Clustering analysis of geriatric and acute characteristics in a cohort of very old patients on admission to ICU

Oded Mousai1, Lola Tafoureau1, Tamar Yovell1, Hans Flaatten2, Bertrand Guidet3, Christian Jung4, Dylan de Lange5, Susannah Leaver6, Wojciech Szczeklik7, Jesper Fjolner8, Peter Vernon van Heerden9, Leo Joskowicz1, Michael Beil10, Gal Hyams1 and Sigal Sviri10*

© 2022 Springer-Verlag GmbH Germany, part of Springer Nature

Abstract

Purpose: The biological and functional heterogeneity in very old patients constitutes a major challenge to prognostication and patient management in intensive care units (ICUs). In addition to the characteristics of acute diseases, geriatric conditions such as frailty, multimorbidity, cognitive impairment and functional disabilities were shown to influence outcome in that population. The goal of this study was to identify new and robust phenotypes based on the combination of these features to facilitate early outcome prediction.

Methods: Patients aged 80 years old or older with and without limitations of life-sustaining treatment and with complete data were recruited from the VIP2 study for phenotyping and from the COVIP study for external validation. The sequential organ failure assessment (SOFA) score and its sub-scores taken on admission to ICU as well as demographic and geriatric patient characteristics were subjected to clustering analysis. Phenotypes were identified after repeated bootstrapping and clustering runs.

Results: In patients from the VIP2 study without limitations of life-sustaining treatment (n = 1977), ICU mortality was 12% and 30-day mortality 19%. Seven phenotypes with distinct profiles of acute and geriatric characteristics were identified in that cohort. Phenotype-specific mortality within 30 days ranged from 3 to 57%. Among the patients assigned to a phenotype with pronounced geriatric features and high SOFA scores, 50% died in ICU and 57% within 30 days. Mortality differences between phenotypes were confirmed in the COVIP study cohort (n = 280).

Conclusions: Phenotyping of very old patients on admission to ICU revealed new phenotypes with different mortality and potential need for anticipatory intervention.

Keywords: Clustering analysis, Geriatric patients, Heterogeneity, Intensive care, Phenotyping

Introduction

The outcome of very old patients admitted to the intensive care unit (ICU) is determined by both the severity of the acute condition and the age-related decline of resilience to stress. However, that decline is characterised by substantial inter-individual heterogeneity which, of note, is considered a hallmark of the ageing process [1]. Thus, it is unsurprising that disease severity scores developed
for critical conditions in younger cohorts show poorer prognostic discrimination in cohorts of very old individuals [2]. Geriatric conditions in very old patients, such as frailty, multimorbidity, cognitive impairment and functional deficits in daily life, were shown to delineate groups with different prognoses in that population [3–6]. The clinical frailty scale (CFS) was introduced into clinical practice as a prognostic aid to triaging individuals for ICU admission during the initial surge of the coronavirus disease 2019 (COVID-19) pandemic [7]. However, the variability and complexity of very old patients appears to require a more holistic approach, using a broader spectrum of patient characteristics for prognostic stratification [8].

Substantial heterogeneity within a patient population and non-linear relationships between patient characteristics may interfere with the accuracy of techniques for predictive modelling [9, 10]. Clustering analysis is an intuitive approach to utilise these features advantageously to classify individuals by their similarity [11]. This technique is gaining popularity since it is an efficient approach to exploring high-dimensional datasets for identifying phenotypes (clusters) of patients with distinct characteristics, indicating specific needs and divergent outcomes in healthcare [12–14]. The additional information gained from clusters may, thus, enhance prognostic accuracy and help to develop interventions targeting the deficiencies shared by the members of a particular phenotype [15–22]. For example, Seymour et al. [21] identified four sepsis phenotypes with 28-day mortality varying from 5 to 40%. One of these phenotypes representing 27% of patients was characterised by an enhanced inflammatory response which might benefit from tailored interventions in the future. Very few of the past studies, however, focused on very old patients with critical conditions [5].

