nature medicine

Article

https://doi.org/10.1038/s41591-025-03956-5

A consensus immune dysregulation framework for sepsis and critical illnesses

Received: 20 November 2024

Accepted: 12 August 2025

Published online: 30 September 2025

Check for updates

A list of authors and their affiliations appears at the end of the paper

Critical care syndromes such as sepsis, acute respiratory distress syndrome (ARDS) and trauma continue to have unacceptably high morbidity and mortality, with progress limited by the inherent heterogeneity within syndromic illnesses. Although numerous immune endotypes have been proposed for sepsis and critical care, the similarities and differences between these endotypes remain unclear, hindering clinical translation. The SUBSPACE consortium is an international consortium that aims to advance precision medicine in critical care through the sharing of transcriptomic data. Here, evaluating the overlap of existing immune endotypes in sepsis across >7,074 samples from 37 independent cohorts, we developed cell-type-specific gene expression signatures to quantify dysregulation within immune compartments. Myeloid and lymphoid dysregulation were associated with disease severity and mortality across all cohorts. Importantly, this dysregulation was also observed in patients with ARDS, trauma and burns, suggesting a conserved mechanism across various critical illness syndromes. Moreover, analysis of randomized controlled trial data revealed that myeloid and lymphoid dysregulation are associated with differential mortality in patients treated with anakinra in the SAVE-MORE trial ($n = 452$) and corticosteroids in the VICTAS ($n = 89$) and VANISH ($n = 117$) trials, underscoring their prognostic and therapeutic implications. In conclusion, our proposed immunology-based framework for quantifying cellular compartment dysregulation offers a potentially valuable tool for understanding immune dysregulation in critical illness with prognostic and therapeutic significance.

The field of critical care has expanded dramatically since the first intensive care units (ICUs) were developed[1]. However, progress has slowed and reduction in ICU mortality has plateaued[2]. One reason for this plateau is that most ICU admissions are related to physiology-driven syndromic definitions such as sepsis, acute respiratory distress syndrome (ARDS) and trauma, which ignore inherent biological and clinical heterogeneity[3]. Over 100 clinical trials of immune-modulating medications in sepsis, costing hundreds of millions of dollars, have failed to achieve consistent clinical benefits[4]. On the other hand, countless secondary analyses have identified biological subgroups that may benefit from targeted therapies[5–10]. To advance precision medicine in the ICU, we must redefine critical illness based on biological mechanisms, rather than clinical syndromes[11].

In sepsis, transcriptomic and proteomic endotyping schemas have successfully identified subgroups, retrospectively, of patients at higher risk of mortality and those who respond differentially to immune-modulatory therapies[12–20]. Importantly, although these endotypes were developed in 'sepsis', there were substantial differences in patient populations, infectious etiology, severity and the clustering approach. For example, Wong et al. evaluated gene expression in pediatric septic shock and identified two endotypes: one high risk and one low risk[17]. Despite potential age-related differences in the host response, these

e-mail: pkhatri@stanford.edu

endotypes were congruent with two endotypes developed in adult patients with pneumonia by Davenport et al. that were later shown to have a differential response to steroids[5,12]. Scicluna et al. identified four transcriptomic endotypes across two ICUs (medication administration records (MARS) 1–4)[20], whereas Sweeney et al. identified three endotypes (inflammopathic, coagulopathic and adaptive) in both critically ill and noncritically ill patients with bacterial sepsis[14]. Finally, Zheng et al. described four continuous immune severity scores (the severe-or-mild (SoM) signature) that are conserved across a broad array of viral and bacterial infections[10,19].

The convergence of results by these independent research groups on sepsis endotypes offers promise for advancing molecular endotyping. However, how these schemas are related to each other, whether they generalize beyond the cohorts in which they were originally identified and whether they represent the same or different underlying biology remain important questions that must be answered to fundamentally redefine critical illness syndromes and advance precision medicine.

The goal of the subtyping in sepsis and critical illness (SUBSPACE) consortium is to advance precision medicine in sepsis and critical care syndromes by identifying and understanding the underlying biological pathways and redefine critical illnesses based on molecular biology, rather than traditional clinical categorizations. We hypothesized that the comparison and integration of existing transcriptomic endotyping frameworks across multiple critical illness cohorts would reveal distinct molecular pathways and immune cell-specific dysregulation. These biological insights could provide a basis for redefining critical care syndromes, enabling a more precise, biology-driven classification that can inform targeted therapies and improve patient outcomes in critical care.

Results

Unsupervised clustering identifies four consensus gene expression-based clusters

Our primary objective was to evaluate whether the existing gene expression endotyping signatures for patients with sepsis identified similar biology. To ensure validity and reproducibility, we applied the same methods in parallel across the public and SUBSPACE cohorts (Extended Data Fig. 1). First, we assigned standardized severity labels to each of the 1,460 blood samples from 19 independent public studies, encompassing adult and pediatric patients infected with 1 of 15 types of bacterial and viral infections[21–38] (Supplementary Table 1): healthy, mild or moderate infections (those who did not require ICU admission), severe infections (those who required ICU admission) and fatal infections[19]. Next, we used combat co-normalization using controls (COCONUT) to co-normalize these datasets and ensured that there were no batch effects post-normalization (Supplementary Fig. 1).

We excluded two endotype signatures, Cano-Gamez sepsis response signature (SRS) and Davenport SRS, from the public dataset analysis because several of their genes were not measured in all datasets. For the remaining five endotyping schemas (Fig. 1a), we generated continuous endotype scores for all samples and used hierarchical clustering, principal component analysis (PCA) and network analysis to evaluate the overlap between them. Each method identified similar clusters across endotype schemas (Fig. 1b–d), suggesting that these schemas identified the same endotypes despite using different methodologies and populations. Silhouette index analysis found that the ideal number of clusters varied between two and four, depending on the etiology and severity of infections (Extended Data Fig. 2a–e). For instance, when using all infections, irrespective of severity, the optimal number of clusters was three; however, when using only severe infections, the optimal number of clusters was four. Importantly, across all clustering methods, the same set of endotypes grouped together across four clusters, regardless of the optimal number. Bootstrapping with 1,000 repetitions confirmed this result ($P < 0.01$; Extended Data Fig. 2f). Therefore, we carried forward these four consensus clusters to further evaluate whether there were important biological and clinical differences. Among these four clusters, two included endotypes that have been previously associated with worse outcomes, which we refer to as detrimental clusters, whereas the other two included endotypes that have been previously associated with improved outcomes, which we refer to as protective clusters. Importantly, this corroborates prior pathway analyses, which have suggested that the biology underlying these endotypes may overlap despite using different endotype-defining genes (Extended Data Table 1). Finally, analysis of pairwise correlation of genes across these four clusters, showing that genes within each cluster were highly correlated, again suggested that these endotype schemas were identifying common biology despite the use of different genes identified using different methods and patient populations (Extended Data Fig. 3).

Next, we investigated whether these 4 molecular endotypes were reproducible using 4,106 blood samples from 3,380 patients across 12 independent prospective cohorts from 10 centers integrated through the SUBSPACE consortium. These samples represented broad heterogeneity, including pediatric and adult patients, noncritical and critically ill patients and infected and noninfected patients, inclusive of both bacterial and viral sepsis (Fig. 2a and Supplementary Table 2). All gene expression data, except the MESSI cohort due to a lack of healthy participants, were COCONUT co-normalized. Housekeeping genes and uniform manifold approximation and projection (UMAP) showed appropriate co-normalization (Supplementary Fig. 2). We calculated continuous endotype scores from each of the seven gene expression signatures for each sample. Once again, unsupervised hierarchical clustering and network analysis identified four gene expression-based clusters, with the addition of quantitative SRS (SRSq) scores clustering with detrimental endotypes (Fig. 2b–e). In the SUBSPACE cohorts, unsupervised clustering identified identical clusters to the public data whereas network analysis identified similar subgroups with the exception of the Sweeney coagulopathic signature and the Wong signature clustering with the detrimental endotypes now containing SRSq, suggesting some similarities between these endotypes. Importantly, none of the clusters was driven by a single cohort (Fig. 2c).

Collectively, our results demonstrated that, despite the biological, clinical and technical heterogeneity across cohorts, the endotypes identified by different schemas converge to four consensus molecular clusters, henceforth called consensus endotypes. These molecular subgroups separated based on detrimental and protective endotypes. Overall, this suggests that prior sepsis transcriptomic signatures share a biological basis that may be leveraged to better understand sepsis pathogenesis and treatment.

Four consensus endotypes can be explained along the myeloid and lymphoid axes

After identifying the consensus of these discrete endotyping signatures, we next evaluated the immunological underpinnings of these consensus endotypes and developed a more generalizable immune framework. We evaluated the 7 endotyping signatures using single-cell RNA sequencing (scRNA-seq) data by integrating 602,388 immune cells from 258 samples from 4 publicly available COVID-19 and sepsis scRNA-seq datasets that included neutrophil profiles[39–42] (Supplementary Table 3). We identified 14 unique cell types (Fig. 3a and Methods) and found that cell type and severity explained the largest variability (Supplementary Fig. 3a,b). The consensus endotypes separated along cellular origin and detrimental or protective effects, which we defined based on whether the corresponding endotype was associated with worse or improved prognosis (that is, higher severity or mortality) in prior studies. Consensus endotypes included a detrimental myeloid cluster (Sweeney inflammopathic, Yao innate, SoM modules 1 and 2, and MARS2), a protective myeloid cluster (Wong score, MARS4 and

https://doi.org/10.1038/s41591-025-03956-5

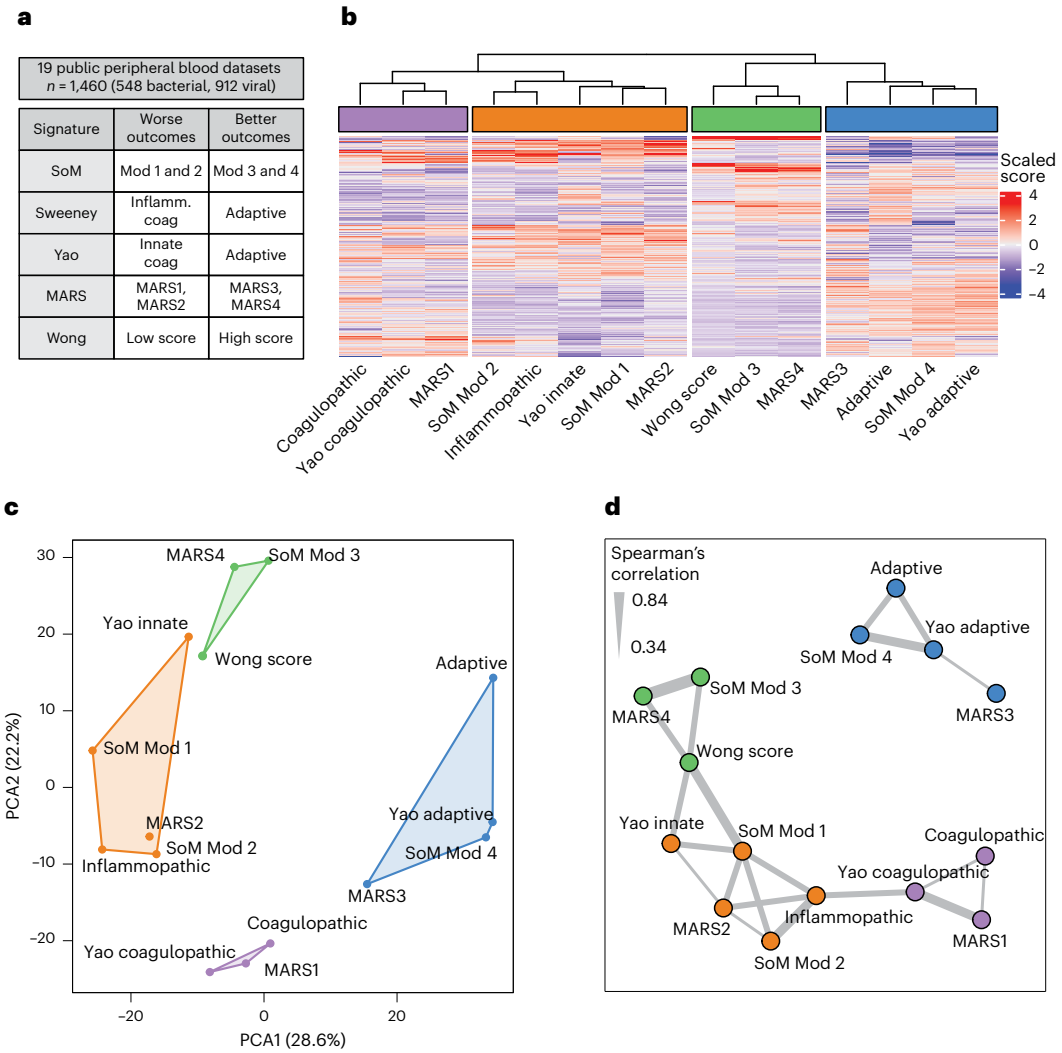


Fig. 1 | Identification of consensus endotypes in public data. a, Peripheral blood gene expression data from 19 cohorts inclusive of 548 samples from bacterially infected patients and 912 samples from virally infected patients co-normalized. We calculated the five sepsis signatures and scaled values for unsupervised clustering. **b**, Unsupervised hierarchical clustering performed by scaled gene expression score (*x* axis) across all samples (*y* axis) identifying four consensus endotypes. **c**, The four identified consensus endotypes separated well in PCA. **d**, Network analysis performed on scaled scores using Spearman's correlation >0.33 to identify edges. Clusters were identified using a greedy forward algorithm, which identified four clusters mirroring those identified by unsupervised hierarchical clustering. The thickness of the line represents correlation between the nodes it connects. Coag, coagulopathic; Inflamm., inflammopathic; Mod, module.

SoM module 4), a protective lymphoid cluster (Sweeney adaptive, Yao adaptive, SoM module 4 and MARS3) and a mixed myeloid or lymphoid cluster (Sweeney coagulopathic, Yao coagulopathic and MARS1) (Fig. 3b).

Although predominant cell-type expression explained these consensus endotypes to some extent, genes in these signatures were expressed across multiple cell types (Fig. 3b and Supplementary Fig. 4). To isolate myeloid-specific and lymphoid-specific dysregulation scores, we evaluated the cell specificity of all genes used in the 7 signatures and identified 104 genes that were selectively expressed in either myeloid or lymphoid cells. We divided these genes into myeloid detrimental, myeloid protective and lymphoid protective subgroups based on whether the original endotyping signature in which they were included was considered detrimental or protective as previously defined (Fig. 3c and Supplementary Table 4). Importantly, no lymphoid-specific genes were found in any of the detrimental signatures evaluated. We then defined myeloid and lymphoid dysregulation scores as the difference between the geometric mean of detrimental genes (when applicable) and the geometric mean of protective genes, for a given cell lineage. Evaluation of myeloid and lymphoid dysregulation scores using scRNA-seq data confirmed their cell-type specificity (Supplementary Fig. 5a–d). Myeloid and lymphoid dysregulation scores were moderately correlated with each other ($r = 0.49$, $P < 2.2 \times 10^{-16}$; Supplementary Fig. 6) in bulk transcriptome data from the SUBSPACE cohorts.

Overall, scRNA-seq data demonstrated that the four consensus endotypes were associated with distinct expression profiles in myeloid and lymphoid immune cells.

Cell-lineage-specific quantification generates a clinically relevant immune dysregulation framework

Association of consensus molecular endotypes with distinct immune cell types presented an opportunity to define an immune response-based evaluation framework, which we term the human immune dysregulation evaluation framework (Hi-DEF). We hypothesized that use of these gene expression-based scores to quantify myeloid- and lymphoid-specific dysregulation for each patient would reduce between-patient heterogeneity.

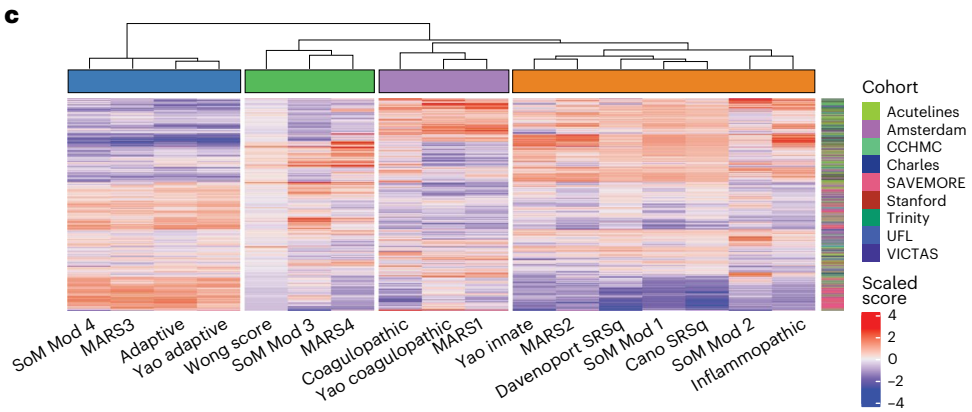
https://doi.org/10.1038/s41591-025-03956-5

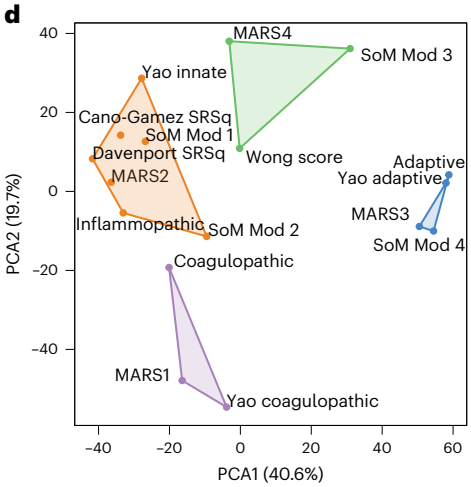
a

Institute	Sample number	Patient characteristics
Acutelines	992	High acuity ED
Amsterdam UMC	1,071	Pneumonia, COVID-19
CCHMC	311	Pediatric sepsis
Charles University	38	Sepsis
SAVE-MORE	752	COVID-19
Stanford University	236	ICU
Trinity College	204	Sepsis
University of Florida	172	Sepsis, trauma
VICTAS	141	Sepsis

b

Signature	Worse outcomes	Better outcomes
Cano-Gamez SRS	High SRSq	Low SRSq
Davenport SRS	High SRSq	Low SRSq
MARS	MARS1, MARS2	MARS3, MARS4
SoM	Modules 1 and 2	Modules 3 and 4
Sweeney endotypes	Inflammopathic coagulopathic	Adaptive
Wong score	Low score	High score
Yao endotypes	Innate coagulopathic	Adaptive





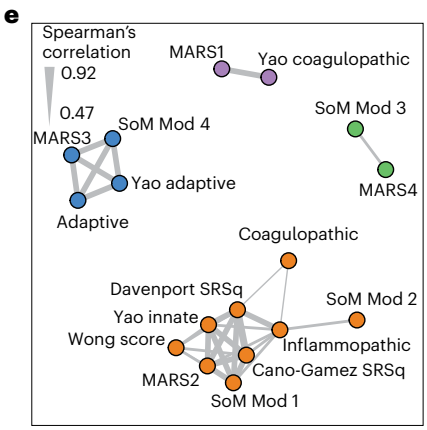


Fig. 2 | Identification of consensus endotypes in SUBSPACE data. a, Peripheral blood gene expression data collected and co-normalized from 3,917 samples across 11 cohorts from 9 international centers. UMC, University Medical Center; CCHMC, Cincinnati Children's Hospital Medical Center. **b**, Application of seven prior sepsis endotyping signatures and scaled signature scores for unsupervised clustering. **c**, Unsupervised hierarchical clustering performed by scaled gene expression score (*x* axis) across all samples (*y* axis) identifying four consensus endotypes. Samples did not cluster together by cohort. **d**, The four identified consensus endotypes separating well on PCA. **e**, Network analysis performed on scaled scores using Spearman's correlation >0.47 (median correlation) to identify edges. The clusters were identified using a greedy forward algorithm, which identified four clusters. The thickness of the line represents correlation between the nodes it connects.

