Journal of Infection xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Infection

journal homepage: www.elsevier.com/locate/jinf

Distinguishing septic shock from non-septic shock in postsurgical patients using gene expression

Pedro Martínez-Paz[a,b], Marta Aragón-Camino[b,c], Esther Gómez-Sánchez[a,b,c], Mario Lorenzo-López[a,b,c], Estefanía Gómez-Pesquera[a,b,c], Alejandra Fadrique-Fuentes[b,d], Pilar Liu[b,c], Álvaro Tamayo-Velasco[b,e], Christian Ortega-Loubon[b,f], Marta Martín-Fernández[b,g], Hugo Gonzalo-Benito[b,h,*], Emilio García-Morán[b,i], María Heredia-Rodríguez[a,b,j], Eduardo Tamayo[a,b,c]

[a] Department of Surgery, Faculty of Medicine, University of Valladolid. 7 Ramón y Cajal Ave, 47005 Valladolid, Spain
[b] BioCritic. Group for Biomedical Research in Critical Care Medicine. Valladolid, Spain
[c] Anesthesiology and Resuscitation Service, University Clinical Hospital of Valladolid. 3 Ramón y Cajal Ave, 47003 Valladolid, Spain
[d] Anesthesiology and Resuscitation Service, Hospital of Medina del Campo. 24 Peñaranda St, 47400 Medina del Campo (Valladolid), Spain
[e] Hematology and Hemotherapy Service, University Clinical Hospital of Valladolid. 3 Ramón y Cajal Ave, 47003 Valladolid, Spain
[f] Cardiovascular Surgery Service, Hospital Clinic of Barcelona. 170 Villarroel St, 08036 Barcelona, Spain
[g] Department of Medicine, Dermatology and Toxicology, Faculty of Medicine, University of Valladolid. 7 Ramón y Cajal Ave, 47005 Valladolid, Spain
[h] Institute of Health Sciences of Castile and Leon (IECSCYL). Santa Clara Sq, 42002 Soria, Spain
[i] Cardiology Service, University Clinical Hospital of Valladolid. 3 Ramón y Cajal Ave, 47003 Valladolid, Spain
[j] Anesthesiology and Resuscitation Service, University Hospital of Salamanca. 182 San Vicente Rd, 37007 Salamanca, Spain

ARTICLE INFO

Article history:
Accepted 28 May 2021
Available online xxx

Keywords:
Sepsis
Shock
Biomarkers
Gene expression profiling
Diagnosis

SUMMARY

Objectives: To obtain a gene expression signature to distinguish between septic shock and non-septic shock in postoperative patients, since patients with both conditions show similar signs and symptoms.

Methods: Differentially expressed genes were selected by microarray analysis in the discovery cohort. These genes were evaluated by quantitative real time polymerase chain reactions in the validation cohort to determine their reliability and predictive capacity by receiver operating characteristic curve analysis.

Results: Differentially expressed genes selected were *IGHG1, IL1R2, LCN2, LTF, MMP8*, and *OLFM4*. The multivariate regression model for gene expression presented an area under the curve value of 0.922. These genes were able to discern between both shock conditions better than other biomarkers used for diagnosis of these conditions, such as procalcitonin (0.589), C-reactive protein (0.705), or neutrophils (0.605).

Conclusions: Gene expression patterns provided a robust tool to distinguish septic shock from non-septic shock postsurgical patients and shows the potential to provide an immediate and specific treatment, avoiding the unnecessary use of broad-spectrum antibiotics and the development of antimicrobial resistance, secondary infections and increase health care costs.

Introduction

Sepsis is defined as organ dysfunction caused by a dysregulated host response to infection.[1] Nowadays, this condition is one of the main health care problems in the intensive care unit (ICU),[2] reaching an incidence of 300–480 cases per 100,000 people and being the leading cause of mortality (20–30%) in the ICU in developed countries.[3,4] Compared to other diseases, the incidence of sepsis remains higher than cerebrovascular accidents (76–119 per 100,000) and other pathologies with important social impact, such as breast cancer (25.9–94.2 per 100,000) or AIDS (24 per 100,000).[5-7] Moreover, sepsis is the first cause of mortality in non-coronary ICUs,[8,9] reaching a mortality rate of 38% for septic shock in Europe and North America.[10]

Similar to other illnesses, such as stroke, ischemic heart disease, or severe trauma, sepsis is a time-dependent disease, so it is important to get a fast and accurate diagnosis of this condition. In this sense, Peltan and colleagues described that the death rate increased by 20% in septic patients for each one hour increase in

* Corresponding author.
E-mail address: hgonzalob@saludcastillayleon.es (H. Gonzalo-Benito).

https://doi.org/10.1016/j.jinf.2021.05.039
0163-4453/

door-to-antimicrobial time.[11] In addition, it has recently been reported that the delay of antibiotic treatment initiation was associated with an increase in mortality among survivors of the initial illness.[12] However, there is no quick and accurate gold standard of sepsis diagnosis that allows differentiating this condition in septic shock and non-septic shock patients following a surgery, since patients with both conditions show similar signs and symptoms.[13] Therefore, the physician indiscriminately uses antibiotics in both conditions in a preventive way to avoid the death of septic shock patients. Thus, to get a quick and accurate diagnosis of septic shock allows an immediate and specific treatment for this condition and avoids the inappropriate use of broad-spectrum antibiotics to minimize the development of antimicrobial resistance, secondary infections and increase health care costs.[14] In addition, the increase of multi-resistant bacteria supposes a great worldwide health problem.[15]

At present, procalcitonin and C-reactive protein (CRP) are the biomarkers most used for the diagnosis and evolution of septic shock because they are more sensitive and specific than other biomarkers, such as lactate or leukocytes.[16] Nevertheless, the procalcitonin levels can be elevated by comorbid disease states, which makes it difficult to diagnose septic shock.[17] Moreover, procalcitonin is a reliable biomarker in medical patients with septic shock but not in surgical patients with septic shock.[18] In the case of CRP, this biomarker is considered sensitive but not specific for this condition and, moreover, their levels can also increase in response to other conditions, such as trauma, ischemia, burns, and other inflammatory conditions.[19] Another classical biomarker used for the diagnosis of sepsis is the leukocyte count; however, Jukic and colleagues described that their count was not a valid biomarker to predict postoperative infection.[20] On the other hand, interleukin-6 and soluble triggering receptor expressed on myeloid cells-1 have been suggested to diagnose sepsis before the onset of symptoms.[21] Therefore, in surgical patients with sepsis, the procalcitonin, CRP, and leukocyte levels are the result of a synergistic effect at the same time of the inflammatory response invoked by surgical damage and bacteria. For this reason, these biomarkers are not the best to differentiate between septic shock and non-septic shock patients following a surgery, so identifying patients correctly can be a challenge for physicians. In this sense, previous works have analyzed the ability of different biomarkers to distinguish septic shock from systemic inflammatory response syndrome; however, the evaluated biomarkers poorly discriminated these conditions. For example, more than 50 circulating biomarkers that had been previously proposed to detect infection, but they demonstrated a modest ability to distinguish septic shock and non-septic shock.[22] Similar results were obtained when cell death biomarkers were analyzed.[23]

Currently, one of the methods used for the design of clinical biomarkers is the analysis of gene expression levels, which allows identifying those genes that are activated during a specific pathological state. However, this methodological approach in sepsis studies is scarce,[24] although in the last years, it has increased the references about the biomarkers sought in sepsis.[25–28] Nevertheless, these reports only analyzed the transcriptomic levels in medical sepsis. Therefore, the aim of the present work is to propose a putative gold standard to help in distinguishing postsurgical septic shock from non-septic shock by identifying the gene expression profile of both conditions and developing an algorithm based on genes that show differences in their expression levels.

