

Early and accurate prediction of COVID-19 hospitalization risk - methodological framework for pandemic preparedness

Corinna Holetschek MSc^{1,2,3 *}, Mehmet Goekkaya PhD^{1,2 *}, Karim Dorgham PhD^{4 *}, Paul Eder¹, Daria Luschkova¹, Amedeo De Tomassi¹, Omaira da Mata-Jardin PhD⁴, Christophe Parizot PhD⁵, Mehmet Tekinsoy MD⁶, Assia Samri PhD⁴, Denise Rauer PhD^{1,2}, Matthias Reiger PhD^{1,2}, Gertrud Hammel DiplSoz^{1,2}, Philipp Steininger MD⁷, Stefanie Gilles PhD¹, Ulrike Protzer MD⁶, Christoph Römmele MD⁸, Guy Gorochoy PhD^{4,5 #}, Claudia Traidl-Hoffmann MD^{1,2,3,9 #} and Avidan U. Neumann PhD^{1,2 #} for the Early-Opt-COVID19 study group

1 Institute of Environmental Medicine and Integrative Health, 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), Sorbonne University, 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, School of Medicine and Health, Technical University of Munich and Helmholtz 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 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, Institute of Environmental Medicine, Helmholtz Munich, Neusaesser Str. 47, D-86156, Augsburg, Germany

E-mail: avidan.neumann@uni-a.de

Supplementary Material

Contents

- 1) Supplementary Methods**
- 2) Supplementary References**
- 3) Supplementary Tables**
- 4) Supplementary Figures**

1. Supplementary Methods

Patients and Study Cohorts

From 1.2.2021 until 14.9.2021, n=116 patients were recruited in Augsburg, of which n=83 were recruited within the Early-Opt-COVID-19 study, n=10 within the COVID-19 Vaccine Consortium (CoVaKo) study¹ and n=23 at the University Hospital Augsburg. We included results from our previous study² for n=73 patients hospitalized (between 17.3.2020 to 11.9.2020), at Pitié-Salpêtrière hospital, Paris. In addition, n=12 patients infected with the Omicron variant were later (14.10.2021 to 5.12.2022) included within the CoVaKo study for validation. Lastly, n=15 control subjects with negative SARS-CoV-2 PCR test were enrolled in Early-Opt-COVID-19 study.

In addition to the patients recruited in the Early-Opt-COVID19 study, patients from three additional studies were added to the cohort. SARS-CoV-2 positive, non-hospitalized patients (N=9) and hospitalized patients (N=1) from the Corona-Vakzin-Konsortium project (CoVaKo)¹ were included. This was a longitudinal multicenter controlled cohort study in Bavaria, Germany. The CoVaKo patients added to our cohort were all from Augsburg, recruited with similar protocol to Early-Opt-COVID19.

Samples from hospitalized patients from the University Hospital Augsburg (N=23) were added from the Biobank Augsburg (COKA). The patients were chosen based on the Early-Opt-COVID19 study inclusion criteria. Additionally, hospitalized patients from our previous study cohort² originating in Paris (n=73) were included in this cohort. The same inclusion criteria as for the other hospitalized patients were applied.

All samples were processed and measured together with the Early-Opt-COVID19 study samples.

Only patient samples collected within 11 DPO were used for the analysis of hospitalization risk prediction (Table 1 and Supplementary Table 2), since at later time points there was a significant decline in many biomarker levels. Only patient samples collected within 6 DPO were used for the analysis of moderate symptom severity prediction in non-hospitalized patients (Supplementary Table 1 and Supplementary Table 2), since the levels of the relevant biomarkers started to decline after day 6 (Supplementary Figure 12). The demographic and clinical aspects are summarized in Table 1 and Supplementary 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=37). Mechanical ventilatory support (MVS, n=46) patients required invasive mechanical ventilation and ECMO patients (n=14) patients required extracorporeal membrane oxygenation.

Definition of Symptom Severity Groups

The symptom severity groups of the non-hospitalized patients were divided into three different symptom groups based on the self-reported questionnaires. Asymptomatic patients had overall no reported symptoms, while mild and moderate patients stated distinct COVID-19 related symptoms (Supplementary Figure 1). Patients with at least 1 out of 6 symptoms (confusion, shortness of breath, fever >38.5°C, fatigue, chest pain, dizziness) linked to dyspnea, which were identified by the WHO as symptoms to immediately seek medical help³, were defined as moderate. Mild patients showed less severe symptoms and fewer symptoms in general compared to moderate patients. Out of 19 symptom categories, mild patients exhibited on average 4 symptoms (IQR: 3.0 – 5.0) while moderate patients had significantly ($P < 0.001$) more symptoms (mean 8.1, IQR: 6.0 – 11.0). Additional details regarding symptom classification can be found in Supplementary Figure 1.

Biological Sampling

For the first visit at the test center only one peripheral blood sample (Serum Gel S-Monovette, SARSTEDT) was drawn from the patients. Four Serum Gel S-Monovettes, a citrate Monovette, and an EDTA Monovette were drawn from the patients on each of the next three visits. In addition, on each of the successive visits, an oropharyngeal swab was taken and stored in a medium containing Guanidine Thiocyanate (PrimeStore MTM) to stabilize the RNA.

Biochemistry and Cell Differentiation

From the two Serum Monovettes, a citrate tube and an EDTA tube, the biochemistry and cell differentiation measurements were done at the laboratory of the University Hospital in Augsburg by standardized procedures. Behind the measured value, the unit and reference values are given in brackets if available. From the serum, LDH (U/l, 0-250), creatine kinase (U/l), and CRP (mg/dl, 0-0.5) were measured. A differential blood count was performed from EDTA whole blood and the mean platelets volume (fl), neutrophils (%; 40-78), and lymphocytes (%; 20-50) were counted. In addition, Ferritin (ng/ml) levels were measured from whole blood.

For a number of patients, there was missing data for LDH (n=14 missing) and CRP (n=23 missing), see Supplementary Table 2.

Immunoassays

Cytokines, chemokines, and interferons (IFN- α , IFN- γ , IL-6, IL-8, IL-10, IL-17A, IL-22, and TNF- α) were measured in serum. Briefly, blood samples were collected in Serum Gel S-Monovettes (SARSTEDT), coagulated, and the serum was recovered by subsequent centrifugation. Samples were stored at -80°C until their analysis.

The concentration of the following cytokines (LOD/LLOQ) was measured using multiplex planar array immunoassay (Quanterix, SP-X, Lexington, MA, USA): IFN- γ (0.007/0.012), IL-6 (0.037/0.59), IL-8 (0.115/0.390), IL-10 (0.012/0.024), IL-22 (0.010/0.024) and TNF- α (0.063/0.098). In brief, pre-coated 96-well plates were incubated with 4 times diluted serum or calibrators at room temperature while shaking at 225 rpm. Following a wash step, a biotinylated

detection antibody was added. After 30 minutes of incubation, unbound antibodies were rinsed off, and Streptavidin-horseradish peroxidase was added. Following a last wash step, a combination of luminol and peroxide solution was added to generate a luminescent signal, which the SP-X Imaging System then detected, and the concentrations were calculated based on the calibration curve. Calibrators were run in duplicate and fitted with a five-parameter logistic (5PL) regression. Batch control samples were included in each plate.

Ultra-sensitive single-plex bead-based assays (Quanterix, HD-1 or HD-X, Lexington, MA, USA) were used for the detection of IFN- α and IL-17A according to the manufacturer's instructions on a fully-automated system. The IL-17A samples were used with a 1:4 dilution as written in the instructions. The LOD was 0.0042 pg/ml and the LLOQ 0.021 pg/ml. For the IFN- α measurements, the samples were diluted 1:2. The LOD was 0.003 pg/ml, and the associated LLOQ is 0.016 pg/ml. Calibrators were run in duplicate and fitted with a five-parameter logistic (5PL) regression. Batch control samples were included in each plate.

SARS-CoV-2 Variant Detection

SARS-CoV-2 variant detection per patient was performed at the Institute of Virology, Technical University of Munich, by the group of Prof. Protzer. Nucleic acid extraction was performed with the SEEPREP32-Kit (Seegene, Germany). Multiplex PCR for the detection of adenoviruses, influenza A and B, metapneumovirus, parainfluenza viruses, rhinoviruses, and respiratory syncytial virus (RSV) was performed using the Seegene Allplex™ RV Master (Seegene, Germany). Viral load was calculated from the CT value compared to a standard derived from German ring trials using an internal laboratory formula.

cDNA synthesis was performed using the extracted RNA from samples with a >100,000 Geq/mL viral load by LunaScript RT SuperMix Kit-based reverse transcriptase PCR. For next-generation sequencing of whole SARS-CoV-2 genomes, an ARTIC-PCR was performed. The cDNA was amplified with 2 different primer pools, which covered the complete sequence. Two

master mixes were prepared with primer pool 1, and two master mixes with primer pool 2 (40 cycles). The amplicon pools were diluted to 0.2 ng/μl for the Nextera XT library preparation (Illumina, USA). Nextera libraries were dual-barcoded and sequenced on an Illumina NextSeq 1000 (Illumina, USA) instrument. Sequencing and data analysis were performed by the Laboratory for Functional Genome Analysis (LaFuGa) at the Gene Center of Ludwig-Maximilians-University of Munich (LMU) using their proprietary bioinformatics pipelines. All variant data were obtained from the "LaFuGa SARS-CoV-2 Database".

SARS-CoV-2 RNA Load Measurement

Viral load was determined by SARS-CoV-2 RNA measurement via standardized RT-qPCR using two different SARS-CoV-2 specific oligonucleotide targets: N-Gene and NSP10, following the protocol by Bustin et al⁴. The primer mix consisted of forward and backward sequences of those targets as well as primer probes.

Oligonucleotides Sequences used for the Primer Mix:

Oligonucleotides name	Sequence
N-gene PROBE	ATG-TCT-GGT-AAA-GGC-CAA-CAA-CAA-CA
N-gene FOR	GCT-GCT-AGA-CAG-ATT-GAA-C
N-gene REV	AGC-AGA-TTT-CTT-AGT-GAC-AGT-TTG
Nsp10 PROBE	CAT-CGT-GTT-GTC-TGT-ACT-GCC-GTT-GCC
Nsp10 FOR	GGA-TCA-AGA-ATC-CTT-TGG-TGG
Nsp10 REV	GTC-ACA-AAA-TCC-TTT-AGG-ATT-TGG-A

Primers were mixed with a ratio of 10:10:4, with 4 for the probe sample. The Quantabio UltraPlex™ 1-Step ToughMix® (4X) was used for the reaction mix. Standards used for quantification were EDX SARS-CoV-2 Standard from Biorad. The qPCR was performed on the CFX384 and C1000 Touch from Biorad. Biorad standard had an original concentration of 200,000 cp/ml. We used a concentration of 1000 cp/5μl, diluted to 7.8/5μl in 1:2 dilution steps. Samples from the triplets with a deviation of more than 1.5 cycles were considered outliers and

removed. From the remaining samples, the mean was calculated. The master mix, standards and samples were pipetted in triplicates either manually onto a 384-well plate or using the Epimotion robotic system. Batch control samples were included in each plate.