To test this hypothesis, we computed the lymphoid and myeloid dysregulation scores as defined above for each sample in the public datasets. Both myeloid and lymphoid dysregulation scores increased significantly with severity across public datasets (Jonckheere–Terpstra (JT) *t*-test $P < 2.2 \times 10^{-16}$ for both scores; Fig. 4a). Next, we defined an abnormal lymphoid or myeloid dysregulation score using the 95th percentile of each score in healthy participants, which corresponds to a *z*-score of 1.65, and defined four quadrants: balanced, lymphoid dysregulation, myeloid dysregulation and system-wide dysregulation (Fig. 4b). Patients with either a myeloid or a lymphoid dysregulation score ≥1.65 had a significantly higher risk of severe infection or mortality (odds ratio (OR) = 5.2, 95% confidence interval (CI) 3.9–7.0, $P < 2.2 \times 10^{-16}$) compared with those with both scores <1.65 (Fig. 4c,d). The risk of severe infection or mortality was highest for patients with system-wide dysregulation, with 51% of these patients experiencing severe infections compared with 24% in the myeloid dysregulation subgroup, 10% in the lymphoid dysregulation subgroup and only 6% in the balanced subgroup ($P < 0.01$ across all comparisons; Fig. 4d).

https://doi.org/10.1038/s41591-025-03956-5

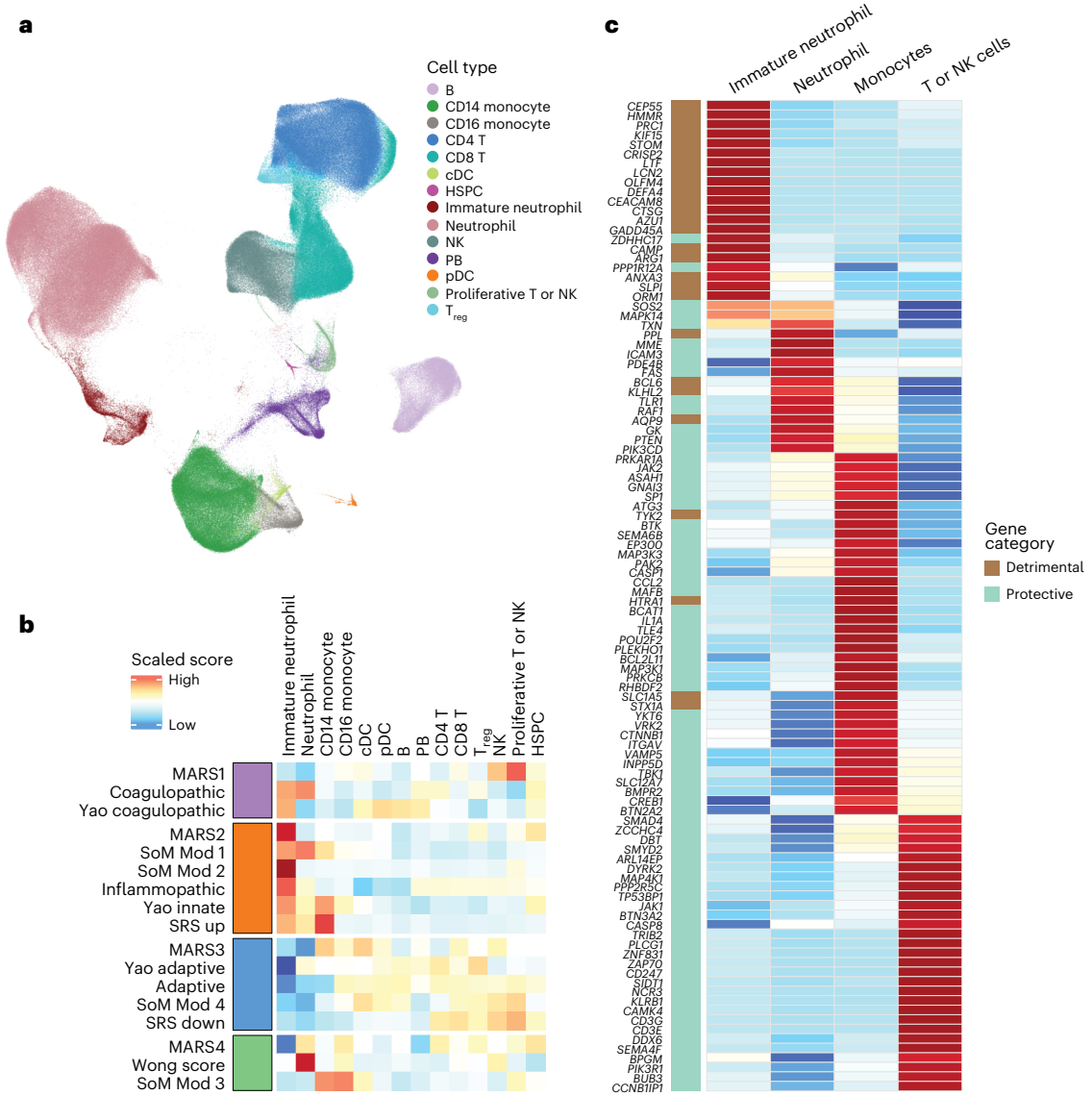


Fig. 3 | Single-cell analysis of consensus endotypes. a, Integration of 4 whole-blood scRNA-seq datasets from patients with COVID-19 and sepsis, inclusive of the neutrophil compartment and identifying 14 unique cell types using the Seurat and Scanpy pathways. The UMAP of cell types is shown. **b**, Evaluation of scaled gene expression signatures across these cell types, showing that the scores included in each consensus molecular endotype were expressed in similar cell types. The red cluster (MARS2, SoM module 1 or 2, Sweeney inflammopathic, Yao innate and SRS signatures) was predominantly expressed with immature neutrophils. The blue cluster (MARS3, Yao adaptive, Sweeney adaptive and SoM module 4) was predominantly expressed in T or NK cells. The purple cluster (MARS1, Sweeney coagulopathic and Yao coagulopathic) was a mix of intermediate expression of neutrophils and T or NK cells. The green cluster (MARS4, Wong score and SoM module 3) was predominantly expressed in mature neutrophils and monocytes. **c**, Development of a cell-type-specific score by evaluating scaled expression of each gene across all endotype signatures and selecting 104 genes that were selectively expressed (defined by >1 s.d. greater than other cell types evaluated) in myeloid or T or NK cell types. We then divided these genes into detrimental or protective genes based on whether the signature from which they were derived was associated with worse or better outcomes in prior studies. cDC, classical dendritic cell; HSPC, hematopoietic stem and progenitor cell; PB, peripheral blood; pDC, plasmacytoid dendritic cell.

These results remained significant across multiple sensitivity analyses, including adult and pediatric cohorts (Supplementary Fig. 7), bacterial and viral infectious etiologies (Supplementary Fig. 8) and US versus non-US cohorts (Supplementary Fig. 9).

Similar to the public datasets, both dysregulation scores increased with severity in the co-normalized SUBSPACE cohorts (JT *t*-test $P < 2.2 \times 10^{-16}$ for both scores; Fig. 4e). Both the myeloid and the lymphoid dysregulation scores were associated with 30-d mortality across all cohorts with an OR of 1.9 (95% CI 1.3–2.0, $P < 0.001$) and 1.6 (95% CI 1.6–2.8, $P < 0.001$), respectively (Fig. 4f). Myeloid dysregulation was most significantly ($P < 0.05$) associated with mortality in predominantly ICU and bacterially infected cohorts (Stanford and VICTAS), whereas the lymphoid dysregulation score had a more significant ($P < 0.05$) association with mortality in cohorts with predominantly viral infections (Amsterdam PANAMO and SAVE-MORE), a trend that was further highlighted when we evaluated differences in outcomes solely in virally or bacterially infected patients (Extended Data Fig. 4).

Using a *z*-score of 1.65 relative to healthy partcipants as a dysregulation threshold, both dysregulation scores ≥1.65 were significantly associated with severe illness requiring ICU level of care or dying within 30 d (OR = 7.1, 95% CI 5.6–8.9, $P < 2.2 \times 10^{-16}$; Fig. 4g,h) in the SUBSPACE cohorts. When considering only 30-d mortality, it was significantly higher in those with either myeloid or lymphoid dysregulation (OR = 3.5, 95% CI 2.3–5.4, $P = 5.3 \times 10^{-12}$;

Extended Data Fig. 5), where the system-wide dysregulation subgroup had the highest mortality rate (18%). Importantly, Hi-DEF also validated in the MESSI cohort, although it was not co-normalized with the SUBSPACE cohorts (Supplementary Fig. 10).

We evaluated whether Hi-DEF was differentiable clinically. Although there were differences in age and white blood counts across subgroups (Supplementary Table 5), there was substantial overlap. In the Stanford cohort, dysregulation was associated with vital and laboratory derangements (Supplementary Table 6); however, overlap would again limit clinical detection. Although the myeloid and lymphoid scores were correlated with the neutrophil-to-lymphocyte ratio (Supplementary Fig. 11), both myeloid (OR = 2.1, 95% CI 1.7–2.7, $P < 0.001$) and lymphoid (OR = 2.8, 95% CI 2.3–3.5, $P < 0.001$) scores remained associated with severity and mortality after adjusting for the neutrophil-to-lymphocyte ratio.

Hi-DEF demonstrates the need for flexible context-dependent subgrouping

A key limitation of existing transcriptomic sub-phenotyping schemas is that they often 'force' subgroupings and thus lack generalizability beyond the populations in which they were developed. For instance, an 'appropriate' immune response to an upper respiratory tract infection or viral pneumonia may be inadequate for Gram-negative bacteremia; however, current endotyping schemas do not allow this nuanced evaluation.

To illustrate this issue, we analyzed differences in myeloid and lymphoid dysregulation by severity (Fig. 5a), recruitment location (Supplementary Fig. 12) and infectious etiology (Extended Data Fig. 6a). These analyses confirmed substantial differences in the magnitude and range of immune dysregulation. For example, patients enrolled in emergency department or non-ICU settings had lower dysregulation scores than those enrolled in an ICU (Supplementary Fig. 12). Next, analysis of the proportion of healthy participants, and those with mild, severe or fatal illness across myeloid and lymphoid score quintiles, found that the composition of patients varied substantially across these quintiles (Fig. 5b,c). These results show that differences in cohort composition (that is, a mix of critically and noncritically ill individuals versus solely critically ill patients) affect the dysregulation measured and subsequent results of endotyping schemas, which is in line with the analysis that showed differences in 'ideal' cluster number by severity and infectious etiology (Extended Data Fig. 2). Importantly, we found that the thresholds for optimal sensitivity and specificity varied depending on whether the goal was to differentiate mild disease from healthy participants compared with differentiating severe or fatal disease from mild cases (Fig. 5d,e). In addition, these thresholds varied depending on whether a patient had a viral or a bacterial infection (Extended Data Fig. 6b,c). Together, these results further suggest that the differing number of endotypes identified by prior unsupervised approaches, even though they identified similar biology, may stem from the differences in cohort composition. Overall, these results also demonstrate the need for a flexible, generalizable framework to better evaluate immune dysregulation across these diverse clinical contexts.

Hi-DEF generalizes to other critical illness syndromes

Prior studies have highlighted the similar pathobiology underlying systemic inflammation in sepsis, burns and trauma[3]. We evaluated whether Hi-DEF could offer insights into other critical illness syndromes. We first examined the Glue Grant cohort[43], which included 430 noninfected, critically ill patients with trauma or burns. We integrated gene expression data from the Glue cohort with SUBSPACE data using COCONUT co-normalization and defined dysregulation as myeloid or lymphoid scores greater than or equal to the median scores across all SUBSPACE patients who required ICU level of care. Higher myeloid or lymphoid dysregulation scores were significantly associated with severe outcomes, defined as multi-system organ failure or mortality (OR = 2.0, 95% CI 1.1–3.7, $P = 0.02$; Extended Data Fig. 7a,b). This association was predominantly driven by myeloid dysregulation and remained significant after adjusting for sex and Acute Physiology and Chronic Health Evaluation II (APACHE II) score (adjusted myeloid dysregulation score: OR = 1.3, 95% CI 1.0–1.8, $P = 0.045$), whereas lymphoid dysregulation was not associated with multi-system organ failure or mortality.

Next, we evaluated whether Hi-DEF was associated with ARDS in the Stanford cohort using the same cut-off as the Glue Grant cohort. Higher myeloid or lymphoid dysregulation scores were significantly associated with the presence of ARDS (OR = 2.7, 95% CI 1.3–6.0, $P = 0.005$; Extended Data Fig. 7c,d). After adjusting for sex and APACHE II score, the lymphoid dysregulation score was significantly associated

