SHOCK, Vol. 58, No. 1, pp. 20–27, 2022

OPEN

EARLY DIFFERENTIATION BETWEEN SEPSIS AND STERILE INFLAMMATION VIA URINARY GENE SIGNATURES OF METABOLIC DYSREGULATION

Sabyasachi Bandyopadhyay,[*†] Tyler J. Loftus,[*‡] Ying-Chih Peng,[§‖] Maria-Cecilia Lopez,[¶] Henry V. Baker,[¶] Mark S. Segal,[§] Kiley Graim,[*,#] Tezcan Ozrazgat-Baslanti,[*§] Parisa Rashidi,[*†] and Azra Bihorac[*§]

**Intelligent Critical Care Center, University of Florida, Gainesville, Florida; †J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, Florida; ‡Department of Surgery, University of Florida, Gainesville, Florida; §Division of Nephrology, Hypertension and Renal Transplantation, Department of Medicine, University of Florida, Gainesville, Florida; ‖Department of Industrial and Systems Engineering, University of Florida, Gainesville, Florida; ¶Department of Molecular Genetics and Microbiology, College of Medicine, University of Florida, Gainesville, Florida; and #Department of Computer and Information Science and Engineering, University of Florida, Gainesville, Florida*

Received 4 Mar 2022; first review completed 20 Apr 2022; accepted in final form 26 May 2022

ABSTRACT—**Objective:** The aim of this study was to characterize early urinary gene expression differences between patients with sepsis and patients with sterile inflammation and summarize in terms of a reproducible sepsis probability score. **Design:** This was a prospective observational cohort study. **Setting:** The study was conducted in a quaternary care academic hospital. **Patients:** One hundred eighty-six sepsis patients and 78 systemic inflammatory response syndrome (SIRS) patients enrolled between January 2015 and February 2018. **Interventions:** Whole-genome transcriptomic analysis of RNA was extracted from urine obtained from sepsis patients within 12 hours of sepsis onset and from patients with surgery-acquired SIRS within 4 hours after major inpatient surgery. **Measurements and Main Results:** We identified 422 of 23,956 genes (1.7%) that were differentially expressed between sepsis and SIRS patients. Differentially expressed probes were provided to a collection of machine learning feature selection models to identify focused probe sets that differentiate between sepsis and SIRS. These probe sets were combined to find an optimal probe set (UrSepsisModel) and calculate a urinary sepsis score (UrSepsisScore), which is the geometric mean of downregulated genes subtracted from the geometric mean of upregulated genes. This approach summarizes the expression values of all decisive genes as a single sepsis score. The UrSepsisModel and UrSepsisScore achieved area under the receiver operating characteristic curves 0.91 (95% confidence interval, 0.86–0.96) and 0.80 (95% confidence interval, 0.70–0.88) on the validation cohort, respectively. Functional analyses of probes associated with sepsis demonstrated metabolic dysregulation manifest as reduced oxidative phosphorylation, decreased amino acid metabolism, and decreased oxidation of lipids and fatty acids. **Conclusions:** Whole-genome transcriptomic profiling of urinary cells revealed focused probe panels that can function as an early diagnostic tool for differentiating sepsis from sterile SIRS. Functional analysis of differentially expressed genes demonstrated a distinct metabolic dysregulation signature in sepsis.

KEYWORDS—Gene expression, machine learning, RNA, sepsis, SIRS, urine

INTRODUCTION

Sepsis is a time-sensitive condition associated with significant mortality, morbidity, and health care costs, especially when the diagnosis is delayed. Clinicians often fail to accurately differentiate between sepsis and a sterile systemic inflammatory response syndrome (SIRS) among patients who incur sterile tissue damage from major surgery (1). Sepsis is driven by a dysregulated host response to pathogens; sterile SIRS is driven primarily by tissue damage (2). Sepsis and SIRS converge on similar systemic inflammation pathways, leading to similar clinical manifestations and diagnostic uncertainty (3,4). Early accurate differentiation between sepsis and SIRS has important implications for patient outcomes. For sepsis, failure to initiate early antibiotic therapy and intravenous fluid resuscitation is associated with increased mortality; for SIRS, administering unnecessary antibiotics and large intravenous fluid volumes are associated with multidrug-resistant infections and increased mortality (5–7).

Address reprint requests to Azra Bihorac, MD, MS, Division of Nephrology, Hypertension, and Renal Transplantation, Department of Medicine, PO Box 100224, Gainesville, FL 32610. E-mail: abihorac@ufl.edu

Conflicts of Interest and Source of Funding: A.B. and T.O.-B. were supported by R01 GM110240 from the National Institute of General Medical Sciences. A.B. T.O.-B., M.-C.L., H.V.B., and M.S.S. were supported by Sepsis and Critical Illness Research Center Award P50 GM-111152 from the National Institute of General Medical Sciences. A.B. and P.R. were supported by R01 GM110240 from the National Institute of General Medical Sciences (NIH/NIGMS), by 1R01EB029699 and 1R21EB027344 from the National Institute of Biomedical Imaging and Bioengineering (NIH/NIBIB), and 1R01NS120924 from the National Institute of Neurological Disorders and Stroke (NIH/NINDS). P.R. was supported by the National Science Foundation CAREER award 1750192. A.B. was supported by UF Research AWD09458. T.O.-B. was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health grant K01 DK120784 and by UF Research AWD09459 and the Gatorade Trust, University of Florida. T.J.L. was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award K23 GM140268. The research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under University of Florida Clinical and Translational Science Awards UL1TR000064 and UL1TR001427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The authors report no conflicts of interest.

Author Contributions: A.B. and M.S.S. conceived the original idea of the study and sought and obtained funding. A.B. and T.O.-B. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the analysis. S.B., Y.-C.P., M.-C.L., and T.O.-B. performed the analyses. S.B., A.B., Y.-C.P., T.J.L., T.O.-B., and P.R. were involved in interpreting the analyses. The article was written by S.B., T.J.L., A.B., and T.O.-B. with input from all coauthors. All authors participated in critically revising the manuscript for important intellectual content and gave final approval of the version to be published. A.B., P.R., and T.O.-B. served as senior authors. A.B. is guarantor for this article.

Supplemental digital content is available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.shockjournal.com).

DOI: 10.1097/SHK.0000000000001952

Given the importance of early recognition of sepsis and inherent difficulties in differentiating between sepsis and postoperative SIRS using clinical criteria alone, it may be helpful to incorporate gene signatures in diagnostic tests. Transcriptomic profiling of whole blood has been used to understand pathophysiologic mechanisms of sepsis and sterile SIRS (8,9). Blood-based gene microarrays have demonstrated efficacy in differentiating sepsis from SIRS (10,11).

Besides blood samples, urine is often collected from critically ill patients as standard clinical care and may offer unique insights regarding inflammatory disease etiologies. We have previously demonstrated that whole-genome transcriptomic profiling of urinary cellular mRNA is different between sepsis and noninfected controls (12). However, the same has not been done between sepsis and SIRS where the true clinical utility lies. Therefore, the diagnostic utility of urine sepsis biomarkers in distinguishing sepsis from SIRS has not been established.

