JCI The Journal of Clinical Investigation

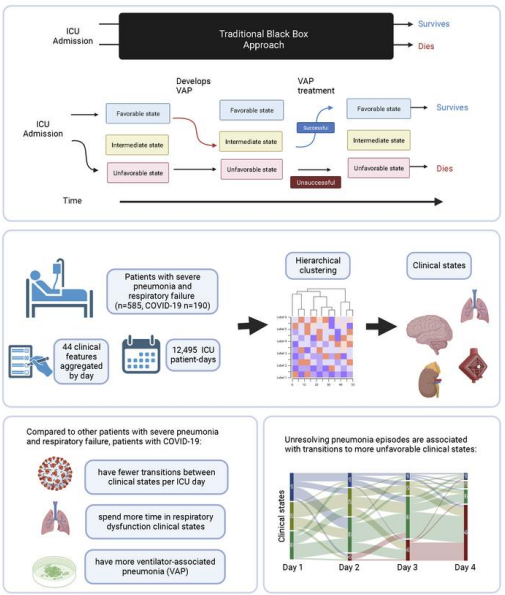
Machine learning links unresolving secondary pneumonia to mortality in patients with severe pneumonia, including COVID-19

Catherine A. Gao, … , Alexander V. Misharin, Benjamin D. Singer

J Clin Invest. 2023. https://doi.org/10.1172/JCI170682.

Clinical Medicine | In-Press Preview | Infectious disease | Pulmonology

Graphical abstract



Machine learning links unresolving secondary pneumonia to mortality in patients with severe pneumonia, including COVID-19

Catherine A. Gao[1] *, Nikolay S. Markov[1] *, Thomas Stoeger[2]*, Anna Pawlowski[3], Mengjia Kang[1], Prasanth Nannapaneni[3], Rogan A. Grant[1], Chiagozie Pickens[1] , James M. Walter[1] , Jacqueline M. Kruser[1,4], Luke Rasmussen[5], Daniel Schneider[3], Justin Starren[5], Helen K. Donnelly[1], Alvaro Donayre[1], Yuan Luo[5], GR Scott Budinger[1,6] **, Richard G. Wunderink[1,6] **, Alexander V. Misharin[1,6] **, Benjamin D. Singer[1,6] **,
The NU SCRIPT Study Investigators

*co-first authors
**co-senior authors

[1] Division of Pulmonary and Critical Care Medicine, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
[2] Department of Chemical and Biological Engineering, Northwestern University, McCormick School of Engineering, Evanston, IL, USA
[3] Northwestern Medicine Enterprise Data Warehouse, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
[4] Division of Allergy, Pulmonary and Critical Care, Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA
[5] Division of Health and Biomedical Informatics, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
[6] Simpson Querrey Lung Institute for Translational Science (SQLIFTS), Northwestern University Feinberg School of Medicine, Chicago, IL, USA

**Corresponding authors:
GR Scott Budinger, Richard G. Wunderink, Alexander V. Misharin, Benjamin D. Singer
303 E. Superior Street
Simpson Querrey, 5th Floor
Chicago, Illinois 60611, USA
Emails: s-buding@northwestern.edu, r-wunderink@northwestern.edu, a-misharin@northwestern.edu, benjamin-singer@northwestern.edu

Conflict of interests:
BDS holds US patent 10,905,706, “Compositions and methods to accelerate resolution of acute lung inflammation,” and serves on the Scientific Advisory Board of Zoe Biosciences, for which he holds stock options.
Other authors declare no conflicting interests.

Funding:
SCRIPT is supported by NIH/NIAID U19AI135964.
This study was also supported by the Simpson Querrey Lung Institute for Translational Science (SQLIFTS) and the Canning Thoracic Institute of Northwestern Medicine.
CAG is supported by NIH/NHLBI Training Grant T32HL076139 and F32HL162377.
TS is supported by K99AG068544.
YL is supported by NIH grants R01LM013337, U01TR003528.
GRSB was supported by NIH grants U19AI135964, P01AG049665, R01HL147575, P01HL071643, R01HL154686, Veterans Affairs grant I01CX001777, a grant from the Chicago Biomedical Consortium and a Northwestern University Dixon Translational Science Award.
RGW is supported by NIH grants U19AI135964, U01TR003528, P01HL154998, R01HL14988, R01LM013337

AVM is supported by NIH awards U19AI135964, P01AG049665, R21AG075423, R01HL158139, R01HL153312, P01HL154998.
BDS is supported by NIH awards R01HL149883, R01HL153122, P01HL154998, P01AG049665, and U19AI135964.

Role of funding source: our funding sources did not have a role in the design, execution, or prior review of the study or data presented in this manuscript. Opinions expressed in this work do not necessarily reflect those of the funding sources.

Abstract

Background: Despite guidelines promoting the prevention and aggressive treatment of ventilator-associated pneumonia (VAP), the importance of VAP as a driver of outcomes in mechanically ventilated patients, including patients with severe COVID-19, remains unclear. We aimed to determine the contribution of unsuccessful treatment of VAP to mortality in patients with severe pneumonia.

Methods: We performed a single-center prospective cohort study of 585 mechanically ventilated patients with severe pneumonia and respiratory failure, 190 of whom had COVID-19, who underwent at least one bronchoalveolar lavage. A panel of ICU physicians adjudicated pneumonia episodes and endpoints based on clinical and microbiologic data. Given the relatively long ICU length of stay among patients with COVID-19, we developed a machine learning approach called *CarpeDiem*, which groups similar ICU patient-days into clinical states based on electronic health record data.

Results: *CarpeDiem* revealed that the long ICU length of stay among patients with COVID-19 is attributable to long stays in clinical states characterized primarily by respiratory failure. While VAP was not associated with mortality overall, mortality was higher in patients with one episode of unsuccessfully treated VAP compared with successfully treated VAP (76.4% versus 17.6%, $p < 0.001$). In all patients, including those with COVID-19, *CarpeDiem* demonstrated that unresolving VAP was associated with transitions to clinical states associated with higher mortality.

Conclusions: Unsuccessful treatment of VAP is associated with greater mortality. The relatively long length of stay among patients with COVID-19 is primarily due to prolonged respiratory failure, placing them at higher risk of VAP.

Funding: U19AI135964

Brief summary

A machine learning approach identified unresolving ventilator-associated pneumonia as a major contributor to mortality in critically ill patients with pneumonia, including due to SARS-CoV-2.

Introduction

Several groups, including ours, have reported that durations of ICU stay and mechanical ventilation are more than twice as long among patients with severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) pneumonia compared to patients with respiratory failure complicating pneumonia due to other pathogens and patients with other causes of the acute respiratory distress syndrome (ARDS) (1–15). Based on analysis of peripheral blood samples from patients with severe compared with mild COVID-19, some investigators have hypothesized that the long ICU length of stay (LOS) among patients with SARS-CoV-2 pneumonia is secondary to multiple-organ dysfunction (reviewed in reference (16)). This hypothesis poorly explains the parallel observation that, despite their longer LOS, mortality is similar in patients with severe SARS-CoV-2 pneumonia compared to patients with pneumonia and respiratory failure secondary to other etiologies (2–4, 6). Intercurrent ICU events that disproportionately affect patients with COVID-19 might explain the disconnect between ICU LOS and mortality.

In a review of autopsy samples stored from the 1918 influenza A pandemic, Fauci and colleagues suggested an unexpectedly important contribution of secondary bacterial infection to mortality after severe viral pneumonia (17). Recent data suggest that secondary pneumonia is present in up to 40% and pneumonia or diffuse alveolar damage is present in over 90% of autopsy specimens obtained from patients with acute SARS-CoV-2 infection (18). Consistent with these observations, we and others found high rates of ventilator-associated pneumonia (VAP) in patients with SARS-CoV-2 pneumonia requiring mechanical ventilation, suggesting that bacterial superinfections such as VAP may contribute to mortality in patients with COVID-19 (7, 19–22). These findings prompt an alternative hypothesis that a relatively low mortality rate directly attributable to primary SARS-CoV-2 infection is offset by a greater risk of death attributable to unresolving VAP (23).

Testing the hypothesis that unresolving VAP explains the disconnect between ICU LOS and mortality in patients with severe SARS-CoV-2 pneumonia poses two challenges. First, traditional methods to compare ICU outcomes standardize severity of illness on ICU admission and treat the entirety of the ensuing ICU stay as a single event

(24–26). This approach fails to capture ICU complications, such as VAP, that by definition are rarely present on ICU admission but likely alter the trajectory of the patient toward unfavorable outcomes, particularly when ICU LOS is long. Indeed, few ICU studies have attempted to examine the effect of late ICU interventions and complications on patient outcomes (27–29). Second, if they resolve, VAP episodes may contribute to prolonged ICU LOS while not worsening outcomes. Nevertheless, most studies use insensitive methods to diagnose VAP and measure response to therapy (30).

We analyzed the contribution of VAP to mortality in 585 patients with severe pneumonia and respiratory failure, including 190 patients with severe SARS-CoV-2 pneumonia, enrolled in the Successful Clinical Response in Pneumonia Therapy (SCRIPT) study. All patients underwent bronchoalveolar lavage (BAL) sampling paired with comprehensive microbiological diagnostics at the time of study enrollment and whenever pneumonia was clinically suspected over the course of their intubation. Clinicians performed BAL sampling as part of routine clinical care and used BAL fluid studies to guide antimicrobial therapy (7). To disentangle the effect of VAP on outcomes over the course of the ICU stay, we developed a machine learning approach, *CarpeDiem*, which clustered individual patient-days in the ICU using clinical parameters extracted from the electronic health record (EHR). Because key clinical data fed the *CarpeDiem* algorithm, these clusters represented clinical states that were differentially associated with hospital mortality. The *CarpeDiem* framework allowed us to examine transitions between clinical states associated with favorable (lower mortality) or unfavorable (higher mortality) outcomes. Indeed, *CarpeDiem* revealed that the long ICU LOS among patients with COVID-19 relative to patients with pneumonia secondary to other pathogens resulted from excess days in clinical states characterized by severe hypoxemic respiratory failure with significantly fewer transitions between states when normalized for their longer LOS. Unresolving episodes of VAP were associated with transitions to clinical states associated with greater mortality. These data suggest mortality associated with severe SARS-CoV-2 pneumonia is more often associated with respiratory failure that increases the risk of unresolving VAP, and is less frequently associated with multiple-organ dysfunction.

Results

Demographics. Of 601 patients enrolled in SCRIPT between June 2018 and March 2022, 585 had an adjudicated pneumonia category and clinical endpoints at the time of analysis (**Figure 1**): 190 had COVID-19, 50 had pneumonia secondary to other respiratory viruses, 252 had other pneumonia (bacterial), and 93 were initially suspected of having pneumonia yet were subsequently adjudicated as having respiratory failure unrelated to pneumonia (non-pneumonia controls). Except for BMI, demographics such as age and gender were similar between groups (**Figure 2A-C** and **Supplemental Table 1**, which also includes a description of patient comorbidities). Severity of illness as measured by the Acute Physiology Score (APS) from APACHE IV (31) and the SOFA score (24, 32) in the first two days of admission did not differ between groups (**Figure 2D-E**). Patients across pneumonia categories underwent intubation following a similar duration of time in the ICU with a trend toward later intubation in patients with COVID-19 (**Supplemental Figure 1A** and **Supplemental Table 2**). At the time of intubation, SOFA scores were similar for patients with COVID-19 compared to other patients in the cohort who were intubated after admission to our hospital (**Supplemental Table 2**). On the first day of intubation, patients with COVID-19 had lower oxygen saturation levels despite a higher FiO_2 (**Supplemental Table 2**). They required higher levels of positive end-expiratory pressure (PEEP) but exhibited lower heart rates and were receiving lower doses of norepinephrine (**Supplemental Table 2**). Despite similar overall severity of illness on ICU admission, durations of intubation and ICU stay were more than twice as long among patients with SARS-CoV-2 pneumonia compared with any other group, reflected by a higher frequency of tracheostomy and longer ICU LOS (**Figure 2F-H** and **Supplemental Figure 2**). The longer ICU LOS persisted when patients who received extracorporeal membrane oxygenation (ECMO) support or who were received in external transfer (31.4% of the cohort) were excluded from the cohort (**Supplemental Figure 1B-C**). Hospital mortality did not differ between groups (**Figure 2I**). A similar fraction of patients with SARS-CoV-2 pneumonia received corticosteroids during their ICU stay compared with the rest of the cohort, but patients with SARS-CoV-2 pneumonia received higher cumulative doses (**Supplemental Table 1**). Patients with SARS-CoV-2 pneumonia were more likely to receive IL-6 receptor antagonists and remdesivir (**Supplemental Table 1**).

CarpeDiem*: a machine learning approach to time-series data in the ICU.* To address the challenge of comparing intercurrent ICU events between groups with different ICU LOS, we developed a machine learning approach, *CarpeDiem,* to discretize each patient-day in the ICU. For all 12,495 ICU patient-days in the cohort, we extracted clinical data from the EHR describing 44 key clinical parameters, including flags for organ failures requiring mechanical support (e.g., mechanical ventilation, renal replacement therapy, and ECMO), continuously recorded clinical parameters (e.g., vital signs and doses of norepinephrine), and commonly measured laboratory values (**Supplemental Figure 3A**). Variables used to calculate the SOFA score are a subset of these parameters. Importantly, patient-intrinsic variables (e.g., demographics, BMI, tracheostomy, and diagnosis), biochemical and microbiological analyses of BAL fluid studies, and adjudication of VAP episodes were not included in the model. Correlation analysis identified expected associations between mathematically or physiologically coupled variables (e.g., plateau pressure, PEEP, and lung compliance; $PaCO_2$ and bicarbonate) and revealed clinically recognizable correlated features (e.g., ECMO, D-dimer, and lactate dehydrogenase [LDH]) (**Supplemental Figure 3B**). After reducing the weight of these highly correlated features, we performed clustering using several methods, all of which yielded similar results (**Supplemental Figure 4A-C**). We designed a clustering strategy based on the similarity between patient-days (see details in Supplemental Methods) and selected the number of clusters by choosing a near-maximal difference in mortality between pairwise comparisons of clusters (**Supplemental Figure 5A**) while limiting the number of cluster breaks to those that were determined to be clinically meaningful by four ICU physicians (CAG, GRSB, RGW, BDS). To explore the stability of our clustering approach, we randomly excluded patients from our cohort and independently re-clustered this subset. While the overall patterns of clustering were similar, the assignment of patient-days to specific clusters differed (**Supplemental Figure 5B**). We visualized the resulting 14 clusters using heatmaps (**Figure 3A-B**) and UMAP plots (**Supplemental Figure 6**). Median SOFA scores for the days in each cluster are shown in **Supplemental Figure 7**. Every cluster contained patients and patient-days from each pneumonia category, including COVID-19 (**Supplemental Figure 8A-B**). Thus, the clinical states defined by *CarpeDiem* are useful to compare patient-days within a given cohort but do not represent *a priori* states to which patient-days can be prospectively assigned.

As *CarpeDiem* uses physiological parameters and laboratory values evaluated by clinicians to develop a daily plan of care, the clusters generated by *CarpeDiem* are recognizable as clinical states. To visualize these data, we arranged the parameters into six physiological groups (neurologic, respiratory, shock, renal, inflammatory, and ventilator instability) and sorted the clusters in order of increasing mortality. The resulting heatmaps (**Figure 3A-B**) and spider plots (**Figure 4**) revealed an association between patient-days characterized by multiple-organ failure and mortality, findings consistent with published scoring systems (24–26). We compared mortality for each clinical state identified by *CarpeDiem* on the first, median, and last day in the ICU. On the first day of the ICU stay, only two of the clinical states were significantly associated with outcome (**Supplemental Figure 9A**). In contrast, the same analysis for the median and last ICU day for each patient revealed eight and nine significant associations, respectively, between clinical state and outcome (**Supplemental Figure 9B-C**), supporting the construct validity of the *CarpeDiem*-generated clusters and the rationale to use *CarpeDiem* in an unsupervised fashion to evaluate all days of the ICU stay.

Critical care physicians (CAG, GRSB, RGW, BDS) used these visualizations to interpret the clinical states. For example, clinical state 12 represents patient-days with very severe respiratory failure (mostly days spent receiving ECMO support), moderately high levels of sedation, an intermediate level of shock without substantial renal failure, and with relatively stable ventilator settings. Importantly, while enriched for patients receiving ECMO support, clinical state 12 contained days spanning the duration of the ICU stay (**Supplemental Figure 10**), supporting that ECMO is a marker of persistent, severe respiratory failure rather than a salvage or peri-mortem intervention applied at the end of the ICU stay. An illustration of time series data and transitions between clinical states over a selected patient's ICU course is provided in **Supplemental Figure 6J**.

Validation of the CarpeDiem *approach in the MIMIC-IV dataset*. We next determined whether the *CarpeDiem* approach could be used to analyze an external dataset. Within the MIMIC-IV database of ICU patients(33), we identified the subset of 1,284 ICU stays similar to those in our cohort. The *CarpeDiem* approach applied to 15,642 ICU patient-days using 27 clinical parameters, a subset of the 44 used above that were readily available in the MIMIC-IV database, identified 12 clusters (**Supplemental Figure 11A-C**). Similar to our observations in

the SCRIPT cohort, *CarpeDiem*-generated clusters in MIMIC-IV were clinically recognizable with increasing organ failure associated with mortality (**Supplemental Figure 11D-E**). While these results support the generalizability of the *CarpeDiem* approach, the clinical states observed in the MIMIC-IV cohort were not identical to those in the SCRIPT cohort. This observation might be expected as, for example, MIMIC-IV had very few patients receiving ECMO, underscoring the concept that clinical states cannot be assigned *a priori* in a given cohort.

CarpeDiem *reveals that the long LOS among patients with COVID-19 is associated with prolonged stays in clinical states characterized by severe respiratory failure.* We reasoned that *CarpeDiem* could provide insight into the reasons why patients with severe SARS-CoV-2 pneumonia have longer ICU LOS relative to patients with pneumonia and respiratory failure secondary to other etiologies despite similar hospital mortality. We posited that this observation could result from 1) longer stays in a given clinical state with similar numbers of transitions between states, as would be observed for prolonged respiratory failure or 2) similar durations of stay in any given clinical state with a balanced increase in the number of transitions between favorable and unfavorable states, as might be observed in patients developing multiple-organ dysfunction. Although the absolute number of transitions between clinical states was higher among patients with SARS-CoV-2 pneumonia when compared with all other patient groups in the cohort (**Figure 5A** and **Supplemental Figure 12A**), the frequency of transitions was significantly lower (**Figure 5B** and **Supplemental Figure 12B**). The longer ICU LOS experienced by patients with severe SARS-CoV-2 pneumonia resulted from significantly prolonged stays in four clinical states (**Figure 5C**). Clusters that were enriched in days from patients with COVID-19 had higher respiratory severity scores (**Figure 3, Figure 4**, and **Figure 5D**), illustrating that patients with COVID-19 spent a disproportionate amount of time in clusters characterized by hypoxemic respiratory failure. Time spent in clinical state 12, characterized by severe hypoxemic respiratory failure, accounted for 29.9% of the difference in ICU LOS experienced by patients with COVID-19. Overall, since some clusters were deficient in patients with COVID-19, time spent in the four clinical states that were significantly enriched in patients with COVID-19 accounted for over 100% of the difference in ICU LOS between patients with and without COVID-19.

To examine the robustness of our findings to changes in the composition of the cohort, we randomly excluded 20% of the cohort and re-clustered patient-days 500 times. As shown in **Supplemental Figure 13**, the main conclusions drawn from the full dataset hold after random subsampling, including the finding that patients with COVID-19 experience fewer transitions per day irrespective of outcome (as in **Figure 5B**) and experience longer stays in clusters with high respiratory severity scores (as in **Figure 5D**).

To explore the potential utility of the *CarpeDiem* approach within the context of a randomized controlled trial, we analyzed the 10 patients within SCRIPT who were also enrolled in a randomized placebo-controlled trial of the IL-6 receptor antagonist sarilumab for treatment of patients with respiratory failure secondary to COVID-19. The results of randomized controlled trials of IL-6 receptor antagonists in patients with COVID-19 have been mixed (34), with some trials reporting benefit, while others, including this trial (35), did not. We calculated the sum of *CarpeDiem*-defined clinical state transitions occurring three and five days following randomization to sarilumab (n = 6) or placebo (n = 4). Even within this very small group, we observed significantly more favorable transitions in patients who received sarilumab compared to those who received placebo in the three days after drug administration (**Supplemental Figure 14A-B**). In contrast, no statistically significant difference was evident five days after randomization (**Supplemental Figure 14C**).

Unresolving VAP drives poor outcomes in patients with severe pneumonia, including due to SARS-CoV-2. Nearly all patients (97.4%) underwent transitions between clinical states over the course of their ICU stay (median [IQR] of 4[2,7] transitions per patient). We defined transitions as favorable if the mortality associated with the destination clinical state was lower than the originating state and *vice versa*. While the number of unfavorable transitions was similar in patients with SARS-CoV-2 pneumonia and other patients in the cohort, the number of favorable transitions was nominally lower in patients with SARS-CoV-2 pneumonia (**Figure 6** and **Supplemental Figure 15A-B**).

We hypothesized that VAP would at least in part explain the disconnect between ICU LOS and mortality in patients with COVID-19. Overall, 35.5% of patients in the cohort developed at least one episode of VAP during

their ICU stay (25.0% among patients without COVID-19 compared to 57.4% among patients with COVID-19, $p < 0.001$) (**Figure 7A**). 8.7% of patients in the cohort experienced more than one episode of VAP (3.5% among patients without COVID-19 compared to 19.5% among patients with COVID-19, $p < 0.001$) (**Figure 7B**). Mortality in patients with VAP has been reported to increase substantially with each ensuing episode, approaching 100% in patients with three or more episodes (36). In contrast, we found that the mortality associated with a single VAP episode did not differ from the mortality associated with multiple VAP episodes (48.6% with a single episode, 53.6% with two episodes, 50.0% with three episodes; p = not significant) (**Figure 7C**), suggesting that cure can be achieved even in patients with multiple VAP episodes. Nevertheless, the relatively small number of patients with multiple VAP episodes limits power to detect small differences (**Figure 7D**).