Methods

Patient selection and clinical data

The present work was performed in the ICU of Hospital Clínico Universitario de Valladolid (Spain) with two study cohorts of adult patients following surgery recruited sequentially between January 2014 and December 2018. For the discovery cohort, a total of 133 patients undergoing surgery were recruited and diagnosed with septic shock ($n = 80$) and non-septic shock ($n = 33$) to perform the microarray analysis. In the case of the validation cohort, a set of 167 patients were recruited, also following surgery, and diagnosed with septic shock ($n = 107$) and non-septic shock ($n = 55$) by quantitative real-time polymerase chain reaction. Fifteen healthy volunteers of similar ages to the patients were recruited for gene expression data normalization. The diagnosis of septic shock was established according to the definition of the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3 definition) and the criteria for the diagnosis of non-septic shock were the same for septic shock without infection.[1] Two physicians reviewed all the cases before the inclusion of the patients in each group. Confirmation of the microbiology cultures by the Microbiology Service from Hospital Clínico Universitario de Valladolid was necessary for inclusion of the patients in the septic shock group. On the other hand, non-Caucasians, pregnant women, patients in agonizing state, and those in a state of limitation of the therapeutic efforts were excluded from the final analysis. All research involving human participants was approved by the Scientific Committee for Clinical Research of Hospital Clínico Universitario de Valladolid, and the patient or his/her legal representative provided informed written consent before recruitment. This study followed the code of ethics of the World Medical Association (Declaration of Helsinki).

Patient selection and clinical data

Microbial diagnosis

When infection was suspected, samples were routinely gram stained and cultured on general purpose media (blood agar, chocolate agar, and the differential media McConkey agar and Chapman agar) by the Microbiology Service from Hospital Clínico Universitario de Valladolid (Spain). For fungal infections, sputum samples were grown on Sabouraud agar containing chloramphenicol. Isolation of microorganisms or clinical suspicion of infection, without administration of antibiotics, was not considered as infection. In this sense, we followed the practice guideline of the Infectious Diseases Society of America and the American Society for Microbiology Clinical Infectious Diseases.[29]

Sample collection and RNA extraction

Whole blood samples (2.5 mL) were collected from healthy controls and septic shock and non-septic shock patients using PAXgene venous blood vacuum collection tubes (Becton Dickinson, USA). Blood samples were obtained in septic shock and non-septic shock patients within 24 h of diagnosis, usually 6–8 h after antibiotic treatment start.

Total RNA was extracted and purified from blood samples using the PAXgene Blood RNA System (PreAnalytix, Switzerland) and RNeasy Mini Kit (Qiagen, Germany) following the manufacturerś protocol. The quality of the total RNA was assessed by RNA Experion Bioanalyser (Bio-Rad, USA), and the quantity was evaluated by absorbance in a NanoDrop 1000 Spectrophotometer (NanoDrop Technologies, USA). Up to 1,750 ng of each RNA sample was concentrated with the RNeasy MinElute Cleanup Kit (QIAGEN, Hilden, Germany) and eluted in a final volume of 10 μL according to the manufacturer's instructions. Finally, the purified RNA was stored at -80 °C.

Microarray processing and data analysis

In all, 100 µg of total RNA was used to produce cyanine 3-CTP-labelled cRNA using the Quick Amp Labelling Kit (Agilent, USA)

Table 1
Primers used for qRT-PCR of genes from humans.

Gene	Forward (5′-3′)	Reverse (5′-3′)	Efficiency
Actin	CCTTGCACATGCCGGAG	ACAGAGCCTCGCCTTTG	87.2%
IGHG1	ACGGCGTGGAGGTGCATAATG	GTTGGAGACCTTGCACTTGTACTC	96.9%
IL1R2	GCATCTGTATTCTCAAAAACTCTGA	GGTGCTCTGTGGCTTCTG	96.9%
LCN2	TCTCCCAGCTCCCTCAATG	AACTTCATCCGCTTCTCCAA	97.2%
LTF	CTGCAAGACAGTGACATCTTTC	TCTGTGCTCTGTGTATTGGC	94.3%
MMP8	GCATCAACTTTACTCTCTATTCCTG	CATACTTCTTTGTAAATGACCAATTCTG	94.5%
OLFM4	TGCTGATGTTCACCACACC	CTGAAGACCAAGCTGAAAGAGT	92.2%

according to the manufacturer's instructions. Following the One-Colour Microarray-Based Gene Expression Analysis protocol Version 5.7 (Agilent, USA), 3 µg of labelled cRNA was hybridized with the Whole Human Genome Oligo Microarray Kit (Agilent, USA), containing 41,000 unique human genes and transcripts. Arrays were scanned in an Agilent G2565BA Microarray Scanner System (Agilent, USA) according to the manufacturer's protocol, and data were extracted using Agilent Feature Extraction Software 9.5.3 following the Agilent protocol GE1-v5_95_Feb07 and the QC Metric Set GE1_QCMT_Jan08. Raw data files were imported into R-Bioconductor programming environment using the read.maimages from the limma package. Repeat probes were aggregated by their median value. Pre-processing was continued by background correction. The normexp ('saddle') method was used with an offset value of 50. Normalization between the arrays was performed by the quantile method. The expression matrix was summarized for further analysis by the selection of the top decile of probes in variance. Differential expression analysis was continued using the lmFit function from the limma package in order to obtain logFold changes between the septic shock and non-septic shock groups. The microarray dataset has been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE131761. Functional enrichment of the most relevant genes was performed using GO-BP ontology with the help of clusterProfiler R package. This software relies on the database org.Hs.eg.db and the cutoff values were for *p*-values 0.01 and for *q*-values 0.05.

Quantitative real-time polymerase chain reaction (qRT-PCR)

To validate the reliability of data obtained by microarray analysis, six genes were selected based on fold changes and *p*-value in microarray and their expression levels were evaluated by qRT-PCR in the validation cohort. To obtain the cDNA, reverse transcription was performed with the iScript Advanced cDNA Synthesis Kit (Bio-Rad) using the isolated RNA from the patients. The cDNA was used as a template for the qRT-PCR to evaluate the mRNA expression profile of the representative genes *IGHG1, IL1R2, LCN2, LTF, MMP8*, and *OLFM4* from control, septic shock and non-septic shock patients. qRT-PCR was performed on a CFX96 thermocycler (Bio-Rad) using PrimeTime Gene Expression Master Mix and the following cycling conditions: initial denaturation at 95 °C for 3 min and 45 cycles of denaturation at 95 °C for 15 s; annealing and elongation at 62 °C for 15 s. In each case, the gene expression patterns of septic shock patients were compared with those observed in non-septic shock patients after normalization with the actin gene, which was employed as the reference gene. The sequence of primers for selected genes are listed in Table 1. In the case of *IGHG3, IGHG2, IGHA1*, and *IGHA2* genes, a primer pair was designed to amplify all IgG classes.