SARS-CoV-2 N-Antigen Load Measurement

Serum N-antigen levels were measured by ELISA using the COV-QUANTO® kit (AAZ) according to the manufacturer's instructions. Anti-SARS-CoV-2 N protein antibody was used to coat the microplates and successive additions of the sample and of the anti-N protein SARS-CoV-2 antibody labeled with biotin. If the sample is positive (presence of the N protein), an immune complex "antibody-N antigen-antibody" labeled with biotin is formed. After washing the microplate, horseradish peroxidase (HRP)/streptavidin is added to bind to the preformed immune complex. Unbound substances are then washed away. A solution of chromogenic substrate containing TMB is then added to the microplate. The wells are blue in color and after stopping the reaction with the Stop Solution they may turn yellow. The intensity of the yellow coloration produced is proportional to the concentration of antigen in the serum sample. The absorbance value is to be read using a microplate reader at a wavelength of 450 nm using a wavelength between 620nm and 690nm as a reference. The concentration of SARS-CoV-2 N protein in the sample is calculated using the SARS-CoV-2 calibrators. The absorbance was measured using the Sunrise Tecan, with the corresponding software, the standard curve was analyzed, and levels in pg/ml were calculated from the measured fluorescent levels. Batch control samples were included in each plate.

Bioinformatics analysis

Descriptive and statistical data analyses were performed using R, version 4.2.2 (Rstudio Inc., Boston, USA) and Python version 3.9.15. Cytokine concentrations below the limit of detection were set to the corresponding LOD. Batch correction of the cytokine levels was applied to

mitigate batch effects between results previously measured in the Dorgham et al³² study and those currently measured. For that, we used repeated measurements of samples with a wide range of concentrations, which resulted in parallel linear regression, and calculated the fold change between the 2 linear regression functions.

Principal component analysis (PCA) over the Z-scaled normalized levels of the 8 cytokines was used to compute principal components weights (Supplementary Table 3) for each cytokine (using R packages ggbiplot and prcomp), for all patients together ($DPSO \leq 11$) for the prediction of hospitalization risk or only for non-hospitalized patients ($DPSO \leq 6$) for the prediction of symptomatic status. A smaller number (2-4) of cytokines with the largest PC factor (either in PC1 and/or PC2) was then selected. The raw values of the selected cytokines were transformed into a 2-dimensional space using the PC factors, and tested for the predictive value of the combination. This was done either by using the parameters of the linear separatrix line separating the relevant groups to construct a 1-dimensional function (IFN-INF function), or by using an OR function between 2 biomarker combinations being each larger than a corresponding threshold (e.g., IMM-INF and LDH-or-IL-6).

Receiver Operating Characteristic (ROC) curves for the OR function biomarker combinations was performed by calculating for each patient the minimum distance between the (X,Y) values to the 2 corresponding thresholds (when both values were higher than both thresholds a chi-square distance to intersection of thresholds was calculated), thus a minimum distance of 0.0 separates the positive versus negative prediction. The thresholds for each prediction were obtained by the largest Youden J-index (sensitivity+specificity-1) on the corresponding ROC curve.

Longitudinal data was analyzed by piecewise linear regression (using the pwlf package⁵ in Python version 3.9.15). The average of the negative control group samples was used as a baseline at day -3, based on the results of the SARS-CoV-2 human challenge by Killingley et

al⁶, for all other groups. The mean of all patients per group was fitted with 3 breakpoints, corresponding to the average time of the visits.

Statistical analysis

Statistical significance between the 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 less than 0.05 was considered statistically significant. Since 18 biomarkers were tested, p-value lower than 0.0027 was considered significant after multiple testing correction using the Bonferroni rule.

2. Supplementary References

1. Prelog M, Jeske SD, Asam C, et al. Clinical and immunological benefits of full primary COVID-19 vaccination in individuals with SARS-CoV-2 breakthrough infections: A prospective cohort study in non-hospitalized adults. *Journal of Clinical Virology* 2024; **170**: 105622.
2. Dorgham K, Quentric P, Gökkaya M, et al. Distinct cytokine profiles associated with COVID-19 severity and mortality. *Journal of Allergy and Clinical Immunology* 2021; **147**(6): 2098-107.
3. WHO. Coronavirus disease (COVID-19). 2023. [https://www.who.int/news-room/fact-sheets/detail/coronavirus-disease-\(covid-19\)](https://www.who.int/news-room/fact-sheets/detail/coronavirus-disease-(covid-19)) (accessed 07.01.2025).
4. Bustin, S., Coward, A., Sadler, G. et al. CoV2-ID, a MIQE-compliant sub-20-min 5-plex RT-PCR assay targeting SARS-CoV-2 for the diagnosis of COVID-19. *Sci Rep* 2020; **10**, 22214.
5. Jekel C, Venter G. pwlf: A Python Library for Fitting 1D Continuous Piecewise Linear Functions; 2019.
6. Killingley B, Mann AJ, Kalinova M, et al. Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults. *Nat Med* 2022; **28**(5): 1031-41.

3. Supplementary Tables

Supplementary Table 1: Patient demographic and clinical parameters by severity groups (DPSO≤6).

DPSO ≤6	Non Hospitalized (N=59) ⁽¹⁾		Hospitalized (N=40) ⁽¹⁾			All (N=99)	p-Value ⁽⁵⁾ Mild vs Moderate	p-Value ⁽⁵⁾ Hosp vs Non-Hosp
	Mild ⁽²⁾ (N=17)	Moderate (N=42)	No MVS (N=17)	MVS (N=18)	ECMO (N=5)			
Augsburg (N) ⁽³⁾	17	42	5	6	0	70		
Paris (N) ⁽⁴⁾	0	0	12	12	5	29		
Age(y), median (IQR)	40.06 (29.0-57.0)	41.43 (30.5-54.25)	69.88 (59.0-84.0)	69.22 (64.5-79.5)	41.2 (27.0-52.0)	51.12 (33.0-67.0)	ns	ns
Sex: male, N (%)	9 (52.94)	20 (47.62)	12 (70.59)	12 (66.67)	5 (100.0)	58 (58.59)	ns	0.024
Days post symptom onset, Median in days (IQR)	3.82 (2.0-5.0)	3.74 (3.0-5.0)	2.82 (1.0-5.0)	4.78 (4.25-6.0)	4.6 (3.0-6.0)	3.83 (2.0-6.0)	ns	ns
Past medical history, N(%)								
Obesity (≥30)	14 (82.35)	32 (76.19)	13 (76.47)	12 (66.67)	2 (40.0)	73 (73.74)	ns	0.086
Diabetes Type 2	0 (0)	2 (4.76)	8 (47.06)	8 (44.44)	1 (20.0)	19 (19.19)	ns	<0.001

(1) Hospitalized patients are patients being hospitalized due to their SARS-CoV-2 infection while non-hospitalized patients were recovering at home

(2) N=1 asymptomatic patient included here within the Mild group.

(3) Patients included in Augsburg from the Early-Opt-COVID19 (n=83), CoVaKo-Augsburg (n=9)¹ and University Hospital Augsburg (n=23) cohorts.

(4) Patients included in Paris from our Dorgham et al² study.

(5) Hospitalized versus non-hospitalized and mild versus moderate patients were compared using Mann-Whitney or Fisher Exact tests as mentioned in methods.

Supplementary Table 2: Number of samples available per measurement type of predictive biomarkers by COVID-19 severity groups and $DPSO \leq 11$ (A) or $DPSO \leq 6$ (B).

A $DPSO \leq 11$

Samples	Asymptomatic	Mild	Moderate	No MVS	MVS	ECMO	N
Included	5	21	66	37	46	14	189
Cytokines	5	21	66	37	46	14	189
LDH	5	20	66	30	41	14	176
CRP	5	20	66	33	36	6	166
Serum Antigen	3	19	61	34	36	12	165
Oral RNA	3	17	59	2	1	0	82

B $DPSO \leq 6$

Samples	Asymptomatic	Mild	Moderate	No MVS	MVS	ECMO	N
Included	1	16	42	17	18	5	99
Cytokines	1	16	42	17	18	5	99
LDH	1	11	31	14	14	5	76
CRP	1	11	31	15	15	2	75
Serum Antigen	1	16	42	17	12	4	92
Oral RNA	1	9	29	1	1	0	41

Supplementary Table 3: Principle component (PC) factors obtained for PCA analysis of the 8 measured cytokines for prediction of hospitalization risk and moderate disease course (Supplementary Figure 7).

	Hospitalization Prediction (N=189)		Moderate Prediction (N=59)	
PCA Factors	PC1	PC2	PC1	PC2
IFN- α	0.110788	0.758937	0.393323	0.065625
IL-17A	0.121007	0.511536	0.186871	-0.59644
IFN- γ	0.360554	0.222802	0.426565	0.059791
IL-6	0.428916	-0.13569	0.467709	-0.06956
IL-8	0.392847	-0.26154	0.185748	-0.45294
IL-22	0.380393	0.033769	0.283404	-0.29289
TNF- α	0.413332	-0.15364	0.262736	0.523958
IL-10	0.435023	-0.03381	0.475202	0.257045

4. Supplementary Figures

eFigure 1

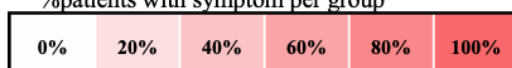
A Patients with DPSO ≤ 11

Symptoms	Asymp (N=5)	Mild (N=21)	Moderate (N=66)	P Value (mild vs moderate)
Abdominal Pain	0%	0%	9%	ns
Diarrhea	0%	10%	26%	ns
Loss of Appetite	0%	5%	54%	<0.001
Nausea	0%	0%	22%	0.03
Cough	0%	75%	71%	ns
Hoarseness	0%	25%	37%	ns
Sore Throat	0%	30%	58%	ns
Blocked Nose	0%	65%	65%	ns
Running Nose	0%	55%	43%	ns
Headache	0%	45%	86%	<0.001
Loss of Smell	0%	35%	42%	ns
Loss of Taste	0%	30%	37%	ns
Muscle Pain	0%	25%	57%	0.02
Fever ($\geq 38.5^{\circ}\text{C}$) ⁽¹⁾	0%	0%	11%	ns
Confusion ⁽¹⁾	0%	0%	18%	0.06
Shortness of Breath ⁽¹⁾	0%	0%	23%	0.02
Chest Pain ⁽¹⁾	0%	0%	25%	0.02
Dizziness ⁽¹⁾	0%	0%	31%	0.003
Fatigue ⁽¹⁾	0%	0%	94%	<0.001
Average number of symptoms (IQR)	0 (0)	4.0 (3-5)	8.1 (6-11)	<0.001

