

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

3 Corinna Holetschek^{1,2,3*}, Mehmet Goekkaya^{1,2*}, Karim Dorgham^{4*}, Paul Eder¹, Daria
4 Luschkova¹, Christophe Parizot⁵, Mehmet Tekinsoy⁶, Denise Rauer^{1,2}, Assia Samri^{4,5},
5 Early-Opt-COVID19 study group[§], Matthias Reiger^{1,2}, Gertrud Hammel^{1,2}, Philipp Steininger⁷,
6 Stefanie Gilles¹, Ulrike Protzer⁶, Christoph Römmele⁸, Guy Gorochov^{4,5,9#}, Claudia Traidl-
7 Hoffmann^{1,2,3,10#} and Avidan U. Neumann^{1,2#}

8

9 1 Environmental Medicine, Faculty of Medicine, University of Augsburg, Augsburg,
10 Germany

11 2 Institute of Environmental Medicine, Helmholtz Munich, Augsburg, Germany

12 3 Chair of Environmental Medicine, Technical University of Munich, Augsburg, Germany

13 4 Centre d'Immunologie et des Maladies Infectieuses (CIMI-Paris), National Institute for
14 Health and Medical Research (INSERM), Paris, France

15 5 Département d'Immunologie, Hôpital Pitié-Salpêtrière, Assistance Publique-Hôpitaux de
16 Paris (AP-HP), Paris, France

17 6 Institute of Virology, Technical University of Munich, Munich, Germany

18 7 Institute of Clinical and Molecular Virology, Universitätsklinikum Erlangen, Friedrich-
19 Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

20 8 Internal Medicine III - Gastroenterology and Infectious Diseases, University Hospital of
21 Augsburg, Augsburg, Germany

22 9 Sorbonne Université, Paris, France

23 10 CK-CARE, Christine Kühne Center for Allergy Research and Education, Davos,
24 Switzerland

25 * Authors contributed equally as first authors.

26 # Authors contributed equally as last authors

27

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

31 **\$ Early-Opt-COVID19 study group:**

32 Mustafa, Akaryildiz¹, Linda Bachmann¹, Anna Bayer¹, Barbara Böglmüller¹, Omaira da Mata-
33 Jardin², Amadeo DeTomassi¹, Christin Dhillon³, Victoria Fairweather¹, Marie Freitag³, André
34 Fuchs³, Anja Fusco¹, Carina Gülzow¹, Hans Hauner^{4,5}, Claudia Khalfi¹, Lena Klepper¹,
35 Elisabeth Kling⁶, Silke Kotschenreuther¹, Helmut Messmann³, Aline Metz¹, Anna
36 Muzalyova⁷, Franziska Neitzel¹, Nicole Pochert¹, Claudia Ranieri¹, Monika Seemann¹,
37 Magnus Wangenheim¹, Katharina Zeiser¹

38 1 Environmental Medicine, Faculty of Medicine, University of Augsburg, Augsburg,
39 Germany

40 2 Centre d'Immunologie et des Maladies Infectieuses (CIMI-Paris), National Institute for
41 Health and Medical Research (INSERM), Paris, France

42 3 Internal Medicine III - Gastroenterology and Infectious Diseases, University Hospital of
43 Augsburg, Augsburg, Germany

44 4 Technical University of Munich, Else Kröner-Fresenius-Center for Nutritional Medicine,
45 Chair of Nutritional Medicine, School of Life Sciences

46 5 Institute for Nutritional Medicine, School of Medicine, Technical University of Munich

47 6 Institute for Lab Medicine and Microbiology, Uniklinikum Augsburg, Augsburg, Germany

48 7 Institute for Digital Medicine, Uniklinikum Augsburg, Augsburg, Germany

49

50 **Abstract**

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

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

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

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

75

76 **Source of funding**

77 This study was funded by the Bavarian State Ministry for Science and Art Program for the
78 funding of Corona research (Early-Opt-COVID19project and research networks FOR-COVID
79 and Bay-VOC), by grants from the Helmholtz Association's Initiative and Networking
80 Fund (KA1-Co-02 “COVIPA” to UP; KA1-Co-06“CORAERO” to CH, MG and AUN), by
81 the European Commission FET Open Grant VIROFIGHT (grant no. 899619), and by the
82 Sanddorf foundation. Work in the Gorochov laboratory was supported by Institut National de
83 la Santé et de la Recherche Médicale (INSERM), Sorbonne Université, Fondation pour la
84 Recherche Médicale (FRM), Paris, France, program “Investissement d’Avenir” launched
85 bythe French Government and implemented by the Agence Nationale de la Recherche(ANR)
86 (programme COFIFERON ANR-21-RHUS-08), by EU Horizon HLTH-2021-DISEASE-
87 04UNDINE project, by Fondation pour la Recherche Médicale, Paris, France (programme
88 Equipe FRM 2022) and by the Département Médico-Universitaire de Biologie et Génomique
89 Médicales (DMU BioGen), APHP, Paris, France.

90 **Acknowledgments**

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

95 **Disclosure statement**

96 UP received personal fees from Abbott, Abbvie, Arbutus, Gilead, GSK, Leukocare, J&J,
97 Roche, MSD, Sanofi, Sobi and Vaccitech. UP is a co-founder and share-holder of SCG Cell
98 Therapy. All other authors declare no conflict of interest. The study was independently
99 designed, run, analyzed and summarized by the authors with no involvement from the funding

100 agencies. The manuscript was written by the authors with no involvement of the funding
101 agencies. The funding agencies did not pay for or were involved in any way of writing the
102 manuscript.

103 **Keywords**

104 SARS-CoV-2, COVID-19, Prediction, Biomarkers, Inflammatory cytokines, Type-I

105 Interferon, LDH, CRP, IL-6

106

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

108

109 **Abbreviations**

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

111 **Introduction**

112 The SARS-CoV-2 virus started spreading in December 2019 causing a total of 6 million
113 COVID-19 related deaths by December 31, 2021¹. In the last 2 years, SARS-CoV-2 became
114 endemic, but COVID-19 mortality is still considerably high, with about 200,000 deaths in
115 2023, even with the less severe Omicron variant². Several Anti-viral drugs are available^{3,4,5}
116 and more are under investigation^{6,7}, also for patients with mild-moderate disease. However,
117 these need to be initiated early within 5-7 days post symptoms onset (DPSO) and targeted to
118 patients that are of risk for hospitalization or develop considerable symptomatic disease. Also,
119 immune modulating drugs that inhibit inflammatory cytokines elevated as part of the COVID-
120 19 cytokine storm (e.g., IL-6 inhibitors) are available^{8,9}. However, currently, anti-viral therapy
121 is underused¹⁰ and cytokine inhibitors are only given to hospitalized severe patients, mainly
122 because of lack of biomarkers that allow early stratification of the appropriate target
123 patients¹¹. Therefore, it is important to understand which SARS-CoV-2 infected individuals
124 are at risk for hospitalization and/or to suffer from symptomatic mild to moderate disease, in
125 order to allow early personalized treatment and optimize public health management.

126 COVID-19 shows a versatile range of severity, from asymptomatic, to symptomatic mild-
127 moderate disease and up to critical severity leading to hospitalization, requiring invasive
128 oxygen support, and potentially leading to multi-organ failure and death¹². COVID-19 has a
129 mean incubation period 5.2 days (95% CI: 4.9–5.5) until symptoms onset¹³. Common early
130 symptoms are characterized by fever, cough, fatigue, diarrhea, and dyspnea¹⁴. However, meta-
131 analysis showed that 40-50% of confirmed SARS-CoV-2 positive cases are asymptomatic¹⁵.
132 The Kings College large app study identified the early symptoms that are most associated
133 with subsequent hospitalization and thus allowed stratification of patients into low versus high
134 symptomatic groups¹⁶.

135 Patients' characteristics, such as age and BMI, and underlying comorbidities (e.g., diabetes,
136 asthma, cancer, cardiovascular disorder, chronic kidney disease) are known factors associated
137 with risk for severe COVID-19 disease¹⁷ and are currently used as recommendations for early
138 treatment with anti-viral therapy¹⁸. However, these factors do not allow accurate prediction of
139 hospitalization and symptomatic risk. In addition, a number of biochemistry and
140 hematological blood biomarkers were reported to be risk factors for severe COVID-19 in
141 hospitalized patients. C-reactive protein (CRP) levels were found to be significantly higher in
142 severe hospitalized COVID-19 compared to moderate cases¹⁹, and elevated blood lactate
143 dehydrogenase (LDH) levels were reported to be associated with a worse outcome in
144 hospitalized patients²⁰. Furthermore, low lymphocyte count was also associated with a more
145 severe COVID-19 disease course²¹.

146 Additionally, as severe cases of COVID-19 suffer from a "cytokine storm"^{22,23} which is an
147 abnormal regulation and excessive release of different pro-inflammatory cytokines as a
148 response to the infection, various cytokines were identified as biomarkers predictive of severe
149 COVID-19. IL-6 is known to play a role as a prognostic marker in COVID 19 since most
150 patients show elevated IL-6 levels during a SARS-CoV-2 infection²³⁻²⁵. Several studies have
151 shown high IL-10 levels in severe COVID-19^{26,27}. Similarly, IL-17A²⁸ and tumor necrosis
152 factor alpha (TNF- α)^{25,29} were reported to be associated with disease severity and
153 progression. As to type-I interferon response, it was reported that hospitalized patients with
154 increased IFN- α levels showed improvement in the COVID-19 disease course³⁰. In contrast,
155 high IFN- γ levels in hospitalized cases were shown to be correlated with a worse diagnosis³¹.
156 We had previously reported that the ratio between type-I interferon response and
157 inflammatory cytokines, at the day of hospitalization, shows the highest accuracy for
158 predicting COVID-19 severity and mortality³².

159 However, all the above studies only tested biomarkers predicting severity in patients that were
160 already hospitalized. It would be rather highly beneficial to discover biomarkers for early
161 predictors of severity at the earliest time possible (i.e., as soon as possible after the first
162 positive SARS-CoV-2 diagnosis and within 5-7 DPO), which are feasible to measure at the
163 point-of-testing. Thus, here we are combining biochemistry, hematological and cytokine
164 biomarkers to predict as early as possible the risk of hospitalization, and the symptomatic
165 versus asymptomatic course of infection.

166

167 **Methods**

168 **Patients and study cohorts**

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

181 Visit 1 at the day of PCR test, for patients recruited at the test center, was available for 36
182 patients, while the other patients (n=56) started at visit 2 (1-5 days after PCR test), with
183 following visit 3 (2-3 weeks after) and visit 4 (4-5 weeks after). Only patients with their first

184 sample available at $DPSO \leq 11$ were included in the analysis, since at later time points there
185 was a significant decline in the various biomarker levels. The patients completed a symptoms
186 questionnaire at each visit to classify the disease course.

187 From 1.2.2021 until 14.9.2021, $n=115$ patients were recruited in Augsburg, of which 83 non-
188 hospitalized patients were recruited within the Early-Opt-COVID-19 study, $n=9$ non-
189 hospitalized patients were included within the COVID-19 Vaccine Consortium (CoVaKo)
190 study³³ and 23 hospitalized patients were included at the University Hospital Augsburg. In
191 addition, we used results from our previous study for $n=70$ patients hospitalized at the Pitié-
192 Salpêtrière hospital in Paris³². Additionally, $n=14$ control subjects with negative by SARS-
193 CoV-2 PCR test were enrolled. Only patients with the first visit at most 11 days post
194 symptoms onset (DPSO average 6.2, IQR: 4-9) were included.