In this prospective observational study of SIRS and sepsis patients, we tested the hypothesis that machine learning feature selection from whole-genome transcriptomic urinary RNA signatures can identify gene expression patterns that differentiate between sepsis and sterile SIRS within 12 hours of sepsis onset.

MATERIALS AND METHODS

Participants

Sepsis patients were prospectively recruited between January 2015 and August 2017 from a prospective longitudinal cohort of surgical sepsis patients at UF Health Shands Hospital (NCT02276066); sterile SIRS patients were prospectively recruited between July 2015 and February 2018 from a prospective observational study of patients undergoing cardiac or vascular surgery at the same hospital (NCT02114138) (see Figure, Supplemental Digital Content 1, http://links.lww.com/SHK/B473, which shows the cohort study diagram). The study protocols were finalized, and institutional review board approvals were obtained (IRB201400611 and IRB201400127) before recruiting patients (13). All study participants provided written informed consent. There was no overlap of patients between the two cohorts. This study complied with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guidelines for observational studies (14).

Inclusion criteria for sepsis patients were intensive care unit (ICU) admission, age ≥18 years, and clinical adjudication of sepsis by an attending physician according to the American College of Chest Physicians consensus criteria (15), with subsequent initiation of a computerized sepsis protocol, as previously described (16). Sepsis was initially diagnosed using a modified version of the Modified Early Warning Score–Sepsis Recognition Score (17), which factors in the temperature, heart rate, respiratory rate, blood pressure, and level of consciousness of the patient. Patients who had been identified by the Modified Early Warning Score screening protocol were subsequently assessed directly by a physician or advanced practice provider for bedside clinical adjudication of the presence of sepsis. Patients who were diagnosed with sepsis were started on a computerized sepsis protocol that maps clinical workflows and recommendations to patient physiology and clinical interventions. The sepsis protocol was developed by a multidisciplinary team of surgeons, intensivists, advanced practitioners, nurses, respiratory therapists, pharmacists, pathologists, and computer engineers, based on Surviving Sepsis Campaign guidelines (13). The sepsis assessment protocol is described in more detail in Supplementary Methods (see Text, Supplemental Digital Content 2, http://links.lww.com/SHK/B474, which has the detailed methodology). Patients were excluded if they had urinary tract infection (UTI) as the primary source of sepsis, end-stage renal disease, advanced liver disease, pancreatic disease, or heart disease (see Figure, Supplemental Digital Content 1, http://links.lww.com/SHK/B473, which shows the cohort study diagram). All patients in the SIRS cohort underwent inpatient cardiac or vascular surgery and had no infection before surgery. All patients in the SIRS cohort satisfied the SIRS criteria; that is, they met at least two of the four criteria described by Bone (18).

All clinical data were collected prospectively. Disease severity was measured by Sequential Organ Failure Assessment (SOFA) scores. Clinical patient outcomes, including in-hospital and 12-month mortality, were prospectively collected for both cohorts. For sepsis patients, urine was collected within 12 hours of sepsis onset; for SIRS patients, urine was collected within 4 hours after the end of surgery. A subset of subjects in these cohorts has been used in a previous research study (12).

Discovery and validation cohorts

The discovery cohort includes 145 sepsis patients and 39 SIRS patients prospectively recruited between January 2015 and February 2017. The validation cohort includes 41 sepsis patients and 39 SIRS patients recruited between February 2016 and February 2018. Cohort sample sizes ensure that at least 85% of the probes have power greater than 80% to detect a twofold change between the mean RNA expressions of sepsis and SIRS patients, using a two-sided, independent *t* test with Bonferroni corrections. A subset of the sepsis cohort has been used in a prior research study (12).

Processing and purification of urine samples

Previously described protocols were used to separate cell pellets from urine supernatant and to isolate total cellular RNA from the cellular urine pellet (12). These methods are elaborated in the Supplementary Methods (see Text, Supplemental Digital Content 2, http://links.lww.com/SHK/B474, which has the detailed methodology). The quantity (absorbance at 260 nm) and purity (ratio of the absorbance at 260 and 280 nm) of RNA isolated from the urine cell pellet were measured using the Take3 Multi-Volume Plate and Synergy HT Multi-Detection Microplate Reader (BioTek, Winooski, Vermont, USA). An RNA sample passed quality control if the optical density 260-to-280 ratio was between 1.5 and 2.2, and the final concentration was at least 8.7 μg/mL (see Table, Supplemental Digital Content 3, http://links.lww.com/SHK/B475, which shows quality, quantity, and concentration of RNA samples) (19). Figure 1A shows how the urine containing immune cells and pathogen/damage-associated molecular patterns was centrifuged and separated into cellular mRNA, exosomal mRNA, and metabolites. In this study, the cellular mRNA was used to find an early sepsis signature.

Microarrays

Biotin-labeled sense strand complementary DNA was prepared and was hybridized to GeneChip Human Transcriptome Array (HTA 2.0) array (catalog no. 902162; Affymetrix, Santa Clara, California, USA) using previously described protocols further described in Supplementary Methods (see Text, Supplemental Digital Content 2, http://links.lww.com/SHK/B474, which has the detailed methodology) (12). Furthermore, array scanning, image analysis, and probe quantification were performed using previous protocols as described in Supplementary Methods (see Text, Supplemental Digital Content 2, http://links.lww.com/SHK/B474, which has the detailed methodology). Transcriptome Analysis Console (TAC) version 4.0.1 (Thermo Fisher Scientific, Santa Clara, California, USA) was used for microarray signal summarization and normalization using robust multiarray averages (Fig. 1B) (20). The final microarray data set consisted of log2-transformed expression values for 67,528 probes, of which 33,494 were mapped to at least one gene. Raw and normalized expression data are available under GSE168443, GSE168442, and GSE168440, GEO series accessions.

Identification of cell-specific transcripts

The 33,494 probes mapped to genes were used to estimate the samples' immune and kidney cell composition (Fig. 1B). The immune response *in silico* (IRIS) repository of 1622 genes, classified by their specific expression in multiple immune cell lineages and previously described transcript sets of 637 genes for kidney-specific cell lineages, were used to estimate the immune and renal cell composition in urine (21,22), respectively.

Identifying genes that discriminate between sepsis and SIRS

We applied an empirical Bayes method in LIMMA (Linear Models for Microarray Analysis) to identify probes that differentiate between sepsis and SIRS (23). The significance threshold was adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR). Probes were considered differentially expressed if they had an FDR-adjusted probability of $Q \leq 0.01$ and an absolute fold change ≥2. Gene expression patterns were illustrated using Euclidean distance heatmaps with ComplexHeatmap. Sepsis endotypes were explored by projecting the differentially expressed probes onto a two-dimensional t-stochastic neighbor embedded (t-SNE) manifold and labeling by sepsis severity, primary diagnosis, endotype, mortality, sepsis onset time, and demographics (age, sex, race). Ingenuity Pathway Analysis (IPA) software (http://www.ingenuity.com) was used to identify significantly enriched biological functions, pathways, molecular networks, and regulatory molecules associated with the differentially expressed genes (Fig. 1B).

