Supplementary Material

1. Study samples

Figure S1 depicts the flowcharts for the study samples from the VIP2 [S1] and COVIP [S2] studies cohorts.

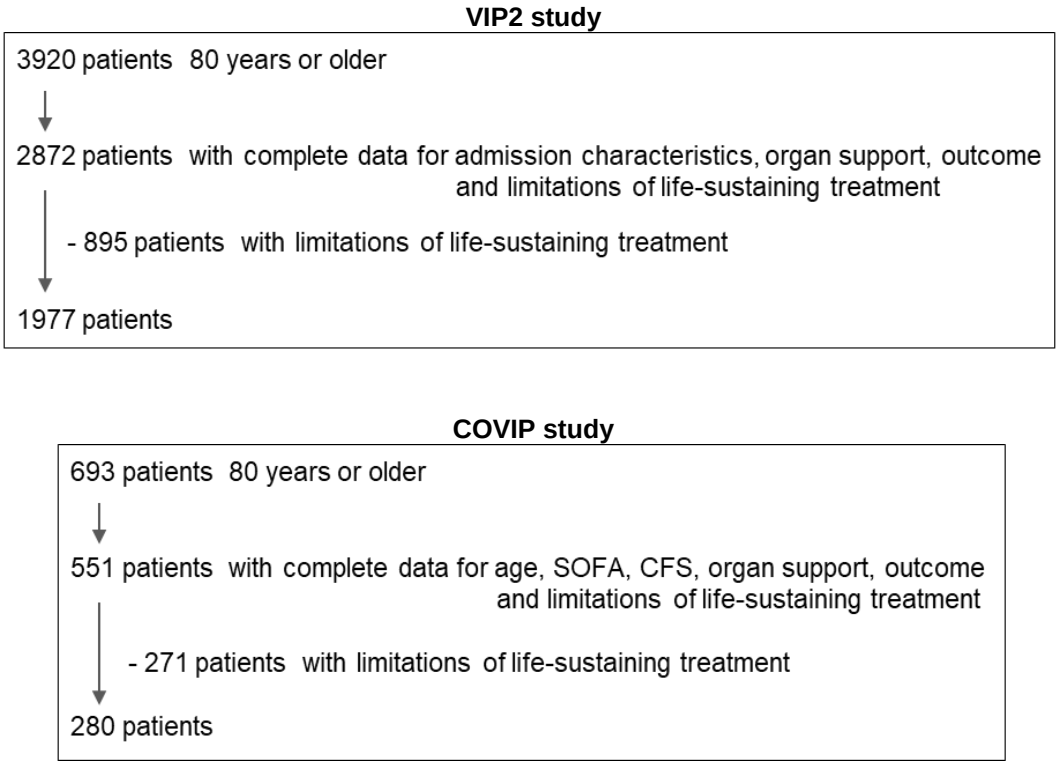

Figure S1. Selection of patients from the VIP2 and COVIP study cohorts.

2. Geriatric patient characteristics

The following geriatric patient characteristics were documented in the VIP2 study on admission to ICU [S1]:

- frailty assessed by the Clinical Frailty Scale (CFS - [S3]),
- cognitive impairments measured by the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE - [S4]),
- activities of daily living evaluated by the Katz index [S5],
- pre-existing chronic conditions summarised by the Comorbidity and Polypharmacy Score (CPS - [S6]).

The CFS is composed of nine categories ranging for the very fit (CFS 1) to the terminally ill (CFS 9) with frail patients having a CFS of 5 or above. The Katz index is based on classifying the performance in six daily activities (bathing, dressing, toileting, transfering, feeding, continence) according to the required level of support (0 - dependent, 1 - independent). The index summarises this assessment, e.g. a Katz index of 6 indicates independence in daily living. The IQCODE is based on a questionnaire that was filled in by caregivers with intimate knowledge of the patient. Cognitive impairment was defined as an overall score of 3.5 or higher [S7]. The CPS is determined by summing up all chronic comorbid conditions and medications.

3. Phenotyping by clustering analysis

The workflow of phenotyping is depicted in Figure S2.