195 Demographic and clinical patient characteristics are summarized in Table 1. Hospitalized
196 patients were divided into three groups according to the oxygen support needed. Patients
197 ventilated with nasal cannula, oxygen mask or without oxygen support were grouped as No-
198 MVS ($n=36$). Mechanical ventilatory support (MVS, $n=43$) patients required invasive
199 mechanical ventilation and ECMO patients ($n=14$) patients required extracorporeal membrane
200 oxygenation³². The non-hospitalized patients were divided into 3 symptom groups (Table 1
201 and Supplementary Figure 1): no symptoms (asymptomatic), mild and few flu-like symptoms
202 (low-symptomatic), and at least 2 out of 5 severe (confusion, shortness of breath, fever,
203 fatigue, diarrhea) symptoms (high-symptomatic), according to the Kings College¹⁶. Low-
204 symptomatic patients exhibited on average 3.75 symptoms, significantly ($p<0.001$) lower than
205 7.6 symptoms in high-symptomatic patients. More details in supplementary methods.

206 **Biological sampling**

207 On visit 1, only one serum Gel S-Monovette (SARSTEDT) was drawn from the patients. On
208 visits 2-4, four serum Gel S-, a citrate, and an EDTA Monovettes of blood were drawn. On

209 visits 2-4 an oropharyngeal swab was taken and stored in Guanidine Thiocyanate (PrimeStore
210 MTM) medium for RNA stabilization. More details in supplementary methods.

211

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

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

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

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

224 More details in supplementary methods.

225

226 **Bioinformatics and statistical analysis**

227 Descriptive and statistical analyses were performed using R and Python. Batch correction was
228 applied to mitigate batch effect bias. Principal component analysis (PCA) was used to cluster
229 patients and compute principal component weights for each parameter. Predictive value of the
230 various biomarkers was evaluated by the Youden J-index (Informedness, Specificity +
231 Sensitivity -1), since accuracy is to be taken only indicatively due to the ratios between
232 number of patients in the different groups does not represent real world ratios. The threshold
233 used for prediction was selected as the point with the highest J-index on the ROC curve for
234 each biomarker. Longitudinal data was analyzed using piecewise linear regression³⁴ and mean

235 negative control data was used as a baseline. Statistical significance between groups was
236 evaluated using the nonparametric Mann-Whitney-U test for continuous variables and the
237 Fisher-Exact test for discrete variables. Correlations were evaluated using the nonparametric
238 Spearman test. A two-sided p-value lower than 0.05 was considered statistically significant.
239 Since 19 biomarkers and 5 biomarker combinations were tested, by the Bonferroni rule a p-
240 value of lower than 0.003 was considered significant after multiple testing correction.
241 More details in supplementary methods.

242

243 **Results**

244 **Hospitalization risk prediction**

245 Of all blood biochemistry and hematology biomarkers measured at the first sample per patient
246 (0-11 days post symptoms onset), LDH and CRP showed the most significant ($p < 0.001$)
247 elevation in hospitalized versus non-hospitalized patients (Fig 1A-B), and were the only ones
248 to show accurate prediction of hospitalization risk with a Youden J-index larger than 0.8
249 (LDH: J-index=91.8%, sensitivity=96.2% and specificity=95.6%; CRP: J-index=87.8%,
250 sensitivity=94.4% and specificity=93.4%; Fig 1E, Table 2). D-dimers, ferritin, creatinine
251 kinase and neutrophils percentage were also significantly elevated, while lymphocytes
252 percentage was significantly lower, in hospitalized patients, but none of those showed a good
253 prediction (Supp Figures 2,4,7).

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

260 hospitalized patients, especially in ECMO patients, versus symptomatic patients (Supp
261 Figures 3-4).

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

$$270 \quad \text{IFN-INF} = \log (2 * \text{IL-6} * \text{IL-10} * \text{TNF-}\alpha / \text{IFN-}\alpha)$$

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

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

282 While LDH, CRP, IL-6, and the IFN-INF function, are all correlated with age and BMI, their
283 predictive value is independent of these factors or of gender and diabetes status (Supp Fig 8-

284 9). Importantly, these 4 predictors are not correlated with, and their predictive value is
285 independent of, the time the sample was taken at the range of 0-11 DPO or the study site
286 (Supp Figures 8-9). Also, the separation between hospitalized and non-hospitalized patients
287 for these 4 biomarkers is observed over all SARS-CoV-2 variants in the study (Supp Figures
288 8-9).

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

297

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

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

305 Among the cytokines, IFN- α , IFN- γ , IL-6, IL-10, IL17-A and IL-22 were significantly
306 elevated in symptomatic versus asymptomatic patients (Fig 2B-C, Supp Fig 3-4). However,
307 only a modest (J-index= 64-72%) predictive value for symptomatic risk was found only for

308 IFN- α and IL-6, but for none of the other cytokine nor for the IFN-INF cytokine ratio (Fig 2E,
309 Table 2, Supp Fig 7C).

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

$$319 \quad \text{IL-symp} = -1.6*(0.175*\log(\text{IFN-}\gamma)+\text{IL-6})+(0.3*\log(\text{IFN-}\alpha)+\log(\text{IL-17A}))$$

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

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

332

333 **Decision tree for combined prediction of hospitalization risk and symptomatic status**

334 Thus, by measuring only LDH and 4 cytokines (IL-6, IL-17A, IFN- α and IFN- γ) it is possible
335 to obtain a combined decision tree (Fig 3), predicting in our study hospitalized patients with
336 sensitivity of 100% (using the IL-6 and LDH combination) and among those predicted not to
337 be hospitalized a prediction of symptomatic patients (using IL-symp) with sensitivity of
338 97.5%.

339

340 **Longitudinal profiles of cytokine and biochemistry biomarkers**

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

352

353 **Discussion**

354 Our findings indicate that it is possible to predict COVID-19 disease progression with
355 biomarkers feasibly measurable at point-of-testing already as early as 0-7 days post symptoms

356 onset and even at the time of the first positive diagnosis of SARS-CoV-2 infection. Most
357 importantly, we have shown that highly accurate (J-index >90%) early prediction of COVID-
358 19 hospitalization risk is possible either by high LDH levels, or high IL-6 levels, or by the
359 ratio between IFN- α and the inflammatory cytokines IL-6, IL-10 and TNF- α , termed here the
360 IFN-INF predictor function. Moreover, combining the LDH with IL-6 biomarkers enables a
361 predictor with sensitivity of 100%. It should be noted that all above predictors are not
362 statistically significantly different in their predictive values. While all above biomarkers have
363 been previously shown to predict COVID-19 severity in hospitalized patients^{19,20,23-31}, our
364 results now show that they are also good early predictors for hospitalization risk.

365 Furthermore, we have shown that a combination of the IFN- α , IFN- γ , IL-6 and IL-
366 17 cytokines, termed here the IL-symp prediction function, is an early and accurate (J-
367 index=89.8%) predictor of symptomatic versus asymptomatic course of infection.
368 Interestingly, neither any single cytokine alone, nor the IFN-INF cytokine combination, nor
369 any of the other biochemistry or hematological parameters, were a good predictors for
370 symptoms risk. While the prediction of symptomatic status is clinically less important, still it
371 could be used to guide personalized treatment with SARS-CoV-2 anti-viral therapy, as these
372 are proven in clinical trials to reduce the associated quality-of-life impairment and public-
373 health burden in mild-moderate patients³⁻⁷.

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

379 We also found that the risk of hospitalization and the risk of symptomatic status are associated
380 with older age and higher BMI, similar to previous studies showing these factors associated

381 with COVID-19 severity^{17,18}. However, we show that the predictors we have identified are
382 independent of such factors as age, gender, BMI or diabetes, probably since these are already
383 factorized into the early levels of the biomarkers in response to SARS-CoV-2 infection.

384 Our results also shed further light on the role of the different cytokines in response to SARS-
385 Cov-2 infection. While high levels of the inflammatory cytokines (e.g. IL-6, IL-8 and TNF-
386 α), and also interestingly IL-10, are associated with both the risk for hospitalization and
387 higher risk for symptomatic course of infection, conversely IFN- α , IFN- γ and IL-17 are less
388 associated with hospitalization risk but rather with the appearance of symptoms. In fact, an
389 early high ratio of IFN- α compared to inflammatory cytokines seems to decrease the risk of
390 hospitalization, similarly to our previous findings about COVID-19 severity in hospitalized
391 patients³². It is important to note that since we aimed at finding the minimum number of
392 cytokines that allow a good prediction, some cytokines (e.g. IL-8) were not included in our
393 prediction function since they do not add significant predictive value, probably because of
394 high inter-correlations, but still are associated with risk of hospitalization or of symptomatic
395 disease. Also, we found that the IFN-INF cytokine combination function is positively
396 correlated with LDH and CRP, indicating a biological link between the different processes
397 that these are biomarkers for. Interestingly, there is no strong correlation between the IL-symp
398 function, predictive of symptoms' risk, with biochemistry or hematological biomarkers or
399 lymphocyte counts.

400 Our longitudinal results show that the kinetics of the cytokines IL-6, TNF- α and IL-10 differ
401 in hospitalized patients, showing higher expression at an early timeframe (already around 0-7
402 days post symptoms onset) and staying high even 4 weeks after. While it was shown that
403 inflammatory cytokines are high in severe COVID-19 patients²³⁻²⁷, we show here that in some
404 patients they are already high very early on, thus indicating that if immune modulatory
405 treatment is to be successful it should be initiated early personalized according to the cytokine

406 profile¹¹. As expected, we see also an early rise in IFN- α and IFN- γ , but these are not as
407 significantly different between hospitalized and symptomatic patients and they decline within
408 2-3 weeks also in hospitalized patients.

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

422 Based on our results, we suggest a combined predictive pipeline for early prediction of
423 patients that will become asymptomatic versus symptomatic versus hospitalized. As early as
424 possible after a positive SARS-CoV-2 test, patients should get blood drawn and LDH, IL-6,
425 IL-17A, IFN- α and IFN- γ should be measured, which could be performed at point of testing
426 by a multiplex ELISA or other methods. Using these biomarkers the hospitalization risk could
427 be evaluated by LDH combined with IL-6, and the symptomatic risk could be evaluated by
428 the IL-symp function, therefore allowing for personalized treatment with either anti-viral
429 therapy or with cytokine-inhibitor treatment, which can be also guided by the cytokine levels.

430 With these two populations predicted, personalized therapy approaches can be targeted to the
431 patients with high risk in regards to hospitalization and a symptomatic course.

