

1 **Title**

2 Clustering ICU patients with sepsis based on the patterns of their circulating
3 biomarkers: a secondary analysis of the CAPTAIN prospective multicenter
4 cohort study.

5 **Short title**

6 Clustering patients with sepsis

7
8 Benoît Misset^{1,2*}, François Philippart^{3,4}, Catherine Fitting⁵, Jean-Pierre Bedos⁶, Jean-Luc
9 Diehl^{7,8}, Olfa Hamzaoui⁹, Djillali Annane^{10,11}, Didier Journois¹², Marianna Parlato⁵, Virginie
10 Moucadel¹³, Jean-Marc Cavaillon⁵, Joël Coste^{14,15}, for the Captain Study Group[^]

11
12 ¹ Intensive Care Department, Liège University Hospital, Liège, Belgium

13 ² Infection, immunity and inflammation Research Unit, GIGA I3, Liège University
14 Research Unit, Liège, Belgium

15 ³ Service de Médecine Intensive et Réanimation, Groupe hospitalier Paris Saint-Joseph, Paris,
16 France

17 ⁴ Unité endotoxines, structures et réponse de l'hôte, Département de Microbiologie, Institut de
18 Biologie Intégrative de la Cellule, Paris Saclay, Saclay, France

19 ⁵ Unit Cytokines & Inflammation, Institut Pasteur, Paris, France

20 ⁶ Service de Réanimation, Hôpital André Mignot, Versailles, France

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

21 ⁷ Service de Réanimation Médicale, Hôpital Européen Georges Pompidou, Assistance Publique
22 - Hôpitaux de Paris, Paris, France

23 ⁸ INSERM UMR S1140, Université Paris Descartes, Paris Sorbonne Cité, Paris France

24 ⁹ Service de Réanimation, Hôpital Antoine Bécclère, Assistance Publique - Hôpitaux de Paris,
25 Clamart, France

26 ¹⁰ Service de Réanimation, Hôpital Raymond Poincaré, Assistance Publique - Hôpitaux de
27 Paris, Garches, France

28 ¹¹ Université Versailles Saint-Quentin, Versailles, France

29 ¹² Service de Réanimation Chirurgicale, Hôpital Européen Georges Pompidou, Assistance
30 Publique - Hôpitaux de Paris, Paris, France

31 ¹³ BioMérieux SA, Lyon, France

32 ¹⁴ Unité de Biostatistiques et d'Epidémiologie, Hôpital Cochin, Assistance Publique - Hôpitaux
33 de Paris, Paris, France

34 ¹⁵ Université de Paris, Paris, France

35

36 *Corresponding author:

37 ORCID iD : 0000000164660065

38 e-mail: benoit.misset@chuliege.be (BM)

39

40 **Authors' contributions:**

41 Drs Misset, Philippart, Cavaillon and Coste had full access to all the data in the study and take
42 responsibility for the integrity of the data and the accuracy of the data analysis.

43 Study concept and design: Misset, Philippart, Cavaillon and Coste

44 Acquisition of data: Philippart, Fitting, Parlato, Moucadel, Bedos, Journois, Hamzaoui,
45 Annane, Diehl

46 Access to data, analysis and interpretation: Misset, Philippart, Parlato, Moucadel, Cavaillon and
47 Coste

48 Drafting of the manuscript: Misset, Philippart, Cavaillon and Coste

49 Critical revision of the manuscript for important intellectual content: Misset, Philippart,
50 Annane, Moucadel, Cavaillon and Coste

51 Statistical analysis: Coste

52 Obtained funding: Misset, Moucadel

53 Administrative, technical, or material support: Fitting

54

55

56 ^Members of the Combined Approach for The eArly diagnosis of INfection in sepsis

57 (CAPTAIN) study group is provided in the Acknowledgments (collaborators section).

58 **Abstract**

59 **Background**

60 Although sepsis is a life-threatening condition, its heterogeneous presentation likely explains
61 the negative results of most trials on adjunctive therapy. This study in patients with sepsis aimed
62 to identify subgroups with similar immune profiles and their clinical and outcome correlates.

63 **Methods**

64 A secondary analysis used data of a prospective multicenter cohort that included patients with
65 early assessment of sepsis. They were described using Predisposition, Insult, Response, Organ
66 failure sepsis (PIRO) staging system. Thirty-eight circulating biomarkers (27 proteins, 11
67 mRNAs) were assessed at sepsis diagnosis, and their patterns were determined through
68 principal component analysis (PCA). Hierarchical clustering was used to group the patients and
69 *k*-means algorithm was applied to assess the internal validity of the clusters.

70 **Results**

71 Two hundred and three patients were assessed, of median age 64.5 [52.0-77.0] years and SAPS2
72 score 55 [49-61] points. Five main patterns of biomarkers and six clusters of patients (including
73 42%, 21%, 17%, 9%, 5% and 5% of the patients) were evidenced. Clusters were distinguished
74 according to the certainty of the causal infection, inflammation, use of organ support, pro- and
75 anti-inflammatory activity, and adaptive profile markers.

76 **Conclusions**

77 In this cohort of patients with suspected sepsis, we individualized clusters which may be
78 described with criteria used to stage sepsis. As these clusters are based on the patterns of

79 circulating biomarkers, whether they might help to predict treatment responsiveness should be
80 addressed in further studies.

81 **Trial registration**

82 The CAPTAIN study was registered on clinicaltrials.gov on June 22, 2011, # NCT01378169.

83

84 **Introduction**

85 Sepsis is a clinical picture of organ dysfunctions elicited by an infection, and associated with
86 immune dysregulation [1]. Its mortality varies between 25 and 60 % [2, 3]. The intensity of the
87 organ dysfunctions are usually assessed by the SOFA score [4]. Immune dysregulation is
88 complex and not fully deciphered [5], follows from the activation by both pathogen and danger-
89 associated molecular patterns [6], and is associated with a variety of immune pathways
90 including inflammation, compensatory anti-inflammation, and low adaptive profile [7, 8]. It is
91 likely dependent on underlying diseases [9], genetic predisposition [10] and the causal agent of
92 infection [11]. All these dimensions of sepsis are included in the Predisposition, Insult,
93 Response, Organ failure sepsis (PIRO) classification system, a tool proposed in 2001 to
94 characterize and stage sepsis [12]. As numerous trials failed to improve unselected cohorts of
95 patients with sepsis [13, 14], “endotypes” describing patient groups with similar genetic,
96 epigenetic or proteomic pattern, have been proposed [15]. They are expected to help to predict
97 treatment responsiveness and not just differences in prognosis [16]. Endotyping aims at
98 categorizing the different pathways involved [14, 16] in order to select patients as potential
99 targets of specific treatments [17, 18]. Data on endotypes are scarce because the collection of
100 many biomarkers is not available in daily routine [19]. To comply with the recommendations
101 of the surviving sepsis campaign [3], sepsis must be suspected and treated before the infection
102 is confirmed, and biomarkers might help to differentiate patients with bacterial infection from

103 those with other causes of immune dysregulation [20]. A better understanding of subgroups
104 within the heterogeneous host response to infection is important both for a better understanding
105 of the biology of sepsis but also for the next generation of trials of more precise interventions
106 for sepsis.

107 In a multicenter prospective cohort called CAPTAIN that included patients with suspected
108 sepsis for whom circulating proteins or mRNAs from circulating leukocytes were assessed,
109 these biomarkers were not able to discriminate patients with *versus* without a documented
110 causal infection [21]. Then, we hypothesized that a clustering approach may help defining
111 subgroups of similar patients in multidimensional populations.

112 In the present study, using the same cohort of patients with sepsis, we aimed at identifying
113 homogeneous subgroups in terms of circulating biomarkers, and clinical phenotypes and
114 mortality correlates.

115

116 **Methods**

117 **Study design**

118 This study is a secondary analysis of the observational multicenter prospective CAPTAIN study
119 (Combined Approach for The eArly diagnosis of INfection in sepsis) [21]. It was designed and
120 conducted according to STROBE (STrengthening the Reporting of OBServational studies in
121 Epidemiology) guidelines (S1 Table) [22].

122

123 **Ethics and study registration**

124 The protocol was approved by the "Comité de Protection des Personnes Ile de France XI"
125 (#2010-A00908-31-10056) on September 13, 2010 and registered on clinicaltrials.gov
126 (NCT01378169) <https://clinicaltrials.gov/ct2/show/NCT01378169>. According to French national
127 regulations, written consent of the patients was required but waived for the unarousable ones,
128 and obtained if the study still required specific samples when the patient awoke.

129

130 **Setting**

131 Patients were recruited from December 2011 to April 2013 in seven ICUs from five hospitals
132 in Paris area.

133

134 **Participants**

135 Eligible ICU patients were those patients with suspected sepsis. The inclusion criteria were
136 hypothermia (below 36.0°C) or hyperthermia (over 38.0 °C), and at least one criterion of
137 systemic inflammatory response syndrome (SIRS) [12] as soon as the physician considered
138 antibiotic therapy. Other inclusion criteria were age over 18 years, no treatment limitation and
139 no obvious immunosuppression.

