Clinical Infectious Diseases

MAJOR ARTICLE

Severe Dysbiosis and Specific *Haemophilus* and *Neisseria* Signatures as Hallmarks of the Oropharyngeal Microbiome in Critically Ill Coronavirus Disease 2019 (COVID-19) Patients

Juliana de Castilhos,[1,22,a] Eli Zamir,[1,a] Theresa Hippchen,[2,a] Roman Rohrbach,[1,a] Sabine Schmidt,[1] Silvana Hengler,[1] Hanna Schumacher,[1] Melanie Neubauer,[3] Sabrina Kunz,[4] Tonia Müller-Esch,[4] Andreas Hiergeist,[5] André Gessner,[5] Dina Khalid,[6] Rogier Gaiser,[1] Nyssa Cullin,[1] Stamatia M. Papagiannarou,[1] Bettina Beuthien-Baumann,[7] Alwin Krämer,[8] Ralf Bartenschlager,[9,10] Dirk Jäger,[11] Michael Müller,[12] Felix Herth,[12] Daniel Duerschmied,[13] Jochen Schneider,[14] Roland M. Schmid,[14] Johann F. Eberhardt,[15] Yascha Khodamoradi,[15] Maria J. G. T. Vehreschild,[15,16] Andreas Teufel,[17] Matthias P. Ebert,[17] Peter Hau,[18] Bernd Salzberger,[19] Paul Schnitzler,[6] Hendrik Poeck,[4,20] Eran Elinav,[1,21,a] Uta Merle,[2,a] and Christoph K. Stein-Thoeringer[1,11,a]

[1]German Cancer Research Center (DKFZ), Research Division Microbiome and Cancer, Heidelberg, Germany; [2]Department of Gastroenterology and Infectious Diseases, University Clinic Heidelberg, Heidelberg, Germany; [3]Department of Medicine II, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany; [4]Department of Internal Medicine III, University Clinic Regensburg, Regensburg, Germany; [5]Institute of Clinical Microbiology and Hygiene, University Clinic Regensburg, Regensburg, Germany; [6]Department of Virology, University Clinic Heidelberg, Heidelberg, Germany; [7]German Cancer Research Center (DKFZ), Research Division Radiology, Heidelberg, Germany; [8]German Cancer Research Center (DKFZ), Research Division Molecular Hematology/Oncology, Heidelberg, Germany; [9]Department of Infectious Diseases, Molecular Virology, Heidelberg University, Heidelberg, Germany; [10]German Cancer Research Center (DKFZ), Research Division Virus-associated Carcinogenesis, Heidelberg; [11]National Center for Tumor Diseases (NCT) Heidelberg, Heidelberg, Germany; [12]Thoraxklinik and Translational Lung Research Center, Heidelberg University, Heidelberg, Germany; [13]Department of Internal Medicine III, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany; [14]Department of Internal Medicine II, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany; [15]Department of Internal Medicine, Infectious Diseases, University Hospital Frankfurt, Frankfurt, Germany; [16]German Center for Infection Research (DZIF), Partner Site Bonn-Cologne, Cologne, Germany; [17]Department of Medicine II, Section of Hepatology, University Medical Center Mannheim, University of Heidelberg, Mannheim, and Center for Preventive Medicine and Digital Health Baden-Württemberg, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany; [18]Wilhelm Sander-NeuroOncology Unit and Department of Neurology, University Clinic Regensburg, Regensburg, Germany; [19]Department of Infectious Disease, University Clinic Regensburg, Regensburg, Germany; [20]National Center for Tumor Diseases (NCT) WERA, [21]Weizmann Institute of Science, Rehovot, Israel; and [22]Vale do Rio dos Sinos University (UNISINOS), Sao Leopoldo, Brazil

Background. At the entry site of respiratory virus infections, the oropharyngeal microbiome has been proposed as a major hub integrating viral and host immune signals. Early studies suggested that infections with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) are associated with changes of the upper and lower airway microbiome, and that specific microbial signatures may predict coronavirus disease 2019 (COVID-19) illness. However, the results are not conclusive, as critical illness can drastically alter a patient's microbiome through multiple confounders.

Methods. To study oropharyngeal microbiome profiles in SARS-CoV-2 infection, clinical confounders, and prediction models in COVID-19, we performed a multicenter, cross-sectional clinical study analyzing oropharyngeal microbial metagenomes in healthy adults, patients with non-SARS-CoV-2 infections, or with mild, moderate, and severe COVID-19 (n = 322 participants).

Results. In contrast to mild infections, patients admitted to a hospital with moderate or severe COVID-19 showed dysbiotic microbial configurations, which were significantly pronounced in patients treated with broad-spectrum antibiotics, receiving invasive mechanical ventilation, or when sampling was performed during prolonged hospitalization. In contrast, specimens collected early after admission allowed us to segregate microbiome features predictive of hospital COVID-19 mortality utilizing machine learning models. Taxonomic signatures were found to perform better than models utilizing clinical variables with *Neisseria* and *Haemophilus* species abundances as most important features.

Conclusions. In addition to the infection per se, several factors shape the oropharyngeal microbiome of severely affected COVID-19 patients and deserve consideration in the interpretation of the role of the microbiome in severe COVID-19. Nevertheless, we were able to extract microbial features that can help to predict clinical outcomes.

Keywords. SARS-CoV-2; COVID-19; microbiome; dysbiosis; machine learning.

The microbiome constitutes a signaling hub regulating host immunity, mucosal homeostasis, and defense against pathogens [1]. The oropharyngeal and lung microbiome has been previously studied during respiratory syncytial virus (RSV) and influenza respiratory tract infection [2, 3]. At these body sites, the microbiome was found to modulate viral infections through adaptions of mucosal immunity or inhibition of mucosal virus adherence [4], which eventually affect the

Received 19 July 2021; editorial decision 4 October 2021; published online 25 October 2021.
[a]J. d. C., E. Z., T. H., R. R., E. E., U. M., and C. K. S.-T. contributed equally to this work.
Correspondence: C. K. Stein-Thoeringer, Microbiome and Cancer Research Division, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 242, 69120 Heidelberg, Germany (c.stein-thoeringer@dkfz.de).

