hospitalized patients, but none of those showed a good prediction (Supp Figures 2,4,7).

Among the 8 cytokines measured, serum IL-6 level was the most significantly (p<0.001) elevated in hospitalized patients (Fig 1C), and the only cytokine that accurately predicted hospitalization risk (J-index=90.3%, sensitivity=93.6% and specificity=96.7%; Fig 1E and Table 2). IL-8, IL-10, IL-17A, IL-22, TNF-α and IFN-γ were also significantly elevated in hospitalized versus non-hospitalized patients, but none of those showed a good prediction of hospitalization risk (Supp Figures 3,4,7). IFN-α rather shows a trend for lower levels in
hospitalized patients, especially in ECMO patients, versus symptomatic patients (Supp Figures 3-4).

Next, we used principal component analysis (PCA) for a multidimensional analysis of all the cytokines normalized (Z-scaled) serum levels. Starting with a PCA including all patients and all 8 cytokines, a clear separation was observed between hospitalized and non-hospitalized patients (Supp Figure 5A). Dimension reduction showed a good separation also in PCA with only 4 cytokines (IL-6, IL-10, TNF-α and IFN-α), which was also validated also for the cytokine raw values (Supp Figure 5B). The separatrix line between hospitalized and non-hospitalized patients was then translated to a prediction function of the ratio between the inflammatory cytokines and IFN-α, (termed here the IFN-INF prediction function),

\[
\text{IFN-INF} = \log \left( \frac{2 \times \text{IL-6} \times \text{IL-10} \times \text{TNF-α}}{\text{IFN-α}} \right)
\]

which was significantly (p<0.001) higher in hospitalized patients (Fig 1D) and resulted in accurately predicting hospitalization risk with J-index=91.3% (sensitivity=98.9% and specificity=92.4%; Fig 1E and Table 2). The combination of 3 inflammatory cytokines as a ratio to IFN-α in the IFN-INF function somewhat increases the predictive value compared to IL-6 alone.

ROC analysis comparison of the prediction obtained by each cytokine alone or each of the biochemistry and hematology parameters as well as the IFN-INF combined cytokines prediction function (Fig 1E, Supplementary Figure 7), shows that LDH and the IFN-INF function have the largest area under the curve, although not statistically significantly better than CRP or IL-6 alone. Age and BMI are significantly higher in hospitalized patients but show significantly lower predictive value for hospitalization risk (Supp Figures 7-8).

While LDH, CRP, IL-6, and the IFN-INF function, are all correlated with age and BMI, their predictive value is independent of these factors or of gender and diabetes status (Supp Fig 8-
Importantly, these 4 predictors are not correlated with, and their predictive value is independent of, the time the sample was taken at the range of 0-11 DPSO or the study site (Supp Figures 8-9). Also, the separation between hospitalized and non-hospitalized patients for these 4 biomarkers is observed over all SARS-CoV-2 variants in the study (Supp Figures 8-9).

Lastly, we investigated if combining the biochemistry and cytokine biomarkers increases the prediction accuracy. The IFN-INF cytokine ratio, IL-6, LDH and CRP are all significantly correlated with each other (Supp Figure 6). Nevertheless, combining (with an OR function) LDH with either IL-6 or the IFN-INF function allows the increase of sensitivity to 100% with only a slight reduction of specificity to 92.3-94.5% (Table 2). Similarly, combining (LDH or CRP) gives sensitivity=100% and specificity=94.5%, although it needs to be noted that CRP data was not available for 12% of patients. All other double or triple biomarker combinations do not add predictive value to the corresponding single biomarker predictions.

**Symptom risk prediction in non-hospitalized patients**

Next, we focused on non-hospitalized patients to discover a biomarker predicting symptomatic versus asymptomatic course of infection. CRP, but not LDH or any other biochemistry biomarker, was significantly elevated in symptomatic versus asymptomatic patients (Fig 2A, Supp Figures 2,4). Platelet count was significantly lower in symptomatic patients (Supp Fig 2,4). None of the biochemistry, hematology or differential blood count biomarkers was predictive for symptoms risk in non-hospitalized patients (Supp Fig 7D).

Among the cytokines, IFN-α, IFN-γ, IL-6, IL-10, IL17-A and IL-22 were significantly elevated in symptomatic versus asymptomatic patients (Fig 2B-C, Supp Fig 3-4). However, only a modest (J-index= 64-72%) predictive value for symptomatic risk was found only for
IFN-α and IL-6, but for none of the other cytokine nor for the IFN-INF cytokine ratio (Fig 2E, Table 2, Supp Fig 7C).