140 Demographics, reasons for ICU admission, underlying diseases, simplified acute severity score
141 (SAPS 2) [20], physiological data, Sequential Organ Failure Assessment (SOFA) score [21]
142 and length of organ failure support were collected at admission to the ICU, at inclusion in the
143 study and over the ICU stay. The population characteristics have been published previously
144 [21]. Briefly, 363 patients were screened and 279 included. Based on the data obtained in the
145 72 hours after inclusion, infection could not be ascertained in one third of the patients after

146 adjudication by two investigators who were blind to the biomarkers. The biomarkers were found
147 to discriminate poorly between patients with *versus* without a documented causal infection [21].
148 The present analysis focused on those patients who were still in ICU after day 3, because we
149 wanted to describe their clinical phenotypes during at least the first three days of the syndrome.
150 We excluded those patients having not at least one available value for all the 38 biomarkers
151 collected in the first two days of inclusion because principal component analysis (see below)
152 does not handle with missing data, leading to 203 patients available for analysis. All included
153 patients had a recent increase of the total SOFA score and fulfilled the characteristics of the
154 Sepsis-3 definition [1] despite being included before its publication.

155

156 **Biomarkers assessment**

157 We collected whole blood samples at day 0 and 1 of inclusion to assess 38 biomarkers, reported
158 as potential indicators of infection or mortality during sepsis [23]. They were measured using
159 various techniques: ELISA for galectin, peptidoglycan, visfatin (previously known as pre-B
160 cell colony-enhancing factor (PBEF), soluble B7-H6 (sB7-H6), soluble Triggering Receptor
161 Expressed on Myeloid cells-1 (sTREM-1), soluble urokinase-type Plasminogen Activator
162 Receptor (suPAR) and Pancreatic Stone Protein (PSP); Bioplex technique for C-Reactive
163 Protein (CRP), ferritin, Granulocyte-Colony Stimulating Factor (G-CSF), chemokines: GRO
164 alpha (GRO- α /CXCL1), Monocyte Chemo-attractant Protein 1 (MCP-1/CCL2), Macrophage
165 Inflammatory Protein-1 alpha (MIP-1 α /CCL3), Macrophage Inflammatory Protein-1 beta
166 (MIP-1 β /CCL4), Regulated upon Activation Normal T-cell Expressed and Secreted
167 (RANTES/CCL5), IFN- γ -induced protein 10 (IP-10/CXCL10); Interferon gamma (IFN- γ),
168 Interleukin-1 Receptor antagonist (IL-1Ra), Interleukins 6, 8, 10, 15, and 18 (IL-6, IL-8, IL-10,
169 IL-15 and IL-18), Macrophage-Colony Stimulating Factor (M-CSF), Macrophage Migration

170 Inhibitory Factor (MIF), Matrix Metallo Proteinase-8 (MMP-8), Procalcitonin (PCT), and
171 Tumor Necrosis Factor (TNF); and Real Time quantitative Polymerase Chain Reaction (RT-
172 qPCR) for the following whole blood mRNAs: Cluster of Differentiation 3 δ and 74 (CD3D and
173 CD74), CX3 chemokine receptor 1 (CX3CR1), Human Leucocyte Antigen DR-alpha chain
174 (HLA-DR), High Mobility Group Box 1 protein (HMGB1), IL-1 β , IL-10, Leukocyte
175 Immunoglobulin-Like Receptor subfamily B member 2 (LILRB2), S100 calcium-binding
176 protein A9 (S100A9) and TNF. Of these biomarkers, IL-1Ra, IL-10 and IL-10 mRNA are
177 considered as anti-inflammatory, HLA-DR mRNA, CD74 mRNA, CD3 mRNA and LILRB2
178 mRNA as markers of adaptive immunity, and peptidoglycan as a pathogen-associated
179 molecular pattern. All other above-mentioned biomarkers are considered as mediators or
180 markers of inflammation.

181 Plasma (EDTA) was prepared at reception, aliquoted and frozen. Concentrations of plasma
182 markers were determined at the end of the study by sandwich ELISA or by Multiplex analysis
183 according to manufacturers' recommendations (S2 Table). All the plasma and cell surface
184 markers were purchased from providers, except Soluble B7-H6 (S2 Table). All primers and
185 probe for the RNA markers were designed internally and purchased from Eurogentec. However,
186 primers and probe (different designs) can also be bought from some providers like
187 ThermoFisher. Limits of quantification are provided in S3 Table. For biomarkers whose value
188 was below the lower limit of quantification (LLoQ), we attributed a value of $LLoQ / \sqrt{2}$. For
189 biomarkers whose value was over the upper limit of quantification (ULoQ), we attributed the
190 ULoQ value. The distribution of missing values and determination of the cut-off to create binary
191 variables are provided in S1 Fig. PAXgene[®] blood samples were incubated at room
192 temperature for 2 hours before freezing. RNA was extracted using the PAXgene[®] blood RNA
193 kit (PreAnalytix) and frozen at -80°C. Two hundred nanograms RNA were reverse transcribed
194 (RT) using the SuperScript[®] VILO[™] cDNA Synthesis Kit (Life Technologies) at the end of

195 the study. A RNA calibrator made from PAXgene® samples collected from a pool of healthy
196 volunteers and stimulated *ex vivo* by LPS was used in each RT run. cDNA was then diluted at
197 1/20 and stored at -20°C. Polymerase Chain Reaction (PCR) was performed on a LightCycler
198 instrument using the standard Taqman Fast Advanced Master Mix PCR kit according to the
199 manufacturer's instructions (Roche Molecular Biochemicals). Thermocycling was performed
200 in a final volume of 20 µL containing 0.5 µM of primers and 0.1 µM of probe (see primer and
201 probe designs in S4 Table). PCR was performed with an initial denaturation step of 10 min at
202 95°C, followed by 45 cycles of a touchdown PCR protocol (10 sec at 95°C, 29 sec annealing
203 with 68°C for the first cycle and decrease of 0.5°C for each cycle until reaching 58°C, and 1
204 sec extension at 72°C). The Second Derivative Maximum Method was used by the LightCycler
205 software to automatically determine the crossing point (cycle threshold, Ct). Ct were converted
206 to Calibrated Normalized Relative Quantity (CNRQ) taking HPRT1 and PPIB as reference
207 genes [24].

208

209 **Clinical phenotypes description**

210 We described the patients clinical phenotypes according to the PIRO classification system [12,
211 25] where predispositions (P) related with chronic status and disease, insult (I) with the cause
212 of the suspected sepsis, response (R) with clinical response to this cause, and organ dysfunction
213 (O) with the nature and severity of the organ dysfunctions. P items were age, gender, body mass
214 index (BMI), Mac Cabe score, chronic lung, cardiac, renal or hepatic insufficiencies, diabetes
215 or malignancy. I items were bacterial infection of the lung, abdomen or urinary tract. R items
216 were body temperature, blood lymphocyte and platelet counts, prothrombin time and serum
217 lactates as indicators of inflammatory response, coagulation activation and tissue hypoxia;
218 pneumonia or bacteremia occurring after day 5 of the ICU stay as indicators of immune

219 dysfunction. O items were sequential organ failure assessment (SOFA) and each of its sub-
220 components (respiratory, nervous, cardiovascular, liver, coagulation and kidney dysfunctions)
221 within two days of inclusion. Outcome was defined as the mortality at the end of the ICU stay.
222 Infections were confirmed *a posteriori*, based on criteria which confirm infection as much as
223 possible, either with or without positive cultures. The definition of infection and its causal link
224 with organ dysfunction required medical interpretation [26] and were based on IDSA guidelines
225 [27]. They were adjudicated blindly to the studied biomarkers, by two investigators (FP and
226 BM). They reviewed the patients' records, including clinical history, results of routine
227 morphologic, biological, or microbiological tests, and response to therapies during the days
228 following inclusion. Strains were considered as infecting, colonizing or contaminants. Infection
229 could be considered as present despite the absence of a positive microbiological sample, for
230 example in cases of abscess or pneumonia [27]. When bacteremia was present, it was linked to
231 most probable anatomical focus of infection. Viruses were only searched in case of Influenza
232 suspicion and were classified as non-septic SIRS. Disagreements on classification were
233 resolved after discussion between the two adjudicators.

234

235

236 **Statistical analysis**

237 The determination of classes in numerical taxonomy is generally achieved by cluster analysis
238 of a resemblance matrix, which is a combination of similarities (or distances) between all pairs
239 of objects, *e.g.* patient's biomarkers. Here, such a simple process appeared inadequate because
240 of the large and heterogeneous scales of biomarkers involved. Transformation and reduction of
241 data were necessary. After adequate transformation of the data, we therefore performed a first

242 factor analysis and further used a number of standardized factor scores for each individual as
243 input (in the resemblance matrix) to the clustering method.