Clinical Infectious Diseases® 2021;XX(XX):1–9

https://doi.org/10.1093/cid/ciab902

clinical course of these infections [5]. Viruses can also facilitate pathogen colonization, eventually leading to bacterial superinfections and acute respiratory distress syndrome [6, 7]. A few studies have already explored associations of the nasal, oropharyngeal or lung microbiome with severity of coronavirus disease 2019 (COVID-19) or its role as a prognostic biomarker with so far inconclusive results [8–13]. In particular, most of these studies investigated microbiome profiles of patients with severe COVID-19 that were admitted to a hospital or even an intensive care unit (ICU) for emergency medical treatment and not mildly or moderately affected cases. Upper versus lower respiratory tract sampling, collection timepoint after admission, or sampling while on antibiotics or during invasive mechanical ventilation are major factors that can determine diversity and composition of the microbiome at these sites [14, 15]. Considering that treatment- and patient-related factors can contribute to extreme dysbiosis and expansion of pathobionts in patients with critical illness, confounder analyses are mandatory in these studies to elucidate associations beyond noise and to identify potential biomarkers for disease outcomes.

METHODS

Study Cohort Description

Recruitment

Seven German medical centers participated in the recruitment of inpatient subjects (both sexes, 21–93 years of age) with laboratory confirmed COVID-19 (n = 102): University Clinic Heidelberg (Thoraxklinik and Department of Gastroenterology and Infectious Diseases), University Hospital Mannheim, University Clinic Regensburg, University Clinic Frankfurt, Klinikum rechts der Isar of Technical University Munich, and University Heart Center Freiburg. The oropharyngeal specimens of inpatients were collected prospectively between 30 March 2020 and 14 November 2020.

Recruitment of outpatient subjects (n = 148, both sexes, 18–86 years of age) was performed in COVID-19 test centers run by the Heidelberg public health department. Testing was initiated by health authorities because of emerging, new influenza-like symptoms, individual contact with SARS-CoV-2 infected people or travel to COVID-19 outbreak areas. Adult healthy controls (n = 72, both sexes, 23–70 years of age; no SARS-CoV-2 infection, no symptoms for upper respiratory tract infections) were recruited among volunteers from DKFZ, University Clinic Heidelberg and the National Center for Tumor Diseases Heidelberg. After quality control of metagenomic sequencing reads including rarefaction (for details see Supplemental Information), samples from 24 patients were excluded from the analysis. Only data from study subjects with samples included in the final analysis (n = 322) are presented in the article.

Participant Metadata

Epidemiological, clinical, laboratory, treatment, and outcome data of inpatient study participants were extracted from medical records using a standardized version of the World Health Organization (WHO) case record form for severe acute respiratory infections [16]. Baseline epidemiologic characteristics, symptoms of infection and antibiotic use were captured from healthy volunteers and outpatient participants.

Assessment of COVID-19 severity according to the WHO guidance document for clinical management of COVID-19 was performed (27 May 2020). Symptomatic patients who were tested positive for SARS-CoV-2 infection without evidence of pneumonia or hypoxia were categorized as mild COVID-19. These were largely patients that were screened for the infection in the outpatient test center. Patients admitted to the hospital with signs of pneumonia but no signs of severe pneumonia (including SpO_2 ≥ 93% on room air) and tested positive for SARS-CoV-2 were considered to have a moderate disease. Severe COVID-19 was assigned to cases with severe pneumonia (respiratory rate > 30 breaths/minute; or SpO_2 < 93% on room air), acute respiratory distress syndrome, sepsis, or septic shock. For 6 patients in the severe COVID-19 group that received mechanical ventilation, we do not have gender, age, or clinical outcome information. Clinical outcomes were monitored until 22 December 2020.

Specimen Collection, Processing, Metagenome Sequencing, and Analysis

Oropharyngeal samples for microbiome analyses were collected cross-sectionally primarily around admission or in a few cases at later time points (see Results section). Laboratory confirmation of SARS-CoV-2 was performed by reverse transcription polymerase chain reaction (RT-PCR) in certified diagnostic departments of the participating university hospitals or in our laboratory in accordance with the protocol established by the WHO. Shotgun metagenome sequencing was carried out on DNA extracted from oropharyngeal biospecimen following standard Illumina workflows of our laboratory. Metagenomic data processing was performed after demultiplexing using Kraken 2/ Bracken for microbial taxonomy profiling and HUMAnN 3.0 for profiling microbial genes and metabolic pathways.

A detailed description of the methods is available online as Supplemental Materials.

RESULTS

To study the role of the human oropharyngeal microbiome in SARS-CoV-2 infection and illness, we collected oropharyngeal biospecimens from a multicenter patient and healthy volunteer cohort in Germany between March and November 2020 (n = 322). We recruited healthy adult individuals, patients with mild, moderate, or severe COVID-19, and SARS-CoV-2 negative patients with mild upper respiratory infections or with critical pneumonia. Demographic and clinical characteristics are reported in Table 1

Table 1. Baseline Demographic Characteristics of Healthy Subjects, Individuals With an Upper Respiratory Tract (URT) Infection, or With COVID-19

	Healthy Individuals	URT Infection	Mild COVID-19	Moderate COVID-19	Severe COVID-19
Subjects	n = 72	n = 112	n = 36	n = 37	n = 65
SARS-CoV-2 status	Negative	Negative	Positive	Positive	Positive
Age					
Median (IQR), y	36.0 (30.0–53.0)	46.0 (30.9–56.7)	50.0 (37.9–54.8)	57.0 (49.0–70.4)	65.0 (55.0–73.1)
Distribution, no. (%)					
18–49 y	54 (73.0)	64 (57.1)	17 (47.2)	10 (27.0)	9 (13.8)
50–64 y	19 (25.7)	36 (32.1)	17 (47.2)	14 (37.8)	22 (33.9)
≥ 65 y	1 (1.4)	12 (10.7)	2 (5.6)	13 (35.1)	34 (52.3)
Gender, no. (%)					
Female	48 (64.9)	75 (67.0)	24 (66.7)	11 (29.7)	15 (23.1)
Male	26 (35.1)	37 (33.0)	12 (33.3)	26 (70.3)	50 (76.9)

Significant effect of age across all groups is shown (Kruskal-Wallis test: $P < .0001$).

Abbreviations: COVID-19, coronavirus disease 2019; IQR, interquartile range; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

and the Methods section. Mild COVID-19 was assigned to patients presenting with symptoms of upper respiratory tract infections that were not hospitalized (Supplementary Table 1). Patients with moderate or severe disease were hospitalized with lower respiratory tract symptoms. In our cohort, 65 cases with severe COVID-19 required intensive care unit (ICU) therapy, and 42 out of them deteriorated over the course of hospitalization and required invasive mechanical ventilation or extracorporeal membrane oxygenation (ECMO); 22 patients (21.6%) died by the conclusion of our study. Supplementary Table 2 presents preexisting conditions and medical treatments of the hospitalized patients of our study cohort.