Therefore, we investigated whether another combination of cytokines could predict the symptomatic risk. PCA performed for only non-hospitalized (asymptomatic versus symptomatic) patients with all 8 cytokines shows some separation between these groups (Supp Figure 5C). Using the same methods and dimension reduction as for the IFN-INF prediction, we found that IFN-α, IL-17A, IL-6, and IFN-γ are the most important cytokines for separation between asymptomatic patients and symptomatic patients. Translation into raw values of these cytokines shows clear separation between asymptomatic and symptomatic ((Supp Figure 5D), and using the PC factors and the separatrix function parameters we obtained the IL-symp prediction function

\[
\text{IL-symp} = -1.6\times(0.175\times\log(\text{IFN-γ})+\text{IL-6})+(0.3\times\log(\text{IFN-α})+\log(\text{IL-17A}))
\]

being significantly (p<0.0001) elevated in symptomatic patients (Fig 2D). The symptomatic risk is accurately (J-index= 89.8%) predicted by the IL-symp cytokine function with a sensitivity of 97.5% and a specificity of 92.3% (Fig 2E, Table 2). IL-symp prediction function is also predictive of hospitalization risk but less accurately than LDH, IFN-INF, IL-6 or CRP (Supp Fig 7).

While IL-symp is correlated with age and BMI, its predictive value is independent of these factors or of gender and diabetes status (Supp Fig 10). IL-symp value is somewhat decreased after 7 DPSO, but its predictive value is independent of the time the sample was taken, within the range of 0-11 DPSO, or of the SARS-CoV-2 variant (Supp Fig 10). CRP, LDH or other biochemistry parameters are not strongly correlated with IL-symp, and the combination of IL-symp (or any single cytokine) with the biochemistry biomarkers does not improve its predictive value.
Decision tree for combined prediction of hospitalization risk and symptomatic status

Thus, by measuring only LDH and 4 cytokines (IL-6, IL-17A, IFN-α and IFN-γ) it is possible to obtain a combined decision tree (Fig 3), predicting in our study hospitalized patients with sensitivity of 100% (using the IL-6 and LDH combination) and among those predicted not to be hospitalized a prediction of symptomatic patients (using IL-symp) with sensitivity of 97.5%.

Longitudinal profiles of cytokine and biochemistry biomarkers

The longitudinal measurements starting as early as the day of (max 5, median 4.6, days after) the PCR test and up to 5 weeks DPSO, reveal differences in the kinetic profiles of different biomarkers between hospitalized as well as between non-hospitalized and symptomatic and asymptomatic patients. Using piecewise linear regression for each group of patients (except for ECMO patients that had no longitudinal data) the average biomarker kinetics profile per severity or symptom group is given in Figure 4. LDH, CRP, IL-6, IL-10 and TNF-α are elevated in hospitalized in comparison to non-hospitalized patients already at 0-7 DPSO, and stay significantly elevated up to 28 DPSO, with no significant difference between the low- and high-symptomatic patients. On the other hand, IFN-α, IFN-γ and IL-17A levels show an increase with the degree of symptoms, but no significant difference between hospitalized patients and high-symptomatic non-hospitalized patients.

Discussion

Our findings indicate that it is possible to predict COVID-19 disease progression with biomarkers feasibly measurable at point-of-testing already as early as 0-7 days post symptoms.
onset and even at the time of the first positive diagnosis of SARS-CoV-2 infection. Most importantly, we have shown that highly accurate (J-index >90%) early prediction of COVID-19 hospitalization risk is possible either by high LDH levels, or high IL-6 levels, or by the ratio between IFN-α and the inflammatory cytokines IL-6, IL-10 and TNF-α, termed here the IFN-INF predictor function. Moreover, combining the LDH with IL-6 biomarkers enables a predictor with sensitivity of 100%. It should be noted that all above predictors are not statistically significantly different in their predictive values. While all above biomarkers have been previously shown to predict COVID-19 severity in hospitalized patients, our results now show that they are also good early predictors for hospitalization risk.

Furthermore, we have shown that a combination of the IFN-α, IFN-γ, IL-6 and IL-17 cytokines, termed here the IL-symp prediction function, is an early and accurate (J-index=89.8%) predictor of symptomatic versus asymptomatic course of infection. Interestingly, neither any single cytokine alone, nor the IFN-INF cytokine combination, nor any of the other biochemistry or hematological parameters, were a good predictors for symptoms risk. While the prediction of symptomatic status is clinically less important, still it could be used to guide personalized treatment with SARS-CoV-2 anti-viral therapy, as these are proven in clinical trials to reduce the associated quality-of-life impairment and public-health burden in mild-moderate patients.

For a large fraction (39.1%) of the non-hospitalized patients the above predictions are based on samples taken already at the day of the first diagnosis of a positive SARS-CoV-2 at the test center, but none of the patients recruited at the test centers were finally hospitalized. Meanwhile, we show that the above predictions of both hospitalization and symptoms risk are independent of the time of sampling.

We also found that the risk of hospitalization and the risk of symptomatic status are associated with older age and higher BMI, similar to previous studies showing these factors associated
with COVID-19 severity\textsuperscript{17,18}. However, we show that the predictors we have identified are independent of such factors as age, gender, BMI or diabetes, probably since these are already factorized into the early levels of the biomarkers in response to SARS-CoV-2 infection.