B Patients with DPSO ≤ 6

Symptoms	Mild* (N=17)	Moderate (N=42)	P Value (mild vs moderate)
Abdominal Pain	0%	10%	ns
Diarrhea	12%	29%	ns
Loss of Appetite	6%	67%	<0.001
Nausea	0%	19%	0.09
Cough	71%	81%	ns
Hoarseness	18%	48%	0.04
Sore Throat	24%	64%	0.009
Blocked Nose	76%	74%	ns
Running Nose	47%	50%	ns
Headache	53%	93%	0.001
Loss of Smell	29%	40%	ns
Loss of Taste	29%	40%	ns
Muscle Pain	24%	69%	0.003
Fever ($\geq 38.5^{\circ}\text{C}$) ⁽¹⁾	0%	7%	ns
Confusion ⁽¹⁾	0%	21%	0.05
Shortness of Breath ⁽¹⁾	0%	24%	0.05
Chest Pain ⁽¹⁾	0%	26%	0.02
Dizziness ⁽¹⁾	0%	36%	0.003
Fatigue ⁽¹⁾	0%	95%	<0.001
Average number of symptoms (IQR)	3.9 (3-5)	8.9 (6.3-11)	<0.001

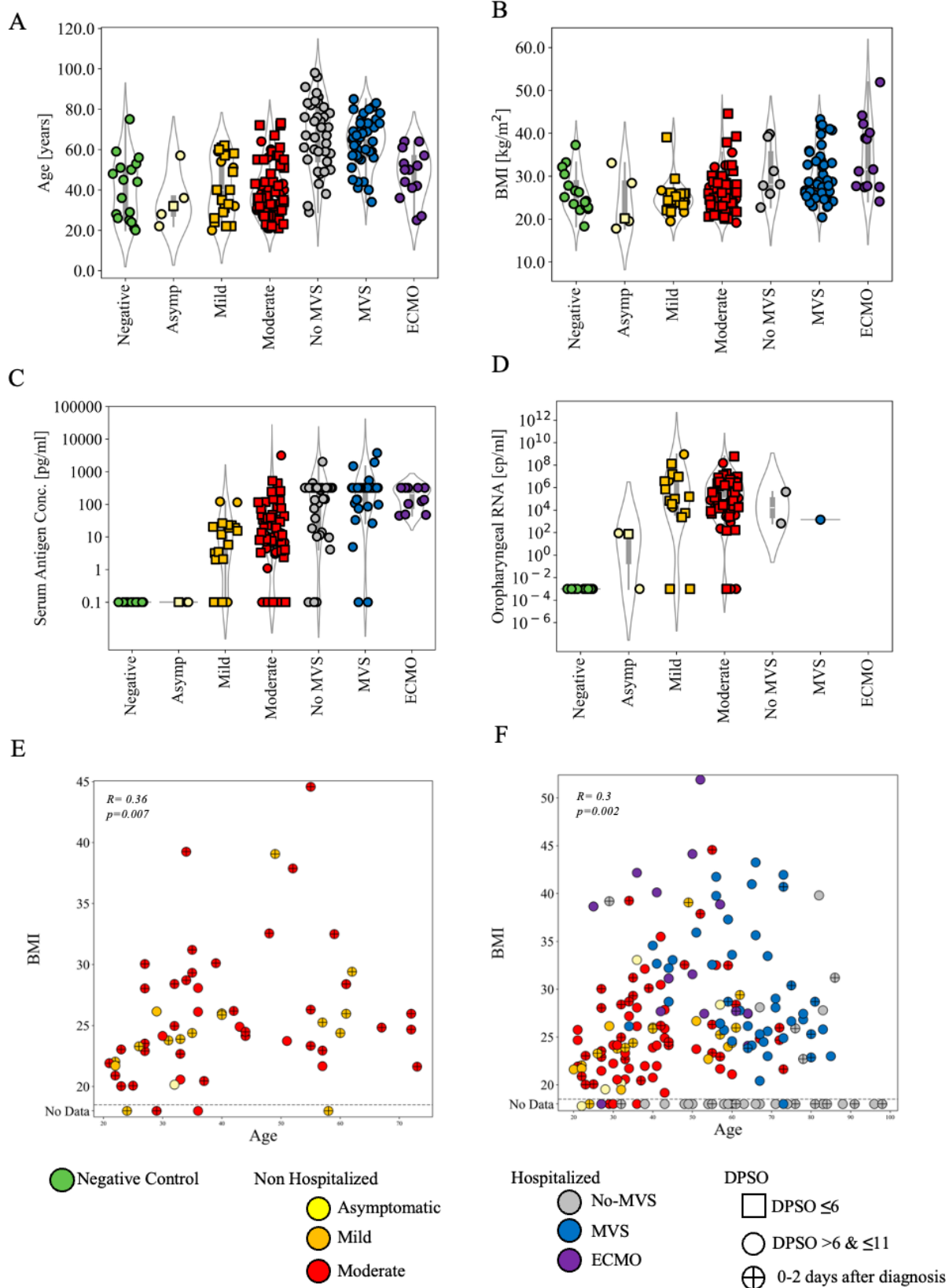
%patients with symptom per group



Supplementary Figure 1: Symptom distribution in different non-hospitalized symptom severity groups. The distribution of 14 different symptoms across asymptomatic, mild and moderate patients in the non-hospitalized group. Patients with $DPSO \leq 11$ are shown in (A), with all 3 subgroups, contrary to (B), where only patients with $DPSO \leq 6$ are shown and one asymptomatic patient is included in the mild group (marked with *). Significant differences are observed between mild and moderate patients for some of the symptoms, as well as for the number of symptoms a patient exhibits. Mild patients exhibited, on average, 4.0 (or 3.9) symptoms, significantly ($p < 0.001$) lower than 8.1 (or 8.9) symptoms in moderate patients.

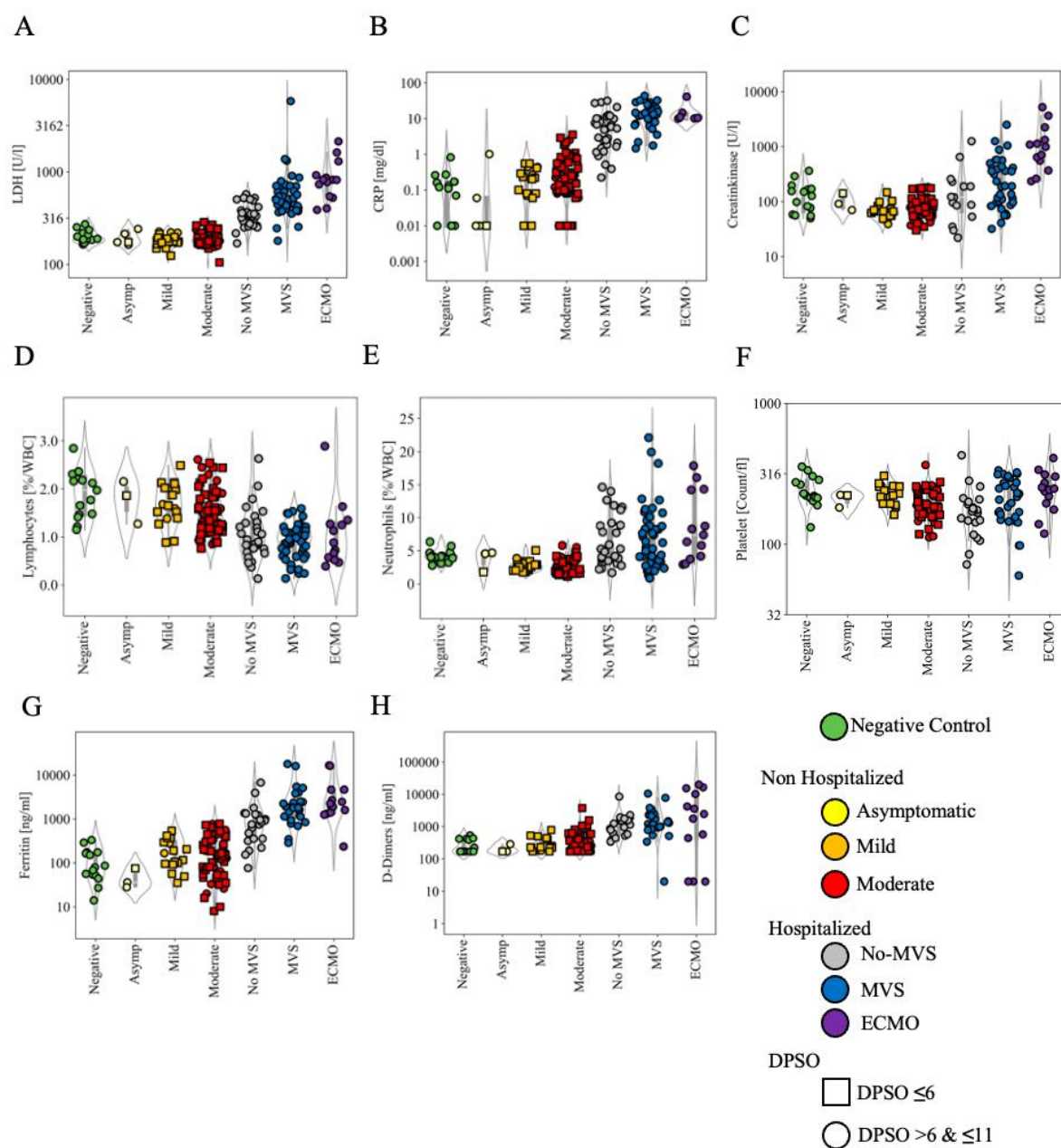
(1) Moderate disease course was defined as having 1 out of 6 of the symptoms most linked to risk of hospitalization (Fever $> 38.5^{\circ}\text{C}$, Confusion, Shortness of Breath, Chest Pain, Dizziness, Fatigue).