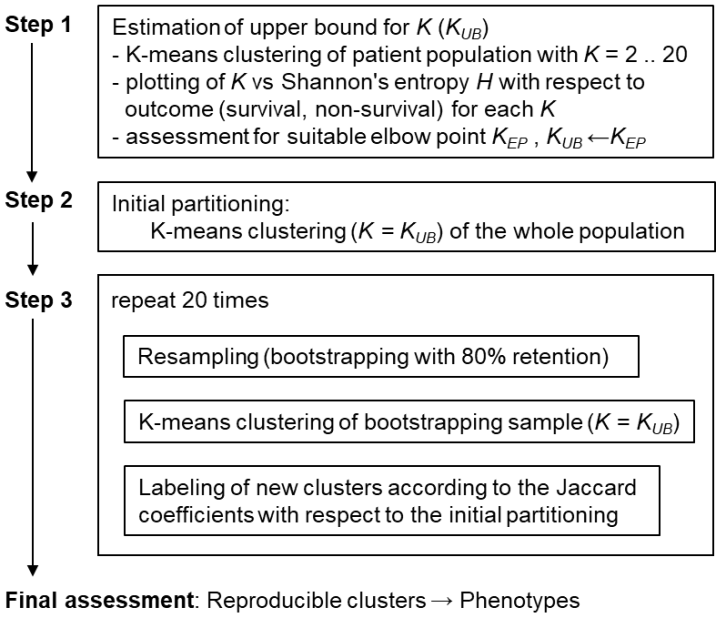

Figure S2. Workflow of phenotyping by clustering analysis. Please see Figure S3 for further details about Step 1.

Variables were first normalised by subtracting the mean and dividing by the standard deviation. We used the K-Prototypes algorithm, which is a variant of the K-means algorithm for handling both numerical and categorical variables [S8], to cluster patients according to the similarity of their n characteristics. This algorithm partitions the n-dimensional dataset into a predefined number K of clusters [S9]. Each patient is assigned to the cluster with the nearest, i.e. least distant, mean. Distances $D(X,Y)$ between two patients X and Y having n characteristics $x_{i=1..n}$ and $y_{i=1..n}$ are calculated as $D(X,Y) = \sum_{i \in numerical} (x_i - y_i)^2 + \sum_{j \in categorical} \delta(x_j - y_j)$. The first term is the Euclidean distance for numerical values and the second term is the distance between categorical variables which is the number of unmatched categories with $\delta(x_j, y_j) = \{0$ if $x_j = y_j$, 1 if $x_j \neq y_j$.

Determining a suitable number of phenotypes is important for an appropriate representation of the diversity within a patient population. Since there is no 'gold standard' for estimating that number, previous studies on phenotyping patients by K-means clustering applied different techniques for that purpose [e.g. S10, S11]. In this study, we have implemented a method (Figure S2) that first obtains an estimate for K, considered an upper bound K_{UB} (Figure S3), then performs K-means clustering with $K = K_{UB}$ in consecutive 20 bootstrapping samples (80% retention rate) and finally selects only those clusters from the K_{UB} clusters which are found to be robust against random variations of the data [S12] (Figure S4). This process determines the number of reproducible clusters, which are considered phenotypes, by using intrinsic properties of the data and requires only an initial estimation of K_{UB} .

In step 1, K_{UB} is estimated by finding an appropriate elbow point in the plot of K against the information in K clusters (Figure S3). For each clustering run resulting in K clusters, the information provided by a single cluster i is estimated with respect to outcome (ICU mortality) by Shannon's entropy H_i [S13] with p_i representing the probability of death within that cluster i: $H_i = - p_i \log_2 p_i - (1-p_i) \log_2 (1-p_i)$. The total entropy TE for each clustering run with fixed K is determined as the size-weighted mean of H for all clusters $i = 1..K$ of size n_i : $TE = {}^1/_N \sum_{i=1..K} n_i H_i$.

Figure S3. Total entropy TE plotted against the number of clusters *K*. *K*=12 was chosen as an appropriate elbow point in that plot which represents the upper bound K_{UB} for the VIP2 dataset.

In step 2, we obtained an initial partitioning of the patient population into $K = K_{UB}$ clusters. After each resampling (bootstrapping with 80% retention) and reclustering run (step 3, Figure S2), the new K_{UB} clusters detected by the K-means algorithm were labeled according to their maximum overlap with the clusters from the initial partitioning as measured by the Jaccard similarity coefficient [S14]. Finally, sets of patients which were reproducibly assigned to the same initial cluster during repeated resampling and reclustering runs and represented more than 1% of the whole cohort were considered phenotypes (Figure S4). The role of specific patient characteristics in that process was estimated by the 'permutation feature importance' technique [S15] (Figure S5).

