Table S1. Characteristics of patients with sepsis based on CETAT database, stratified by risk status

Characteristic	Risk- N=987	Risk+ N=373	P Value	OR	OR*
Male	625 (63.3%)	238 (63.8%)	0.92	0.98 (0.76,1.25)	--
Age, median (IQR)	70.0 (55.0, 81.0)	68.0 (52.0, 78.0)	<0.01	0.9930 (0.9866,0.9996)	--
Transported by ambulance	603 (61.1%)	263 (70.5%)	<0.01	0.66 (0.51,0.85)	0.65 (0.50,0.84)
Coma	111 (11.2%)	145 (38.9%)	<0.01	5.02 (3.77,6.70)	5.19 (3.89,6.94)
Body temperature, median (IQR)	36.5 (36.3, 37.0)	36.5 (36.3, 37.1)	0.58	1.001 (0.872,1.144)	0.996 (0.867 1.139)
Heart rate, median (IQR)	97.0 (82.0, 113.0)	99.0 (82.0, 119.0)	0.04	1.0049 (0.9999 1.0099)	1.0045 (0.9995,1.0095)
Respiration rate, median (IQR)	21.0 (20.0, 23.0)	21.0 (20.0, 25.0)	0.42	0.96 (0.94,0.98)	0.96 (0.94,0.98)
Systolic pressure, median (IQR)	131.0 (113.0, 152.5)	124.0 (103.0, 146.0)	<0.01	0.99 (0.99,0.99)	0.99 (0.99,0.99)
Diastolic pressure, median (IQR)	78.0 (66.0, 91.0)	75.0 (61.0, 87.0)	<0.01	0.98 (0.98,0.99)	0.98 (0.98,0.99)
SpO2, median (IQR)	95.0 (89.0, 97.0)	94.0 (84.0, 97.0)	0.02	0.98 (0.97,0.99)	0.98 (0.97,0.99)
WBC, median (IQR)	10.2 (7.1, 13.9)	11.6 (8.3, 17.1)	<0.01	1.005 (0.996, 1.015)	1.005 (0.996, 1.015)
NE%, median (IQR)	84.3 (75.8, 89.9)	86.9 (79.3, 91.0)	<0.01	1.02 (1.01,1.03)	1.02 (1.01,1.03)

HGB, median (IQR)	120.0 (101.0, 138.0)	116.0 (94.0, 136.0)	0.06	0.9965 (0.9925 1.0006)	0.9960 (0.9919 1.0002)
PLT, median (IQR)	189.0 (138.0, 250.5)	172.0 (117.0, 229.0)	<0.01	0.998 (0.997 0.999)	0.998 (0.997 0.999)
ALT, median (IQR)	22.3 (13.8, 40.8)	31.3 (17.0, 70.1)	<0.01	1.002 (1.001 1.003)	1.002 (1.001 1.003)
AST, median (IQR)	27.8 (19.0, 48.1)	38.0 (24.0, 87.9)	<0.01	1.002 (1.001 1.002)	1.002 (1.001 1.002)
Albumin, median (IQR)	36.3 (31.5, 40.3)	35.0 (30.4, 39.8)	<0.01	0.98 (0.96,0.99)	0.98 (0.96,0.99)
TBIL, median (IQR)	13.1 (8.6, 21.0)	15.5 (9.5, 27.7)	<0.01	1.004 (1.001,1.007)	1.006 (1.002,1.010)
Creatinine, median (IQR)	76.6 (57.0, 113.0)	84.0 (58.4, 157.5)	<0.01	1.0005 (0.9998,1.0011)	1.0004 (0.9998,1.0010)
CO2CP, median (IQR)	22.7 (19.5, 25.9)	21.3 (17.8, 26.0)	<0.01	0.97 (0.96,0.99)	0.98 (0.96,0.99)
Potassium, median (IQR)	4.0 (3.6, 4.4)	4.0 (3.6, 4.6)	0.30	1.17 (1.01,1.37)	1.17 (1.01,1.38)
Sodium, median (IQR)	137.7 (134.0, 140.5)	138.9 (135.0, 142.0)	<0.01	1.04 (1.02,1.06)	1.04 (1.02,1.06)
Phosphorus, median (IQR)	1.1 (0.9, 1.4)	1.2 (0.9, 1.6)	<0.01	1.57 (1.27,1.94)	1.56 (1.26,1.93)
AG, median (IQR)	16.0 (13.2, 19.1)	17.3 (13.2, 21.7)	<0.01	1.04 (1.02,1.06)	1.04 (1.02,1.06)
Osmotic pressure, median (IQR)	281.5 (274.4, 289.8)	286.6 (279.5, 296.5)	<0.01	1.02 (1.01,1.03)	1.02 (1.01,1.03)
Glucose, median (IQR)	7.2 (6.0, 9.4)	7.9 (6.2, 10.6)	<0.01	1.018	1.018

				(0.997,1.039)	(0.997,1.039)
Prothrombin time, median (IQR)	13.3 (12.3, 14.7)	13.8 (12.6, 15.7)	<0.01	1.0198 (1.0007,1.0403)	1.0196 (1.0004,1.0403)
APTT, median (IQR)	34.5 (28.9, 39.3)	34.8 (29.1, 41.3)	0.08	1.014 (1.005,1.024)	1.014 (1.005,1.024)
Fibrinogen, median (IQR)	4.1 (3.0, 5.5)	4.2 (2.8, 5.7)	0.86	1.01(0.95,1.08)	1.01(0.95,1.08)
TT, median (IQR)	16.7 (15.8, 17.9)	16.9 (15.7, 18.0)	0.50	1.007 (0.999 1.015)	1.007 (0.999 1.015)
qSOFA, median (IQR)	0.0 (0.0, 1.0)	1.0 (0.0, 1.0)	<0.01	2.23 (1.88,2.67)	2.29 (1.93,2.74)
SIRS, median (IQR)	2.0 (1.0, 2.0)	2.0 (1.0, 3.0)	<0.01	1.19 (1.07,1.34)	1.19 (1.07,1.34)
NEWS, median (IQR)	4.0 (2.0, 7.0)	7.0 (4.0, 9.0)	<0.01	1.20 (1.16,1.25)	1.20 (1.16,1.25)