Our results also shed further light on the role of the different cytokines in response to SARS-CoV-2 infection. While high levels of the inflammatory cytokines (e.g. IL-6, IL-8 and TNF-\(\alpha\)), and also interestingly IL-10, are associated with both the risk for hospitalization and higher risk for symptomatic course of infection, conversely IFN-\(\alpha\), IFN-\(\gamma\) and IL-17 are less associated with hospitalization risk but rather with the appearance of symptoms. In fact, an early high ratio of IFN-\(\alpha\) compared to inflammatory cytokines seems to decrease the risk of hospitalization, similarly to our previous findings about COVID-19 severity in hospitalized patients\textsuperscript{32}. It is important to note that since we aimed at finding the minimum number of cytokines that allow a good prediction, some cytokines (e.g. IL-8) were not included in our prediction function since they do not add significant predictive value, probably because of high inter-correlations, but still are associated with risk of hospitalization or of symptomatic disease. Also, we found that the IFN-INF cytokine combination function is positively correlated with LDH and CRP, indicating a biological link between the different processes that these are biomarkers for. Interestingly, there is no strong correlation between the IL-symp function, predictive of symptoms’ risk, with biochemistry or hematological biomarkers or lymphocyte counts.

Our longitudinal results show that the kinetics of the cytokines IL-6, TNF-\(\alpha\) and IL-10 differ in hospitalized patients, showing higher expression at an early timeframe (already around 0-7 days post symptoms onset) and staying high even 4 weeks after. While it was shown that inflammatory cytokines are high in severe COVID-19 patients\textsuperscript{23-27}, we show here that in some patients they are already high very early on, thus indicating that if immune modulatory treatment is to be successful it should be initiated early personalized according to the cytokine
profile\textsuperscript{11}. As expected, we see also an early rise in IFN-\(\alpha\) and IFN-\(\gamma\), but these are not as significantly different between hospitalized and symptomatic patients and they decline within 2-3 weeks also in hospitalized patients.

The patients analyzed here were all non-vaccinated and recruited prior to the omicron variant appearance. While most of the hospitalized patients in this study were infected with the wild-type virus and the non-hospitalized patients infected with other variants (mostly alpha as well as beta, eta, iota and delta), we also have some hospitalized patients infected with other variants and in general our predictors were not affected by the different variants. Nevertheless, because of the design of our study, we had a bias towards a larger number of symptomatic patients as compared to non-symptomatic patients. Moreover, since none of the patients recruited at the test centers were finally hospitalized, hospitalized patients were recruited at the hospital rather than at the test center, with a larger than realistic ratio of hospitalized to non-hospitalized patients. Due to these imbalances we test our predictors only using sensitivity and specificity, and their combined Youden J-index (Informedness), rather than using accuracy. Considering the above limitations, a larger sequential recruitment study is needed to validate our findings.

Based on our results, we suggest a combined predictive pipeline for early prediction of patients that will become asymptomatic versus symptomatic versus hospitalized. As early as possible after a positive SARS-CoV-2 test, patients should get blood drawn and LDH, IL-6, IL-17A, IFN-a and IFN-g should be measured, which could be performed at point of testing by a multiplex ELISA or other methods. Using these biomarkers the hospitalization risk could be evaluated by LDH combined with IL-6, and the symptomatic risk could be evaluated by the IL-symp function, therefore allowing for personalized treatment with either anti-viral therapy or with cytokine-inhibitor treatment, which can be also guided by the cytokine levels.
With these two populations predicted, personalized therapy approaches can be targeted to the patients with high risk in regards to hospitalization and a symptomatic course.

References


27. Jafrin S, Aziz MA, Islam MS. Elevated Levels of Pleiotropic Interleukin-6 (IL-6) and Interleukin-10 (IL-10) are Critically Involved With the Severity and Mortality of COVID-19: An Updated Longitudinal Meta-Analysis and Systematic Review on 147 Studies. *Biomarker Insights* 2022.

28. Chang Y, Bai M, You Q. Associations between Serum Interleukins (IL-1β, IL-2, IL-4, IL-6, IL-8, and IL-10) and Disease Severity of COVID-19: A Systematic Review and Meta-Analysis. *BioMed Research International* 2022; 2022: 2755246.


Table 1: Patient demographics and clinical parameters by Covid-19 symptoms and severity groups.

<table>
<thead>
<tr>
<th></th>
<th>Non Hospitalized&lt;sup&gt;(1)&lt;/sup&gt;</th>
<th>Hospitalized&lt;sup&gt;(1)&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Asymptomatic (n=13)</td>
<td>Low Symptomatic (n=44)</td>
</tr>
<tr>
<td>Augsburg (n)&lt;sup&gt;(2)&lt;/sup&gt;</td>
<td>13</td>
<td>44</td>
</tr>
<tr>
<td>Paris (n)&lt;sup&gt;(3)&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Age(y), median (IQR)</td>
<td>36.46 (32.0-40.0)</td>
<td>40.39 (27.0-52.0)</td>
</tr>
<tr>
<td>Sex: male, n (%)</td>
<td>6 (46.15)</td>
<td>20 (45.45)</td>
</tr>
<tr>
<td>Days since symptom start, Median in days (IQR)</td>
<td>7.54 (6.0-10.0)</td>
<td>5.02 (2.75-7.0)</td>
</tr>
<tr>
<td>Past medical history, n (%)&lt;sup&gt;(4)&lt;/sup&gt;</td>
<td>10 (76.92)</td>
<td>35 (79.55)</td>
</tr>
<tr>
<td>Obesity (≥ 30)</td>
<td>0 (0)</td>
<td>1 (2.27)</td>
</tr>
<tr>
<td>Diabetes Typ 2</td>
<td>0 (0)</td>
<td>1 (2.27)</td>
</tr>
</tbody>
</table>