Feature selection

The differentially expressed probes were subjected to feature selection using random forest, recursive feature elimination using support vector machine (SVM)

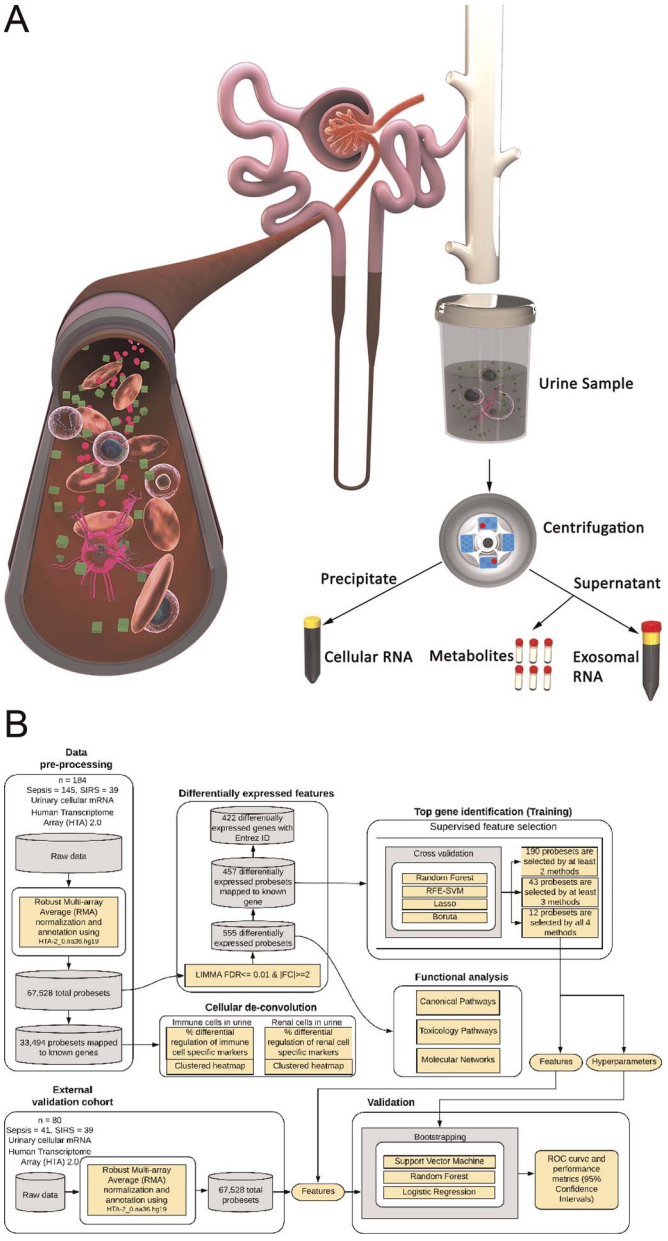

FIG. 1. **Workflow.** A, Process flow for isolation of urinary markers. B, Conceptual workflow from data acquisition to analysis. Panel A is adapted from Bandyopadhyay et al. (12), 2020; Copyright © 2020 The Authors. Published by Wolters Kluwer Health, Inc., on behalf of the Society of Critical Care Medicine.

classifier, logistic regression with Lasso, and Boruta (24) (Fig. 1B) machine learning techniques to generate four different lists of selected features. Here, random forest selected 100 features, Boruta selected 30 features, recursive feature elimination SVM selected 300 features, and Lasso selected 157 features. Details on feature selection methodology are provided in Supplementary Methods (see Text, Supplemental Digital Content 2, http://links.lww.com/SHK/B474, which has the detailed methodology). All machine learning methods except for Boruta were parametrized inside a fivefold cross-validation design. Finally, a simple voting strategy was used to identify features appearing in two, three, or all four of the four feature lists.

UrSepsis model validation

The final list of probes was validated using the microarrays of 41 sepsis patients and 39 SIRS patients, normalized separately from the discovery cohort. We used each of the three feature selection techniques that were trained and tuned with discovery cohort cross-validation (i.e., support vector machine, random forest, and logistic regression) to calculate the following performance metrics: area under the receiver operating characteristics curve (AUC), accuracy, F1 score, sensitivity, specificity, and positive predictive value. Ninety-five percent confidence intervals (CIs) for each performance metric in each model were estimated by bootstrapping the validation cohort without replacement 100 times. These methods compose the UrSepsis model.

UrSepsisScore calculation

The UrSepsisScore was calculated by the geometric mean of downregulated genes subtracted from the geometric mean of upregulated genes, similar to the sepsis score developed by Sweeney and Khatri (25). The UrSepsisScore summarizes expression values of decisive genes in a single value. We determined the threshold for partitioning between sepsis and SIRS patients by maximizing AUC on the discovery cohort. We then evaluated UrSepsisScore performance in the independent validation cohort. Bioconductor (version 3.7, Buffalo, New York, USA) in R (version 3.4.2, Vienna, Austria) and scikit-learn (version 0.19.2, Paris, France) in Python (version 3.8, Fredericksburg, Virginia, USA) were used in this project.

RESULTS

Patient characteristics

Compared with sepsis patients, SIRS patients in the validation cohort had advanced age (65 vs. 55 years) and a greater proportion of current or former smokers (67% vs. 41%). Comorbidities were similar between sepsis and SIRS patients (Table 1). Among SIRS patients, all urine samples were obtained within 4 hours after the end of surgery. Among sepsis patients, all samples were collected within 12 hours of sepsis onset. The median interval between sepsis onset and urine collection was 7 hours. Patients who had UTI as the primary cause of sepsis were excluded from the study as they had a significantly larger RNA load, which is indicative of a higher cell count in their urine. A single-tailed *t* test carried out between RNA quantities of sepsis patients with and without UTI as their primary cause of infection showed a $P = 0.008$, assuming equal variance. Their variances were found to be identical by an *F* value of 2.41E−6, where the *F* critical was 1.714. A sensitivity analysis showed that including patients with UTI in the sepsis cohort resulted in 10% more differentially expressed probes. No samples were excluded from the study based on their RNA quantity as they all passed the quality-control step in Transcriptome Analysis Console analysis. At the time of urine sampling, sepsis and SIRS cohorts had similar SOFA scores. As expected, white blood cell counts were higher in sepsis patients. None of the SIRS patients developed sepsis within 7 days of surgery. Sepsis patients had longer median lengths of stay in the ICU (4–5 days longer) and hospital (8 days longer). The different surgical procedures that SIRS patients underwent are presented in the Supplementary (see Table, Supplemental Digital Content 4, http://links.lww.com/SHK/B476, which lists surgical procedures undergone by SIRS subjects).