Overall, mortality was not significantly different in patients who developed VAP compared to those who did not (**Figure 8A**). To further explore the association between VAP and ICU outcomes, we used the validated clinical adjudication results from the SCRIPT study to compare patients with successful treatment of VAP (cured) with those who experienced an indeterminate outcome or unsuccessful treatment (not cured). Examining these endpoints among patients who had only a single VAP episode, we found that mortality was lowest among patients with successful treatment (cured), intermediate among those with an indeterminate outcome, and highest among those with unsuccessful treatment (not cured) (**Figure 8B**). Among these patients, the rate of unfavorable outcomes (hospice or death) was 17.6% in patients with a cured episode and 76.5% in patients with unsuccessful treatment (intermediate or not cured episode), $p < 0.001$. We also observed a similar pattern among the subset of patients with COVID-19 (**Supplemental Figure 16A**). Patients with COVID-19 experienced longer durations of VAP episodes (**Figure 8C**). Unresolving VAP episodes (those with an indeterminate outcome or that were not cured) were of longer duration than cured episodes (**Figure 8D**). Since survival is included in our definition of successful VAP treatment, we performed a sensitivity analysis on VAP episodes experienced by patients who survived for at least 14 days following VAP diagnosis. Even among this group biased toward better outcomes, we found that unresolving VAP was associated with a higher mortality rate (**Supplemental Figure 16B**).

CarpeDiem corroborates the clinical adjudication analysis, identifying an association between unresolving VAP episodes and transitions to unfavorable clinical states associated with greater hospital mortality. We then used the transition analyses provided by *CarpeDiem* to test whether unresolving VAP was associated with a subsequent trajectory toward progressively unfavorable clinical states. To visualize the transitions surrounding the diagnosis of VAP, we generated Sankey diagrams that show the clinical state and transitions encountered preceding and following the diagnosis of VAP. Successful treatment of VAP was associated with an higher likelihood of favorable subsequent transitions (**Figure 9A**). In contrast, indeterminate episodes demonstrated a flat trajectory (**Figure 9B**). Not cured episodes were associated with a greater risk of unfavorable subsequent transitions (**Figure 9C** and **Supplemental Figure 17**). The robustness of the propensity for patients with cured VAP to undergo more favorable transitions than patients without cured VAP was confirmed in subsampling analysis (**Supplemental Figure 13E**). We then used the sum of transitions occurring in the seven days following a diagnosis of VAP as a summative measure of trajectory and examined the distribution of trajectories to define favorable, intermediate, and unfavorable trajectory categories **(Figure 10A**). Favorable trajectories were significantly enriched in cured VAP episodes with significantly higher proportions of indeterminate and not cured episodes in intermediate and unfavorable trajectory categories, respectively (**Figure 10B**). Finally, we examined the trajectory categories preceding a VAP diagnosis compared with the average inter-day trajectory across the cohort. We identified an increase in unfavorable transitions one day ahead of a VAP diagnosis, presumably reflecting the clinical events that prompted the diagnostic BAL procedure, that was not associated with the duration of the ensuing VAP episode (**Supplemental Figure 18A-B**).

To assess whether the same associations could be revealed independent of the *CarpeDiem* approach, we added flags denoting the development of VAP and its outcome to a standard model of ICU mortality prediction based on clinical parameters measured early (in the first two days) of ICU admission. Using gradient boosting, we found only a nominal increase in the predictive ability of early clinical parameters with addition of the VAP flags (**Supplemental Figure 19A**). These findings are possibly explained by the disconnect between clinical parameters measured early in a clinical course and the fact that VAP, by definition, occurs later in an ICU stay.

Expectedly, the same clinical parameters applied to the median two days or final two days of the ICU stay had intermediate and excellent predictive capability, respectively, but were similarly unmodified by addition of the VAP flags (**Supplemental Figure 19B-C**).

Discussion

The ICU course of patients with severe SARS-CoV-2 pneumonia is more than twice as long as the duration among similarly ill patients with pneumonia and respiratory failure due to other etiologies (1–15). Despite significantly longer durations of critical illness, mortality in patients with COVID-19 is similar to patients with other causes of pneumonia and respiratory failure (2–4, 6). We and others have reported unexpectedly high rates of VAP complicating the ICU course of patients with SARS-CoV-2 pneumonia (7, 23). In this large prospective observational cohort study, we used state-of-the-art microbiological analysis of serially-collected BAL samples (7, 30, 37) over the course of the ICU stay combined with validated clinical adjudications to identify VAP episodes and clinical endpoints. We found that unresolving episodes of VAP were associated with mortality, including among patients with COVD-19. Accordingly, we suggest that the discordance between ICU LOS and mortality in patients with severe SARS-CoV-2 pneumonia results from a low mortality attributable to the primary viral pneumonia that is offset by an increased risk of mortality from unresolving VAP or other ICU complications.

Our cohort included large numbers of patients with COVID-19 and similarly ill patients with pneumonia secondary to other pathogens, providing an opportunity to determine whether and how VAP differentially contributes to outcomes in patients with COVID-19. Compared to patients with pneumonia secondary to other pathogens, we found that patients with COVID-19 had a longer duration of mechanical ventilation and ICU stay and higher rates of VAP yet similar mortality. Because VAP, the duration of mechanical ventilation, and mortality are interrelated, we developed a data-driven machine learning approach to disentangle these features. *CarpeDiem* uses data extracted from the EHR to discretize days in the ICU and generate clusters of patient-days with similar physiological and laboratory features, paralleling the practice of daily ICU rounds. As key clinical data drove the *CarpeDiem* algorithm, the resulting clusters represented clinical states that were associated with differential hospital mortality. As patients improve or worsen, they undergo transitions to more favorable (lower mortality) or less favorable (higher mortality) clinical states. We reasoned that if multiple-organ failure drives prolonged ICU LOS among patients with COVID-19, *CarpeDiem* would identify frequent transitions between clinical states associated with more organ failures. Instead, *CarpeDiem* showed that the long ICU LOS among patients with

SARS-CoV-2 pneumonia is attributable to significantly longer stays in clinical states primarily characterized by severe hypoxemic respiratory failure. When normalized for ICU LOS, patients with SARS-CoV-2 pneumonia experienced fewer transitions between clinical states than other patients over the course of their ICU stay. This finding provides clinical support for emerging models of SARS-CoV-2 pneumonia pathobiology. In these models, severe SARS-CoV-2 pneumonia results from a slowly progressive but spatially localized pulmonary infection that unfolds over days to weeks (6, 16, 38), leading to prolonged respiratory failure and higher rates of VAP.

Guidelines adopted by professional societies recommend a host of interventions to prevent and treat known or suspected VAP in patients requiring mechanical ventilation, implicitly acknowledging the importance of VAP in determining outcomes (39, 40). Nevertheless, we are unaware of prior studies demonstrating an association between unresolving VAP with poor ICU outcomes. We used a rigorous clinical and microbiological adjudication procedure to show that unresolving VAP was associated with mortality, including in patients with SARS-CoV-2 pneumonia. Furthermore, *CarpeDiem* demonstrated that unresolving VAP was associated with transitions toward unfavorable clinical states, providing independent, complementary, and unsupervised support for our adjudication procedures and findings. Perhaps as importantly, we found that successfully treated VAP was associated with improved outcomes and favorable transitions in all patients with severe respiratory failure. These findings suggest that improved strategies to diagnose and successfully treat VAP episodes, including pathogen-directed therapy guided by BAL fluid analysis, may improve ICU outcomes.

The importance of VAP as a driver of mortality in patients with COVID-19 has been underestimated, likely because bronchoscopic sampling has been uncommon during the pandemic, use of antibiotics is ubiquitous, and clinical criteria and biomarkers do not accurately distinguish between primary SARS-CoV-2 pneumonia and secondary bacterial pneumonia (41). For example, only one episode of secondary pneumonia was reported in the 403 patients included in the REMAP-CAP trial of hydrocortisone for COVID-19, and no episodes were reported in the 6,425 patients included in the RECOVERY trial of dexamethasone therapy for COVID-19 (1, 15). If unresolving episodes of VAP rather than the primary viral pneumonia contribute to mortality in a substantial fraction of patients with severe SARS-CoV-2 pneumonia, it might explain why therapies that attenuate the host

response (e.g., corticosteroids, IL-6 receptor antagonists, JAK2 inhibitors, and CRAC channel inhibitors) are more effective when administered early in the clinical course, before patients are intubated and at risk for VAP (1, 34, 42–44).

Our study has important limitations. First, as ours is an observational study, we cannot exclude unmeasured confounders that link unresolving VAP to poor outcomes. Other processes of care, such as ventilator and antibiotic management strategies, and host factors, such as exposure to immunomodulatory therapies and alterations in the microbiome, likely drive VAP outcomes. Second, we used state-of-the art clinical microbiological analysis of distal lung samples to diagnose VAP, and we have shown that clinicians in our center use this information to optimize, narrow, or discontinue antibiotic therapy (7), minimizing its harmful effects (45). The reasons underlying the failure of appropriate antimicrobial therapy in some patients cannot easily be determined from our study, raising important questions about the drivers of unresolving VAP despite targeted antimicrobial therapy. Potential drivers of unresolving VAP include pharmacodynamic and pharmacokinetic properties of pathogen-targeted antibiotics, dysregulated microbiome composition, and an inappropriate balance between host immune responses that favor ongoing inflammation and injury versus resolution and repair. Further studies of the pathogen, microbiome, and host response using high-resolution next-generation sequencing approaches applied to BAL fluid and spatial profiling applied to lung tissue may provide insights into these mechanisms. Ultimately, causal validation of these mechanisms will need to come from in vitro systems, experimental animal models of pneumonia, and randomized controlled trials in patients. Third, it is important to note that the clustering tools used in *CarpeDiem*, necessarily driven by patients with a longer LOS, will generate different clusters as the composition of the cohort changes. Therefore, *CarpeDiem* is primarily useful to compare ICU diagnoses, interventions, and complications within a single cohort, while application of the clinical states from one cohort to another, or prospective assignment of clinical states to new data in the same cohort, remains to be investigated. In our analysis of patients enrolled in a randomized controlled trial of sarilumab who were also included in our cohort, we demonstrated the potential utility of the *CarpeDiem* approach for generating hypotheses about mechanisms that may underlie negative findings in the trial. In our example, the results suggested that the negative outcome might have resulted from a lack of repeated sarilumab dosing. If confirmed

in larger numbers of patients, these results might inform the design of subsequent trials. Importantly, *CarpeDiem* uses only data collected as part of routine clinical care; therefore, it could be retrospectively applied to EHR data from multiple centers to analyze the results of this or other clinical trials in the ICU. Fourth, *CarpeDiem* uses a limited number of parameters to define clinical states, potentially neglecting important determinants of outcome and information that might be found in missing data (e.g., reduced monitoring and ordering of laboratory tests as patients improve or move toward comfort-focused care). Similarly, intermittently measured biomarkers associated with outcomes, for example those used by Calfee et al. to define hyper- and hypo-inflammatory states in the ICU, are incompletely represented in *CarpeDiem* (46). Future iterations of the tool can incorporate these data with a goal of improving the association between clinical states and mortality in both observational and interventional studies. To this end, we have made de-identified data from the SCRIPT dataset, as well as detailed code, freely available to the research community.

Methods

Study setting. Patients were enrolled in the Successful Clinical Response in Pneumonia Therapy (SCRIPT) Systems Biology Center, a single-site, prospective cohort study of patients hospitalized in the ICUs of Northwestern Memorial Hospital (NMH) with suspected severe pneumonia (severe pneumonia defined as lower respiratory tract infection requiring mechanical ventilation), all of whom underwent at least one BAL procedure. A subset of patients were co-enrolled in a randomized placebo-controlled trial of the IL-6 receptor antagonist sarilumab (NCT04315298).

Study procedures. ICU physicians at NMH routinely obtain bronchoscopic or non-bronchoscopic BAL samples from mechanically ventilated patients whenever pneumonia is suspected (47). In SCRIPT, patients were screened for enrollment when the clinical team decided to perform the first BAL procedure. For all BAL samples, NMH clinical laboratories perform quantitative bacterial culture and antimicrobial susceptibility testing. Many of the samples in SCRIPT were also analyzed by multiplex PCR (BioFire® FilmArray® Pneumonia (PN) Panel) with results provided to the clinical team within three hours. We previously reported that our physicians initiate guideline-recommended antimicrobial therapy when pneumonia is suspected and use data obtained from analysis of BAL fluid to appropriately narrow or discontinue empirical guideline-recommended antimicrobial therapy (7).

Data extraction and analysis. Demographics, clinical data, and outcomes were extracted from the EHR via the Northwestern Medicine Enterprise Data Warehouse (48). For the *CarpeDiem* machine learning approach, we trialed three different computational strategies that involved hierarchical clustering of 44 clinical features. We chose the number of clusters by optimizing clinical interpretability and reasonable between-cluster differential mortality. We used UMAP (49) for visualization. We externally validated the *CarpeDiem* approach in a suspected pneumonia cohort derived from the MIMIC-IV database (33), using code from the MIMIC Code Repository (50). Selection criteria in the MIMIC-IV cohort included admission to and discharge from a medical ICU, respiratory failure requiring mechanical ventilation, and pneumonia as defined by ICD-9 codes. We used XGBoost (51) to

model outcomes based on clinical features from the first two days (similar to most clinical prediction models), median two days, and last two days of the admission. See Supplemental Materials for extended methods.

Definition of pneumonia episodes. A panel of six critical care physicians used a prospectively generated, standardized score sheet (**Supplemental File 1**) to manually review each patient's EHR, including clinical notes, to identify and categorize pneumonia episodes and adjudicate whether these episodes were successfully treated (i.e., resolved). A detailed description of our adjudication protocol is available (52). Pneumonia episodes were captured up to 99 days following the enrollment BAL procedure and categorized as non-pneumonia controls, other pneumonia (bacterial), other viral pneumonia, or COVID-19. By definition, VAP was considered to be an incident pneumonia that was diagnosed by a BAL performed after at least 48 hours of mechanical ventilation (39); only VAP episodes that were adjudicated to be due to bacteria were included in the analysis. VAP duration was defined as the time interval between the diagnostic BAL procedure and clinical cure, discontinuation of antibiotics, or death, whichever was the shortest. Endpoints for VAP episodes were adjudicated at day 7-8, day 10, and day 14 following the diagnostic BAL procedure. See Supplemental Materials for detailed definitions.

Statistics. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. Categorical values were compared using Fisher's Exact tests with FDR correction using the Benjamini-Hochberg procedure. A p-value or q-value < 0.05 was our threshold for statistical significance.

Study approval. This study was approved by the Northwestern University Institutional Review Board with study ID STU00204868. The sarilumab trial was approved by the Northwestern University Institutional Review Board with study ID STU00212239.

Code and data availability. Programming was performed in Python (version 3.9). A detailed description of all data extraction and computational procedures, including code, are available at https://github.com/NUSCRIPT/carpediem and in Supplemental Materials. A de-identified version of all SCRIPT

cohort data used in this manuscript is available on PhysioNet at https://doi.org/10.13026/5phr-4r89 (53, 54). A demo interactive data browser illustrating the features of *CarpeDiem* is available on our website, https://nupulmonary.org/carpediem, and the full browser is available on PhysioNet.

Author contributions

CAG, NSM, TS, GRSB, RGW, AVM, BDS conceived, designed, interpreted results, and wrote/edited the manuscript.
CAG, NSM, TS performed programming and analysis.
CAG, CP, JMW, JMK, RGW, BDS performed clinical pneumonia episode adjudication.
HKD and AD enrolled patients and compiled clinical adjudication data.
AP, PN, RAG, CAG, MK, LR, DS, JS, YL compiled the dataset.
CAG, NSM, TS, MK, LR, DS, JS have directly accessed and verified the underlying data reported in the manuscript.
All authors confirm that they had full access to all the data in the study and accept responsibility to submit for publication. All authors read and approved the final draft of the manuscript.

The order of the three co-first authors was determined by contributions to text of manuscript.

The NU SCRIPT Study Investigators are listed in **Supplemental File 2.**

Acknowledgements

The authors would like to thank Malte Luecken and Neal Ravindra for valuable discussion. The graphical abstract was created with BioRender.com.

References

1. RECOVERY Collaborative Group et al. Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary Report [Internet]. *N. Engl. J. Med.* [published online ahead of print: July 17, 2020]; doi:10.1056/NEJMoa2021436

2. Ferrando C et al. Clinical features, ventilatory management, and outcome of ARDS caused by COVID-19 are similar to other causes of ARDS. *Intensive Care Med.* 2020;46(12):2200–2211.

3. Gamberini L et al. Factors influencing liberation from mechanical ventilation in coronavirus disease 2019: multicenter observational study in fifteen Italian ICUs. *J. Intensive Care Med.* 2020;8(1):80.

4. Roedl K et al. Mechanical ventilation and mortality among 223 critically ill patients with coronavirus disease 2019: A multicentric study in Germany [Internet]. *Aust. Crit. Care* [published online ahead of print: October 27, 2020]; doi:10.1016/j.aucc.2020.10.009

5. Guan W-J et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* 2020;382(18):1708–1720.

6. Grant RA et al. Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia [Internet]. *Nature* [published online ahead of print: January 11, 2021]; doi:10.1038/s41586-020-03148-w

7. Pickens CO et al. Bacterial Superinfection Pneumonia in Patients Mechanically Ventilated for COVID-19 Pneumonia. *Am. J. Respir. Crit. Care Med.* 2021;204(8):921–932.

8. Zhu N et al. A novel Coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 2020;382(8):727–733.

9. Arons MM et al. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N. Engl. J. Med.* 2020;382(22):2081–2090.

10. Baggett TP, Keyes H, Sporn N, Gaeta JM. Prevalence of SARS-CoV-2 infection in residents of a large homeless shelter in Boston. *JAMA* 2020;323(21):2191–2192.

11. Louie JK et al. Lessons from mass-testing for Coronavirus disease 2019 in long-term care facilities for the elderly in San Francisco. *Clin. Infect. Dis.* 2021;72(11):2018–2020.

12. Kasper MR et al. An outbreak of Covid-19 on an aircraft carrier. *N. Engl. J. Med.* 2020;383(25):2417–2426.

13. Letizia AG et al. SARS-CoV-2 Transmission among Marine Recruits during Quarantine. *N. Engl. J. Med.* 2020;383(25):2407–2416.

14. Wu Z, McGoogan JM. Characteristics of and important lessons from the Coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese center for disease control and prevention. *JAMA* 2020;323(13):1239–1242.

15. Tomazini BM et al. Effect of Dexamethasone on Days Alive and Ventilator-Free in Patients With Moderate or Severe Acute Respiratory Distress Syndrome and COVID-19: The CoDEX Randomized Clinical Trial. *JAMA* 2020;324(13):1307–1316.

16. Budinger GRS, Misharin AV, Ridge KM, Singer BD, Wunderink RG. Distinctive features of severe SARS-CoV-2 pneumonia [Internet]. *J. Clin. Invest.* 2021;131(14). doi:10.1172/JCI149412

17. Morens DM, Taubenberger JK, Fauci AS. Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness. *J. Infect. Dis.* 2008;198(7):962–970.

18. Stein SR et al. SARS-CoV-2 infection and persistence in the human body and brain at autopsy [Internet]. *Nature* [published online ahead of print: December 14, 2022]; doi:10.1038/s41586-022-05542-y

19. Maes M et al. Ventilator-associated pneumonia in critically ill patients with COVID-19. *Crit. Care* 2021;25(1):25.

20. Rouzé A et al. Relationship between SARS-CoV-2 infection and the incidence of ventilator-associated lower respiratory tract infections: a European multicenter cohort study. *Intensive Care Med.* 2021;47(2):188–198.

21. Blonz G et al. Epidemiology and microbiology of ventilator-associated pneumonia in COVID-19 patients: a multicenter retrospective study in 188 patients in an un-inundated French region. *Crit. Care* 2021;25(1):72.

22. Vacheron C-H et al. Increased incidence of ventilator-acquired pneumonia in Coronavirus disease 2019 patients: A multicentric cohort study. *Crit. Care Med.* 2022;50(3):449–459.

23. Vacheron C-H et al. Attributable Mortality of Ventilator-associated Pneumonia Among COVID-19 Patients [Internet]. *Am. J. Respir. Crit. Care Med.* [published online ahead of print: May 10, 2022]; doi:10.1164/rccm.202202-0357OC

24. Vincent JL et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on “sepsis-related problems” of the European Society of Intensive Care Medicine. *Crit. Care Med.* 1998;26(11):1793–1800.

25. Zimmerman JE, Kramer AA, McNair DS, Malila FM, Shaffer VL. Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV. *Crit. Care Med.* 2006;34(10):2517–2529.

26. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA* 2001;286(14):1754–1758.

27. Viglianti EM et al. Late vasopressor administration in patients in the ICU: A retrospective cohort study. *Chest* 2020;158(2):571–578.

28. Viglianti EM et al. Late organ failures in patients with prolonged intensive care unit stays. *J. Crit. Care* 2018;46:55–57.

29. Iwashyna TJ et al. Timing of onset and burden of persistent critical illness in Australia and New Zealand: a retrospective, population-based, observational study. *Lancet Respir. Med.* 2016;4(7):566–573.

30. Pickens CI, Wunderink RG. Principles and Practice of Antibiotic Stewardship in the ICU. *Chest* 2019;156(1):163–171.

31. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit. Care Med.* 2006;34(5):1297–1310.