To characterize the oropharyngeal microbiome profiles in our cohort, we used shotgun metagenome sequencing on oropharyngeal samples from all study subjects that were collected either during SARS-CoV-2 screening of suspected cases or during/after admission of patients with confirmed or suspected COVID-19. The number of metagenomic reads obtained from the oropharyngeal biospecimens and the fractions of human, bacterial and viral reads are listed in Supplementary Table 3 (Supplementary Table 4 for complete list of individual bacterial reads). Analyzing alpha diversity using the Shannon index, hospitalized COVID-19 patients showed a lower bacterial diversity (Figure 1A), but the differences were not significant after false discovery rate (FDR) adjustment of *P*-values (mild vs moderate: $P = .73$; moderate vs severe: $P = .08$). The Simpson alpha diversity index, which also weighs taxonomic abundances, did not differ significantly between groups (Supplementary Figure 1). Richness and Chao index, a richness estimator, were found reduced in the microbiome of hospitalized patients (Supplementary Figure 1A). Differences in microbiome compositions between each group were most apparent in a principal coordinate analysis (PCoA) with significant separations according to COVID-19 severity (PERMANOVA: $F = 10.6$, $P < .001$; Figure 1B). We further classified samples with taxonomic compositions highly unlike those of healthy controls as "dysbiotic"[15] as measured by Bray-Curtis beta-diversity distances.[16] We observed that patients with moderate and severe COVID-19 showed a significantly higher dysbiosis index of the oropharyngeal microbiome compared to healthy controls, URT patients, and mild COVID-19 patients, respectively (Figure 1C). We further asked whether other patient anthropometrics and clinical variables contribute to the variation of the oropharyngeal microbiome than COVID-19 severity. As shown in Figure 1D, antibiotic exposure and severity of the infection explained most variance in the taxonomy and to only a minor extend to the microbial gene repertoire or metabolic pathways encoded in the microbiome. No differences in the SARS-CoV-2 virus load in the upper respiratory tract of patients with moderate or severe disease or with low vs high microbial dysbiosis were observed in our cohort (Supplementary Figure 1B).

Hospitalized COVID-19 patients, notably severe cases, are frequently treated with antibiotics [17, 18]. In our inpatient cohort, 79.4% of all patients received broad-spectrum antibiotic treatments (Supplementary Table 2), mainly piperacillin/tazobactam (72.8%), carbapenems (21.0%), ampicillin/sulbactam and amoxicillin/clavulanic acid (9.9%), or macrolides (8.6% of all antibiotics administered). Administration of these antibiotics during oropharyngeal sample collection significantly affected the compositional diversity of the microbiome (Figure 2A and Supplementary Figure 2A; PERMANOVA: $F = 25.5$, $P < .001$). This was reflected by a significantly higher dysbiosis index of samples collected during antibiotic exposure in patients with severe COVID-19 (Figure 2B), with similar trends in patients with moderate disease ($P = .22$). The alpha diversity was not affected by antibiotic administration in hospitalized patients (Figure 2B). The time spent in an ICU has also been reported to affect the structure of the oral and upper respiratory tract microbiome with longer ICU stays being associated with lower diversity [19]. In our study, most of the oropharyngeal samples were collected at or early after hospital admission, and only a few swabs were taken at later timepoints

