Supplementary Methods.

1.Metabolomics profiling

Untargeted metabolomics analysis was based on the Discovery HD4 metabolomics platform from Metabolon (Durham, North Carolina, USA). The detailed method was described previously.[1] In brief, 100 μL of serum was extracted, centrifuged, and then divided into four aliquots, each of which was analyzed using a distinct ultra-high performance liquid chromatography/tandem mass spectrometry (UHPLC/MS/MS) method. Metabolite identification, quantitation, curation, normalization, and quality control were performed using the platform's software. The major steps are as follows:

1) Sample preparation: Samples were extracted with methanol to precipitate proteins. The supernatant was divided into four equal aliquots by four different UPLC-MS/MS methods: two for analysis by two separate reverse phase (RP) UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, and one for analysis by hydrophilic interaction liquid chromatography (HILIC)/UPLC-MS/MS with negative ion mode ESI.

2) UPLC-MS/MS: All methods employed a Waters ACQUITY ultra-performance liquid chromatography (UPLC) system coupled to a Thermo Scientific Q-Exactive high-resolution/accurate mass spectrometer equipped with a heated electrospray ionization (HESI-II) source. One aliquot was analyzed under acidic positive ion conditions on a C18 column. A second aliquot was analyzed under the same acidic positive ion conditions on the same C18 column, using mobile phases optimized for more hydrophobic compounds. A third aliquot was analyzed under basic negative ion-optimized conditions on the same type of C18 column. The fourth aliquot was analyzed in negative ionization mode on a HILIC column. Across all methods, the scan cycle consisted of alternating full MS scans and data-dependent MS^n scans, incorporating dynamic exclusion.

3) Quality control criteria: A pooled plasma quality control (QC) sample was prepared by combining a small fraction from each experimental sample. QC replicates were inserted among every 6 experimental samples to monitor assay performance. The stability of the assay was assessed by the variability among QC samples, and metabolites with a coefficient of variation > 0.3 in QC samples were excluded. In addition, metabolites with missing values > 80% were excluded. The remaining missing values were imputed using half of the minimal detected values for the corresponding metabolite. For batch effect correction, we ensured that each analytical batch (run day) consisted of 80-100 experimental samples, and samples from each experimental group were evenly distributed across batches. Within each batch, the experimental samples were analyzed in a randomized order. The median level of each metabolite in the QC sample was used to normalize and merge the data from each batch.

4) Metabolite Identification and Quantification: We identified metabolites by comparison with a library of purified standards that records the retention time/index (RI), mass to charge ratio (m/z), and chromatographic (including MS/MS spectral data)

data for all molecules covered by each method. Peaks were quantified by integrating the area under the curve (AUC) of the primary MS ion for each metabolite.

2. Win ratio analysis

The unmatched win ratio method was used to analyze the hierarchical composite endpoint of infection-related prognostic events comprising five outcomes. They were: all-cause death, organ failure, sepsis, new onset acute decompensation and systemic inflammatory response syndrome (SIRS).[2,3] These endpoints were prioritized for evaluation within 90 days in the following hierarchical order of clinical importance:

1) Time to all-cause death
2) Organ failure or not (>=2 in ACLF cohort, >=1 in non-ACLF cohort)
3) Time to sepsis
4) New onset acute decompensation or not (within 1 month)
5) Time to SIRS

Using the ACLF cohort as an example, and following Pocock et al's methodology,[4,5] each patient in the high infection-related risk group is compared with every patient in the low infection-related risk group in an unmatched win- ratio analysis, resulting in a total of 73 × 63 pairwise comparisons. Each pair is compared sequentially, first on the basis of the most important event (e.g. death). Any ties on this event are resolved by comparing the next most important event. The unmatched win ratio calculation proceeded as follows:

Workflow 1: All patient pairs (4599 pairs) are compared for the time to death, with follow-up truncated at 90 days.

1) If both participants die, the "winner" will be the one who had a shorter time until death (indicating a worse outcome).
2) If one participant dies but another survived, the "winner" will be the one with a shorter censored time.
3) Otherwise, the match is tied and proceeds to workflow 2.

Workflow 2: The tied pairs from the workflow 1 will be compared based on the occurrence of two or more organ failures, truncated at 90 days.

1) If one participant experiences two or more organ failures but another does not, the "winner" will be the one with two or more organ failures.
2) Otherwise, the match is tied and proceeds to workflow 3.

Workflow 3: The tied pairs from the workflow 2 will be compared for the time to sepsis, truncated at 90 days.

1) If both participants develop sepsis, the "winner" will be the one who had a shorter time until sepsis.
2) If one participant develops sepsis but another does not, the "winner" will be the one with a shorter censored time.
3) Otherwise, the match is tied and proceeds to workflow 4.

Workflow 4: The tied pairs from the workflow 3 will be compared based on the occurrence of new onset acute decompensation, truncated at 60 days.

1) If one participant experiences new-onset acute decompensation but another does not, the “winner” will be the one with a new onset acute decompensation.

2) Otherwise, the match is tied and proceeds to workflow 5.

Workflow 5: The tied pairs from the workflow 4 will be compared for the time to SIRS, truncated at 90 days.

1) If both participants develop SIRS, the “winner” will be the one who had a shorter time until SIRS.

2) If one participant has a SIRS but another does not, the “winner” will be the one with a shorter censored time.

3) Otherwise, the match is tied.

The win ratio and win odds were calculated using the WINS package in R statistical software version 4.4.3 (http://www.r-project.org). The win ratio was calculated as: (the total number of wins in the high infection-related risk group) / (the total number of wins in low infection-related risk group). The win odds was calculated as: (the total number of wins in the high infection-related risk group + half number of all ties) / (the total number of wins in low infection-related risk group + half number of all ties). The *P*-value in the win ratio analysis was from the method based on U-statistics described by Dong et al.[6]

3. Boruta algorithm

The Boruta algorithm selects all relevant features through an iterative comparison with randomness. Its core idea is to identify features genuinely important to the outcome by comparing them to randomized "shadow" features. It operates as follows:

1) Shadow feature creation: The algorithm begins by extending the original data set. For each original feature, a corresponding "shadow feature" is created by randomly shuffling its values across observations. Shadow features are combined with the original features to form a new feature matrix.

2) Importance assessment: A Random Forest classifier is run on the extended dataset (containing both original and shadow features). The importance of each feature is quantified using a Z-score. The Z-score is calculated, for a single feature, as the mean loss of classification accuracy caused by the random permutation divided by the standard deviation of that loss across the Random Forest.

3) Threshold determination: The maximum Z-score among all shadow features (MZSA) is established as the benchmark threshold for importance in that iteration. A two-sided test of equality is then performed for each original feature of undetermined importance against this MZSA.

4) Feature classification: based on the test:

Features with importance significantly higher than the MZSA are deemed “Confirmed";

Features with importance significantly lower than the MZSA are deemed "Unimportant" and are permanently removed from consideration;

Features whose importance is not significantly different from the MZSA remain "Tentative".

5) Iteration: All shadow features are removed. Steps 1-4 are repeated with the remaining feature set (Confirmed and Tentative features), until each feature is assigned an importance or a preset maximum number of runs is reached.

Ultimately, the Boruta algorithm returns a subset of features confirmed as important, after rigorous iterative comparison against random noise through hypothesis testing.