<sup>(1)</sup> Hospitalized versus non-hospitalized and asymptomatic versus symptomatic patients were compared using Mann-Whitney or Fisher Exact tests accordingly.

<sup>(2)</sup> Patients included in Augsburg from the Early-Opt (n=83), CoVaKo-Augsburg (n=9)<sup>(3)</sup> and University Hospital Augsburg (n=23) cohorts.

<sup>(3)</sup> Patients included in Paris from our Dorgham et al<sup>(32)</sup> study.

<sup>(4)</sup> Past medical history given here concisely only for the 2 most prevalent conditions, more details in Supplementary Methods section.
Table 2: Biomarkers and biomarker combinations allowing early accurate prediction of COVID-19 hospitalization risk (A) or of symptoms risk in non-hospitalized patients (B).

### A

<table>
<thead>
<tr>
<th>Hospitalization risk predictor</th>
<th>J-Index</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>p-Value</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDH &gt;250</td>
<td>91.8%</td>
<td>96.2%</td>
<td>95.6%</td>
<td>95.9%</td>
<td>&lt;0.0001</td>
<td>169</td>
</tr>
<tr>
<td>IFN-INF &gt;4.25</td>
<td>91.3%</td>
<td>98.9%</td>
<td>92.4%</td>
<td>95.7%</td>
<td>&lt;0.0001</td>
<td>185</td>
</tr>
<tr>
<td>IL-6 &gt;7.46</td>
<td>90.3%</td>
<td>93.6%</td>
<td>96.7%</td>
<td>95.1%</td>
<td>&lt;0.0001</td>
<td>185</td>
</tr>
<tr>
<td>CRP &gt;1.13</td>
<td>87.8%</td>
<td>94.4%</td>
<td>93.4%</td>
<td>93.8%</td>
<td>&lt;0.0001</td>
<td>185</td>
</tr>
<tr>
<td>IL-symp &gt;0.74</td>
<td>84.9%</td>
<td>92.5%</td>
<td>92.4%</td>
<td>92.4%</td>
<td>&lt;0.0001</td>
<td>185</td>
</tr>
<tr>
<td>IL-6 &gt;7.2 or LDH &gt;260</td>
<td>94.5%</td>
<td>100%</td>
<td>94.5%</td>
<td>97.0%</td>
<td>&lt;0.0001</td>
<td>169</td>
</tr>
<tr>
<td>CRP &gt;2.5 or LDH &gt;250</td>
<td>94.5%</td>
<td>100%</td>
<td>94.5%</td>
<td>96.8%</td>
<td>&lt;0.0001</td>
<td>154</td>
</tr>
<tr>
<td>IFN-INF &gt;4.25 or LDH &gt;285</td>
<td>92.3%</td>
<td>100%</td>
<td>92.3%</td>
<td>95.8%</td>
<td>&lt;0.0001</td>
<td>169</td>
</tr>
<tr>
<td>IFN-INF or CRP</td>
<td>Combination does not add predictive value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>162</td>
</tr>
<tr>
<td>IL-6 or CRP</td>
<td>Combination does not add predictive value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>162</td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th>Symptoms risk predictor</th>
<th>J-Index</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>p-Value</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-symp &gt;-2</td>
<td>89.8%</td>
<td>97.5%</td>
<td>92.3%</td>
<td>96.7%</td>
<td>&lt;0.0001</td>
<td>92</td>
</tr>
<tr>
<td>IFN-IF &gt;0.06</td>
<td>71.9%</td>
<td>94.9%</td>
<td>76.9%</td>
<td>92.4%</td>
<td>&lt;0.0001</td>
<td>91</td>
</tr>
<tr>
<td>IL-6 &gt;1.1</td>
<td>64.4%</td>
<td>79.8%</td>
<td>84.6%</td>
<td>80.4%</td>
<td>&lt;0.0001</td>
<td>92</td>
</tr>
</tbody>
</table>

(1) The threshold used for prediction was selected as the point with highest J-index on the ROC curve.

(2) Youden J-Index (Informedness, defined by Specificity + Sensitivity -1) was used to assess predictive value.

(3) Accuracy should be taken only indicatively since the ratio of hospitalized to non-hospitalized, and ratio of symptomatic to asymptomatic, patient numbers in this study does not represent real world ratios. For the same reason, Odds Ratio (OR), Positive Predictive Value (PPV) and Negative Predictive Value (NPV) were not used.