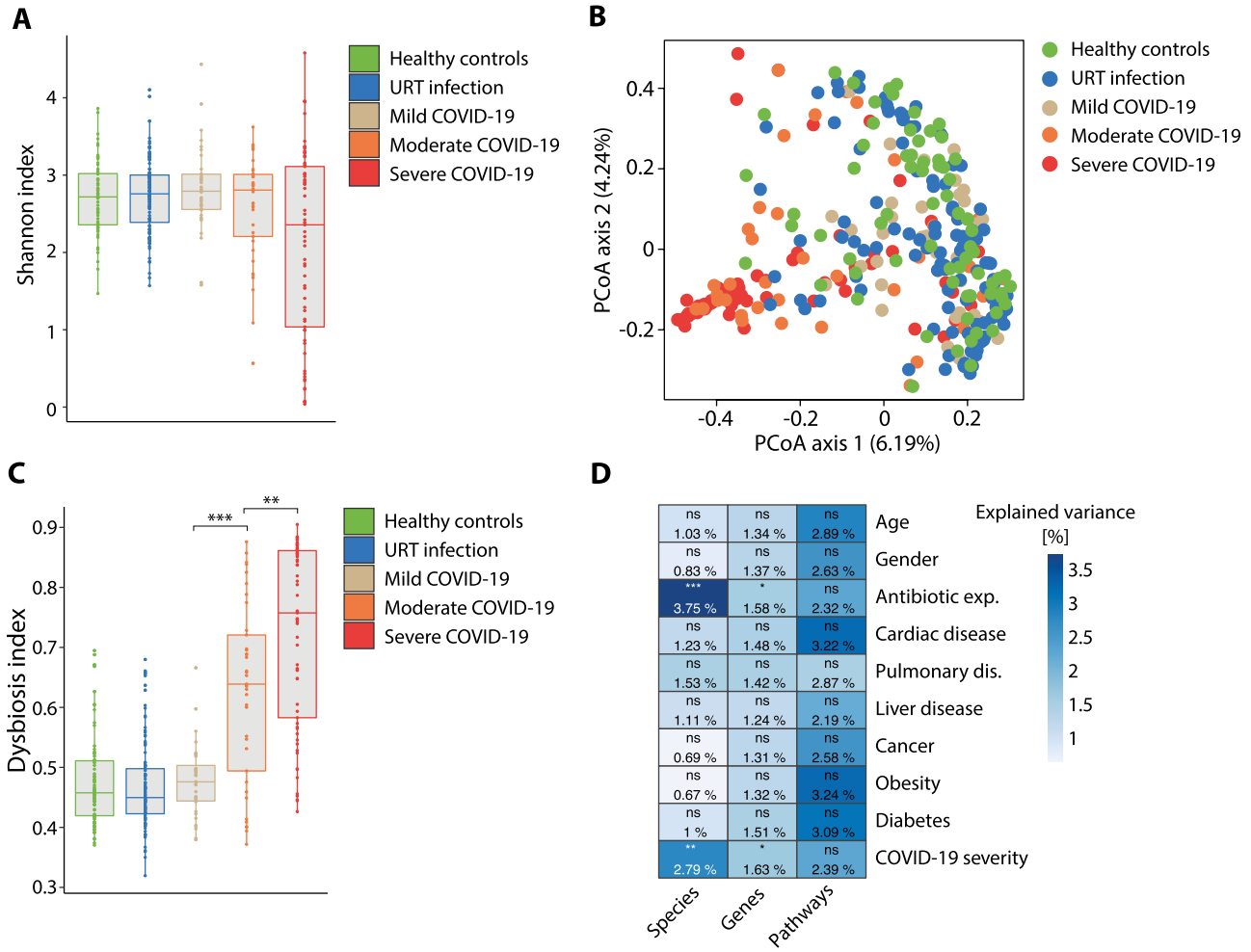

Figure 1. Alpha and beta diversity metrices of the oropharyngeal microbiome. *A*, Alpha diversity displayed as Shannon indexes for the oropharyngeal microbiomes sampled from healthy controls (n = 72), patients with SARS-CoV-2 negative URT infections (n = 112), and SARS-CoV-2 positive patients with mild, moderate, or severe COVID-19 (n = 36, 37, and 65). *B*, PCoA based on species-level Bray-Curtis dissimilarity for microbiome samples collected from controls and patients with percent of variance explained per PCoA axis. *C*, Dysbiosis indexes defined per sample as its Euclidian distance to the centroid of the healthy controls' samples in the 5-dimensional PCoA-projection space of Bray-Curtis dissimilarities. *D*, Effects of patient-derived features or phenotypes are shown regarding inter-individual variation of taxonomy, genes and pathways encoded in the oropharyngeal microbiome (stars show FDR-corrected statistical significance: $*P < .05$, $**P < .01$, $***P < .001$). Variance is estimated for each feature independently. *A–C*: $**P < .01$, $***P < .001$ by Wilcoxon test after FDR correction for pairwise testing. Abbreviations: COVID-19, coronavirus disease 2019; FDR, false discovery rate; ns, not significant; PCoA, principal coordinates analyses; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; URT, upper respiratory tract.

(Supplementary Figure 2B). Stratifying samples collected early (day 0–3, n = 51), intermediate (day 4–14, n = 34) or late after hospital admission (n = 17) had a strong effect on the microbial community structure as observed in a PCoA (Figure 2C; PERMANOVA: F = 3.6, $P < .001$). A loss of diversity and an increased dysbiosis that is associated with the length of hospital stay was only observed in severe COVID-19 cases, that is, those admitted to an ICU, and only in samples collected at a late timepoint (Figure 2D). As this community disruption was seen only in severe COVID-19, we hypothesized that intensive medical care with intubation and mechanical ventilation affects the structure of the oropharyngeal microbiome. As presented in Figure 2E, mechanical ventilation at the time of specimen collection has a significant effect on the compositional diversity of the microbiome in hospitalized COVID-19 patients (PERMANOVA: F = 4.9, $P < .001$). Although trending toward lower levels, especially in samples from late collection timepoints, the alpha diversity was not significantly reduced in the oropharyngeal microbiome collected during invasive mechanical ventilation (Figure 2F). In contrast, samples taken at late timepoints and during invasive ventilation had a significantly increased dysbiosis score (Figure 2F). The factors age and center had only modest effects, whereas ICU admission a significant effect on the composition of the oropharyngeal microbiome in hospitalized COVID-19 patients (Supplementary Figure 2C–2E; PERMANOVA: age: F = 1.04, $P = .362$; center: F = 1.1, $P = .282$; ICU admission: F = 2.9, $P < .01$).

So far, we found that antibiotic exposure, sampling timepoint, intubation, and invasive ventilation determine structure and