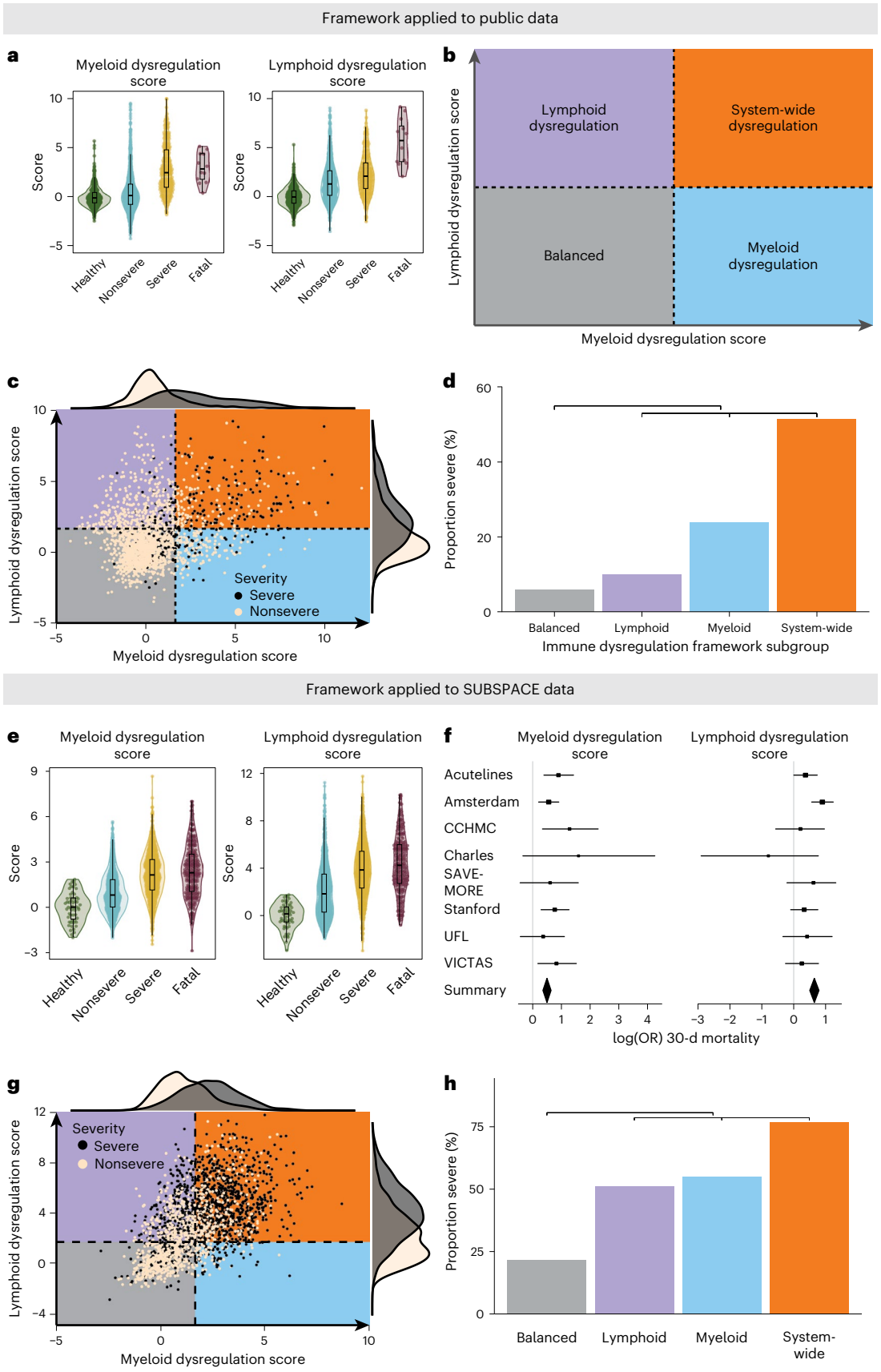
Fig. 4 | Evaluation of the immune dysregulation framework in public and SUBSPACE data. a, Myeloid (left) and lymphoid (right) scores (calculated as *z*-scores relative to healthy participants) calculated for all public samples ($n = 2{,}096$). The box plots represent the median and interquartile range (IQR) whereas whiskers represent the range of data excluding outliers (1.5× the IQR). The association of increasing scores (*y* axis) with increasing severity (*x* axis) was calculated using the JT *t*-test. Myeloid and lymphoid dysregulation scores were associated with severity ($P < 2.2 \times 10^{-16}$ for both scores). **b**, Theoretical framework for defining immune dysregulation with myeloid dysregulation on one axis and lymphoid dysregulation on the other. This provides a means of subgrouping patients into four subgroups depending on the level of dysregulation present: (1) balanced: both dysregulation scores low; (2) lymphoid dysregulation: only lymphoid dysregulation score elevated; (3) myeloid dysregulation: only myeloid dysregulation score elevated; and (4) system-wide dysregulation: both myeloid and lymphoid dysregulation scores elevated. **c**, The immune dysregulation framework applied to public co-normalized data ($n = 2{,}096$). Cut-offs are defined by a *z*-score of 1.65 relative to healthy participants. The black dots represent patients with severe infections (defined by ICU admission), whereas the tan dots represent nonsevere infections. **d**, Barplot representing the proportion of severe infections (*y* axis) by immune dysregulation framework subgroup (*x* axis). The OR was calculated using two-sided Fisher's exact test unadjusted for multiple comparisons, comparing patients with dysregulation on any axis relative to the balanced subgroup. Dysregulation on either the myeloid or the lymphoid axis (inclusive of lymphoid dysregulation ($n = 449$), myeloid dysregulation ($n = 197$) and system-wide dysregulation ($n = 259$) subgroups) was associated with severe infections with an OR of 5.2 (95% CI 3.9–7.0, $P < 2.2 \times 10^{-16}$) compared with patients in the balanced subgroup ($n = 1{,}191$). **e**, Myeloid (left) and lymphoid (right) scores (calculated as *z*-scores relative to healthy participants) were calculated for baseline SUBSPACE samples with phenotype data available ($n = 2{,}212$). The box plots represent the median and IQR whereas the whiskers represent the range of data excluding outliers (1.5× the IQR). The association of increasing scores (*y* axis) with severity (*x* axis) was calculated using the JT *t*-test. Myeloid dysregulation and lymphoid dysregulation scores were associated with severity ($P < 2.2 \times 10^{-16}$ for both scores). **f**, Forest plots showing log(OR) and 95% CIs of 30-d mortality across each site, calculated using logistic regression and showing strong association with mortality: myeloid dysregulation score (left) and lymphoid dysregulation score (right). Patient numbers by site were: ACUTELINES ($n = 275$), Amsterdam ($n = 717$), Cincinnati Children's Hospital Medical Center ($n = 184$), Charles University ($n = 12$), SAVE-MORE ($n = 452$), Stanford University ($n = 236$), University of Florida ($n = 172$) and VICTAS ($n = 137$). **g**, The immune dysregulation framework applied to SUBSPACE co-normalized data ($n = 2{,}212$). Cut-offs were defined by a *z*-score of 1.65 relative to healthy patients. The black dots represent critically ill patients, whereas the tan dots represent noncritically ill patients. **h**, Barplot representing the proportion of severe infections (*y* axis) by immune dysregulation framework subgroup (*x* axis). The OR was calculated using two-sided Fisher's exact test unadjusted for multiple comparisons, comparing patients with dysregulation on any axis relative to the balanced subgroup. Dysregulation on either the myeloid or the lymphoid axis (inclusive of lymphoid dysregulation ($n = 615$), myeloid dysregulation ($n = 133$) and system-wide dysregulation ($n = 1{,}050$) subgroups) was associated with severe infections with an OR of 7.1 (95% CI 5.6–8.9, $P < 2.2 \times 10^{-16}$) compared with patients in the balanced subgroup ($n = 564$).

https://doi.org/10.1038/s41591-025-03956-5

with ARDS (adjusted OR = 1.2, 95% CI 1.02–1.37, $P = 0.03$), but the myeloid dysregulation score was not. Together, these results suggest that Hi-DEF may provide insights into similarities and differences across diverse critical illness syndromes.

Hi-DEF generalizes to immunocompromised patients

To investigate the generalizability of Hi-DEF in immunosuppressed patients, we evaluated two cohorts (Stanford ICU and MESSI) that recruited from quaternary care center ICUs with substantial



https://doi.org/10.1038/s41591-025-03956-5

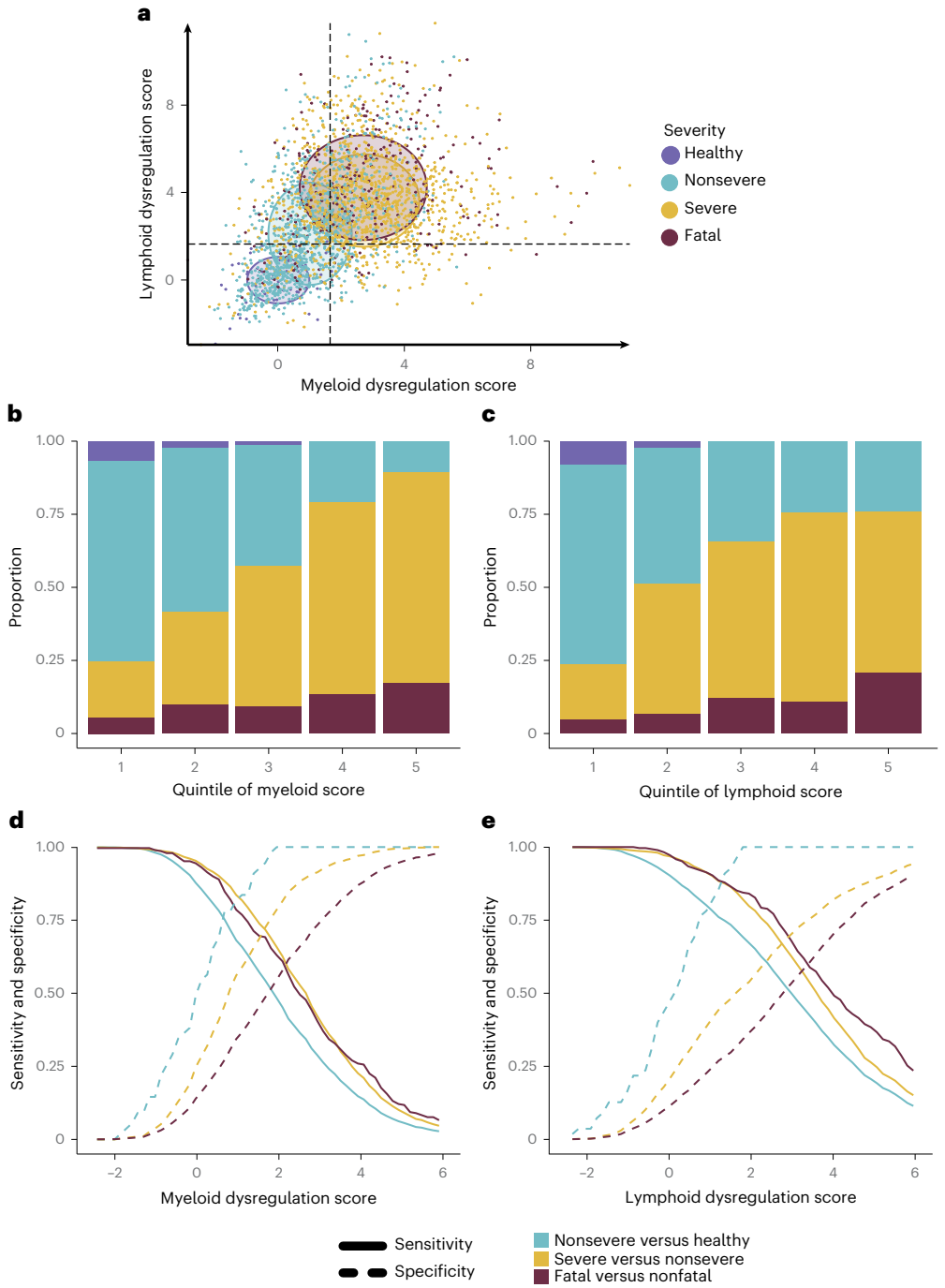


Fig. 5 | Ideal threshold for immune subgrouping identification depends on clinical cohort and question. a, Mean and s.d. of the myeloid (*x* axis) and lymphoid (*y* axis) dysregulation scores represented graphically, stratified by severity (defined by critical versus noncritical illness). The individual dots represent patient samples colored by severity and circles the mean and s.d. of lymphoid dysregulation scores. Together this shows the differences that patient cohort composition may have on endotype-defining gene expression signatures. **b**, Evaluation of the proportion of severity subgroups within the quintiles of myeloid dysregulation scores, indicating how different thresholds would affect the patient population above each threshold. **c**, Evaluation of the proportion of severity subgroups within the quintiles of lymphoid dysregulation scores, indicating how different thresholds would affect the patient population above each threshold. **d**, Evaluation of the sensitivity (solid line) and specificity (dashed line) of different myeloid dysregulation score thresholds when identifying mild or moderate infections from healthy participants (blue), severe from mild or moderate infections (yellow) and fatal from nonfatal (red) cases. These results show that the ideal dysregulation threshold will depend on the clinical question. **e**, Evaluation of the sensitivity (solid line) and specificity (dashed line) of different lymphoid dysregulation score thresholds when identifying mild or moderate infections from healthy participants (blue), severe from mild or moderate infections (yellow) and fatal from nonfatal (red) cases. These results show that the ideal dysregulation threshold will depend on the clinical question.

immunosuppressed populations. In the Stanford and MESSI cohorts, 28% and 46% of patients, respectively, were immunocompromised. Myeloid and lymphoid dysregulation scores were significantly higher in immunocompromised patients in the Stanford cohort (Wilcoxon's $P = 0.002$, $P = 0.02$, respectively) but were not different in the MESSI cohort. In the Stanford cohort, immunocompromised patients were more likely to be dysregulated (OR = 2.8, 95% CI 1.3–6.7, $P = 0.006$), but not in the MESSI cohort (Extended Data Fig. 8a–d). In both cohorts,

https://doi.org/10.1038/s41591-025-03956-5

although immunocompromised status was associated with worse outcomes, this did not differ significantly by assigned subgroup (Extended Data Fig. 8e,f). In both cohorts, the myeloid dysregulation score remained significantly associated with 30-d mortality after adjustment for immune status ($P = 0.004$ and $P < 0.001$, respectively). Overall, these results suggest that Hi-DEF is not significantly affected by baseline immunocompromise as defined in these cohorts and can be used to further substratify this high-risk population.

Hi-DEF is associated with differential response to immune-modulating medications across critical illnesses

Numerous clinical trials of immune-modulating agents in critical illness have been negative, which is often attributed to underlying biological heterogeneity causing differential treatment response[4]. We hypothesized that our proposed Hi-DEF would reduce the biological heterogeneity and be associated with a differential treatment response.

To test this hypothesis, we first used the SAVE-MORE cohort, a randomized controlled trial of anakinra in noncritically ill, hospitalized patients with COVID-19 and elevated soluble urokinase plasminogen-activating receptor, which showed a mortality benefit of anakinra in the entire cohort[44]. In this study, dysregulation was defined by a *z*-score greater than or equal to the median dysregulation score in noncritically ill, infected patients across all SUBSPACE cohorts. Patients with lymphoid dysregulation at baseline treated with anakinra had a significantly lower rate of 28-d mortality (2.2%) compared with those treated with placebo (20.8%, Fisher's $P = 0.02$, P (interaction) = 0.05; Fig. 6a). There was no difference in 28-d mortality rate in patients without baseline lymphoid dysregulation ($P = 0.77$). It is interesting that the subgroup of patients with lymphoid dysregulation experienced the highest mortality benefit from anakinra, but those with only myeloid dysregulation did not (Supplementary Fig. 13a,b). This survival benefit in patients with lymphoid dysregulation remained significant even after adjustment for age, sex and baseline sequential organ failure assessment (SOFA) score (adjusted hazard ratio (HR) = 0.06, 95% CI 0.008–0.53, $P = 0.01$; Fig. 6b). Together these results suggest that anakinra preferentially benefits patients with baseline lymphoid dysregulation.

Next, we evaluated whether Hi-DEF was associated with a differential response to corticosteroids using two independent studies. The VICTAS trial was a randomized controlled trial of hydrocortisone, vitamin C and thiamine in 501 patients with sepsis[45]. A subset of patients ($n = 141$) had blood transcriptome data available. We excluded the 52 (37%) patients who received open-label steroids (and were thus randomized only to receipt of thiamine and vitamin C versus placebo). Patients were divided into subgroups based on whether dysregulation scores were greater than or equal to median scores across all infected, critically ill patients in the SUBSPACE consortium. In this limited cohort of patients with available RNA-seq data, there was a trend toward mortality benefit (26% mortality rate in placebo versus 11% with the three-drug active treatment; Fisher's $P = 0.11$). Again, this apparent benefit was driven by the patients with lymphoid dysregulation at baseline. In patients with high lymphoid dysregulation scores, those treated with hydrocortisone had significantly lower mortality rates compared with those in the placebo arm (11% versus 39%, OR = 0.20, 95% CI 0.03–1.06, Fisher's $P = 0.03$; Fig. 6c). This survival difference was robust after adjustment for age and sex (adjusted HR = 0.22, 95% CI 0.06–0.85, $P = 0.03$; Fig. 6d). The differences were not observed in patients with myeloid dysregulation (Supplementary Fig. 13c,d).

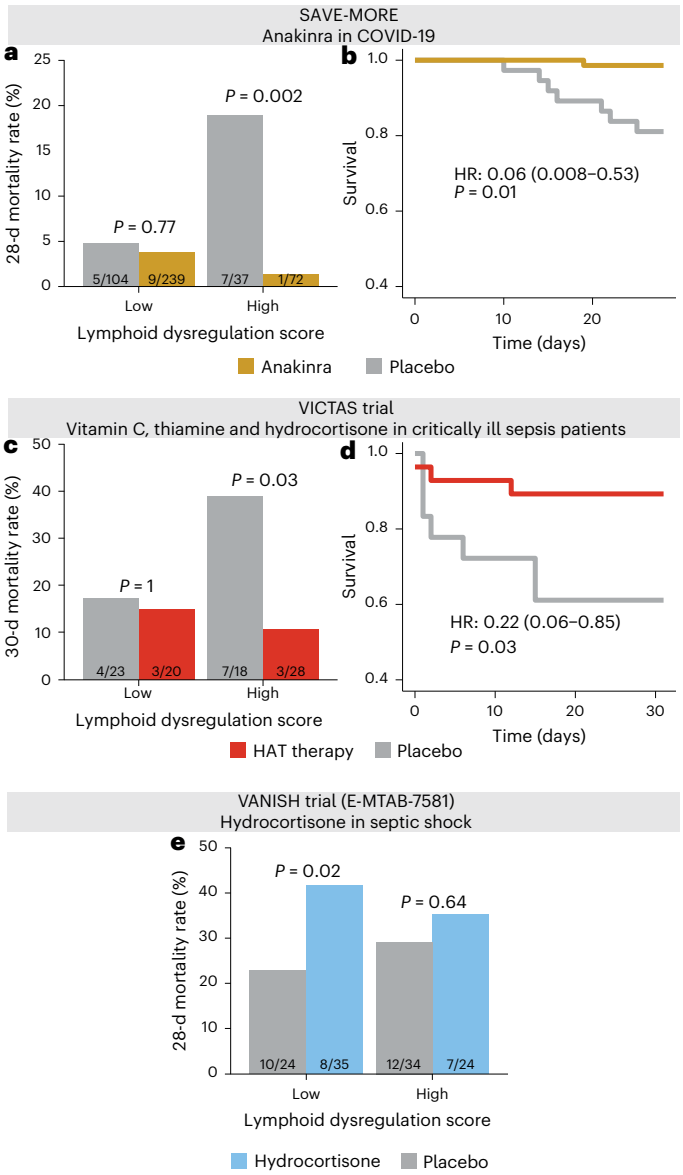


Fig. 6 | Association of lymphoid immune dysregulation with treatment. **a**, Evaluation of 28-d mortality rate on the *y* axis stratified by high and low lymphoid dysregulation scores (defined by *z*-score ≥ 1.65) and anakinra (gold) versus placebo (gray) treatment in patients with COVID-19 in the SAVE-MORE clinical trial, using two-sided Fisher's exact test unadjusted for multiple comparisons. Dysregulation was defined based on median scores across all noncritically ill, infected patients in SUBSPACE. Lymphoid dysregulation is associated with a disproportionate benefit from anakinra therapy relative to patients with low (balanced) lymphoid responses. **b**, Evaluation of Kaplan–Meier survival curve for 28-d survival in patients with lymphoid dysregulation stratified by anakinra (gold) and placebo (gray) in the SAVE-MORE trial. Cox's proportional HR is adjusted for age, sex and SOFA score. **c**, Evaluation of 30-d mortality (*y* axis) in the VICTAS trial (a randomized controlled trial of vitamin C, thiamine and hydrocortisone in critically ill patients with sepsis) stratified by high and low lymphoid dysregulation score and treatment (red) versus placebo (gray). Dysregulation was defined by the median score across all infected, critically ill patients in SUBSPACE, and the significance was assessed using two-sided Fisher's exact test unadjusted for multiple comparisons. The results indicate that lymphoid dysregulation was associated with disproportionate benefit from steroids, vitamin C and thiamine therapy. **d**, Evaluation of Kaplan–Meier survival curve for 30-d survival in patients with lymphoid dysregulation stratified by treatment (red) versus placebo (gray) in the VICTAS trial. Cox's proportional HR is adjusted for age and sex. **e**, Evaluation of the 28-d mortality rate (*y* axis) in the VANISH trial (a clinical trial of hydrocortisone in patients with septic shock) stratified by high and low lymphoid dysregulation score (defined by median score) and randomized steroid treatment (red). The significance was assessed using two-sided Fisher's exact test unadjusted for multiple comparisons. Patients with a low (balanced) lymphoid dysregulation score were disproportionately harmed by steroid therapy.