244

245 **Data transformation**

246 We determined the maximum value of the blood levels (Cmax) obtained at day 0 and day 1 of
247 inclusion for each biomarker. For all biomarkers, these values were log-transformed to
248 normalize their distribution and further standardized to mean 0 and unit variance.

249

250 **Principal component analysis**

251 To evidence patterns, Cmax of biomarkers were used to build correlation matrices (Pearson
252 coefficients), then studied by principal component analysis (PCA), followed by varimax
253 rotations of retained components. The patterns obtained were uncorrelated linear combinations
254 of normalized and standardized biomarkers, and sorted by decreasing variance of rates
255 explained, whose coefficients, the “loadings”, are interpretable as correlation coefficients
256 between patterns and original biomarkers. These loadings help identify the “nature” or
257 “meaning” of the patterns: “loadings” > 0.40 are usually considered to indicate substantial
258 *correlation*. The number of components-patterns to retain was determined by the Horn and
259 Velicer methods as recommended [28].

260

261 **Clustering**

262 The scores of the patients on the components-patterns retained were selected for cluster
263 analysis. Hierarchical cluster analysis (Ward method) was used to obtain the initial cluster

264 grouping because of the lack of *a priori* knowledge of the number of clusters involved. The
265 number of clusters selected was based on standard statistical criteria (optimal values of R2,
266 pseudo-F, pseudo-t2 and cubic clustering criterion, which all reflect some balance between
267 within and between cluster variances), as recommended [29].

268

269 **Cluster internal validity**

270 Two methods were used to evaluate the stability and the replicability of the hierarchical cluster
271 solution; (1) a *k*-means algorithm: this method does not assume a hierarchical relationship
272 among clusters and allows for relocation of cases throughout the clustering process (reducing
273 the risk of misassignment common to hierarchical cluster method [29]); (2) a subsample
274 analysis: the hierarchical cluster analysis was repeated with a random 50% sample of the initial
275 population to investigate whether subjects clustered similarly when they were distributed in
276 subsamples [30].

277

278 **Description of the clusters**

279 The clusters obtained were finally compared for each individual circulating biomarker and for
280 the clinical variables usually used to describe patients with sepsis. These variables were
281 classified according to the different categories of the PIRO system to facilitate reading and
282 interpretation. For each phenotype qualitatively described, we defined two groups of patients
283 according to its presence or absence. We compared the proportions of patients in each cluster
284 with these phenotypes using non-parametric Fisher exact test. For each phenotype defined with
285 a quantitative value, we assessed correlations of the phenotype with each cluster using a
286 Spearman rank test and we compared the values of each phenotype within each cluster using

287 Kruskal-Wallis statistics. The quantitative values are displayed as median [Q1-Q3] and the
288 qualitative values as n (%). We considered a p-value below 0.05 for statistical significance.
289 SAS 9.4 package was used for all analysis (SAS Inst., Cary, NC, USA).

290

291 **Results**

292 **Patients**

293 Out of 363 patients screened for biomarkers, 279 patients were included in the Captain cohort,
294 but 33 died or were discharged from the ICU before day 3, leading to 246 eligible patients. Of
295 these, 43 had a least one missing value among the 38 biomarkers, leading to 203 evaluable
296 patients (Fig 1). The population characteristics according to the PIRO system, are reported in
297 Table 1. Their median age was 64.5 [52.-77.0] years, median SAPS II score, 55 [49-61] points,
298 and they were included 23 [11-45] hours after ICU. Among them, 189 (77%) were in the “>5-
299 year life expectancy” category of the Mac Cabe score, 116 (47%) had underlying diseases, the
300 suspected infection was confirmed for 171 (70%). After day 5 of ICU, 26 (13%) had acquired
301 pneumonia and 8 (4%) bacteremia. Mechanical ventilation, vasopressors, renal replacement
302 therapy and low-dose steroids were used in 182 (74%), 95 (38%), 19 (8%) and 22 (9%) patients,
303 respectively, and 58 (29%) patients died in the ICU. The levels of each single biomarker for the
304 total cohort have been published previously [21].

305

306 **Fig 1. Flowchart**

307

308 **Table 1 : Characteristics of the patients, organized according to the PIRO system and**
 309 **outcome**

PIRO			n (%) or med [Q1-Q3]
category		Variable	
Predisposition		Age (years)	64.5 [52.0-77.0]
		Male sex	160 (65)
		BMI (kg/m ²)	25.7 [21.6-30.0]
		Mc Cabe score, % prediction > 5 years	189 (77)
		COPD	46 (19)
		Cardiac insufficiency	25 (10)
		Diabetes	53 (22)
		Chronic renal insufficiency	22 (9)
		Solid tumor	34 (14)
		Hematologic malignancy	4 (2)
		Chronic hepatic insufficiency	18 (7)
		Any prior disease	115 (47)
Insult	At inclusion	Infection due to Gram positive bacteria	76 (31)
		Infection due to Gram negative bacteria	124 (50)
		Pneumonia	123 (50)
		Intra-abdominal infection	14 (6)
		Urinary tract infection	19 (8)
		Confirmed infection	171 (70)
Response	At inclusion	Temperature (°C)	38.2 [37.5-38.8]
		Lymphocyte count (/mm ³)	905 [640-1390]
		Blood platelets (10 ³ /mm ³)	187 [134-268]

		Prothrombin time (%)	66 [59-82]
		Blood lactates (meq/L)	1.6 [1.0-2.3]
		PaO2 (mmHg)	88 [73-145]
		FiO2 (%)	40 [30-60]
		PaCO2 (mmHg)	39 [34-46]
		Serum creatinin (μmol/L)	92 [71-182]
		Blood hematocrit (%)	32.2 [28.5-38.6]
		White blood cell count (/mm ³)	13,200 [9,480-18,600]
		Respiratory rate (/min)	26 [22-33]
		Heart rate (/min)	106 [95-125]
		Mean arterial pressure (mmHg)	73 [63-94]
		Urinary output (L/24h)	1.30 [0.81-2.00]
		SAPS II score (points)	55 [49-61]
	After ICU day 5	ICU acquired pneumonia	26 (13)
		ICU acquired bacteremia	8 (4)
Organ failure	At inclusion	Total SOFA score (points)	6 [3-9]
		Respiratory SOFA score (points)	2 [0-3]
		Neurological SOFA score (points)	0 [0-2]
		Circulatory SOFA score (points)	0 [0-1]
		Hepatic SOFA score (points)	0 [0-0]
		Coagulation SOFA score (points)	0 [0-1]
		Kidney SOFA score (points)	0 [0-2]
	During the ICU stay		182 (74)
		Mechanical ventilation	
		Non invasive ventilation	15 (6)

Vaso-active drugs	95 (38)
Renal replacement therapy	19 (8)
Low doses steroid therapy	22 (9)

310

311 **Patterns of circulating biomarkers and clusters of patients**

312 PCA of the 38 biomarkers provided five main components-patterns, which explained 30.7%,
313 9.7%, 8.0%, 5.5% and 4.7% (total 58.6%) of the variance, respectively. The biomarkers with a
314 high loading ($|\text{loading}| > 0.40$) are displayed in S5 Table. Pattern #1 gathers circulating
315 biomarkers and mRNAs linked to both pro- and anti-inflammatory response and to altered
316 immunity. Pattern #2 gathers only circulating biomarkers associated with pro- and anti-
317 inflammatory response, and shares numerous cytokines and chemokines with pattern #1 (i.e.,
318 IL-1Ra, IL-6, IL-8; MCP-1, G-CSF, GM-CSF, and MIP-1 β). Pattern #3 gathers only mRNA
319 markers linked to both pro- and anti-inflammatory response. Pattern #4 displays specific
320 biomarkers not shared with any other clusters (i.e., Galectin-9, SuPAR, MIF, and Ferritin).
321 Similarly, pattern #5 has its specific biomarkers (RANTES, sTREM-1). These 5 patterns
322 allowed to build 6 clusters of patients with homogeneous biological profiles. The S6 Table
323 displays the scores of the 5 patterns within each cluster of patients. The use of a *k*-means
324 algorithm (non-hierarchical method), with the number of clusters set to 6, led to similar
325 clustering as with the Ward method, with satisfactory agreement (Cramer's $V = 0.63$). Similar
326 clustering solutions were found with analysis of a random 50% of the sample (Cramer's $V =$
327 0.59). These results support the robustness of the six clusters.