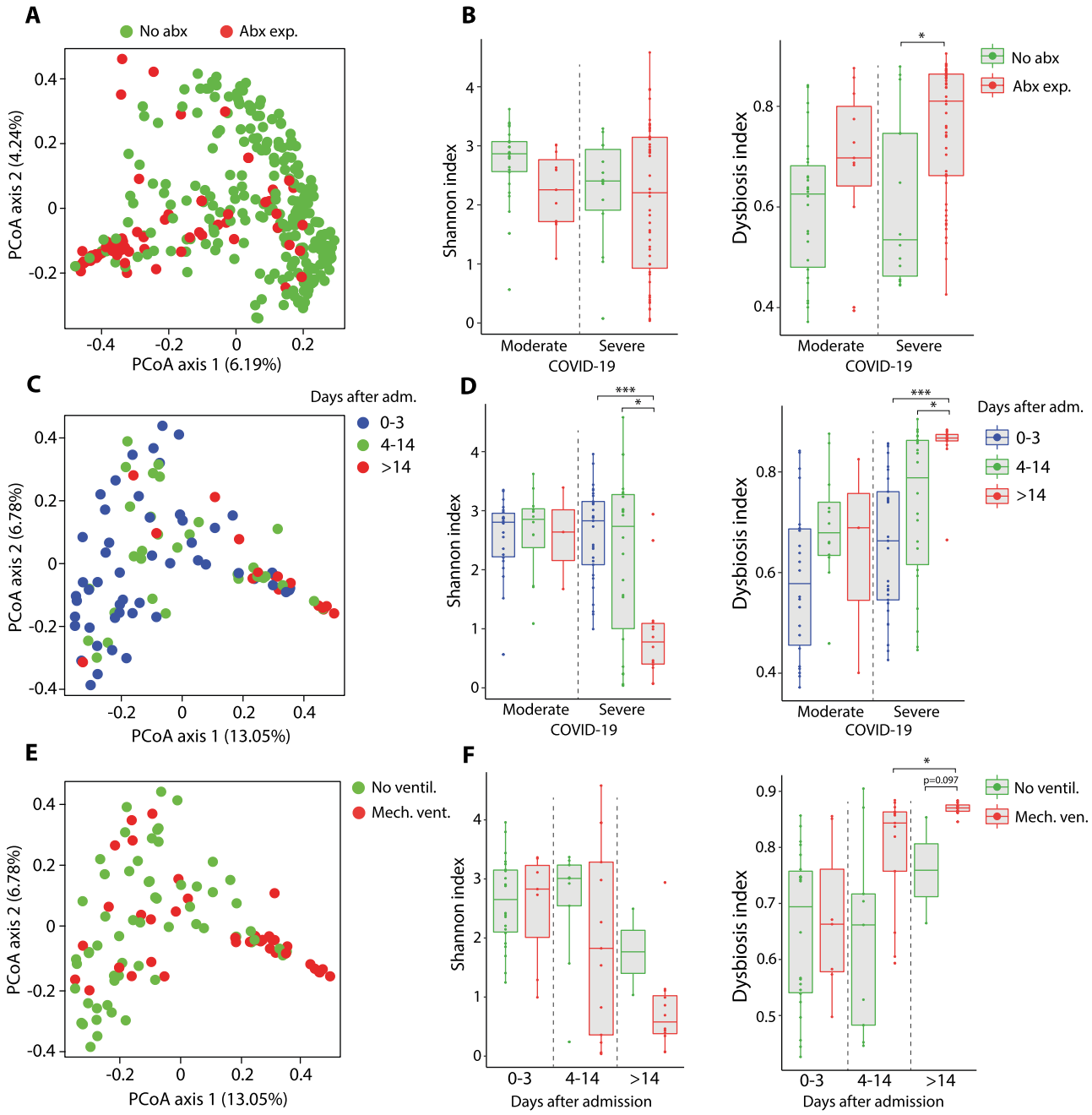

Figure 2. Clinical covariates affecting diversity and dysbiosis of the oropharyngeal microbiome in COVID-19 patients. *A*, PCoA based on species-level Bray-Curtis dissimilarity distinguishing samples collected while a patient was treated with antibiotics (Abx) or > 7 days off antibiotics (all subjects included). *B*, Shannon and dysbiosis index for specimen collected on vs off antibiotics in hospitalized patients. *C*, PCoA of samples from hospitalized patients collected between days 0–3 (n = 51), days 4–14 (n = 34) or > day 14 after admission (n = 17), and (*D*) presentation of Shannon and dysbiosis index. *E*, PCoA of samples from hospitalized patients collected while a patient received invasive mechanical ventilation, and (*F*) presentation of Shannon and dysbiosis index of stratified by ventilation and days after admission (only patients with severe COVID-19 included). **P* < .05, ****P* < .001 by Wilcoxon test after FDR correction for pairwise testing. Abbreviations: Abx. exp., antibiotic expression; adm., administration; COVID-19, coronavirus disease 2019; FDR, false discovery rate; Mech. ven., mechanical ventilator; PCoA, principal coordinates analyses; ventil., ventilator.

composition of the oropharyngeal microbiome. This is reflected by mono-domination patterns of individual bacterial species observed in patients with moderate and severe COVID-19, whereas a rather diverse ecology of microbes was observed in healthy controls, patients with non-SARS-CoV-2 URT, or mild COVID-19 (Supplementary Figure 3A). In order to circumvent these cases of extreme dysbiosis induced by late sampling that may confound clinical association studies, we focused on specimens collected early after hospital admission (day 0–3) for further species vs outcome analyses (n = 51 hospitalized patients; Supplementary Figure 3B: overview of the taxonomic composition). We aimed to determine whether specific taxa are

able to discriminate mortality as the major outcome of hospitalized COVID-19 patients. To address this question, we applied a machine learning (ML) analysis which provides a powerful approach for unsupervised biomarker discovery in microbiome research [20]. Here we utilized random forest (RF)-based ML models that are little prone to overfitting, can handle effectively high-dimensional data and heterogenous types of input features (e.g., integrating taxonomic and clinical variables as combined input) and provide interpretable quantification of the importance of each input feature for the performance of the model [21]. We aggregated the microbiome and clinical data of the subset of hospitalized patients and initially assessed how well the mortality in these patients can be predicted based on their species-level oropharyngeal microbiome. After partitions of the data into training and validation sets of patients, RF models were trained using the training sets and tested for their power to predict mortality in the validation sets (see Methods section in the Supplementary Materials for details). Microbiome features enabled the RF models to discriminate mortality, even with a better performance than RF models solely based on the clinical variables age, gender, obesity, and antibiotic treatment (Figure 3A and 3B). Given the power of the obtained RF models to discriminate outcomes based on the oropharyngeal microbiome, we next analyzed the importance of the bacterial species in these models and found that *Haemophilus influenzeae*, *parainfluenzae, pittmaniae,* and *Neisseria subflava* were among the most important species that segregated patients succumbed to COVID-19 versus survivors in our cohort (Figure 3C and 3D). We further assessed the contribution of high vs low abundances of these bacteria on survival as determined by splitting along the median of their relative abundances. A low abundance of all four species identified in the RF models significantly increased the risk for mortality using Cox proportional-hazards models, and *Neisseria subflava* showed the strongest effect (HR 16, 95% confidence interval [CI]: 2.0–121.6) (Figure 3E and Supplementary Figure 4A). The risk for mortality in patients with low *Neisseria subflava* abundance remained significantly elevated even after multivariate adjustment of the model for the covariates age, gender, and center (ie, hospital in which patients were treated) (Supplementary Figure 4B). Administration of antibiotics, or sampling during invasive ventilation did not affect high vs low *Neisseria subflava* abundance (Supplementary Figure 4C). The SARS-CoV-2 viral load was also not affected by high vs low *Neisseria subflava* abundance (Supplementary Figure 4D). We further analyzed the general architecture of the oropharyngeal microbiome in samples with high vs low *Neisseria* abundance, but we did not observe a significant effect on Shannon diversity or on the dysbiosis index (Figure 3F); however, there were a few positive and negative correlations with other species in the oropharyngeal microbiome as found in a network analysis (Supplementary Figure 4E).

Stratification of hospitalized patients into those with high vs low Shannon diversity or high vs low dysbiosis index of oropharyngeal microbiomes collected between days 0 and 3 after admission was not associated with survival differences (Shannon index: HR 0.91, 95% CI: .3–.7, $P = .87$; dysbiosis index: HR 0.96, 95% CI: .3–2.9, $P = .95$). Finally, we looked into secondary infections records of our hospitalized patients, and primarily *Streptococcus* and *Staphylococcus spp.* were diagnosed in microbiological tests of tracheal secretes or bronchoalveolar lavages (Supplementary Table 4). These species, however, were not found to dominate the oropharyngeal microbiome, and this discrepancy can be explained by site-specificity of potentially pathogenic microbes and methodological differences.