432

433 **References**

- 434 1. Larkin HD. Global COVID-19 Death Toll May Be Triple the Reported Deaths. *JAMA* 2022;
435 **327**: 1438.
- 436 2. Medicine JeCbJHU. COVID-19 Dashboard by the Center for Systems Science and
437 Engineering (CSSE) at Johns Hopkins University (JHU). 2023. <https://coronavirus.jhu.edu/map.html>
438 (accessed 20 November 2023).
- 439 3. Hammond J, Leister-Tebbe H, Gardner A, et al. Oral nirmatrelvir for high-risk,
440 nonhospitalized adults with Covid-19. *New England Journal of Medicine* 2022; **386**(15): 1397-408.
- 441 4. Wang Y, Zhang D, Du G, et al. Remdesivir in adults with severe COVID-19: a randomised,
442 double-blind, placebo-controlled, multicentre trial. *The lancet* 2020; **395**(10236): 1569-78.
- 443 5. Wahl A, Gralinski LE, Johnson CE, et al. SARS-CoV-2 infection is effectively treated and
444 prevented by EIDD-2801. *Nature* 2021; **591**(7850): 451-7.
- 445 6. Gupta A, Gonzalez-Rojas Y, Juarez E, et al. Effect of sotrovimab on hospitalization or death
446 among high-risk patients with mild to moderate COVID-19: a randomized clinical trial. *Jama* 2022;
447 **327**(13): 1236-46.
- 448 7. Yotsuyanagi H, Ohmagari N, Doi Y, et al. Efficacy and Safety of 5-Day Oral Ensitrelvir for
449 Patients With Mild to Moderate COVID-19: The SCORPIO-SR Randomized Clinical Trial. *JAMA*
450 *Network Open* 2024; **7**(2): e2354991-e.
- 451 8. Zizzo G, Tamburello A, Castelnovo L, et al. Immunotherapy of COVID-19: Inside and
452 beyond IL-6 signalling. *Frontiers in Immunology* 2022; **13**: 795315.
- 453 9. Galván-Román JM, Rodríguez-García SC, Roy-Vallejo E, et al. IL-6 serum levels predict
454 severity and response to tocilizumab in COVID-19: An observational study. *Journal of Allergy and*
455 *Clinical Immunology* 2021; **147**(1): 72-80. e8.
- 456 10. Rubin R. Paxlovid Is Effective but Underused—Here’s What the Latest Research Says About
457 Rebound and More. *JAMA* 2024; **331**(7): 548-51.
- 458 11. Neumann AU, Goekkaya M, Dorgham K, Traidl-Hoffmann C, Gorochov G. Tocilizumab in
459 COVID-19 therapy: who benefits, and how? *The Lancet* 2021; **398**(10297): 299-300.
- 460 12. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease
461 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center
462 for Disease Control and Prevention. *jama* 2020; **323**(13): 1239-42.
- 463 13. Alene M, Yismaw L, Assemie MA, Ketema DB, Gietaneh W, Birhan TY. Serial interval and
464 incubation period of COVID-19: a systematic review and meta-analysis. *BMC Infectious Diseases*
465 2021; **21**(1).
- 466 14. Alimohamadi Y, Sepandi M, Taghdir M, Hosamirudsari H. Determine the most common
467 clinical symptoms in COVID-19 patients: a systematic review and meta-analysis. *Journal of*
468 *preventive medicine and hygiene* 2020; **61**(3): E304.
- 469 15. Ma Q, Liu J, Liu Q, et al. Global percentage of asymptomatic SARS-CoV-2 infections among
470 the tested population and individuals with confirmed COVID-19 diagnosis: a systematic review and
471 meta-analysis. *JAMA network open* 2021; **4**(12): e2137257-e.
- 472 16. Sudre CH, Lee KA, Lochlainn MN, et al. Symptom clusters in Covid19: A potential clinical
473 prediction tool from the COVID Symptom study app. Cold Spring Harbor Laboratory; 2020.
- 474 17. Control CfD, Prevention. Underlying medical conditions associated with higher risk for severe
475 COVID-19: information for healthcare professionals. 2022.
- 476 18. Control CfD, Prevention. Interim clinical considerations for COVID-19 treatment in
477 outpatients. 2023.
- 478 19. Yitbarek GY, Walle Ayehu G, Asnakew S, et al. The role of C-reactive protein in predicting
479 the severity of COVID-19 disease: A systematic review. *SAGE Open Medicine* 2021; **9**.

- 480 20. Henry BM, Aggarwal G, Wong J, et al. Lactate dehydrogenase levels predict coronavirus
481 disease 2019 (COVID-19) severity and mortality: A pooled analysis. *Am J Emerg Med* 2020; **38**(9):
482 1722-6.
- 483 21. Huang I, Pranata R. Lymphopenia in severe coronavirus disease-2019 (COVID-19):
484 systematic review and meta-analysis. *Journal of intensive care* 2020; **8**: 1-10.
- 485 22. Hu B, Huang S, Yin L. The cytokine storm and COVID-19. *Journal of medical virology*
486 2021; **93**(1): 250-6.
- 487 23. Moore JB, June CH. Cytokine release syndrome in severe COVID-19. *Science* 2020;
488 **368**(6490): 473-4.
- 489 24. Potere N, Batticciotto A, Vecchié A, et al. The role of IL-6 and IL-6 blockade in COVID-19.
490 *Expert Review of Clinical Immunology* 2021; **17**(6): 601-18.
- 491 25. Del Valle DM, Kim-Schulze S, Huang HH, et al. An inflammatory cytokine signature predicts
492 COVID-19 severity and survival. *Nat Med* 2020; **26**(10): 1636-43.
- 493 26. Jafrin S, Aziz MA, Islam MS. Elevated Levels of Pleiotropic Interleukin-6 (IL-6) and
494 Interleukin-10 (IL-10) are Critically Involved With the Severity and Mortality of COVID-19: An
495 Updated Longitudinal Meta-Analysis and Systematic Review on 147 Studies. *Biomarker Insights*
496 2022.
- 497 27. Chang Y, Bai M, You Q. Associations between Serum Interleukins (IL-1 β , IL-2, IL-4, IL-6,
498 IL-8, and IL-10) and Disease Severity of COVID-19: A Systematic Review and Meta-Analysis.
499 *BioMed Research International* 2022; **2022**: 2755246.
- 500 28. Francesco M, Gian Marco C, Federica R, et al. Interleukin-17A (IL-17A): A silent amplifier of
501 COVID-19. *Biomedicine & Pharmacotherapy* 2021; **142**: 111980.
- 502 29. Halim C, Mirza AF, Sari MI. The association between TNF- α , IL-6, and vitamin D levels and
503 COVID-19 severity and mortality: a systematic review and meta-analysis. *Pathogens* 2022; **11**(2):
504 195.
- 505 30. Contoli M, Papi A, Tomassetti L, et al. Blood Interferon- α Levels and Severity, Outcomes,
506 and Inflammatory Profiles in Hospitalized COVID-19 Patients. *Frontiers in Immunology* 2021; **12**.
- 507 31. Gadotti AC, de Castro Deus M, Telles JP, et al. IFN- γ is an independent risk factor associated
508 with mortality in patients with moderate and severe COVID-19 infection. *Virus research* 2020; **289**:
509 198171.
- 510 32. Dorgham K, Quentric P, Gökkaya M, et al. Distinct cytokine profiles associated with COVID-
511 19 severity and mortality. *Journal of Allergy and Clinical Immunology* 2021; **147**(6): 2098-107.
- 512 33. Prelog M, Jeske SD, Asam C, et al. Clinical and immunological benefits of full primary
513 COVID-19 vaccination in individuals with SARS-CoV-2 breakthrough infections: A prospective
514 cohort study in non-hospitalized adults. *Journal of Clinical Virology* 2024; **170**: 105622.
- 515 34. Jekel C, Venter G. A python library for fitting 1D continuous piecewise linear functions. 2019.

516

517

518

519

520

521

522

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

	Non Hospitalized ⁽¹⁾			Hospitalized ⁽¹⁾			All (n=185)	P-Value Asymp vs Symp	P-Value Hosp vs Non-Hosp
	Asymptomatic (n=13)	Low Symptomatic (n=44)	High Symptomatic (n=35)	No MVS (n=36)	MVS (n=43)	ECMO (n=14)			
Augsburg (n)⁽²⁾	13	44	35	14	9	0	115		
Paris (n)⁽³⁾	0	0	0	22	34	14	70		
Age(y), median (IQR)	36.46 (32.0-40.0)	40.39 (27.0-52.0)	41.74 (33.0-48.0)	65.42 (53.75-78.75)	63.21 (57.0-71.0)	46.07 (38.5-54.0)	50.97 (35.0-64.0)	0.51	<0.001
Sex: male, n (%)	6 (46.15)	20 (45.45)	18 (51.43)	27 (75.0)	28 (65.12)	11 (78.57)	110 (59.46)	0.96	0.61
Days since symptom start, Median in days (IQR)	7.54 (6.0-10.0)	5.02 (2.75-7.0)	5.29 (4.0-7.0)	6.47 (4.0-10.0)	7.26 (6.0-9.0)	7.57 (6.0-10.0)	6.24 (4.0-9.0)	0.02	0.61
Past medical history, n (%)⁽⁴⁾									
Obesity (≥ 30)	10 (76.92)	35 (79.55)	25 (71.43)	24 (66.67)	25 (58.14)	6 (42.86)	125 (67.57)	0.88	0.29
Diabetes Typ 2	0 (0)	1 (2.27)	2 (5.71)	12 (33.33)	20 (46.51)	6 (42.86)	41 (22.16)	0.47	<0.001

524

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

526 (2) Patients included in Augsburg from the Early-Opt (n=83), CoVaKo-Augsburg (n=9)³³ and University Hospital Augsburg (n=23) cohorts.

527 (3) Patients included in Paris from our Dorgham et al ³² study.

528 (4) Past medical history given here concisely only for the 2 most prevalent conditions, more details in Supplementary Methods section.

529

530

531 **Table 2: Biomarkers and biomarker combinations allowing early accurate prediction of**
 532 **COVID-19 hospitalization risk (A) or of symptoms risk in non-hospitalized patients (B).**

533 **A**

Hospitalization risk predictor ¹	J-Index ²	Sensitivity	Specificity	Accuracy ³	p-Value ⁴	N
LDH >250	91.8%	96.2%	95.6%	95.9%	<0.0001	169
IFN-INF >4.25	91.3%	98.9%	92.4%	95.7%	<0.0001	185
IL-6 >7.46	90.3%	93.6%	96.7%	95.1%	<0.0001	185
CRP >1.13	87.8%	94.4%	93.4%	93.8%	<0.0001	162
IL-symp >0.74	84.9%	92.5%	92.4%	92.4%	<0.0001	185
IL-6 >7.2 or LDH >260	94.5%	100%	94.5%	97.0%	<0.0001	169
CRP >2.5 or LDH >250	94.5%	100%	94.5%	96.8%	<0.0001	154
IFN-INF >4.25 or LDH >285	92.3%	100%	92.3%	95.8%	<0.0001	169
IFN-INF or CRP	Combination does not add predictive value					162
IL-6 or CRP	Combination does not add predictive value					162

534

535 **B**

Symptoms risk predictor ¹	J-Index ²	Sensitivity	Specificity	Accuracy ³	p-Value ⁴	N
IL-symp >-2	89.8%	97.5%	92.3%	96.7%	<0.0001	92
IFN- γ > 0.06	71.9%	94.9%	76.9%	92.4%	<0.0001	91
IL-6 >1.1	64.4%	79.8%	84.6%	80.4%	<0.0001	92

536

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

537

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

538

(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

539

reason, Odds Ratio (OR), Positive Predictive Value (PPV) and Negative Predictive Value (NPV) were not

540

used.

541

used.

542

(4) Statistical significance of the prediction was assessed by Fisher-Exact Test on the confluence table.

543 LDH [log U/l]; CRP [log mg/dl]; All cytokines [log pg/ml].

544 IFN-INF function is defined by: $2(\log(\text{IL-6})+\log(\text{IL-10})+\log(\text{TNF-}\alpha) - \log(\text{IFN-}\alpha))$

545 IL-symp function is defined by: $-1.6(\log(\text{IL-6})+0.175*\log(\text{IFN-}\gamma)) + (\log(\text{IL-17A})+0.3*\log(\text{IFN-}\alpha))$

546

547 **Figure Captions**

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

549 Early levels of LDH (A), CRP (B) and IL-6 (C), depicted as function of symptoms and severity
550 groups, show significant differentiation between hospitalized and non-hospitalized patients. Also, the
551 ratio between inflammatory cytokine (IL-6, IL-10 and TNF- α) levels compared to IFN- α (IFN-INF
552 function, see Supp Figure 5) is significantly higher in hospitalized patients (D). Receiver Operating
553 Characteristic (ROC) curves for IFN-INF, LDH, IL-6 and CRP show significantly larger AUC for the
554 prediction of hospitalization risk as compared to age and BMI (E). Solid horizontal black lines (A, B,
555 C, D) represent the threshold for each biomarker for which the hospitalization risk prediction has the
556 largest Youden J-index (see Table 2). All other biochemistry and cytokine biomarkers show lower
557 prediction of hospitalization risk (Supp Figures 2,3,4,7). Crossed circles depict patients measured
558 already at the day of the COVID-19 PCR test or within 2 days after. IFN-INF function is defined by:
559 $\log(2 * \text{IL-6} * \text{IL-10} * \text{TNF-}\alpha / \text{IFN-}\alpha)$.