(4) Statistical significance of the prediction was assessed by Fisher-Exact Test on the confluence table.
LDH [log U/l]; CRP [log mg/dl]; All cytokines [log pg/ml].

IFN-INF function is defined by:
\[2(\log(IL-6)+\log(IL-10)+\log(TNF-\alpha) - \log(IFN-\alpha)\]

IL-symp function is defined by:
\[-1.6(\log(IL-6)+0.175*\log(IFN-\gamma)) + (\log(IL-17A)+0.3*\log(IFN-\alpha))\]

Figure Captions

**Figure 1:** Biochemistry and cytokine biomarkers predictive of COVID-19 hospitalization risk.

Early levels of LDH (A), CRP (B) and IL-6 (C), depicted as function of symptoms and severity groups, show significant differentiation between hospitalized and non-hospitalized patients. Also, the ratio between inflammatory cytokine (IL-6, IL-10 and TNF-\(\alpha\)) levels compared to IFN-\(\alpha\) (IFN-INF function, see Supp Figure 5) is significantly higher in hospitalized patients (D). Receiver Operating Characteristic (ROC) curves for IFN-INF, LDH, IL-6 and CRP show significantly larger AUC for the prediction of hospitalization risk as compared to age and BMI (E). Solid horizontal black lines (A, B, C, D) represent the threshold for each biomarker for which the hospitalization risk prediction has the largest Youden J-index (see Table 2). All other biochemistry and cytokine biomarkers show lower prediction of hospitalization risk (Supp Figures 2,3,4,7). Crossed circles depict patients measured already at the day of the COVID-19 PCR test or within 2 days after. IFN-INF function is defined by:
\[
\log(2 \times IL-6 \times IL-10 \times TNF-\alpha / IFN-\alpha).
\]

**Figure 2:** Biochemistry and cytokine biomarkers predictive of COVID-19 symptoms risk. Early levels of CRP (A), IFN-\(\alpha\) (B) and IL-6 (C), depicted as function of symptoms groups, show significant differentiation between symptomatic and asymptomatic patients. Also, the product of IL-6, IL-17A, IFN-\(\gamma\) and IFN-\(\alpha\) cytokine levels (IL-symp function, see Supp Figure 5) is significantly higher in symptomatic patients (D). Receiver Operating Characteristic (ROC) curves of the prediction for symptoms risk show largest AUC for IL-symp as compared to IFN-\(\alpha\), IL-6, CRP, age and BMI (E). Solid horizontal black lines (A, B, C, D) represent the threshold for each biomarker for which the symptoms risk prediction has the largest Youden J-index (see Table 2). All other biochemistry and cytokine biomarkers show no prediction of symptoms risk (Supp Figures 2,3,4,7). Crossed circles
depict patients measured already at the day of the COVID-19 PCR test or within 2 days after. IL-symp
function is defined by: -1.6*(log(IL-6)+0.175*log(IFN-γ)) + (log(IL-17A)+0.3*log(IFN-α)).

Figure 3: Schematic flowchart of the decision tree for the combined prediction for hospitalization risk and symptoms risk. A flowchart showing the stratification of patients in the study to those predicted (using IL-6 or LDH combination) for hospitalization (with 100% sensitivity) versus non-hospitalization, and the latter further stratified to those predicted (with the IL-symp function) to have symptomatic (sensitivity=97.5%) versus asymptomatic course of infection. Only 169 patients are shown since 26 patients were missing LDH values. IL-symp function is defined by: -1.6*(log(IL-6)+0.175*log(IFN-γ)) + (log(IL-17A)+0.3*log(IFN-α)).

Figure 4: Kinetics of the main cytokine and biochemistry biomarkers stratified by the different severity and symptom groups. Cytokine and biochemistry biomarkers in SARS-CoV-2 positive patients, as well as in the negative control group, were measured 3-4 times over the span of 28 days post symptoms onset (DPSO). Different kinetic profiles for LDH (A), CRP (B), IL-6 (C), TNF-α (D), IL-10 (E), IFN-α (F), IFN-γ (G) and IL-17A (H) are observed in hospitalized, high-symptomatic, low-symptomatic and asymptomatic patients as compared to negative controls, using the average piecewise linear regression of the patients in each group (thick lines). The dotted thin lines in the background represent the kinetics of the single patients.

Supplementary figure captions

Supplementary Figure 1: Symptom distribution in different non-hospitalized symptom groups. The distribution of 14 different COVID-19 related symptoms across asymptomatic, low symptomatic and high symptomatic patients in the non-hospitalized cohort. High symptomatic was defined as having at least 2 out of 5 of the symptoms most linked to risk of hospitalization (marked by *).
Significant differences are observed between low and high-symptomatic patients in the frequency of several symptoms as well as for the number of symptoms a patient exhibits.