The VIP2 study examined the association of survival with acute and geriatric features in ICU patients aged 80 years or older [23]. The current study investigated the heterogeneity within that population by clustering analysis of patient characteristics recorded on admission to the ICU. We hypothesised that acute and geriatric characteristics delineate distinct phenotypes which can assist early prognostication and indicate options for anticipatory interventions, such as early rehabilitation in frail patients with reversible conditions or palliative care at the predicted end of life. To this end, particular emphasis was put on the robustness and information content of phenotypes.

**Take-home message**

Clustering analysis of very old patients in critical condition reveals predictive phenotypes on admission to the intensive care unit. Information about geriatric characteristics can be used for planning tailored interventions.

**Patients and methods**

**Patients**

The Very elderly Intensive care Patient (VIP)2 study was a prospective observational study in patients aged 80 years or older admitted as emergencies to ICUs in 22 countries [23]. The study objective was to investigate the impact of acute and geriatric patient characteristics on survival in ICU and after 30 days. National coordinators were responsible for registering ICUs and obtaining ethics committee approval in their countries. Participating units recruited consecutive patients who met the above demographic and clinical criteria during any 6-month period between May 2018 and May 2019. This new study included all patients from the VIP2 dataset with complete data and, initially, without limitations (withholding or withdrawing) of life-sustaining treatment (LST) (Figure S1, supplementary material). The latter criterion was introduced to minimise the impact of subjective and variable decision-making on outcome [24]. For sensitivity analysis, we later included patients with limitations of LST.

To validate the findings in the VIP2 population with an independent dataset, we extracted a subset of patients from the COVID-19 in VIP (COVIP) study, a prospective observational study of old patients admitted to the ICU with COVID-19 in 14 countries between March and December 2020 [25]. This subset encompassed all patients aged 80 years and older with complete data for age, sequential organ failure assessment (SOFA) score, CFS and outcome and without limitations of LST (Figure S1, supplementary material).

**Measurements**

The following patient characteristics were recorded in the VIP2 study on admission to the ICU [23] and used for phenotyping in the current study: age, gender, residence prior to admission, reason for admission to ICU, SOFA score and its sub-scores (Table 1) as well as the following geriatric features: CFS, informant questionnaire on cognitive decline in the elderly (IQCODE), Katz score to assess activities of daily living, comorbidity and polypharmacy score (CPS) (Sect. 2 in the supplementary material). Ventilatory and vasopressor support and renal replacement therapy were recorded as interventions. Death in ICU and within 30 days, regardless of place, and
length of stay in ICU were documented as outcome. The COVIP study recorded a similar set of variables, except IQCODE and CPS [25].

**Phenotyping by clustering analysis**

Clustering is a method for grouping of patients who are assigned different labels based on the (dis)similarity of their characteristics [11]. We implemented a variant of the K-means algorithm which partitions patient populations into a pre-defined number $K$ of clusters in which distances between the patients’ data points are minimised [11]. This algorithm was embedded in a three-step procedure to obtain robust clusters (phenotypes): (1) estimation of an upper bound $K_{UB}$ for the total number of clusters, (2) initial clustering of the patient population with $K = K_{UB}$ and (3) repeated re-sampling by bootstrapping and re-clustering to identify those clusters which are robust against random variations of the data. Although the initial number $K$ of clusters was set, the final number of phenotypes was determined by the intrinsic properties of the data. Moreover, we examined if the number of phenotypes can be reduced without substantial loss of information by grouping clusters according to their similarity. Additional details about these methods are provided in the supplementary material.