eFigure 2



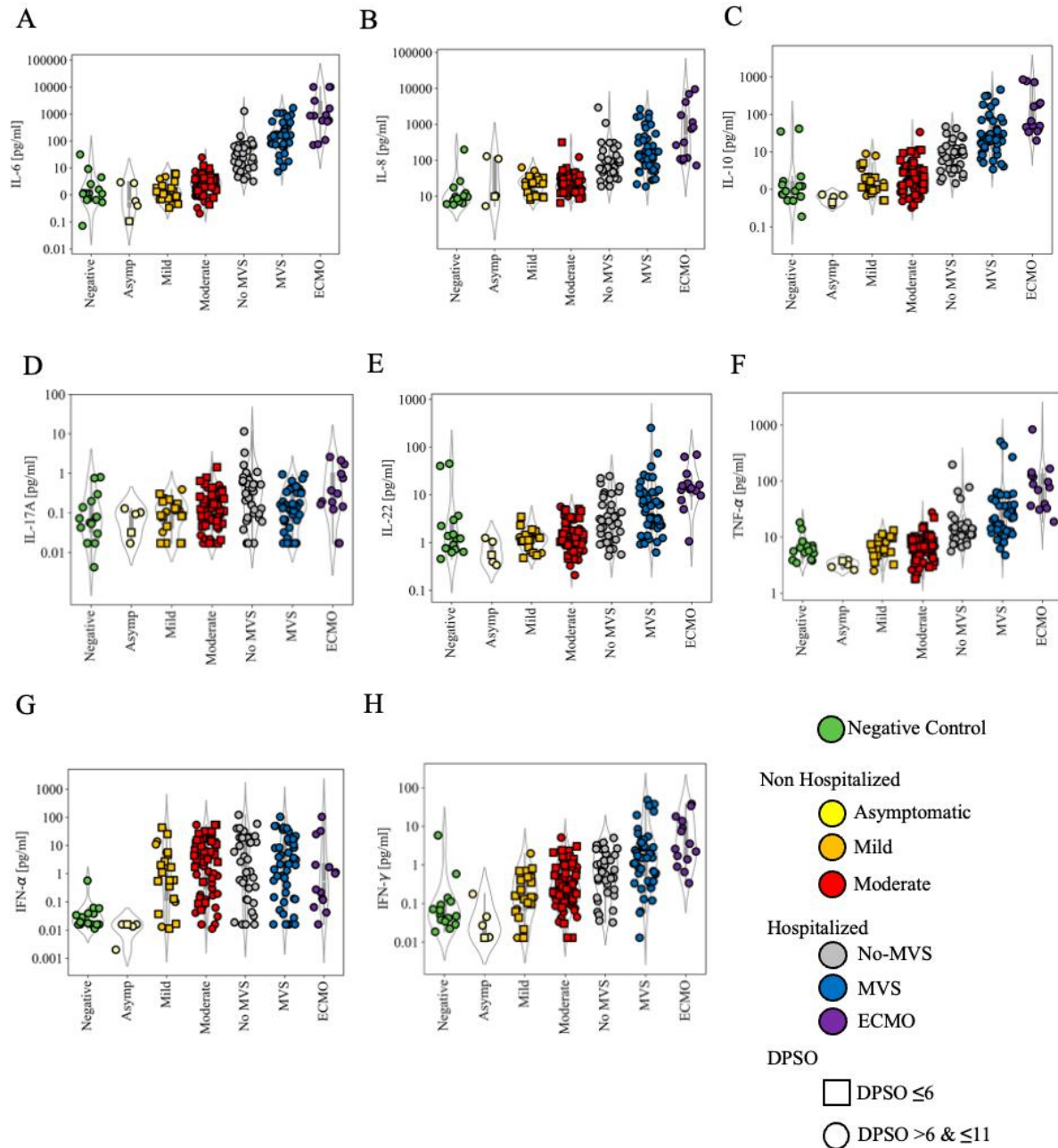
Supplementary Figure 2: Viral load levels, Age, and BMI across non-hospitalized symptoms groups and hospitalized severity classes as well as a negative control group. The viral load levels serum Antigen (C), oropharyngeal RNA (D), as well as age (A) and BMI (B), show significantly different levels between the severity classes and symptom groups (see Supplementary Figure 5). A wide range of levels for the single markers is also found within the negative patients control group, except for the obviously negative viral load. Circles mark patients with $DPSO > 6 \leq 11$, and squares mark patients with $DPSO \leq 6$. No additive predictive value comes out of the interaction between age and BMI, as seen when plotting the relationship between the two for $DPSO \leq 6$ non-hospitalized (E) and $DPSO \leq 11$ for all patient groups (F).

eFigure 3



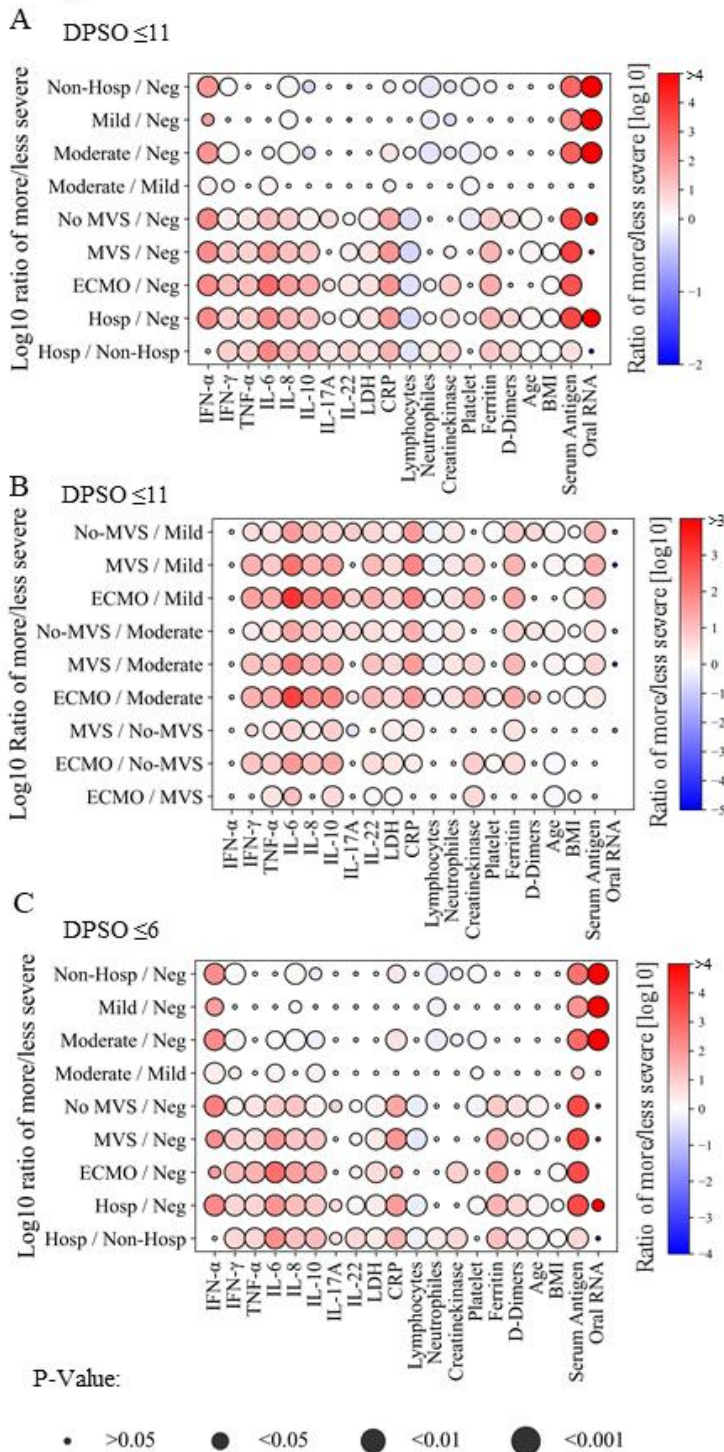
Supplementary Figure 3: Biochemistry biomarker levels across non-hospitalized symptoms groups and hospitalized severity classes as well as negative control group. The single biomarkers: LDH (A), CRP (B), Ceratinkinase (C), Lymphocytes (D), Neutrophils (E), Platelets (F), Ferritin (G) and D-Dimers (H), show significantly different levels between the severity classes and symptom groups (see Supplementary Figure 5). A wide range of levels for the single markers is also found within the negative patients control group. Circles mark patients with $DPSO > 6 \leq 11$ and squares mark patients with $DPSO \leq 6$.

eFigure 4



Supplementary Figure 4: Single Cytokine levels across non-hospitalized symptom groups and hospitalized severity classes as well as negative control group. The single cytokines: IL-6 (A), IL-8 (B), IL-10 (C), IL-17A (D), IL-22 (E), IFN- α (F), IFN- γ (G), and TNF- α (H), show significantly different levels between the severity classes and symptom groups (see eFigure 5). A wide range of levels for the single markers are also found within the negative patients control group. Circles mark patients with $DPSO > 6 \leq 11$ and squares patients with $DPSO \leq 6$.

eFigure 5

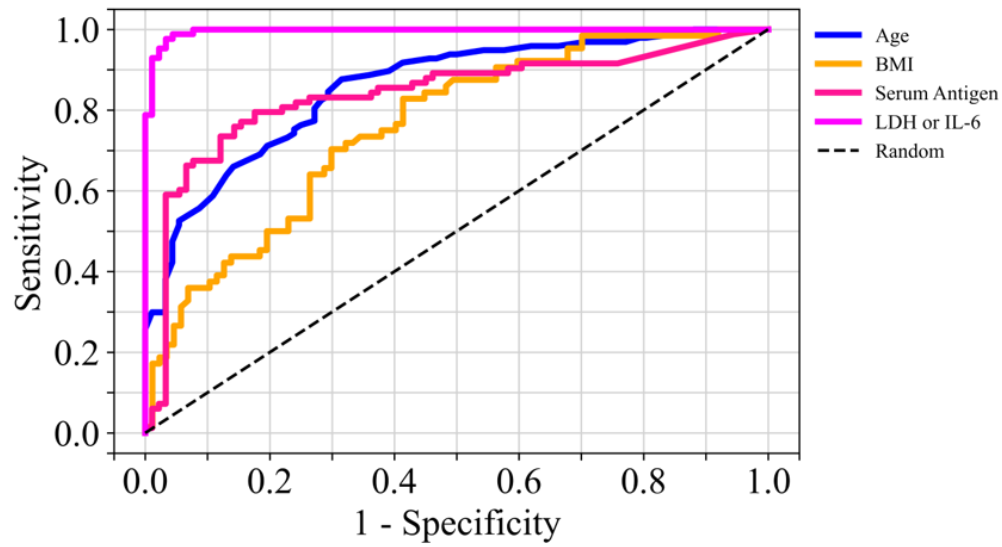


Supplementary Figure 5: Differences between the levels across non-hospitalized symptom groups and hospitalized severity classes, as well as the negative control group, for age, BMI, and the cytokines, biochemistry and viral load biomarkers. The color of the bubbles indicates the log10 ratio (more/less severe) differences between the various severity groups. The more red, the higher are the values in the more severe group, and, vice versa, the more blue, the higher are the values in the less severe group. The order for the groups severity is: negative,