We also evaluated whether lymphoid dysregulation was associated with a differential response to corticosteroids in the VANISH cohort, a randomized controlled, factorial trial that evaluated norepinephrine versus vasopressin and hydrocortisone versus placebo in patients with septic shock[46], with no difference in mortality rate related to hydrocortisone administration in the overall trial of 409 patients. In the subset of 176 patients with RNA expression data, hydrocortisone treatment was associated with a 38% mortality rate in those treated with hydrocortisone, compared with 22% in those not treated with hydrocortisone ($P = 0.03$). Unlike the SAVE-MORE and VICTAS trials, where the benefit of treatment was limited to subgroups with lymphoid dysregulation, this difference in the VANISH cohort was driven largely by increased mortality in patients with low (balanced) lymphoid dysregulation scores treated with hydrocortisone relative to those who did not receive steroids (28-d mortality rate 42% versus 16%, OR = 3.8, 95% CI 1.2–12.6, $P = 0.02$; Fig. 6e). This difference was not seen based on myeloid dysregulation (Supplementary Fig. 13e,f).

Collectively, these results demonstrated that Hi-DEF provides flexibility for context-specific evaluation and has the potential to identify appropriate immunomodulatory treatment for patients with critical illness, reducing the heterogeneity of treatment effects.

Discussion

In this study, we demonstrated that previously defined sepsis endotypes are biologically similar, leading to the identification of four consensus endotypes. We further found that these endotypes were defined by detrimental and protective immune responses and cellular origin (myeloid or lymphoid). Based on these results, we proposed an immune dysregulation evaluation framework, Hi-DEF, that is defined by two continuous scores and generalizes to both infectious and noninfectious critical illnesses, including sepsis, burn, trauma and ARDS, irrespective of patient age and immunosuppression status. Finally, we demonstrated that Hi-DEF could identify molecularly homogeneous groups of patients with a differential response to anakinra in COVID-19 and steroids in sepsis, suggesting its potential application for targeted therapeutic intervention.

Our findings mirror those identified by Scicluna et al.[47], who used independent cohorts of predominantly bacterial sepsis to assess overlap among a subset of endotyping schemas analyzed here (SRS, MARS and Sweeney). They identified three consensus transcriptomic subtypes (CTSs) that align with our findings: CTS1 includes SRS1, MARS2 and inflammopathic endotypes; CTS2 includes MARS1 and coagulopathic endotypes; and CTS3 includes SRS2, MARS3 and adaptive endotypes. CTS1, CTS2 and CTS3 correspond broadly to myeloid dysregulation, lymphoid dysregulation and lymphoid protective clusters in our results. The fourth myeloid protective cluster defined in our analysis, which included MARS4 and SoM module 3, is predominantly driven by mature neutrophil and monocyte populations and interferon responses. Its identification in this study is likely due to the inclusion of additional protective myeloid scores (SoM module 3 and the Wong signature) as well as inclusion of numerous viral infections, where monocyte interferon responses play a crucial role, further highlighting the importance of context in identifying clinically relevant subgroups.

Importantly, we found that the number of endotypes varied based on disease etiology and severity, suggesting the number of 'clinically relevant' endotypes may depend on the clinical question posed. For instance, if the goal is prognostication, two endotypes (high risk versus low risk) based on lower dimensional 'system-wide' dysregulation may be sufficient, such as is seen for ARDS latent class analysis subphenotypes, the TriVerity severity score and the original SRS1 and SRS2 endotypes[12,48,49]. However more nuanced clinical trial designs may rely on and benefit from sub-phenotyping based on specified myeloid or lymphoid dysregulated endotypes. Finally, any immune response is context dependent, with appropriate responses depending on the severity of the insult and/or pathogen. Therefore, we believe that Hi-DEF, in which axes of myeloid and lymphoid dysregulation may be used in isolation or collectively, has the potential to define immune dysregulation across critical illness syndromes and allow for rapid advancements in the field of critical care.

Our findings provide further evidence that neutrophils, particularly immature neutrophils, and loss of protective T and natural killer (NK) cells are associated with infection severity[50–53]. It is interesting that lymphoid dysregulation was more strongly associated with severe viral infections, whereas myeloid dysregulation showed a stronger link to severity in bacterial infections, burns or trauma. This aligns with the distinct immune pathways activated by viral versus bacterial pathogens and highlights the utility of Hi-DEF's flexibility. Notably, we found that ARDS was more closely linked with lymphoid dysregulation, even though myeloid dysregulation was more associated with mortality. Prior studies have suggested that T cell dysregulation may be associated with ARDS and these results warrant further evaluation across other ARDS cohorts[54].

In addition, the association of lymphoid dysregulation with a differential treatment response to corticosteroids is in line with prior studies[5,18]. As interleukin-1 is predominantly released by myeloid cells, however, the differential treatment effect to anakinra seen in this subset of patients is counterintuitive. Overall, 'lymphoid dysregulation' in this study is related to loss of protective lymphocyte subsets, potentially correlating with the 'immune exhaustion' state that has previously been described[12]. In particular, T or NK cells are known to play an important role in mitigating inflammation in the macrophage activation syndrome; thus loss or dysfunction of these cells likely plays a key role in the uncontrolled inflammation that anakinra is targeting[55,56]. Our results suggest that limiting cytokine activation in this 'exhausted' immune state may be beneficial and show that further study into the mechanism of anakinra's benefits in this subgroup is indicated.

Our study has several strengths. To our knowledge, this is the largest multi-cohort analysis to date aimed at understanding the biology of critical illness. Integrating >7,000 samples with rich metadata enabled robust evaluation of the similarities between endotyping schemas, their comparison to clinical markers and their association with outcomes. The SUBSPACE cohorts and gene expression data represent a monumental step forward for critical care transcriptomic research because these 4,106 samples enriched for high severity patients have not previously been evaluated. The significance of these findings across multiple gene expression measurement techniques, cohorts, age groups, etiologies and disease states adds to the credibility and generalizability of these findings. The use of healthy participants, when possible, to define dysregulation further increases the generalizability of these findings and could facilitate cross-platform quantification and endotyping. Inclusion of noninfectious critical care data provides evidence of the similarities in systemic dysregulation. The inclusion of single-cell data allowed for nuanced evaluation for the underlying biology of these findings. The inclusion of treatment data shows the association of these scores with treatment outcomes and the potential of this framework to advance precision medicine.

Our study has several limitations. Hi-DEF aims to provide a unifying template for quantifying immune dysregulation by consolidating published endotyping signatures and is purposefully simplistic. In particular, the continuous scores outlined here were designed as a proof of concept and are a conglomerate of genes derived from the signatures used to identify these consensus endotypes. We included all genes that met inclusion criteria and defined detrimental and protective genes based on empirical evidence from prior studies. Thus, although these genes in aggregate were somewhat selective for myeloid and lymphoid compartments, more selective cell-type-specific genes may be identified. More discovery-oriented approaches may also identify other relevant genes or cell-specific or pathway-specific dysregulation. Finally,

the genes and thresholds used in this study were used to demonstrate the potential for clinical utility. These scores and thresholds require further fine-tuning depending on the patient, disease and treatment contexts. In addition, the immune dysregulation quantified by Hi-DEF is correlated with, but may not be causative of, severe outcomes. These samples were also predominantly collected within the first 24–48 h of hospitalization or ICU admission and disproportionately represent patients from the United States of America and Europe. Finally, across all clinical trials assessed for differential treatment response, only a subgroup of patients in the broader studies underwent gene expression analysis, which may introduce bias. Furthermore, the analyses of treatment responsiveness were post-hoc and exploratory in these small subgroups. Thus, these results are preliminary and should be interpreted with caution. Future prospective studies across broader population and infectious etiologies are needed to explore the mechanistic underpinnings of these results and validate these findings.

In summary, Hi-DEF provides a launching point for further prospective multi-omic investigations into critical illness immunobiology, as well as a more readily translatable, biology-based schema for developing precision medicine tools in the ICU. The Hi-DEF framework provides a shared, measurable foundation for biological endotyping and addresses several key issues observed with other sub-phenotyping schemas to facilitate clinical translation. First, by pinpointing biological pathways that are specifically enriched within each of the four consensus endotypes, it has the potential to help streamline candidate therapeutic targets for future hypothesis-driven mechanistic studies. Second, identifying cell-specific drivers of these consensus endotypes could pave the way for precision therapies aimed at modulating particularly detrimental cell subsets or states. In addition, our ability to quantify immune dysregulation on a continuous score further allows for adjustments based on patient, infection or insult type and the question being asked (that is, diagnosis, prognosis, treatment). Third, Hi-DEF is flexible and could be further refined and extended. For example, future research into more specific immature neutrophils, monocytes or T or NK cell subsets or addition of other cell types may enhance this framework. Perhaps more importantly, the methodology used here to build an endotyping consensus could provide a scaffold to build similar consensus and biological frameworks in heterogeneous disease states. Finally, using multiple gene expression measurement platforms (microarrays and RNA-seq) and defining dysregulation based on healthy participants provides an opportunity to facilitate translation to a rapid point-of-care measurement platform.

Together, these factors highlight the clinical potential of Hi-DEF as well as a path forward for this and other endotyping signatures. Although cross-platform validation is difficult and will require prospective validation, recent advances in RNA-seq technology have enabled the development of point-of-care, multivariate, gene expression signatures such as the TriVerity or Myrna platform, which is a sepsis diagnostic test that quantifies expression of 29 genes in approximately 30 min[49]. Future studies should focus on three key factors for clinical translation: (1) development of the framework into a parsimonious gene signature and identification of appropriate weighting and thresholding based on context to facilitate development and interpretation of point-of-care testing; (2) prospective validation of the utility of this framework and integration with existing clinical decision tools; and (3) measurement in therapeutic and platform trials to facilitate precision medicine clinical trials in critical care.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-025-03956-5.

References

1. Berthelsen, P. G. & Cronqvist, M. The first intensive care unit in the world: Copenhagen 1953. *Acta Anaesthesiol. Scand.* **47**, 1190–1195 (2003).
2. Zimmerman, J. E., Kramer, A. A. & Knaus, W. A. Changes in hospital mortality for United States intensive care unit admissions from 1988 to 2012. *Crit. Care* **17**, R81 (2013).
3. Maslove, D. M. et al. Redefining critical illness. *Nat. Med.* **28**, 1141–1148 (2022).
4. Marshall, J. C. Why have clinical trials in sepsis failed? *Trends Mol. Med.* **20**, 195–203 (2014).
5. Antcliffe, D. B. et al. Transcriptomic signatures in sepsis and a differential response to steroids. From the VANISH randomized trial. *Am. J. Respir. Crit. Care Med.* **199**, 980–986 (2019).
6. Calfee, C. S. et al. ARDS subphenotypes and differential response to simvastatin: secondary analysis of a randomized controlled trial. *Lancet Respir. Med.* **6**, 691–698 (2018).
7. Famous, K. R. et al. Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. *Am. J. Respir. Crit. Care Med.* **195**, 331–338 (2017).
8. Sinha, P. et al. Latent class analysis reveals COVID-19-related acute respiratory distress syndrome subgroups with differential responses to corticosteroids. *Am. J. Respir. Crit. Care Med.* **204**, 1274–1285 (2021).
9. Meyer, N. J. et al. Mortality benefit of recombinant human interleukin-1 receptor antagonist for sepsis varies by initial interleukin-1 receptor antagonist plasma concentration. *Crit. Care Med.* **46**, 21–28 (2018).
10. Ganesan, A. et al. A conserved immune dysregulation signature is associated with infection severity, risk factors prior to infection, and treatment response. *Immunity* **58**, 1–16 (2025).
11. Stanski, N. L. & Wong, H. R. Prognostic and predictive enrichment in sepsis. *Nat. Rev. Nephrol.* **16**, 20–31 (2020).
12. Davenport, E. E. et al. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *Lancet Respir. Med.* **4**, 259–271 (2016).
13. Sweeney, T. E. et al. A community approach to mortality prediction in sepsis via gene expression analysis. *Nat. Commun.* **9**, 694 (2018).
14. Sweeney, T. E. et al. Unsupervised analysis of transcriptomics in bacterial sepsis across multiple datasets reveals three robust clusters. *Crit. Care Med.* **46**, 915–925 (2018).
15. Wong, H. R. et al. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med.* **7**, 34 (2009).
16. Cano-Gamez, E. et al. An immune dysfunction score for stratification of patients with acute infection based on whole-blood gene expression. *Sci. Transl. Med.* **14**, eabq4433 (2022).
17. Wong, H. R. et al. Developing a clinically feasible personalized medicine approach to pediatric septic shock. *Am. J. Respir. Crit. Care Med.* **191**, 309–315 (2015).
18. Yao, L. et al. Gene expression scoring of immune activity levels for precision use of hydrocortisone in vasodilatory shock. *Shock* **57**, 384–391 (2022).
19. Zheng, H. et al. Multi-cohort analysis of host immune response identifies conserved protective and detrimental modules associated with severity across viruses. *Immunity* **54**, 753–768 (2021).
20. Scicluna, B. P. et al. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. *Lancet Respir. Med.* **5**, 816–826 (2017).
21. Waddington, C. S. et al. An outpatient, ambulant-design, controlled human infection model using escalating doses of *Salmonella typhi* challenge delivered in sodium bicarbonate solution. *Clin. Infect. Dis.* **58**, 1230–1240 (2014).

22. Tang, B. M. et al. Neutrophils-related host factors associated with severe disease and fatality in patients with influenza infection. *Nat. Commun.* **10**, 3422 (2019).
23. Wallihan, R. G. et al. Molecular distance to health transcriptional score and disease severity in children hospitalized with community-acquired pneumonia. *Front. Cell. Infect. Microbiol.* **8**, 382 (2018).
24. Blohmke, C. J. et al. Diagnostic host gene signature for distinguishing enteric fever from other febrile diseases. *EMBO Mol. Med.* **11**, e10431 (2019).
25. Yu, J. et al. Host gene expression in nose and blood for the diagnosis of viral respiratory infection. *J. Infect. Dis.* **219**, 1151–1161 (2019).
26. Ambite, I. et al. Fimbriae reprogram host gene expression—divergent effects of P and type 1 fimbriae. *PLoS Pathog.* **15**, e1007671 (2019).
27. Petzke, M. M. et al. Global transcriptome analysis identifies a diagnostic signature for early disseminated Lyme disease and its resolution. *mBio* **11**, e00047-20 (2020).
28. Berdal, J.-E. et al. Excessive innate immune response and mutant D222G/N in severe A (H1N1) pandemic influenza. *J. Infect.* **63**, 308–316 (2011).
29. Banchereau, R. et al. Host immune transcriptional profiles reflect the variability in clinical disease manifestations in patients with Staphylococcus aureus infections. *PLoS ONE* **7**, e34390 (2012).
30. Mejias, A. et al. Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection. *PLoS Med.* **10**, e1001549 (2013).
31. Alder, M. N., Opoka, A. M., Lahni, P., Hildeman, D. A. & Wong, H. R. Olfactomedin-4 Is a candidate marker for a pathogenic neutrophil subset in septic shock. *Crit. Care Med.* **45**, e426–e432 (2017).
32. Heinonen, S. et al. Rhinovirus detection in symptomatic and asymptomatic children: value of host transcriptome analysis. *Am. J. Respir. Crit. Care Med.* **193**, 772–782 (2016).
33. Jaggi, P. et al. Whole blood transcriptional profiles as a prognostic tool in complete and incomplete Kawasaki disease. *PLoS ONE* **13**, e0197858 (2018).
34. Zhai, Y. et al. Host transcriptional response to influenza and other acute respiratory viral infections—a prospective cohort study. *PLoS Pathog.* **11**, e1004869 (2015).
35. Lindow, J. C. et al. Cathelicidin insufficiency in patients with fatal leptospirosis. *PLoS Pathog.* **12**, e1005943 (2016).
36. de Steenhuijsen Piters, W. A. A. et al. Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. *Am. J. Respir. Crit. Care Med.* **194**, 1104–1115 (2016).
37. Thair, S. A. et al. Transcriptomic similarities and differences in host response between SARS-CoV-2 and other viral infections. *iScience* **24**, 101947 (2021).
38. Soares-Schanoski, A. et al. Systems analysis of subjects acutely infected with the Chikungunya virus. *PLoS Pathog.* **15**, e1007880 (2019).
39. Schulte-Schrepping, J. et al. Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* **182**, 1419–1440 (2020).
40. Wilk, A. J. et al. Multi-omic profiling reveals widespread dysregulation of innate immunity and hematopoiesis in COVID-19. *J. Exp. Med.* **218**, e20210582 (2021).
41. Combes, A. J. et al. Global absence and targeting of protective immune states in severe COVID-19. *Nature* **591**, 124–130 (2021).
42. Sinha, S. et al. Dexamethasone modulates immature neutrophils and interferon programming in severe COVID-19. *Nat. Med.* **28**, 201–211 (2022).
43. Tompkins, R. G. Genomics of injury: the Glue Grant experience. *J. Trauma Acute Care Surg.* **78**, 671–686 (2015).
44. Kyriazopoulou, E. et al. Early treatment of COVID-19 with anakinra guided by soluble urokinase plasminogen receptor plasma levels: a double-blind, randomized controlled phase 3 trial. *Nat. Med.* **27**, 1752–1760 (2021).
45. Sevransky, J. E. et al. Effect of vitamin C, thiamine, and hydrocortisone on ventilator- and vasopressor-free days in patients with sepsis: the VICTAS randomized clinical trial. *JAMA* **325**, 742–750 (2021).
46. Gordon, A. C. et al. Effect of early vasopressin vs norepinephrine on kidney failure in patients with septic shock: the VANISH randomized clinical trial. *JAMA* **316**, 509–518 (2016).
47. Scicluna, B. P. et al. A consensus blood transcriptomic framework to advance precision medicine in sepsis. *Nat. Med.* https://doi.org/10.1038/s41591-025-03964-5 (2025).
48. Calfee, C. S. et al. Latent class analysis of ARDS subphenotypes: analysis of data from two randomized controlled trials. *Lancet Respir. Med.* **2**, 611–620 (2014).
49. Liesenfeld, O. et al. Rapid and accurate diagnosis and prognosis of acute infections and sepsis from whole blood using host response mRNA amplification and result interpretation by machine-learning classifiers. *Nat. Med.* https://doi.org/10.1038/s41591-025-03933-y (2025).
50. Prozan, L. et al. Prognostic value of neutrophil-to-lymphocyte ratio in COVID-19 compared with influenza and respiratory syncytial virus infection. *Sci. Rep.* **11**, 21519 (2021).
51. Nahm, C. H., Choi, J. W. & Lee, J. Delta neutrophil index in automated immature granulocyte counts for assessing disease severity of patients with sepsis. *Ann. Clin. Lab. Sci.* **38**, 241–246 (2008).
52. Park, B. H. et al. Delta neutrophil index as an early marker of disease severity in critically ill patients with sepsis. *BMC Infect. Dis.* **11**, 299 (2011).
53. Kwok, A. J. et al. Neutrophils and emergency granulopoiesis drive immune suppression and an extreme response endotype during sepsis. *Nat. Immunol.* **24**, 767–779 (2023).
54. Neyton, L. P. A. et al. Host and microbe blood metagenomics reveals key pathways characterizing critical illness phenotypes. *Am. J. Respir. Crit. Care Med.* **209**, 805–815 (2024).
55. Villanueva, J. et al. Natural killer cell dysfunction is a distinguishing feature of systemic onset juvenile rheumatoid arthritis and macrophage activation syndrome. *Arthritis Res. Ther.* **7**, R30–37 (2005).
56. Schulert, G. S. & Grom, A. A. Macrophage activation syndrome and cytokine-directed therapies. *Best Pract. Res. Clin. Rheumatol.* **28**, 277–292 (2014).