PCR amplification efficiency was established using calibration curves. For each gene, a standard curve based on five dilutions from an equimolar mix of cDNA samples was produced in triplicate to verify the amplification efficiency (Table 1). Each sample was run in triplicate wells. The cycle threshold (C_t) values were obtained with Bio-Rad CFX Maestro software and converted to relative gene expression levels using the $2^{-\Delta\Delta Ct}$ method.

Statistical analysis

All the statistical analysis was performed using SPSS Statistics version 24 (IBM, USA) and R statistical package version 3.4.1 (The R Foundation, Austria). Categorical variables were evaluated using Pearson's χ^2 test, and continuous variables were analyzed by Student's t-test to find qualitative statistical significance between septic shock and non-septic shock patients. Normal distribution and variance homogeneity of data were assessed using the Kolmogorov-Smirnov and Levene's tests, respectively. For microarray analysis, *p*-values were adjusted by False Discovery Rate method following the Benjamini-Hochberg procedure. A receiver operating characteristic (ROC) curve analysis with the area under the curve (AUC) was performed to quantify the sensitivity and specificity for gene expression levels, procalcitonin, CRP, and neutrophils. A forward multivariate logistic model process was developed using bilirubin, CRP, glucose, and neutrophils as adjust variables to add the best-performing clinical parameters to our model. In all cases, a p-value ≤ 0.05 was considered to indicate statistical significance. The optimal operating point (OOP) was calculated by Pythagorean theorem, where this value was the point on the curve that had the minimum distance to the upper left corner (sensitivity = 1 and specificity = 1); $OOP = \sqrt{(1 - \text{sensitivity})^2 + (1 - \text{specificity})^2}$.

Results

Patient characteristics

The characteristics of the postsurgical patients enrolled in this study are described in Table 2. Septic shock and non-septic shock patients in the discovery cohort and validation cohort were largely similar in demographics and comorbidities. The length of hospital stay was significantly different between both cohorts; however, the mortality did not differ between them. Regarding the kind of shock in non-septic shock group of discovery cohort, 63% was caused by cardiogenic shock, 26% by hypovolemic shock, and 11% by distributive shock. For validation cohort, 62% was caused by cardiogenic shock, 13% by hypovolemic shock, and 25% by distributive shock. On the other hand, the most common surgery in septic shock patients was general surgery (45% for discovery cohort and 60.75% for validation cohort), followed by cardiac surgery (40% and 26.17%) and other types of surgeries (15% and 13.08%). For the non-septic shock patients, the most common surgery was cardiac surgery (81.82% and 78.18%), followed by general surgery (9.09% and 10.91%) and other types of surgeries (9.09% and 10.91%).

In the sepsis group, the blood was the most common site of infection (46% for discovery cohort and 37% for validation cohort), followed by the respiratory tract (40% and 30%), abdomen (31% and 30%) and urinary tract (21% and 18%). On the other hand, the most common microorganism isolated was gram-negative bacteria (76%

Table 2
Characteristics of postsurgical patients.

Parameter	Discovery cohort			Validation cohort		
	Septic shock (n = 80)	Non-septic shock (n = 33)	p	Septic shock (n = 107)	Non-septic shock (n = 55)	p
Characteristics						
Age	70.68	65.52	0.056	69.15	66.09	0.155
Male [% (n)]	58.75 (47)	48.48 (16)	0.318	57.94 (62)	67.27 (37)	0.249
Comorbidities, % (n)						
High blood pressure	66.25 (53)	60.61 (20)	0.568	62.62 (67)	58.18 (32)	0.583
Chronic cardiovascular disease	51.25 (41)	57.57 (19)	0.540	33.64 (36)	23.63 (13)	0.189
Chronic respiratory disease	15.00 (12)	9.09 (3)	0.400	18.69 (20)	16.36 (9)	0.714
Chronic renal failure	11.25 (9)	6.06 (2)	0.397	6.54 (7)	9.09 (5)	0.557
Chronic hepatic failure	1.25 (1)	0.00 (0)	0.519	1.87 (2)	1.81 (1)	0.982
Diabetes mellitus	23.75 (19)	15.15 (5)	0.310	23.36 (25)	12.73 (7)	0.107
Cancer	28.75 (23)	12.12 (4)	0.059	26.17 (28)	14.55 (8)	0.092
Time course and outcome						
Length of hospital stay	33.60	23.36	0.025	31.05	21.55	0.020
Length of ICU stay	12.24	7.30	0.350	10.34	7.89	0.149
Mortality [% (7 days)]	13.58 (11)	15.15 (5)	0.706	11.21 (12)	20.00 (11)	0.129
Mortality [% (28 days)]	30.86 (25)	30.30 (10)	0.869	28.04 (30)	32.73 (18)	0.536
Mortality [% (90 days)]	45.68 (37)	30.30 (10)	0.184	36.45 (39)	34.55 (19)	0.811
Type of surgery % (n)						
Cardiac surgery	40.00 (32)	81.82 (27)	0.000	26.17 (28)	78.18 (43)	0.000
General surgery	45.00 (36)	9.09 (3)	0.000	60.75 (65)	10.91 (6)	0.000
Others	15.00 (12)	9.09 (3)	0.776	13.08 (14)	10.91 (6)	0.387
Source of infection, % (n)						
Respiratory tract	40.00 (32)	0.00 (0)	0.000	29.91 (32)	0.00 (0)	0.000
Abdomen	31.25 (25)	0.00 (0)	0.000	29.91 (32)	0.00 (0)	0.000
Urinary tract	21.25 (17)	0.00 (0)	0.000	17.76 (19)	0.00 (0)	0.000
Surgical site	35.00 (28)	0.00 (0)	0.000	35.51 (38)	0.00 (0)	0.000
Bacteraemia	46.25 (37)	0.00 (0)	0.000	37.38 (40)	0.00 (0)	0.000
Microbiology, % (n)						
Gram +	65.00 (52)	0.00 (0)	0.000	62.62 (67)	0.00 (0)	0.000
Gram −	76.25 (61)	0.00 (0)	0.000	64.49 (69)	0.00 (0)	0.000
Fungi	28.75 (23)	0.00 (0)	0.000	28.04 (30)	0.00 (0)	0.000
Measurements at diagnosis, [median (IQR)]						
SOFA score	9.00 (4)	9.00 (3)	0.957	8.00 (3)	9.00 (3)	0.188
APACHE score	14.00 (6)	10.00 (6)	0.002	14.00 (7)	12.50 (5)	0.042
Total bilirubin (mg/dL)	0.80 (1.43)	0.82 (1.80)	0.987	0.80 (1.21)	1.10 (1.37)	0.101
Glucose (mg/dL)	157 (70)	166 (87)	0.113	160 (75)	178 (97)	0.025
Platelet count (cell/mm^3)	148500 (173250)	100000 (75000)	0.000	180000 (188000)	144000 (86000)	0.002
$ScvO_2$ (%)	70.90 (12.30)	68.70 (21.90)	0.031	72.30 (13.00)	61.90 (17.65)	0.000
C-reactive protein (mg/L)	220.95 (198.58)	90.60 (158.85)	0.000	231.30 (183.00)	106.00 (184.00)	0.000
Procalcitonin (ng/mL)	5.99 (20.61)	1.42 (3.99)	0.000	4.03 (21.40)	2.49 (4.18)	0.000
Lactate (mM)	2.83 (2.00)	3.55 (4.00)	0.156	2.55 (2.11)	3.66 (4.33)	0.018
White blood cells (cells/mm^3)	16575 (11997)	13620 (4580)	0.070	16330 (12507)	13880 (7500)	0.100