The acute urinary molecular response to sepsis

We identified a distinct urinary transcriptomic profile including 422 candidates out of 23,956 genes (1.7%) that were differentially expressed between sepsis and SIRS patients, defined here as FDR ≤1% and absolute fold change ≥2 (Fig. 2, A and B). Principal component analysis showed that the sepsis profile was distinct from the SIRS profile, as illustrated in Figure 2C, despite there being some overlap between the two groups. Patients with SIRS were in the bottom right of a two-dimensional plot between principal component (PC) 1 and PC 2, and the top-right part in a PC 1 versus PC 3 plot (see Figure, Supplemental Digital Content 5, http://links.

TABLE 1. **Clinical characteristics of patients in discovery and validation cohorts**

	Discovery cohort			Validation cohort		
Variables	Sepsis patients (n = 145)	SIRS patients (n = 39)	*P*	Sepsis patients (n = 41)	SIRS patients (n = 39)	*P*
Baseline characteristics						
Female sex, n (%)	67 (46)	14 (36)	0.279	17 (41)	17 (44)	1
Age, mean (SD), y	59 (15)	70 (10)	**<0.001**	55 (18)	65 (11)	**0.004**
Age ≥65 y, n (%)	55 (38)	30 (77)	**<0.001**	16 (39)	23 (59)	0.117
Race, n (%)			**0.036**			0.553
White	130 (90)	34 (87)		37 (90)	33 (85)	
African American	12 (8)	1 (3)		4 (10)	4 (10)	
Other	3 (2)	4 (10)		0 (0)	2 (5)	
BMI, median (25th, 75th)	29 (25, 34)	25 (22, 34)	**0.049**	29 (25, 40)	29 (25, 33)	0.283
Comorbidities, n (%)						
Chronic kidney disease	19 (13)	6 (15)	0.793	6 (15)	12 (31)	0.178
Hypertension*	102 (70)	30 (77)	0.548	29 (71)	33 (85)	0.183
Diabetes†	43 (30)	8 (21)	0.316	9 (22)	13 (33)	0.319
Chronic pulmonary disease	51 (35)	15 (38)	0.71	9 (22)	14 (36)	0.219
Congestive heart failure	23 (16)	8 (21)	0.478	6 (15)	11 (28)	0.176
Current or former smoker*	74 (51)	33 (85)	**<0.001**	17 (41)	26 (67)	**0.027**
Acuity at the time of sampling						
SOFA score, median (25th, 75th)	6 (3, 8)	4 (2, 8)	**0.042**	6 (3, 7)	6 (4, 8)	0.882
Primary sepsis source, n (%)						
Intra-abdominal sepsis	61 (42)	NA		18 (44)	NA	
Pneumonia	31 (21)	NA		8 (20)	NA	
Necrotizing soft tissue infection	26 (18)	NA		7 (17)	NA	
Surgical site infection	19 (13)	NA		1 (2)	NA	
UTI	0 (0)	NA		0 (0)	NA	
Other†	8 (6)	NA		7 (17)	NA	
Sepsis stage on enrollment, n (%)						
Sepsis/severe sepsis	112 (77)	NA		33 (80)	NA	
Septic shock	33 (23)	NA		8 (20)	NA	
Lactate, median (25th, 75th), mmol/L	1.8 (1.3, 2.9)	2.1 (1.3, 3.4)	0.612	1.7 (1.2, 2.5)	2.8 (1.8, 6)	**0.002**
Serum creatinine, median (25th, 75th), mg/dL	1.0 (0.7, 1.5)	1.1 (0.9, 1.4)	0.31	1.1 (0.9, 1.7)	1.0 (0.9, 1.3)	0.252
Urea nitrogen, median (25th, 75th), mg/dL	19 (12, 32)	16 (14, 21)	0.094	24 (17, 36)	19 (13, 24)	**0.013**
White blood cell count, median (25th, 75th), $\times 10^3/\mu L$	17 (12, 22)	13 (9, 15)	**<0.001**	19 (14, 26)	15 (11, 18)	**0.007**
Hematocrit, median (25th, 75th), %	27 (23, 32)	27 (25, 32)	0.408	26 (24, 31)	26 (23, 29)	0.229
Outcomes						
Hospital mortality, n (%)	11 (8)	1 (3)	0.466	6 (15)	0 (0)	**0.026**
Discharge to home, n (%)	72 (50)	27 (69)	**0.032**	17 (41)	25 (64)	**0.048**
ICU LOS, median (25th, 75th), d	8 (4, 18)	5 (4, 8)	0.056	10 (5, 15)	5 (3, 11)	0.064
ICU ≥14 d, n (%)	49 (34)	6 (15)	**0.03**	16 (39)	6 (15)	**0.024**
Hospital LOS, median (25th, 75th), d	18 (9, 28)	10 (6, 18)	**0.009**	17 (11, 30)	9 (6, 13)	**<0.001**

Significance level is set to be 0.05.
Boldface values represent statistical significance.
*Percentages are calculated based on available values due to missing values.
†Other primary sepsis source includes catheter-related bloods, empyema, bacteremia, and esophageal perforation.

lww.com/SHK/B477, which separately shows the comparison of PC 1 with PC 2 and PC 3). The t-SNE projections did not result in any clustering within the sepsis cohort, indicating the absence of any definite sepsis endotypes (see Figure, Supplemental Digital Content 6, http://links.lww.com/SHK/B478, which shows two-dimensional t-SNE projections of differentially expressed genes of sepsis patients colored by different disparities). Ingenuity Pathway Analysis functional analyses showed downregulation of functional pathways related to amino acid metabolism, lipid metabolism, and energy production in sepsis. Among the top 20 pathways, 19 were underexpressed in sepsis, and 1 pathway was overexpressed. The five most underexpressed pathways in sepsis were valine degradation, fatty acid β-oxidation, phenylalanine degradation, oxidative phosphorylation, and the TCA (tricarboxylic acid) cycle. Conversely, LPS/IL-1–mediated inhibition of RXR function was upregulated in sepsis (Fig. 3A). Ingenuity Pathway Analysis multiorgan toxicology analysis revealed that sepsis patients had significantly higher cell death in the liver, notably associated with steatosis (Fig. 3B), and higher cell death in the kidney, especially in the renal tubule. Hepatocyte nuclear factor 4α, peroxisome proliferator–activated receptor α, hepatocyte nuclear factor 1α, and LIM homeobox 1 were key upstream regulator molecules (see Table, Supplemental Digital Content 7, http://links.lww.com/SHK/B479, which lists important upstream regulator molecules). Information about the specific genes present in each of the IPA functional pathways is available (see Table, Supplemental Digital Content 8, http://links.lww.com/SHK/B480, which gives additional information about genes in functional pathways). Information about specific genes in the toxicology pathways is provided (see Table, Supplemental Digital Content 9, http://links.lww.com/SHK/B481, which gives additional information about genes in toxicology pathways). The top coexpression network is amino acid metabolism, which is