Andrew R. Moore[1,2,3], Hong Zheng[2,3], Ananthakrishnan Ganesan[2,3], Yehudit Hasin-Brumshtein[4], Manoj V. Maddali[1,3], Joseph E. Levitt[1], Tom van der Poll[5,6], Jingyi Lu[7,8], Hjalmar R. Bouma[7,8,9], Brendon P. Scicluna[10], Evangelos J. Giamarellos-Bourboulis[11,12], Antigone Kotsaki[11,12], Ignacio Martin-Loeches[13,14], Alexis Garduno[13], Jeremiah Hinson[15], Richard E. Rothman[15], Jonathan Sevransky[16], David W. Wright[17], Mihir R. Atreya[18], Lyle L. Moldawer[19], Philip A. Efron[19], Marcela Kralovcova[20], Thomas Karvunidis[20], Heather M. Giannini[21], Nuala J. Meyer[21], Timothy E. Sweeney[4,22], Angela J. Rogers[1,22] & Purvesh Khatri[2,3,22]

[1]Division of Pulmonary, Allergy and Critical Care Medicine, Stanford University, Stanford, CA, USA. [2]Institute for Immunity, Transplantation and Infection, Stanford University, Stanford, CA, USA. [3]Center for Biomedical Informatics Research, Department of Medicine, Stanford University, Stanford, CA, USA. [4]Inflammatix Inc., Sunnyvale, CA, USA. [5]Center of Experimental and Molecular Medicine, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands. [6]Division of Infectious Diseases, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands. [7]Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [8]Department of Acute Care, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [9]Department of Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [10]Department of Applied Biomedical Science, Faculty of Health Sciences, and Centre for Molecular Medicine and Biobanking, University of Malta, Msida, Malta. [11]4th Department of Internal Medicine, National and Kapodistrian University of Athens, Medical School, Athens, Greece. [12]Hellenic Institute for the Study of Sepsis, Athens, Greece. [13]Department of Intensive Care Medicine, Multidisciplinary Intensive Care Research Organization (MICRO), St James's Hospital, Dublin, Ireland. [14]Hospital Clinic, Universitat de Barcelona, IDIBAPS, CIBERES, Barcelona, Spain. [15]Department of Emergency Medicine, The Johns Hopkins University, Baltimore, MD, USA. [16]Division of Pulmonary, Allergy, Critical Care and Sleep and Emory Critical Care Center, Emory University, Atlanta, GA, USA. [17]Department of Emergency Medicine, Emory University, Atlanta, GA, USA. [18]Division of Critical Care Medicine, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA. [19]Sepsis and Critical Illness Research Center and the SPIES Consortium, University of Florida College of Medicine, Gainesville, FL, USA. [20]1st Department of Internal Medicine, Faculty of Medicine, Teaching Hospital and Biomedical Center in Pilsen, Charles University, Pilsen, Czech Republic. [21]Division of Pulmonary, Allergy, and Critical Care Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [22]These authors jointly supervised this work: Timothy E. Sweeney, Angela J. Rogers, Purvesh Khatri. e-mail: pkhatri@stanford.edu

Methods

Inclusion and ethics

All samples were collected in accordance with site-specific institutional review board (IRB) protocols and complied with ethical principles set forth by the Helsinki Declaration of 1975. Individual approvals for each site are listed below:

- ACUTELINES: the Medical Ethics Board and the Central Review Board of the University Medical Center (UMC) Groningen have evaluated and approved the protocol of Acutelines (number 2019/589).
- Amsterdam: the Medical Ethics Committee of the Amsterdam UMC, location AMC has given approval for the conduct of the ELDER-BIOME study (number NL57847.018.16), the OPTIMACT study (nos 2016/280 and NL57923.018.16) and the PANAMO study (IRB number 2020_067#B2020179).
- Charles University: the Charles University IRB approved sampling, handling, research and storage or biobanking of genetic material of participants in the IMHOTEP clinical study evaluated here (decision or approval number 251/2022 attached).
- Cincinnati Children's Hospital Medical Center: the study protocol was approved by the IRBs of the primary site (Cincinnati Children's Hospital, Genomic Analysis of Pediatric Systemic Inflammatory Syndrome, IRB number 2008-0558).
- SAVE-MORE: the protocol was approved by the National Ethics Committee of Greece (approval number 161/20) and the Ethics Committee of the National Institute for Infectious Diseases Lazzaro Spallanzani (RCCS) in Rome (1 February 2021) (EudraCT number 2020-005828-11; ClinicalTrials.gov: NCT04680949).
- Stanford University: the Stanford Biorepository study protocol was approved by the Stanford University IRB (number 28205).
- Trinity University: the study protocol was approved by the Tallaght University Hospital IRB (study title: Sepsis immunosuppression in critically ill patients; project ID: sjh428).
- University of Florida: the University of Florida IRB approved the SPIES clinical study evaluated here (IRB number 202000924).
- University of Pennsylvania: the MESSI cohort study protocol was approved by the University of Pennsylvania (IRB number 808542).
- VICTAS: the VICTAS clinical trial and subsequent sample storage and handling were approved by the Johns Hopkins University IRB (nos 00102528 and 00164053).

Consent was obtained from individuals or their legally authorized representatives per each study's protocols. Participants were not compensated for involvement in this study. All age groups, races, ethnicities and sex or genders that met inclusion criteria with available gene expression data were included.

Statistics and reproducibility

This study was designed as a large-scale biological evaluation of existing sepsis signatures. Analysis was split across three primary cohorts: public data, single-cell data and SUBSPACE data (Extended Data Table 1). Public data were collected through systematic review of publicly available whole-blood and peripheral blood mononuclear cell gene expression data from the Genome Expression Omnibus (GEO) and ArrayExpress. All samples across all ages, with healthy controls and necessary severity metadata and gene expression data, were included in the analysis with an expectation of collecting >1,000 samples to be able to adequately assess the biological overlap of signatures. SUBSPACE was a consortium formed to collate existing biobanked whole-blood gene expression data from international collaborators with the intention to collect >4,000 samples. Samples were inclusive of all ages from neonate to >80-year-old individuals. Sex was self-reported across all sites.

All signatures, including the framework, were calculated blinded to patient phenotypes and clinical outcomes. Unsupervised clustering analysis methodologies (hierarchical clustering, PCA and network analysis) were pre-planned and performed in parallel across public and SUBSPACE data to ensure replicability with iterations across subgroups to ensure robustness of results.

Single-cell samples were collected from publicly available peripheral blood single-cell data inclusive of the neutrophil compartment in patients with infections. Sepsis signatures were calculated blinded to patient phenotype and collated according to the predefined clusters to evaluate the single-cell biology leading to sepsis dysregulation. Criteria to select genes for Hi-DEF were predefined and all genes meeting inclusion criteria were included for all subsequent analyses.

Subgroup cut-offs within Hi-DEF were predefined and subgroups were identified blinded to clinical outcomes. The analytical plan of outcomes and differential response to treatment were predefined. All samples that included necessary phenotypic information were included in primary analyses and performed in parallel when possible across public and SUBSPACE data to ensure reproducibility. Sensitivity analyses by site and clinical characteristics were performed to ensure robustness.

Publicly available data curation and co-normalization

We performed a systematic review of publicly available whole-blood and peripheral blood mononuclear cell gene expression data from infected individuals in the GEO and ArrayExpress (Supplementary Table 1). Cohorts were assessed and studies were excluded if they did not have healthy participants, the necessary severity metadata or the gene expression data needed for score calculation. Patients were assigned as severe versus nonsevere, with severe being defined based on ICU admission requirement or mortality.

Cohorts were co-normalized using COCONUT co-normalization. Co-normalization was assessed by evaluating expression of housekeeping genes and through UMAP analysis.

The SUBSPACE consortium: curation, sequencing and co-normalization

The SUBSPACE consortium is an international consortium of researchers focused on developing a better understanding of the underlying biology behind sepsis endotypes. Institutions and patient characteristics are outlined in Supplementary Table 2.

Before processing, samples in PAXgene Blood RNA tubes were removed from −80 °C to thaw at room temperature for 2 h. The samples were then inverted several times to achieve homogeneity. RNA was isolated using the PAXgene Blood miRNA Kit (QIAGEN) according to the manufacturer's instructions with an elution volume of 80 µl.

The library preparation was done using the QIAseq Stranded Total RNA Library Kit with QIAseq FastSelect rRNA and globin depletion. The amount of 100 ng of starting material was heat fragmented. QIAseq FastSelect rRNA Globin or HMR was used to reduce the amount of unwanted RNA species. After first-strand and second-strand synthesis, the complementary DNA was end-repaired and 3′-adenylated. Sequencing adapters were ligated to the overhangs. Adapted molecules were enriched by using 18 cycles of PCR and purified by a bead-based cleanup. Library preparation was quality controlled using capillary electrophoresis (Tape D1000). High-quality libraries were pooled based on equimolar concentrations. The library pool(s) were quantified using quantitative PCR and the optimal concentration of the library pool used to generate the clusters on the surface of a flow cell before sequencing on a NovaSeq 6000 (Illumina Inc.) instrument (on four S4 flowcells, 2× 75, 2× 10), according to the manufacturer's instructions.

Data were co-normalized using COCONUT co-normalization. The MESSI cohort was excluded from co-normalization due to lack of healthy participants and used as a separate validation cohort. Co-normalization was assessed through evaluation of housekeeping genes and UMAP analysis.

Transcriptomic signature calculation

We applied a total of seven previously defined gene expression sepsis endotyping signatures: Sweeney endotype signature, Yao endotype signature, Davenport SRS, Cano-Gamez SRS, Wong score, MARS endotype signature and the SoM signature. Continuous scores were calculated based on prior publications and scaled for analysis[12,14,16–20].

Clustering

We first performed unsupervised hierarchical clustering analysis by applying the Ward method to Euclidean distances between scaled scores. The optimal number of clusters across infectious etiologies and severities were assessed by silhouette width. Significance was assessed by generating bootstrap *P* values with 1,000 repetitions. We then performed network analysis to identify interrelatedness of scores. Edges were defined based on a Spearman's correlation greater than the median or 0.33, whichever was greater. Score clusters were generated by a cluster-greedy forward algorithm.

Single-cell data analysis

To evaluate the immune cell origin of molecular endotypes, four peripheral blood scRNA-seq datasets inclusive of the neutrophil compartment were integrated. Integration was performed using the Seurat and Scanpy pathways. Cell assignments were made based on canonical cell markers cross-referenced with Seurat cluster assignments. Scaled scores were calculated for each individual cell, results were assessed by UMAP and conglomerate results of scaled scores by cell type were plotted to assess trends across sepsis signatures.

Development of the immune dysregulation framework

After identifying the cell type of origin, we then set out to develop a more granular score to interrogate specific parts of the immune response. We first separated single-cell expression data into four cell types of interest: immature neutrophils, neutrophils, monocytes and T or NK cells. We then evaluated scaled gene expression by cell line for all genes used across the seven signatures. To ensure cell specificity, a gene was included as part of the myeloid or lymphoid dysregulation score only if its scaled gene expression was >1 s.d. higher than other cell lines. Genes were then divided into detrimental and protective, based on whether the signature from which these genes were derived was previously defined as a detrimental or a protective cluster.

After identifying myeloid and lymphoid protective and detrimental genes, myeloid and lymphoid dysregulation scores were calculated as the geometric mean of detrimental genes minus the geometric mean of protective genes. Cell specificity was assessed using scaled scores overlaid on UMAPs.

Evaluation of clinical outcomes

To evaluate the association of myeloid and lymphoid scores with clinical outcomes, we first evaluated the performance of the continuous myeloid and lymphoid dysregulation scores. We evaluated the association of these scores across all severity levels using the JT *t*-test. We evaluated the association of these scores with severe infections and mortality using logistic regression.

We then set out to evaluate whether clinically meaningful cut-offs for myeloid and lymphoid dysregulation could be developed. To develop theoretical cut-offs, we evaluated scores relative to healthy participants. Within healthy participants, myeloid and lymphoid scores were generated as above and the population mean and s.d. were calculated. We then used this mean and s.d. to calculate a *z*-score for non-healthy individuals. Dysregulation was defined as a *z*-score ≥ 1.65 across the SUBSPACE consortium, indicative of a score in the 95th percentile of healthy patients. This then allowed for subgrouping of patients into four theoretical subgroups: balanced (myeloid and lymphoid *z*-score < 1.65), myeloid dysregulation (myeloid *z*-score ≥ 1.65, lymphoid *z*-score < 1.65), lymphoid dysregulation (lymphoid *z*-score ≥ 1.65, myeloid *z*-score < 1.65) and system-wide (myeloid and lymphoid *z*-scores ≥ 1.65). For more specific, cohort-level questions, *z*-score thresholds were defined across the subspace dataset based on median dysregulation scores for the same patient population. For instance, when evaluating ICU cohorts, abnormal dysregulation was defined as greater than or equal to the median dysregulation across all ICU-level patients in the SUBSPACE consortium dataset. When healthy participant gene expression was not available to allow for co-normalization, dysregulation was defined based on median myeloid and lymphoid scores within the cohort. Using these cut-offs, we evaluated the association of each subgroup with severity and mortality using Fisher's exact test.

Demonstration of the need for framework flexibility

To evaluate why prior subgrouping schemas identified different numbers of endotypes, we evaluated how myeloid and lymphoid dysregulation differed across patient cohorts. We evaluated the differences in mean and s.d. of myeloid and lymphoid dysregulation by severity and infectious etiology. We then evaluated the sensitivity and specificity of different myeloid and lymphoid dysregulation scores for discriminating mild, severe and fatal diseases.

Evaluation of treatment responsiveness

We then tested whether myeloid and lymphoid dysregulation was associated with a differential treatment response to immune modulation. We first evaluated treatment response to anakinra in the SAVE-MORE trial, which was included in the SUBSPACE consortium. The SAVE-MORE trial was a randomized controlled trial of anakinra in hospitalized patients with COVID-19 who had elevated soluble urokinase plasminogen-activating receptor levels. We evaluated the differential mortality in patients with myeloid or lymphoid dysregulation, defined based on dysregulation scores greater than or equal to those present in infected, noncritically ill patients, using Fisher's exact test. Interaction terms were generated using logistic regression, adjusting for age, sex and the SOFA score. Cox's proportional HRs were calculated, adjusting for age sex and the SOFA score.

To evaluate the association of steroid treatment with differential outcomes, we turned to two randomized controlled trials: VICTAS, a randomized controlled trial of hydrocortisone, thiamine and vitamin C in critically ill sepsis patients[45]; and VANISH[46], a randomized controlled factorial trial comparing norepinephrine versus vasopressin and hydrocortisone versus placebo. In VICTAS, which had healthy participants for co-normalization, dysregulation was defined as greater than or equal to the median dysregulation score across all infected, critically ill patients. In VANISH, which did not have healthy participants to allow for co-normalization, dysregulation was defined by the median score within the cohort. We evaluated differential outcomes among myeloid and lymphoid dysregulated patients using Fisher's exact test. For both cohorts, logistic regression was performed, adjusting for sex and APACHE II score. In the VICTAS cohort, where survival data were available, Cox's proportional HRs were calculated, adjusting for sex and APACHE II score.

Glossary or key definitions

Cluster: a broad definition denoting a group of similar patients or samples. In general, cluster is used in this paper to broadly denote samples that are identified as similar based on their biomarker profiles before evaluating biological relevance.

Endotype: a subgroup within a syndrome, based on distinct pathways, profiles or signatures that drive the disease.

Signature: a predefined analytical tool to identify patient endotypes. In general, signature is used in this paper to denote a priori defined methods using gene expression data to identify sepsis endotypes.

Modules: a set of independent units that are combined to calculate a single more complex signature or score. In this paper, this is used in the context of the SoM endotyping signature, which is divided into four modules: two detrimental modules (1 and 2) and two protective modules (3 and 4) that together define a patient's SoM score.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The expression levels of the genes used to calculate all signatures and phenotyping information to recreate figures are available for every sample from the SUBSPACE and public datasets via GitHub at https://github.com/Khatri-Lab/SUBSPACE and will remain open access. Access to comprehensive gene expression data for public cohorts can be made through the provided accessions in Supplementary Table 1. The SUBSPACE consortium bylaws forbid sharing consortium-wide complete gene expression data. For each cohort within the SUBSPACE consortium, access to complete gene expression data may be granted on a case-by-case basis at the discretion of individual site principal investigators. Requests can be directed to the corresponding author (pkhatri@stanford.edu), who will direct them as appropriate to the individual site principal investigator within 1 week.