560 **Figure 2: Biochemistry and cytokine biomarkers predictive of COVID-19 symptoms risk.**

561 Early levels of CRP (A), IFN- α (B) and IL-6 (C), depicted as function of symptoms groups, show significant
562 differentiation between symptomatic and asymptomatic patients. Also, the product of IL-6, IL-17A,
563 IFN- γ and IFN- α cytokine levels (IL-symp function, see Supp Figure 5) is significantly higher in
564 symptomatic patients (D). Receiver Operating Characteristic (ROC) curves of the prediction for
565 symptoms risk show largest AUC for IL-symp as compared to IFN- α , IL-6, CRP, age and BMI (E).
566 Solid horizontal black lines (A, B, C, D) represent the threshold for each biomarker for which the
567 symptoms risk prediction has the largest Youden J-index (see Table 2). All other biochemistry and
568 cytokine biomarkers show no prediction of symptoms risk (Supp Figures 2,3,4,7). Crossed circles

569 depict patients measured already at the day of the COVID-19 PCR test or within 2 days after. IL-symp
570 function is defined by: $-1.6 * (\log(\text{IL-6}) + 0.175 * \log(\text{IFN-}\gamma)) + (\log(\text{IL-17A}) + 0.3 * \log(\text{IFN-}\alpha))$.

571

572 **Figure 3: Schematic flowchart of the decision tree for the combined prediction for**
573 **hospitalization risk and symptoms risk.** A flowchart showing the stratification of patients in the
574 study to those predicted (using IL-6 or LDH combination) for hospitalization (with 100% sensitivity)
575 versus non-hospitalization, and the latter further stratified to those predicted (with the IL-symp
576 function) to have symptomatic (sensitivity=97.5%) versus asymptomatic course of infection. Only 169
577 patients are shown since 26 patients were missing LDH values. IL-symp function is defined by: -
578 $1.6 * (\log(\text{IL-6}) + 0.175 * \log(\text{IFN-}\gamma)) + (\log(\text{IL-17A}) + 0.3 * \log(\text{IFN-}\alpha))$.

579

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

588

589 **Supplementary figure captions**

590 **Supplementary Figure 1: Symptom distribution in different non-hospitalized symptom groups.**
591 The distribution of 14 different COVID-19 related symptoms across asymptomatic, low symptomatic
592 and high symptomatic patients in the non-hospitalized cohort. High symptomatic was defined as
593 having at least 2 out of 5 of the symptoms most linked to risk of hospitalization (marked by *).

594 Significant differences are observed between low and high-symptomatic patients in the frequency of
595 several symptoms as well as for the number of symptoms a patient exhibits.

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

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

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

617 **Supplementary Figure 5: Derivation of the multi-dimensional cytokine combination prediction**
618 **functions for hospitalizaion risk and symptoms status.** PCA of the eight measured cytokines (A)
619 shows differentiation between hospitalized and non-hospitalized patients. After a dimension reduction
620 to determine the minimum number of most predictive cytokines, the ratio between the raw levels of

621 inflammatory cytokine (IL-6, IL-10 and TNF- α) compared to IFN- α (B) shows a clear separation of
622 the 2 groups (solid line), giving rise to the IFN-INF prediction function showing significant prediction
623 of hospitalization risk (Figure 1 and Table 2). A grey zone (between the solid and dashed lines), is
624 observable for IFN-INF function with only No-MVS and high-symptomatic patients (B). Next, PCA
625 of the eight measured cytokines in only non-hospitalized patients (C), shows separation between
626 asymptomatic and symptomatic patients. After a dimension reduction, the product of IL-6, IL-17A,
627 IFN- γ and IFN- α raw levels (D) shows a clear separation of the two groups (solid line), giving rise to
628 the IL-symp prediction function, allowing accurate and significant prediction of symptoms risk
629 (Figure 2 and Table 2). A grey zone (between the solid and dashed lines), is observable for IL-symp
630 function with only low-symptomatic patients (D). Crossed circles depict patients measured already at
631 the day of the COVID-19 PCR test or within 2 days after. PCA ellipses (D) represent 69% of patient
632 distribution in each group. IFN-INF prediction function is defined by: $\log(2*IL-6*IL-10*TNF-\alpha/IFN-\alpha)$.
633 IL-symp function is defined by: $-1.6*(\log(IL-6)+0.175*\log(IFN-\gamma)) + (\log(IL-17A)+0.3*\log(IFN-\alpha))$.

634

635 **Supplementary Figure 6: Prediction of hospitalization risk by combining biochemistry and**
636 **cytokine biomarkers.** The combinations of LDH with IL-6 (A) or with the IFN-INF function (B)
637 show a clear separation of hospitalized versus non-hospitalized patients. Consequently, combining the
638 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
639 square in B) allow for high sensitivity prediction of hospitalization risk. CRP, on the other hand, does
640 not have additive predictive value for IL-6 (C) or IFN-INF function (D). A number of patients are
641 missing LDH and/or CRP data (marked No Data below or to the left of the dotted lines accordingly).

642

643 **Supplementary Figure 7: Comparison of the predictive values for hospitalization risk and for**
644 **symptomatic risk for all measured biochemistry, hematology and cytokine biomarkers.** Receiver
645 Operating Characteristic (ROC) curves for the eight measured different single cytokines, as well as the
646 IFN-INF and IL-IEM cytokine combination functions, as predictors for hospitalization risk (A) and for
647 symptomatic risk (B). ROC curves for all biochemistry biomarkers, as well as age and BMI, as

648 predictors for hospitalization risk (C) and for symptomatic risk (D). IFN-INF function is defined by:
649 $\log(2*IL-6*IL-10*TNF-\alpha/IFN-\alpha)$. IL-symp function is defined by: $-1.6*(\log(IL-6)+0.175*\log(IFN-\gamma)) +$
650 $(\log(IL-17A)+0.3*\log(IFN-\alpha))$.

651

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

662

663 **Supplementary Figure 9: Independence of IFN-INF and IL-6 as predictors for hospitalization**
664 **risk as function of different known risk factors for COVID-19 severity.** Scatterplots show the
665 independence from other factors of the prediction of hospitalization risk by IFN-INF (A, C, E and G)
666 and IL-6 (B, D, F and H). Solid horizontal black lines represent the threshold for each biomarker for
667 which the symptoms risk prediction has the largest Youden J-index (see Table 2). Although both IFN-
668 INF and IL-6 are significantly correlated with BMI (A,B) and age (C,D), prediction by IFN-INF and
669 IL-6 is independent of BMI (A,B), having diabetes (A,B), age (C,D), or gender (C,D). Importantly, no
670 differences in the prediction power by IFN-INF or IL-6 are observed as function of time from
671 symptom onset ranging 0-11 DPO (E,F), or between the study sites (E, F). Lastly, the prediction of
672 hospitalization risk by IFN-INF and IL-6 held for different SARS-CoV-2 variants (G,H). IFN-INF
673 prediction function is defined by: $\log(2*IL-6*IL-10*TNF-\alpha/IFN-\alpha)$

674

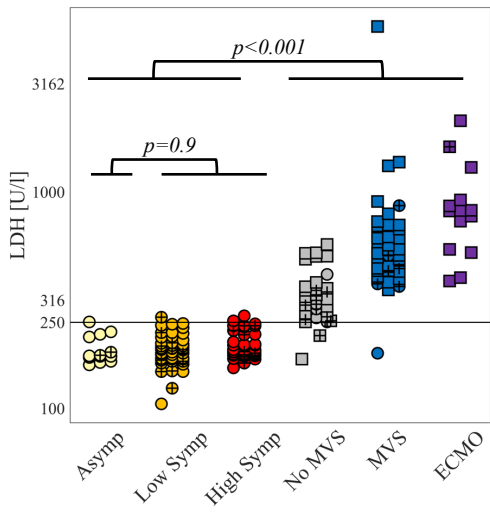
675 **Supplementary Figure 10: Independence of IL-symp as predictor for symptoms risk as function**
676 **of different known risk factors for COVID-19 severity.** Scatterplots show the independence from
677 other factors of the prediction of symptoms risk by IL-symp. Solid horizontal black lines represent the
678 IL-symp threshold for which the symptoms risk prediction has the largest Youden J-index (see Table
679 2). Prediction by IL-symp is independent of BMI (A), having diabetes (A), age (B), or gender (B).
680 Importantly, no difference in the prediction power by IL-symp is observed as function of time from
681 symptom onset ranging 0-11 DPSO (C). Lastly, the prediction of symptoms risk by IL-symp held for
682 different SARS-CoV-2 variants (D). IL-symp function is defined by: $-1.6*(\log(\text{IL}-6)+0.175*\log(\text{IFN}-\gamma))$
683 $+ (\log(\text{IL}-17\text{A})+0.3*\log(\text{IFN}-\alpha))$.