328 The Table 2 shows the criteria of the PIRO profile of sepsis. "Predisposition" items of the PIRO
329 system differed moderately across clusters. Among the "insult", "response", and "organ failure"
330 categories, the most different items between clusters were infection certainty, blood lactate

331 levels, serum creatinine levels, urinary output, survival, circulatory and renal SOFA sub-scores,
332 and use of renal replacement therapy, vaso-pressors and steroids. The Table 3 shows that the
333 levels of the 38 individual circulating biomarkers differed markedly across the clusters and that
334 biomarkers of the same category (pro-inflammatory, anti-inflammatory, adaptive immunity)
335 displayed consistent values within each cluster. In these tables, the color code (from dark red
336 to high level, to dark blue for low level) illustrates the differences within each item of the
337 clusters. Based on these comparisons, six clusters can be distinguished according to the
338 certainty level of the causal infection, the existence of inflammation, use of renal and/or
339 hemodynamic support, pro-inflammatory and anti-inflammatory activity, and markers of
340 adaptive profile. Four clusters (clusters #2, #3, #4 and #6) were associated with high mortality
341 (> 30%) and a low adaptive profile (Table 2), and represented 53% of the cohort (Fig 2).
342 Clusters #2 and #6 exhibited both high levels of inflammatory and anti-inflammatory mediators,
343 but differed with regards to CRP and ferritin, #3 displayed anti-inflammatory mediators at low
344 level, and #4 displayed both inflammatory and anti-inflammatory mediators at low level (Fig
345 2).

346

347 **Fig 2. Distribution and description of the clusters in the cohort.** All the clusters display
348 different characteristics in terms of parameters of the PIRO system, of circulating biomarkers
349 and outcome.

350 Table 2 : Value of each clinical criterion of the PIRO profile of sepsis and outcome in each cluster

		Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		
PIRO category		86		43		34		18		11		11		
Variable	unit	med or %	[Q1-Q3]	med or %	[Q1-Q3]	med or %	[Q1-Q3]	med or %	[Q1-Q3]	med or %	[Q1-Q3]	med or %	[Q1-Q3]	
Predisposition														<i>p value</i>
Age	years	64	[52-74]	63	[52-77]	70	[58-78]	54	[38-69]	74	[49-84]	77	[63-87]	0.02
BMI	kg/m ²	26.1	[21.4-31.4]	24.8	[20.9-27.8]	26.3	[23.9-30.4]	24.5	[20.2-24.5]	27.2	[22.8-37.2]	24.6	[22.7-28.8]	0.35
Male sex	%	62%		77%		67%		83%		45%		64%		0.18
Mc Cabe score (% > 5 years)	%	72%		84%		73%		83%		91%		64%		0.19
COPD	%	24%		28%		21%		6%		9%		9%		0.36
Cardiac insufficiency	%	9%		12%		18%		6%		9%		0%		0.66
Diabetes	%	20%		12%		29%		17%		28%		28%		0.42
Chronic renal insufficiency	%	5%		5%		26%		0%		0%		27%		0.001
Solid tumor	%	15%		12%		18%		6%		18%		18%		0.82
Hematologic malignancy	%	15%		20%		17%		0%		0%		0%		0.83
Chronic hepatic insufficiency	%	10%		2%		9%		11%		0%		9%		0.51
Any prior disease	%	56%		46%		79%		39%		54%		64%		0.03
Insult														
at inclusion	GPC infection	%	7%		16%		12%		11%		9%		9%	0.65
	GNB infection	%	37%		65%		65%		67%		54%		45%	0.01
	Pneumonia	%	44%		56%		56%		67%		45%		18%	0.2
	Intra-abdominal infection	%	2%		12%		3%		11%		9%		9%	0.12
	UTI	%	3%		12%		18%		0%		0%		27%	0.01

Confirmed infection as a cause	%	53%		88%		85%		83%		64%		54%		0.0001	
Response															
at inclusion	Temperature	°C	38.2	[37.5-38.7]	37.9	[37.4-38.7]	38.3	[37.8-38.9]	39,0	[38.0-39.3]	38.6	[38.2-40.0]	38.2	[37.5-38.9]	0.04
	Lymphocyte count	/mm ³	960	[690-1,540]	820	[500-1,190]	910	[525-1,200]	990	[860-1,200]	700	[540-1,390]	1,085	[685-1,695]	0.29
	Blood platelets	10 ³ /mm ³	170	[129-252]	171	[131-215]	222	[145-306]	252	[203-379]	225	[145-293]	95	[72-156]	0.008
	Prothrombin time	%	71	[62-84]	69	[57-80]	69	[58-78]	74	[61-82]	83	[76-87]	35	[28-40]	0.001
	Blood lactates	meq/L	1.35	[0.90-1.80]	2,00	[1.40-2.70]	1.55	[1.35-2.80]	1.35	[0.90-2.30]	1.45	[0.80-2.30]	2.9	[2.30-4.50]	0.0002
	PaO2	mmHg	94	[74-123]	79	[62-123]	85	[74-106]	86	[77-100]	71	[57-125]	95	[81-168]	0.34
	FiO2	%	35	[30-50]	47	[30-80]	40	[30-55]	40	[30-50]	30	[24-39]	50	[40-70]	0.05
	PaCO2	mmHg	41	[34-47]	40	[34-48]	39	[34-49]	39	[34-42]	31	[27-36]	36	[31-47]	0.03
	Serum creatinin	μmol/L	90	[70-150]	120	[75-195]	195	[90-280]	85	[80-140]	140	[80-220]	260	[220-520]	0.0007
	Blood hematocrit	%	34.2	[29.0-39.4]	35.7	[29.9-40.0]	29.6	[27.2-33.0]	28.4	[25.1-30.4]	32.8	[28.5-38.1]	31.2	[29.2-38.5]	0.001
	White blood cell count	/mm ³	12,380	[8,700-15,450]	13,450	[10,220-18,700]	14,350	[11,300-19,300]	15,600	[9,200-18,800]	13,900	[8,200-21,700]	14,000	[10,140-40,300]	0.32
	Respiratory rate	/min	25	[21-31]	25	[22-30]	29	[23-33]	30	[27-33]	37	[28-40]	26	[21-34]	0.006
	Heart rate	/min	100	[91-111]	112	[98-128]	115	[98-139]	112	[102-121]	111	[99-116]	110	[98-148]	0.03
	Mean arterial pressure	mmHg	75	[67-89]	67	[60-97]	71	[59-98]	82	[70-101]	81	[66-113]	59	[46-65]	0.009
	Urinary output	L/24h	1,200	[700-1,700]	1,442	[930-1,990]	1,075	[400-2,075]	1,900	[1,500-2,275]	2,125	[1,750-2,300]	85	[5-400]	0.0001
	SAPS II score	points	55	[49-61]	52	[49-61]	55	[49-60]	51	[50-57]	53	[44-59]	59	[56-61]	0.34
over the ICU stay	acquired pneumonia	%	8%		14%		12%		22%		9%		9%		0.59
	acquired bacteremia	%	2%		2%		3%		6%		9%		18%		0.12
Organ failure															
over the ICU stay	Total SOFA score	points	5	[2-5]	6	[4-10]	5	[4-9]	3	[2-7]	7	[4-8]	11	[10-14]	0.0001
	Respiratory SOFA score	points	2	[1-3]	2	[2-3]	2	[1-3]	2	[2-3]	2	[1-3]	2	[2-3]	0.2
	Neurological SOFA score	points	0	[0-1]	0	[0-3]	0	[0-1]	0	[0-0]	1	[0-4]	0	[0-2]	0.24
	Circulatory SOFA score	points	0	[0-0]	0	[0-3]	0	[0-3]	0	[0-0]	0	[0-0]	4	[1-4]	0.0001

Hepatic SOFA score	<i>points</i>	0	[0-0]	0	[0-0]	0	[0-0]	0	[0-0]	0	[0-2]	2	[0-2]	0.13
Coagulation SOFA score	<i>points</i>	0	[0-1]	0	[0-1]	0	[0-0]	0	[0-0]	0	[0-1]	1	[3-4]	0.1
Kidney SOFA score	<i>points</i>	0	[0-1]	0	[0-2]	1	[1-2]	0	[0-0]	1	[1-4]	4	[3-4]	0.0001
Mechanical ventilation	%	65%		86%		76%		94%		54%		78%		0.02
Non invasive ventilation	%	7%		2%		3%		0%		27%		11%		0.06
Vaso-active drugs	%	30%		58%		47%		12%		9%		73%		0.0001
Renal replacement therapy	%	2%		2%		24%		0%		0%		45%		0.0001
Low doses steroid therapy	%	1%		19%		15%		0%		9%		27%		0.0005
Outcome														
Death at ICU discharge	%	15%		33%		33%		44%		18%		82%		0.0001

351 The clinical criteria are sorted according to the PIRO classification system. In each row, the dark red color indicates the highest value and dark

352 blue color indicates the lowest value, for those variables which are the most significantly different across the clusters.