Code availability

All analyses were performed with R studio v.4.1.1. Packages used in analyses included: COCONUT (v.1.0.2, archived), tidyverse (1.3.1), data.table (1.17.8), pals (1.7), rawr (1.0.1), pROC (1.18.0), ggpubr (0.6.1), ggbeeswarm (0.7.2), ggrastr (1.0.2), interactions (1.1.5), ggcorrplot (0.1.4), ComplexHeatmap (2.10.0), cluster (2.1.2), factoextra (1.0.7), dendextend (1.17.1), stats (4.1.1), scales (1.3.0), pvclust (2.2.0), clustMixType (0.3.14), FactoMineR (2.9), igraph (1.2.7), gtsummary (1.7.2), ggalluvial (0.12.5), survival (3.5.5), ggsurvfit (0.3.0), ggdendro (0.2.0), grid (4.1.1), forestplot (3.1.3), lme4 (1.1.33), ggExtra (0.10.1), ggplotify (0.2), ggforce (0.4.2), SepstratifieR (1.0), cowplot (1.1.1), grid (4.1.1) and gridExtra (2.3). All code needed for reproduction of analyses is available via GitHub at https://github.com/Khatri-Lab/SUBSPACE.

Acknowledgements

We thank H. Wong for his outstanding achievements and lasting legacy in the field of transcriptomics and endotyping in sepsis. K. Harmon maintained the Sepsis Genomics Collaborative Biorepository at Cincinnati Children's Hospital Medical Center. A.R.M. received funding through the Stanford Training Program in Lung Biology: National Heart, Lung, and Blood Institute (grant no. 5T32HL129970-07). M.R.A. received funding from the National Institutes of Health (NIH) through grants (nos R35GM155165, R21GM150093 and R21GM151703). The Stanford Biorepository is supported by the NIH (grant no. R01HL152083 to A.J.R.). The MESSI cohort was funded by grant no. HL161196 (to N.J.M.). P.K. was funded in part by the National Institute of Allergy and Infectious Diseases (grant nos U19AI167903 and 2U19AI057229-21) and the Brennan family. The VICTAS trial was funded by the Marcus Foundation. The SAVE-MORE trial was sponsored by the Hellenic Institute for the Study of Sepsis and funded by Swedish Orphan BioVitrum AB. The SPIES study was funded by the NIH (grant no. 5R01GM139046-02 to L.L.M.).

Author contributions

P.K., A.R.M. and T.E.S. conceived of the study. A.R.M. and P.K. designed the analyses. P.K. supervised the study. Y.H.-B. processed the data from the SUBSPACE cohort. A.R.M. collected, annotated, processed and analyzed the data. H.Z. and A. Ganesan collected, integrated and analyzed the scRNA-seq data. A.R.M., A.J.R. and P.K. interpreted the results and wrote the paper. M.V.M., J.E.L., T.v.d.P., J.L., H.R.B., B.P.S., E.J.G.-B., A.K., I.M.-L., A. Ganesan, J.H., R.E.R., J.S., D.W.W., M.R.A., L.L.M., P.A.E., M.K., T.K., H.M.G. and N.J.M. contributed data and aided in editing the paper.

Competing interests

Y.H.-B. is employed by Inflammatix, Inc. P.K. is co-founder of and consultant to Inflammatix, Inc. T.E.S. is co-founder and CEO of Inflammatix, Inc. Gene expression sequencing was funded by Inflammatix, Inc.; however Inflammatix employees were not involved in the analytical plan, performance or supervision of analyses. N.J.M. reports consulting fees from Novartis, Inc., AstraZeneca, Inc. and Endpoint Health, Inc. (>2 years ago). E.J.G. has received honoraria from Abbott Products Operations AG, bioMérieux, Brahms GmbH, GSK, InflaRx GmbH, Sobi and Xbiotech Inc.; independent educational grants from Abbott Products Operations, bioMérieux Inc, InflaRx GmbH, Johnson & Johnson, MSD, UCB and Swedish Orphan Biovitrum AB; and funding from the Horizon 2020 European Grants ImmunoSep and RISCinCOVID and the Horizon Health grant EPIC-CROWN-2, Homi-LUNG and POINT (granted to the Hellenic Institute for the Study of Sepsis or to the National and Kapodistrian University of Athens). H.R.B. has received unrestricted institutional funding from Becton Dickinson, Inflammatix, Octapharma, OSAsense and MeMed. J.H. is employed by Danaher Diagnostics. Danaher Diagnostics played no role in the study. The other authors declare no competing interests.

Additional information

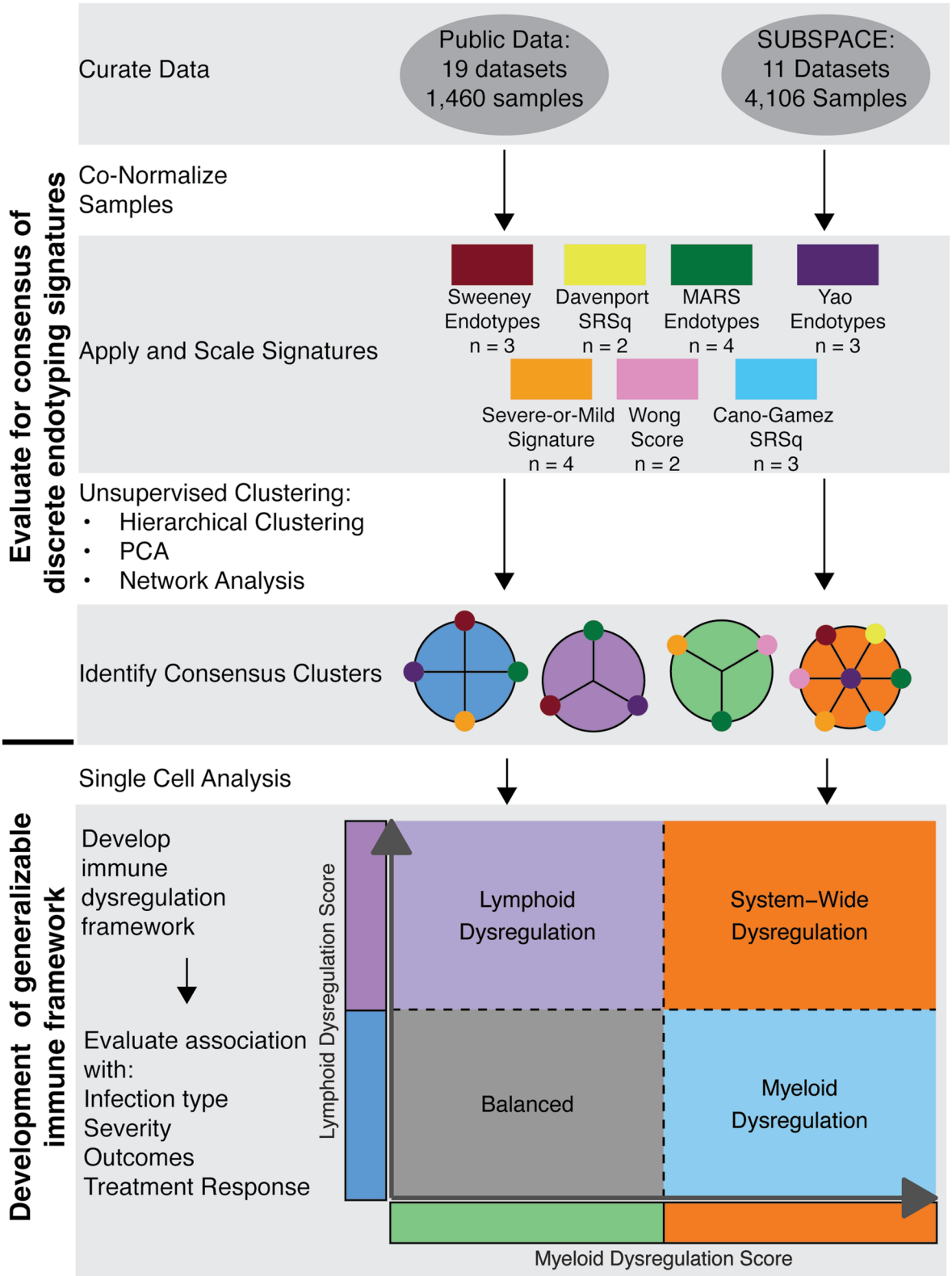
Extended data is available for this paper at https://doi.org/10.1038/s41591-025-03956-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-025-03956-5.

Correspondence and requests for materials should be addressed to Purvesh Khatri.

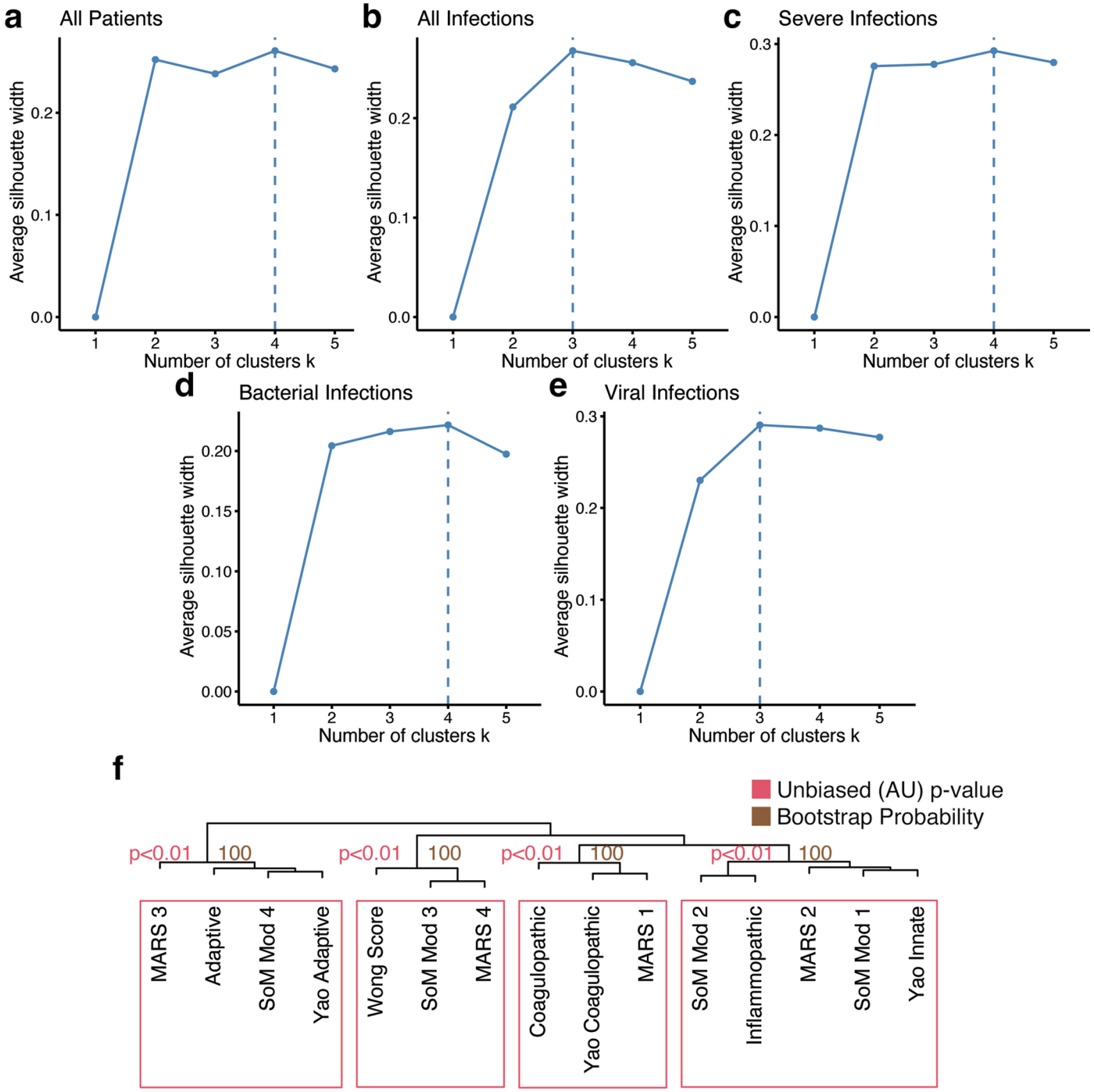
Peer review information *Nature Medicine* thanks Robert Stevens and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Saheli Sadanand, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-025-03956-5



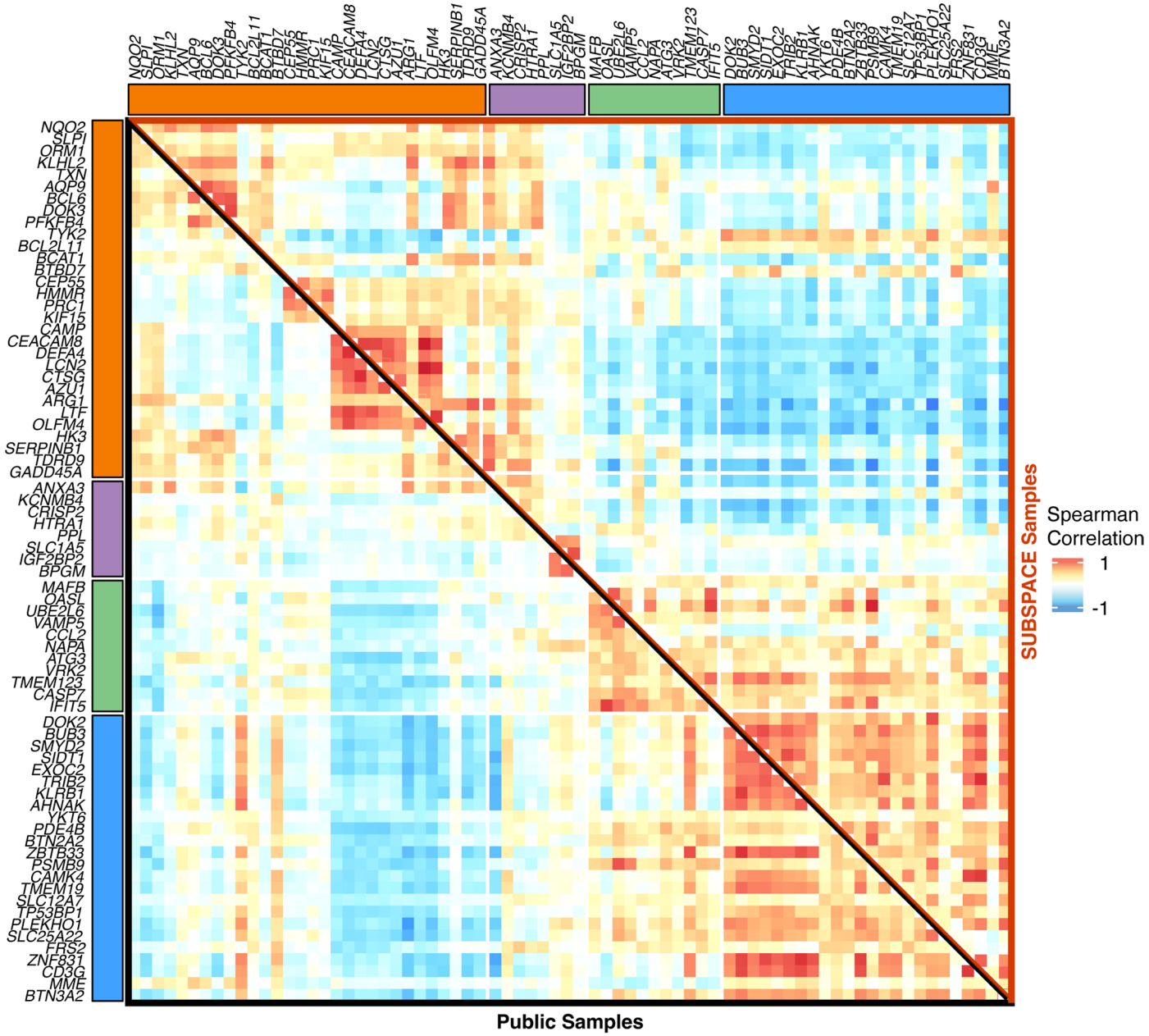
Extended Data Fig. 1 | Analysis overview. *Consensus clusters and how they are utilized to define the framework are denoted by the color scheme (blue, green, purple, and red). This color scheme is found throughout applicable clustering figures.