32. Vincent JL et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* 1996;22(7):707–710.

33. Johnson A et al. MIMIC-IV [Internet]2022; doi:10.13026/7VCR-E114

34. Angriman F et al. Interleukin-6 receptor blockade in patients with COVID-19: placing clinical trials into context [Internet]. *Lancet Respir Med* [published online ahead of print: April 27, 2021]; doi:10.1016/S2213-2600(21)00139-9

35. Sivapalasingam S et al. Efficacy and Safety of Sarilumab in Hospitalized Patients With Coronavirus Disease 2019: A Randomized Clinical Trial. *Clin. Infect. Dis.* 2022;75(1):e380–e388.

36. Silver DR, Cohen IL, Weinberg PF. Recurrent Pseudomonas aeruginosa pneumonia in an intensive care unit [Internet]. *Chest* 1992;101(1). doi:10.1378/chest.101.1.194

37. Pickens C et al. A multiplex polymerase chain reaction assay for antibiotic stewardship in suspected pneumonia. *Diagn. Microbiol. Infect. Dis.* 2020;98(4):115179.

38. Muus C et al. Single-cell meta-analysis of SARS-CoV-2 entry genes across tissues and demographics. *Nat. Med.* 2021;27(3):546–559.

39. Torres A et al. International ERS/ESICM/ESCMID/ALAT guidelines for the management of hospital-acquired pneumonia and ventilator-associated pneumonia: Guidelines for the management of hospital-acquired pneumonia (HAP)/ventilator-associated pneumonia (VAP) of the European Respiratory Society (ERS), European Society of Intensive Care Medicine (ESICM), European Society of Clinical Microbiology and Infectious Diseases (ESCMID) and Asociación Latinoamericana del Tórax (ALAT) [Internet]. *Eur. Respir. J.*

2017;50(3). doi:10.1183/13993003.00582-2017

40. Kalil AC et al. Management of Adults With Hospital-acquired and Ventilator-associated Pneumonia: 2016 Clinical Practice Guidelines by the Infectious Diseases Society of America and the American Thoracic Society. *Clin. Infect. Dis.* 2016;63(5):e61–e111.

41. Gao CA et al. Bronchoscopy on Intubated COVID-19 Patients is Associated with Low Infectious Risk to Operators [Internet]. *Ann. Am. Thorac. Soc.* [published online ahead of print: January 15, 2021]; doi:10.1513/AnnalsATS.202009-1225RL

42. Angus DC et al. Effect of Hydrocortisone on Mortality and Organ Support in Patients With Severe COVID-19: The REMAP-CAP COVID-19 Corticosteroid Domain Randomized Clinical Trial. *JAMA* 2020;324(13):1317–1329.

43. Kalil AC et al. Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19 [Internet]. *N. Engl. J. Med.* [published online ahead of print: December 11, 2020]; doi:10.1056/NEJMoa2031994

44. Bruen C et al. Auxora vs. placebo for the treatment of patients with severe COVID-19 pneumonia: a randomized-controlled clinical trial. *Crit. Care* 2022;26(1):101.

45. Kollef MH et al. A randomized trial of 7-day doripenem versus 10-day imipenem-cilastatin for ventilator-associated pneumonia. *Crit. Care* 2012;16(6):R218.

46. Calfee CS et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *Lancet Respir Med* 2014;2(8):611–620.

47. Walter JM, Helmin KA, Abdala-Valencia H, Wunderink RG, Singer BD. Multidimensional assessment of alveolar T cells in critically ill patients [Internet]. *JCI Insight* 2018;3(17). doi:10.1172/jci.insight.123287

48. Starren JB, Winter AQ, Lloyd-Jones DM. Enabling a Learning Health System through a Unified Enterprise Data Warehouse: The Experience of the Northwestern University Clinical and Translational Sciences (NUCATS) Institute. *Clin. Transl. Sci.* 2015;8(4):269–271.

49. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction [Internet]. *arXiv [stat.ML]* 2018;http://arxiv.org/abs/1802.03426. cited

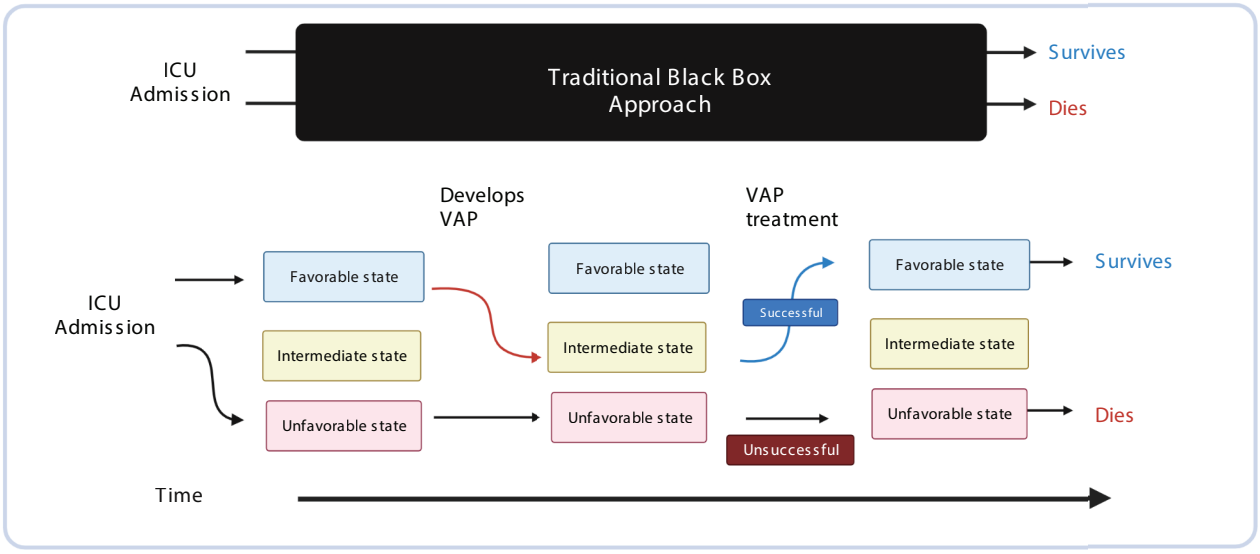
50. Johnson AE, Stone DJ, Celi LA, Pollard TJ. The MIMIC Code Repository: enabling reproducibility in critical care research. *J. Am. Med. Inform. Assoc.* 2018;25(1):32–39.

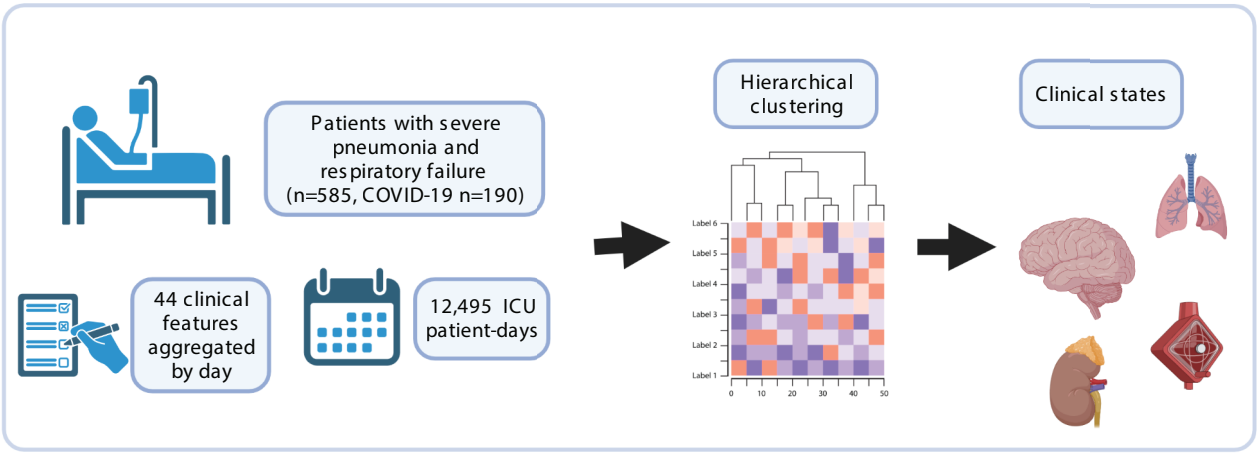
51. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: Association for Computing Machinery; 2016:785–794

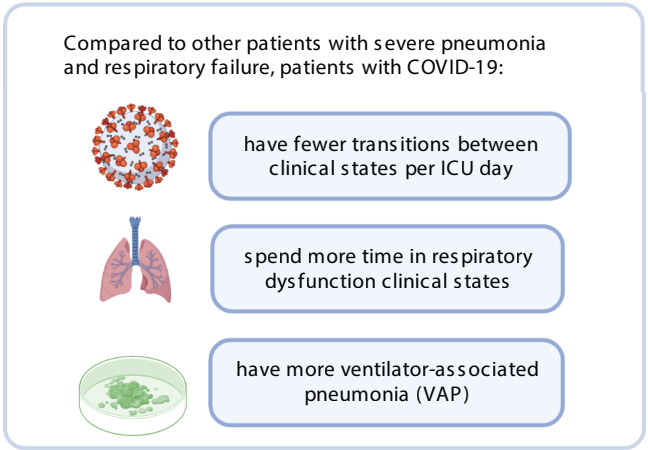
52. Pickens CI et al. An adjudication protocol for severe bacterial and viral pneumonia [Internet]. *bioRxiv* 2022; doi:10.1101/2022.10.26.22281461

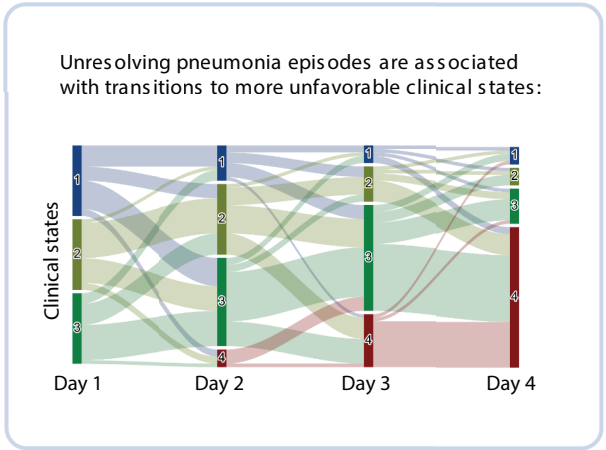
53. Goldberger AL et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000;101(23):E215-20.

54. Markov N et al. SCRIPT CarpeDiem Dataset: demographics, outcomes, and per-day clinical parameters for critically ill patients with suspected pneumonia [Internet]2023; doi:10.13026/5PHR-4R89









Graphical Abstract

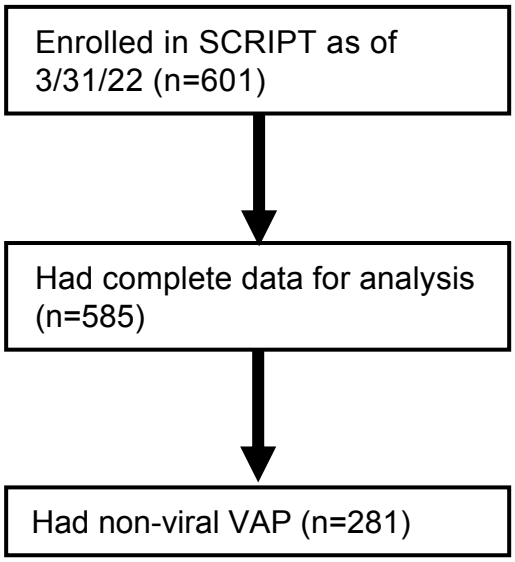


Figure 1. **CONSORT diagram of study participants and analysis.**
SCRIPT = Successful Clinical Response in Pneumonia Therapy study, VAP = ventilator-associated pneumonia.

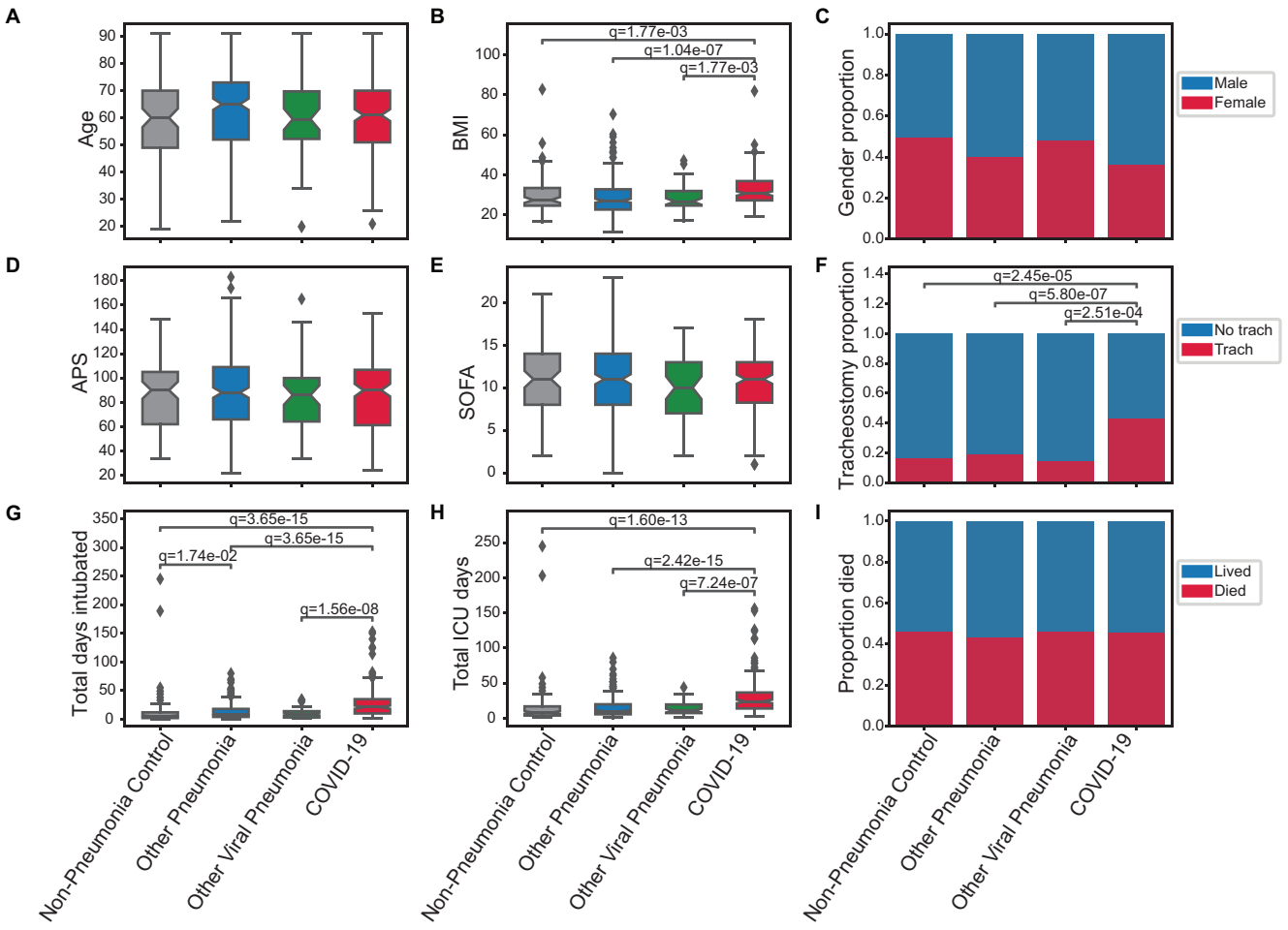


Figure 2. Demographics and outcomes of the cohort grouped by pneumonia category. Distribution of (**A**) patient age in years, (**B**) BMI in kg/m^2,* (**C**) gender, (**D**) APS, (**E**) SOFA score, (**F**) tracheostomy placement, (**G**) duration of intubation**, (**H**) length of ICU stay**, and (**I**) hospital mortality***.

* One patient did not have BMI data available.

** Total days intubated (G) and total ICU days (H) include only days at our hospital and do not capture intubation duration or ICU LOS at a transferring hospital.

*** Lived includes dispositions of discharge to home, acute inpatient rehabilitation, a long-term acute care hospital (LTACH), or a skilled nursing facility (see Supplemental Table 1). Died includes patients who died, underwent lung transplantation for refractory respiratory failure, or who were transferred to home or inpatient hospice.

BMI = body mass index, APS = Acute Physiology Score from APACHE IV (score calculated from worst value within the first two ICU days), SOFA = Sequential Organ Failure Assessment (score calculated from worst value within the first two ICU days).

Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Notches are bootstrapped 95% confidence interval of median. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. Categorical values were compared using Fisher's Exact tests with FDR correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

Numerical values and additional details are available in Supplemental Tables 1 and 2.

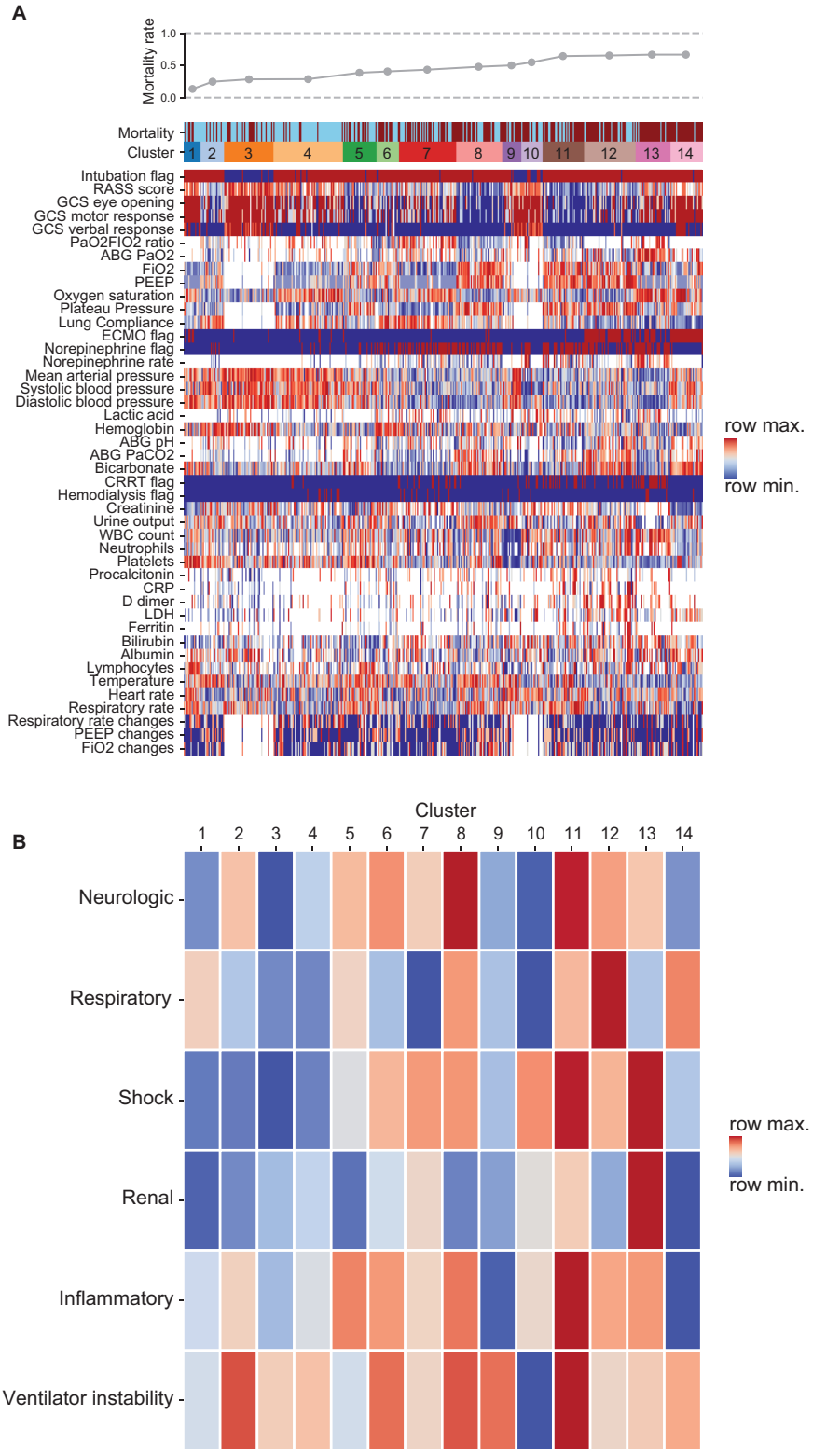


Figure 3. *CarpeDiem* groups patient-days into clusters representing clinical states associated with differential hospital mortality. (**A**) Heatmap of 44 clinical parameters with columns (representing 12,495 ICU patient-days from 585 patients) grouped into *CarpeDiem*-generated clusters (clinical states) ordered from lowest to highest mortality. Rows are sorted into physiologically related groups. The top row signifies hospital mortality of the patient shown in the column (blue = lived, red = died). The hospital mortality rate associated with each

cluster is shown above the heatmap. (**B**) Heatmap of the composite signal from each cluster and physiological group with ordering same as (A).