Supplementary Figure 2: Biochemistry biomarker levels across non-hospitalized symptoms groups and hospitalized severity classes as well as negative control group. The single biomarkers LDH (A), CRP(B), Ceratinkinase (C), Lymphocytes (D), Neutrophiles (E), Platelets (F), Ferritin (G) and D-dimers (H) show significantly different levels between the severity classes and symptom groups. A wide range of levels for the single markers are also found within the negative patient control group. Circles mark patients from the Augsburg study site and squares for Paris.

Supplementary Figure 3: Single cytokine levels across non-hospitalized symptoms groups and hospitalized severity classes as well as negative control group. The levels of single cytokines IL-6 (A), IL-8 (B), IL-10 (C), IL-17A (D), IL-22 (E), TNF-α (F), IFN-α (G) and IFN-γ (H) show significantly different levels between the severity classes and symptom groups. A wide range of levels for the single markers are also found within the negative patient control group. Circles mark patients from the Augsburg study site and squares for Paris.

Supplementary Figure 4: Differences between the levels of cytokine and biochemistry biomarkers across non-hospitalized symptoms groups and hospitalized severity classes as well as negative control group. The difference between the groups is measured by the log2 ratio of more/less severe. The bubble plot depicts the ratio magnitude (bubble color scale) and P-value (bubble size) of the difference between the severity and symptom groups. The order for the groups is asymptomatic, low-symptomatic, high-symptomatic, No-MVS, MVS and ECMO. (A) The differences between the main groups: negative, asymptomatic, symptomatic, non-hospitalized, and hospitalized. (B) The differences between all groups as compared to the negative control groups. (C) Comparison between all sub-groups.

Supplementary Figure 5: Derivation of the multi-dimensional cytokine combination prediction functions for hospitalization risk and symptoms status. PCA of the eight measured cytokines (A) shows differentiation between hospitalized and non-hospitalized patients. After a dimension reduction to determine the minimum number of most predictive cytokines, the ratio between the raw levels of...
inflammatory cytokine (IL-6, IL-10 and TNF-α) compared to IFN-α (B) shows a clear separation of
the 2 groups (solid line), giving rise to the IFN-INF prediction function showing significant prediction
of hospitalization risk (Figure 1 and Table 2). A grey zone (between the solid and dashed lines), is
observable for IFN-INF function with only No-MVS and high-symptomatic patients (B). Next, PCA
of the eight measured cytokines in only non-hospitalized patients (C), shows separation between
asymptomatic and symptomatic patients. After a dimension reduction, the product of IL-6, IL-17A,
IFN-γ and IFN-α raw levels (D) shows a clear separation of the two groups (solid line), giving rise to
the IL-symp prediction function, allowing accurate and significant prediction of symptoms risk
(Figure 2 and Table 2). A grey zone (between the solid and dashed lines), is observable for IL-symp
function with only low-symptomatic patients (D). Crossed circles depict patients measured already at
the day of the COVID-19 PCR test or within 2 days after. PCA ellipses (D) represent 69% of patient
distribution in each group. IFN-INF prediction function is defined by: log(2*IL-6*IL-10*TNF-α/IFN-α).
IL-symp function is defined by: -1.6*(log(IL-6)+0.175*log(IFN-γ)) + (log(IL-17A)+0.3*log(IFN-α)).

Supplementary Figure 6: Prediction of hospitalization risk by combining biochemistry and
cytokine biomarkers. The combinations of LDH with IL-6 (A) or with the IFN-INF function (B)
show a clear separation of hospitalized versus non-hospitalized patients. Consequently, combing the
conditions that (LDH>260 or IL-6>7.2, out of square in A) or by (LDH>285 or IFN-INF>4.25, out of
square in B) allow for high sensitivity prediction of hospitalization risk. CRP, on the other hand, does
not have additive predictive value for IL-6 (C) or IFN-INF function (D). A number of patients are
missing LDH and/or CRP data (marked No Data below or to the left of the dotted lines accordingly).

Supplementary Figure 7: Comparison of the predictive values for hospitalization risk and for
symptomatic risk for all measured biochemistry, hematology and cytokine biomarkers. Receiver
Operating Characteristic (ROC) curves for the eight measured different single cytokines, as well as the
IFN-INF and IL-IEM cytokine combination functions, as predictors for hospitalization risk (A) and for
symptomatic risk (B). ROC curves for all biochemistry biomarkers, as well as age and BMI, as
predictors for hospitalization risk (C) and for symptomatic risk (D). IFN-INF function is defined by:

\[ \log(2 \times \text{IL-6} \times \text{IL-10} \times \text{TNF-} \alpha / \text{IFN-} \alpha) \]. IL-symp function is defined by:

\[-1.6 \times (\log(\text{IL-6}) + 0.175 \times \log(\text{IFN-} \gamma)) + (\log(\text{IL-17A}) + 0.3 \times \log(\text{IFN-} \alpha)) \].