SOFA, sequential organ failure assessment; APACHE, acute physiology and chronic health evaluation; INR, international normalized ratio; ScvO2, central venous oxygen saturation. Quantitative data are expressed as medians with interquartile range (IQR). Qualitative data are presented as percentages and absolute numbers. A p-value $\leq$ 0.05 was considered to indicate significant differences.

and 64%), followed by gram-positive bacteria (65% and 63%) and fungi (29% and 28%).

Identification of biomarker genes discriminating septic shock from non-septic shock patients

We performed an exploratory data analysis by computing the principal components of the gene expression matrix. The plot of PC1 versus PC2 is shown in Fig. 1A, where every dot represents a patient and its color codes which group it belongs to. This first reduction of dimensionality suggests that the healthy control group obviously diverges in gene expression, and although overlap between septic shock and non-septic shock patients exist, the former patients tend to lie lower across PC1 and PC2. This differential gene expression was confirmed by the results of the multiple linear regression model, which provided a log fold change and a statistically significant p-value for every transcript. The simultaneous representation of -log10(p) versus log fold changes provides the volcano plot (Fig. 1B), where every dot is a transcript. A small group of transcripts lies above the signification threshold (p-value $\leq$ 0.05) and beyond a positive log fold change of 1.5 (Table S1). From this ranking, we selected the first nine genes that were differentially overexpressed in septic shock patients: *LCN2, IGHG3, IGHG2, LTF, OLFM4, IGHA1, IGHA2, MMP8*, and *IL1R2*. The expression of these genes is shown for septic shock patients besides non-septic shock patients in the heat map (Fig. 1C). Finally, the enriched functional characterization of these nine genes is presented in Fig. 1D and Table S2. All significant enriched tags are related to the immune system, where the neutrophil mediated immunity, neutrophil degranulation, neutrophil activation involved in immune response and neutrophil activation are the most shared features.

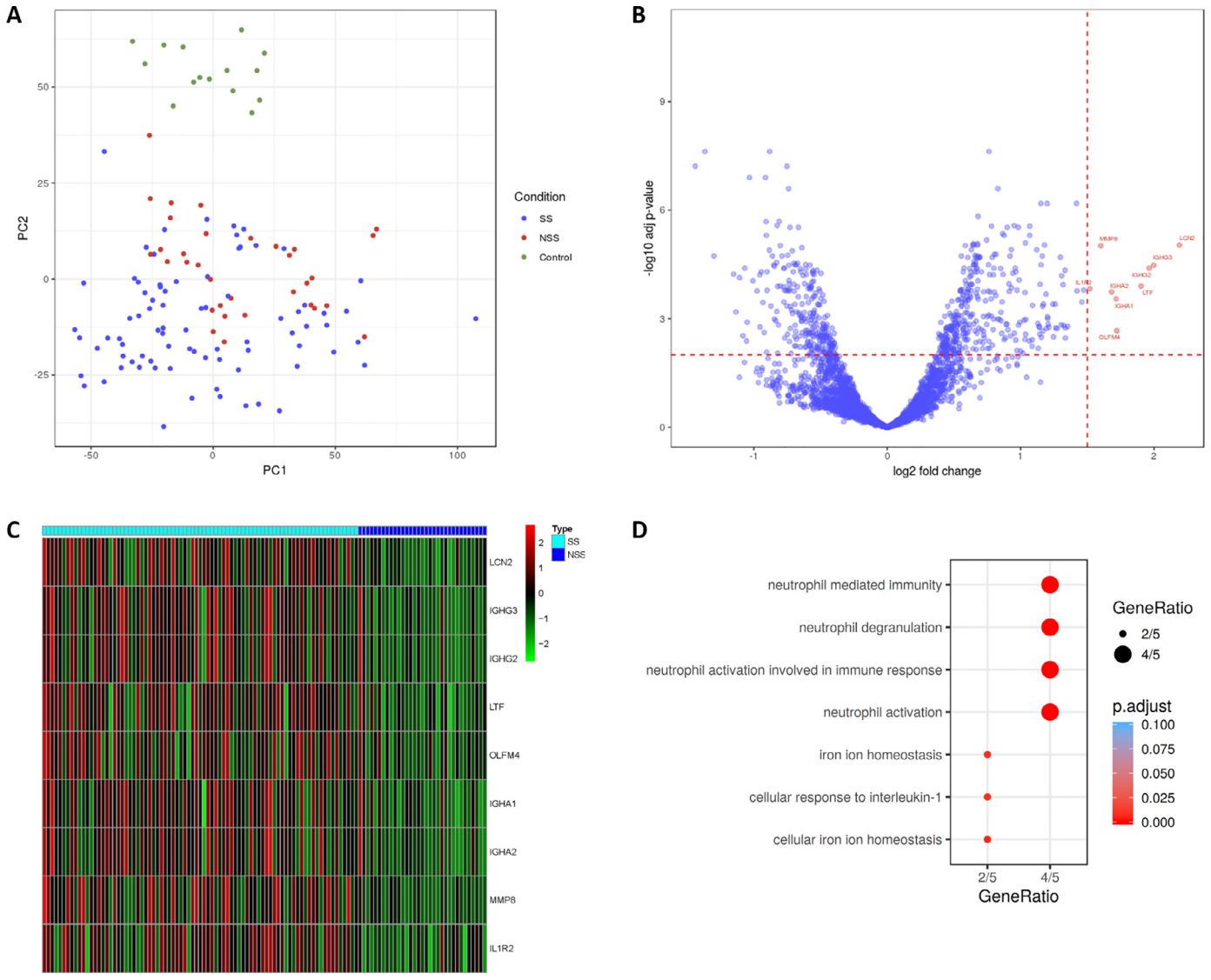

Fig. 1. Variation in the shock transcriptional response according to shock origin. (A) Scatterplot of the first two principal components (PC1 and PC2) obtained from the principal component analysis (PCA) performed on the DE transcripts. PCA plots reveal three distinct patient groups. Variation was 0.3244 for PC1 and 0.1961 for PC2. (B) Volcano plot. Dashed red lines indicate the threshold value for fold changes >1.5 and *p*-value < 0.01, blue dots represent the genes that are below the threshold and red dots indicate the genes that are above the threshold (C) Heat map plot from genes of interest (*LCN2, IGHG3, IGHG2, LTF, OLFM4, IGHA1, IGHA2, MMP8*, and *IL1R2*). Row represents the gene expression value and column represents the sample. The colour conventions are as follows: red represents over-expressed transcripts and green indicates under-expressed transcripts; septic shock patients are represented by light blue and non-septic shock patients are represented by dark blue. (D) Enriched biological functions: The size of the dots represents the number of genes in the significant differentially expressed gene list associated with the GO term and the color of the dots represent the *p*-adjusted values. Figure legends: SS, septic shock; NSS, non-septic shock.