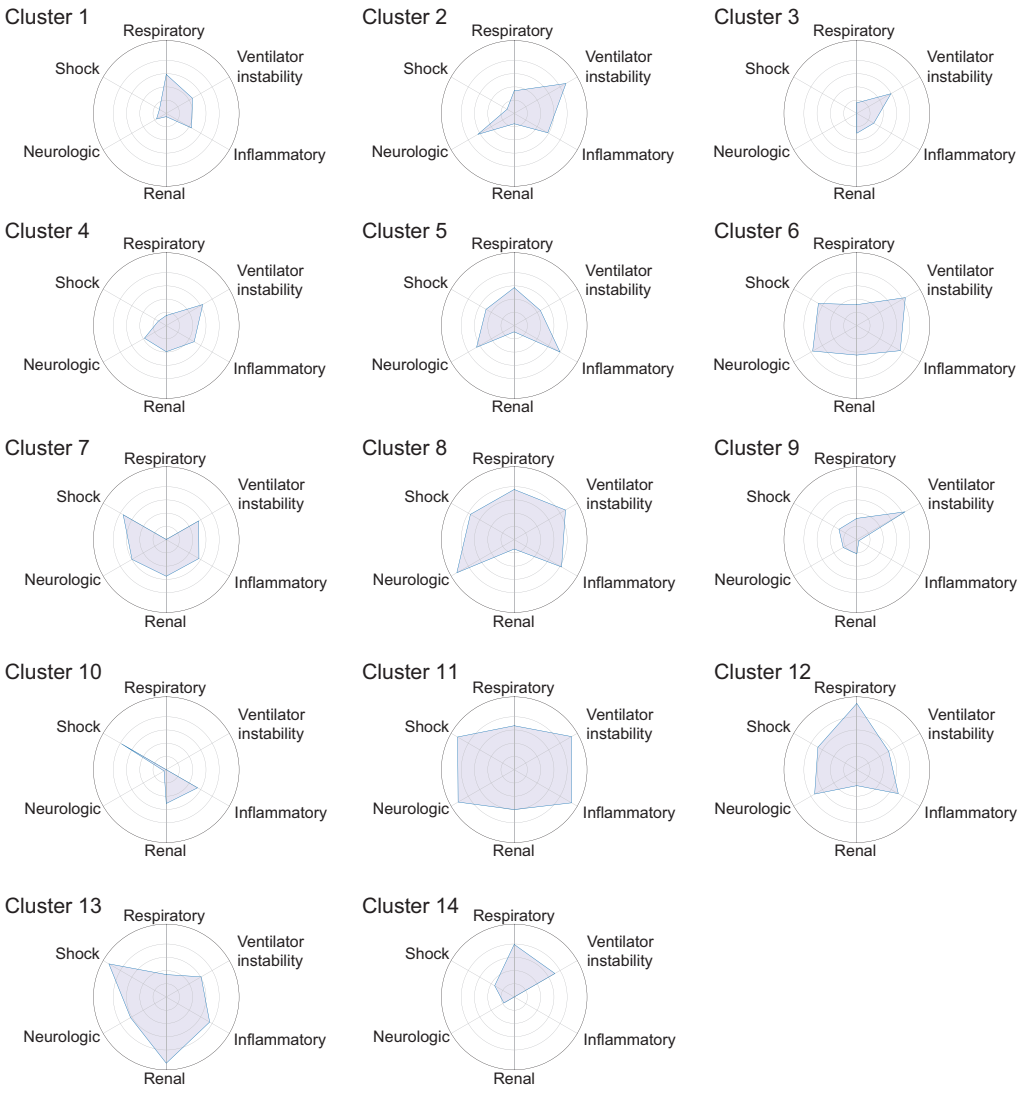


Figure 4. *CarpeDiem* clinical states have different patterns of organ dysfunction. Spider plots of min-max normalized composite features from Figure 3B for each clinical state. Circles indicate values of 0.2 (inner-most), 0.4, 0.6, 0.8, and 1 (outer-most).

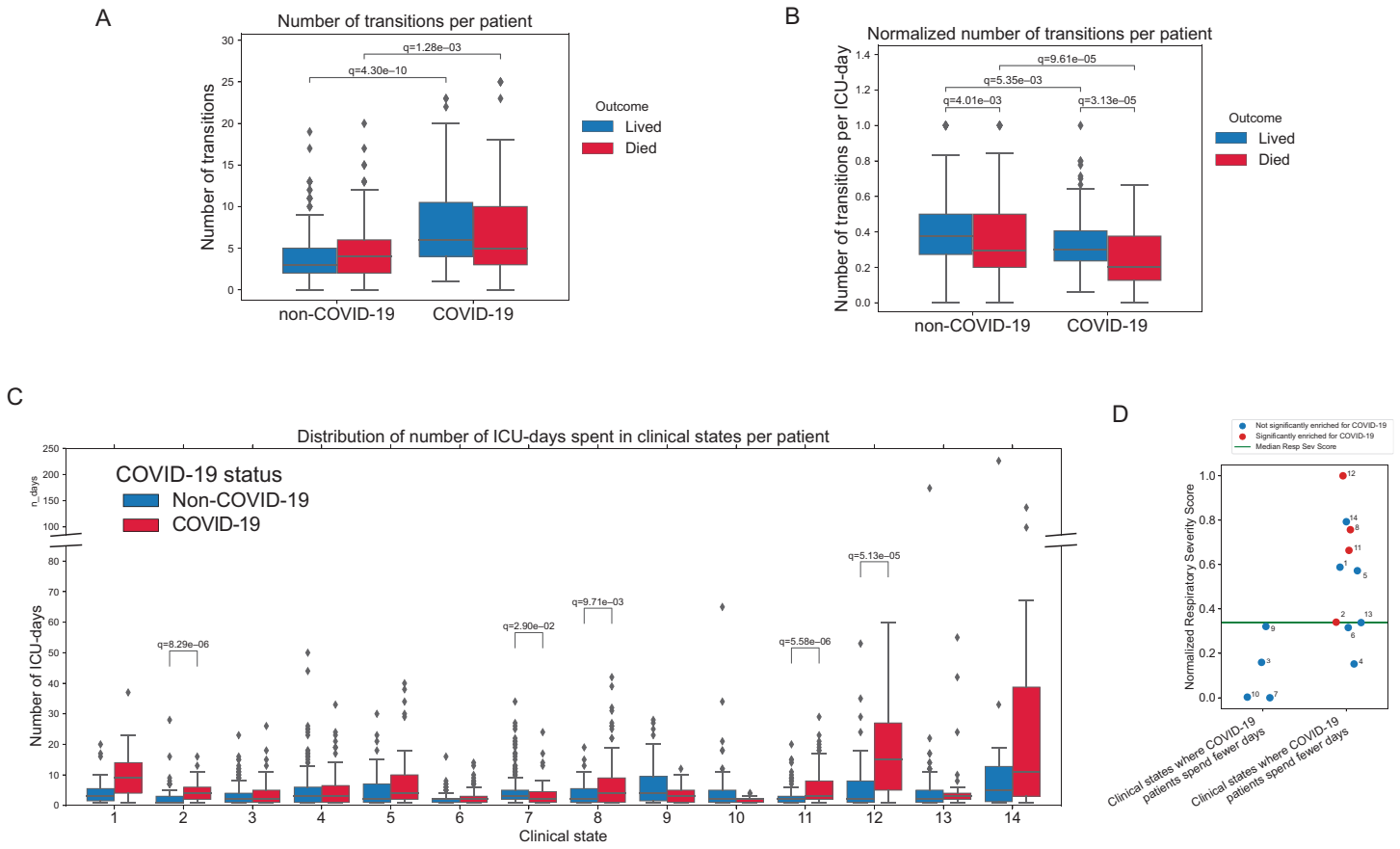


Figure 5. The long length of stay among patients with COVID-19 is driven by a lower frequency of transitions, resulting in longer durations of time spent in certain clinical states. (**A**) Distribution of transitions per patient. (**B**) Distribution of transitions normalized by ICU LOS. (**C**) Distribution of ICU-days spent in each clinical state per patient. Y-axis is discontinuous to accommodate all data points. (**D**) Respiratory severity score per clinical state, which are numbered next to each point, split by whether that cluster is enriched in patient-days from patients with COVID-19. The green line indicates the median respiratory severity score for the cohort. Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

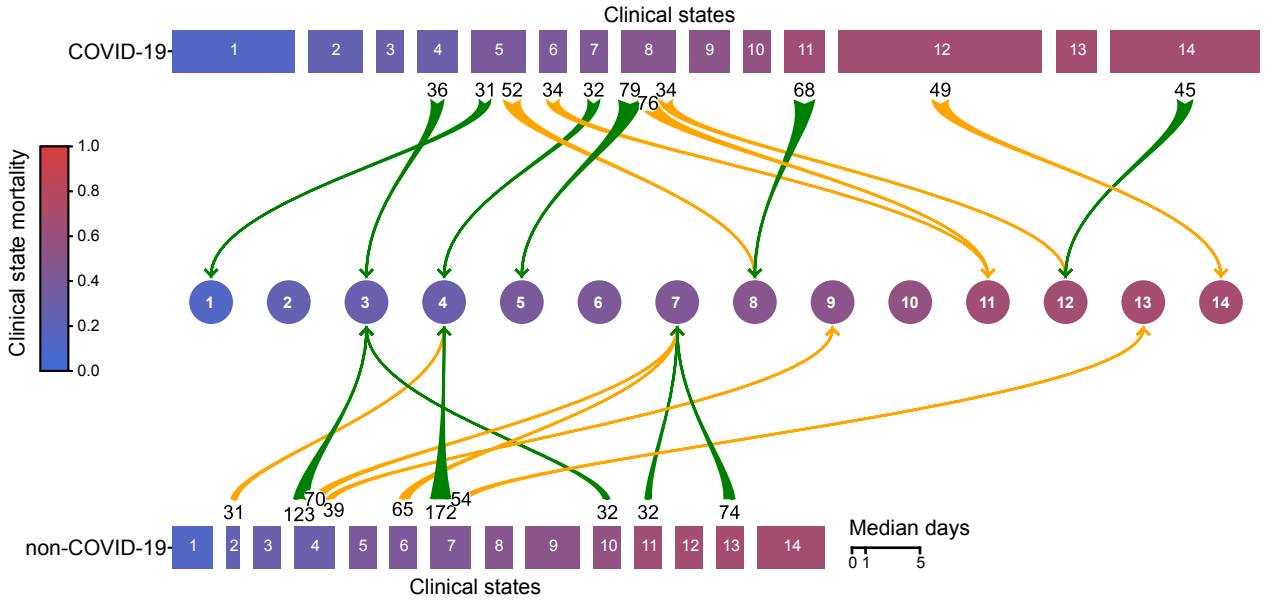


Figure 6. Patients with SARS-CoV-2 pneumonia have a longer length of stay and fewer transitions between clinical states per day compared to patients with non-COVID-19-related respiratory failure. Clinical states are ordered and numbered 1-14 by their associated mortality (blue to red). Rectangle width reflects median days per clinical state. Transitions marked by green arrows are to a more favorable (lower mortality) clinical state; yellow arrows mark transitions to a less favorable (higher mortality) clinical state. Numbers at the arrow bases represent the number of transitions between the two clinical states connected by the arrow. Only transitions that occurred more than 30 times are shown.

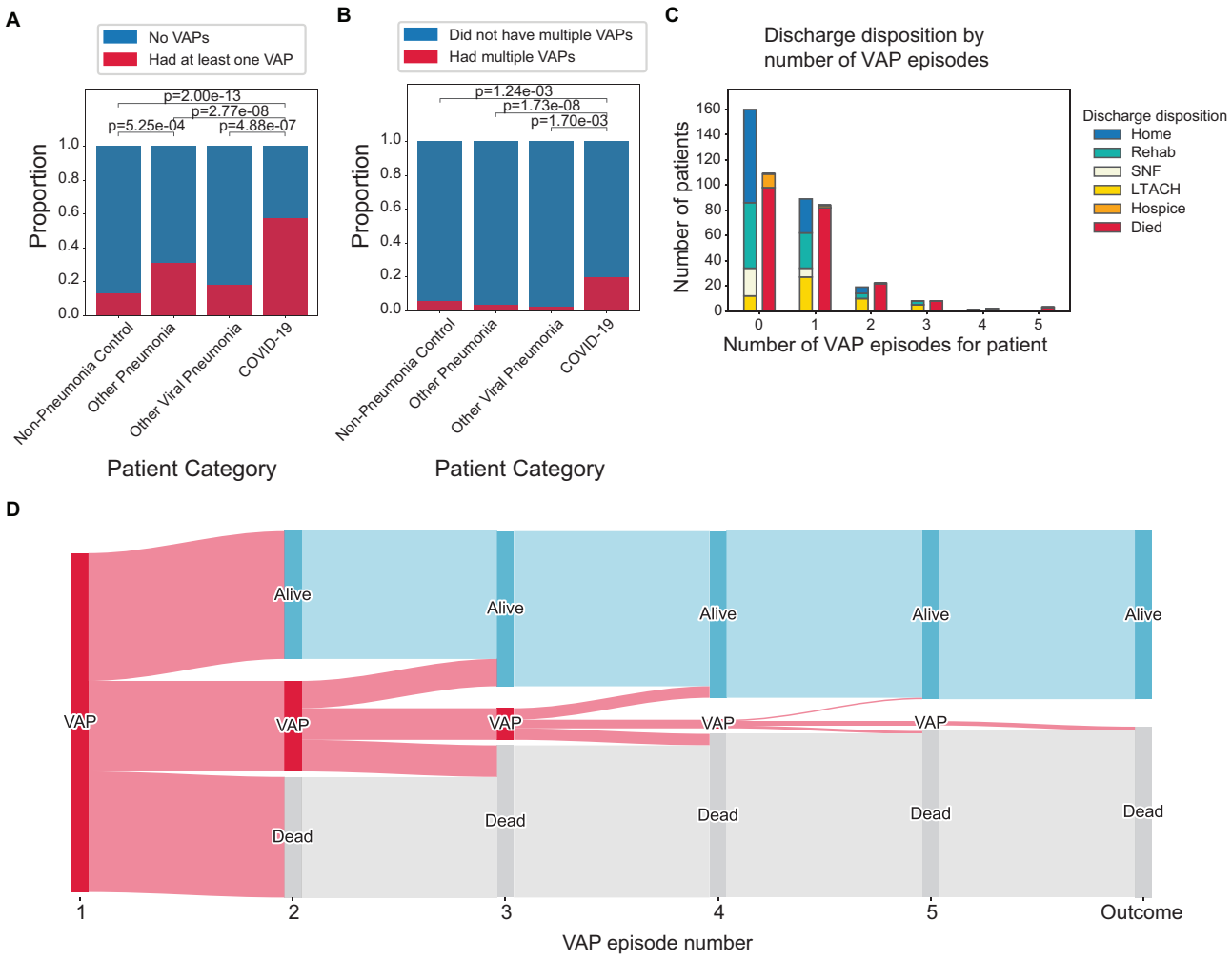


Figure 7. Patients with COVID-19 experience more VAP episodes compared to patients without COVID-19. (A) Proportion of patients having at least one VAP. (**B**) Proportion of patients having more than one VAP. (**C**) Outcomes for patients experiencing different numbers of VAP episodes. Outcomes are displayed in two columns: the first column aggregates favorable discharge dispositions (Home, Rehab, SNF, LTACH), the second column aggregates unfavorable discharge dispositions (Hospice, Died). (**D**) Sankey diagram of VAP episodes and outcomes for each VAP episode.
SNF = skilled nursing facility, LTACH = long-term acute care hospital.
Categorical values were compared using Fisher's Exact tests with FDR correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

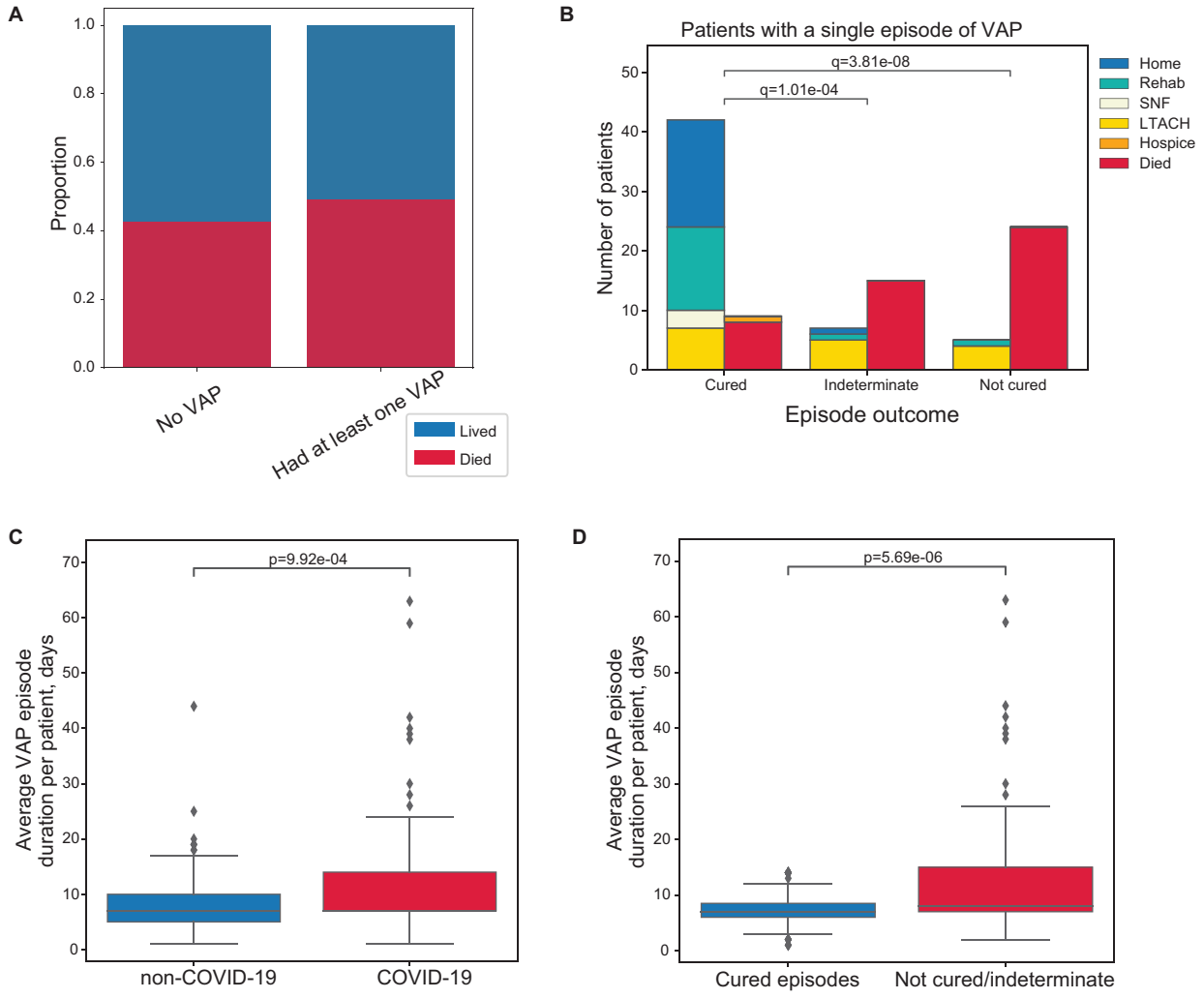


Figure 8. Unresolving VAP is associated with worse outcomes. (**A**) Mortality associated with having at least one episode of VAP. (**B**) Outcomes for patients who experienced one episode of VAP that was cured, of indeterminate cure status, or not cured by 14 days following diagnosis. Outcomes are displayed in two columns: the first column aggregates favorable discharge dispositions (Home, Rehab, SNF, LTACH), the second column aggregates unfavorable discharge dispositions (Hospice, Died). (**C**) VAP episode duration in patients with COVID-19 compared to patients without COVID-19. (**D**) VAP episode duration in patients who were cured or not cured/indeterminate cure status.
SNF = skilled nursing facility, LTACH = long-term acute care hospital.
Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. Categorical values were compared using Fisher's Exact tests with FDR correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

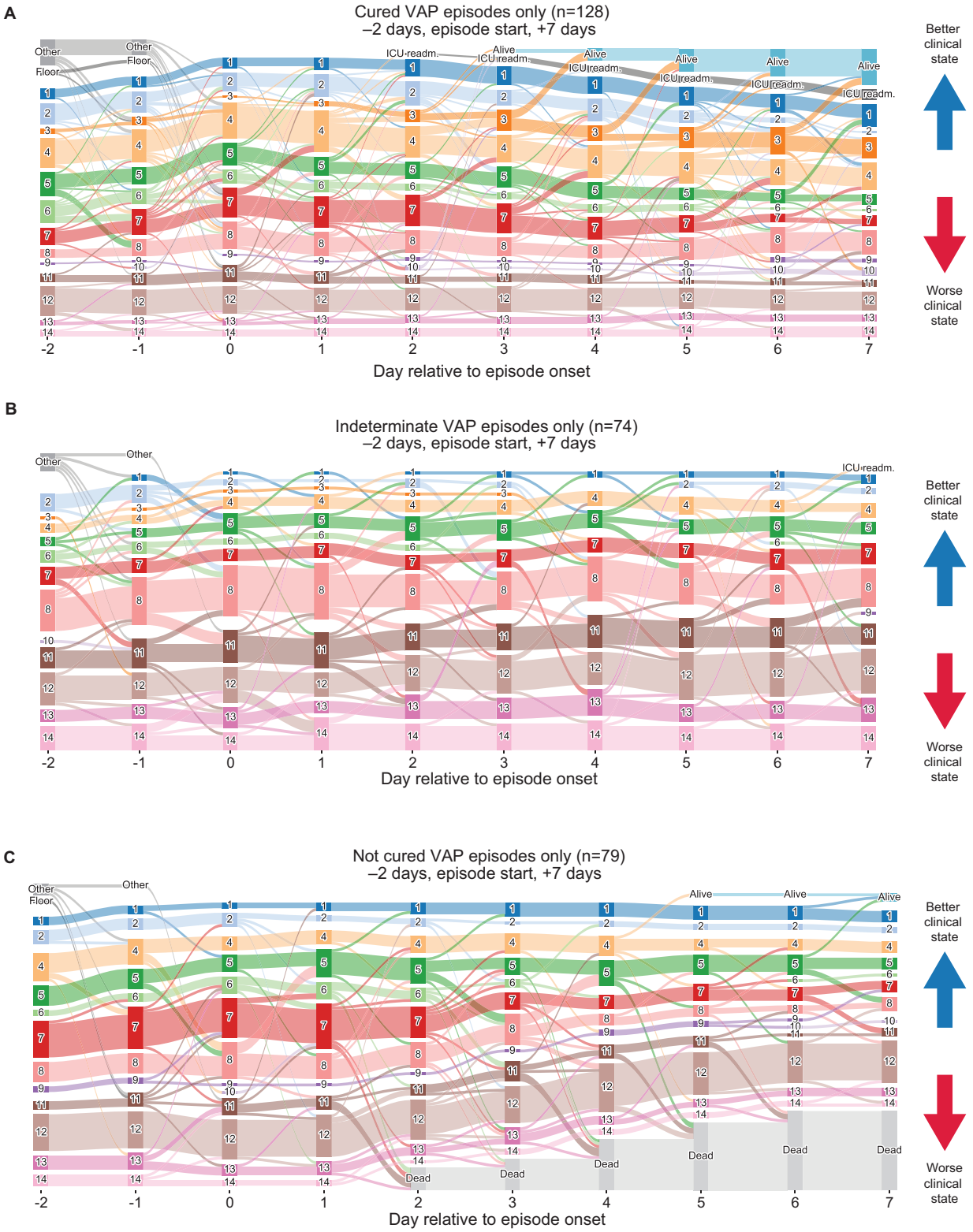


Figure 9. Trajectory analysis reveals that unresolving VAP is associated with transitions to progressively unfavorable clinical states. On these Sankey diagrams, day 0 represents the day that a BAL procedure was performed to evaluate VAP adjudicated as (**A**) cured, (**B**) indeterminate, or (**C**) not cured. More favorable (lower

mortality) clinical states are at the top of the graphs, with leaving the ICU alive being the highest, and less favorable (higher mortality) clinical states are at the bottom, with death being the lowest. Graphs start at two days prior to episode onset; patients who were not in our ICU are labeled as ‘Other’ (patients received in external transfer, chronically ventilated patients) or ‘Floor’ (within 48 hours of extubation, or chronically ventilated patients).