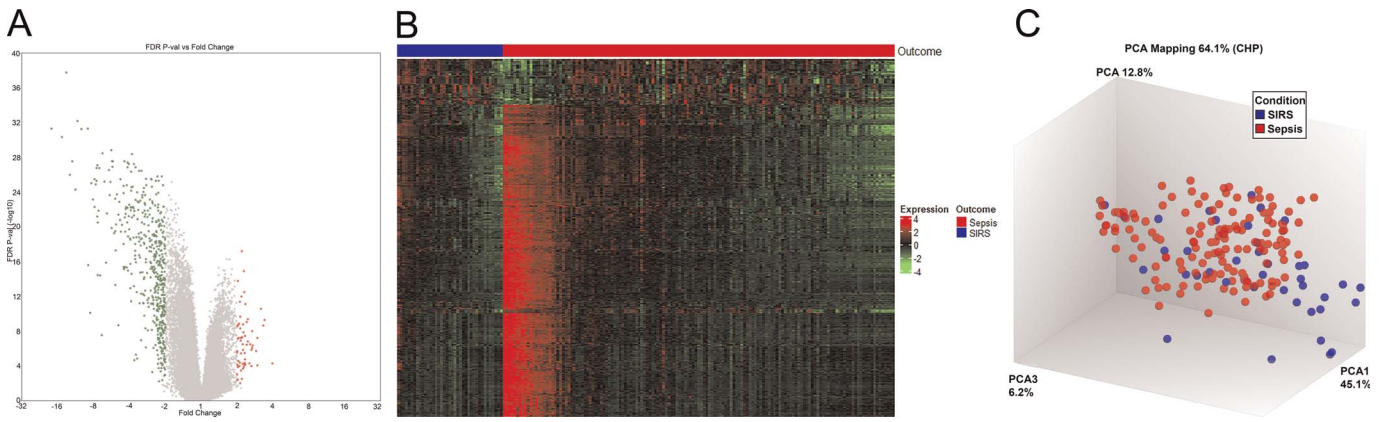

Fig. 2. **The early transcriptomic response to sepsis in the cells retrieved from the urine pellet.** A, Volcano plot demonstrates the degree of differential expression of 67,528 probes of Human Transcriptome 2.0 chip. A total of 555 (0.8%) of 67,528 probes were differentially expressed in the sepsis patients. Green dots indicate probes below the 0.01 FDR cutoff and with fold change ≤–2; red dots indicate probes below the 0.01 FDR cutoff and with fold change ≥2 between sepsis and control patients. B, Heatmap of expression values of 555 differentially expressed probes. The sepsis cohort is highlighted in red and the SIRS cohort is highlighted in blue. C, Principal component analysis of differentially expressed genes in acute phase of sepsis (within 12 hours) compared with SIRS patients. Sepsis patients are generally well separated from SIRS patients with a small proportion of SIRS patient overlapped with the sepsis cohort.

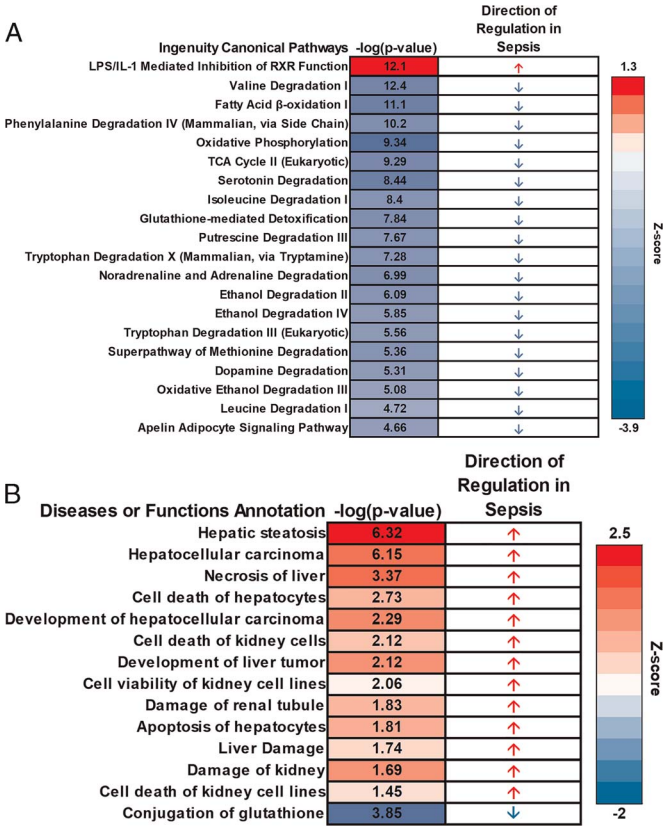

Fig. 3. **Pathways and biofunctions in the acute response to sepsis (within 12 hours of sepsis onset), compared with control patients.** A, Ingenuity Pathway Analysis of differentially expressed probes showed downregulation of pathways mainly related to amino acid metabolism, lipid metabolism, and cellular energy production in sepsis compared with SIRS. The only pathway that was significantly upregulated in sepsis is associated with IL-1–mediated inhibition of retinoid X receptors. *P* values are calculated by IPA using right-tailed Fisher exact test to measure likelihood that pathways or functions are overrepresented by molecules in data set. B, Ingenuity toxicology analysis of differentially expressed probes shows that hepatic steatosis is the top toxicological response of the dysregulated transcriptome followed by repeated occurrence of cell death in liver and in kidney. The only downregulated toxicology pathway in sepsis is conjugation of glutathione.

downregulated in sepsis (see Figure, Supplemental Digital Content 10, http://links.lww.com/SHK/B482, which demonstrates the most significant gene coexpression network). The other significantly different gene coexpression networks were cellular assembly and organization, energy production, both being downregulated in sepsis.

Immune and kidney cell–specific transcripts in the urine

Deconvolution analysis identified overall significant upregulation of marker genes for neutrophils and monocytes and an overall significant downregulation of T-cell marker genes (Fig. 4, A and B). Analysis of average expression of these marker genes revealed that whereas monocyte and neutrophil marker genes showed increased average expression in sepsis patients, average expression of T-cell markers was identical in the two groups (Fig. 4C). In the IRIS article, the authors described that their results create a reliable representation of cellular populations (21). Applying the same methodology to kidney lineage–specific cells, we found no significant difference in concentration of such cells in sepsis compared with SIRS patients (see Figure, Supplemental Digital Content 11, http://links.lww.com/SHK/B483, http://links.lww.com/SHK/B484, http://links.lww.com/SHK/B485, which shows cellular deconvolution of kidney-specific genes).