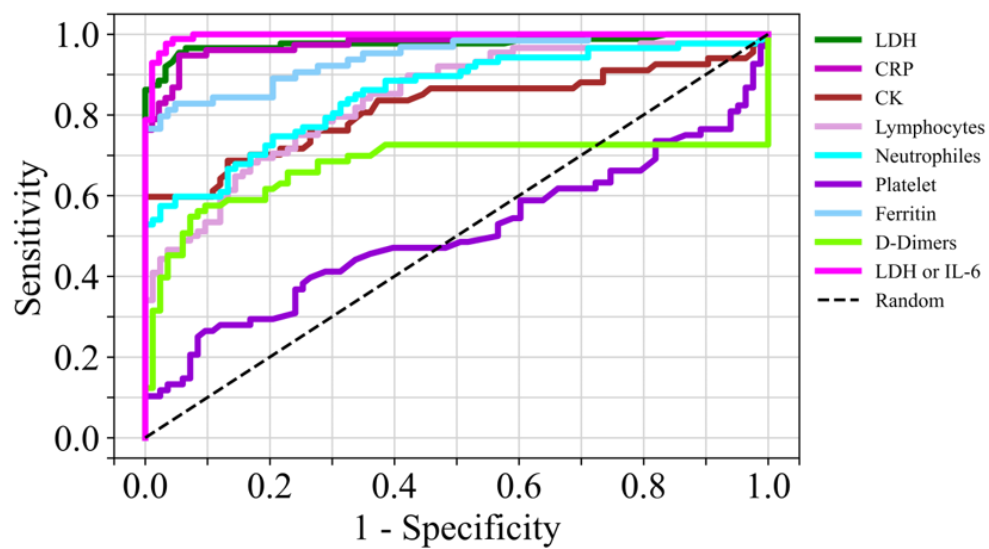
then non-hospitalized (Non-Hosp) ranging from mild (including asymptomatic) to moderate, and then hospitalized (Hosp, severe), ranging from No-MVS to MVS, and to ECMO. The bubble size shows the significance of the P-value for the differences. The differences between the main groups (moderate/mild and hospitalized/non-hospitalized), and between all groups to the negative control group, are shown in (A) for patients with $DPSO \leq 11$, and in (C) for patients with $DPSO \leq 6$. In (B) are shown the differences between all sub-groups, for patients with $DPSO \leq 11$. Actual values are depicted in Supplementary Figures 2-4.

eFigure 6

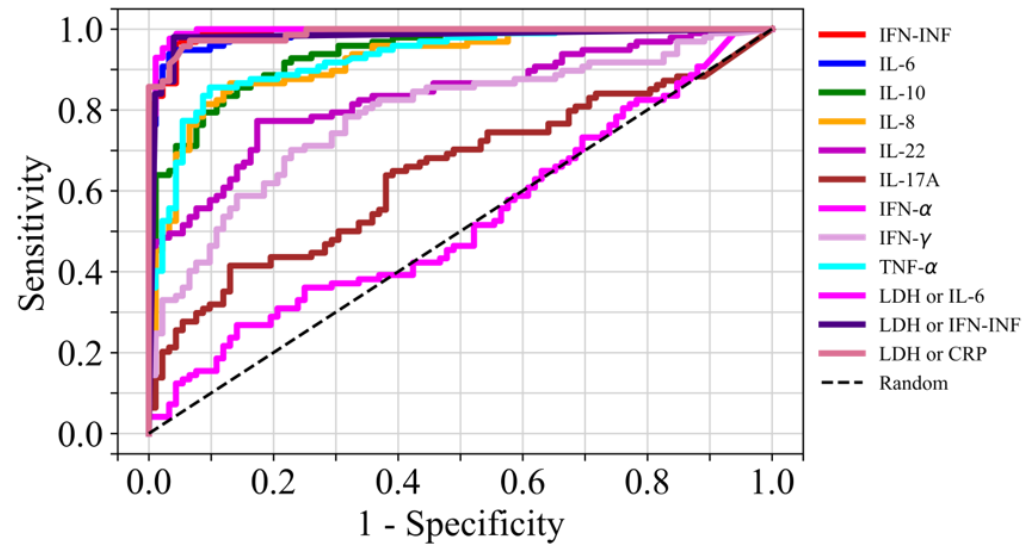
A



B



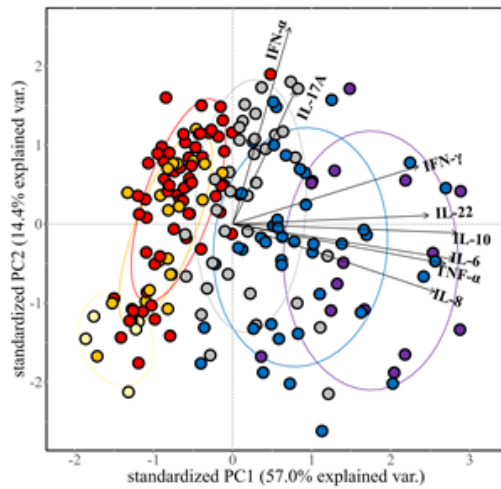
C



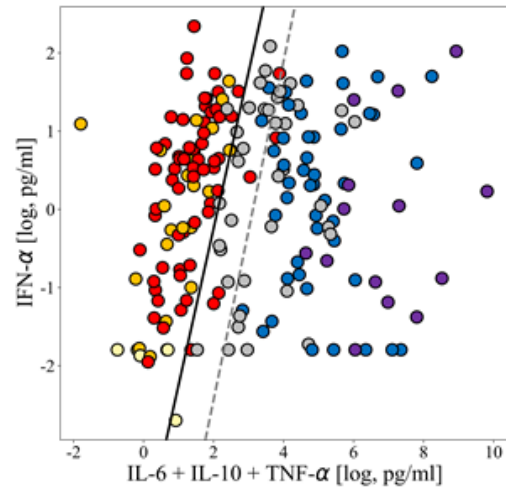
Supplementary Figure 6: ROC analysis comparison of the predictive values of age, BMI, SARS-CoV-2 N-antigen, and all measured biochemistry, hematology and cytokine biomarkers for prediction of hospitalization risk. Receiver Operating Characteristic (ROC) curves for age, BMI, and serum N-antigen (A) in comparison to LDH-or-IL-6 show a significantly higher AUC for the LDH-or-IL-6 predictor for hospitalization risk. ROC curves for the prediction of hospitalization risk using any of the measured biochemistry values (B) and any of the eight measured single cytokines as well as the IFN-INF function and various combinations of biomarkers (C), measured at the first visit per patient (DPSO $\leq$ 11). IFN-INF prediction function is defined by: $\log(2*IL-6*IL-10*TNF-\alpha/IFN-\alpha)$.