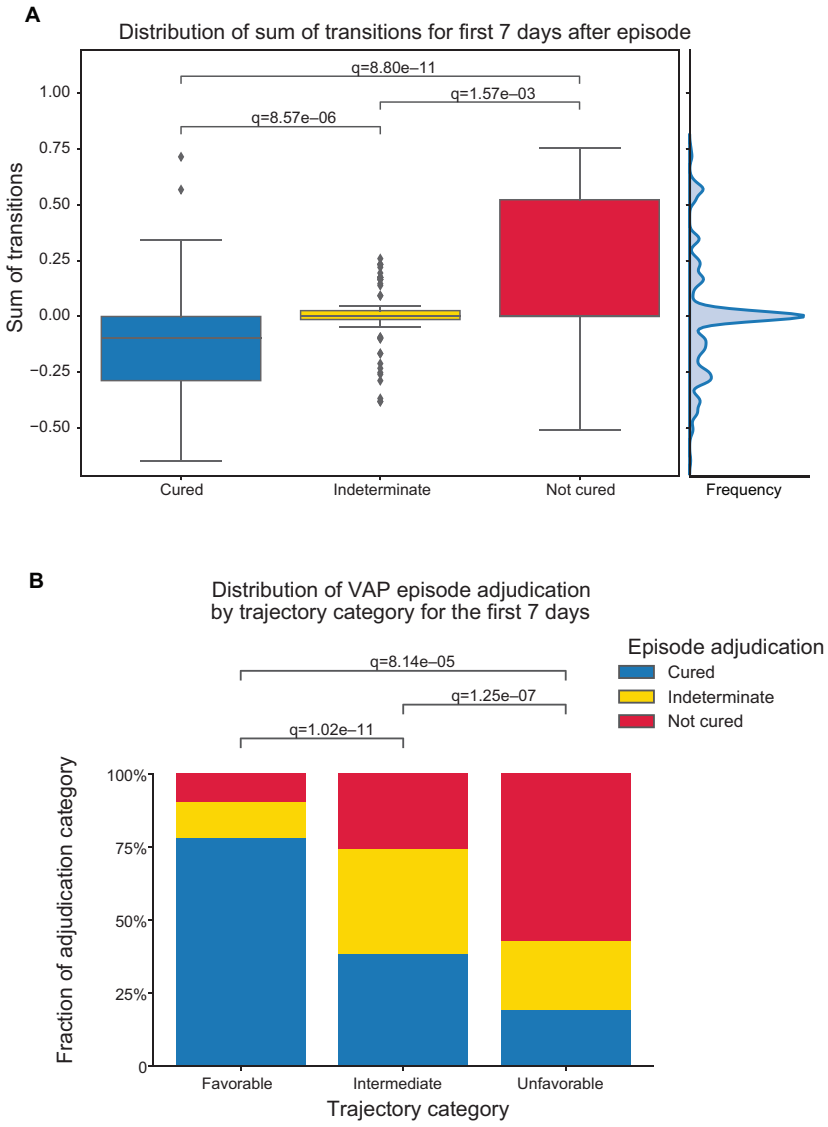


Figure 10. Unresolving VAP episodes are associated with unfavorable clinical states. (A) Distribution of the sum of transitions for the seven days following VAP diagnosis by episode outcome, identifying a breakpoint of 0.1 in the middle of the distribution (shown by the cumulative data histogram along the right axis). Higher sums of transitions reflect transitions to unfavorable (higher mortality) clusters. (**B**) Proportion of VAP episode outcomes in each trajectory category. Trajectories were grouped into favorable (sum of transitions < -0.1), indeterminate (-0.1–0.1), and unfavorable (>0.1) categories.
Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. Categorical values were compared using Chi-squared tests with FDR correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

Supplemental Table 1. Demographics and outcomes data for the cohort, grouped by pneumonia category.

[A] Racial groups with fewer than five individuals were classified as 'Unknown or Not Reported' to protect patient anonymity.

[B] One patient did not have BMI documented.

[C] Total days intubated and total ICU days include only days at our hospital and do not capture intubation duration or ICU LOS at a transferring hospital.

[D] Died included those who died or underwent lung transplantation for refractory respiratory failure.

BMI = body mass index, APS = Acute Physiology Score (score calculated from worst value within the first two ICU days), SOFA = Sequential Organ Failure Assessment (score calculated from worst value within the first two ICU days).

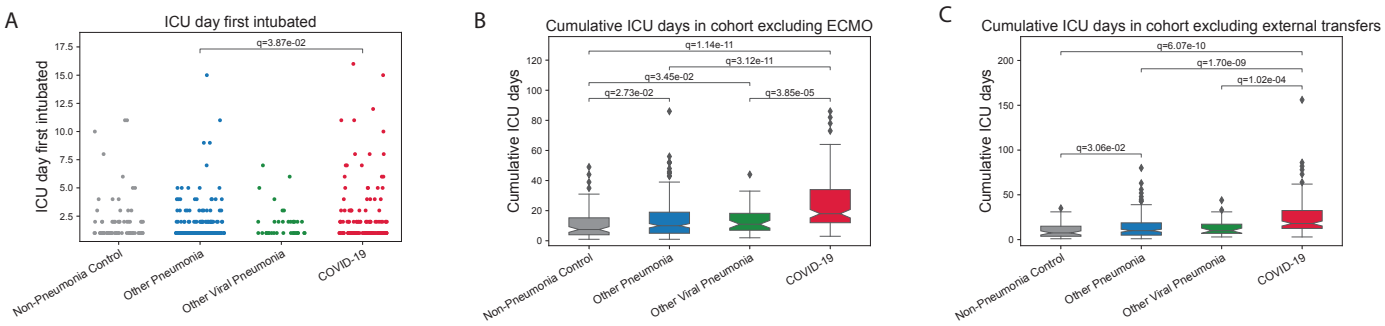
Feature	Overall	Non-Pneumonia Control	Other Pneumonia	Other Viral Pneumonia	COVID-19
n	585	93	252	50	190
Age, median [Q1,Q3]	62.0 [51.0,72.0]	60.0 [49.0,70.0]	65.0 [52.0,73.0]	59.5 [52.2,69.8]	61.0 [51.0,70.0]
Ethnicity, n (%)					
Hispanic or Latino	122 (20.9)	12 (12.9)	24 (9.5)	10 (20.0)	76 (40.0)
Not Hispanic or Latino	438 (74.9)	77 (82.8)	218 (86.5)	37 (74.0)	106 (55.8)
Unknown or Not Reported	25 (4.3)	4 (4.3)	10 (4.0)	3 (6.0)	8 (4.2)
Gender, n (%)					
Female	239 (40.9)	46 (49.5)	100 (39.7)	24 (48.0)	69 (36.3)
Male	346 (59.1)	47 (50.5)	152 (60.3)	26 (52.0)	121 (63.7)
Race, n (%) [A]					
Asian	17 (2.9)	3 (3.2)	7 (2.8)	3 (6.0)	4 (2.1)
Black or African American	119 (20.3)	17 (18.3)	55 (21.8)	8 (16.0)	39 (20.5)
Unknown or Not Reported	105 (17.9)	14 (15.1)	33 (13.1)	5 (10.0)	53 (27.9)
White	344 (58.8)	59 (63.4)	157 (62.3)	34 (68.0)	94 (49.5)
Smoking status, n (%)					
Current Smoker	48 (8.2)	9 (9.7)	30 (11.9)	5 (10.0)	4 (2.1)
Never Smoker	230 (39.3)	40 (43.0)	94 (37.3)	22 (44.0)	74 (38.9)
Past Smoker	150 (25.6)	25 (26.9)	78 (31.0)	19 (38.0)	28 (14.7)
Unknown Smoking Status	157 (26.8)	19 (20.4)	50 (19.8)	4 (8.0)	84 (44.2)
BMI, median [Q1,Q3] [B]	28.7 [24.6,34.1]	27.4 [24.6,33.4]	27.0 [22.6,32.7]	26.6 [24.6,31.9]	30.6 [27.2,36.9]
Admit APS score, median [Q1,Q3]	89.0 [64.0,107.0]	90.0 [62.0,105.0]	88.0 [66.0,109.0]	86.0 [64.2,100.0]	90.0 [61.2,106.8]
Admit SOFA score, median [Q1,Q3]	11.0 [8.0,13.0]	11.0 [8.0,14.0]	11.0 [8.0,14.0]	10.0 [7.0,13.0]	11.0 [8.2,13.0]
Cumulative ICU days, median [Q1,Q3] [C]	14.0 [6.0,26.0]	8.0 [4.0,17.0]	10.0 [5.8,20.0]	11.0 [7.5,19.8]	24.0 [14.0,36.8]
Number of ICU stays, median [Q1,Q3]	1.0 [1.0,1.0]	1.0 [1.0,1.0]	1.0 [1.0,2.0]	1.0 [1.0,2.0]	1.0 [1.0,1.0]
Cumulative pred equivalents during admission, median [Q1,Q3]	150.0 [0.0,390.0]	130.0 [0.0,720.0]	114.0 [0.0,300.0]	118.5 [26.5,310.0]	240.0 [26.2,437.5]
Received steroids during admission, n (%)	409 (69.9)	64 (68.8)	160 (63.5)	38 (76.0)	147 (77.4)

Received tocilizumab during admission, n (%)	17 (2.9)	-	1 (0.4)	-	16 (8.4)
Received sarilumab during admission, n (%)	1 (0.2)	-	-	-	1 (0.5)
Sarilumab study drug during admission, n (%)	12 (2.1)	-	-	-	12 (6.3)
Received remdesivir during admission, n (%)	76 (13.0)	-	4 (1.6)	-	72 (37.9)
Remdesivir study drug during admission, n (%)	11 (1.9)	-	-	-	11 (5.8)
Congestive heart failure, n (%)	170 (29.1)	37 (39.8)	89 (35.3)	10 (20.0)	34 (17.9)
Peripheral vascular disease, n (%)	121 (20.7)	19 (20.4)	62 (24.6)	12 (24.0)	28 (14.7)
Cerebrovascular disease, n (%)	123 (21.0)	21 (22.6)	61 (24.2)	14 (28.0)	27 (14.2)
Chronic pulmonary disease, n (%)	209 (35.7)	33 (35.5)	97 (38.5)	21 (42.0)	58 (30.5)
Peptic ulcer disease, n (%)	53 (9.1)	10 (10.8)	25 (9.9)	8 (16.0)	10 (5.3)
Liver disease, n (%)	150 (25.6)	24 (25.8)	76 (30.2)	18 (36.0)	32 (16.8)
Diabetes, n (%)	208 (35.6)	25 (26.9)	85 (33.7)	24 (48.0)	74 (38.9)
Renal disease, n (%)	158 (27.0)	33 (35.5)	77 (30.6)	14 (28.0)	34 (17.9)
Cancer, n (%)	196 (33.5)	42 (45.2)	90 (35.7)	25 (50.0)	39 (20.5)
Immunocompromised flag, n (%)	162 (27.7)	31 (33.3)	78 (31.0)	22 (44.0)	31 (16.3)
Tracheostomy flag, n (%)	151 (25.8)	15 (16.1)	48 (19.0)	7 (14.0)	81 (42.6)
Cumulative intubation days, median [Q1,Q3]	10.0 [4.0,23.0]	5.0 [2.0,12.0]	8.0 [4.0,18.0]	9.0 [3.0,14.0]	21.0 [10.0,35.0]
Discharge disposition, n (%)					
Died [D]	243 (41.5)	37 (39.8)	99 (39.3)	20 (40.0)	87 (45.8)
Home	133 (22.7)	27 (29.0)	49 (19.4)	10 (20.0)	47 (24.7)
LTACH	59 (10.1)	6 (6.5)	27 (10.7)	4 (8.0)	22 (11.6)
Rehab	97 (16.6)	11 (11.8)	48 (19.0)	12 (24.0)	26 (13.7)
SNF	34 (5.8)	6 (6.5)	19 (7.5)	1 (2.0)	8 (4.2)
Hospice	19 (3.2)	6 (6.5)	10 (4.0)	3 (6.0)	-

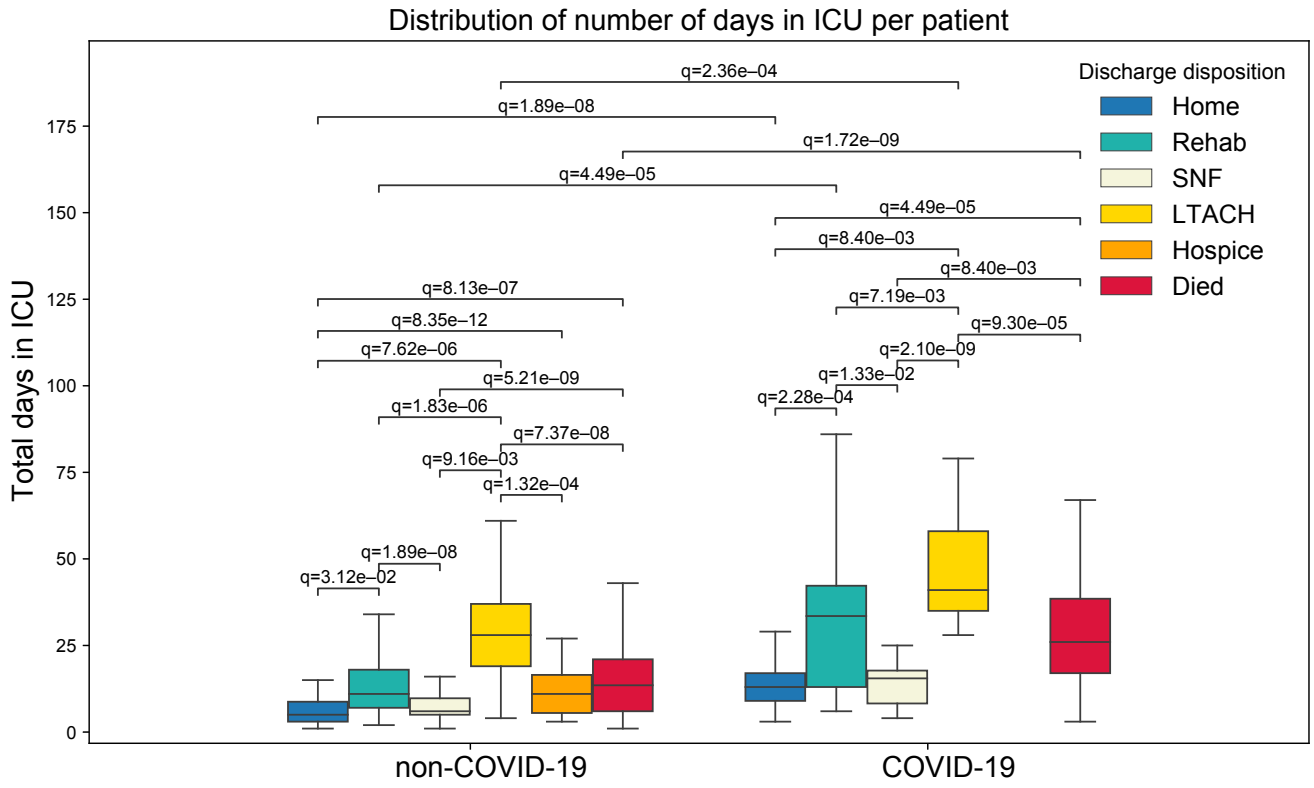
Supplemental Table 2. **Descriptive features from first day of intubation in patients, excluding patients received in external transfer who were intubated on ICU day 1, grouped by pneumonia category.** Of our cohort of 585 patients, 184 patients (31.4% of the cohort) were received in transfer from another hospital. Of these patients, 139 (75.5%) were intubated at time of transfer or during the first day in our ICU and are excluded from this table.

Feature	Non-Pneumonia Control	Other Pneumonia	Other Viral Pneumonia	COVID-19
n	72	200	36	138
ICU day, median [Q1,Q3]	1.0 [1.0,2.0]	1.0 [1.0,2.0]	1.0 [1.0,2.0]	1.0 [1.0,2.0]
SOFA score, median [Q1,Q3]	12.0 [9.0,14.2]	12.0 [9.0,14.0]	11.5 [7.8,15.0]	11.0 [9.0,13.0]
ECMO flag, n (%)	1 (1.4)	1 (0.5)	-	-
Intubation flag, n (%)	72 (100.0)	200 (100.0)	36 (100.0)	138 (100.0)
Hemodialysis flag, n (%)	3 (4.2)	5 (2.5)	1 (2.8)	-
CRRT flag, n (%)	6 (8.3)	18 (9.0)	5 (13.9)	5 (3.6)
Temperature, median [Q1,Q3]	98.2 [97.6,99.5]	98.3 [97.5,99.4]	98.7 [98.0,99.7]	98.9 [98.3,99.8]
Heart rate, median [Q1,Q3]	88.7 [77.3,107.8]	95.2 [82.0,109.9]	99.8 [86.8,112.9]	85.1 [75.5,96.6]
Systolic blood pressure, median [Q1,Q3]	112.2 [104.3,130.8]	112.8 [104.1,124.7]	111.7 [106.6,124.9]	117.9 [110.6,126.4]
Diastolic blood pressure, median [Q1,Q3]	61.1 [56.1,68.8]	59.8 [55.4,67.7]	59.6 [54.8,65.8]	62.9 [57.4,68.1]
Mean arterial pressure, median [Q1,Q3]	61.0 [56.0,68.0]	60.0 [53.0,65.0]	59.5 [55.0,66.2]	61.0 [57.2,66.0]
Norepinephrine rate, median [Q1,Q3]	0.2 [0.1,0.3]	0.2 [0.1,0.3]	0.2 [0.1,0.2]	0.1 [0.1,0.2]
Norepinephrine flag, n (%)	46 (63.9)	126 (63.0)	20 (55.6)	98 (71.0)
Respiratory rate, median [Q1,Q3]	22.6 [19.3,26.5]	22.3 [19.8,26.1]	23.7 [20.1,27.3]	25.0 [22.0,28.1]
Oxygen saturation, median [Q1,Q3]	96.5 [94.9,98.1]	96.7 [95.2,98.0]	96.6 [94.9,98.7]	94.8 [93.6,96.1]
Urine output, median [Q1,Q3]	857.5 [461.2,1558.8]	647.5 [244.0,1293.2]	800.0 [190.0,1237.5]	900.0 [483.0,1400.0]
GCS eye opening, median [Q1,Q3]	1.0 [1.0,3.0]	1.0 [1.0,3.0]	1.0 [1.0,2.0]	1.0 [1.0,3.0]
GCS motor response, median [Q1,Q3]	4.0 [1.0,5.0]	4.0 [1.0,5.0]	4.0 [1.0,4.0]	2.0 [1.0,5.0]
GCS verbal response, median [Q1,Q3]	1.0 [1.0,1.0]	1.0 [1.0,1.0]	1.0 [1.0,1.0]	1.0 [1.0,1.0]
RASS score, median [Q1,Q3]	-2.0 [-4.0,-1.0]	-2.0 [-3.5,-1.0]	-3.0 [-4.0,-2.0]	-3.0 [-4.0,-2.0]
PEEP, median [Q1,Q3]	5.0 [5.0,8.1]	5.0 [5.0,8.9]	5.0 [5.0,7.8]	11.0 [10.0,14.0]
FiO2, median [Q1,Q3]	50.0 [45.0,70.5]	59.0 [45.0,75.0]	58.8 [46.7,70.0]	73.3 [60.6,85.0]
Plateau Pressure, median [Q1,Q3]	22.4 [18.8,25.6]	21.5 [17.5,26.3]	20.8 [16.8,23.2]	24.5 [21.8,28.7]
Lung Compliance, median [Q1,Q3]	31.0 [21.5,38.0]	33.5 [25.0,40.5]	30.0 [25.2,40.8]	34.5 [25.8,43.7]
PEEP changes, median [Q1,Q3]	1.0 [1.0,2.0]	1.0 [1.0,2.0]	1.0 [1.0,2.0]	2.0 [1.0,2.0]
Respiratory rate changes, median [Q1,Q3]	2.0 [1.0,3.0]	2.0 [1.0,2.0]	2.0 [1.0,2.0]	2.0 [1.0,2.0]