Sepsis diagnostic model using urinary molecular signature

Sepsis diagnostic model consisted of three probe sets (12, 43, and 190 probes) comprising probes common to all four, at least three, and at least two of the machine learning models described previously. The performance of each of the three probes sets was evaluated in an independent validation cohort comprising 41 sepsis and 39 SIRS patients. The best performance in the validation cohort was achieved using the set of 43 probes with support vector machines. This approach yielded AUC of 0.91 (95% CI, 0.86–0.96), accuracy of 0.82 (95% CI, 0.76–0.89), F1 score of 0.83 (95% CI, 0.75–0.89), sensitivity of 0.83 (95% CI, 0.74–0.92), and specificity of 0.83 (95% CI, 0.72–0.92). The set of 43 probes mapped to 30 genes (see Table, Supplemental Digital Content 12, http://links.lww.com/SHK/B486, which shows gene

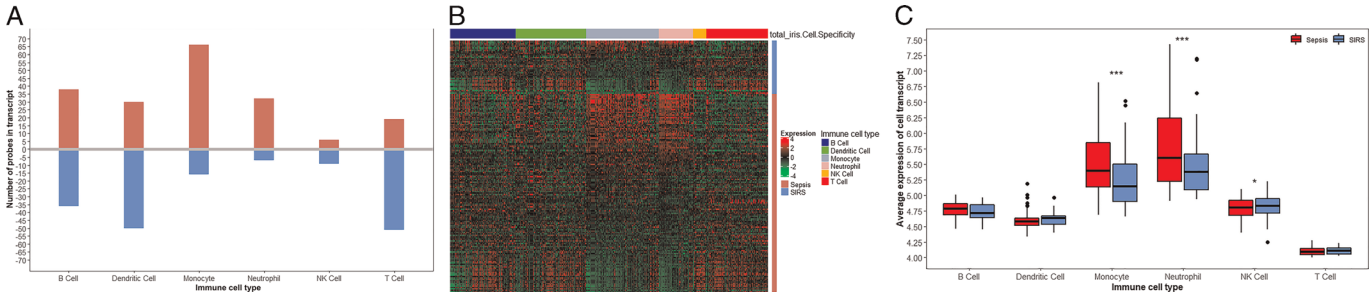

FIG. 4. **Immune cell–specific transcript changes in the acute response to sepsis.** A, Immune cell deconvolution showing the overall % differential regulation of immune cell–specific markers (selected from the 1622 genes from IRIS resource [see Materials and Methods]) between sepsis and control patients. There was predominant upregulation of neutrophil and monocyte markers; a mixed response in B cells, dendritic cells, and natural killer (NK) cells; and downregulation of T cells. B, Heatmap of immune cell–specific/enriched markers (selected from the 823 genes from IRIS resource) in the sepsis and control patients. Most of the signature genes in T cells and NK cells are underexpressed, and most of the signature genes in neutrophils and monocytes are overexpressed in sepsis compared with SIRS. C, Average expression of immune cell–specific transcripts confirmed the deconvolution results. There were increased numbers of neutrophils and monocytes in the acute sepsis window. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

annotation of 43 probes). Henceforth, this model is described as the UrSepsisModel. For each patient, we calculated a simplified urine sepsis score (UrSepsisScore) using the expression values of 43 probes that appeared in at least two of the gene sets. The score value that maximized AUC in the discovery cohort was selected as the threshold cutoff value. The UrSepsisScore above this threshold (0.80) showed good performance in the validation cohort: AUC of 0.80 (95% CI, 0.70–0.88), F1 score of 0.72 (95% CI, 0.55–0.83), sensitivity of 73% (95% CI, 0.42–0.99), and specificity of 68% (95% CI, 0.28–0.95). Figure 5 illustrates the application of the UrSepsisScore to a randomly selected bootstrap of the validation cohort. The values of multiple performance metrics for the three probe sets (12, 43, and 190 probe sets) using all candidate models are presented (see Table, Supplemental Digital Content 13, http://links.lww.com/SHK/B487, which summarizes performances of selected probe sets over all machine learning models).

DISCUSSION

These findings suggest that urinary gene signatures of metabolic dysregulation differentiate between sepsis and sterile SIRS within hours of disease onset. Functional analyses demonstrated metabolic dysregulation manifesting as reduced oxidative phosphorylation, decreased amino acid metabolism, and decreased oxidation of lipids and fatty acids. Machine learning modeling identified an optimal subset of features that simultaneously discriminate between sepsis and SIRS and preserve relevant nonlinear relationships among input features that arise from underlying pathophysiology and not from the chosen model or feature selection procedure. When the selected genes were aggregated to make the UrSepsisScore, good performance was retained. The UrSepsisScore can be generated without the use of any machine learning model, which is intended to facilitate clinical application. Because urine samples were obtained within 12 hours of sepsis onset or 4 hours of surgery, these metabolic signatures and sepsis classifications can be applied early after clinical manifestations of systemic inflammation, when diagnostic uncertainty is greatest.

In a previous study, we demonstrated that urine contains sufficient transcriptomic information to differentiate between septic patients and uninfected controls (12). However, the clinical challenge lies in distinguishing sepsis from uninfected SIRS, especially in surgical patients who incur tissue damage and may or may not have superimposed infection. Urine biomarkers have been previously used for identifying sepsis patients and differentiating them from SIRS patients. Su et al. (26) showed that urinary s-TREM1 can differentiate sepsis from SIRS patients with AUC of 0.797 (95% CI, 0.711–0.884). Later, Su et al. (27) demonstrated that urinary sCD163 could differentiate between sepsis and SIRS with AUC of 0.83 (95% CI, 0.72–0.94). Kustán et al. (28) reported that the ratio between urine orosomucoid and urine creatinine differentiates between severe sepsis and SIRS within 24 hours of sepsis diagnosis with AUC of 0.954, but used relatively small sample sizes (severe sepsis: n = 43, SIRS: n = 13, control: n = 30). In addition, these previous studies are subject to overfitting due to a lack of independent validation cohorts. Predictive performance in the validation cohort of the present study is greater than or equal to that of comparative assays for distinguishing sepsis from SIRS made using blood (10,11,29). A more elaborate comparison with existing diagnostic tests using urine was made in the Supplementary (see Text, Supplemental Digital Content 2, http://links.lww.com/SHK/B474, which compares current study with previous urinary sepsis diagnostic tests).

Beyond the potential advantages of urine gene expression as an early diagnostic tool, functional analysis elucidated underlying pathophysiologic signatures. In the present study, 12 of the top 20 functional biomolecular pathways were amino acid metabolic pathways. Urinary isoleucine, leucine, tryptophan, tyrosine, and valine degradation were downregulated. Decreased amino acid metabolism is well documented in severe sepsis, suggesting a therapeutic role for amino acid supplementation (30,31). The second most underexpressed metabolic pathway in our analysis was fatty acid β-oxidation, which is known to be significantly impaired in sepsis nonsurvivors (30,32). The most significant upregulated pathway was IL-1/LPS–mediated inhibition of RXR function, suggesting impaired metabolism, transport, and biosynthesis of lipids and bile acids (33,34). Collectively, these findings are consistent with observations by Langley et al. (32) that several fatty acids were upregulated in the blood of sepsis nonsurvivors and that steatosis of the liver, and damage of renal tubule with associated cell death occurs after the onset of sepsis as is outlined in this review (35). Therefore, the mechanisms of metabolic failure identified in this study are consistent with known pathophysiology of severe sepsis or in sepsis nonsurvivors but appear in the urine within 12 hours of sepsis diagnosis, which is much earlier than the time frames presented in the above blood-based studies.