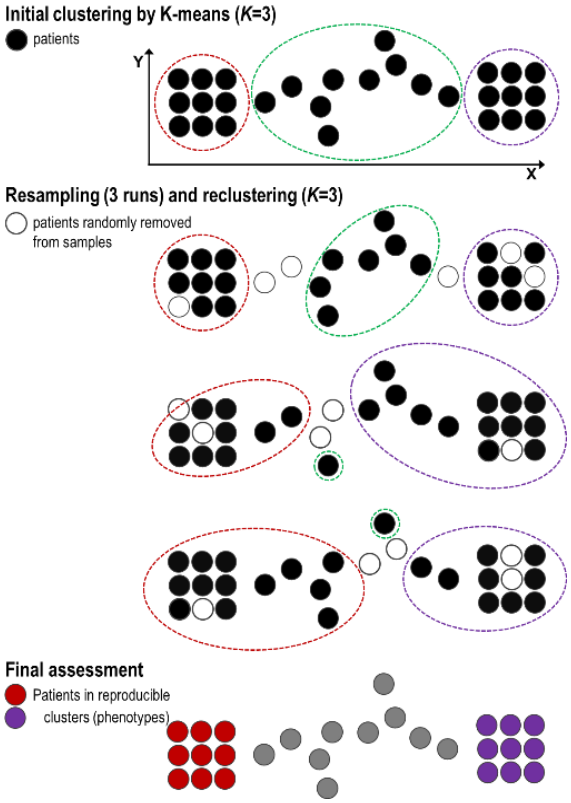

Figure S4. Illustration of the clustering of the patient population by their phenotypical similarity based on two characteristics x and y. Sets of patients which are very similar, i.e. close to each other in the 2-dimensional data space, and reproducibly assigned to the same labels (depicted in colours) despite random variations of the dataset are designated as phenotypes.

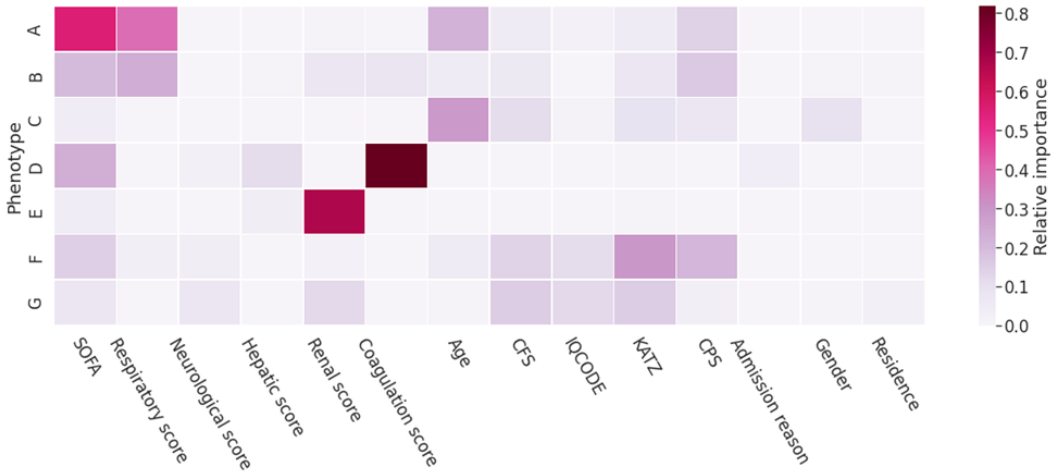

Figure S5. Relative importance of patient characteristics for clustering analysis.

In an additional step, grouping of clusters was performed to examine if the number of phenotypical categories can be reduced by relaxing the reproducibility criterion for identifying phenotypes. Clusters which shared more than 5% of the patient population during repeated bootstrapping and reclustering runs are merged into phenotypical groups. Figure S4 depicts an example where all dots in the lower row would be merged into a single phenotypical group during that process due to significant overlaps. Note that this algorithm can lead to sequential merging of clusters in the data space with subsequent loss of differential information. For example, when considering the x-axis in Figure S4 as representing a specific patient characteristic, distinctions that exist between the two colour-labeled phenotypes (interpreted as 'low x' vs 'high x' phenotypes) disappear after grouping was performed. The impact of that effect on geriatric characteristics in the VIP2 study cohort is demonstrated in Figure S6. After differences in the geriatric characteristics and age were attenuated during grouping, the SOFA score becomes the main determinant for differences between groups I and II.