DISCUSSION

COVID-19 severity has been linked to several immunological changes such as lymphocytopenia [22], disinhibited release of proinflammatory cytokines [16], hyperactivation of monocytes and macrophages [23], dampened type-I interferon (IFN) signaling [24], or a hypercoagulable state [25]. Given that secondary bacterial infections are associated with increased risk for mortality in viral respiratory infections, the upper and lower respiratory tract microbiome has been explored in COVID-19 patients, first in case series and later in more systematic, but mono-centric clinical studies with overall varying and inconsistent results [11, 26–28].

In our study, we asked whether a SARS-CoV2 infections and its severity induced changes in diversity measures of the microbiome at the entry site of infection. We observed that mild illness or infections in outpatients was not associated with significant changes in diversity or in taxonomic composition compared to healthy controls or patients with other viral upper respiratory tract infections. Similar observations were previously published in an Italian cohort of mildly affected COVID-19 cases [28]. However, patients admitted to the hospital, and notably to an ICU, presented with reduced alpha diversity and increased dysbiosis of the oropharyngeal microbiome. In a recent study of patients admitted to an ICU for various medical conditions, McDonald et al longitudinally explored the oral and gut microbiome and found patterns of microbial dysbiosis at both sites. Of note, the disruption of the microbiome was related to the length of ICU stay and, secondly, attributed to ventilation and antibiotic treatment [19]. In our cohort of severely affected COVID-19 patients, specimen collection at prolonged timepoints during ICU stay, administration of broad-spectrum antibiotics, and invasive mechanical ventilation were identified to destabilize the oropharyngeal microbiome with a loss of diversity and severe dysbiosis, which goes along previous results on disruption of the lung microbiome in critically ill COVID-19 patients [12].

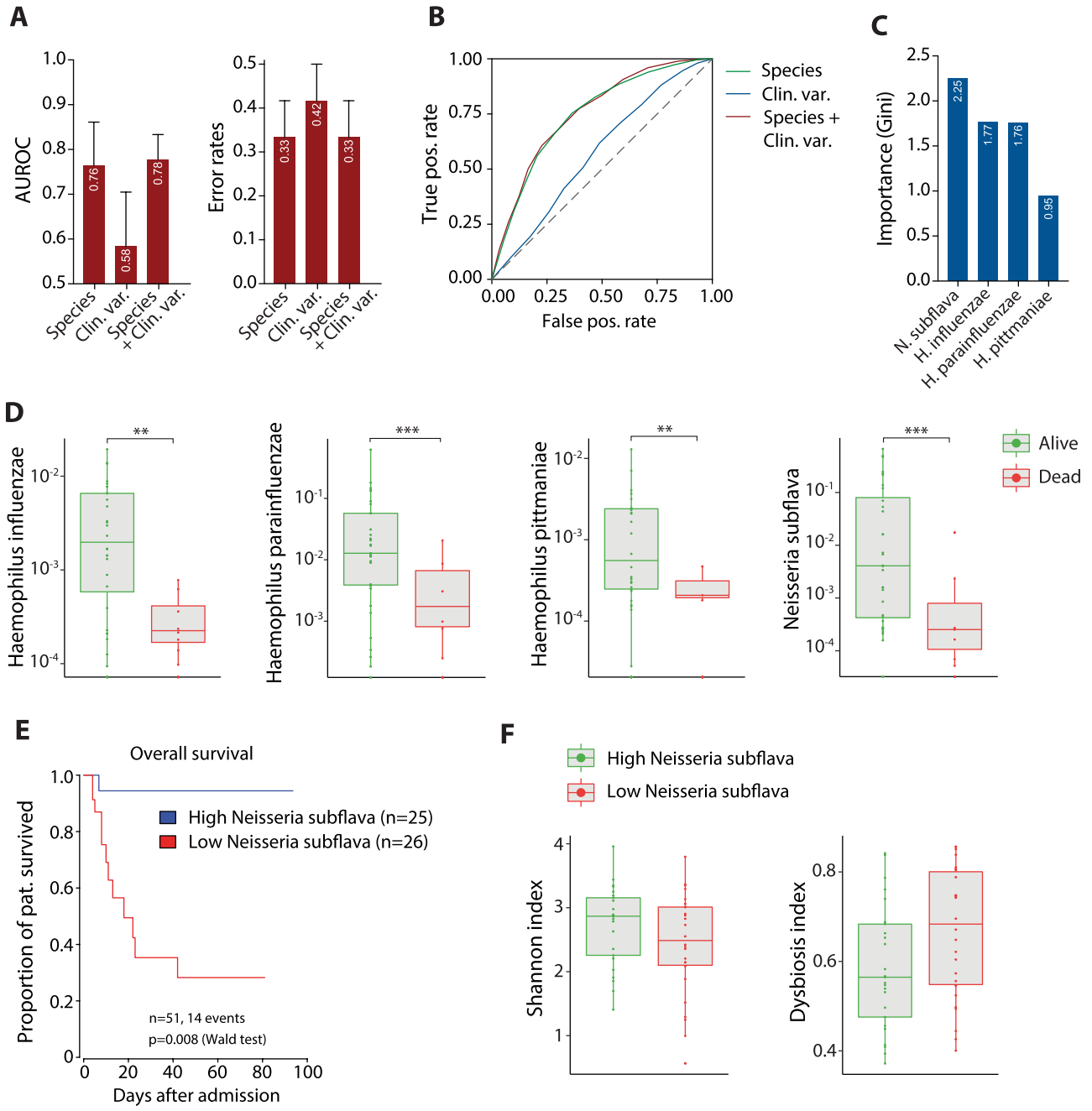

Figure 3. Prediction of mortality in hospitalized patients by RF models. *A*, Performance of the RF models using either taxonomic features (species), specific clinical parameters (clin. var.: age, gender, obesity, and antibiotics treatments), or both features combined (species + clin. var.) was assessed regarding mortality as major clinical outcome in balanced test data sets (n = 51, only patients with samples collected < day 4 after admission included). Bar plots indicate the average area under the receiver operating characteristic curves (AUROC, *left*), and error rates (*right*) obtained from the different sets of input features and corresponding RF models (median ± interquartile range, n = 100). *B*, True and false positive rates for mortality based on species, clinical variables, or a combination of both. *C*, Bar plot showing the final median Gini importance (n = 100) of the species found to be important for the prediction of mortality by the RF models (based on a median Gini importance > 0.05 and an FDR-corrected *P*-value < .05 for a positive impact on the accuracy of the classification, for each of the 3 different rarefying random sub-sampling of the sequencing data sets, n = 100 forests for each). Final Gini importance was obtained by RF models using only these selected species for predicting mortally. *D*, Relative abundances of important species in patients who succumbed to COVID-19 or were alive until discharge from the ICU or the hospital. *E* Survival analysis of hospitalized patients stratified low vs high *Neisseria subflava* relative abundance (only patients with samples from day 0 to 3 included). *F*, Shannon alpha diversity and dysbiosis index of samples stratified according to high vs low *Neisseria* abundances in samples of hospitalized patients collected between day 0 and 3 after admission (n = 51; Shannon index: *P* = .26; dysbiosis index: *P* = .07 [Wilcoxon tests]). ***P* < .01, ****P* < .001 by Wilcoxon test. Abbreviations: clin. var., clinical variance; COVID-19, coronavirus disease 2019; FDR, false discovery rate; pat., patients; pos., positive; ICU, intensive care unit; RF, random forest.