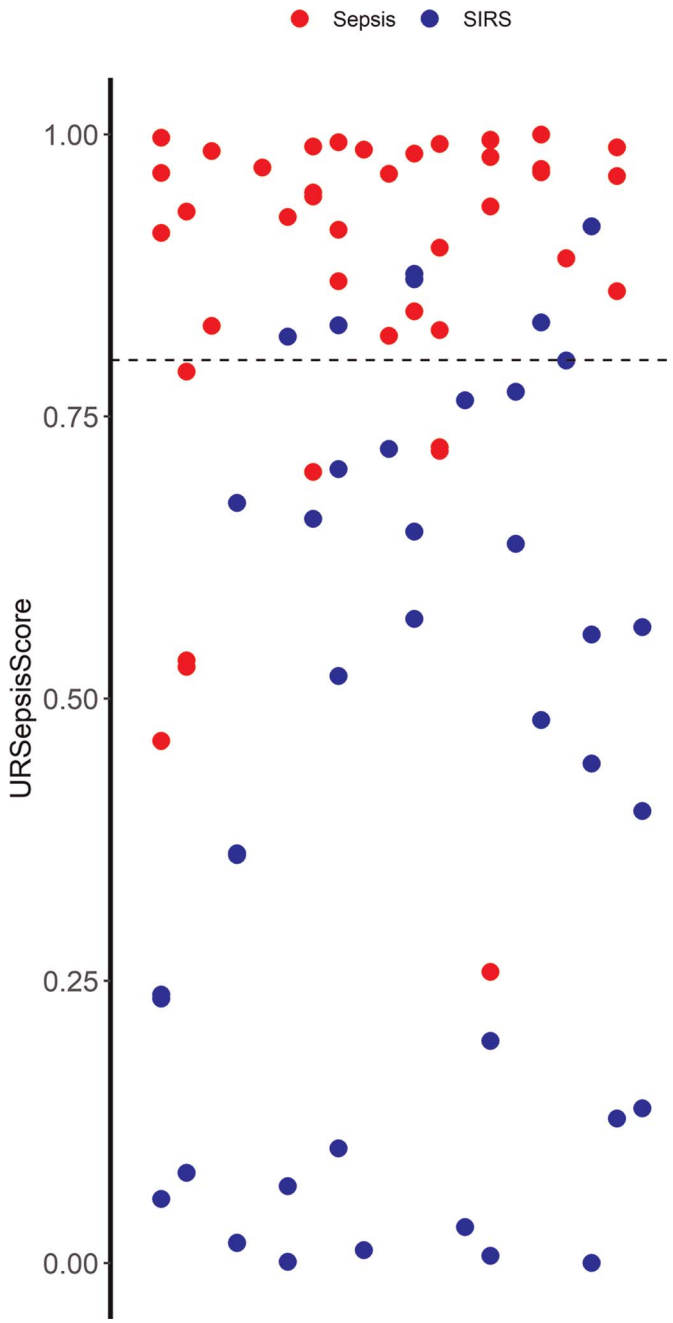

Fig. 5. **Scatterplot of UrSepsisScore values determined using 43 machine learning selected probes in external validation cohort.** UrSepsisScore values for the external validation cohort were calculated using expression values of 43 probes selected by consensus voting of the feature lists generated by our machine learning models. This score represents the geometric mean of downregulated genes subtracted from the geometric mean of upregulated genes. UrSepsisScore values for sepsis and SIRS patients are represented in this scatterplot using red and blue dots, respectively. UrSepsisScore values for sepsis patients show less variance than those of SIRS patients. The dashed line at UrSepsisScore = 0.8 is the Youden index, which is used as the threshold for classification. There are six false positives and seven false negatives based on this classification.

Further description of the functional and toxicology pathways is provided in the Supplementary (see Text, Supplemental Digital Content 2, http://links.lww.com/SHK/B474, which discusses functional and toxicology pathways and their implications).

This study was limited by the small sample size and limited generalizability due to its restrictive inclusion criteria, time constraints of obtaining urine samples, and single-institution design. A drawback to using a single-point estimate of sepsis probability (UrSepsisScore) is that the CIs are large. Further investigation is required to elucidate the role of small RNAs present in the data set and determine whether our methods could be augmented using urinary exosomal RNA retrieved from the supernatant after centrifugation. Finally, to improve predictive performance, gene expression signatures could be integrated with other modalities, such as metabolomics and clinical data.

CONCLUSIONS

Whole-genome transcriptomic analysis of urinary cells demonstrated metabolic dysregulation in sepsis relative to sterile SIRS. Machine learning models identified a stable, consistent, and focused probe set for differentiating sepsis from SIRS and validated its performance in an independent validation cohort. These probes are used to calculate UrSepsisScore, which uses geometric means to summarize the expression values of all decisive genes as a single sepsis probability score.

ACKNOWLEDGMENTS

The authors thank the staff of the University of Florida Sepsis and Critical Illness Research Center and Laboratory for Inflammation Biology and Surgical Science for their invaluable contributions, including Zhongkai Wang, Gabriela Ghita, Jennifer Lanz, Ruth Davis, Jillianne Brakenridge, Ashley McCray, Ricardo Ungaro, Dina Nacionales, Brian Bouverat, Marvin Dirain, Angela Avery, Tabitha Johns, and Ada Malcolm and PRISMA^{P} research group, including George Omalay, Matthew Ruppert, Nicole Hamza-Fields, and Victoria Hough. They acknowledge the University of Florida Integrated Data Repository and the UF Health Office of the Chief Data Officer for providing the analytic data set for this project.