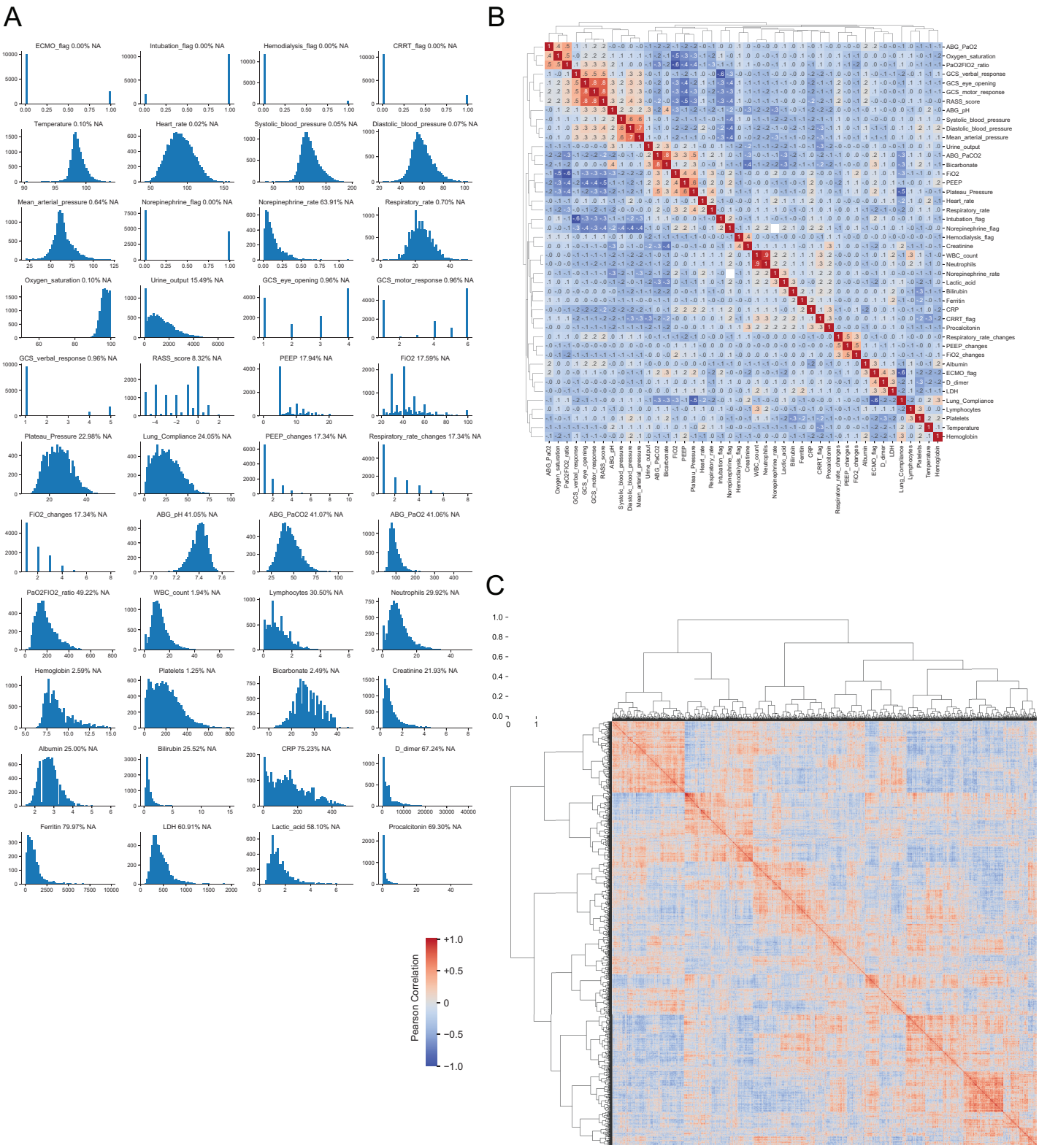
FiO2 changes, median [Q1,Q3]	2.0 [1.8,3.0]	2.0 [2.0,3.0]	2.0 [2.0,3.0]	2.0 [2.0,3.0]
ABG pH, median [Q1,Q3]	7.4 [7.3,7.4]	7.3 [7.3,7.4]	7.3 [7.3,7.4]	7.4 [7.3,7.4]
ABG PaCO2, median [Q1,Q3]	37.7 [30.8,42.0]	39.5 [32.9,46.2]	37.0 [33.9,47.7]	41.0 [37.0,45.0]
ABG PaO2, median [Q1,Q3]	103.8 [85.1,126.4]	106.1 [86.1,131.5]	108.7 [85.0,125.0]	99.1 [84.8,116.2]
PaO2FIO2 ratio, median [Q1,Q3]	140.0 [91.0,237.1]	136.0 [90.0,200.2]	129.5 [99.5,228.2]	98.6 [71.1,126.6]
WBC count, median [Q1,Q3]	10.7 [7.0,19.2]	13.0 [8.0,17.0]	9.6 [6.7,14.8]	9.8 [6.7,12.5]
Lymphocytes, median [Q1,Q3]	0.5 [0.3,1.1]	0.9 [0.5,1.6]	0.8 [0.4,1.6]	0.8 [0.6,1.3]
Neutrophils, median [Q1,Q3]	7.6 [3.4,14.1]	9.7 [5.2,14.4]	7.8 [3.3,12.7]	7.8 [5.4,10.4]
Hemoglobin, median [Q1,Q3]	9.8 [8.1,11.1]	9.8 [8.2,11.8]	8.4 [7.5,10.4]	12.1 [10.9,13.2]
Platelets, median [Q1,Q3]	129.0 [54.5,223.0]	184.0 [103.5,269.8]	196.0 [61.5,274.5]	215.5 [155.8,304.2]
Bicarbonate, median [Q1,Q3]	23.0 [19.7,26.1]	22.2 [19.0,25.0]	23.2 [20.4,25.0]	23.0 [21.0,26.0]
Creatinine, median [Q1,Q3]	1.3 [0.9,2.0]	1.4 [0.8,1.9]	1.1 [0.7,1.7]	0.9 [0.8,1.4]
Albumin, median [Q1,Q3]	3.0 [2.6,3.5]	3.1 [2.7,3.6]	2.6 [2.3,3.1]	3.3 [3.0,3.5]
Bilirubin, median [Q1,Q3]	0.9 [0.6,1.7]	0.8 [0.6,1.5]	0.7 [0.5,1.0]	0.6 [0.5,0.8]
CRP, median [Q1,Q3]	137.8 [57.0,281.0]	54.2 [8.2,95.3]	-	165.0 [93.2,243.3]
D dimer, median [Q1,Q3]	1382.0 [468.0,3222.0]	2230.5 [483.2,6851.1]	1564.0 [876.0,1823.0]	570.0 [290.2,2086.5]
Ferritin, median [Q1,Q3]	387.8 [159.5,3622.1]	535.5 [105.5,1143.9]	746.5 [465.7,1027.2]	729.7 [423.8,1178.0]
LDH, median [Q1,Q3]	378.0 [248.0,532.8]	332.0 [251.5,541.2]	299.0 [187.5,330.0]	454.0 [347.2,583.5]
Lactic acid, median [Q1,Q3]	2.2 [1.4,3.0]	1.9 [1.3,3.1]	1.7 [1.3,2.0]	1.4 [1.2,1.9]
Procalcitonin, median [Q1,Q3]	0.7 [0.1,2.3]	0.6 [0.1,2.3]	1.1 [0.2,9.6]	0.3 [0.2,1.0]



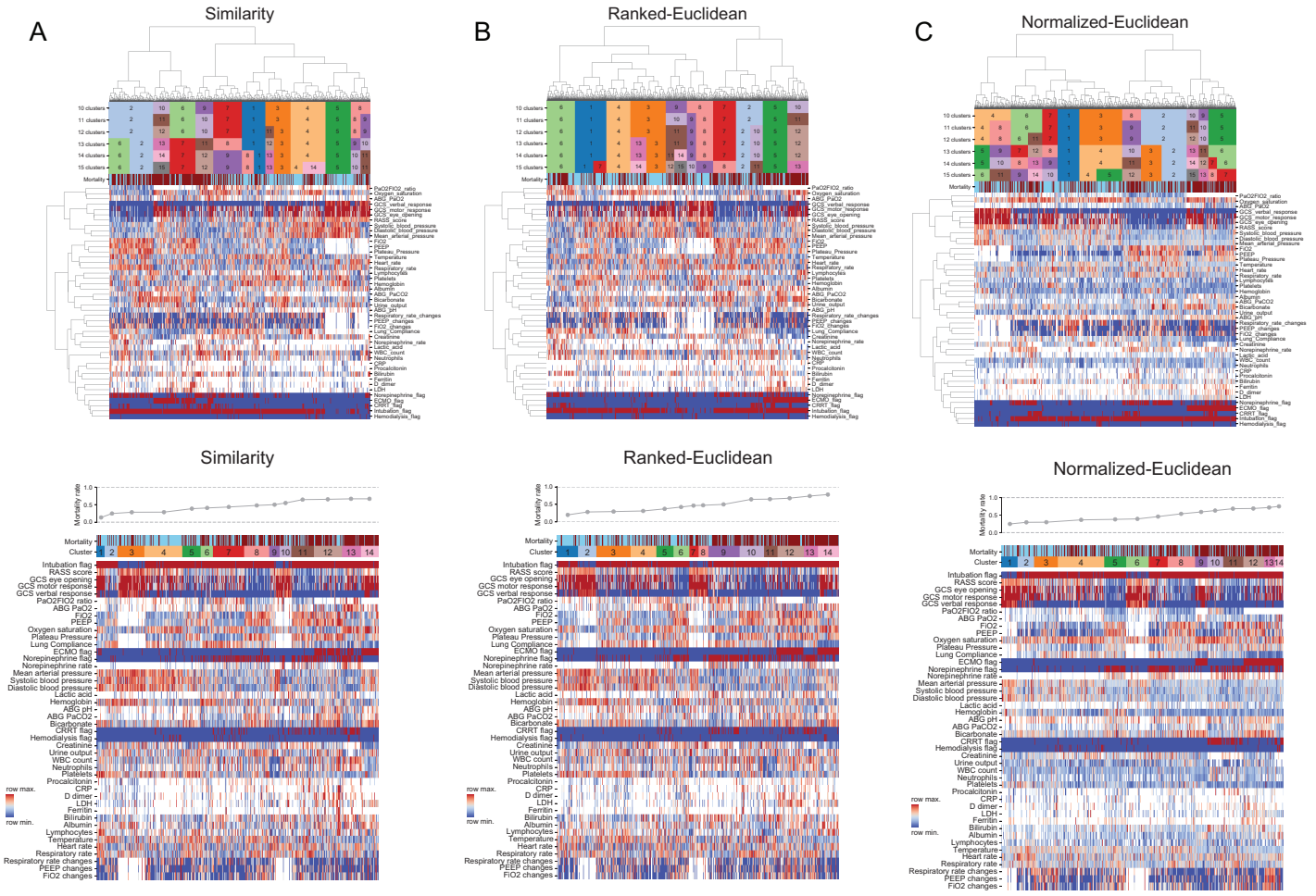
Supplemental Figure 1. Additional demographics and outcomes of the cohort, grouped by pneumonia category. (**A**) ICU day of first intubation, excluding patients received in external transfer (total patients received in external transfer: 184 patients, 31.4% of the cohort), who were intubated on ICU day 1 at our hospital. 139 patients (75.5%) of the patients received in external transfer were intubated at the time of transfer or during the first day in our ICU. (**B**) Cumulative ICU days excluding patients who received ECMO support. Data include only days at our hospital and do not capture ICU LOS from a transferring hospital. (**C**) Cumulative ICU days excluding patients received in external transfer. Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Notches are bootstrapped 95% confidence interval of median. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.



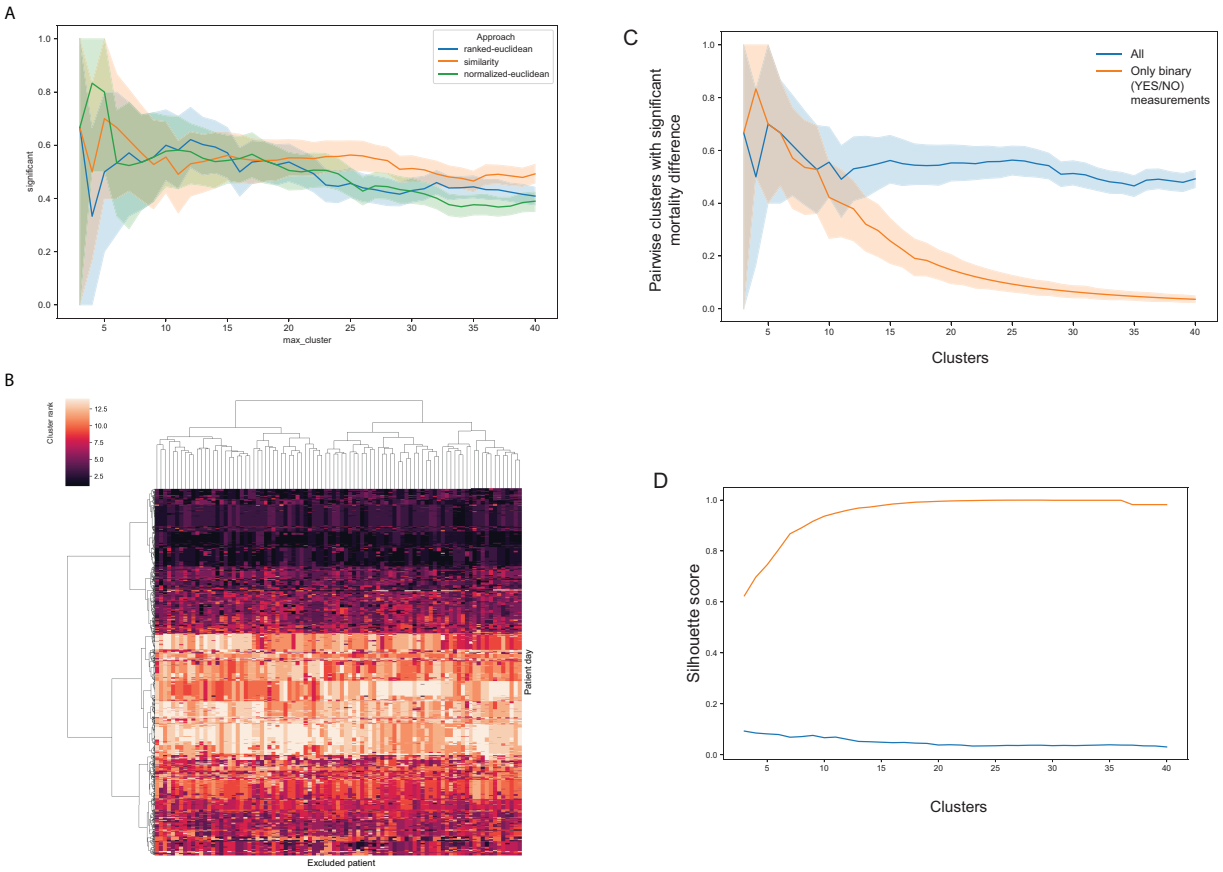
Supplemental Figure 2. Length of stay by discharge disposition among patients with and without COVID-19. Patients with severe SARS-CoV-2 pneumonia experienced significantly longer ICU LOS for all discharge disposition groups, except Skilled Nursing Facility (SNF; not significantly different) and Hospice (no patients in COVID-19 group). No patient was transferred back to their referring center intubated. Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.



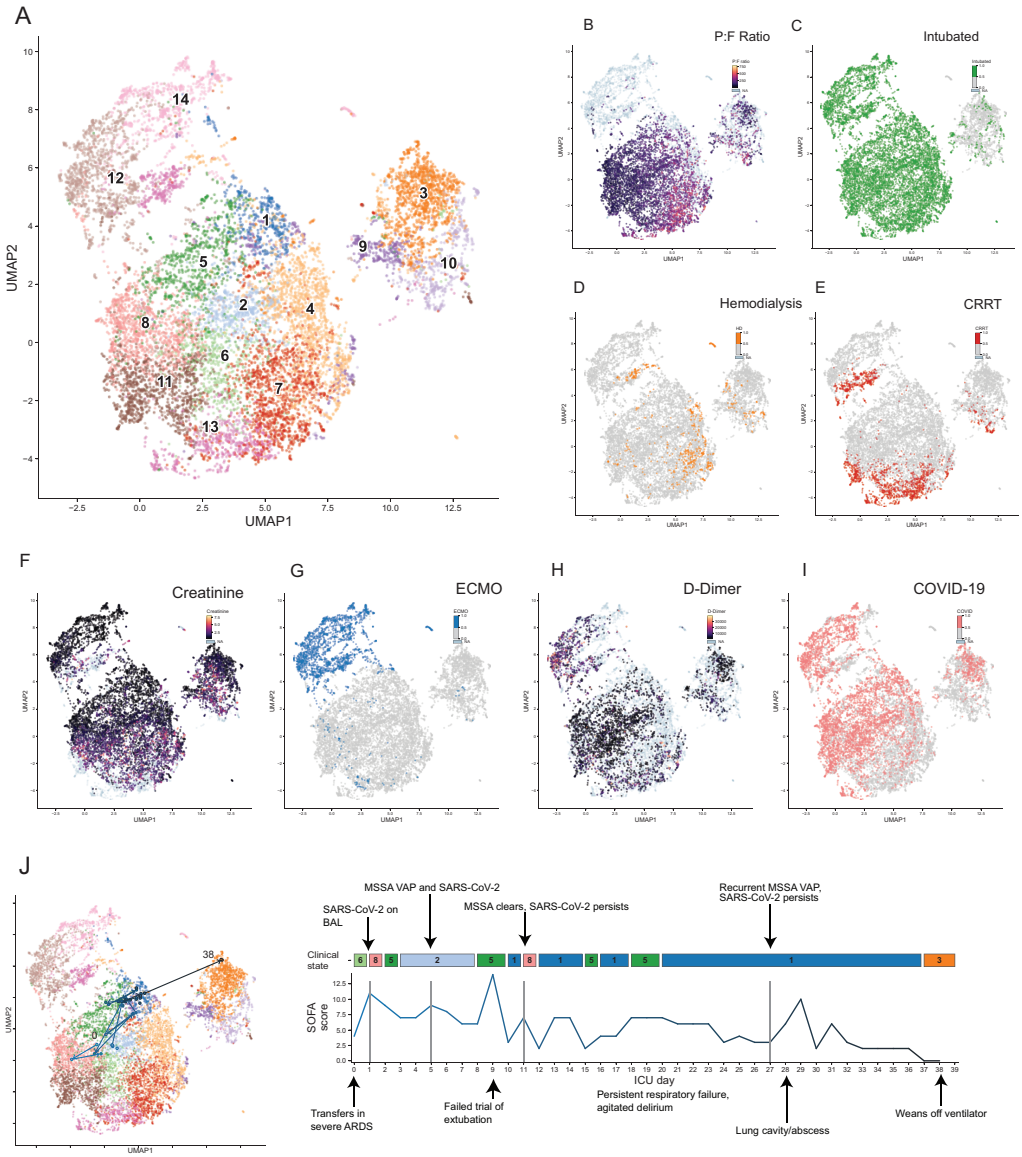
Supplemental Figure 3. Clinical parameters and their correlation. (**A**) Histogram display of 44 clinical values extracted from the EHR. (**B**) Pearson correlation matrix of clinical parameters; some measurements in our dataset displayed a high correlation due to mathematical or physiological coupling (e.g., plateau pressure, PEEP and lung compliance; $PaCO_2$ and bicarbonate). (**C**) Pearson correlation between different patient-days as used as an intermediate step for the “Similarity” clustering strategy.



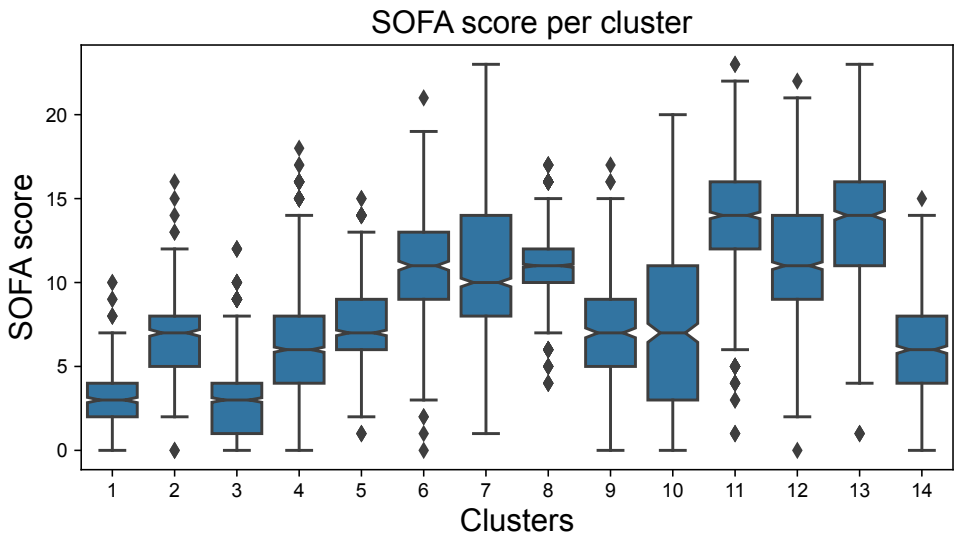
Supplemental Figure 4. Comparing three different clustering strategies. (**A**) Similarity, (**B**) Ranked-Euclidean, and (**C**) Normalized-Euclidean strategies. For each method, hierarchical clustering of clinical parameters (rows) and columns (patient-days) is shown on the top, grouping patient-days into 10-15 separate clusters. The bottom panels show re-ordered clustering with columns organized into clusters and sorted by ascending cluster mortality and rows organized into physiologically similar groups. Cluster mortality is shown above the heatmaps.