For internal validation, supervised classifiers (gradient boosting) were used to verify the discrimination of phenotypes. The classifiers were trained on the phenotype labels and based on the same set of patient characteristics used for clustering. The area under the receiver operating characteristic (AUROC) was obtained as a measure of discrimination between patients of a specific phenotype and all other patients. External validation of predictive phenotyping was based on comparing the mortality of VIP2 phenotypes with that of similar patients from the COVIP study. Similarity was approximated by measuring the distance between the COVIP patient’s data points to the centroids of VIP2 phenotypes in the data space using three simple and informative patient characteristics (age, SOFA, CFS) having the smallest number of missing values [23].

**Statistical analysis**

Descriptive characteristics were reported as medians with inter-quartile ranges (IQRs) for continuous variables and proportions for nominal variables. One-way ANOVA test was used to examine differences of continuous variables and Fisher’s exact test and its variants for nominal variables. The resulting $p$ values were adjusted to control for false discovery rates by applying the Benjamini–Hochberg procedure [27]. Odds ratios with 95% confidence intervals were calculated for binary variables (outcome, interventions) for each phenotype using the phenotype with the lowest mortality as reference [28]. All analyses were performed using Python 3 (Python Software Foundation, Beaverton, OR, USA).

### Table 1  Admission characteristics, interventions and outcome for patients aged 80 years or older without limitations of life-sustaining treatment from the VIP2 and COVIP studies

| Variable                                      | VIP2 study | COVIP study | $p$ value |
|-----------------------------------------------|------------|-------------|-----------|
| Number of patients                            | 1977       | 280         |           |
| Age—median (IQR)                              | 83 (81–86) | 83 (81–85)  | $<0.001$  |
| Gender—number (% female)                      | 989 (45%)  | 103 (37%)   | 0.01      |
| Residence—number (%)                          |            |             |           |
| Own home                                      | 1495 (76%) | 182 (65%)   | $<0.001$  |
| Home with family or caregivers                | 208 (10%)  | 48 (17%)    | 0.003     |
| Nursing home or other hospital                | 251 (12%)  | 44 (16%)    | 0.22      |
| Other                                         | 23 (1%)    | 0           | 0.13      |
| Admission reason—number (%)                  |            |             |           |
| Respiratory/cardiovascular                    | 905 (46%)  | (COVID-19)  |           |
| Sepsis                                        | 243 (12%)  |             |           |
| Emergency surgery                             | 311 (16%)  |             |           |
| Other                                         | 518 (26%)  |             |           |
| SOFA score—median (IQR)                       | 5 (3–9)    | 5 (3–8)     | 0.7       |
| sub-scores:                                   |            |             |           |
| Respiratory                                   | 2 (1–3)    | 3 (2–3)     | $<0.001$  |
| Cardiovascular                                | 1 (0–3)    | 0 (0–3)     | $<0.001$  |
| Hepatic                                       | 0 (0–0)    | 0 (0–0)     | 0.01      |
| Coagulation                                   | 0 (0–1)    | 0 (0–0)     | 0.66      |
| Renal                                         | 1 (0–2)    | 0 (0–1)     | 0.03      |
| Neurological                                  | 0 (0–2)    | 0 (0–1)     | 0.06      |
| Geriatric features—median (IQR)               |            |             |           |
| CFS                                           | 4 (3–5)    | 4 (3–6)     | 0.07      |
| Katz index                                    | 6 (5–6)    | [6 (3–6)*]  |           |
| IQCODE                                        | 3.1 (3–3.5)| n/a         |           |
| CPS                                           | 10 (7–14)  | n/a         |           |
| Interventions—number (%)                      |            |             |           |
| Invasive ventilation                          | 934 (47%)  | 175 (62%)   | $<0.001$  |
| Vasopressors                                  | 1099 (56%) | 160 (57%)   | 0.7       |
| Renal replacement therapy                     | 215 (11%)  | 36 (13%)    | 0.55      |
| ICU length of stay (days)—median (IQR)        | 4 (2–8)    | 7 (3–13)    | 0.001     |
| Outcome—number (%)                            |            |             |           |
| Died in ICU                                   | 231 (12%)  | 117 (42%)   | $<0.001$  |
| Died within 30 days                           | 384 (19%)  | 128 (46%)   | $<0.001$  |