354 **Table 3 : Value of each circulating biomarker of sepsis in each cluster.**

		Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		
Biomarker category		n=86		n=43		n=34		n=18		n=11		n=11		
Variable	unit	med	[Q1-Q3]	med	[Q1-Q3]	med	[Q1-Q3]	med	[Q1-Q3]	med	[Q1-Q3]	med	[Q1-Q3]	<i>p value</i>
Inflammatory mediators or biomarkers														
<i>Cytokines</i>														
<i>TNF-α</i>	<i>ng/L</i>	10	[10-10]	10	[10-10]	10	[10-10]	10	[10-10]	10	[10-10]	10	[10-220]	0.0001
<i>TNF-α RNA</i>	<i>CNRQ</i>	0.015	[0.012-0.018]	0.017	[0.011-0.022]	0.013	[0.007-0.019]	0.008	[0.006-0.010]	0.018	[0.014-0.023]	0.009	[0.004-0.016]	0.0001
<i>IL-1β RNA</i>	<i>CNRQ</i>	0,008	[0.006-0.011]	0.008	[0.005-0.008]	0.006	[0.005-0.013]	0.004	[0.003-0.005]	0.013	[0.009-0.018]	0.004	[0.002-0.009]	0.0001
<i>IL-18</i>	<i>ng/L</i>	45.2	[16.1-80.7]	53.8	[35.9-97.4]	77.4	[38.7-197.1]	46.6	[29.2-72.3]	123.5	[123.5-368.0]	365.0	[333.7-924.3]	0.0001
<i>IL-15</i>	<i>ng/L</i>	2.5	[2.50-2.50]	2.5	[2.50-2.50]	2.5	[2.50-2.50]	2.5	[2.50-2.50]	48.4	[2.5-90.9]	25.7	[6.7-107.4]	0.0001
<i>IL-6</i>	<i>ng/L</i>	45	[15-149]	1,298	[327-5,344]	119	[44-342]	59	[24-485]	873	[193-2107]	8,333	[1,259-85,509]	0.0001
<i>GM-CSF</i>	<i>ng/L</i>	9.8	[4.0-28.9]	609.1	[64.0-1,754.9]	13.5	[4.0-53.9]	4.0	[4.0-43.2]	18.2	[4.0-575.8]	411.5	[293.0-13,027.0]	0.0001
<i>Chemokines and receptors</i>														
<i>MCP-1</i>	<i>ng/L</i>	49	[5-85]	173	[98-927]	54	[5-84]	75	[44-192]	240	[48-379]	342	[127-11,415]	0.0001
<i>MIF</i>	<i>μg/L</i>	8.5	[4.4-14.2]	5.1	[2.7-7.5]	15.5	[6.7-22.2]	11.7	[7.7-23.9]	12.5	[8.7-24.3]	42.0	[23.3-48.6]	0.0001
<i>Rantes CCL5</i>	<i>μg/L</i>	14.2	[9.6-14.2]	10.4	[7.3-16.3]	12.1	[8.1-15.7]	15.8	[10.4-19.1]	57.6	[22.4-115.1]	6.6	[4.1-31.6]	0.0001
<i>IP-10</i>	<i>ng/L</i>	165	[59-363]	284	[117-552]	505	[216-904]	282	[185-800]	6,034	[1,631-21,478]	836	[580-60,981]	0.0001
<i>IL-8</i>	<i>ng/L</i>	12.5	[4.0-23.8]	73.4	[31.6-337.2]	26.0	[213.5-47.9]	14.3	[9.4-26.0]	24.4	[4.0-137.0]	494.0	[187.0-1,494.0]	0.0001
<i>MIP-1β</i>	<i>ng/L</i>	39	[20-60]	88	[40-243]	62	[24-132]	53	[37-73]	305	[70-450]	149	[119-1,229]	0.0001
<i>CX3CR1 RNA</i>	<i>CNRQ</i>	14.1	[10.9-18.9]	7.2	[4.0-12.5]	9.5	[4.7-16.0]	5.7	[4.1-7.0]	10.4	[6.8-16.4]	4.2	[1.7-7.6]	0.0001
<i>Others</i>														

<i>C reactive protein</i>	mg/L	146	[68-203]	296	[241-331]	265	[194-346]	235	[154-318]	310	[265-454]	208	[153-249]	0.0001
<i>Procalcitonin</i>	µg/L	1.3	[1.2-1.3]	2.7	[1.4-6.3]	2.2	[1.6-4.5]	1.8	[1.5-2.2]	1.6	[1.4-5.3]	7.2	[2.8-18.9]	0.0001
<i>SuPAR</i>	ng/L	6.14	[4.35-8.34]	7.27	[5.41-11.49]	15.33	[12.89-18.27]	11.58	[6.4-14.2]	10.4	[7.0-14.1]	20.6	[11.8-35.0]	0.0001
<i>Visfatin</i>	µg/L	5.56	[4.52-7.00]	4.98	[4.22-5.85]	5.47	[4.42-6.75]	4.45	[4.07-5.34]	5.17	[4.50-6.03]	38.46	[10.31-140.11]	0.0001
<i>PSP</i>	ng/L	69	[45-139]	325	[139-613]	341	[145-1,121]	144	[111-321]	115	[66-257]	780	[349-1,538]	0.0001
<i>sB7-H6</i>	ng/L	24.6	[16.5-27.7]	27.5	[18.3-31.4]	23.9	[15.9-29.1]	19.2	[11.2-26.5]	28.2	[17.7-35.8]	0.0	[0.0-0.0]	0.0004
<i>MMP-8</i>	µg/L	20	[10-45]	77	[47-167]	85	[30-192]	46	[20-110]	74	[64-197]	198	[93-212]	0.0001
<i>sTREM-1</i>	µg/L	2.2	[1.3-4.0]	4.1	[2.0-5.5]	4.7	[3.1-6.5]	1.0	[0.6-2.6]	2.3	[1.4-4.3]	4.3	[3.3-8.6]	0.0001
<i>HMGB1 RNA</i>	CNRQ	3.64	[3.14-4.22]	3.58	[3.03-4.05]	3.89	[3.19-3.89]	3.22	[2.47-3.61]	3.35	[2.53-4.03]	3.81	[3.00-4.45]	0.03
<i>Ferritin</i>	µg/L	1	[1-2]	2	[1-2]	211	[117-434]	248	[165-501]	273	[178-516]	960	[732-1,075]	0.0001
<i>Galectin 9</i>	µg/L	5.7	[3.4-7.4]	5.8	[3.2-8.5]	12.5	[9.4-17.6]	8.2	[5.0-12.8]	8.1	[6.6-12.5]	25.7	[19.2-35.9]	0.0001
<i>S100A9 RNA</i>	CNRQ	15.4	[11.4-15.4]	28.2	[21.7-36.5]	24.1	[17.0-28.3]	14.6	[9.7-20.3]	24.3	[14.1-32.7]	30.2	[18.9-37.4]	0.0001
Anti-inflammatory mediators														
<i>IL-1Ra</i>	CNRQ	15.0	[15.0-15.0]	18.3	[15.0-620.4]	15.0	[15.0-16.7]	15.0	[15.0-15.0]	215.4	[15.0-632.3]	3,174.8	[927.0-5,062.0]	0.0001
<i>IL-10</i>	ng/L	4.0	[4.0-4.0]	4.0	[4.0-10.9]	4.0	[4.0-4.0]	4.0	[4.0-4.0]	4.0	[4.0-4.0]	107.9	[51.3-321.7]	0.0001
<i>IL-10 RNA</i>	CNRQ	0.17	[0.12-0.24]	0.60	[0.41-1.13]	0.23	[0.12-0.40]	0.09	[0.07-0.12]	0.41	[0.27-0.78]	0.37	[0.19-0.57]	0.0001
Adaptive immunity														
<i>HLA-DR RNA</i>	CNRQ	0.34	[0.27-0.42]	0.15	[0.10-0.20]	0.24	[0.11-0.41]	0.18	[0.15-0.21]	0.29	[0.27-0.42]	0.28	[0.04-0.36]	0.0001
<i>CD74 RNA</i>	CNRQ	0.59	[0.46-0.71]	0.28	[0.18-0.40]	0.36	[0.19-0.63]	0.27	[0.27-0.32]	0.59	[0.45-0.72]	0.26	[0.08-0.53]	0.0001
<i>LILRB2 RNA</i>	CNRQ	0.83	[0.67-0.99]	0.93	[0.82-1.28]	0.76	[0.68-0.94]	0.42	[0.35-0.54]	1.12	[0.71-1.32]	0.79	[0.48-1.09]	0.0001
<i>CD3 RNA</i>	CNRQ	1.18	[0.87-1.51]	0.56	[0.34-0.74]	0.49	[0.33-1.06]	0.56	[0.44-1.12]	0.79	[0.49-1.02]	0.21	[0.20-0.53]	0.0001
Pathogen associated molecular patterns														
<i>Peptidoglycan</i>	µg/L	2.19	[0.75-3.31]	2.79	[1.63-3.79]	2.63	[1.75-3.90]	2.64	[2.01-3.53]	4.44	[2.76-6.08]	1.89	[0.75-2.71]	0.001

355 The individual biomarkers are sorted according to their role in inflammation, anti-inflammation or adaptive immune profile. In each row, the dark

356 red color indicates the highest value and dark blue color indicates the lowest value. CNRQ = Calibrated Normalized Relative Quantity

357 **Discussion**

358 In a prospective cohort of ICU patients suspected of sepsis, through the levels of circulating
359 biomarkers indicative of pro-inflammation, anti-inflammation or adaptive immunity and the
360 use of unsupervised statistical approaches, we individualized six different clusters of patients
361 with homogeneous profiles regarding sepsis clinical staging. These clusters presented with
362 different immune and clinical profiles, making them potential targets for individualized
363 therapies.