eFigure 7

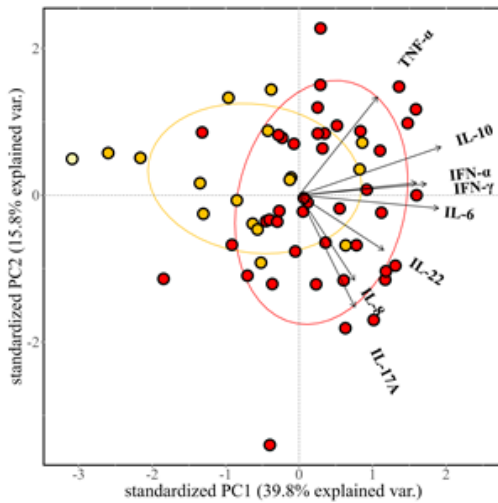
A



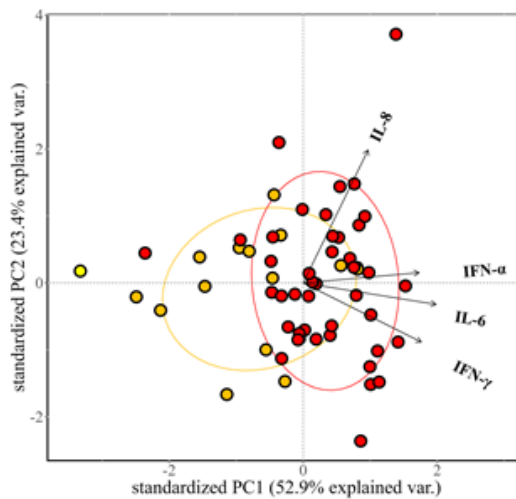
B



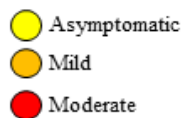
C



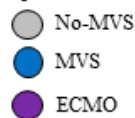
D



Non Hospitalized



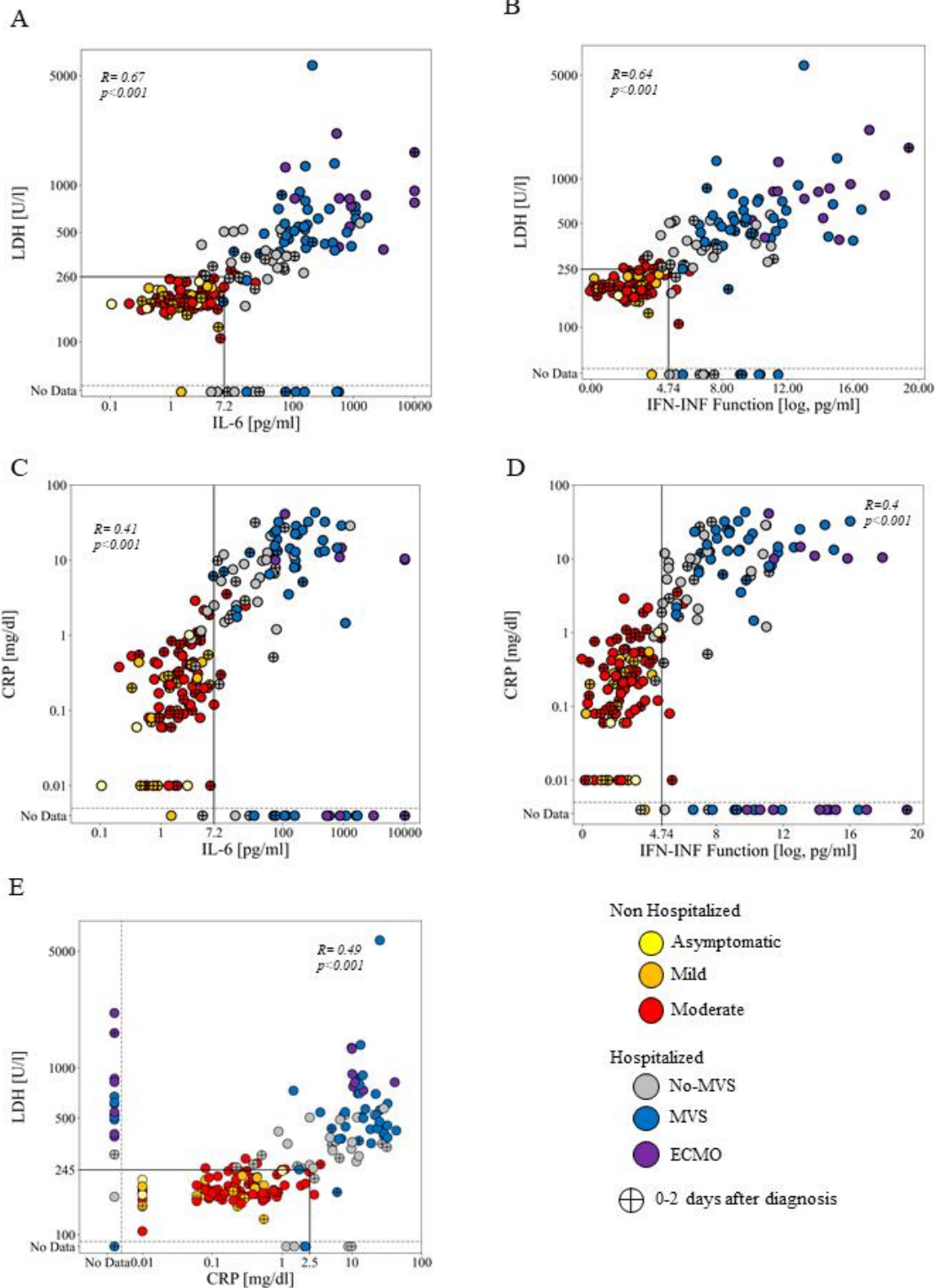
Hospitalized



Supplementary Figure 7: PCA analysis and dimension reduction for the prediction of hospitalization risk and moderate disease course. A PCA of all measured cytokines (A) in all patients ($DPSO \leq 11$) shows separation between hospitalized and non-hospitalized patients (PC factors in Supplementary Table 3). Dimension reduction indicates that the ratio between the inflammatory cytokines (IL-6, IL-10, and TNF- α) to IFN- α raw values (B) shows a significant separation between hospitalized and non-hospitalized patients (separatrix by solid

line in B), thus leading to the formalization of the IFN-INF function (Figure 1). In contrast, PCA of all measured cytokines in only non-hospitalized patients with $DPSO \leq 6$ (PC factors in Supplementary Table 3), shows no clear separation between mild and moderate severity (C). Dimension reduction to 4 cytokines (IL-6, IL-8, IFN- α , IFN- γ) (D), shows a better separation between the mild and moderate severity groups, thus leading to the formalization of the IMM-INF function (Figure 2).

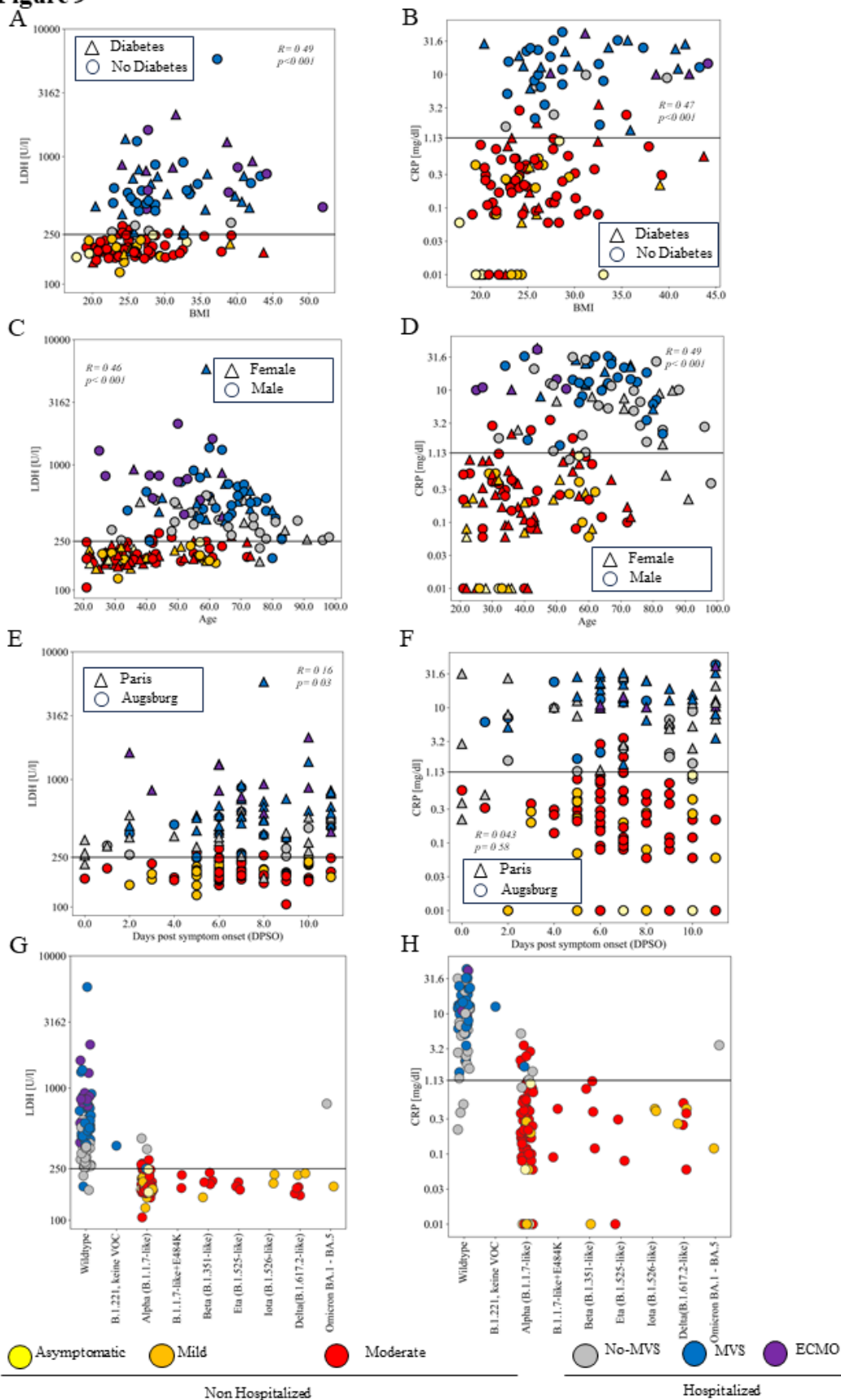
eFigure 8



Supplementary Figure 8: Combinations of LDH, CRP, or cytokines biomarkers for prediction of hospitalization risk. The levels of LDH, CRP, IL-6 and the IFN-INF function at the first visit per patient (DPSO $\leq$ 11) are all significantly inter-correlated. Nevertheless, LDH

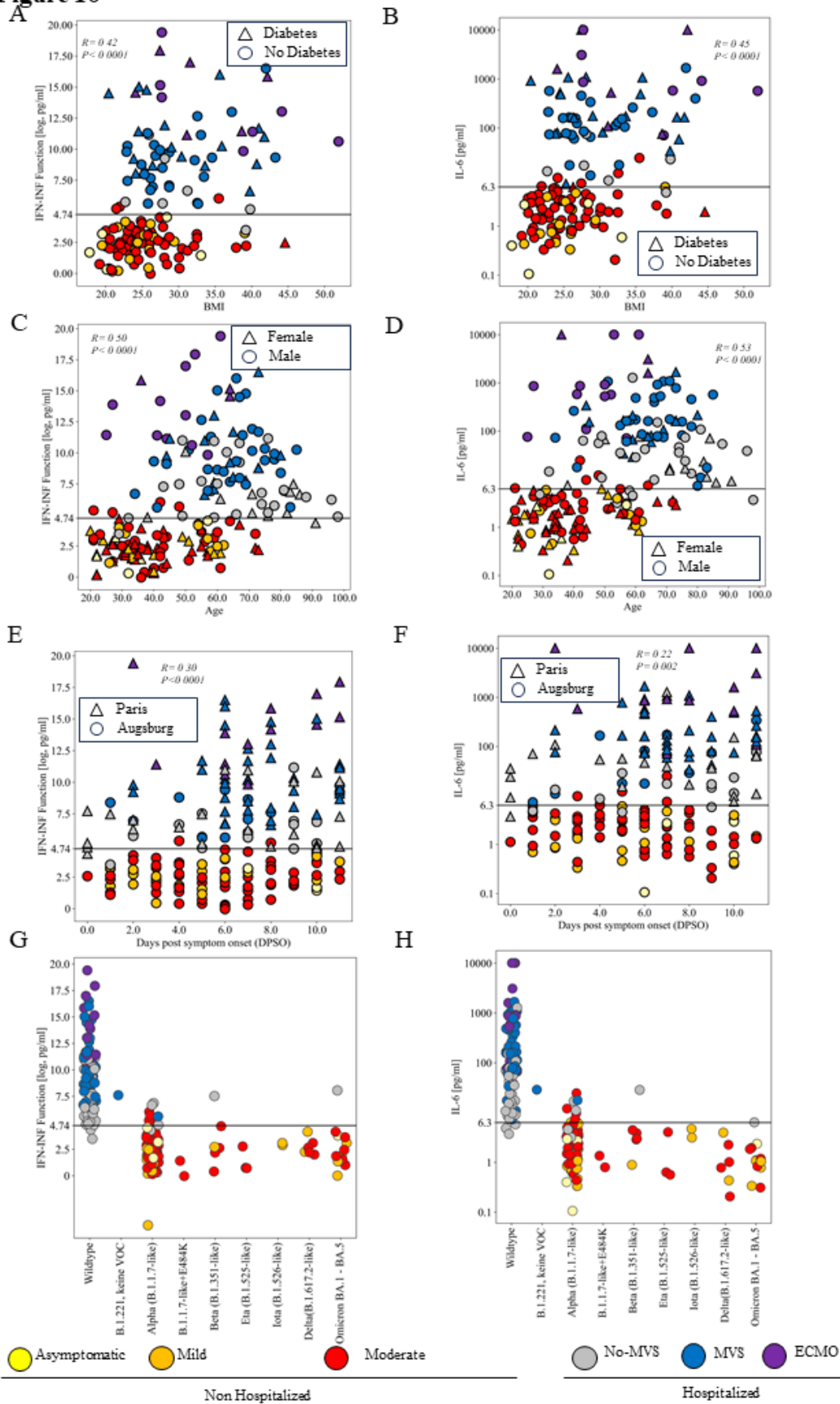
can be combined with IL-6 (A), IFN-INF Function (B) or CRP (E) to give more accurate prediction than each biomarker alone (J-Index >0.9, see Table 2). A few patients are missing LDH values (points below the horizontal dotted line, Supplementary Table 2), but these would not significantly change the prediction. However, in addition, more patients were lacking also CRP data (points to the right or below dotted line, Supplementary Table 2), and therefore the LDH+CRP combination is less robust. The combination of CRP with IL-6 (C) or with IFN-INF function (D) does not add predictive value to IL-6 or IFN-INF alone. The solid black lines show the thresholds used for prediction (with largest J-index). The dotted grey lines depict the points where data was missing for LDH and/or CRP (Supplementary Table 2).