Note the percentage of missing values for the Katz index in the COVIP study cohort

*23% missing values

CFS, clinical frailty scale; COVID-19, coronavirus disease 2019; CPS, comorbidity and polypharmacy score; ICU, intensive care unit; IQCODE, informant questionnaire on cognitive decline in the elderly; IQR, interquartile range; SOFA, sequential organ failure assessment
Table 2  Admission characteristics, interventions and outcome for phenotypes in the VIP2 study cohort

| Variable                  | Phenotype                  | p value |
|---------------------------|----------------------------|---------|
|                           | Non-geriatric              | Geriatric |
|                           | A-very low SOFA failure    | B-respiratory failure | C-oldest old SOFA | D-moderate SOFA | E-renal failure | F-low SOFA | G-high SOFA |<0.001 |
| Number of patients        | 159                        | 94      | 27      | 57      | 101         | 46          | 26         |<0.001 |
| Age—median (IQR)          | 82 (81–84)                 | 83 (81–84) | 93 (90–95) | 83 (81–86) | 83 (81–84) | 84 (82–87) | 87 (83–90) |<0.001 |
| Gender—number (% female)  | 84 (53%)                   | 41 (44%) | 18 (67%) | 17 (30%) | 29 (29%)   | 25 (54%)   | 12 (46%)   |<0.001 |
| Residence—number (%)      | Own home 140 (88%)         | 79 (84%) | 17 (63%) | 48 (84%) | 86 (85%)   | 32 (69%)   | 9 (34%)    |<0.001 |
|                           | Home with family or caregivers 8 (5%) | 7 (7%) | 5 (33%) | 8 (7%) | 6 (13%) | 6 (13%) | 9 (34%) |<0.001 |
|                           | Nursing home or other hospital 9 (5%) | 8 (8%) | 1 (3%) | 7 (6%) | 5 (10%) | 8 (30%) | 0.0017 |
| Other                     | 23 (1%)                    | 2 (1%)   | 0       | 0       | 0          | 3 (6%)    | 0         |0.0051 |
| Admission reason—number (%) | Respiratory/cardiovascular | 28 (17%) | 67 (71%) | 13 (48%) | 22 (38%) | 48 (37%) | 25 (54%) | 13 (50%) |<0.001 |
|                           | Sepsis 4 (2%)              | 6 (6%)   | 1 (3%)  | 9 (15%) | 18 (17%)  | 3 (6%)    | 10 (38%)  |<0.001 |
|                           | Emergency surgery Other    | 31 (19%) | 10 (10%) | 3 (11%) | 12 (21%)  | 5 (10%)  | 0          |0.015 |
| SOFA score—median (IQR)   | 1 (0–2)                    | 4 (3–4)  | 4 (2–4) | 6 (5–7) | 6 (5–8)   | 3 (2–4)   | 10 (8–13) |<0.001 |
| Geriatric features—median (IQR) | CFS 3 (2–3) | 4 (3–4) | 4 (3–5) | 3 (2–4) | 4 (3–4) | 6 (6–7) | 7 (7–7) |<0.001 |
|                           | Katz index 6 (6–6)         | 6 (6–6)  | 6 (5–6) | 6 (5–6) | 6 (5–6)   | 2 (1–3)   | 1 (0–1)   |<0.001 |
|                           | IQCODE 3 (3–3.2)           | 3.1 (3–3.2) | 3.3 (3.1–3.6) | 3 (3–3.2) | 3 (3–3.4) | 3.8 (3.3–4.1) | 4.5 (4.1–5) |<0.001 |
|                           | CPS 7 (5–9)                | 12 (11–14) | 9 (6–10) | 9 (7–12) | 11 (8–15) | 15 (13–19) | 11 (7–15) |<0.001 |
| Interventions—number (%)  | Invasive ventilation 24 (15%) | 34 (36%) | 7 (25%) | 16 (28%) | 23 (22%) | 8 (17%) | 23 (88%) |<0.001 |
|                           | Vasopressors 26 (16%)      | 24 (25%) | 9 (33%) | 24 (42%) | 47 (46%) | 13 (28%) | 25 (96%) |<0.001 |
|                           | Renal replacement therapy 2 (1%) | 0     | 0       | 2 (3%)  | 45 (45%) | 3 (6%)   | 4 (15%)   |<0.001 |
| Length of stay (days)—median (IQR) | 2 (1–4)  | 5 (2–8) | 2 (1–4) | 3 (2–6) | 5 (3–8) | 3 (2–5) | 4 (2–12) |<0.001 |
| Outcome—number (%)        | Died in ICU 3 (2%)         | 3 (3%)  | 1 (3%)  | 4 (7%)  | 8 (7%)   | 4 (8%)    | 13 (50%)  |<0.001 |
|                           | Died within 30 days 5 (3%) | 11 (11%) | 2 (7%)  | 8 (14%) | 13 (12%) | 8 (17%) | 15 (57%) |<0.001 |
Results