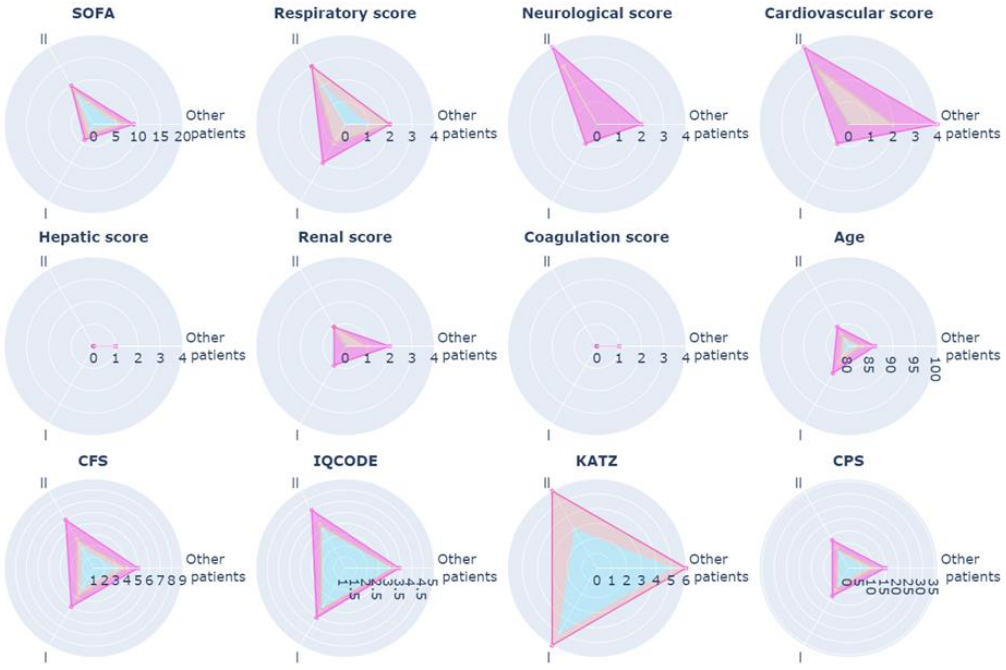

Figure S6. Grouping of clusters in the VIP2 study cohort (n=1977). Distributions of patient characteristics are shown for two phenotypical groups (I, II). The colours blue, orange and pink represent the 25th, 50th and 75th percentile of the distribution, respectively. Note the loss of differential information for geriatric characteristics and age between groups I amd II in comparison to the profiles of seven phenotypes shown in Figure 1.

4. Contribution of individual countries to VIP2 phenotypes

Figure S7 shows the relative contribution of the countries, which participated in the VIP2 study, to the seven phenotypes A - G.