364 Sepsis is a life-threatening condition elicited by various infectious conditions, with a
365 heterogeneous presentation, and an outcome impacted by both the pathogen and host
366 characteristics [1]. This phenotypic polymorphism led to the proposal of the PIRO classification
367 and staging system in 2001 to help individualize future therapies [12]. A better understanding of
368 subgroups within the heterogeneous host response to infection is important both for a better
369 understanding of the biology of sepsis but also for the next generation of trials of more precise
370 interventions for sepsis.

371 Given the heterogeneity of both clinical and immune presentation of sepsis and the multiple
372 failures of trials in unselected populations [13], a cluster approach has been used by several
373 authors [31]. In these studies, clustering was based either on phenotypes to describe different
374 clinical profiles without addressing immune mechanisms of sepsis, and mostly provide
375 differences in prognosis, or on endotypes, deriving subclasses from genome-wide expression
376 profiling [31]. The latter option, as endotypes are markers of pathophysiological pathways, may
377 help to predict treatment responsiveness [16]. Wong *et al.* identified three pediatric septic shock
378 subclasses named A, B and C [32]. Subclass A displayed a depressed expression of adaptive
379 immune system, of glucocorticoid signaling and zinc-related biology and associated with higher
380 severity and mortality. Scicluna *et al.* identified four sepsis subclasses named Mars 1 to 4 [15].

381 Subclass Mars1, with a higher mortality, displayed reduced expression of genes involved in
382 innate and immune functions. Subclass Mars3, with a better survival, displayed increased
383 expression of adaptive immune or T-cell functions. Davenport *et al.* identified two sepsis
384 subclasses named SRS1 and SRS2 [33], with SRS1 characterized by a higher mortality and
385 decreased expression of immune functions. Finally, based on gene activation profiles, Sweeney
386 *et al.* identified three sepsis subclasses named inflammopathic, adaptive and coagulopathic
387 [34]. The adaptive subgroup was associated with lower severity and mortality, and the
388 coagulopathic subgroup with higher mortality and clinical coagulopathy. In our study, we also
389 derived clusters from endotypes, but these were assessed with biomarkers made of molecules
390 previously documented to play a role the pathophysiology of sepsis. We used unsupervised
391 statistical approaches to set up clusters, because they explore data without *a priori* classification
392 [35]: principal component analysis, to determine different patterns, hierarchical clustering to
393 group the patients and *k*-means algorithm to assess the internal validity of the clusters.

394 In our cohort, two clusters (#1 and #5) had a low level of organ dysfunction and mortality.
395 Clusters #2 and #3 displayed high level of infection certainty and inflammation, and differed
396 by their anti-inflammatory status, consistent with the concept of compensatory anti-
397 inflammation and its heterogeneity [8]. Cluster #4 displayed a low level of innate response
398 despite high severity and high ferritin levels. Lastly, two clusters, #5 and #6, were associated
399 with very specific phenotypes, one (#5) with pro and anti-inflammatory high-level profile
400 despite low mortality, and the second one (#6) with high immune alteration and particularly
401 high level of ferritin.

402 Half our cohort belongs to clusters associated with a high mortality rate. They are characterized
403 by their low adaptive profile at sepsis diagnosis. They differ between themselves by their
404 respective levels of certainty of the causal infection, of CRP and ferritin levels, of renal and
405 hemodynamic level of support, and of pro-inflammatory and anti-inflammatory activities. Each

406 of these characteristics may be available at bedside in parallel of the assessment of infection,
407 organ dysfunction, pro- (for example TNF α and/or IL-18) and anti-inflammatory (for example
408 IL-1Ra and/or IL-10) cytokines, and markers of adaptive function (for example HLA-DR).
409 These criteria are relatively simple and should be validated in external cohorts before they can
410 be used as inclusion criteria in prospective trials.

411 Our study has limitations. First, several biomarkers, including some recently described, were
412 not assessed in this cohort, and should be assessed in similar conditions. Especially, biomarkers
413 more specific of endothelial dysfunction or coagulation activation were underrepresented in our
414 panel. Second, although at risk of sepsis, several categories of patients were not included in the
415 cohort, particularly those with prior immune suppression, whose innate and adaptive responses
416 are likely different from the immunocompetent patients. These patients should be investigated
417 using a similar approach. Third, while the generalizability of our study may be reinforced by
418 its multicenter design and the use of internal validity assessments, we did not perform external
419 validation in a separate cohort. This is particularly important for the groups with small numbers
420 of patients in our cohort.

421

422 **Conclusion**

423 In a prospective cohort of ICU patients with suspected sepsis, we individualized clusters of
424 patients which may be described with criteria commonly used to stage sepsis in routine practice.
425 As these clusters are based on the patterns of circulating biomarkers, whether they might help
426 to predict treatment responsiveness should be addressed in further studies.

427

428

429

430 **Acknowledgements**

431 We sincerely thank Drs Anne-Françoise Rousseau, Bernard Lambermont and François Jouret
432 for their critical review, Céline Féger, MD (EMIBiotech) and Séverine Marck for their editorial
433 support for this manuscript.

434 **Collaborators**

435 Members of the Combined Approach for The eArly diagnosis of INfection in sepsis
436 (CAPTAIN) study group:

437 Sébastien JACQMIN, Didier JOURNOIS, Alix LAGRANGE, Gabrielle PINOT de
438 VILLECHENON (*Hôpital Européen Georges Pompidou, AP – HP - Université Paris*
439 *Descartes, Service d’Anesthésie-Réanimation*), Nadia AISSAOUI, Jean-Luc DIEHL,
440 Emmanuel GUEROT, Marion VENOT (*Hôpital Européen Georges Pompidou, AP – HP -*
441 *Université Paris Descartes, Service de Réanimation Médicale*), Olfâ HAMZAOUI, Dominique
442 PRAT, Benjamin SZTRYMF (*Hôpital Antoine Béchère, AP – HP - Université Paris Sud,*
443 *Clamart*), Djillali ANNANE, Virginie MAXIME, Andrea POLITO (*Hôpital Raymond*
444 *Poincaré, AP – HP - Université Paris Ile de France Ouest , Service de Réanimation Médico-*
445 *chirurgicale*), Belaïd BOUHEMAD, Cédric BRUEL, Frédéric ETHUIN, Julien FOURNIER,
446 Maïté GARROUSTE-ORGEAS, Charles GREGOIRE, Nicolas LAU, Benoît MISSET,
447 François PHILIPPART (*Groupe Hospitalier Paris Saint Joseph, Service de Réanimation*),
448 Jean-Pierre BEDOS, Pierrick CROSNIER, Virginie LAURENT, Sybille MERCERON
449 (*Hôpital André Mignot, Versailles, Service de Réanimation médico-chirurgicale*), Elsa
450 BOURNAUD, Laurence LECOMTE, Jean-Marc TRELUYER, (*Hôpital Cochin, AP-HP -*
451 *Université Paris Descartes, Unité de Recherche Clinique*), Alexandre PACHOT, Javier

452 YUGUEROS-MARCOS, Laurent ESTEVE, Sophie BLEIN, Virginie MOUCADEL
453 (bioMérieux, Lyon & Grenoble), Myriam BEN BOUTIEB, Alexandra ROUQUETTE, Joël
454 COSTE (*Hôpital Cochin, AP-HP - Université de Paris, Unité de Biostatistique et*
455 *d’Epidémiologie*), Minou ADIB-CONGUY, Jean-Marc CAVAILLON, Catherine FITTING,
456 Marianna PARLATO, Virginie PUCHOIS, Fernando SOUZA-FONSECA-GUIMARAES
457 (Institut Pasteur, Paris, Unit Cytokines & Inflammation).

458

459 **References**

460 1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The
461 Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*.
462 2016;315:801–810.

463 2. Cecconi M, Evans L, Levy M, Rhodes A. Sepsis and septic shock. *Lancet*. 2018;392:75–87.

464 3. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. Surviving Sepsis
465 Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016.
466 *Intensive Care Med*. 2017;43:304–377.

467 4. Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, et al. Use of the
468 SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results
469 of a multicenter, prospective study. Working group on “sepsis-related problems” of the
470 European Society of Intensive Care Medicine. *Crit Care Med*. 1998;26:1793–1800.

471 5. Rubio I, Osuchowski MF, Shankar-Hari M, Skirecki T, Winkler MS, Lachmann G, et al.
472 Current gaps in sepsis immunology: new opportunities for translational research. *Lancet Infect*
473 *Dis*. 2019;19:e422–e436.