Table 1 depicts the characteristics of the patient cohorts from the VIP2 \( (n=1977) \) and COVIP \( (n=280) \) studies without limitations of LST. Based on the complete set of acute and geriatric characteristics recorded on admission to ICU, clustering analysis of the VIP2 cohort identified seven phenotypes encompassing 26% of this patient population (Table 2). Internal validation of phenotyping by supervised classification yielded AUROCs greater than 0.95 for distinguishing each of these phenotypes from all other patients in that cohort. Figure S7 (supplementary material) shows the distribution of countries with respect to these phenotypes.

Phenotypes F and G (Table 2, Fig. 1A) are considered ‘geriatric’ since most of these patients are frail (CFS \( \geq 5 \)), have limitations in their activities of daily living (Katz \( < 5 \)) or cognitive impairments (IQCODE \( \geq 3.5 \)). However, these two phenotypes significantly differ in their SOFA scores (\( p<0.001 \)). In contrast, phenotype A is characterised by the absence of geriatric features and very low SOFA scores. The other ‘non-geriatric’ phenotypes B–E are distinguishable by age (phenotype C), respiratory and cardiovascular (phenotypes B and D) or renal SOFA subscores (phenotype E). Forty-five percent of patients in the latter phenotype received renal replacement therapy in the ICU.

Sensitivity analyses were performed with respect to the inclusion of patients with limitations of LST and the number of phenotypical categories. Phenotyping of all VIP2 patients with complete data, i.e. after adding patients with limitations of LST \( (n=895) \) to the main study cohort \( (n=1977) \), resulted in the same number of phenotypes with similar profiles (Fig. 1B). Major differences were only found for phenotype D where the inclusion of patients with limitations of LST led to an increase of the fraction of patients with severe respiratory or cardiovascular failure. An attempt to reduce the total number of phenotypical categories by relaxing the reproducibility criterion for clusters and grouping overlapping clusters caused a loss of information about geriatric characteristics (Sect. 3, supplementary material).

Phenotyping was based on information available on admission to ICU and did not consider later events. The phenotypes, however, were found to be associated with variable rates of organ support in ICU as well as a different ICU and 30-day mortality (Tables 2 and 3, Fig. 2). Phenotype A was characterised by very low mortality (2% in ICU, 3% after 30 days) and the lowest rates of organ support. The highest mortality was found for phenotype G (50% in ICU, 57% within 30 days) that also showed the highest rate of respiratory and cardiovascular support. Odds ratios for mortality and organ support were calculated with phenotype A as reference (Table 3). Note that large odds ratios and wide confidence intervals for phenotype G are related to relatively high event rates in a small sample.