Validation of the biomarker genes in a validation cohort

To evaluate the robustness of the selected genes as candidate biomarkers, we tested it by qRT-PCR in an independent study cohort (validation cohort). In concordance with microarray data from the discovery cohort, the transcriptional activity of *LCN2, IGHG3, IGHG2, LTF, OLFM4, IGHA1, IGHA2, MMP8*, and *IL1R2* in septic shock patients were significantly increased as compared with non-septic shock patients (Fig. 2). In addition, expression patterns of these genes were evaluated in healthy controls with the aim to define a threshold (Fig. S1).

Assessment of the biomarker genes in a validation cohort

With the aim of comparing our results with the most used biomarkers for the diagnosis and evolution of septic shock, we assessed the ability of gene expression levels, procalcitonin, CRP, and neutrophils to discriminate between septic shock and non-septic shock patients. The clinical parameter values were restricted at the moment of blood collection for gene expression assays. In a first step, these biomarkers were evaluated using ROC curve analysis in the validation cohort, with AUC to quantify its accuracy. As shown in Fig. 3A, the highest reported AUC was for *IGHG1* (0.797; 0.727–0.867) followed by *LTF* (0.719; 0.634–0.803), *OLFM4* (0.698; 0.611–0.786), *LCN2* (0.677; 0.590–0.764), *MMP8* (0.637; 0.548–0.726) and *IL1R2* (0.632; 0.540–0.725). On the other hand, we performed a multivariate regression model including all the genes evaluated, improving the AUC up to 0.841 (0.779–0.904) (Fig. 3B). In the case of the classical biomarkers, their AUCs were not as good as expected in order to differentiate the two patient categories: 0.705 (0.605–0.804) for CRP, 0.605 (0.505–0.706) for neutrophils, and 0.598 (0.504–0.693) for procalcitonin (Fig. 3B). In an effort to further improve the AUC, we performed a multivariate regression model including in the analysis the variables bilirubin, CRP, glu-

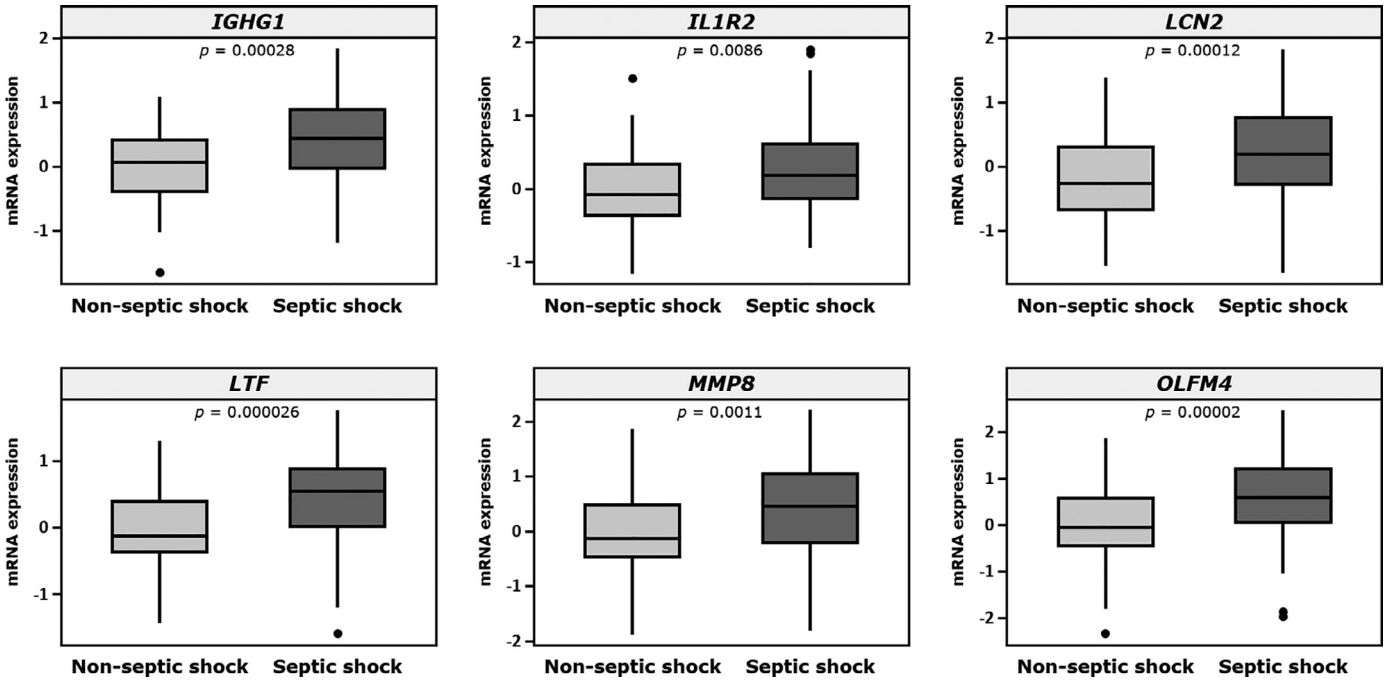

Fig. 2. Transcriptional patterns of the candidate gene biomarkers. Relative mRNA levels of *IGHG1, IL1R2, LCN2, LTF, MMP8*, and *OLFM4* in non-septic shock and septic shock patients as measured by qRT-PCR with primers and reference genes as indicated in Materials and methods. The horizontal line within the box indicates the median, the boundaries of the box indicate the 25th and 75th percentiles, and the whiskers indicate the highest and lowest values. The Y-axis represents the RNA expression levels in arbitrary units and logarithmic scale.

cose, and neutrophils. The AUCs were better than the AUCs presented previously for all the biomarkers analyzed (Fig. 3C). The multivariate regression model for gene expression had very good accuracy in the studied patients, with AUC = 0.922 (0.874–0.970) (Fig. 3D). OOP, sensitivity and specificity values from each ROC curves are showed in Table S3.

Discussion

In this study, we assessed the transcriptomic profile in patients who developed septic shock and non-septic shock after a major surgery. The comparison of the gene expression patterns between these patients revealed: i) gene expression patterns could distinguish these conditions; ii) each of six genes were identified as single efficient biomarkers that discriminated between both kinds of shock; and iii) the expression of a few genes could discern between septic shock and non-septic shock patients better than other biomarkers used to discriminate cases of these conditions.