https://doi.org/10.1038/s41591-025-03956-5



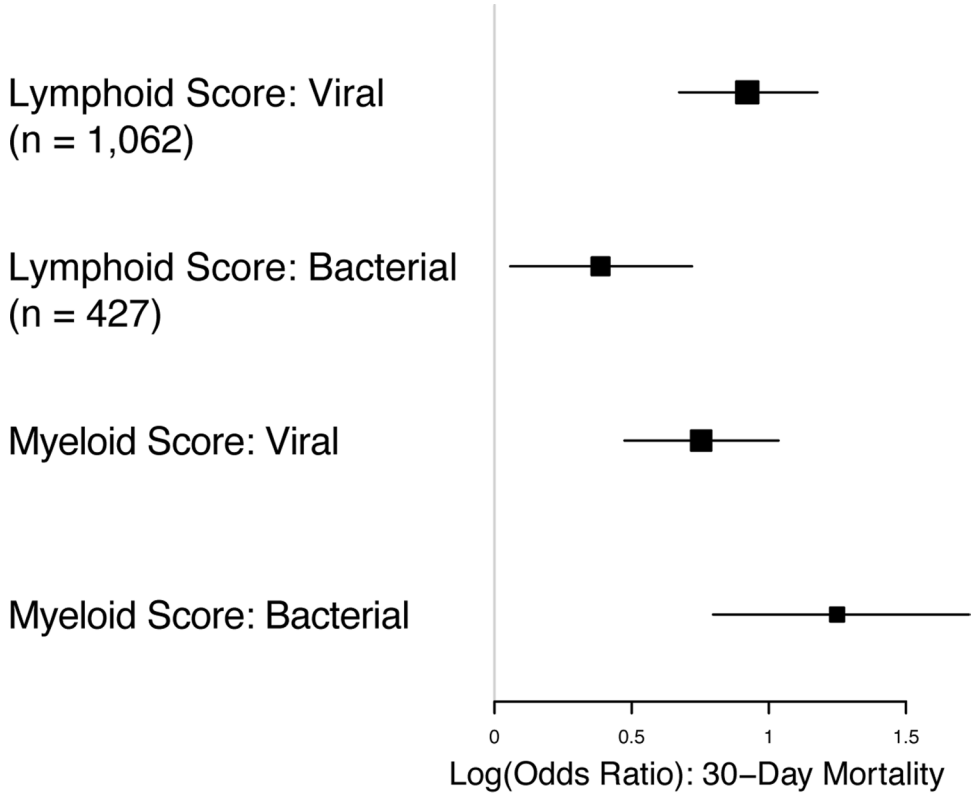
Extended Data Fig. 2 | Ideal number of clusters identified by unsupervised hierarchical clustering varies by inclusion criteria. Silhouette width index identifies differing "ideal" number of clusters depending on patient inclusion criteria. (**a**) Silhouette width index if all patient's are included (healthy and infected). (**b**) Silhouette width index if only patients with infections are included. (**c**) Silhouette width index if only severe infections (those requiring intensive care unit) are included. (**d**) Silhouette width index if only bacterial infections are included. (**e**) Silhouette width index if only viral infections are included. (**f**) Two-sided bootstrap probabilities generated with 1000 repetitions show significance of these four clusters.

https://doi.org/10.1038/s41591-025-03956-5



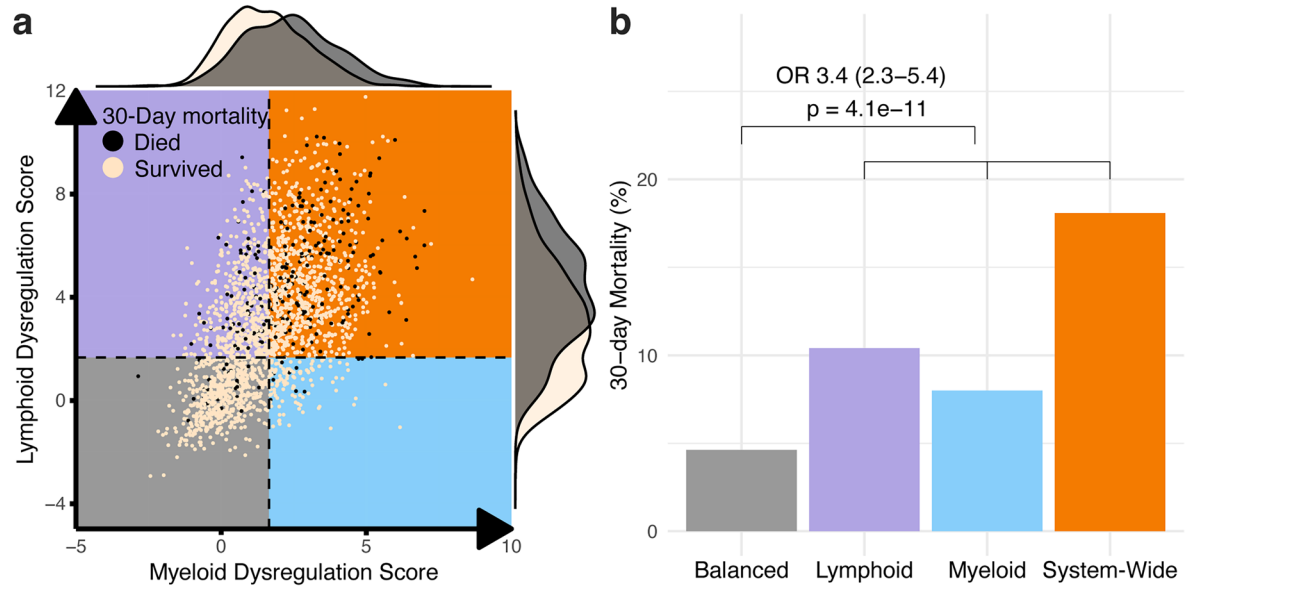
Extended Data Fig. 3 | Correlation of genes across sub-phenotyping schemas. Spearman correlation coefficients of up-regulated genes within each of the 7 endotypes signatures in public (bottom left) and SUBSPACE (top right) data indicates that while different genes were used across the prior endotyping signatures, that these genes are highly correlated.

https://doi.org/10.1038/s41591-025-03956-5



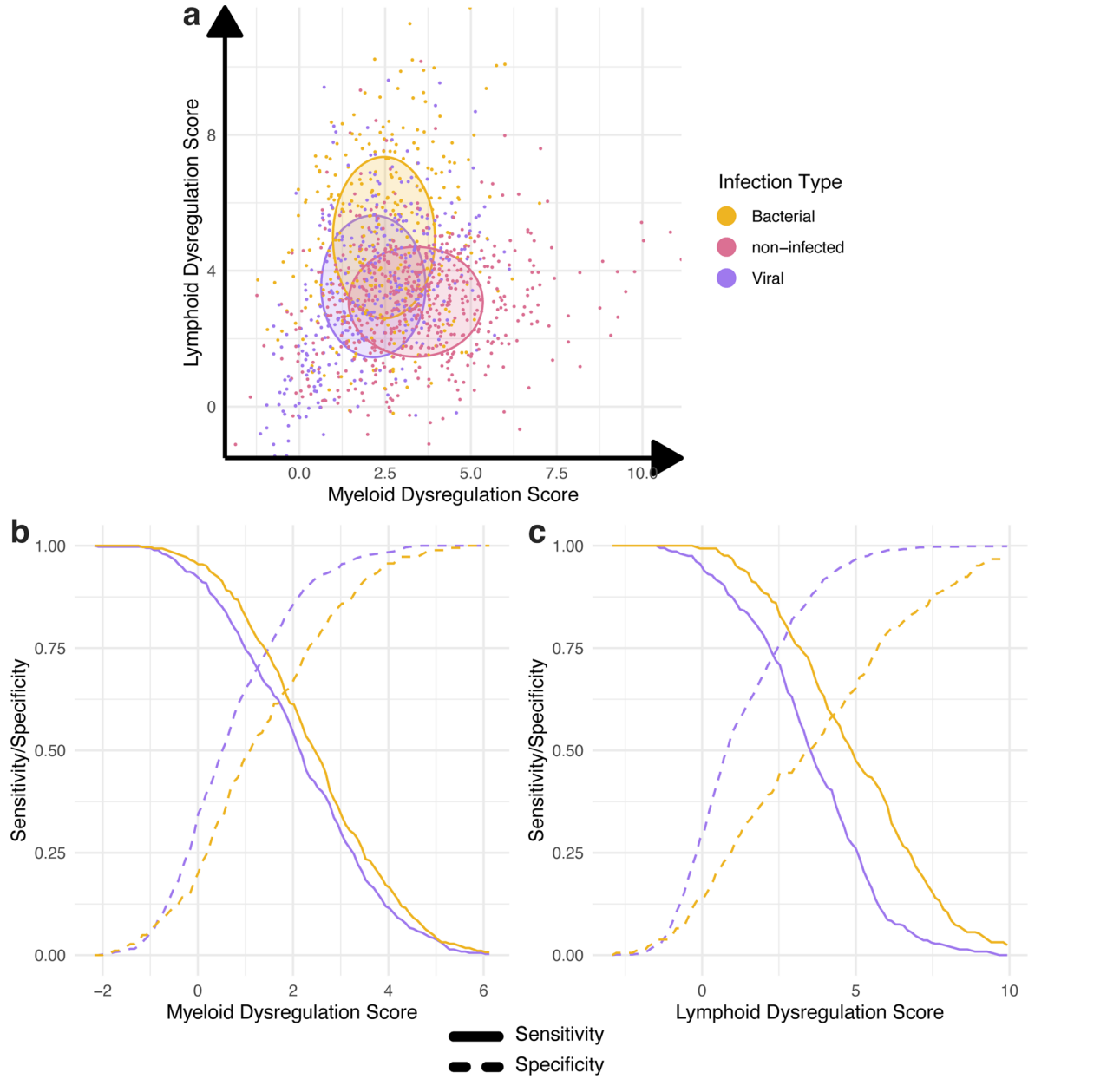
Extended Data Fig. 4 | Association of lymphoid and myeloid dysregulation scores in viral versus bacterial infections. Forest plots showing log(odds ratio) and 95% confidence intervals of 30-day mortality in viral and bacterial infections in the SUBSPACE dataset by myeloid and lymphoid dysregulation scores shows that lymphoid dysregulation is more associated with outcomes in viral infections whereas myeloid dysregulation score is more associated with outcomes in bacterial infections.

https://doi.org/10.1038/s41591-025-03956-5



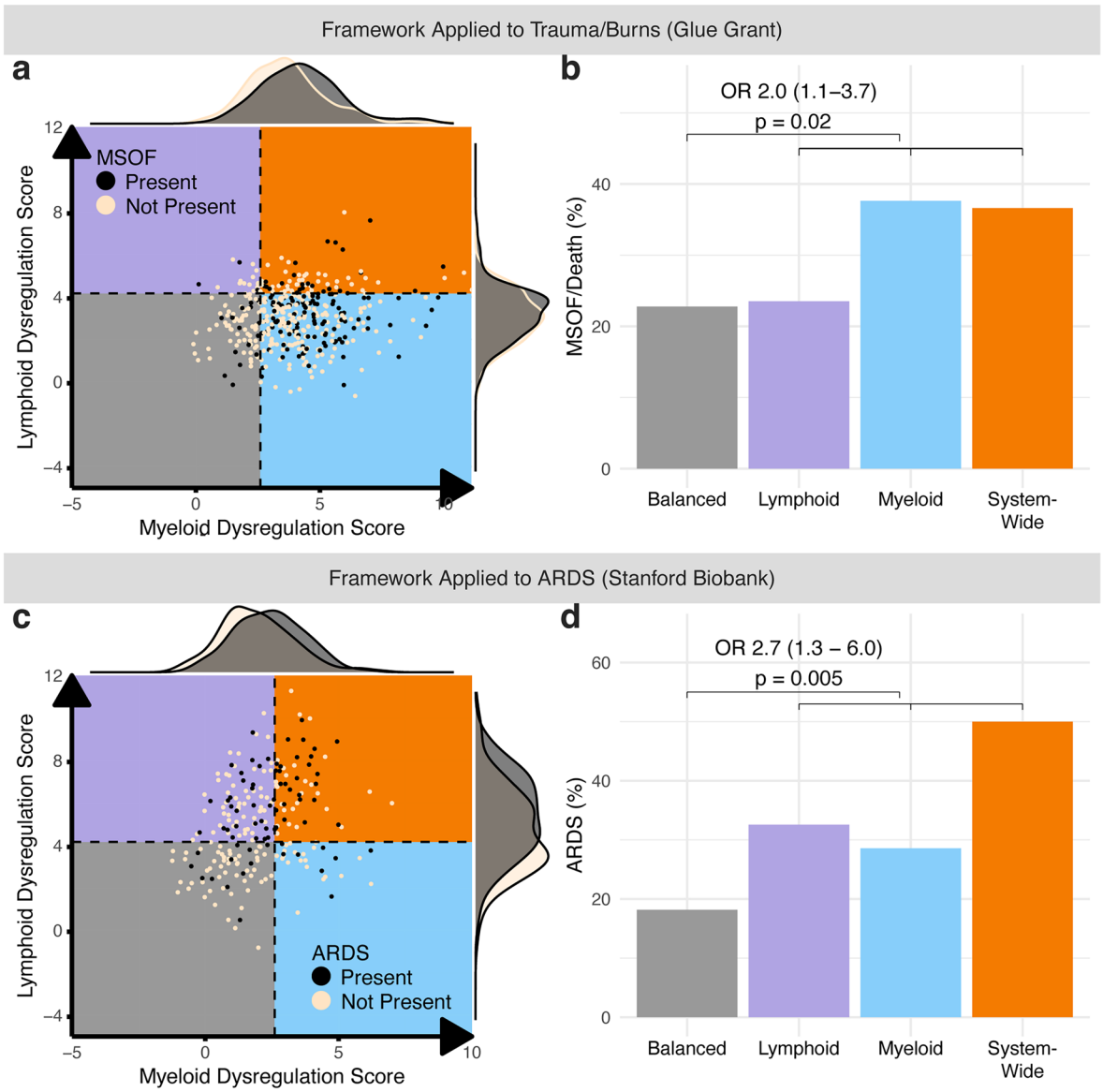
Extended Data Fig. 5 | Immune dysregulation framework applied to 30-day mortality in SUBSPACE datasets. (**a**). The immune dysregulation framework applied to SUBSPACE co-normalized data. Cut-offs are defined by a Z-score of 1.65 relative to healthy patients. Black dots represent patients with severe infectious (defined by ICU admission) while tan dots represent non-severe infections. (**b**). Barplot representing proportion of 30-day mortality (y-axis) by immune dysregulation framework subgroup (x-axis). Odds ratio calculated using two-sided Fisher's exact test unadjusted for multiple comparisons comparing patients with dysregulation on any axis relative to "Balanced" subgroup. Dysregulation on either myeloid or lymphoid axis (inclusive of lymphoid dysregulation (n = 586), myeloid dysregulation (n = 125), and system-wide dysregulation (n = 896) subgroups) was associated with 30-day mortality with an odds ratio of 3.4 (95% CI 2.3 – 5.4, p = 4.1e-11) compared to patients in the balanced subgroup (n = 606).

https://doi.org/10.1038/s41591-025-03956-5



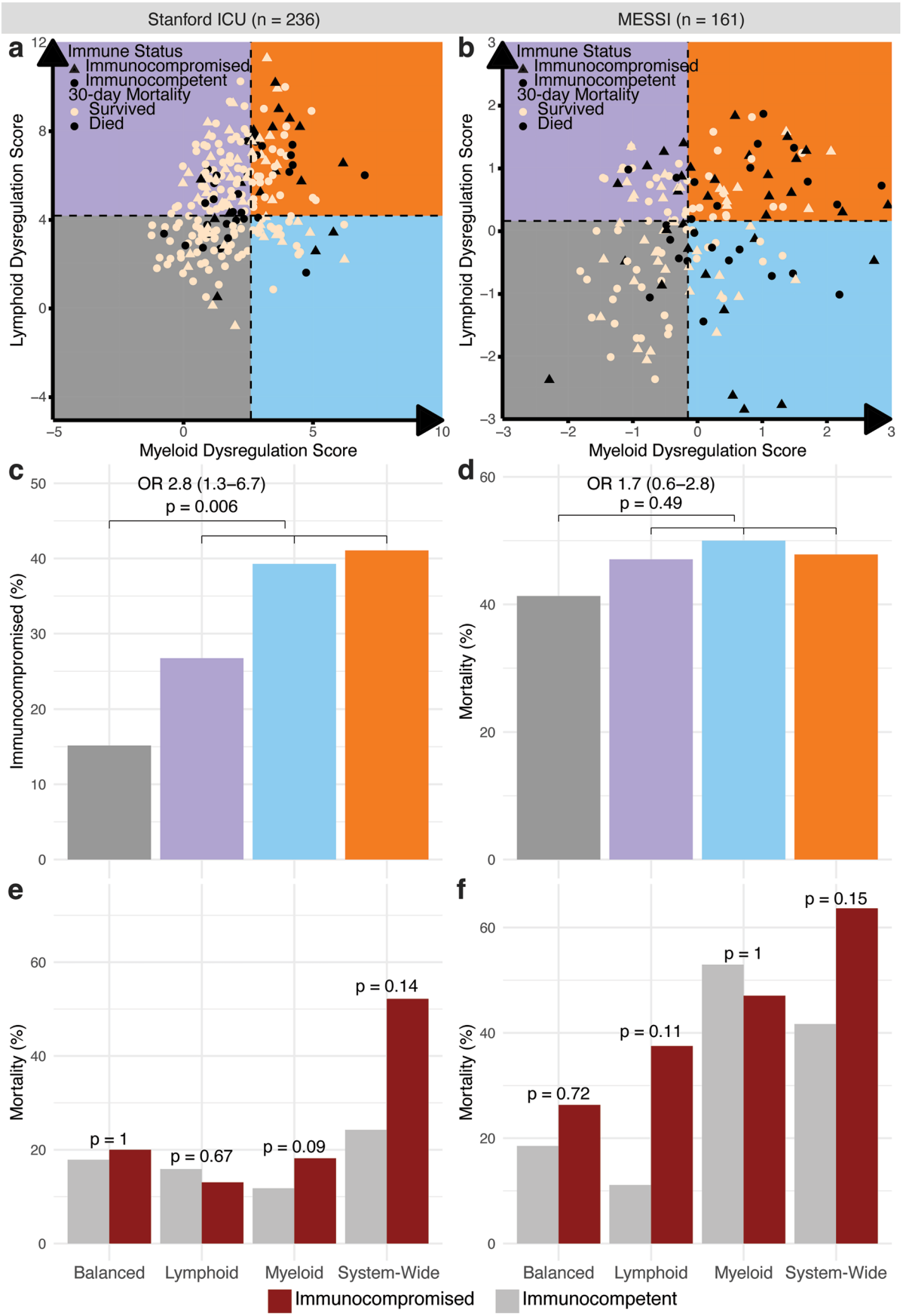
Extended Data Fig. 6 | Lymphoid and myeloid dysregulation vary by infection status. (**a**). Mean and standard deviation of myeloid and lymphoid dysregulation scores by infection status in patients requiring intensive care unit level of care indicating that values differ by infectious etiology even amongst similar severity levels. (**b**). Sensitivity (solid line) and specificity (dashed line) of myeloid dysregulation score for discriminating severe from non-severe infections in viral (purple) vs bacterial (gold) infections shows that "ideal" threshold may vary across infectious individuals. (**c**). Sensitivity (solid line) and specificity (dashed line) of lymphoid dysregulation score for discriminating severe from non-severe infections in viral (purple) vs bacterial (gold) infections shows that "ideal" threshold may vary across infectious individuals.