We also investigated whether the mortality estimates for phenotypes can be used for patients who are similar but not identical to these phenotypes. Figure 2 depicts the distributions of ICU and 30-day mortality for all clusters obtained during repeated bootstrapping runs and associated with specific phenotypes. These clusters can be considered random samples of the immediate neighbourhood of phenotypes within the data space. The small width of these distributions indicates that there is only a small variation of mortality in this neighbourhood suggesting that patients who are formally outside but still similar to these phenotypes have a comparable outcome.

Finally, we examined whether outcome predictions for phenotypes in the VIP2 dataset are generalisable. In the absence of another study with an identical design and set of patient characteristics, we chose the dataset of the recent COVIP study [25] for that purpose. Patients from the COVIP study, who were 80 years or older and admitted to the ICU with COVID-19, were associated with the closest VIP2 phenotype based on similarity with regard to three admission characteristics (age, SOFA, CFS). ICU and 30-day mortality in these COVIP sub-groups is depicted in Fig. 2. Although the absolute mortality differed between the phenotypes from VIP2 and the associated sub-groups from COVIP, the trend, especially with phenotype A representing the lowest and phenotype G the highest mortality, was reproducible.

Discussion

This study investigated the heterogeneity of very old patients on admission to the ICU. Using clustering analysis with rigorous criteria for reproducibility, we identified seven phenotypes in patients from the VIP2 study. These phenotypes represent distinct profiles of acute and geriatric characteristics and are associated with different...
Fig. 1 (See legend on previous page.)
mortality. Considering that up to five distinct phenotypes are required to describe populations of younger patients with sepsis [19, 21], seven phenotypes appear to be an appropriate number to reflect the heterogeneity in very old patients admitted to the ICU with a broad spectrum of conditions. Sensitivity analyses further validated that number. Although the fraction of patients formally represented by these phenotypes was rather small, we demonstrated that their prognostic information can also be used for additional patients who have similar characteristics on admission.

There are considerable differences in outcome across the spectrum of phenotypes with phenotype A having the lowest ICU and 30-day mortality. The two phenotypes with a prominent geriatric profile (phenotypes F and G) exhibit the highest mortality. The combination of a geriatric profile with a high SOFA score (phenotype G) resulted in an ICU mortality of 50% emphasising the major impact of reduced resilience in geriatric patients. Phenotype C includes the oldest old patients with low SOFA scores and only very mild geriatric characteristics. A 30-day mortality of < 10% in this group underlines that advanced age alone should not be considered a negative prognosticator. Younger patients with substantial organ failure but without frailty or other major chronic impairments were found in phenotypes B, D and E with a 30-day mortality between 12 and 14%. These findings underline that mortality in very old ICU patients is determined by geriatric vulnerabilities and the acute illness. Both types of conditions need to be considered when prognosticating in that heterogeneous population.

Previous studies on old patients in ICU did not record the ensemble of geriatric and acute characteristics to a similar extent as the VIP2 study did. In one of the few other investigations, Ferrante et al. [4] showed an association of outcome with markers of organ failure, cognitive impairment and functional disabilities. However, that study did not distinguish between patients with limitations of LST and those without. Survival data were thus impacted by subjective decisions about LST. In contrast, the design of the VIP2 study allowed us to identify and exclude patients with limitations of LST from outcome analysis [23]. Our findings, therefore, mainly reflect the natural course of critical conditions in very old individuals without major interference by the variable assessment of decision-makers [24]. In comparison to VIP2, the COVIP study [25] had a similar design but recorded a smaller set of geriatric features. Using three simple admission characteristics (age, CFS, SOFA), the trend of mortality differences between phenotypes found in VIP2 could be replicated in the COVIP dataset. This indicates generalisability of our findings despite marked differences in absolute mortality between VIP2 and COVIP, which can be attributed to the substantial severity of COVID-19 within the population of very old individuals [29]. The reductionist approach to approximate similarity between patients with only three easy to obtain characteristics may facilitate comparisons in future studies but requires further validation. Of note, phenotyping based on this limited set of variables would not result in identical clusters, since several phenotypes, e.g. renal phenotype E, are mostly defined by variables not included in that set.