Supplementary Figure 8: Independence of LDH and CRP as predictors for hospitalization risk as function of different known risk factors for COVID-19 severity. Scatterplots show the independence from other factors of the prediction of hospitalization risk by LDH (A, C, E and G) and CRP (B, D, F and H). Solid horizontal black lines represent the threshold for each biomarker for which the symptoms risk prediction has the largest Youden J-index (see Table 2). Although both LDH and CRP are significantly correlated with BMI (A,B) and age (C,D), prediction by LDH and CRP is independent of BMI (A,B), having diabetes (A,B), age (C,D), or gender (C,D). Importantly, no differences in the prediction power by LDH or CRP are observed as function of time from symptom onset ranging 0-11 DPSO (E,F), or between the study sites (E, F). Lastly, the prediction of hospitalization risk by LDH and CRP held for different SARS-CoV-2 variants (G,H).

Supplementary Figure 9: Independence of IFN-INF and IL-6 as predictors for hospitalization risk as function of different known risk factors for COVID-19 severity. Scatterplots show the independence from other factors of the prediction of hospitalization risk by IFN-INF (A, C, E and G) and IL-6 (B, D, F and H). Solid horizontal black lines represent the threshold for each biomarker for which the symptoms risk prediction has the largest Youden J-index (see Table 2). Although both IFN-INF and IL-6 are significantly correlated with BMI (A,B) and age (C,D), prediction by IFN-INF and IL-6 is independent of BMI (A,B), having diabetes (A,B), age (C,D), or gender (C,D). Importantly, no differences in the prediction power by IFN-INF or IL-6 are observed as function of time from symptom onset ranging 0-11 DPSO (E,F), or between the study sites (E, F). Lastly, the prediction of hospitalization risk by IFN-INF and IL-6 held for different SARS-CoV-2 variants (G,H). IFN-INF prediction function is defined by:

\[ \log(2 \times \text{IL-6} \times \text{IL-10} \times \text{TNF-} \alpha / \text{IFN-} \alpha) \].
Supplementary Figure 10: Independence of IL-symp as predictor for symptoms risk as function of different known risk factors for COVID-19 severity. Scatterplots show the independence from other factors of the prediction of symptoms risk by IL-symp. Solid horizontal black lines represent the IL-symp threshold for which the symptoms risk prediction has the largest Youden J-index (see Table 2). Prediction by IL-symp is independent of BMI (A), having diabetes (A), age (B), or gender (B). Importantly, no difference in the prediction power by IL-symp is observed as function of time from symptom onset ranging 0-11 DPSO (C). Lastly, the prediction of symptoms risk by IL-symp held for different SARS-CoV-2 variants (D). IL-symp function is defined by: 

\[-1.6*(\log(IL-6)+0.175*\log(IFN-\gamma)) + (\log(IL-17A)+0.3*\log(IFN-\alpha))\]
Figure 1

A) p < 0.001
B) p < 0.001
C) p < 0.001
D) p < 0.001
E) Sensitivity vs. 1 - Specificity for Biomarkers for Hospitalization Risk

- Non Hospitalized
  - Asymptomatic
  - Low-Symptomatic
  - High-Symptomatic

- Hospitalized
  - No-MVS
  - MVS
  - ECMO

- Study site
  - Augsburg
  - Paris

0-2 days after PCR
Figure 2

(A) CRP [mg/dl]

(B) IFN-α [pg/ml]

(C) IL-6 [pg/ml]

(D) IL-sympt Function [logs pg/ml]

(E) ROC Curve

Biomarkers for Symptom Status

- Age
- BMI
- CRP
- IL-6
- IFN-α
- IL-sympt
- Random

Non Hospitalized

- Asymptomatic
- Low-Symptomatic
- High-Symptomatic

0-2 days after PCR
All Patients (N=171)
- Hospitalized (N=80)
- Non-Hospitalized (N=91)
  - Symptomatic (N=78)
  - Asymptomatic (N=13)

IL-6 > 7.2 or LDH > 260

Yes
- Predicted Hospitalized
  - Hospitalized (N=80)
  - Non-Hospitalized (N=5)
  → Sensitivity = 100%

No
- Predicted Non-Hospitalized
  - Hospitalized (N=0)
  - Non-Hospitalized (N=86)
  → Specificity = 94.5%

IL-symp Function > 2
(IL-6, IL-17A, IFN-α, IFN-γ)

Yes
- Predicted Symptomatic
  - Symptomatic (N=71)
  - Asymptomatic (N=1)
  → Sensitivity = 97.5%

No
- Predicted Asymptomatic
  - Symptomatic (N=2)
  - Asymptomatic (N=12)
  → Specificity = 92.3%
Figure 4

(A) LDH (logU/L) vs. Days post symptom onset (DPSO)

(B) CRP (logug/dl) vs. Days post symptom onset (DPSO)

(C) IL-6 (logpg/ml) vs. Days post symptom onset (DPSO)

(D) TNF-α (logpg/ml) vs. Days post symptom onset (DPSO)

(E) IL-10 (logpg/ml) vs. Days post symptom onset (DPSO)

(F) IFN-γ (logpg/ml) vs. Days post symptom onset (DPSO)