In addition to a general characterization of the oropharyngeal microbiome and covariates affecting diversity and composition, we asked whether configurations of the microbiome can be utilized to segregate mortality as major clinical outcome in hospitalized COVID-19 patients applying machine learning models which we have used in the past for the prediction of metabolic responses to diet [29]. We restricted this analysis to samples collected early after admission—at timepoints where we and others found little effects on community disruption by hospitalization and other external factors [19]. In our cohort, taxonomic features were more suitable to discriminate this clinically critical endpoint than selected clinical variables such as age, gender, or obesity that were previously reported as major risk factors for lethal COVID-19 [30, 31]. The abundances of *Neisseria subflava* and *Haemophilus* spp. were found as most important features associated with mortality, all of which are abundant in the local oropharyngeal microbiome [32]. This finding confirms a previous study by Merenstein et al [11] who also reported a significantly lower abundance of *Neisseria* and *Haemophilus* in the oropharyngeal microbiome of patients with severe disease. We found no association of high vs low abundance of *Neisseria* with a general diversity disruption or any association with specific genes or pathways encoded in the microbiome suggesting the *Neisseria* or *Haemophilus* spp. are not a surrogate marker for general community changes, but may be relevant for pathophysiological processes in SARS-CoV-2 infections. For instance, commensal *Neisseria* spp. have been reported to regulate innate immune responses and cytokine production in experimental infection models [33].

Limitations

Due to its cross-sectional design with a 1-time oropharyngeal microbiome collection either in the outpatient setting or during hospital admission or later during the hospital stay, and a lack of follow-up data on outpatients, we cannot infer any specific microbiome configuration at the very beginning of the infection or even in an asymptomatic pre-infectious state that would help to predict the full course of infection including its severity. Therefore, the predictive power of our observations is limited to hospitalized COVID-19 patients. With these limitations notwithstanding, utilization of the readily accessible oral microbiome signature at admission may enable data-driven patient stratification, leading to improved follow-up, tailored management from earlier disease stages and optimization of allocation of critical care supplies and staff prior to clinical deterioration.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. The authors would like to thank the FightCovid@DKFZ initiative and the High Throughput Sequencing unit of the Genomics and Proteomics Core Facility at the German Cancer Research Center for providing sequencing service on Illumina NovaSeq 6000 instruments.

Financial Support. This work was supported by a donation of the Ralph Lauren Foundation to R. B. and C. S. T. (12237). H. P. received funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—grant number 395357507–SFB 1371; M. E. and A. T. were supported by the state of Baden-Wuerttemberg, Ministry of Science, Research and Arts ([Forum Gesundheitsstandort Baden-Wuerttemberg], Project BW-ZDFP) and by The Sino-German Center for Research Promotion (German Research Foundation (DFG) and the National Natural Science Foundation of China (NSFC), Project C-0012).

Author Contributions. J. d. C., E. Z., T. H., and R. R. contributed equally. U. M. and C. S.-T. contributed equally. J. d. C., T. H., R. R., R. G., B. B-B., A. K., D. J., E. E., U. M., and C. S.-T. contributed to study concept and design. J. d. C., T. H., R. R., E. Z., S. S., S. H., H. S., M. N., S. K., T. M-E., A. H., Y. K., A. G., M. M., D. D., J. S., P. H., J. E., and C. S.-T. participated in data collection and analysis. J. d. C., D. K., N. C., R. B., F. H., R. S., M. V., A. T., M. E., B. S., H. P., E. E., and U. M. contributed to literature search, writing the article and data interpretation. J. d. C., E. Z., H. S., and C. S.-T. made figures and tables.

Potential conflicts of interest. A. K. reports support from BMS; consulting fees from BMS, Roche, Daiichi Sankyo, and Abbvie, honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Roche; travel support from BMS, Roche, Daiichi Sankyo, and Abbvie; participation on advisory board or DSMB from Roche, Daiichi Sankyo, and Abbvie, outside the conduct of the study. B. S. reports personal/consulting fees from GSK, Roche, and Sanofi; personal fees from Falk foundation for lectures, presentations, speakers bureaus, manuscript writing or educational events; serving as German Society of Infectious Diseases president, outside the conduct of the study. D. D. reports receiving consulting fees from Bayer Healthcare; honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Bayer Healthcare, Daiichi Sankyo, outside the conduct of the study. E. E. reports consulting fees and stock/stock options from DayTwo, BiomX, Roots GmBH (personal payments, unrelated to this work), outside the conduct of the study. J. S. reports honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Abbvie, Gilead, MSD, and Falk, outside the conduct of the study. M. V. reports grants from 3M, Astellas Pharma, Biontech, Da Volterra, Evonik, Gilead Sciences, Gycom Imunic, Maat Pharma, Merck, MSD, Organobalance, Seres Therapeautics, and Takeda Pharmaceutical, and consulting fees from Alb Fils Klinken GmbH, Arderypharm, Astellas Pharma, Basilea, Bimoreieux, Da Volterra, Farmak Internat. Holding GmbH, Ferring, Gilead Sciences, and SocraTec R&D GmbH, outside the conduct of the study. D. J. reports consulting fees from CureVac AG (Definiens), F. Hoffmann-La Roche Ltd. (Genmab A-S), Life Science Inkubator GmbH (VAXIMM AG), and OncoOne Research and Development Research GmbH (Oncolytics Biotech Inc.), honoraria for lectures, presentations, speakers bureaus, article writing or educational events from SKK Kliniken Heilbronn GmbH (Georg Thieme Verlag), Terrapinn (Touch Medical Media), and BMS GmbH & Co KGaA (MSD); payment for expert testimony from Wilhelm-Sander Stifung, Else-Kroner-Fresenius Stiftung, Schering Stiftung, and Nordforsk; travel support from Amgen Inc (Oryx GmbH), Roche Glycart AG (Parexel.com), and IKTZ HD GmbH (BMS); participation on DSMB or advisory board for CureVac AG (Definiens), F. Hoffmann-La Roche Ltd. (Genmab A-S), Life Science Inkubator GmbH (VAXIMM AG), and OncoOne Research and Development Research GmbH (Oncolytics Biotech Inc.), leadership role with BMS Stiftung Immunonkologi, outside the conduct of the study. All other authors report no potential conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. Cell Res **2020**; 30:492–506.