Microarray analysis has previously been used to predict diagnostic marker genes of different diseases, such as gastric cancer, Parkinsonś disease, Alzheimer's disease or Crohnś disease.[30–33] In this sense, different works have classified sepsis patients using blood transcriptomic profiles,[34–37] but any previous work has stratified septic shock from non-septic shock. Thus, the present work shows a differential gene expression pattern that can reflect the presence of infection in patients with postoperative shock. Based on the blood gene expression profiling data, the microarray analysis revealed a set of genes that were differentially expressed between septic shock and non-septic shock patients. We identified the differentially expressed genes in the discovery cohort of patients that were selected based on their estimated of log fold changes, with the top six upregulated genes being *IGHG1, IL1R2, LCN2, LTF, MMP8*, and *OLFM4*. These genes encode for six proteins involved in the immune system response, being consistent with the key role of the immune system in sepsis and with previous studies that suggest that sepsis is accompanied by overall immune dysregulation.[34,35,38] Some of the genes identified in this work have been previously reported when postsurgical septic patients were compared with postsurgical patients who did not show any signs or symptoms.[34] However, the main strength of this study is the comparison between septic shock and non-septic shock postoperative patients, two similar conditions at their onset, obtaining new biomarker genes that can distinguish these conditions with similar signs and symptoms but different etiology. In addition to providing novel diagnostic markers to distinguish septic shock from non-septic shock patients, the knowledge of these selected genes can provide new opportunities to understand the mechanisms of illness and offer new therapeutic pathways.

Despite previous works having studied different biomarkers for the diagnosis of sepsis, no gold standard has been identified. Nowadays, different non-gene expression biomarkers are routinely used in the diagnosis of infections by physicians. However, there is a controversial discussion on the role of these biomarkers as a diagnostic tool for infection. Thus, previous studies have investigated the diagnostic accuracy of procalcitonin in critically ill patients, with conflicting results.[39,40] Procalcitonin is increased by many inflammatory states and may be influenced by previous comorbidities, which are associated with higher values at baseline. An additional sepsis biomarker is CRP, but even though it is in routine clinical use, its specificity for sepsis diagnosis is often questioned. Different studies have analyzed CRP as a potential biomarker for sepsis diagnosis; however, the results have not been consistent.[41,42] Other emerging sepsis biomarkers are proadrenomedullin, presepsin and endothelial markers; however, all have different limitations to distinguish infected from non-infected shock patients, and further research is needed to analyze their validity and role in sepsis diagnosis to include them in hospital-based clinical laboratories. In this sense, with the aim to obtain gene biomarkers for sepsis diagnosis, we focused on the evaluation of the gene expression of the selected differentially expressed genes based on the results of the microarray analysis. We analyzed the transcriptional responses of *IGHG1, IL1R2, LCN2, LTF, MMP8*, and *OLFM4* in septic shock and

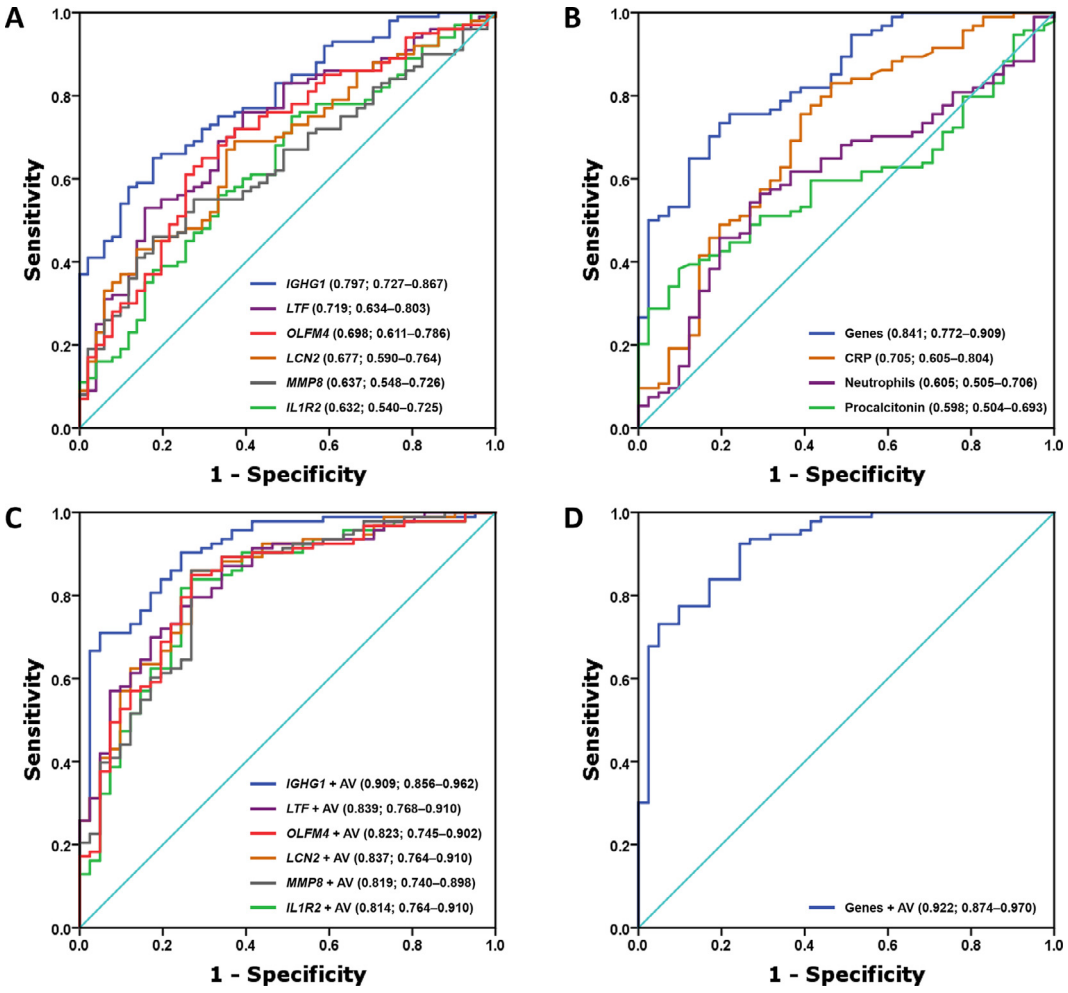

Fig. 3. ROC AUC analysis of biomarkers in the validation cohort. (A) ROC AUC analysis of gene expression. (B) ROC AUC comparison between gene expression regression model, procalcitonin, CRP, and neutrophils. (C) ROC AUC analysis of gene expression regression model including bilirubin, CRP, glucose, and neutrophils as adjust variables. (D) ROC AUC analysis of multivariate regression model that includes all gene expression and bilirubin, CRP, glucose, and neutrophils as adjust variables. Figure legends: AV, adjust variables.

non-septic shock patients after a major surgery in an independent cohort by qRT-PCR. Thus, we calculated the AUC values from mRNA levels of each selected gene and performed a multivariate regression model including the gene expression and different adjust variables. Having demonstrated that the selected genes can distinguish between septic shock and non-septic shock patients, we proceeded to compare our results with procalcitonin, CRP and neutrophils in the same patient cohort. Although these three non-gene expression biomarkers are routinely used in the diagnosis of infection by physicians, our results demonstrate that the set of genes is better than the classical markers, since they present the highest AUC value (0.922), that is considered outstanding to differentiate this kind of patients.[43] Moreover, it would be enough to analyze the expression levels of *IGHG1* gene in clinical routine because it presents sufficient diagnostic accuracy (0.909).