(G) INF-γ (logpg/ml) vs. Days post symptom onset (DPSO)

(H) IL-17A (logpg/ml) vs. Days post symptom onset (DPSO)

Legend:
- **Negative Control**
- **Hospitalized**
- **Non Hospitalized**
  - **Asymptomatic**
  - **Low-Symptomatic**
  - **High-Symptomatic**
<table>
<thead>
<tr>
<th>Symptoms</th>
<th>Asymptomatic (N=13)</th>
<th>Low Symptomatic (N=41)</th>
<th>High Symptomatic (N=36)</th>
<th>P-value (low vs high symp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cough</td>
<td>0%</td>
<td>70%</td>
<td>79%</td>
<td>0.526</td>
</tr>
<tr>
<td>Sore throat</td>
<td>0%</td>
<td>43%</td>
<td>58%</td>
<td>0.277</td>
</tr>
<tr>
<td>Muscle pain</td>
<td>0%</td>
<td>38%</td>
<td>61%</td>
<td>0.093</td>
</tr>
<tr>
<td>Loss of smell</td>
<td>0%</td>
<td>35%</td>
<td>45%</td>
<td>0.45</td>
</tr>
<tr>
<td>Hoarseness</td>
<td>0%</td>
<td>33%</td>
<td>39%</td>
<td>0.624</td>
</tr>
<tr>
<td>Chest pain</td>
<td>0%</td>
<td>23%</td>
<td>24%</td>
<td>0.908</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>0%</td>
<td>0%</td>
<td>12%</td>
<td>0.38</td>
</tr>
<tr>
<td>Headache</td>
<td>0%</td>
<td>60%</td>
<td>97%</td>
<td>0.006</td>
</tr>
<tr>
<td>Loss of appetite</td>
<td>0%</td>
<td>28%</td>
<td>67%</td>
<td>0.004</td>
</tr>
<tr>
<td>Confusion *</td>
<td>0%</td>
<td>0%</td>
<td>27%</td>
<td>0.046</td>
</tr>
<tr>
<td>Shortness of breath *</td>
<td>0%</td>
<td>0%</td>
<td>33%</td>
<td>0.014</td>
</tr>
<tr>
<td>Diarrhea *</td>
<td>0%</td>
<td>3%</td>
<td>39%</td>
<td>0.007</td>
</tr>
<tr>
<td>Fever *</td>
<td>0%</td>
<td>5%</td>
<td>79%</td>
<td>0.001</td>
</tr>
<tr>
<td>Fatigue *</td>
<td>0%</td>
<td>40%</td>
<td>100%</td>
<td>0.001</td>
</tr>
<tr>
<td>Average number of symptoms (IQR)</td>
<td>0 (0)</td>
<td>3.75 (2-5.25)</td>
<td>7.6 (6-9)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

% patients with symptom per group

0% | 20% | 40% | 60% | 80% | 100%
Supplementary Figure 6

A

B

C

D

E

Non Hospitalized
- Asymptomatic
- Low-Symptomatic
- High-Symptomatic

Hospitalized
- No-MVS
- MVS
- ECMO

Study site
- Augsburg
- Paris
- 0-2 days after PCR
Supplementary Figure 7

A. ROC curve for various cytokines and functions:
- IFN-INF Function
- IL-symp Function
- IL-6
- IL-10
- IL-8
- IL-22
- IL-17A
- IFN-α
- IFN-γ
- TNF-α
- Random

B. ROC curve for various cytokines and functions:
- IFN-INF Function
- IL-symp Function
- Age
- BMI
- LDH
- CRP
- Creatine kinase
- Lymphocytes
- Neutrophiles
- Platelet
- Ferritin
- D-Dimers
- Random
Supplementary Figure 8

A

B

C

D

E

F

G

H

Diabetes
No Diabetes

Female
Male

Paris
Augsburg

Non Hospitalized
Hospitalized

Low-Symptomatic
Asymptomatic
High-Symptomatic

All rights reserved. No reuse allowed without permission.

(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. The copyright holder for this preprint this version posted July 5, 2024. ; https://doi.org/10.1101/2024.07.05.24309641
doi: medRxiv preprint
Supplementary Figure 9

A

Diabetes

No Diabetes

R = 0.55

P < 0.0001

B

IL-6 [log pg/ml]

R = 0.41

P < 0.0001

C

Female

Male

R = 0.37

P < 0.0001

D

Female

Male

R = 0.35

P < 0.0001

E

Paris

Augsburg

R = 0.16

P = 0.04

F

Paris

Augsburg

R = 0.24

P = 0.001

G

Days post symptom onset (DPSO)

R = 0.79

P < 0.0001

H

Days post symptom onset (DPSO)

R = 0.79

P < 0.0001
Supplementary Figure 10

A

![Graph A: Diabetes vs. No Diabetes]

B

![Graph B: Female vs. Male]

C

![Graph C: Days post symptom onset (DPSO)]

D

![Graph D: Various comparisons]

- Asymptomatic
- Low-Symptomatic
- High-Symptomatic

Non Hospitalized