684

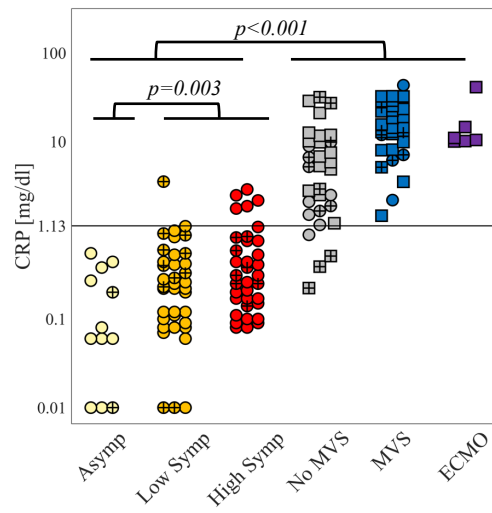
685

686

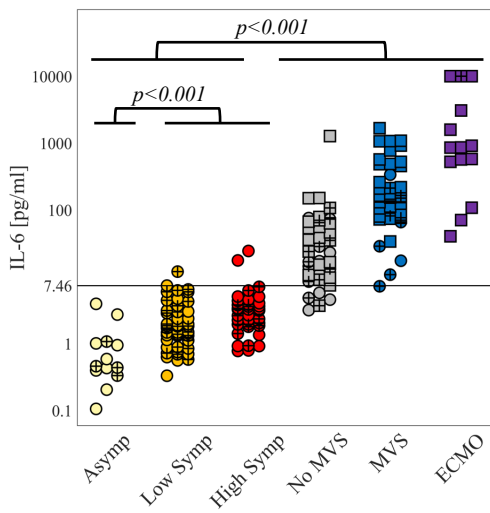
A



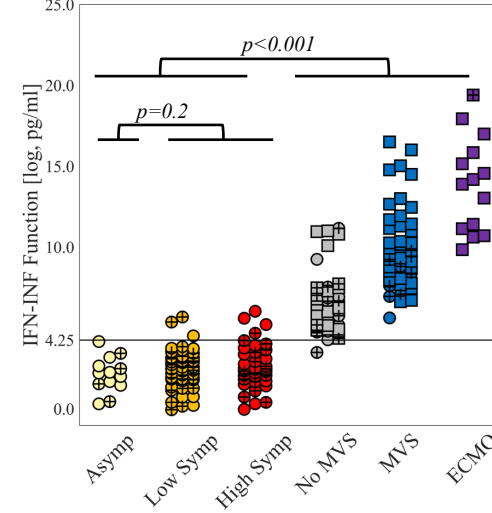
B



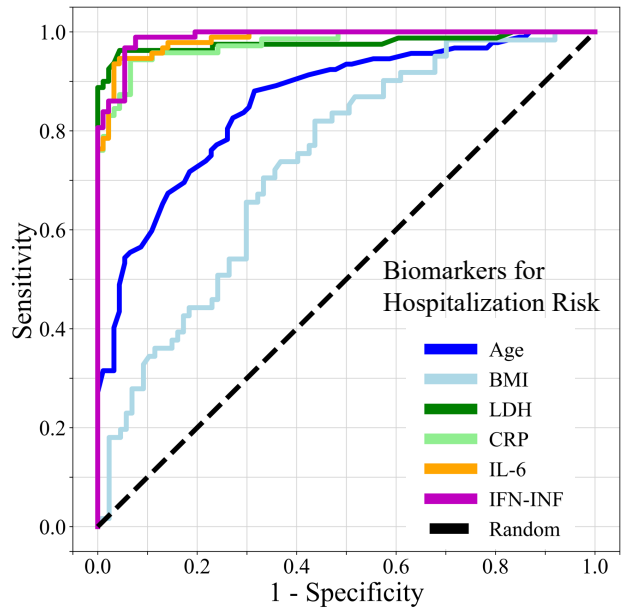
C



D



E



Non Hospitalized

- Asymptomatic
- Low-Symptomatic
- High-Symptomatic

Hospitalized

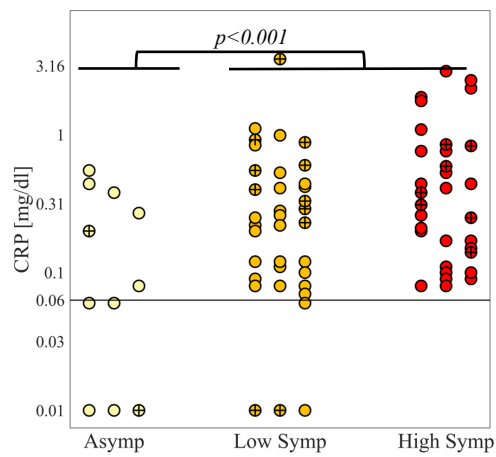
- No-MVS
- MVS
- ECMO

Study site

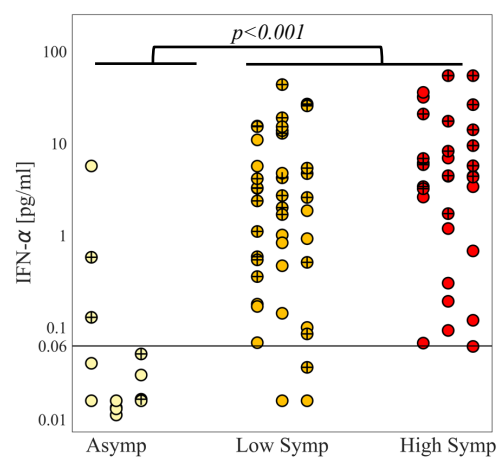
- Augsburg
- Paris
- ⊕ 0-2 days after PCR

Figure 2

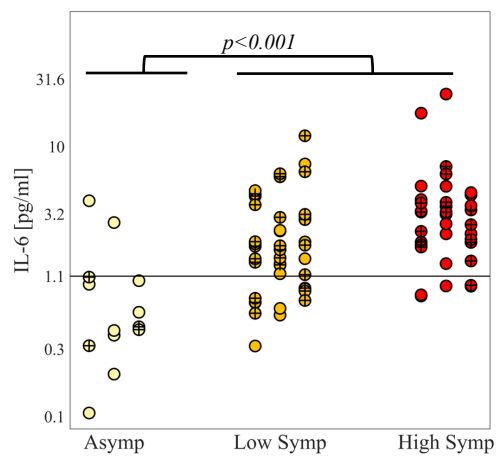
A



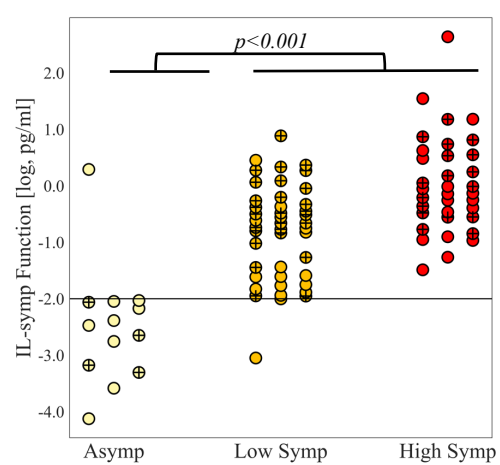
B



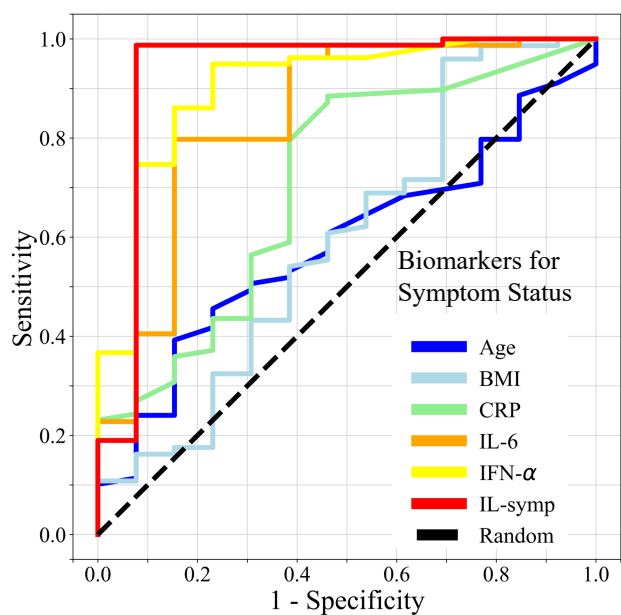
C



D



E



Non Hospitalized

● Asymptomatic

● Low-Symptomatic

● High-Symptomatic

⊕ 0-2 days after PCR

Figure 3

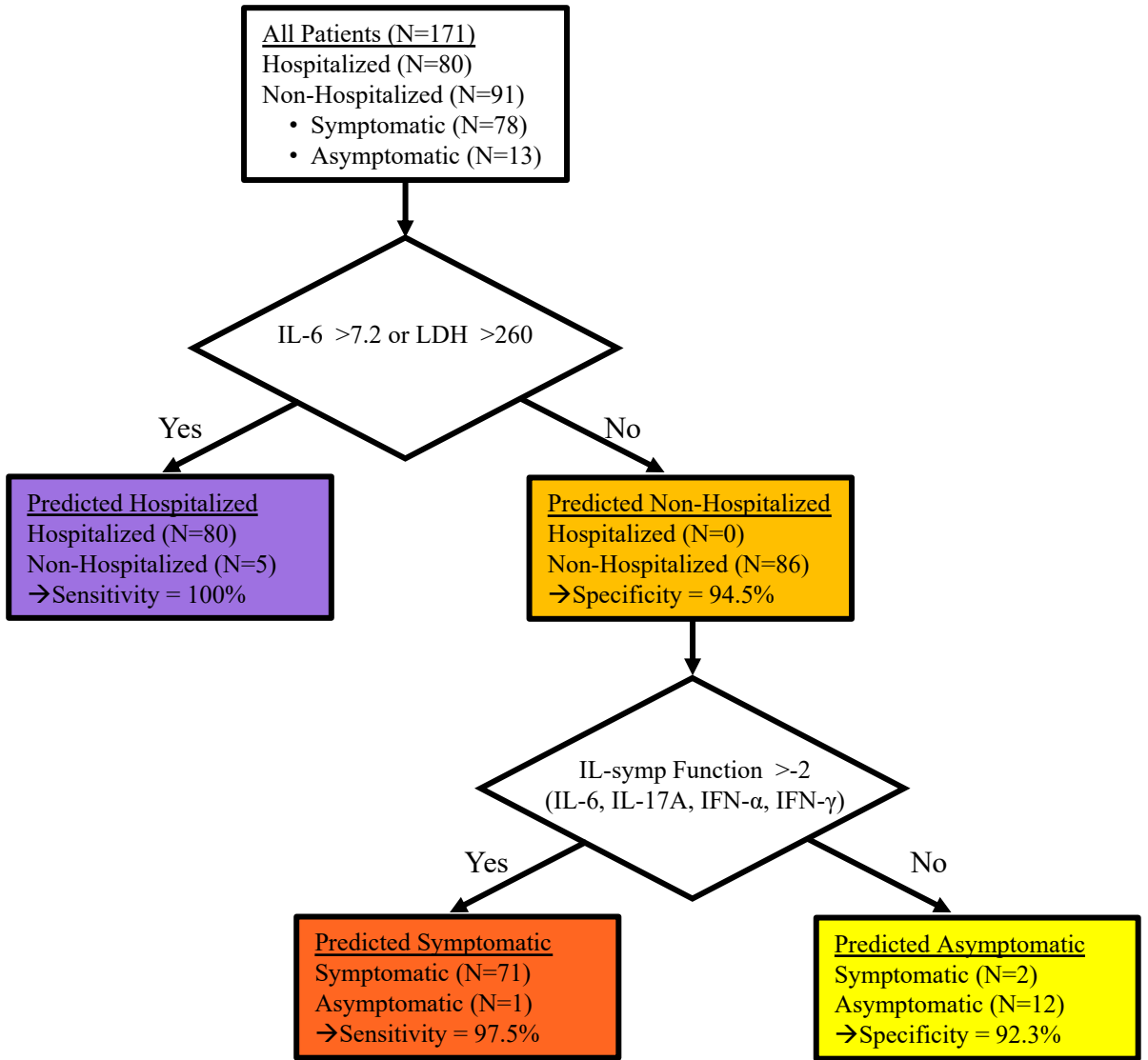
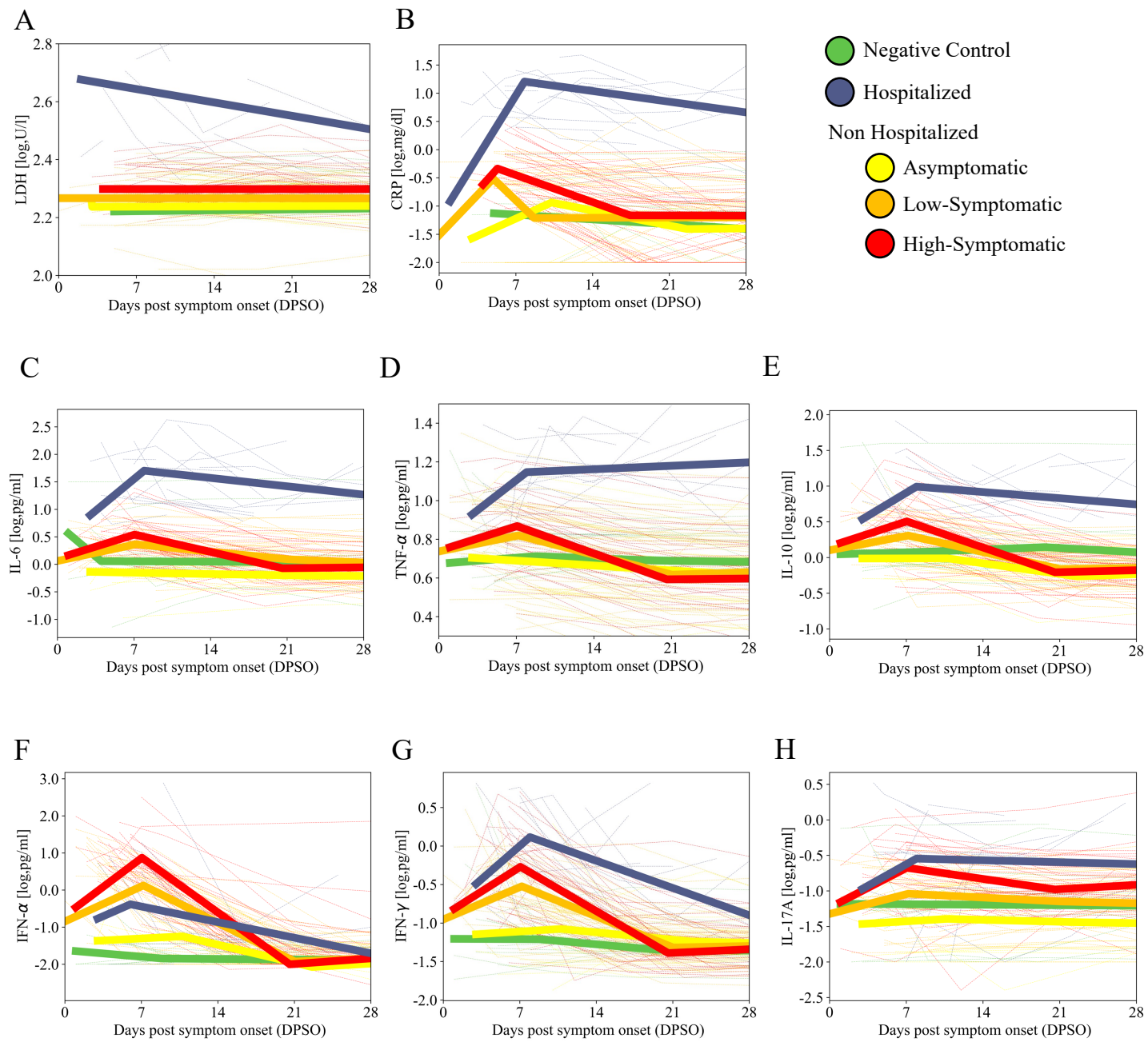