Our study has several limitations that should be considered. First, it is a single-centre study; therefore, a multi-centre study would provide valuable insight into the global transcriptional septic response in shock patients to confirm our results and confer utility in clinical application. Second, the gene expression patterns have been analyzed independent of the pathogen. In this sense, while previous studies have reported a common gene pattern independent of pathogen, other works have supported that the transcriptional response can discriminate the infecting organisms.[44,45] Boldrick and colleagues reported that, despite the striking similarities, there were qualitative and quantitative differences in the responses to different bacteria.[46] Third, the last limitation is that this work did not analyze the evolution of the transcriptomic profile over time. Thus, future studies should keep in mind the pathogen phenotype to analyze the possible relationship with the blood-gene expression patterns and to evaluate the effects of time on gene expression after the shock diagnosis.

In conclusion, the differential expression of some immunological genes can help physicians to solve the challenge of distinguishing septic shock and non-septic shock postsurgical patients. Thus, this study provides a transcript tool based on a small number of genes to classify the patients as septic shock and non-septic shock that would make the procedures easily transferable to clinical laboratories, where PCR is a quick, accurate, cheap, and reliable technique used on a daily basis.

Authors' contributions

PMP, EGM, and ET contributed to the study conception and design. PMP, MAC, EGS, MLL, EGP, AFF, PL, ATV, COL, EGM, HGB, MHR, and ET had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. EGS, MLL, EGP, AFF, PL, ATV, COL, and HGB assisted with data collection and study management. EGM performed the bioin-

formatic analysis. PMP, EGM, and ET drafted the manuscript for intellectual content. ET is responsible for the funding of the study.

Declaration of Competing Interest

None.

Acknowledgments

We would like to thank nursery staff from Research Service of Hospital Clínico Universitario de Valladolid for sampling support of this work. This work was supported by the Instituto de Salud Carlos III from Ministerio de Ciencia, Innovación y Universidades (Spain) [grant number PI15/01451]; and Gerencia Regional de Salud from Junta de Castilla y León (Spain) [grant number 1255/A/16]. PMP was supported by a Postdoctoral Fellowship from Universidad de Valladolid co-financed by European Regional Development Fund [grant number VA321P18].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jinf.2021.05.039.