eFigure 9



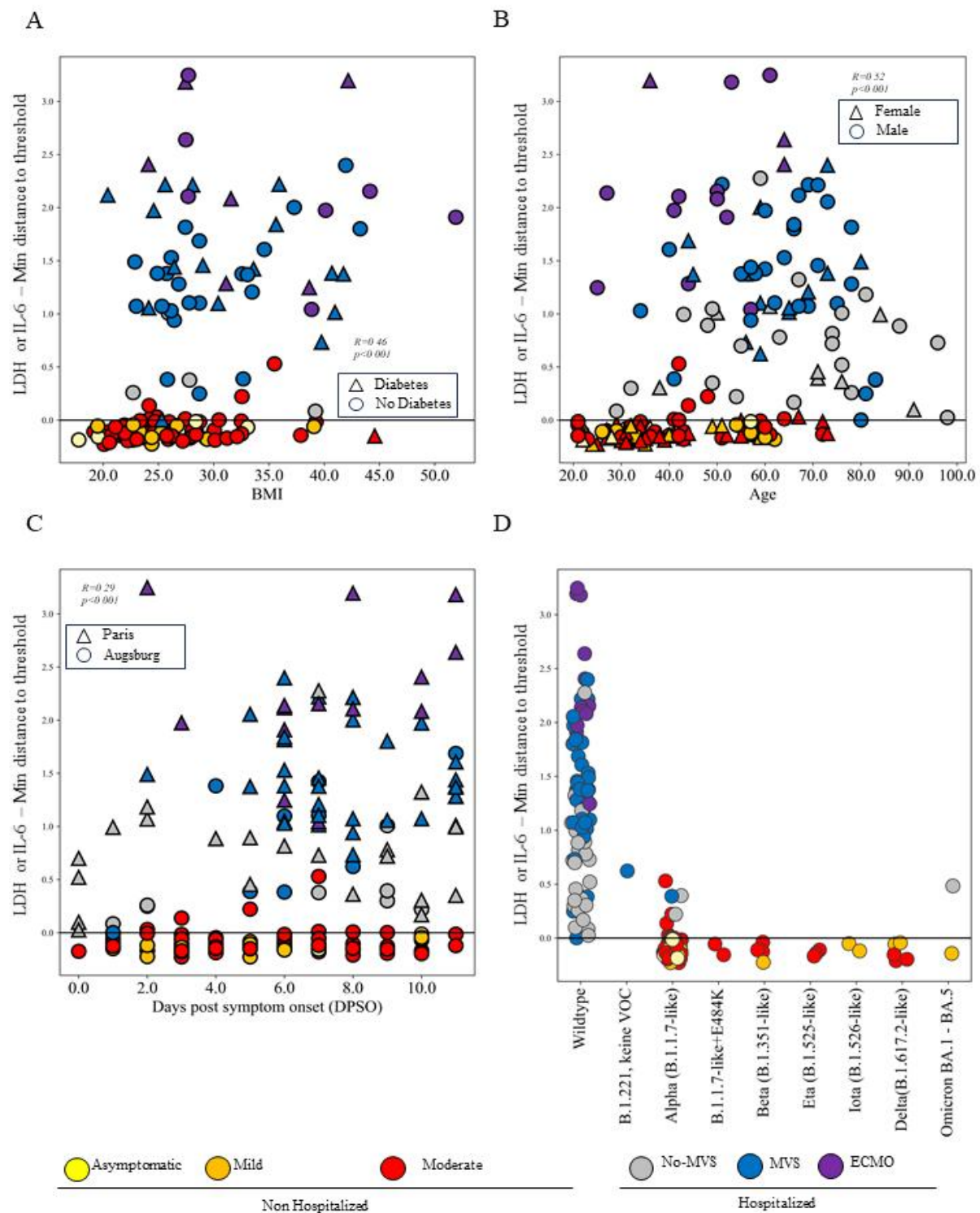
Supplementary Figure 9: Independence of the LDH and CRP biomarkers from different known severity risk factors for the prediction of hospitalization risk. Patients with high BMI (A, B), having diabetes (A, B), are of high age (C, D), and males (C, D) are known for having a higher risk for more severe COVID-19, and also in this study show an association with hospitalization risk. Nevertheless, the LDH and CRP biomarkers, measured at the first visit per patient ($DPSO \leq 11$), are significant and accurate predictors of hospitalization risk independently of those different risk factors, albeit being significantly associated with them (A-D). The predictive value of LDH and CRP is independent of study site (E, F) and of the time the first sample was taken within 0-11 days post symptom onset (E, F). Also, the predictive value of LDH and CRP shows a trend to be independent of SARS-CoV-2 variant (G, H). The horizontal solid lines depict the threshold used for prediction for each biomarker.

eFigure 10



Supplementary Figure 10: Independence of the IFN-INF and IL-6 biomarkers from different known severity risk factors for the prediction of hospitalization risk. The IFN-INF function (A, C) and IL-6 (B, D) biomarkers, measured at the first visit per patient ($DPSO \leq 11$), are significant and accurate predictors of hospitalization risk independently of different risk factors known for having a higher risk for more severe COVID-19, such as: high BMI (A, B), having diabetes (A, B), high age (C, D), or males (C, D). The predictive value of IFN-INF and IL-6 is independent of study site (E, F) and of the time the first sample was taken within 0-11 days post symptom onset (E, F). Also, the predictive value of IFN-INF and IL-6 shows a trend to be independent of SARS-CoV-2 variant (G, H). The horizontal solid lines depict the threshold used for prediction for each biomarker.

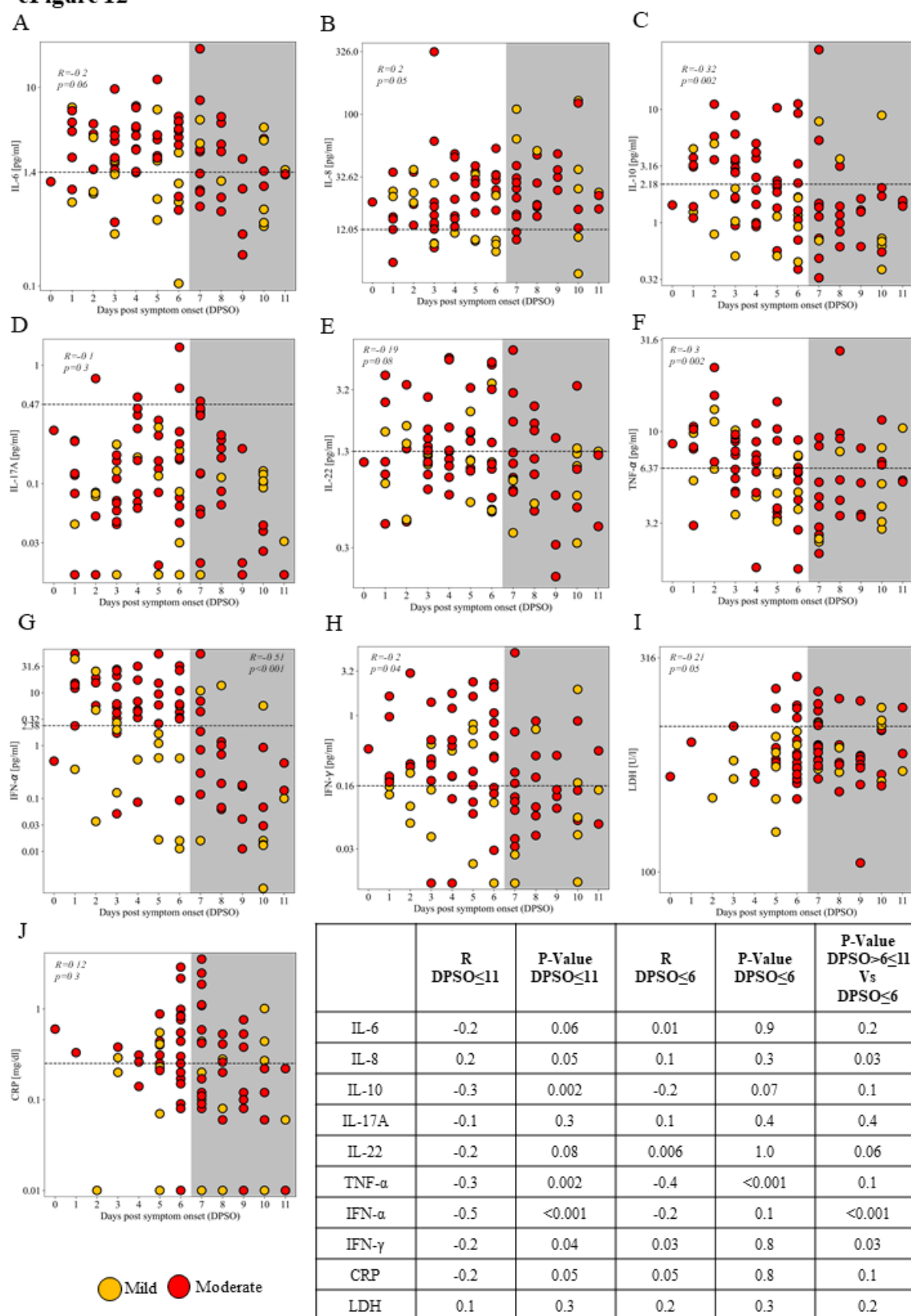
eFigure 11



Supplementary Figure 11: Independence of the LDH-OR-IL-6 prediction function from different known severity risk factors for the prediction of hospitalization risk. The LDH-OR-IL-6 prediction function, measured at the first visit per patient (DPSO≤11), is a significant and accurate predictor of hospitalization risk independently of different risk factors known for

having a higher risk for more severe COVID-19, such as: high BMI (A), having diabetes (A), high age (B), or males (B). The predictive value of the LDH-OR-IL-6 prediction function is independent of study site (C) and of the time the first sample was taken within 0-11 days post symptom onset (C). Also, the predictive value of the LDH-OR-IL-6 function shows a trend to be independent of SARS-CoV-2 variant (D). On Y-axis is depicted the minimum distance between the (LDH,IL-6) coordinates and the corresponding thresholds, thus the horizontal line at 0.0 separates the prediction of hospitalized versus non-hospitalized.