How can the knowledge about phenotypes be used in clinical practice? Predicted 30-day mortality smaller than 10% permit early prognostication of a favourable outcome for phenotypes A and C. These phenotypes represent patients of all age groups with low SOFA scores and without marked geriatric characteristics. Phenotypes B, D and E are close to that prognostic category. Frail patients with advanced cognitive and functional impairments (phenotypes F and G) have a guarded prognosis, especially those with a substantial degree of organ failure (phenotype G). Importantly, phenotype G has an ICU mortality of approximately 50%, i.e. prognostication on admission equals random guessing. This uncertainty is reflected by the substantial variability of ICU length of stay (LOS). Although the fraction of patients formally represented by these phenotypes was rather small, we demonstrated that their prognostic information can also be used for additional patients who have similar characteristics on admission.

### Table 3 Odds ratios with 95% confidence intervals for invasive ventilation, vasopressor support, mortality in ICU and within 30 days analysed as binary variables with phenotype A as reference

| Phenotype | Invasive ventilation | p value | Vasopressors | p value | ICU mortality | p value | 30-day mortality | p value |
|-----------|----------------------|---------|--------------|---------|---------------|---------|------------------|---------|
| A (reference) | 1 (reference) | 1 (reference) | 1 (reference) | 1 (reference) | 1 (reference) | 1 (reference) | 1 (reference) | 1 (reference) |
| B | 3.19 [1.74, 5.83] | < 0.001 | 1.75 [0.94, 3.28] | 0.1 | 1.71 [0.34, 8.67] | 0.6734 | 4.08 [1.37, 12.1] | 0.013 |
| C | 1.97 [0.75, 5.16] | 0.17 | 2.56 [1.04, 6.32] | 0.058 | 2.0 [0.2, 20] | 0.469 | 2.46 [0.45, 13.4] | 0.27 |
| D | 2.2 [1.07, 4.52] | 0.045 | 3.72 [1.9, 7.29] | < 0.001 | 3.92 [0.85, 18.1] | 0.0807 | 5.03 [1.57, 16.1] | 0.0064 |
| E | 1.66 [0.88, 3.13] | 0.14 | 4.45 [2.51, 7.9] | < 0.001 | 4.47 [1.16, 17.3] | 0.0058 | 4.55 [1.57, 13.2] | 0.0045 |
| F | 1.18 [0.49, 2.85] | 0.82 | 2.02 [0.94, 4.34] | 0.088 | 4.95 [1.07, 23] | 0.0463 | 6.48 [2.01, 20.9] | 0.0019 |
| G | 43.1 [12.0, 155] | < 0.001 | 128 [16.6, 986] | < 0.001 | 52 [13.1, 206] | < 0.001 | 42 [12.9, 137] | < 0.001 |

*p values are provided for comparison with phenotype A*
stay in these patients. In this situation, more information, such as about the individual’s response to treatment during a time-limited trial (TLT), will be necessary for more informative outcome predictions [30, 31]. However, the specific management of these patients, including decisions about ICU admission and TLT, may depend on the preferences within a particular healthcare system.

The geriatric components of the phenotypes provide actionable information for anticipatory management in ICU. The higher mortality of phenotypes with marked geriatric features reflect the enhanced vulnerability and reduced resilience to stress [32]. This situation should trigger targeted interventions [13] which may become part of standard care in ICU. This could especially be pertinent to the prevention and management of delirium in patients with pre-existing cognitive impairments in phenotypes F and G [33], dealing with problems of polypharmacy by thoroughly reviewing prescriptions.