References

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM, Hotchkiss RS, Levy MM, Marshall JC, Martin GS, Opal SM, Rubenfeld GD, van der Poll T, Vincent JL, Angus DC. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;**315**:801–10.
2. Vincent JL, Sakr Y, Sprung CL, Ranieri VM, Reinhart K, Gerlach H, Moreno R, Carlet J, Le Gall JR, Payen D. Sepsis in European intensive care units: results of the SOAP study. *Crit Care Med* 2006;**34**:344–53.
3. Angus DC, van der Poll T. Severe sepsis and septic shock. *N Engl J Med* 2013;**369**:840–51.
4. Álvaro-Meca A, Jiménez-Sousa MA, Micheloud D, Sánchez-Lopez A, Heredia-Rodríguez M, Tamayo E, Resino S. Epidemiological trends of sepsis in the twenty-first century (2000-2013): an analysis of incidence, mortality, and associated costs in Spain. *Popul Health Metr* 2018(1):4.
5. Thrift AG, Thayabaranathan T, Howard G, Howard VJ, Rothwell PM, Feigin VL, Norrving B, Donnan GA, Cadilhac DA. Global stroke statistics. *Int J Stroke* 2017;**12**:13–32.
6. World Health Organization. *The global cancer observatory*; 28 April 2020. Available at *http://gco.iarc.fr* Accessed.
7. World Health Organization. *Global health observatory data repository*; 28 April 2020. Available at *http://apps.who.int/gho/data/node.main.HIVINCIDENCE?lang=en* Accessed.
8. Ferrer R, Artigas A, Levy MM, Blanco J, González-Díaz G, Garnacho-Montero J, Ibáñez J, Palencia E, Quintana M, de la Torre-Prados MV. Improvement in process of care and outcome after a multicenter severe sepsis educational program in Spain. *JAMA* 2008;**299**:2294–303.
9. Levy MM, Rhodes A, Phillips GS, Townsend SR, Schorr CA, Beale R, Osborn T, Lemeshow S, Chiche JD, Artigas A, Dellinger RP. Surviving sepsis campaign: association between performance metrics and outcomes in a 7.5-year study. *Crit Care Med* 2015;**43**:3–12.
10. Vincent JL, Jones G, David S, Olariu E, Cadwell KK. Frequency and mortality of septic shock in Europe and North America: a systematic review and meta-analysis. *Crit Care* 2019;**23**:196.
11. Peltan ID, Mitchell KH, Rudd KE, Mann BA, Carlbom DJ, Hough CL, Rea TD, Brown SM. Physician variation in time to antimicrobial treatment for septic patients presenting to the emergency department. *Crit Care Med* 2017;**45**:1011–18.
12. Peltan ID, Brown SM, Bledsoe JR, Sorensen J, Samore MH, Allen TL, Hough CL. ED Door-to-antibiotic time and long-term mortality in sepsis. *Chest* 2019;**155**:938–46.
13. Vincent JL, De Backer D. Circulatory shock. *N Engl J Med* 2013;**369**:1726–34.
14. Naylor NR, Atun R, Zhu N, Kulasabanathan K, Silva S, Chatterjee A, Knight GM, Robotham JV. Estimating the burden of antimicrobial resistance: a systematic literature review. *Antimicrob Resist Infect Control* 2018;**7**:58.
15. Prestinaci F, Pezzotti P, Pantosti A. Antimicrobial resistance: a global multifaceted phenomenon. *Pathog Glob Health* 2015;**109**:309–18.
16. Pierrakos C, Vincent JL. Sepsis biomarkers: a review. *Crit Care* 2010;**14**:R15.
17. Covington EW, Roberts MZ, Dong J. Procalcitonin monitoring as a guide for antimicrobial therapy: a review of current literature. *Pharmacotherapy* 2018;**38**:569–81.
18. Clec'h C, Fosse JP, Karoubi P, Vincent F, Chouahi I, Hamza L, Cupa M, Cohen Y. Differential diagnostic value of procalcitonin in surgical and medical patients with septic shock. *Crit Care Med* 2006;**34**:102–7.
19. Santonocito C, De Loecker I, Donadello K, Moussa MD, Markowicz S, Gullo A, Vincent JL. C-reactive protein kinetics after major surgery. *Anesth Analg* 2014;**119**:624–9.
20. Jukic T, Ihan A, Stubljar D. Dynamics of inflammation biomarkers C-reactive protein, leukocytes, neutrophils, and CD64 on neutrophils before and after major surgical procedures to recognize potential postoperative infection. *Scand J Clin Lab Invest* 2015;**75**:500–7.
21. Dolin HH, Papadimos TJ, Stepkowski S, Chen X, Pan ZK. A novel combination of biomarkers to herald the onset of sepsis prior to the manifestation of symptons. *Shock* 2018;**49**:364–70.
22. Parlato M, Philippart F, Rouquette A, Moucadel V, Puchois V, Blein S, Bedos JP, Diehl JL, Hamzaoui O, Annane D, Journois D, Boutieb MB, Estève L, Fitting C, Treluyer JM, Pachot A, Adib-Conquy M, Cavaillon JM, Misset B. Circulating biomarkers may be unable to detect infection at the early phase of sepsis in ICU patients: the CAPTAIN prospective multicenter cohort study. *Intensive Care Med* 2018;**44**:1061–70.
23. Duplessis C, Gregory M, Frey K, Bell M, Truong L, Schully K, Lawler J, Langley RJ, Kingsmore SF, Woods CW, Rivers EP, Jaehne AK, Quackenbush EB, Fowler VG, Tsalik EL, Clark D. Evaluating the discriminating capacity of cell death (apoptotic) biomarkers in sepsis. *J Intensive Care* 2018;**6**:72.
24. Tang BM, Huang SJ, McLean AS. Genome-wide transcription profiling of human sepsis: a systematic review. *Crit Care* 2010;**14**:R237.
25. McHugh L, Seldon TA, Brandon RA, Kirk JT, Rapisarda A, Sutherland AJ, Presneill JJ, Venter DJ, Lipman J, Thomas MR, Klein Klouwenberg PM, van Vught L, Scicluna B, Bonten M, Cremer OL, Schultz MJ, van der Poll T, Yager TD, Brandon RB. A molecular host response assay to discriminate between sepsis and infection-negative systemic inflammation in critically ill patients: discovery and validation in independent cohorts. *PLoS Med* 2015;**12**:e1001916.
26. Scicluna BP, Klein Klouwenberg PM, van Vught LA, Wiewel MA, Ong DS, Zwinderman AH, Franitza M, Toliat MR, Nürnberg P, Hoogendijk AJ, Horn J, Cremer OL, Schultz MJ, Bonten MJ, van der Poll T. A molecular biomarker to diagnose community-acquired pneumonia on intensive care unit admission. *Am J Respir Crit Care Med* 2015;**192**:826–35.
27. Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time–course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci Transl Med* 2015;**7**:287 ra71.
28. Burnham KL, Davenport EE, Radhakrishnan J, Humburg P, Gordon AC, Hutton P, Svoren-Jabalera E, Garrard C, Hill AVS, Hinds CJ, Knight JC. Shared and distinct aspects of the sepsis transcriptomic response to fecal peritonitis and pneumonia. *Am J Respir Crit Care Med* 2017;**196**:328–39.
29. Miller JM, Binnicker MJ, Campbell S, Carroll KC, Chapin KC, Gilligan PH, Gonzalez MD, Jerris RC, Kehl SC, Patel R, Pritt BS, Richter SS, Robinson-Dunn B, Schwartzman JD, Snyder JW, Telford 3rd S, Theel ES, Thomson Jr RB, Weinstein MP, Yao JD. A guide to utilization of the microbiology laboratory for diagnosis of infectious diseases: 2018 update by the infectious diseases society of america and the american society for microbiology. *Clin Infect Dis* 2018;**67**:e1–e94.
30. Cui J, Chen Y, Chou WC, Sun L, Chen L, Suo J, Ni Z, Zhang M, Kong X, Hoffman LL, Kang J, Su Y, Olman V, Johnson D, Tench DW, Amster IJ, Orlando R, Puett D, Li F, Xu Y. An integrated transcriptomic and computational analysis for biomarker identification in gastric cancer. *Nucleic Acids Res* 2011;**39**:1197–207.
31. Diao H, Li X, Hu S, Liu Y. Gene expression profiling combined with bioinformatics analysis identify biomarkers for Parkinson disease. *PLoS ONE* 2012;**7**:e52319.
32. Zhao Y, Tan W, Sheng W, Li X. Identification of biomarkers associated with Alzheimer's disease by bioinformatics analysis. *Am J Alzheimers Dis Other Demen* 2016;**31**:163–8.
33. Palmieri O, Creanza TM, Bossa F, Latiano T, Corritore G, Palumbo O, Martino G, Biscaglia G, Scimeca D, Carella M, Ancona N, Andriulli A, Latiano A. Functional implications of microRNAs in Crohn's disease revealed by integrating microRNA and messenger RNA expression profiling. *Int J Mol Sci* 2017;**18**:1580.
34. Almansa R, Heredia-Rodríguez M, Gomez-Sanchez E, Andaluz-Ojeda D, Iglesias V, Rico L, Ortega A, Gomez-Pesquera E, Liu P, Aragón M, Eiros JM, Jiménez–Sousa MÁ, Resino S, Gómez-Herreras I, Bermejo-Martín JF, Tamayo E. Transcriptomic correlates of organ failure extent in sepsis. *J Infect* 2015;**70**:445–56.
35. Sweeney TE, Wong HR. Risk stratification and prognosis in sepsis: what have we learned from microarrays? *Clin Chest Med* 2016;**37**:209–18.
36. Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, Rautanen A, Gordon AC, Garrard C, Hill AV, Hinds CJ, Knight JC. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *Lancet Respir Med* 2016;**4**:259–71.
37. Sweeney TE, Perumal TM, Henao R, Nichols M5, Howrylak JA, Choi AM, Bermejo-Martin JF, Almansa R, Tamayo E, Davenport EE, Burnham KL, Hinds CJ, Knight JC, Woods CW, Kingsmore SF, Ginsburg GS, Wong HR, Parnell GP, Tang B, Moldawer LL, Moore FE, Omberg L, Khatri P, Tsalik EL, Mangravite LM, Langley RJ. A community approach to mortality prediction in sepsis via gene expression analysis. *Nat Commun* 2018;**9**:694.
38. Parnell GP, Tang BM, Nalos M, Armstrong NJ, Huang SJ, Booth DR, McLean AS. Identifying key regulatory genes in the whole blood of septic patients to monitor underlying immune dysfunctions. *Shock* 2013;**40**:166–74.
39. Rowland T, Hilliard H, Barlow G. Procalcitonin: potential role in diagnosis and management of sepsis. *Adv Clin Chem* 2015;**68**:71–86.
40. Yunus I, Fasih A, Wang Y. The use of procalcitonin in the determination of severity of sepsis, patient outcomes and infection characteristics. *PLoS ONE* 2018;**13**:e0206527.
41. Zhang H, Wang X, Zhang Q, Xia Y, Liu D. Comparison of procalcitonin and high-sensitivity C-reactive protein for the diagnosis of sepsis and septic shock in the oldest old patients. *BMC Geriatr* 2017;**17**:173.

42. Tan M, Lu Y, Jiang H, Zhang L. The diagnostic accuracy of procalcitonin and C-reactive protein for sepsis: A systematic review and meta-analysis. *J Cell Biochem* 2019;**120**:5852–9.
43. Hosmer DW, Lemeshow S Assessing the fit of the model. In: *Applied logistic regression*. John Wiley and Sons; 2000. p. 160–4.
44. Jenner RG, Young RA. Insights into host responses against pathogens from transcriptional profiling. *Nat Rev Microbiol* 2005;**3**:281–94.
45. Sweeney TE, Wong HR, Khatri P. Robust classification of bacterial and viral infections via integrated host gene expression diagnostics. *Sci Transl Med* 2016;**8**:346 ra91.
46. Boldrick JC, Alizadeh AA, Diehn M, Dudoit S, Liu CL, Belcher CE, Botstein D, Staudt LM, Brown PO, Relman DA. Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc Natl Acad Sci USA* 2002;**99**:972–7.