- 474 6. Matzinger P. Tolerance, danger, and the extended family. *Annu Rev Immunol.* 1994;12:991–
475 1045.
- 476 7. Hotchkiss RS, Monneret G, Payen D. Sepsis-induced immunosuppression: from cellular
477 dysfunctions to immunotherapy. *Nat Rev Immunol.* 2013;13:862–874.
- 478 8. Adib-Conquy M, Cavaillon J-M. Compensatory anti-inflammatory response syndrome.
479 *Thromb Haemost.* 2009;101:36–47.
- 480 9. Angus DC, Burgner D, Wunderink R, Mira J-P, Gerlach H, Wiedermann CJ, et al. The PIRO
481 concept: P is for predisposition. *Crit Care.* 2003;7:248–251.
- 482 10. Toubiana J, Courtine E, Pène F, Viallon V, Asfar P, Daubin C, et al. IRAK1 functional
483 genetic variant affects severity of septic shock. *Crit Care Med.* 2010;38:2287–2294.
- 484 11. Bedos J-P, Varon E, Porcher R, Asfar P, Le Tulzo Y, Megarbane B, et al. Host-pathogen
485 interactions and prognosis of critically ill immunocompetent patients with pneumococcal
486 pneumonia: the nationwide prospective observational STREPTOGENE study. *Intensive Care*
487 *Med.* 2018;44:2162–2173.
- 488 12. Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, et al. 2001
489 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. *Intensive Care*
490 *Med.* 2003;29:530–538.
- 491 13. Gotts JE, Matthay MA. Sepsis: pathophysiology and clinical management. *BMJ.*
492 2016;353:i1585.
- 493 14. Cavaillon J-M, Singer M, Skirecki T. Sepsis therapies: learning from 30 years of failure of
494 translational research to propose new leads. *EMBO Mol Med.* 2020;12:e10128.

- 495 15. Scicluna BP, van Vught LA, Zwinderman AH, Wiewel MA, Davenport EE, Burnham KL,
496 et al. Classification of patients with sepsis according to blood genomic endotype: a prospective
497 cohort study. *Lancet Respir Med.* 2017;5:816–826.
- 498 16. Prescott HC, Calfee CS, Thompson BT, Angus DC, Liu VX. Toward Smarter Lumping and
499 Smarter Splitting: Rethinking Strategies for Sepsis and Acute Respiratory Distress Syndrome
500 Clinical Trial Design. *Am J Respir Crit Care Med.* 2016;194:147–155.
- 501 17. Vincent J-L, Francois B, Zabolotskikh I, Daga MK, Lascarrou J-B, Kirov MY, et al. Effect
502 of a Recombinant Human Soluble Thrombomodulin on Mortality in Patients With Sepsis-
503 Associated Coagulopathy: The SCARLET Randomized Clinical Trial. *JAMA.* 2019;321:1993–
504 2002.
- 505 18. Francois B, Jeannet R, Daix T, Walton AH, Shotwell MS, Unsinger J, et al. Interleukin-7
506 restores lymphocytes in septic shock: the IRIS-7 randomized clinical trial. *JCI Insight.*
507 2018;3:e98960.
- 508 19. Dupuy A-M, Philippart F, Péan Y, Lasocki S, Charles P-E, Chalumeau M, et al. Role of
509 biomarkers in the management of antibiotic therapy: an expert panel review: I - currently
510 available biomarkers for clinical use in acute infections. *Ann Intensive Care.* 2013;3:22.
- 511 20. Miller RR, Lopansri BK, Burke JP, Levy M, Opal S, Rothman RE, et al. Validation of a
512 Host Response Assay, SeptiCyt[®] LAB, for Discriminating Sepsis from Systemic Inflammatory
513 Response Syndrome in the ICU. *Am J Respir Crit Care Med.* 2018;198:903–913.
- 514 21. Parlato M, Philippart F, Rouquette A, Moucadel V, Puchois V, Blein S, et al. Circulating
515 biomarkers may be unable to detect infection at the early phase of sepsis in ICU patients: the
516 CAPTAIN prospective multicenter cohort study. *Intensive Care Medicine.* 2018;44:1061–
517 1070.

- 518 22. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The
519 Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement:
520 guidelines for reporting observational studies. *Ann Intern Med.* 2007;147:573–577.
- 521 23. Parlato M, Cavaillon J-M. Host response biomarkers in the diagnosis of sepsis: a general
522 overview. *Methods Mol Biol.* 2015;1237:149–211.
- 523 24. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative
524 quantification framework and software for management and automated analysis of real-time
525 quantitative PCR data. *Genome Biol.* 2007;8:R19.
- 526 25. Marshall JC. The PIRO (predisposition, insult, response, organ dysfunction) model: toward
527 a staging system for acute illness. *Virulence.* 2014;5:27–35.
- 528 26. Cohen J, Vincent J-L, Adhikari NKJ, Machado FR, Angus DC, Calandra T, et al. Sepsis: a
529 roadmap for future research. *Lancet Infect Dis.* 2015;15:581–614.
- 530 27. IDSA Practice Guidelines. Available from:
531 [http://www.idsociety.org/PracticeGuidelines/?q=&ref=taxonomy_map%3B\(%22%2FIDSA%](http://www.idsociety.org/PracticeGuidelines/?q=&ref=taxonomy_map%3B(%22%2FIDSA%2FSite+Map%2FGuidelines%2FPatient+Care%2FIDSA+Practice+Guidelines%2FInfections+By+Organ+System-81567%2FLower%2FUpper+Respiratory%22)%3BOrgan+System)
532 [2FSite+Map%2FGuidelines%2FPatient+Care%2FIDSA+Practice+Guidelines%2FInfections+](http://www.idsociety.org/PracticeGuidelines/?q=&ref=taxonomy_map%3B(%22%2FIDSA%2FSite+Map%2FGuidelines%2FPatient+Care%2FIDSA+Practice+Guidelines%2FInfections+By+Organ+System-81567%2FLower%2FUpper+Respiratory%22)%3BOrgan+System)
533 [By+Organ+System-81567%2FLower%2FUpper+Respiratory%22\)%3BOrgan+System.](http://www.idsociety.org/PracticeGuidelines/?q=&ref=taxonomy_map%3B(%22%2FIDSA%2FSite+Map%2FGuidelines%2FPatient+Care%2FIDSA+Practice+Guidelines%2FInfections+By+Organ+System-81567%2FLower%2FUpper+Respiratory%22)%3BOrgan+System)
534 Accessed 15 Feb 2018.
- 535 28. Coste J, Bouée S, Ecosse E, Leplège A, Pouchot J. Methodological issues in determining
536 the dimensionality of composite health measures using principal component analysis: case
537 illustration and suggestions for practice. *Qual Life Res.* 2005;14:641–654.
- 538 29. Everitt BS. Unresolved Problems in Cluster Analysis. *Biometrics.* 1979;35:169–181.

539 30. Wastell DG, Gray R. The numerical approach to classification: a medical application to
540 develop a typology for facial pain. *Stat Med*. 1987;6:137–146.

541 31. DeMerle KM, Angus DC, Baillie JK, Brant E, Calfee CS, Carcillo J, et al. Sepsis
542 Subclasses: A Framework for Development and Interpretation. *Crit Care Med*. 2021;49:748–
543 759.

544 32. Wong HR, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Willson DF, et al. Identification
545 of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med*.
546 2009;7:34.

547 33. Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, et al.
548 Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort
549 study. *Lancet Respir Med*. 2016;4:259–271.

550 34. Sweeney TE, Azad TD, Donato M, Haynes WA, Perumal TM, Henao R, et al. Unsupervised
551 Analysis of Transcriptomics in Bacterial Sepsis Across Multiple Datasets Reveals Three Robust
552 Clusters. *Crit Care Med*. 2018;46:915–925.

553 35. Omorou AY, Coste J, Escalon H, Vuillemin A. Patterns of physical activity and sedentary
554 behaviour in the general population in France: cluster analysis with personal and socioeconomic
555 correlates. *J Public Health (Oxf)*. 2016;38:483–492.

556

557 **Supporting information**

558 **S1 Table. STROBE—Checklist of items that should be included in reports of *cohort***
559 ***studies*.** An Explanation and Elaboration article discusses each checklist item and gives
560 methodological background and published examples of transparent reporting. The STROBE

561 checklist is best used in conjunction with this article (freely available on the Web sites of PLoS
562 Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at
563 <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the
564 STROBE Initiative is available at <http://www.strobe-statement.org>.

565

566 **S2 Table. Kits for soluble markers concentration measure.**

567

568 **S3 Table. Lower and Upper limits of Quantification for each plasma biomarker.**

569

570 **S4 Table. Primer and probe designs for mRNA biomarkers.**

571

572 **S5 Table. Circulating biomarkers with a loading > 0.40 or < -0.40 in each main**
573 **independent patterns obtained after principal component analysis.** The 5 patterns where
574 those which explained the largest part of the variance (59 % in total). For each pattern, the
575 individual biomarkers with high loadings are displayed. The individual biomarkers are sorted
576 by the absolute value of their loading. Loadings are interpretable as correlation coefficients
577 between patterns and original biomarkers.