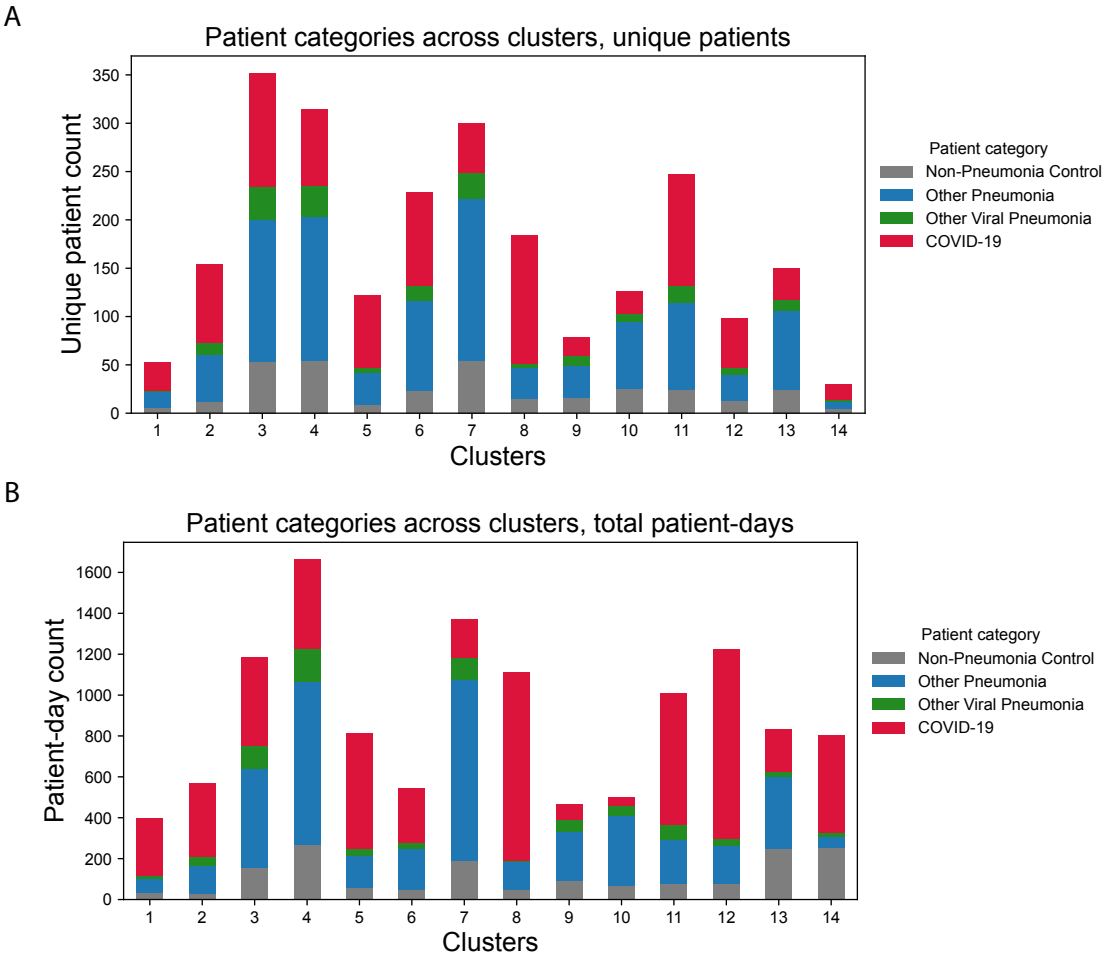
Supplemental Figure 5. Cluster mortality differentiation and robustness against small data perturbations. (**A**) Fraction of all possible pairs between two clusters that show a significantly different mortality at $p < 0.01$. X-axis shows different cutoffs for total number of clusters. Shaded area is the bootstrapped 95th percentile. (**B**) Allocation of individual patient-days (rows) to clusters following 100 randomizations, in which a single patient hospitalization has been excluded (columns). Cluster rank is the rank of cluster based on associated mortality with 1 being lowest and 14 being highest. (**C**) Clustering as described in panel (A) using "Similarity" approach but only using binary YES/NO flags. Evaluation of performance by share of pairwise clusters that show significantly different mortality at $p < 0.01$ as in panel (A). (**D**) Corresponding Silhouette scores for the methods in (C). Note that despite YES/NO flags reaching higher compactness of clusters as indicated by higher Silhouette scores, the clinically-relevant ability to distinguish clusters based on mortality is not increased as seen in (C).



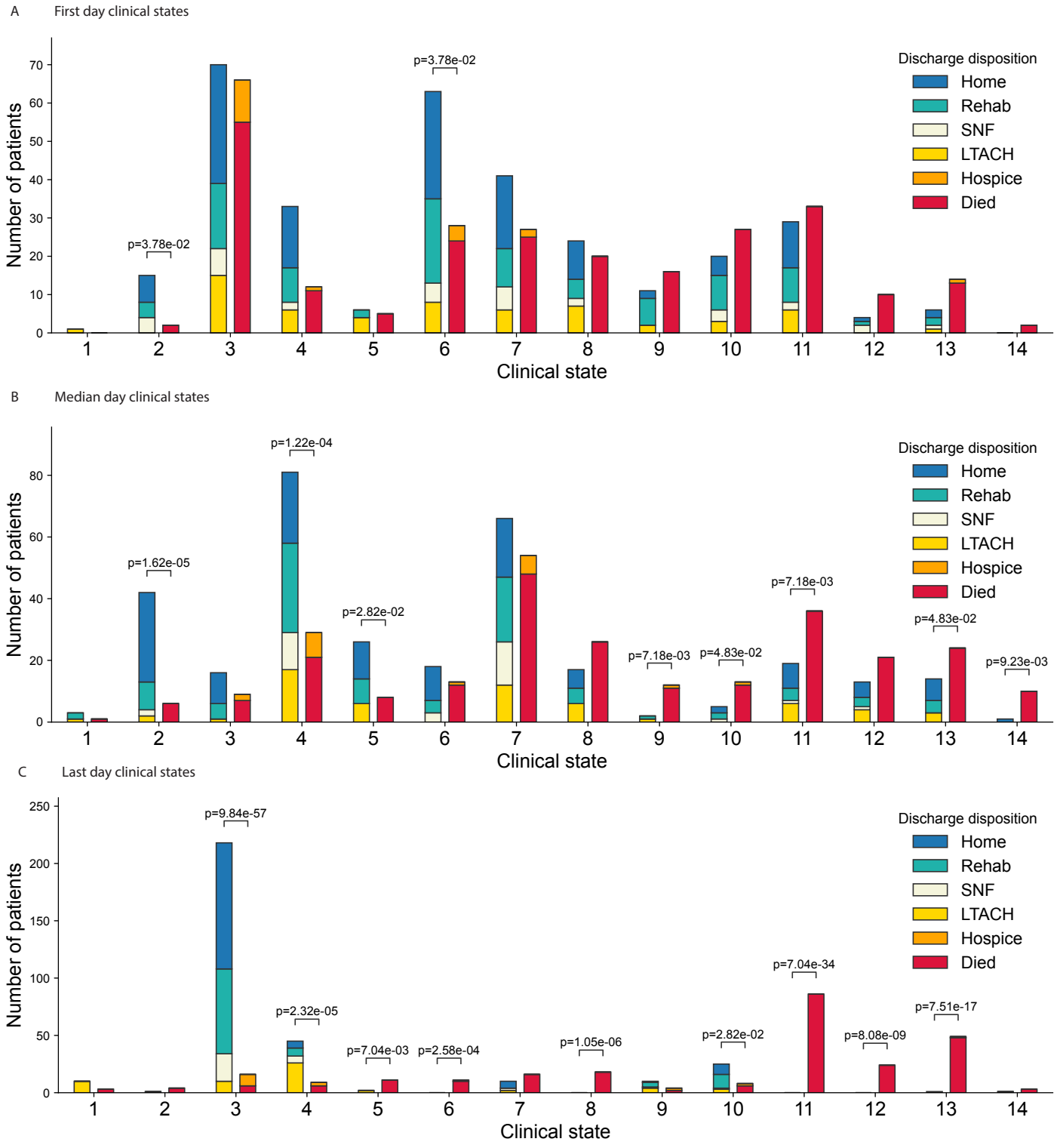
Supplemental Figure 6. UMAP and feature plots. (**A**) UMAP with colors and numbers representing *CarpeDiem*-defined clusters (clinical states). (**B-H**) Feature plots for individual parameters. (**I**) Patient-days from patients with COVID-19. (**J**) Example trajectory of a patient who was liberated from mechanical ventilation and was discharged home. A UMAP on the left demonstrating clinical transitions is labeled with the starting clinical state (day 0) and the day of discharge from the ICU (day 38). On the right, microbiological and clinical events overlaid on clinical state transitions are shown. Vertical lines indicate timepoints of BAL sampling. The beginning of the ICU course is outlined in light blue, whereas the end is dark blue, on both the timeline and the UMAP. More example trajectories are available in an interactive web app available at https://nupulmonary.org/carpediem. HFNC = high-flow nasal cannula, ECMO = extracorporeal membrane oxygenation, CRRT = continuous renal replacement therapy, MSSA = methicillin-sensitive *Staphylococcus aureus*.



Supplemental Figure 7. **SOFA scores per cluster.** Median [IQR] SOFA scores were calculated for the days represented in each cluster. Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Notches are bootstrapped 95% confidence interval of median.

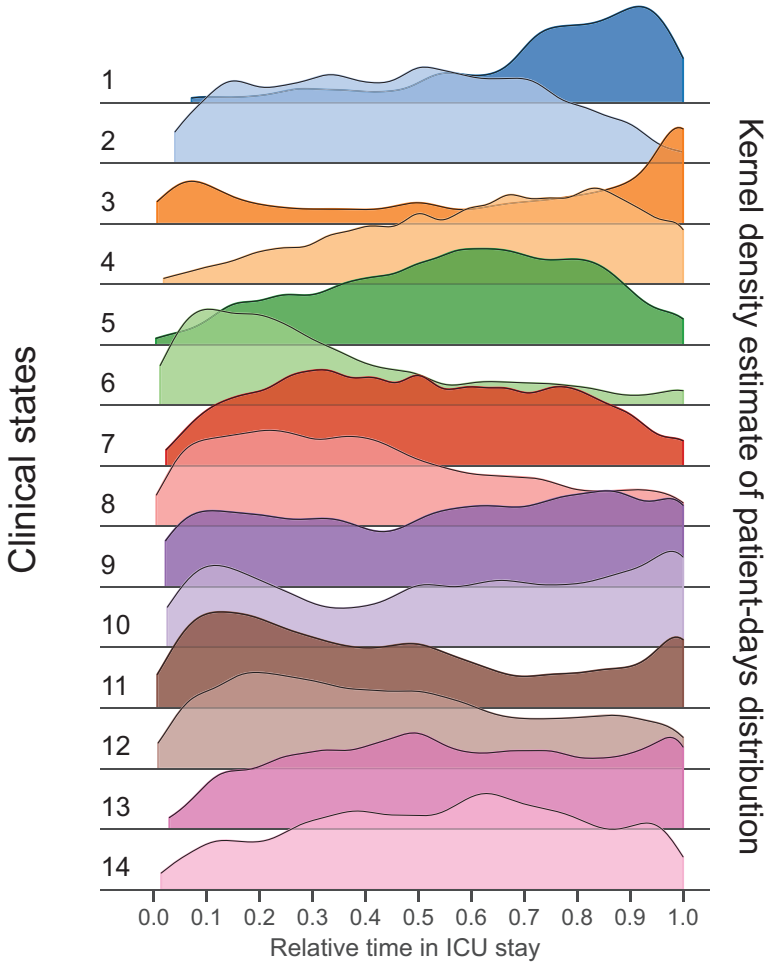


Supplemental Figure 8. Every cluster is comprised of patients from all four pneumonia categories. (**A**) Count of unique patients that contribute to clusters, colored by category. (**B**) Patient-days, colored by category.

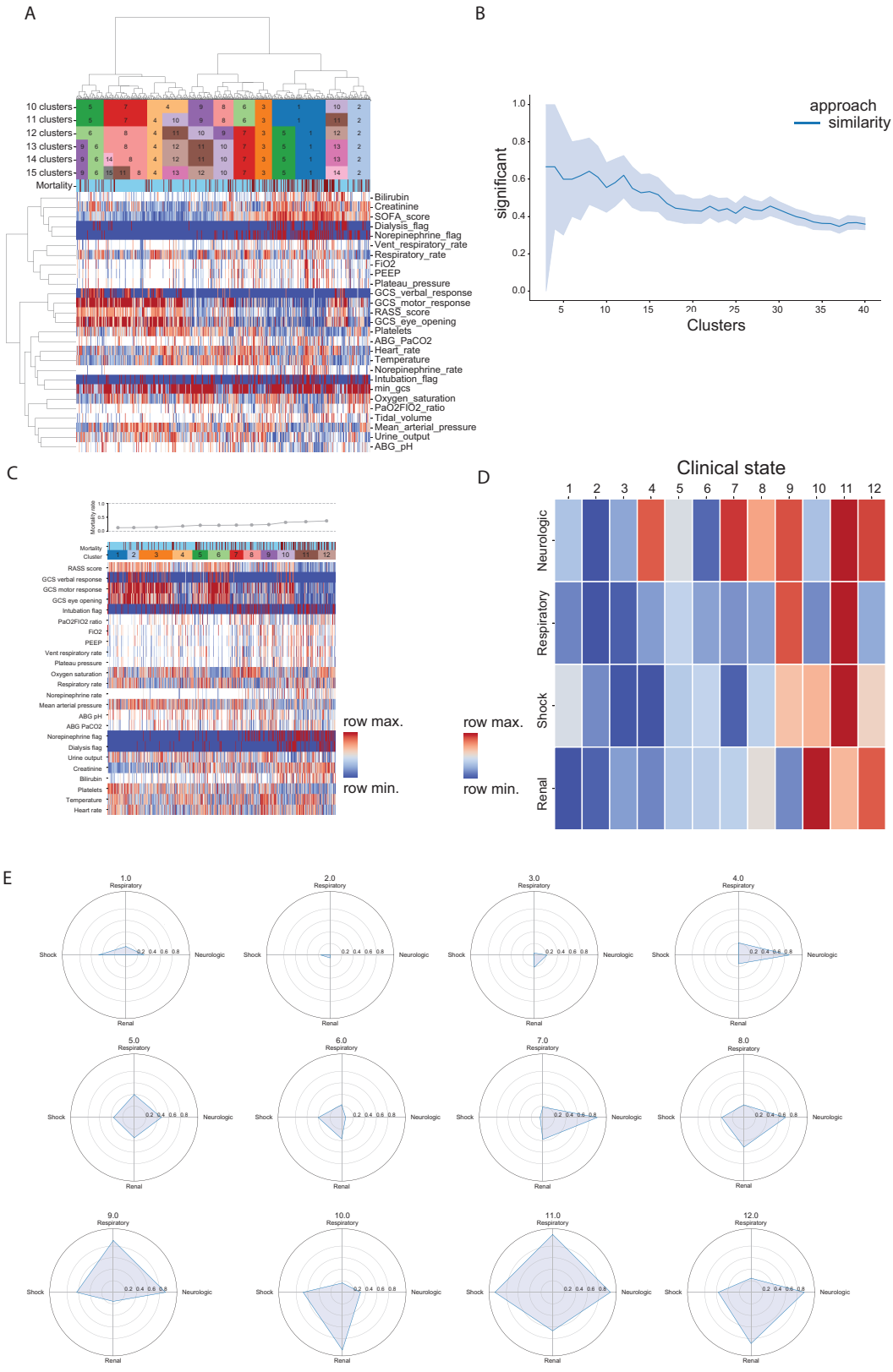


Supplemental Figure 9. Clinical states are associated with outcome. Association between the clinical state occupied by each patient on their first (**A**), median (**B**), or last (**C**) ICU day and their discharge disposition. Outcomes are displayed in two columns: the first column aggregates favorable discharge dispositions (Home, Rehab, SNF, LTACH), the second column aggregates unfavorable discharge dispositions (Hospice, Died). SNF = Skilled Nursing Facility, LTACH = long-term acute care hospital.

Categorical values were compared using Fisher's Exact tests with FDR correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

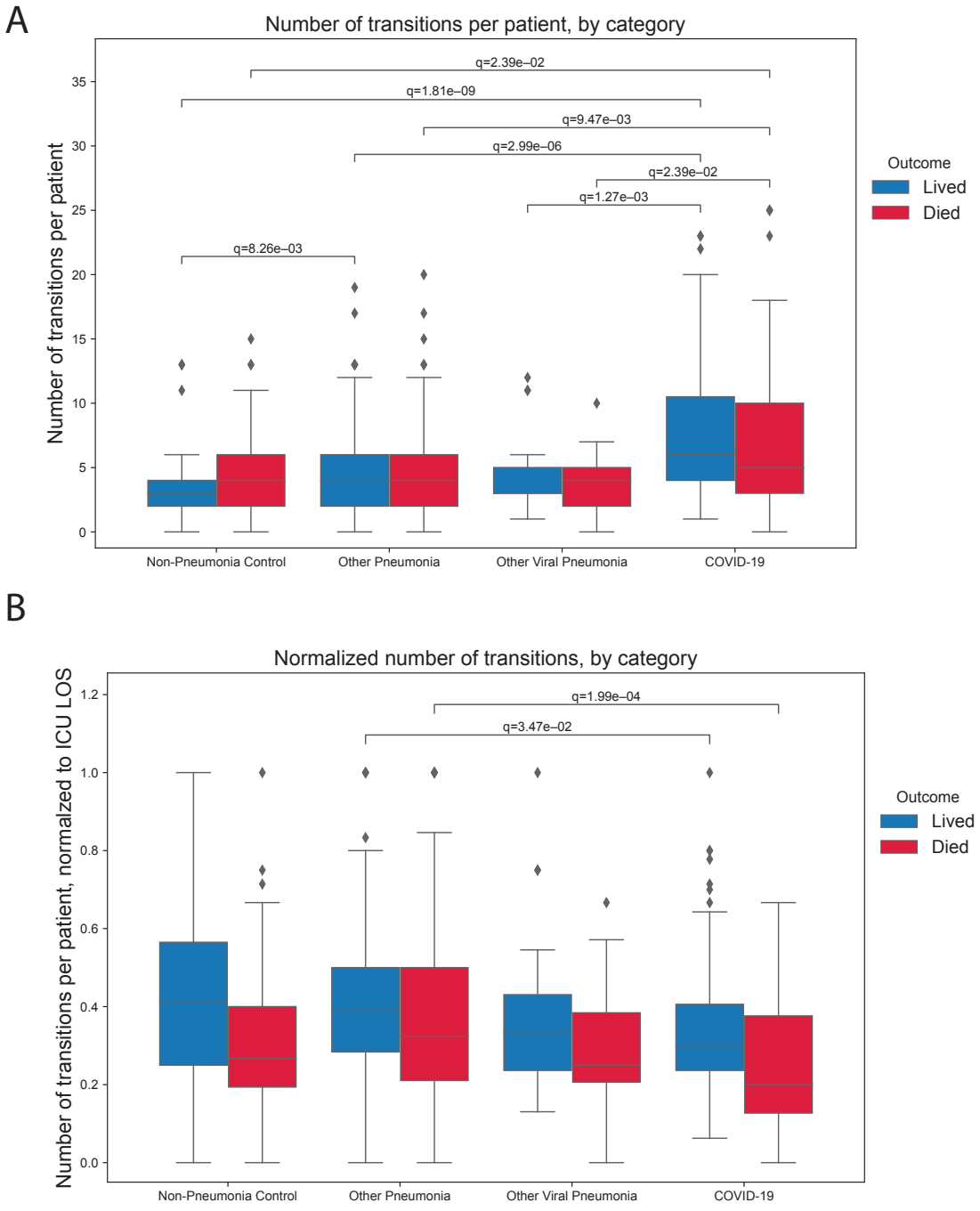


Supplemental Figure 10. Kernel density estimate plots showing relative time through the ICU stay for each clinical state.