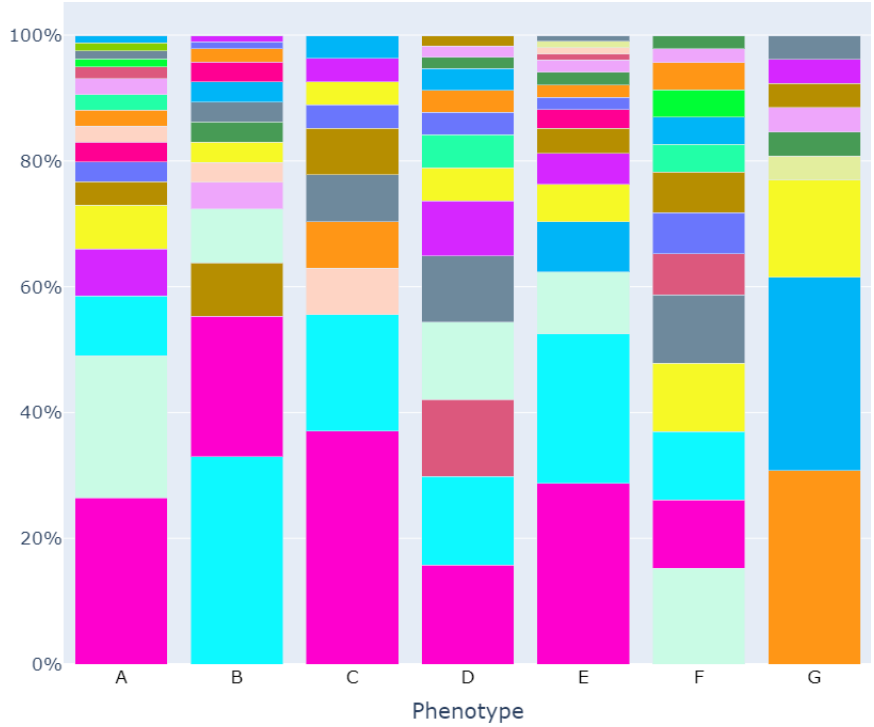

Figure S7. Relative contribution of different countries, distinguished by colour, to the phenotypes A - G identified in the VIP2 study cohort.

References

S1. Guidet B, de Lange DW, Boumendil A, Leaver S, Watson X, Boulanger C, Szczeklik W et al (2020) The contribution of frailty, cognition, activity of daily life and comorbidities on outcome in acutely admitted patients over 80 years in European ICUs: the VIP2 study. Intensive Care Med 46(1):57-69. https://doi/10.1007/s00134-019-05853-1

S2. Jung C, Flaatten H, Fjølner J, Bruno RR, Wernly B, Artigas A, Bollen et al (2021) The impact of frailty on survival in elderly intensive care patients with COVID-19: the COVIP study. Crit Care 25(1):149. https://doi/10.1186/s13054-021-03551-3

S3. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A (2005) A global clinical measure of ftness and frailty in elderly people. CMAJ 173:489-95.

S4. Jorm A (2004) The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE). Int Psychogeriatr. 16:1-19.

S5. Katz S (1983) Assessing self-maintenance: activities of daily living, mobildity, and instrumental activities of daily living. J Am Geriatr Soc. 31:721-7.

S6. Evans DC, Cook CH, Christy JM, Murphy CV, Gerlach AT, Eiferman D, Lindsey DE et al (2012) Comorbidity-polypharmacy scoring facilitates outcome prediction in older trauma patients. J Am Geriatr Soc. 60:1465-70.

S7. Quinn TJ, Fearon P, Noel-Storr AH, Young C, McShane R, Stott DJ. (2014) Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within community dwelling populations. Cochrane Database of Syst Rev. https://doi.org/10.1002/14651858.cd010079.pub2

S8. Huang ZX (1997) Clustering large data sets with mixed numeric and categorical values. Proceedings of the First Pacific Asia Knowledge Discovery and Data Mining Conference, Singapore. 1997: 21-34.

S9. Jain AK (2010) Data clustering. Pattern Recogn Letters. 31(8): 651-66.

S10. Geri G, Vignon P, Aubry A, Fedou AL, Charron C, Silva S, Repessé X, VieillardBaron A (2019) Cardiovascular clusters in septic shock combining clinical and echocardiographic parameters: a post hoc analysis. Intensive Care Med 45(5):657-67. https://doi/10.1007/s00134-019-05596-z

S11. Seymour CW, Kennedy JN, Wang S, Chang CH, Elliott CF, Xu Z, Berry S et al (2019) Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis. JAMA 321(20):2003-17. https://doi/10.1001/jama.2019.5791

S12. Yu, H., Chapman, B., Di Florio, A., Eischen, E., Gotz, D., Jacob, M., & Blair, R. H. (2019). Bootstrapping estimates of stability for clusters, observations and model selection. Computational Statistics. 34(1):349-72. DOI 10.1007/s00180-018-0830-y

S13. Shannon CE, Sloane NJA, Wyner AD (1993). Claude Elwood Shannon: Collected Papers. IEEE Press; ISBN:978-0-7803-0434-5

S14. Hancock JM (2004). Jaccard distance (Jaccard index, Jaccard similarity coefficient). Dictionary of Bioinformatics and Computational Biology. John Wiley & Sons.

S15. Altmann A, Toloşi L, Sander O, Lengauer T (2010) Permutation importance: a corrected feature importance measure, Bioinformatics. 26(10):1340-47. https://doi.org/10.1093/bioinformatics/btq134