578

579 **S6 Table. Value of each pattern of biomarkers in the identified clusters (med [Q1-Q3]).** In
580 each raw, the dark red color indicates the highest value and dark blue color indicates the lowest
581 value.

582

583 **S1 Fig. Distribution of missing values and determination of the cut-off to create binary**
584 **variables.** Lower part of the figure: category of the value of each biomarker for each patient.
585 Color code: white: value between the LLoQ and the ULoQ ; orange = value under the LLoQ;
586 blue = value over the ULoQ; grey = undetermined value.
587 X axis: biomarkers
588 Y axis: patients
589 Upper part of the figure:
590 Y axis: percentage of patients with a value below the LLoQ
591 The plain blue line indicates the separation between biomarkers treated as binary variables (left
592 side of the line) and as continuous variables (right side of the line). The dotted line indicates
593 the cut-off of 40%, over which the biomarker was treated as a binary value.

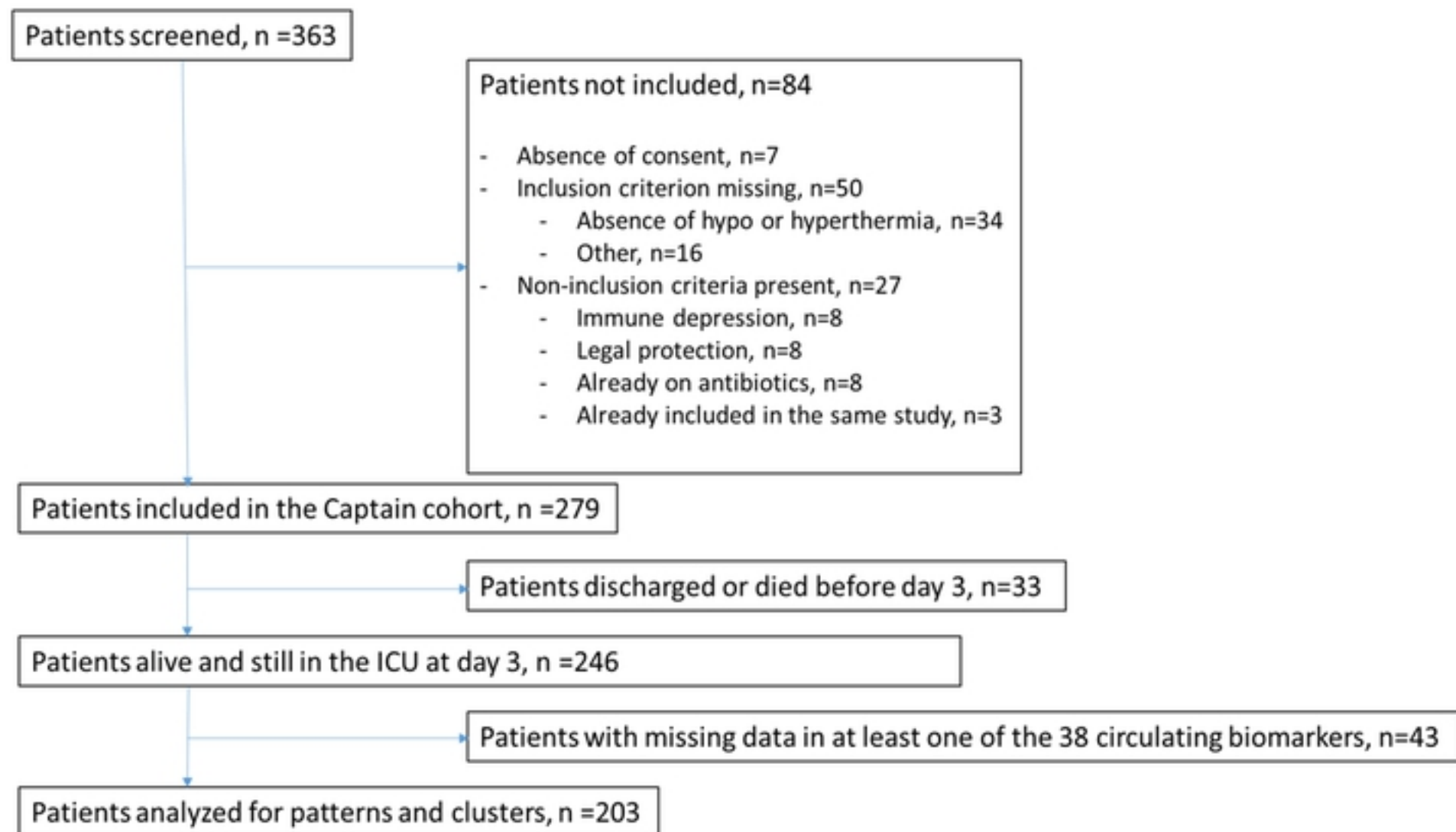


Figure 1

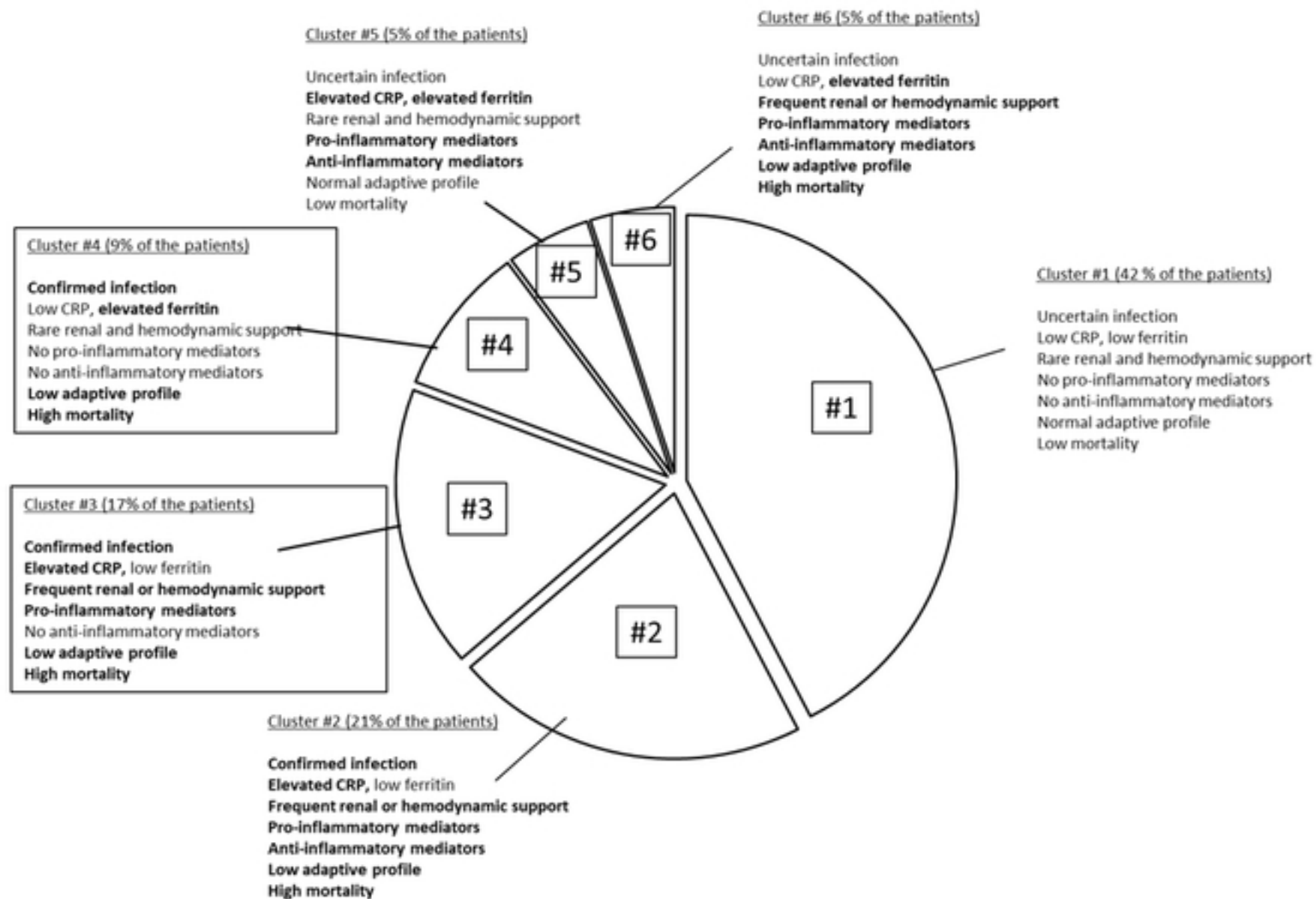


Figure 2