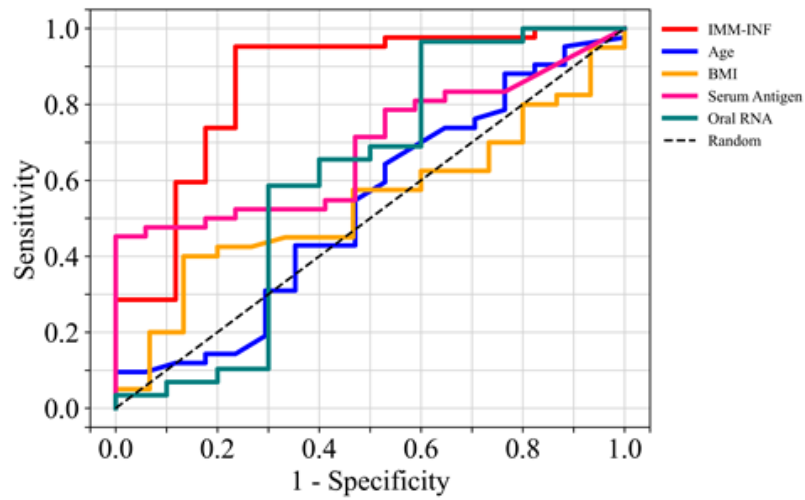
eFigure 12



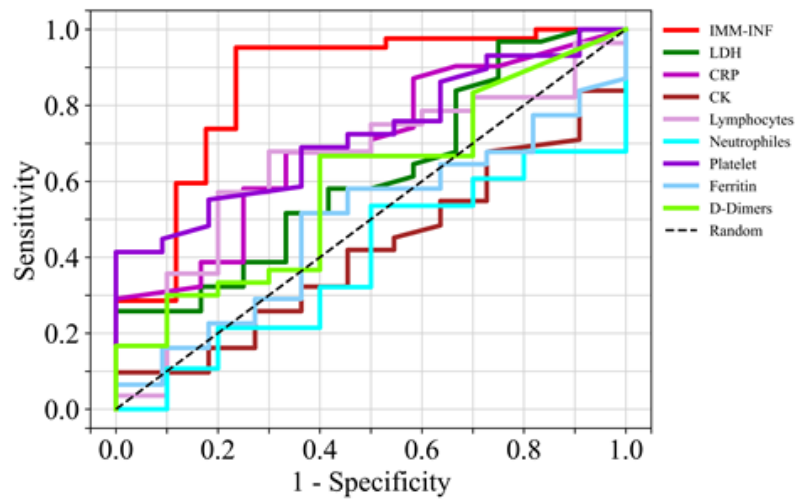
Supplementary Figure 12: Decline of serum cytokines and biochemistry levels over time in non-hospitalized patients. The correlation of serum levels of IL-6 (A), IL-8 (B), IL-10 (C), IL-17A (D), IL-22 (E), TNF- α (F), IFN- α (G), IFN- γ (H), LDH (I) and CRP (J), as measured at the first sample available in non-hospitalized patients, with time (0-11 days post symptom onset), indicates that for some cytokines (e.g., IFN- α and IFN- γ) there is a strong decline over time. Mann-Whitney testing shows that the time threshold where differences are most significant are between 0-6 and 7-11 DPSO, and these correlations are not significant within the period of 0-6 DPSO (inset table). Thus, for the prediction of moderate versus mild disease in non-hospitalized patients only patients with the first sample available at 0-6 DPSO were taken, excluding those with a first sample at 7-11 DPSO (grey boxes).

eFigure 13

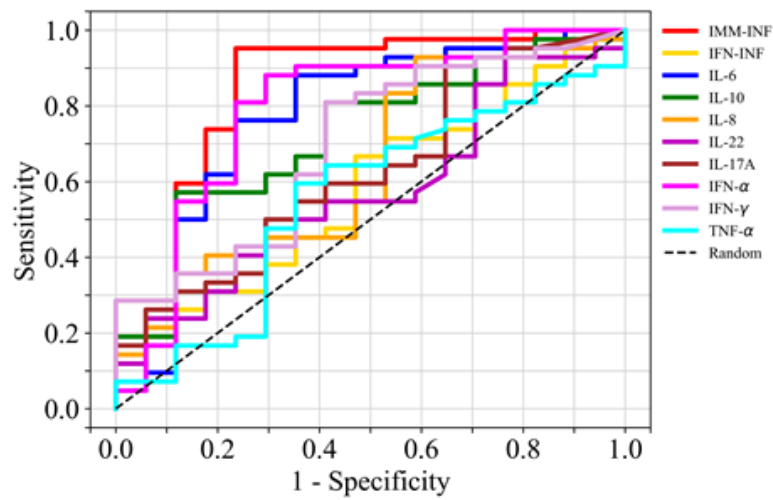
A



B

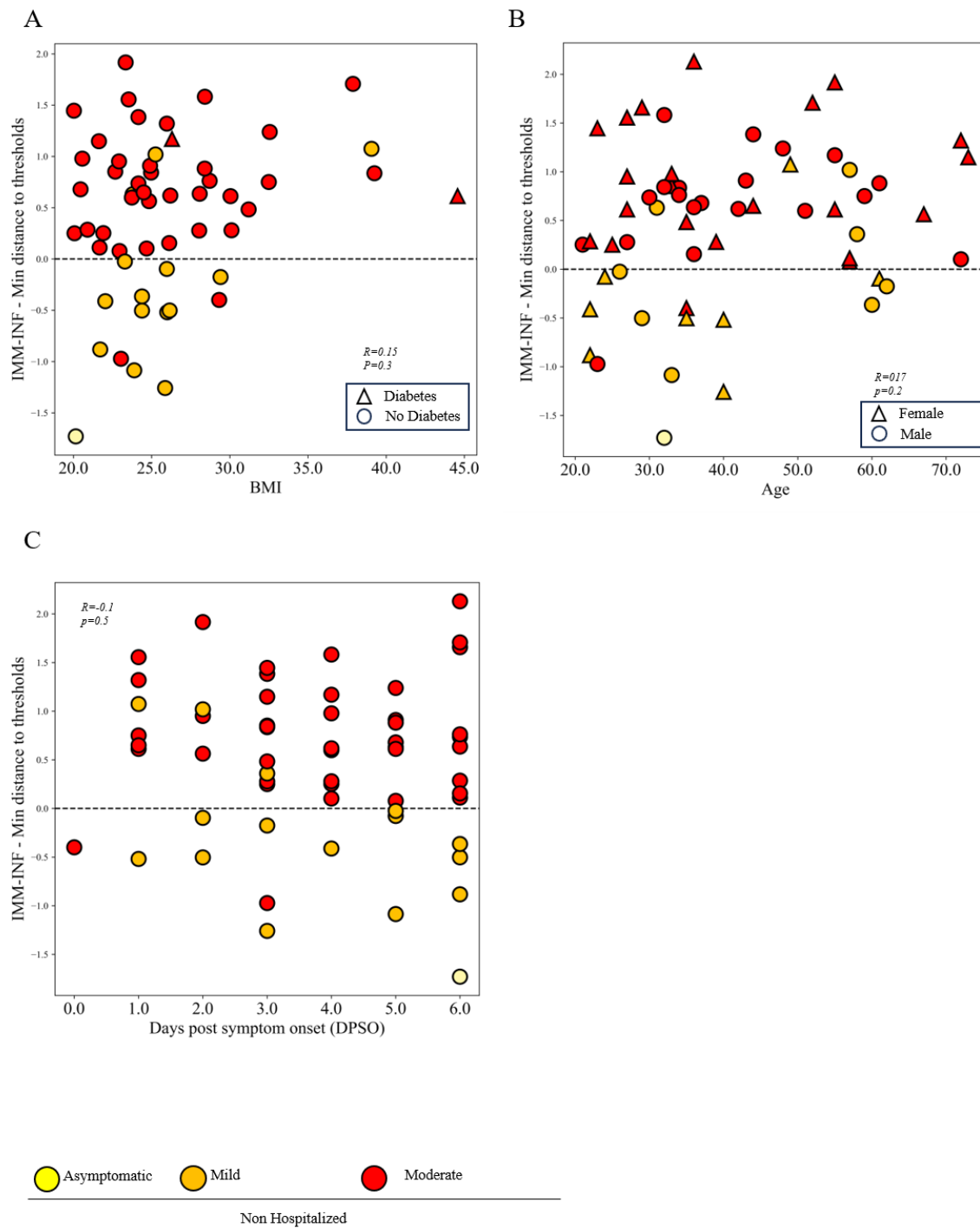


C



Supplementary Figure 13: ROC analysis comparison of the predictive values of age, BMI, SARS-CoV-2 N-antigen, oropharyngeal SARS-CoV-2 RNA, and all measured biochemistry, hematology and cytokine biomarkers for the prediction of moderate versus mild symptom severity. Receiver Operating Characteristic (ROC) curves for age, BMI, oropharyngeal SARS-CoV-2 RNA, and serum N-antigen (A) in comparison to the IMM-INF function show a significantly higher AUC for the IMM-INF function as predictor of moderate versus mild symptom severity. ROC curves for the prediction of moderate versus mild symptom severity using any of the measured biochemistry values (B) and any of the eight measured single cytokines as well as the IMM-INF and IFN-INF functions (C), measured at the first visit per patient ($DPSO \leq 6$). IFN-INF function is defined by: $\log(2 * IL-6 * IL-10 * TNF-\alpha / IFN-\alpha)$. IMM-INF function is defined by: $(\log(IL-6) + \log(IL-8) > 1.75)$ OR $(\log(IFN-\alpha) + \log(IFN-\gamma) > -0.25)$.

eFigure 14



Supplementary Figure 14: Independence of the IMM-INF prediction function from different known severity risk factors for the prediction of moderate versus mild disease.

The IMM-INF prediction function, measured in non-hospitalized patients at the first visit per patient (DPSO $\leq$ 6), is a significant and accurate predictor of moderate versus mild disease independently of different risk factors known for having a higher risk for more severe COVID-

19, such as: high BMI (A), having diabetes (A), high age (B), or males (B). The predictive value of the IMM-INF prediction function is independent of study site (C) and of the time the first sample was taken within 0-6 days post symptom onset (C). On Y-axis is depicted the minimum distance between the (X,Y) coordinates of the IMM-INF function for each patient and the corresponding thresholds, thus the dashed line at 0.0 separates the prediction of moderate versus mild.