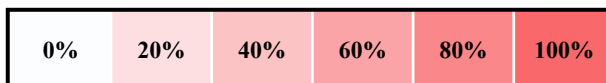
Figure 4



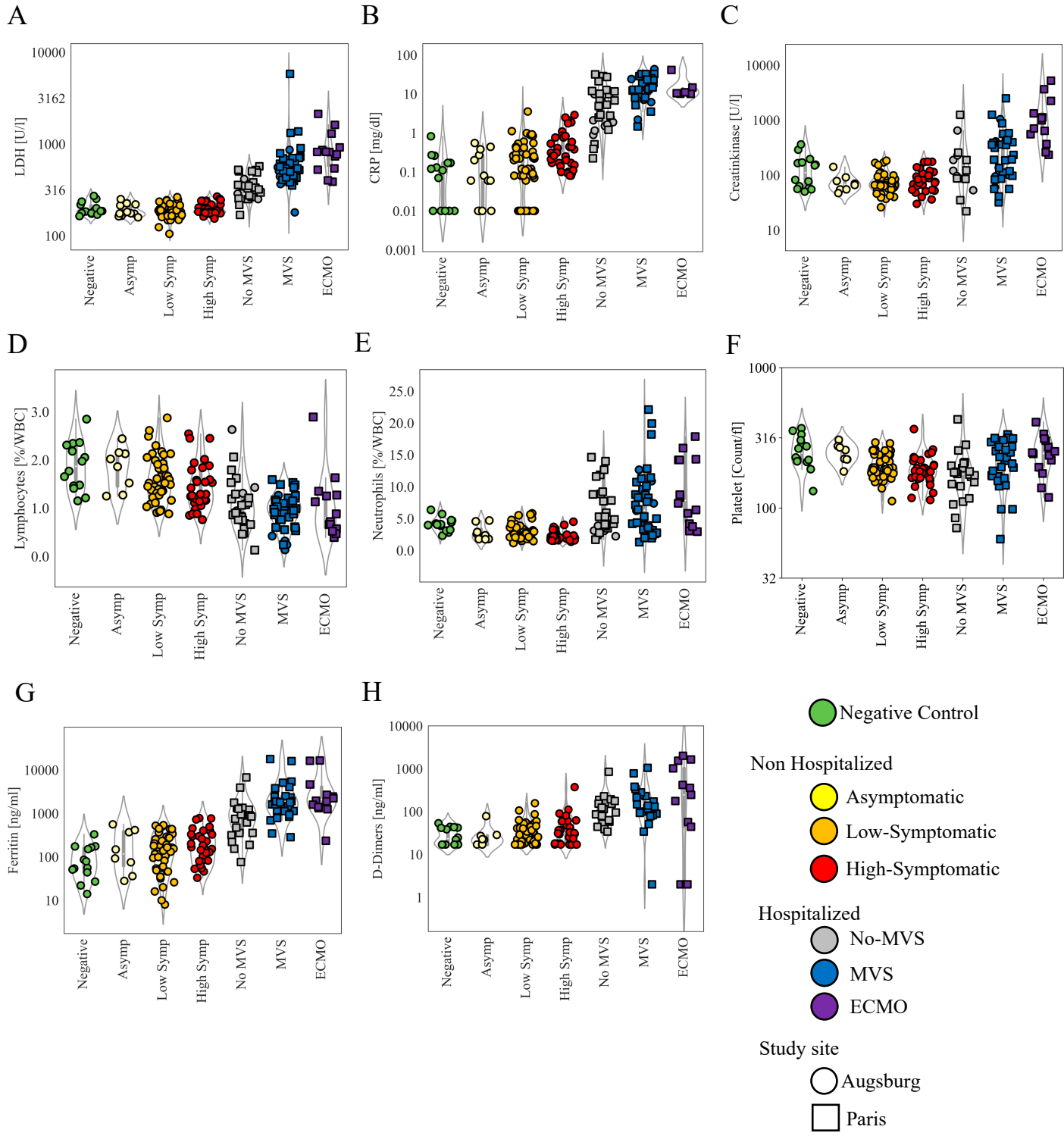
Supplementary Figure 1

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

%patients with symptom per group

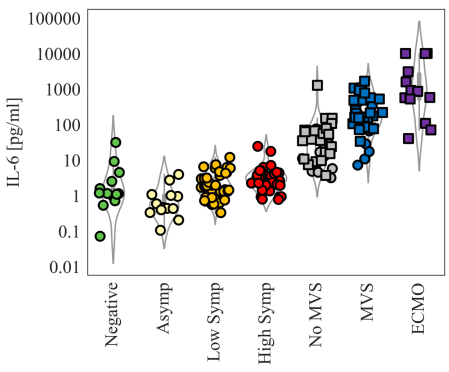


Supplementary Figure 2

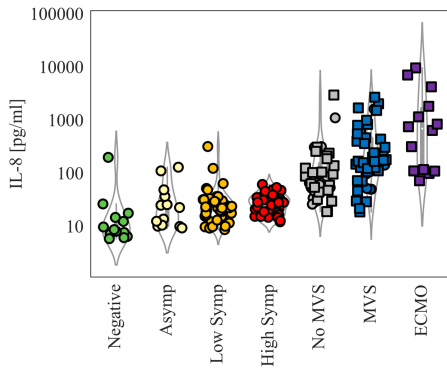


Supplementary Figure 3

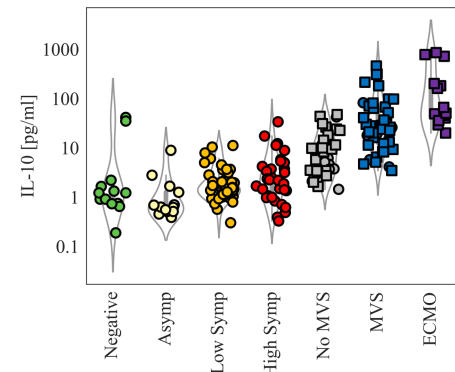
A



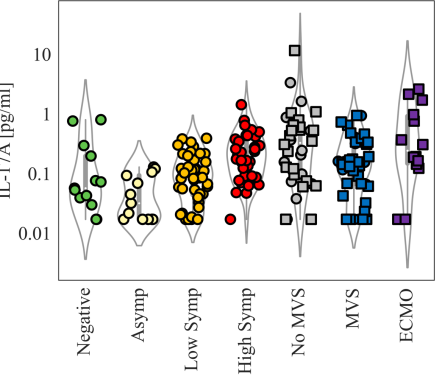
B



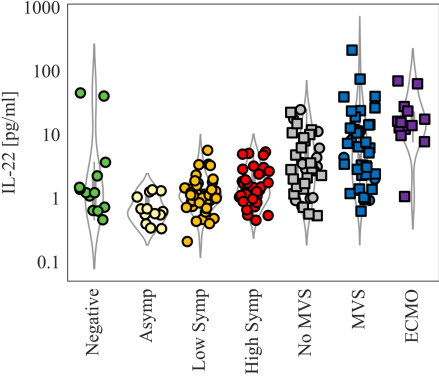
C



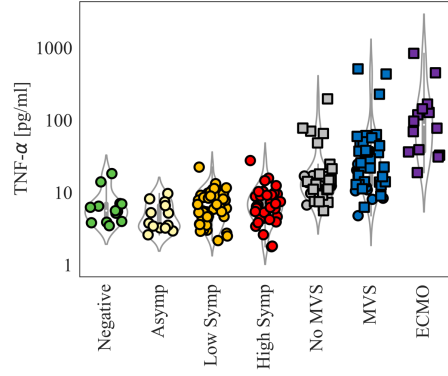
D



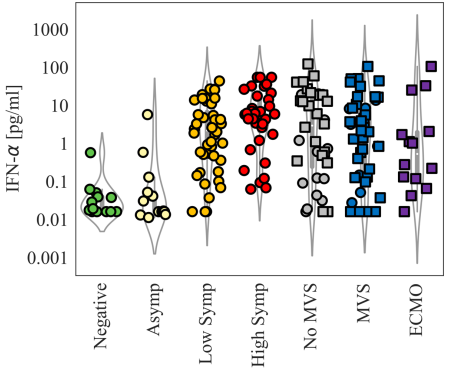
E



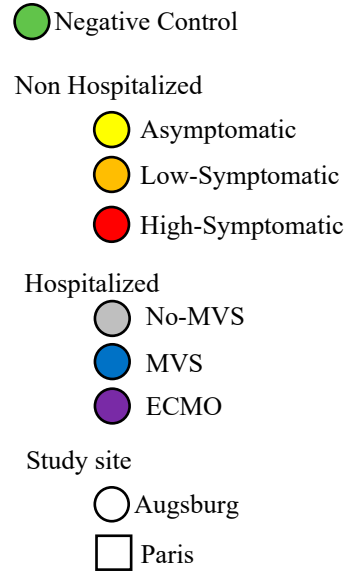
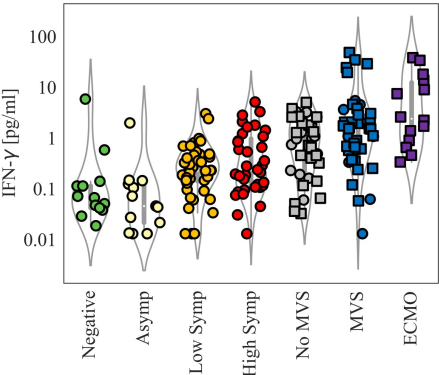
F