REFERENCES

1. Vincent JL: The clinical challenge of sepsis identification and monitoring. *PLoS Med* 13(5):e1002022, 2016.
2. Gentile LF, Moldawer LL: DAMPs, PAMPs, and the origins of SIRS in bacterial sepsis. *Shock* 39(1):113–114, 2013.
3. Castellheim A, Brekke OL, Espevik T, Harboe M, Mollnes T: Innate immune responses to danger signals in systemic inflammatory response syndrome and sepsis. *Scand J Immunol* 69(6):479–491, 2009.
4. Lowry SF: The stressed host response to infection: the disruptive signals and rhythms of systemic inflammation. *Surg Clin North Am* 89(2):311–326, vii, 2009.
5. Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, Suppes R, Feinstein D, Zanotti S, Taiberg L, et al.: Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med* 34(6):1589–1596, 2006.
6. Shaw AD, Schermer CR, Lobo DN, Munson SH, Khangulov V, Hayashida DK, Kellum JA: Impact of intravenous fluid composition on outcomes in patients with systemic inflammatory response syndrome. *Crit Care* 19:334, 2015.
7. Timsit JF, Bassetti M, Cremer O, Daikos G, de Waele J, Kallil A, Kipnis E, Kollef M, Laupland K, Paiva JA: Rationalizing antimicrobial therapy in the ICU: a narrative review. *Intensive Care Med* 45(2):172–189, 2019.
8. Johnson SB, Lissauer M, Bochicchio GV, Moore R, Cross AS, Scalea TM: Gene expression profiles differentiate between sterile SIRS and early sepsis. *Ann Surg* 245(4):611–621, 2007.
9. Wong HR, Cvijanovich N, Allen GL, Lin R, Anas N, Meyer K, Freishtat RJ, Monaco M, Odoms K, Sakthivel B, et al.: Genomic expression profiling across the pediatric systemic inflammatory response syndrome, sepsis, and septic shock spectrum. *Crit Care Med* 37(5):1558–1566, 2009.
10. Sutherland A, Thomas M, Brandon RA, Brandon RB, Lipman J, Tang B, McLean A, Pascoe R, Price G, Nguyen T, et al.: Development and validation of a novel molecular biomarker diagnostic test for the early detection of sepsis. *Crit Care* 15(3): R149, 2011.
11. Tang BM, McLean AS, Dawes IW, Huang SJ, Lin RC: Gene-expression profiling of peripheral blood mononuclear cells in sepsis. *Crit Care Med* 37(3):882–888, 2009.
12. Bandyopadhyay S, Lysak N, Adhikari L, Velez LM, Sautina L, Mohandas R, Lopez MC, Ungaro R, Peng YC, Kadri F, et al.: Discovery and validation of urinary molecular signature of early sepsis. *Crit Care Explor* 2(10):e0195, 2020.

13. Loftus TJ, Mira JC, Ozrazgat-Baslanti T, Ghita GL, Wang Z, Stortz JA, Brumback BA, Bihorac A, Segal MS, Anton SD, et al.: Sepsis and critical illness research center investigators: protocols and standard operating procedures for a prospective cohort study of sepsis in critically ill surgical patients. *BMJ Open* 7(7):e015136, 2017.
14. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 370(9596):1453–1457, 2007.
15. American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med* 20(6):864–874, 1992.
16. Croft CA, Moore FA, Efron PA, Marker PS, Gabrielli PS, Westhoff LS, Lottenberg L, Jordan J, Klink V, Sailors RM, et al.: Computer versus paper system for recognition and management of sepsis in surgical intensive care. *J Trauma Acute Care Surg* 76(2):311–317; discussion 318–319, 2014.
17. Gardner-Thorpe J, Love N, Wrightson J, Walsh S, Keeling N: The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. *Ann R Coll Surg Engl* 88(6):571–575, 2006.
18. Bone RC: Toward an epidemiology and natural history of SIRS (systemic inflammatory response syndrome). *JAMA* 268(24):3452–3455, 1992.
19. Chomczynski P: A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques* 15(3):532–534, 536–537, 1993.
20. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249–264, 2003.
21. Abbas AR, Baldwin D, Ma Y, Ouyang W, Gurney A, Martin F, Fong S, van Lookeren Campagne M, Godowski P, Williams PM, et al.: Immune response *in silico* (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* 6(4):319–331, 2005.
22. Chabardès-Garonne D, Mejéan A, Aude JC, Cheval L, Di Stefano A, Gaillard MC, Imbert-Teboul M, Wittner M, Balian C, Anthouard V, et al.: A panoramic view of gene expression in the human kidney. *Proc Natl Acad Sci U S A* 100(23): 13710–13715, 2003.
23. Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3, 2004.
24. Kursa MB, Rudnicki WR: Feature selection with the Boruta package. *J Stat Softw* 36(11):1–13, 2010.
25. Sweeney TE, Khatri P: Benchmarking sepsis gene expression diagnostics using public data. *Crit Care Med* 45(1):1–10, 2017.
26. Su LX, Feng L, Zhang J, Xiao YJ, Jia YH, Yan P, Feng D, Xie LX: Diagnostic value of urine sTREM-1 for sepsis and relevant acute kidney injuries: a prospective study. *Crit Care* 15(5):R250, 2011.
27. Su L, Feng L, Liu C, Jiang Z, Li M, Xiao K, Yan P, Jia Y, Feng D, Xie L: Diagnostic value of urine sCD163 levels for sepsis and relevant acute kidney injury: a prospective study. *BMC Nephrol* 13:123, 2012.
28. Kustán P, Szirmay B, Horváth-Szalai Z, Ludány A, Kovács GL, Miseta A, Kőszegi T, Mühl D: Urinary orosomucoid: a novel, early biomarker of sepsis with promising diagnostic performance. *Clin Chem Lab Med* 55(2):299–307, 2017.
29. Wong HR, Cvijanovich NZ, Hall M, Allen GL, Thomas NJ, Freishtat RJ, Anas N, Meyer K, Checchia PA, Lin R, et al.: Interleukin-27 is a novel candidate diagnostic biomarker for bacterial infection in critically ill children. *Crit Care* 16(5):R213, 2012.
30. Arts RJW, Gresnigt MS, Joosten LAB, Netea MG: Cellular metabolism of myeloid cells in sepsis. *J Leukoc Biol* 101(1):151–164, 2017.
31. Su L, Li H, Xie A, Liu D, Rao W, Lan L, Li X, Li F, Xiao K, Wang H, et al.: Dynamic changes in amino acid concentration profiles in patients with sepsis. *PLoS One* 10(4):e0121933, 2015.
32. Langley RJ, Tsalik EL, van Velkinburgh JC, Glickman SW, Rice BJ, Wang C, Chen B, Carin L, Suarez A, Mohney RP, et al.: An integrated clinico-metabolomic model improves prediction of death in sepsis. *Sci Transl Med* 5(195):195ra95, 2013.
33. Standage SW, Caldwell CC, Zingarelli B, Wong HR: Reduced peroxisome proliferator–activated receptor α expression is associated with decreased survival and increased tissue bacterial load in sepsis. *Shock* 37(2):164–169, 2012.
34. Mandard S, Müller M, Kersten S: Peroxisome proliferator–activated receptor alpha target genes. *Cell Mol Life Sci* 61(4):393–416, 2004.
35. Garofalo AM, Lorente-Ros M, Goncalvez G, Carriedo D, Ballen-Barragan A, Villar-Fernandez A, Penuelas O, Herrero R, Granados-Carreno R, Lorente JA: Histopathological changes of organ dysfunction in sepsis. *Intensive Care Med Exp* 7-(Suppl 1):45, 2019.