https://doi.org/10.1038/s41591-025-03956-5



Extended Data Fig. 7 | Application of the immune dysregulation framework to other critical illness syndromes. (**a**). The immune dysregulation framework applied to non-infected trauma and burn patients at baseline from the Glue Grant (n = 430). Cut-offs are defined by median dysregulation scores across all ICU-level patients in SUBSPACE. Black dots represent patients with multi-system organ failure or death while tan dots represent survivors without multi-system organ failure. (**b**). Barplot representing proportion of multi-system organ failure or death (y-axis) by immune dysregulation framework subgroup (x-axis). Odds ratio calculated using two-sided Fisher's exact test unadjusted for multiple comparisons comparing patients with dysregulation on any axis relative to "Balanced" subgroup. Dysregulation on either myeloid or lymphoid axis (inclusive of lymphoid dysregulation (n = 17), myeloid dysregulation (n = 263), and system-wide dysregulation (n = 71) subgroups) was associated with multi-system organ failure or death with an odds ratio of 2.0 (95% CI 1.1 – 53.7, p = 0.02) compared to patients in the balanced subgroup (n = 79). (**c**). The immune dysregulation framework applied to Stanford data (n = 236). Cut-offs are defined by median dysregulation scores across all ICU-level patients in SUBSPACE. Black dots represent patients with Acute Respiratory Distress Syndrome (ARDS) while tan dots represent those without ARDS. (**d**). Barplot representing proportion of severe ARDS (y-axis) by immune dysregulation framework subgroup (x-axis). Odds ratio calculated using two-sided Fisher's exact test unadjusted for multiple comparisons comparing patients with dysregulation on any axis relative to "Balanced" subgroup. Dysregulation on either myeloid or lymphoid axis (inclusive of lymphoid dysregulation (n = 86), myeloid dysregulation (n = 28), and system-wide dysregulation (n = 56) subgroups) was associated with severe infections with an odds ratio of 2.7 (95% CI 1.3 – 6.0, p = 0.005) compared to patients in the balanced subgroup (n = 66).

https://doi.org/10.1038/s41591-025-03956-5



Extended Data Fig. 8 | See next page for caption.

https://doi.org/10.1038/s41591-025-03956-5

Extended Data Fig. 8 | Effect of immunocompromise on the immune dysregulation framework in the Stanford and MESSI cohort. (**a**). The immune dysregulation framework applied to Stanford data (n = 236) evaluating 30-day survival and immunocompromised status. Cut-offs are defined by median dysregulation scores across all ICU-level patients in SUBSPACE. Immunocompromised patients (n = 67) are represented by triangles while immunocompetent patients (n = 169) are represented by circles and survivors (n = 187) are indicated by tan shapes relative to those who died within 30 days (n = 49) with black shapes. (**b**). The immune dysregulation framework applied to the MESSI cohort data (n = 160) evaluating 30-day survival and immunocompromised status. Given lack of healthy controls for co-normalization, cut-offs are defined by median dysregulation scores in the MESSI cohort. Immunocompromised patients (n = 74) are represented by triangles while immunocompetent patients (n = 86) are represented by circles and survivors (n = 59) are indicated by tan shapes relative to those who died within 30 days (n = 101) with black shapes. (**c**). Barplot representing proportion of immunocompromised patients (y-axis) by immune dysregulation framework subgroup (x-axis) in the Stanford cohort. Odds ratio calculated using two-sided Fisher's exact test unadjusted for multiple comparisons comparing patients with dysregulation on any axis relative to "Balanced" subgroup. Dysregulation on either myeloid or lymphoid axis (inclusive of lymphoid dysregulation (n = 86), myeloid dysregulation (n = 28), and system-wide dysregulation (n = 56) subgroups) was associated with immunocompromised status with an odds ratio of 2.8 (95% CI 1.3 – 6.7, p = 0.006) compared to patients in the balanced subgroup (n = 66). (**d**). Barplot representing proportion of immunocompromised patients (y-axis) by immune dysregulation framework subgroup (x-axis) in the MESSI cohort. Odds ratio calculated using two-sided Fisher's exact test unadjusted for multiple comparisons comparing patients with dysregulation on any axis relative to "Balanced" subgroup. Dysregulation on either myeloid or lymphoid axis (inclusive of lymphoid dysregulation (n = 34), myeloid dysregulation (n = 34), and system-wide dysregulation (n = 46) subgroups) was not associated with immunocompromised status with an odds ratio of 1.7 (95% CI 0.6 – 2.8, p = 0.49) compared to patients in the balanced subgroup (n = 46). (**e**). Barplot representing proportion of patients who died within 30 days (y-axis) by immune dysregulation framework subgroup (x-axis) and immunocompromised (red) vs immunocompetent (grey) status in the Stanford cohort. Odds ratio calculated using two-sided Fisher's exact test unadjusted for multiple comparisons comparing 30-day mortality between immunocompromised and immunocompetent patients within each subgroup. (**f**). Barplot representing proportion of patients who died within 30 days (y-axis) by immune dysregulation framework subgroup (x-axis) and immunocompromised (red) vs immunocompetent (grey) status in the MESSI cohort. Odds ratio calculated using two-sided Fisher's exact test unadjusted for multiple comparisons comparing 30-day mortality between immunocompromised and immunocompetent patients within each subgroup.

https://doi.org/10.1038/s41591-025-03956-5

Extended Data Table 1 | Pathway analysis of prior gene signatures

Gene Signature	Up-Regulated genes	Down-Regulated Genes	Summary of pathway and gene ontology analysis results
SRS	*TDRD9*	*DYRK2, CCNB1IP1, ZAP70, ARL14EP, MDC1, ADGRE3*	Immature neutrophil activation, endotoxin clearance, T cell exhaustion, downregulation of MHC class II
Sweeney Inflammopathic	*ARG1, LCN2, LTF, OLFM4*	*HLA-DMB*	Immature neutrophils, myeloid-derived suppression, reduction in antigen presentation, anti-inflammatory macrophage activation
Yao Innate	*HK3, SERPINB1*	*EPB42, GSPT1, LAT*	Innate activation
MARS 2	*GADD45A*	*PCGF5*	Increased PAMP activation, cytokines, cell growth, and mobility
SoM Module 1,2	*NQO2, SLPI, ORM1, KLHL2, ANXA3, TXN, AQP9, BCL6, DOK3, PFKFB4, TYK2, BCL2L11, BCAT1, BTBD7, CEP55, HMMR, PRC1, KIF15, CAMP, CEACAM8, DEFA4, LCN2, CTSG, AZU1*		Immature neutrophil activation and myeloid-derived suppression
Sweeney Coagulopathic	*KCNMB4, CRISP2, HTRA1, PPL*	*RHBDF2, ZCCHC4, YKT6, DDX6, SENP5, RAPGEF1, DTX2, RELB*	Coagulopathy, neutrophil activation, reduced adaptive immune cell function
Yao Coagulopathic	*SLC1A5, IGF2BP2, ANXA3*	*GBP2*	Coagulopathy, neutrophil activation
MARS 1	*BPGM*	*TAP2*	Reduction in innate and adaptive immune cell functions, particularly TLRs and T-cell receptor signaling
Wong Score	100 genes upregulated in less severe infections – Listed in full in subtext*		Neutrophil biology and adaptive immunity
SoM Module 3	*MAFB, OASL, UBE2L6, VAMP5, CCL2, NAPA, ATG3, VRK2, TMEM123, CASP7*		Monoctye activation and interferon responses
MARS4	*IFIT5*	*NOP53*	Interferon responses
Sweeney adaptive	*YKT6, PDE4B, TWISTNB, BTN2A2, ZBTB33, PSMB9, CAMK4, TMEM19, SLC12A7, TP53BP1, PLEKHO1, SLC25A22, FRS2*	*GADD45A, CD24, S100A12, STX1A*	Adaptive immune pathways. Decreased neutrophil responses
Yao Adaptive	*ZNF831, CD3G, MME, BTN3A2, HLA-DPA1*	*STOM*	Adaptive immune pathways
SoM module 4	*DOK2, HLA-DPB1, BUB3, SMYD2, SIDT1, EXOC2, TRIB2, KLRB1*		Adaptive immune pathways
MARS3	*AHNAK*	*PDCD10*	Adaptive immune pathways

*Wong score comprises: APAF1, ARPC5, ASAH1, ATP2B2, BCL6, BMPR2, BTK, CAMK2D, CAMK2G, CAMK4, CASP1, CASP2, CASP4, CASP8, CD247, CD3E, CD3G, CD79A, CREB1, CREB5, CSNK1A1, CTNNB1, DAPP1, DBT, EP300, FAS, FCGR2A, FCGR2C, FYN, GK, GNAI3, HDAC4, HLA-DMA, HLA-DOA, ICAM3, IL1A, INPP5D, ITGAM, ITGAV, ITGAX, JAK1, JAK2, KAT2B, LAT2, LYN, MAP2K4, MAP3K1, MAP3K3, MAP3K5, MAP3K7, MAP4K1, MAP4K4, MAPK1, MAPK14, MDH1, MKNK1,NCOA2, NCR3, NFATC1, PAK2, PDPR, PIAS1, PIK3C2A, PIK3C3, PIK3CA, PIK3CD, PIK3R1, PLCG1, POU2F2, PPP1R12A, PPP2R2A, PPP2R5C, PRKAR1A, PRKCB, PSMB7, PTEN, PTPRC, RAF1, RHOT1, ROCK1, SEMA4F, SEMA6B, SMAD4, SOS1, SOS2, SP1, TAF11, TBK1, TGFBR1, TLE4, TLR1, TLR2, TLR8, TNFSF10, TRA@, TYROBP, UBE3A, USP48, ZAP70, ZDHHC17.

nature portfolio

Corresponding author(s): Purvesh Khatri

Last updated by author(s): Aug 7, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All analyses were performed with R studio version 4.1.1. All code needed for reproduction of analyses is available at our GitHub repository https://github.com/Khatri-Lab/SUBSPACE.
Data analysis	All analyses were performed with R studio version 4.1.1. Packages used in analyses included: COCONUT(v1.0.2- archived), tidyverse (1.3.1), data.table (1.17.8), pals (1.7), rawr (1.0.1), pROC (1.18.0), ggpubr (0.6.1), ggbeeswarm (0.7.2), ggrastr (1.0.2), interactions (1.1.5), ggcorrplot (0.1.4), ComplexHeatmap (2.10.0), cluster (2.1.2), factoextra (1.0.7), dendextend (1.17.1), stats (4.1.1), scales (1.3.0), pvclust (2.2.0), clustMixType (0.3.14), FactoMineR (2.9), igraph (1.2.7), gtsummary (1.7.2), ggalluvial (0.12.5), survival (3.5.5), ggsurvfit (0.3.0), ggdendro (0.2.0), grid (4.1.1), forestplot (3.1.3), lme4 (1.1.33), ggExtra (0.10.1), ggplotify (0.2), ggforce (0.4.2), SepstratifieR (1.0), cowplot (1.1.1), grid (4.1.1), gridExtra (2.3). All code needed for reproduction of analyses is available at our GitHub repository https://github.com/Khatri-Lab/SUBSPACE.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The expression levels of the genes used to calculate all signatures as well as phenotyping information to recreate figures are available for every sample from the SUBSPACE and public datasets at our GitHub repository https://github.com/Khatri-Lab/SUBSPACE at the time of publication and will remain open access. Access to comprehensive gene expression data for public cohorts can be accessed through the provided accessions in supplemental table 1. The SUBSPACE consortium bylaws forbid sharing consortium-wide complete gene expression data. For each cohort within the SUBSPACE consortium, access to complete gene expression data may be granted on a case-by-case basis at the discretion of individual site principal investigators. Requests can be directed to the corresponding author (pkhatri@stanford.edu) who will direct them as appropriate to the individual site principal investigator within 1 week.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Self-reported sex were reported for each dataset. All multivariate models performed included sex as a covariate to ensure generalizability.
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity were not routinely collected across all sites. Inclusion of international sites including Low and Middle Income countries was purposely performed and sensitivity analyses were performed across United States and non_united States sites to ensure generalizability of these findings. However further study in the specific effects of race and ethnicity and across more diverse patient populations is needed.
Population characteristics	This study was designed as an large-scale biologic evaluation of existing sepsis signatures. Analysis was split across three primary cohorts: public data, single-cell data, and SUBSPACE data. Public data was collected through systematic review of publicly available whole blood and peripheral blood mononuclear cell gene expression data from the Genome Expression Omnibus (GEO) and ArrayExpress. All samples across all ages with healthy controls and necessary severity metadata and gene expression data were included in analysis .SUBSPACE was a consortium formed to collate existing biobanked whole blood gene expression data from international collaborators with the intention to collect >4,000 samples. Samples were inclusive of all ages from neonate to >80 year-old individuals.
Recruitment	Recruitment was site-specific depending on pre-specified cohort inclusion criteria. General cohort and patient characteristics can be seen in the supplement.
Ethics oversight	All samples were collected in accordance with site-specific Institutional Review Board (IRB) protocols and complied with ethical principles set forth by the Helsinki Declaration of 1975. Individual approvals for each site are listed below: • ACUTELINES: The Medical Ethics Board and the Central Review Board of the University Medical Center Groningen have evaluated and approved the protocol of Acutelines (2019/589) • Amsterdam: The Medical Ethics Committee of the Amsterdam UMC, location AMC has given approval for the conduct of the ELDER-BIOME study(NL57847.018.16), the OPTIMACT study (2016/280, NL57923.018.16), and the PANAMO study (IRB 2020_067#B2020179) • Charles University: The Charles University IRB approved sampling, handling, research and storage/biobanking of genetic material of subjects in the IMHOTEP clinical study evaluated here (decision/approval nr. 251/2022 attached) • Cincinnati Children's Hospital Medical Center: The study protocol was approved by Institutional Review Boards (IRBs) of the primary site (Cincinnati Children's Hospital IR, Genomic Analysis of Pediatric Systemic Inflammatory Syndrome, IRB 2008-0558) • SAVE-MORE: The protocol was approved by the National Ethics Committee of Greece (approval 161/20) and by the Ethics Committee of the National Institute for Infectious Diseases Lazzaro Spallanzani, IRCCS, in Rome (1 February 2021) (EudraCT no. 2020-005828-11; ClinicalTrials.gov NCT04680949). • Stanford University: The Stanford Biorepository study protocol was approved by the Stanford University IRB (IRB 28205) • Trinity University: The study protocol was approved by the Tallaght University Hospital IRB (Study title: Sepsis Immunosuppression in Critically ill Patients, Project ID: sjh428) • University of Florida: The University of Florida IRB approved the SPIES clinical study evaluated here (IRB 202000924) • University of Pennsylvania: The MESSI cohort study protocol was approved by the University of Pennsylvania IRB (IRB 808542) • VICTAS: the VICTAS clinical trial and subsequent sample storage and handling was approved by the Johns Hopkins University IRB (IRB 00102528, IRB 00164053) Consent was obtained from individuals or their legally authorized representatives per each study's protocols. Subjects were not compensated for involvement in this study. All ages groups, races, ethnicities, and sex/genders that met inclusion criteria with available gene expression data were included.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	This study was designed as an large-scale biologic evaluation of existing sepsis signatures. Analysis was split across three primary cohorts: public data, single-cell data, and SUBSPACE data (Extended Data Table 1). Public data was collected through systematic review of publicly available whole blood and peripheral blood mononuclear cell gene expression data from the Genome Expression Omnibus (GEO) and ArrayExpress. All samples across all ages with healthy controls and necessary severity metadata and gene expression data were included in analysis with an expectation of collecting >1,000 samples to be able to adequately assess biologic overlap of signatures. SUBSPACE was a consortium formed to collate existing biobanked whole blood gene expression data from international collaborators with the intention to collect >4,000 samples. Samples were inclusive of all ages from neonate to >80 year-old individuals. Single-cell samples were collected from publicly available peripheral blood single cell data inclusive of the neutrophil compartment in patients with infections.
Data exclusions	All data meeting inclusion criteria were included in broad biologic analyses. In individual outcome analyses, samples may have been excluded if necessary phenotyping outcome for that analysis was not available
Replication	All signatures, including the framework were calculated blinded to patient phenotypes and clinical outcomes. Unsupervised clustering analysis methodologies (hierarchical clustering, principal component analysis, and network analysis) were pre-planned and were performed in parallel across public and SUBSPACE data to ensure replicability with iterations across sub-groups to ensure robustness of results.
Randomization	Subjects were not randomized in this study
Blinding	All signatures, including the framework were calculated blinded to patient phenotypes and clinical outcomes. Unsupervised clustering analysis methodologies (hierarchical clustering, principal component analysis, and network analysis) were pre-planned and were performed in parallel across public and SUBSPACE data to ensure replicability with iterations across sub-groups to ensure robustness of results. For single cell data Sepsis signatures were calculated blinded to patient phenotype and were collated according to the pre-defined clusters to evaluate the single-cell biology leading to sepsis dysregulation. Criteria to select genes for Hi-DEF were pre-defined and all genes meeting inclusion criteria were included for all subsequent analyses. Subgroup cut-offs within Hi-DEF were pre-defined and subgroups were identified blinded to clinical outcomes. The analytic plan of outcomes and differential response to treatment were pre-defined. All samples that included necessary phenotypic information were included in primary analyses and were performed in parallel when possible across public and SUBSPACE data to ensure reproducibility. Sensitivity analyses by site and clinical characteristics were performed to ensure robustness.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	N/A
Study protocol	N/A

Data collection	Data was collected from over 30 cohorts across 15 countries. Details for individual studies can be found through the accessions and individual prospective cohort studies identified in supplemental tables 1 and 2
Outcomes	The primary outcome in this study was identification of signature overlap in transcriptomic data, which was analyzed using unsupervised clustering and network analysis. After generation of the consensus immune dysregulation framework, we performed numerous secondary analyses to identify association with outcomes and severity. Outcomes of interest were association of these scores with severity of disease (ICU admission, multi-system organ failure) or mortality depending on dataset evaluated and phenotyping data available.

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A