G

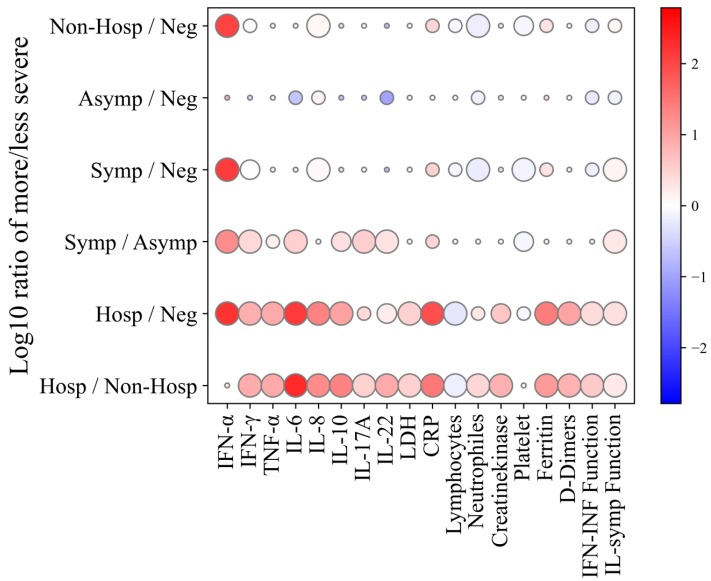


H

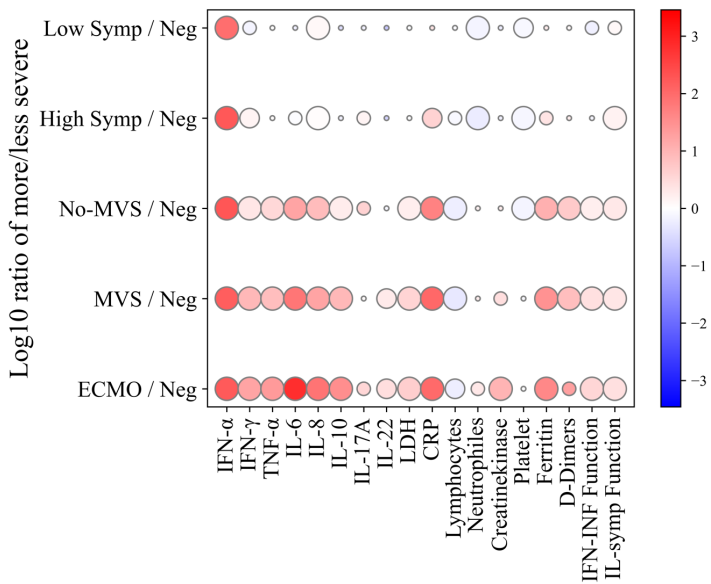


Supplementary Figure 4

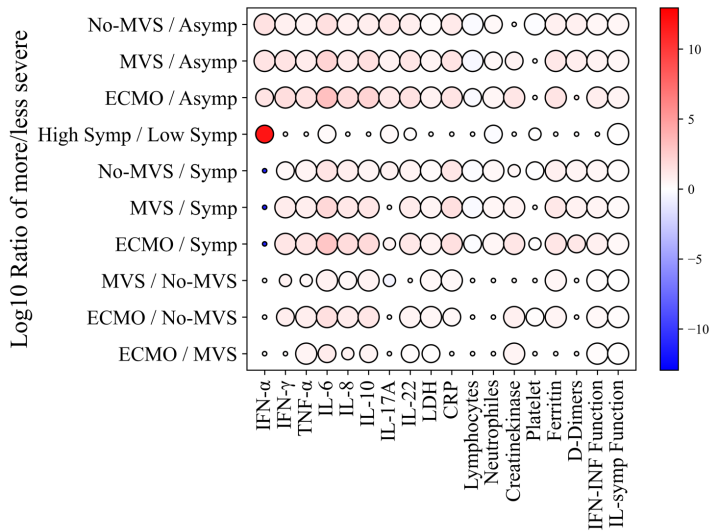
A



B



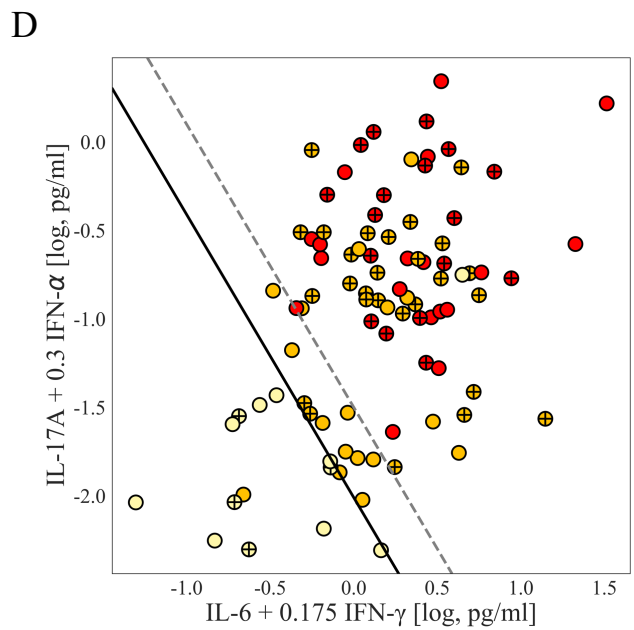
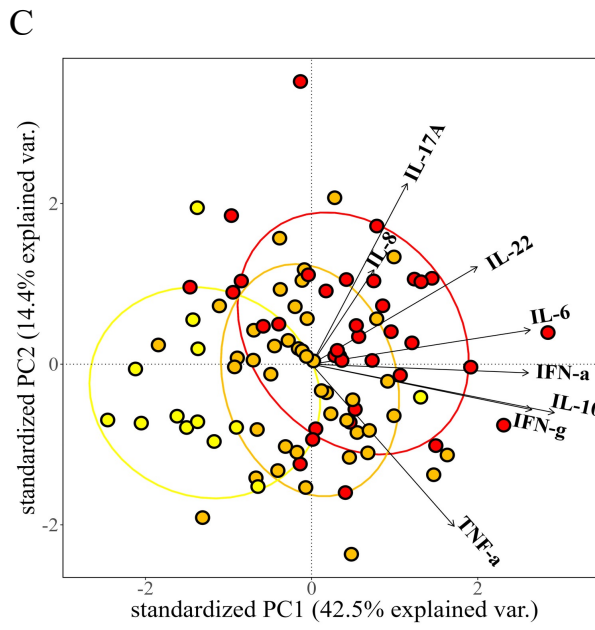
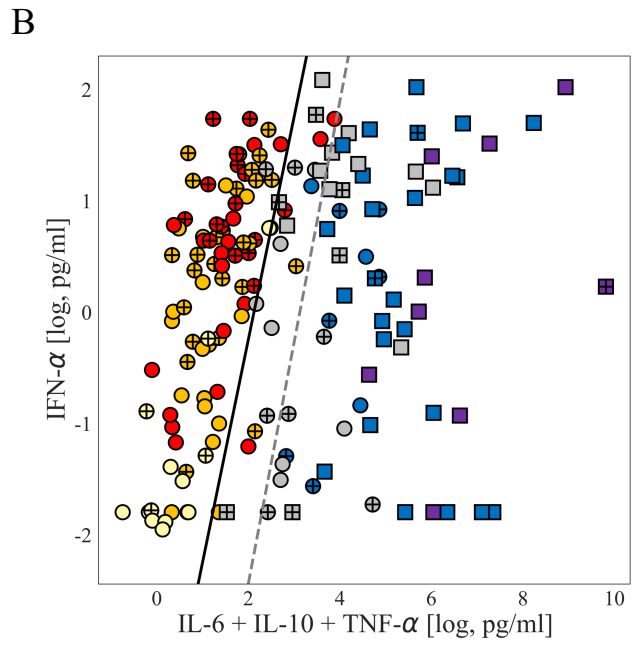
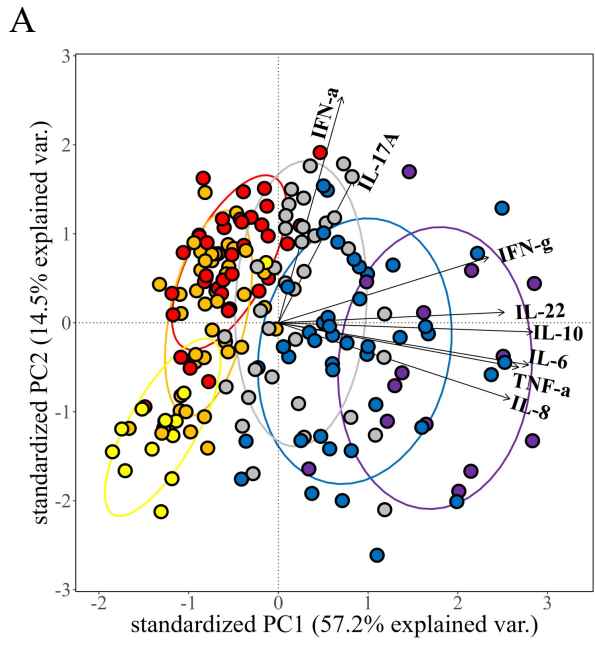
C



P-Value:



Supplementary Figure 5



Non Hospitalized

- Asymptomatic
- Low-Symptomatic
- High-Symptomatic

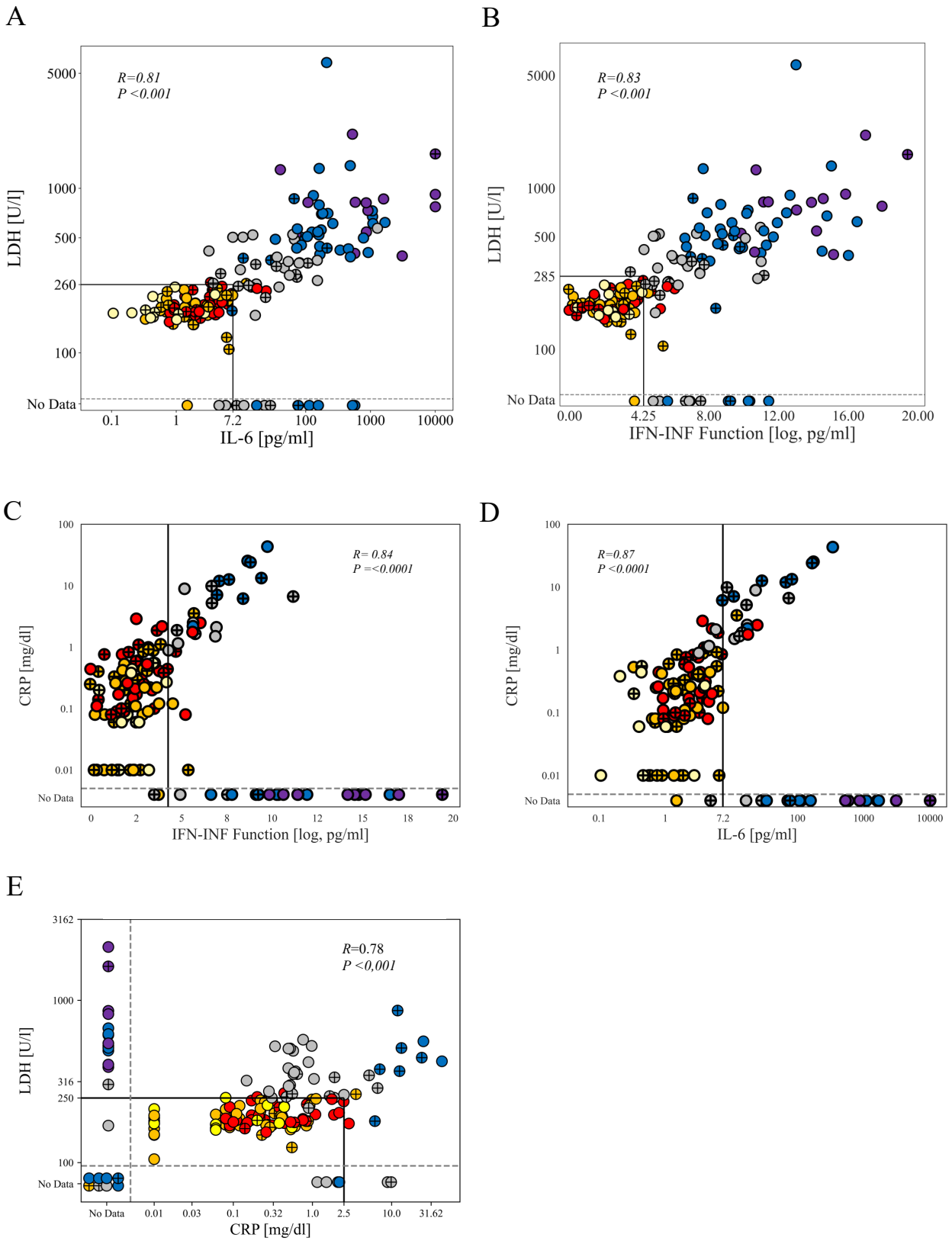
Hospitalized

- No-MVS
- MVS
- ECMO

Study site

- Augsburg
- Paris
- ⊕ 0-2 days after PCR

Supplementary Figure 6



Non Hospitalized

- Asymptomatic
- Low-Symptomatic
- High-Symptomatic

Hospitalized

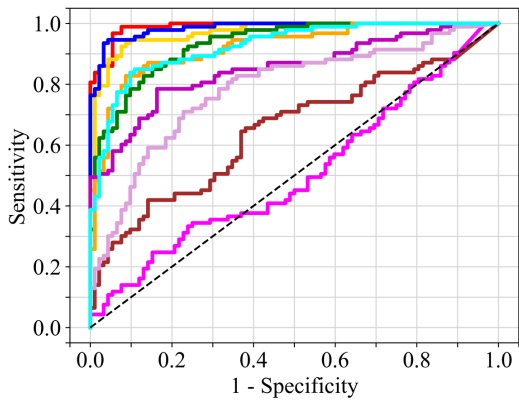
- No-MVS
- MVS
- ECMO

Study site

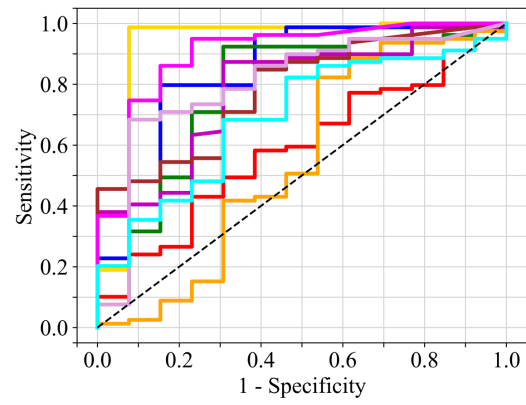
- Augsburg
- Paris
- 0-2 days after PCR

Supplementary Figure 7

A

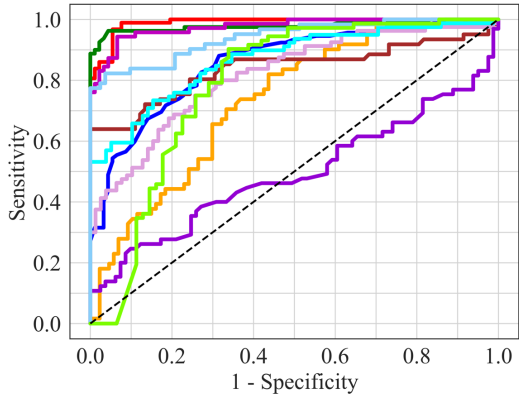


B

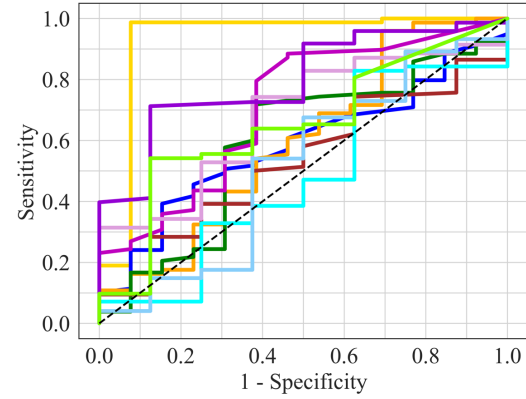


IFN-INF Function IL-symp Function IL-6 IL-10 IL-8
IL-22 IL-17A IFN- α IFN- γ TNF- α --- Random

C

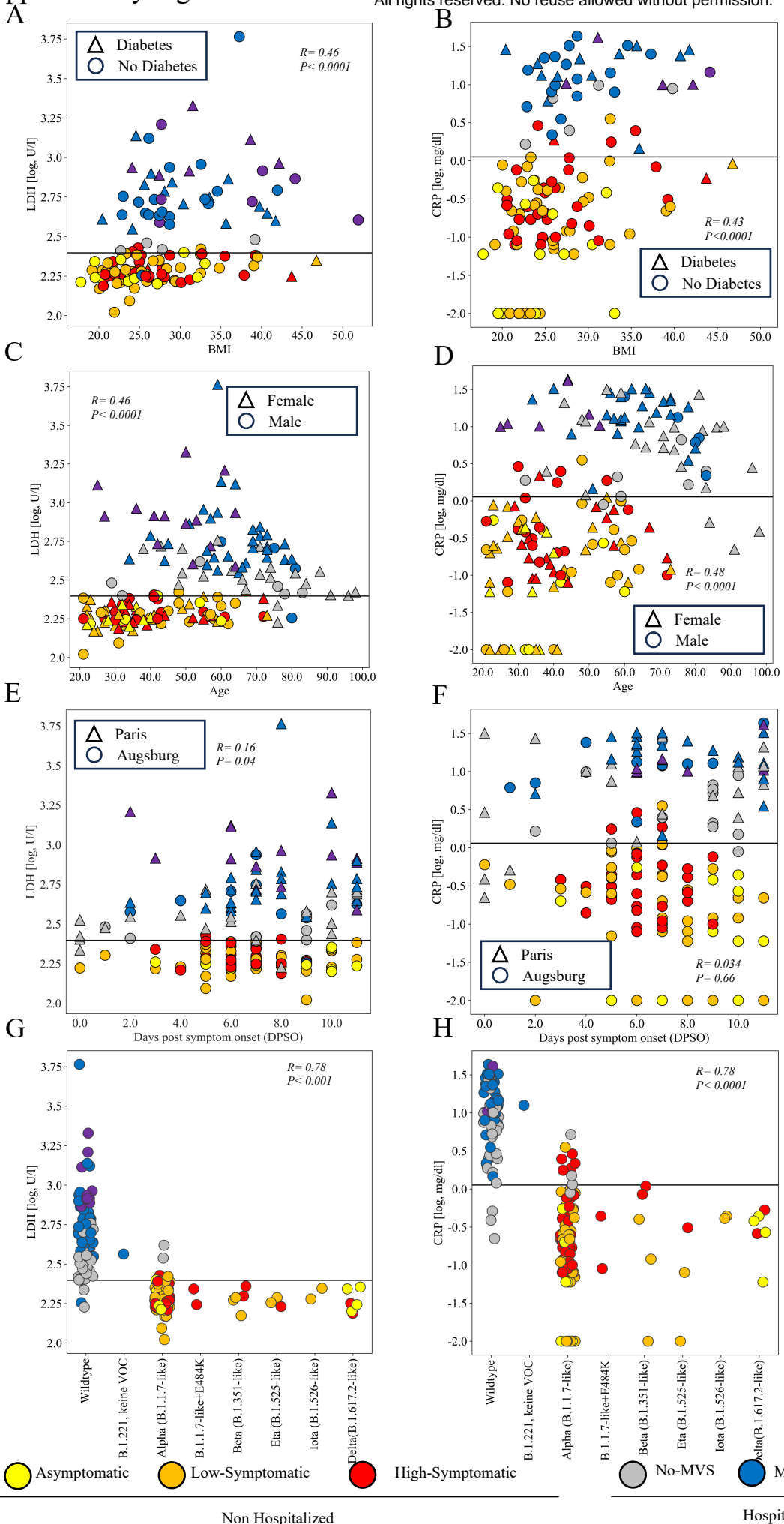


D



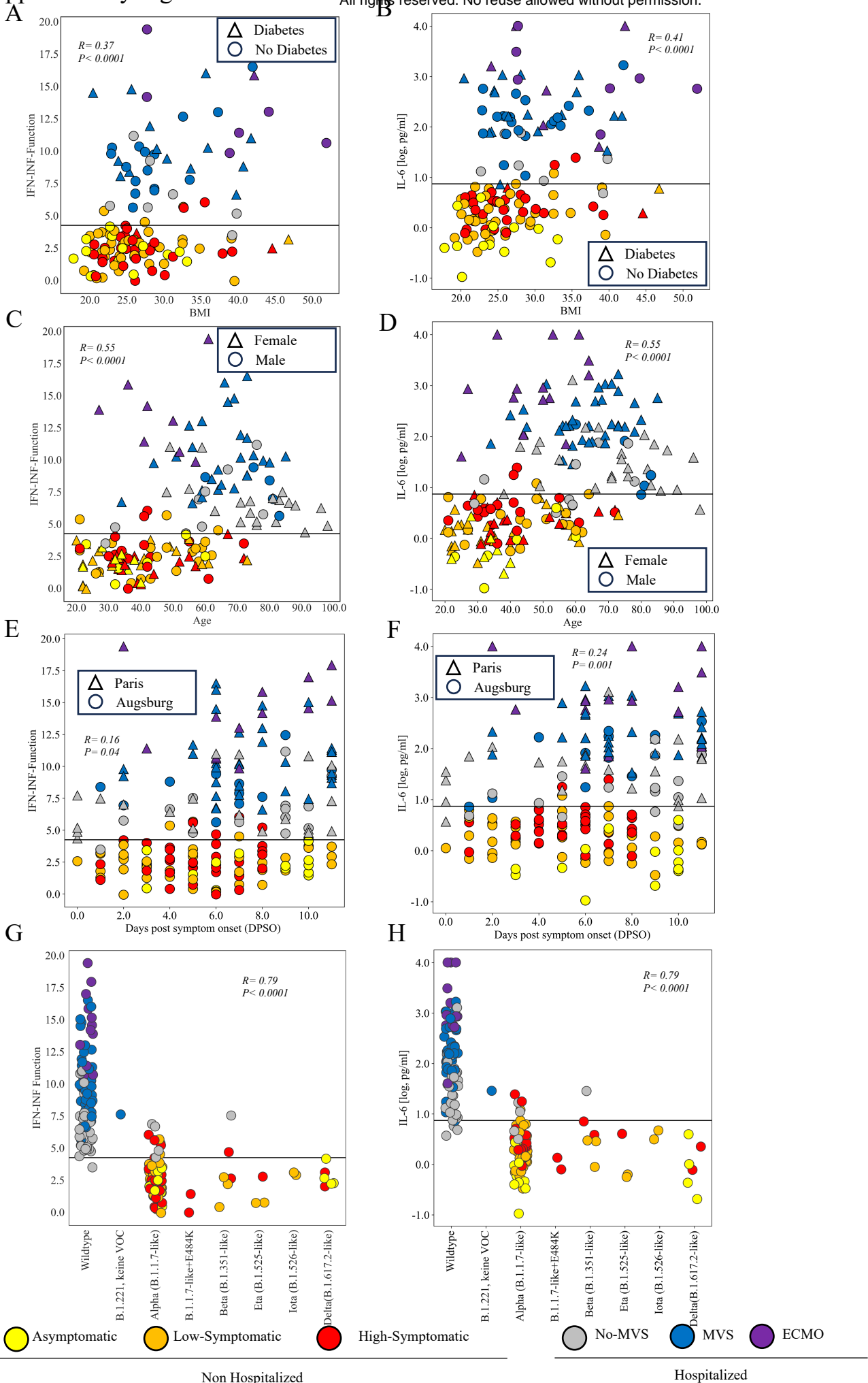
IFN-INF Function IL-symp Function Age BMI LDH CRP

Creatinkinase Lymphocytes Neutrophiles Platelet Ferritin D-Dimers --- Random



Non Hospitalized

Hospitalized



Supplementary Figure 10