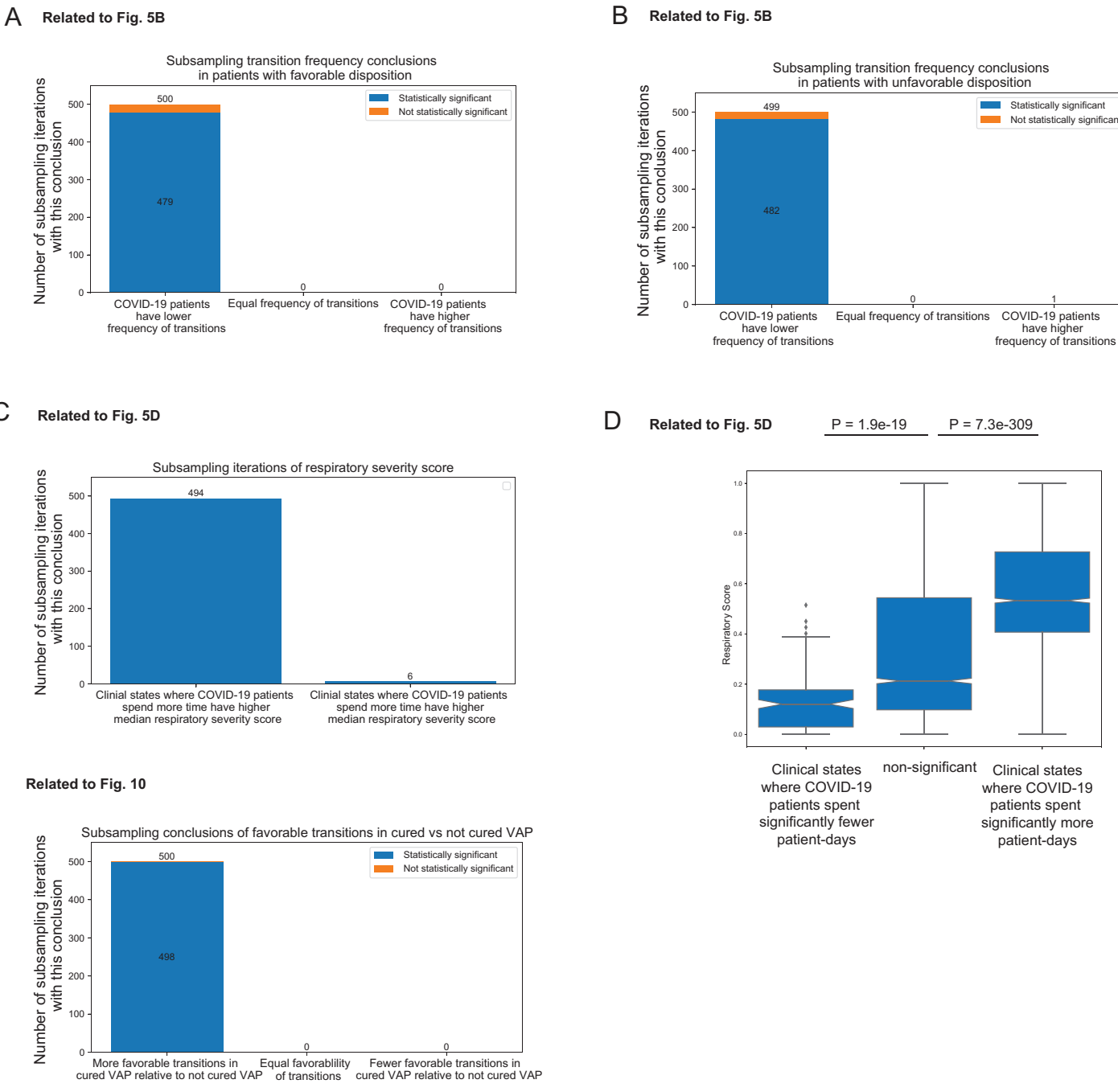


Supplemental Figure 11. Application of *CarpeDiem* to the MIMIC-IV dataset. (**A**) Hierarchical clustering of 27 clinical parameters (rows) with columns representing 15,642 ICU patient-days from 1,284 patients. (**B**) The fraction of pairs between clusters that show a significantly different mortality rate at different numbers of clusters. Shaded area is bootstrapped 95% confidence interval. (**C**) Heatmap of data from (A) re-ordered from lowest to

highest mortality, using 12 clusters. The top strip signifies hospital mortality of the patient shown in the column (blue = survived, red = died). The hospital mortality rate associated with each cluster is shown above the heatmap. (**D**) Heatmap of the composite signal from each cluster and physiologic group with ordering same as (C). (**E**) Spider plots of normalized composite features from (D) for each clinical state.

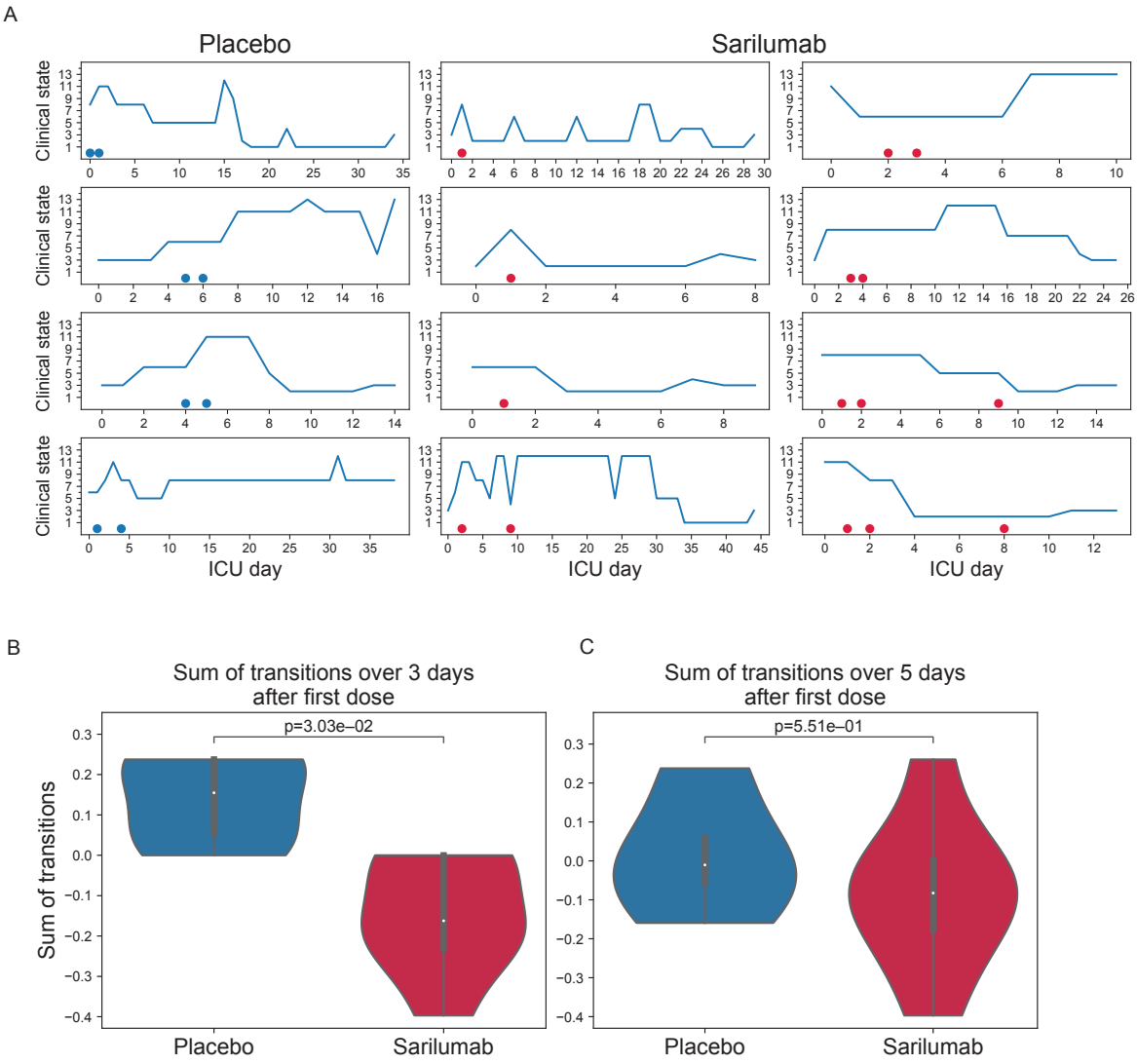


Supplemental Figure 12. Patients with COVID-19 have a higher absolute number of transitions but fewer when normalized for their longer ICU LOS. (**A**) Distribution of transitions per patient. (**B**) Distribution of transitions normalized by ICU LOS. Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.



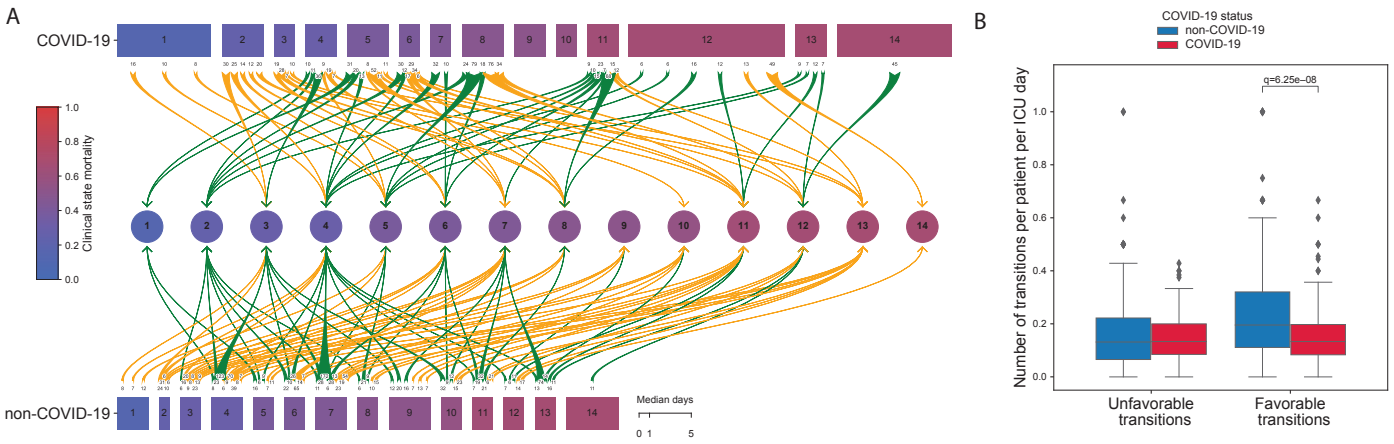
Supplemental Figure 13. Conclusions drawn from the full dataset are robust to random subsampling of 80% of patients over 500 iterations. As pertaining to Figure 5B, the frequency of transitions between clinical states was significantly lower in patients with COVID-19 compared with other patients, both in patients who had a favorable discharge disposition (**A**) and in patients who had an unfavorable discharge disposition (**B**). Patients with COVID-19 had a lower normalized frequency of transitions in 500 of 500 iterations among patients who had a favorable disposition (479 were statistically significant) and in 499 of 500 iterations among patients who had an unfavorable disposition (482 were statistically significant). (**C**) As pertaining to Figure 5D, we found that clinical states enriched in patients with COVID-19 had higher median respiratory severity scores (494 of 500 subsampling iterations). (**D**) Distribution of respiratory severity scores for clusters in which LOS is significantly shorter among patients with COVID-19, in which clusters have no significant difference in LOS, and in which

LOS is significantly longer among patients with COVID-19. (**E**) As pertaining to Figure 10A, 499 of 500 subsampling iterations found that there were more favorable transitions in episodes of cured VAP compared to episodes of VAP that were not cured. In all panels, significance indicates two-sided Mann-Whitney U test $p < 0.05$. Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Notches are bootstrapped 95% confidence interval of median.



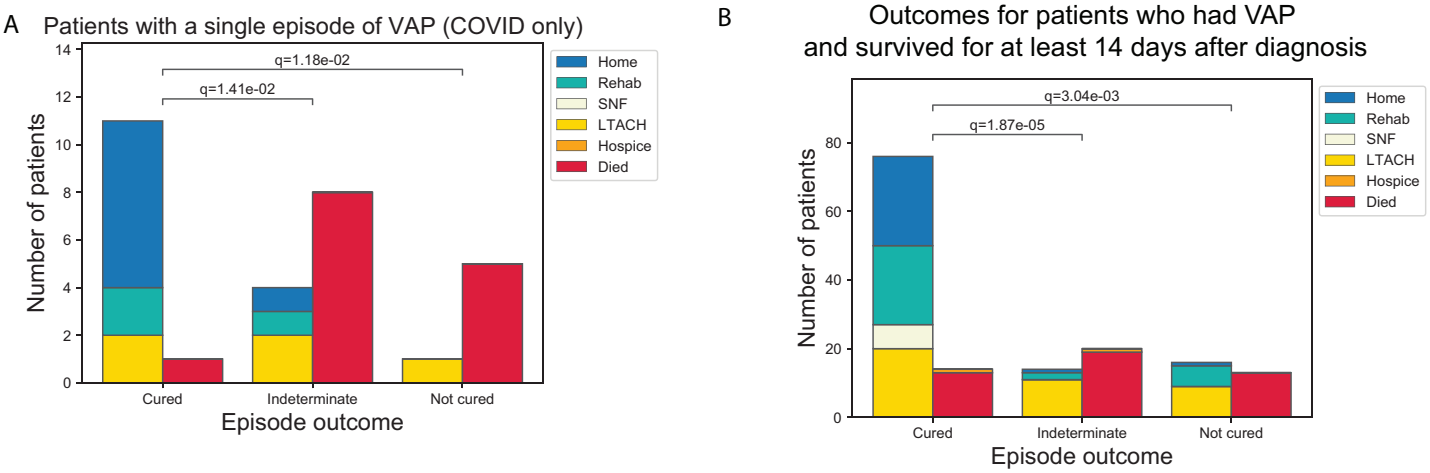
Supplemental Figure 14. *CarpeDiem* provides potential insights in the subset of patients co-enrolled in a randomized placebo-controlled trial of sarilumab for COVID-19 and respiratory failure. (**A**) Timeline of patients who received placebo (first column) or sarilumab (second and third columns) with x-axis showing ICU day and y-axis showing *CarpeDiem* clinical state. Dots indicate day of placebo or sarilumab administration. (**B**) Sum of transitions grouped by placebo versus sarilumab for the three days following the first dose. Higher sums of transitions reflect transitions to unfavorable (higher mortality) clusters. (**C**) Sum of transitions grouped by placebo versus sarilumab for the five days following the first dose.

Violin plots: colored area shows the kernel density estimate of the values distribution; inside, box-and-whisker plots are drawn in black with white dot corresponding to the median. Numerical values were compared by Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.



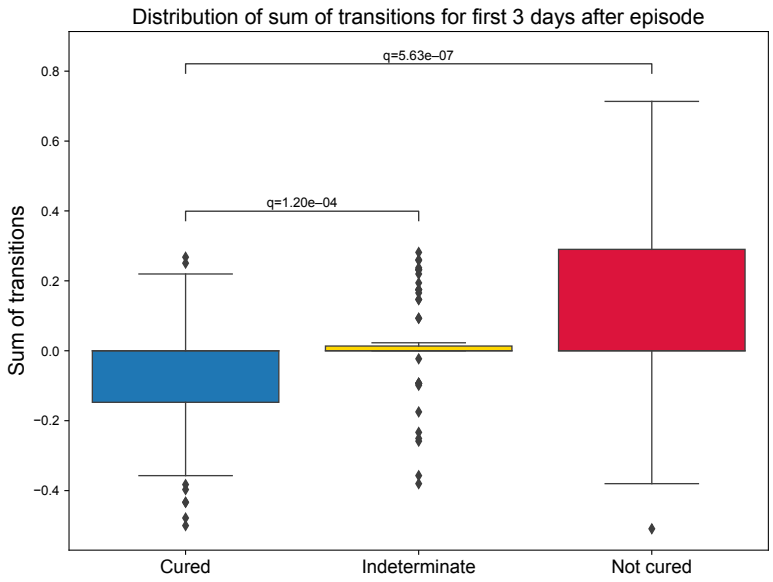
Supplemental Figure 15. Patients with SARS-CoV-2 pneumonia have a longer length of stay and fewer transitions between clinical states per day compared to patients with non-COVID-19-related respiratory failure. (**A**) Clinical states are ordered and numbered 1-14 by their associated mortality rate (blue to red). Rectangle width reflects median days per clinical state. Transitions marked by green arrows are to a more favorable (lower mortality) clinical state; yellow arrows mark transitions to a less favorable (higher mortality) clinical state. Numbers at the arrow bases represent the number of transitions between the two clinical states connected by the arrow. Only transitions that occurred more than five times are shown (cutoff of 30 transitions used in Figure 6). (**B**) Quantification of the number of transitions per patient per ICU day, grouped by COVID-19 and outcome favorability.

Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

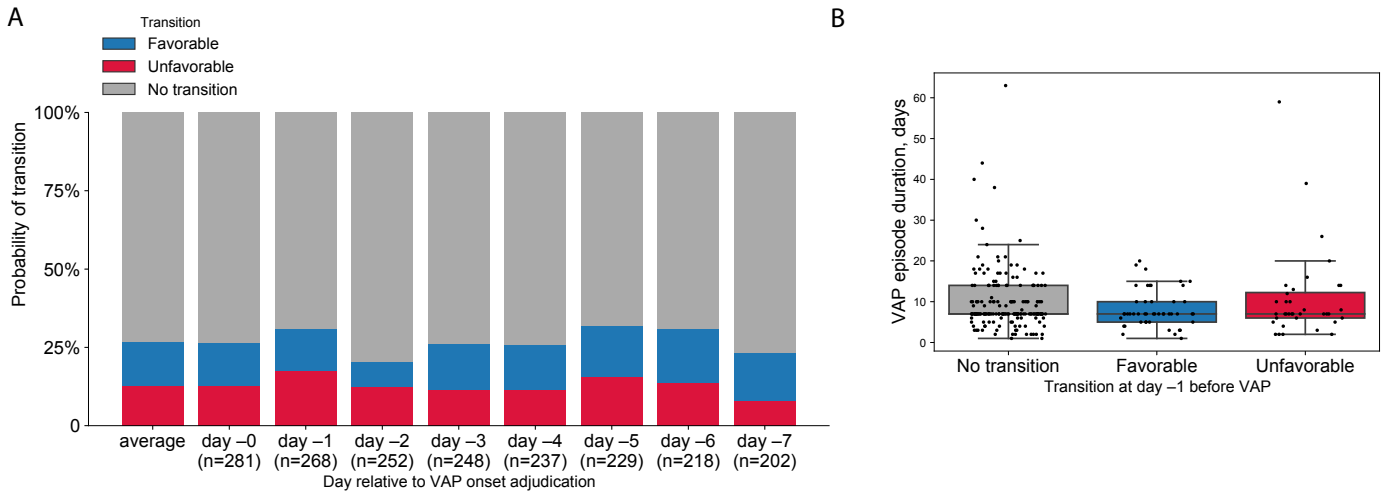


Supplemental Figure 16. Unresolving VAP is associated with worse outcomes. (**A**) Mortality associated with a single episode of VAP among patients with COVID-19. (**B**) Outcomes for patients who did not die within 14 days following the onset of their VAP episode. Outcomes are displayed in two columns: the first column aggregates favorable discharge dispositions (Home, Rehab, SNF, LTACH), the second column aggregates unfavorable discharges (Hospice, Died).
Categorical values were compared using Fisher's Exact tests with FDR correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.

A

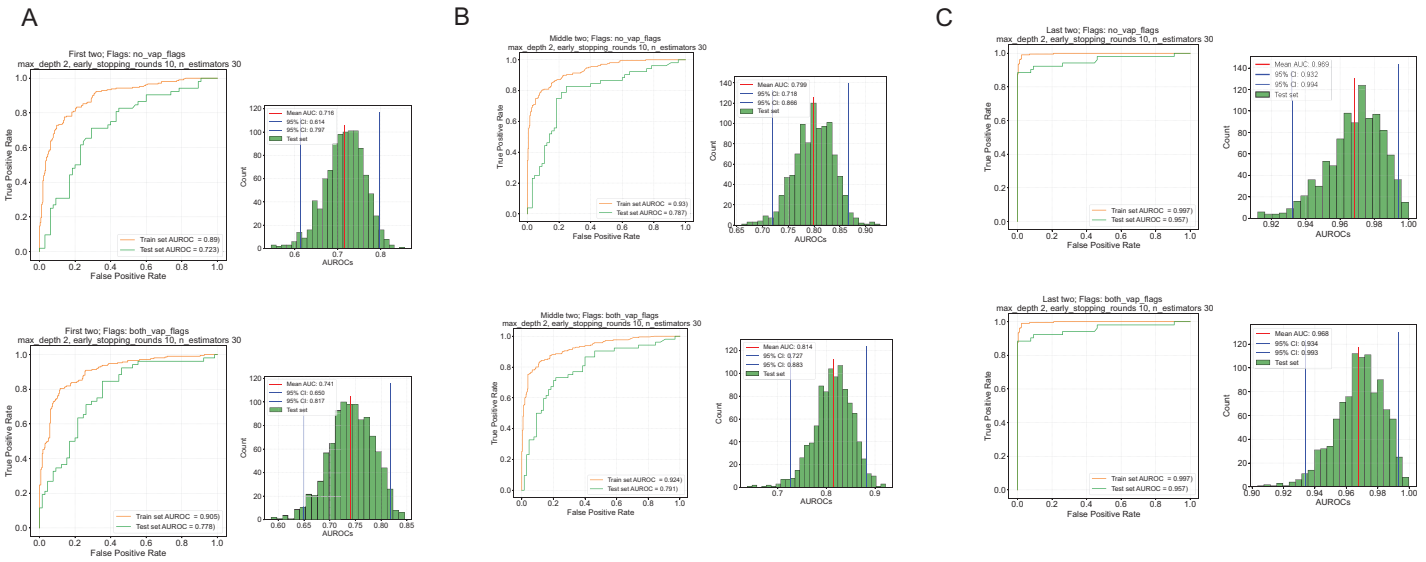


Supplemental Figure 17. Differences in the summative favorability of transitions is evident as early as day 3 following the diagnosis of VAP. (**A**) Sum of transitions grouped by VAP episode outcome for the three days following the diagnosis of VAP. Higher sums of transitions reflect transitions to unfavorable (higher mortality) clusters. Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.



Supplemental Figure 18. Clinical state favorability changes before VAP diagnosis. (**A**) An increase in unfavorable transitions occurs a day before VAP diagnosis (day -**1**). (**B**) Day -1 transitions are not associated with duration of the ensuing VAP episode.

Box-and-whisker plots: box shows quartiles and median, whiskers show minimum and maximum except for outliers, which are shown as individual data points. Numerical values were compared using Mann-Whitney U tests with false-discovery rate (FDR) correction using the Benjamini-Hochberg procedure. A q-value < 0.05 was our threshold for statistical significance.



Supplemental Figure 19. Gradient boosting modeling reveals minimal increase in predictive capability when VAP and VAP cure status are added to clinical parameters measured at the beginning, in the middle, and at the end of the ICU course. Area under the receiver operating characteristics (AUROC) curve values for clinical parameters in the gradient boosting analysis predicting unfavorable hospital outcome, with corresponding confidence interval plots obtained using bootstrapping. Using the worst features from the (**A**) first two days, (**B**) median two days, and (**C**) last two days. Bottom row has addition of two flags, *had_vap* and *vap_interdeterminate_uncured*, to indicate a diagnosis of VAP during the ICU stay and a VAP outcome other than cured.