**Fig. 2** Mortality in ICU (A) and within 30 days (B) for the phenotypes from the VIP2 cohort (dots) and patients from the COVIP study (triangles) who are similar to a specific VIP2 phenotypes according to age, CFS and SOFA score. The box plots show the inter-quartile distribution of mortality for VIP2 patients who are similar to the respective phenotype.
in individuals with multimorbidity, e.g. in phenotype F, [34] and planning early rehabilitation for patients with frailty and functional impairments [35]. For example, the 82-year-old patient with frailty (CFS 6) who requires only limited ventilatory but no other organ support (phenotype F) may substantially benefit from these interventions to return to the previous level of functioning. Importantly, the 95-year-old patient with a low SOFA score, who is only borderline frail (phenotype C) and has an excellent prognosis, might be the one who benefits most from early rehabilitation. Specific interventions though, will differ from those for phenotype F as suggested by different functional baselines (Katz score). Preparing discussions about palliative care might be indicated for individuals matching phenotype G, such as an 85-year-old ventilated patient with cardiovascular as well as renal failure and CFS 7, who already required substantial assistance at home [36]. Importantly, these interventions should be coordinated at an interdisciplinary level and require organisational adjustments, e.g. by facilitating ICU discharge to suitable step-down units. Whether these measures can eventually offer a benefit for survival and functional outcome remains to be elucidated in prospective trials. Recruitment to such trials can now focus on specific phenotypes [37]. This situation is analogous to the current research into new sepsis phenotypes [21].

The most important limitation of our study is the focus on mortality in the short term. Despite the consensus that quality of life is the most important outcome for very old ICU patients, short-term mortality still plays an important role. First, one can argue that evaluating quality of life necessitates physical survival and it is still important for benchmarking [38]. Second, in various countries, triage criteria for admission to ICU during periods of resource constraints, in particular during the COVID-19 pandemic, included the probability of short-term survival [39]. More precise predictions based on both acute and geriatric characteristics are necessary to support a more accurate decision-making in that situation [40].

In conclusion, we detected reproducible phenotypes in very old ICU patients based on both acute and geriatric characteristics available on admission to ICU. These phenotypes are associated with substantial differences in short-term mortality. Future studies can focus on specific interventions tailored for specific phenotypes and their impact on long-term outcome measures.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1007/s00134-022-06868-x.

Author details
1 School of Computer Science and Engineering, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel. 2 Department of Anaesthesia and Intensive Care, Haukeland University Hospital, Bergen, Norway. 3 Assistance Publique-Hôpitaux de Paris, Hôpital Saint-Antoine, Service de Réanimation Médicale, Paris, France. 4 Department of Cardiology, Pulmonology and Vascular Medicine, Faculty of Medicine, Heinrich-Heine-University, Düsseldorf, Germany. 5 Department of Intensive Care Medicine, University Medical Center, University Utrecht, Utrecht, The Netherlands. 6 General Intensive Care, St George's University Hospitals NHS Foundation Trust, London, UK. 7 Center for Intensive Care and Perioperative Medicine, Jagiellonian University Medical College, Krakow, Poland. 8 Department of Anaesthesia and Intensive Care, Viborg Regional Hospital, Viborg, Denmark. 9 General Intensive Care Unit, Department of Anaesthesiology, Critical Care and Pain Medicine, Faculty of Medicine, Hebrew University and Hadassah University Medical Center, Jerusalem, Israel. 10 Department of Medical Intensive Care, Faculty of Medicine, Hebrew University and Hadassah University Medical Center, Jerusalem, Israel.

Acknowledgements
VIP1 and VIP2: Support for implementing the eCRF and running the database was granted by the Department of Epidemiology, University of Aarhus, Denmark. VIP2: Financial support for creation of the eCRF and maintenance of the database was supported by a grant by Western Health Region in Norway. DRC Ile de France and URC Est helped conducting this study in France. For a list of collaborators in VIP2 and COMIP please see [23] and [29].

Declarations
Conflict of interests
The authors do not have any conflict of interest to declare related to this manuscript.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The authors do not have any conflict of interest to declare related to this manuscript.