2. de Steenhuijsen Piters WA, Heinonen S, Hasrat R, et al. Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. Am J Respir Crit Care Med **2016**; 194:1104–15.
3. Tsang TK, Lee KH, Foxman B, et al. Association between the respiratory microbiome and susceptibility to influenza virus infection. Clin Infect Dis **2020**; 71:1195–203.
4. Rowe HM, Meliopoulos VA, Iverson A, Bomme P, Schultz-Cherry S, Rosch JW. Direct interactions with influenza promote bacterial adherence during respiratory infections. Nat Microbiol **2019**; 4:1328–36.
5. Fagundes CT, Amaral FA, Vieira AT, et al. Transient TLR activation restores inflammatory response and ability to control pulmonary bacterial infection in germ-free mice. J Immunol **2012**; 188:1411–20.
6. McCullers JA. The co-pathogenesis of influenza viruses with bacteria in the lung. Nat Rev Microbiol **2014**; 12:252–62.
7. Bahrani-Mougeot FK, Paster BJ, Coleman S, et al. Molecular analysis of oral and respiratory bacterial species associated with ventilator-associated pneumonia. J Clin Microbiol **2007**; 45:1588–93.
8. Ma S, Zhang F, Zhou F, et al. Metagenomic analysis reveals oropharyngeal microbiota alterations in patients with COVID-19. Signal Transduct Target Ther **2021**; 6:191.
9. Gaibani P, Viciani E, Bartoletti M, et al. The lower respiratory tract microbiome of critically ill patients with COVID-19. Sci Rep **2021**; 11:10103.
10. Braun T, Halevi S, Hadar R, et al. SARS-CoV-2 does not have a strong effect on the nasopharyngeal microbial composition. Sci Rep **2021**; 11:8922.
11. Merenstein C, Liang G, Whiteside SA, et al. Signatures of COVID-19 severity and immune response in the respiratory tract microbiome. mBIO **2021**; 12:e0177721.
12. Tsitsiklis A, Zha BS, Byrne A, et al. Impaired antibacterial immune signaling and changes in the lung microbiome precede secondary bacterial pneumonia in COVID-19. medRxiv **2021**: 2021.03.23.21253487.
13. Ren Z, Wang H, Cui G, et al. Alterations in the human oral and gut microbiomes and lipidomics in COVID-19. Gut **2021**; 70:1253–65.
14. Kelly BJ, Imai I, Bittinger K, et al. Composition and dynamics of the respiratory tract microbiome in intubated patients. Microbiome **2016**; 4:7.
15. Fernández-Barat L, López-Aladid R, Torres A. Reconsidering ventilator-associated pneumonia from a new dimension of the lung microbiome. EBioMedicine **2020**; 60:102995.
16. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet **2020**; 395:1054–62.
17. Beović B, Doušak M, Ferreira-Coimbra J, et al. Antibiotic use in patients with COVID-19: a 'snapshot' Infectious Diseases International Research Initiative (ID-IRI) survey. J Antimicrob Chemother **2020**; 75:3386–90.
18. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet **2021**; 397:220–32.
19. McDonald D, Ackermann G, Khailova L, et al. Extreme dysbiosis of the microbiome in critical illness. mSphere **2016**; 1:e00199-16.
20. Zhou YH, Gallins P. A review and tutorial of machine learning methods for microbiome host trait prediction. Front Genet **2019**; 10:579.
21. Topcuoglu BD, Lesniak NA, Ruffin MTt, Wiens J, Schloss PD. A framework for effective application of machine learning to microbiome-based classification problems. mBio **2020**; 11:e00434–20.
22. Liao D, Zhou F, Luo L, et al. Haematological characteristics and risk factors in the classification and prognosis evaluation of COVID-19: a retrospective cohort study. Lancet Haematol **2020**; 7:e671–8.
23. Merad M, Martin JC. Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. Nat Rev Immunol **2020**; 20:355–62.
24. Bost P, Giladi A, Liu Y, et al. Host-viral infection maps reveal signatures of severe COVID-19 patients. Cell **2020**; 181(7): 1475-88 e12.
25. Leentjens J, van Haaps TF, Wessels PF, Schutgens REG, Middeldorp S. COVID-19-associated coagulopathy and antithrombotic agents-lessons after 1 year. Lancet Haematol **2021**; 8:e524–33.
26. Shen Z, Xiao Y, Kang L, et al. Genomic diversity of severe acute respiratory syndrome-coronavirus 2 in patients with coronavirus disease 2019. Clin Infect Dis **2020**; 71: 713–20.
27. Fan J, Li X, Gao Y, et al. The lung tissue microbiota features of 20 deceased patients with COVID-19. J Infect **2020**; 81:e64–7.
28. De Maio F, Posteraro B, Ponziani FR, Cattani P, Gasbarrini A, Sanguinetti M. Nasopharyngeal microbiota profiling of SARS-CoV-2 infected patients. Biol Proced Online **2020**; 22:18.
29. Zeevi D, Korem T, Zmora N, et al. Personalized nutrition by prediction of glycemic responses. Cell **2015**; 163:1079–94.
30. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature **2020**; 584:430–6.
31. Karagiannidis C, Mostert C, Hentschker C, et al. Case characteristics, resource use, and outcomes of 10 021 patients with COVID-19 admitted to 920 German hospitals: an observational study. Lancet Respir Med **2020**; 8:853–62.
32. Liu G, Tang CM, Exley RM. Non-pathogenic Neisseria: members of an abundant, multi-habitat, diverse genus. Microbiology (Reading) **2015**; 161:1297–312.
33. Powell DA, Ma M, So M, Frelinger JA. The commensal *Neisseria musculi* modulates host innate immunity to promote oral colonization. Immunohorizons **2018**; 2:305–13.