Table S2. Characteristics of patients with sepsis+ and sepsis- based on MIMIC-IV database

Characteristic	Sepsis- N=32,872	Sepsis+ N=2,145	P Value	OR	OR*
Coma (GCS* <=7)	901 (2.7%)	60 (2.8%)	0.93	1.02 (0.78,1.32)	0.97 (0.74,1.26)
Systolic pressure, median (IQR)	125.0 (109.0, 142.0)	113.0 (99.0, 132.0)	<0.01	0.98 (0.98,0.98)	0.98 (0.98,0.98)
Diastolic pressure, median (IQR)	68.0 (58.0, 80.0)	62.0 (52.0, 73.0)	<0.01	0.98 (0.98,0.98)	0.98 (0.98,0.99)
Heart rate, median (IQR)	85.0 (74.0, 98.0)	96.0 (82.0, 111.0)	<0.01	1.02 (1.02,1.03)	1.03 (1.02,1.03)
Respiration rate, median (IQR)	18.0 (15.0, 22.0)	21.0 (17.0, 25.0)	<0.01	1.07 (1.06,1.08)	1.07 (1.06,1.08)
SpO2, median (IQR)	98.0 (96.0, 100.0)	97.0 (94.0, 99.0)	<0.01	0.94 (0.93,0.95)	0.95 (0.94,0.95)
Body temperature, median (IQR)	36.7 (36.4, 37.0)	36.8 (36.4, 37.3)	<0.01	1.34 (1.06,1.42)	1.40 (1.32,1.48)
AG, median (IQR)	14.0 (12.0, 17.0)	16.0 (14.0, 20.0)	<0.01	1.07 (1.06,1.08)	1.07(1.07,1.08)
Age, median (IQR)	66.6 (54.4, 77.8)	68.9 (57.4, 81.1)	<0.01	1.01 (1.01,1.01)	--
WBC, median (IQR)	10.5 (7.6, 14.4)	12.9 (8.1, 18.7)	<0.01	1.01 (1.01,1.01)	1.01 (1.01,1.01)
GCS, median (IQR)	15.0 (15.0, 15.0)	15.0 (14.0, 15.0)	<0.01	0.97 (0.95,0.98)	0.97 (0.95,0.99)
qSOFA, median (IQR)	0.0 (0.0, 1.0)	1.0 (0.0, 1.0)	<0.01	2.21 (2.07,2.36)	2.19 (2.05,2.34)
SIRS, median (IQR)	1.0 (0.0, 2.0)	2.0 (1.0, 3.0)	<0.01	1.94 (1.86,2.02)	1.98 (1.90,2.06)
NEWS, median (IQR)	2.0 (1.0, 4.0)	4.0 (2.0, 6.0)	<0.01	1.30 (1.28,1.32)	1.30 (1.28,1.32)

* GCS: Glasgow coma scale.

Table S3. The prediction performance of sepsis with scoring systems

Prediction of sepsis	N	AUC-ROC	95%CI		P Value
qSOFA	35,017	0.623	0.612	0.634	Ref
SIRS	35,017	0.689	0.678	0.700	<0.001
NEWS	35,017	0.690	0.679	0.702	<0.001
Prediction of sepsis	N	AUC-ROC	95%CI		P Value
qSOFA-based LR*	35,017	0.625	0.614	0.637	0.029
SIRS-based LR*	35,017	0.694	0.683	0.705	<0.001
NEWS-based LR*	35,017	0.708	0.697	0.720	<0.001

Prediction performance (AUC-ROC, with 95% confidence interval) for the prediction of sepsis, based on medical scoring system (first scenario) or logistic regression with the available features in medical scoring system (second scenario). (Data from MIMIC-IV database.) For the first prediction scenario, the AUC-ROC for qSOFA was considered as the reference for comparison (to obtain test p-value). For the second prediction scenario, the corresponding AUC-ROC in the first scenario was considered as the reference for comparison (to obtain test p-value).

* logistic regression with the available features in medical scoring system.

Table S4. The prediction performance of scoring systems updated with additional clinical features

Prediction of sepsis	N	AUC-ROC	95%CI		P Value
qSOFA-based LR*	12,799	0.657	0.642	0.673	<0.001
SIRS-based LR*	12,799	0.725	0.711	0.740	<0.001
NEWS-based LR*	12,799	0.756	0.742	0.770	<0.001
Prediction of sepsis with high risk of death	N	AUC-ROC	95%CI		P Value
qSOFA-based LR*	1,360	0.674	0.641	0.706	<0.001
SIRS-based LR*	1,360	0.578	0.544	0.611	0.042
NEWS-based LR*	1,360	0.718	0.686	0.750	<0.001

Prediction performance (AUC-ROC, with 95% confidence interval) for the prediction of sepsis (among study population, first scenario) and the prediction of sepsis with high risk of death (among patients with sepsis, second scenario), based on logistic regression with the available features in medical scoring system. (Data from CETAT database.) For each prediction scenario, the corresponding AUC-ROC in Table 3 was considered as the reference for comparison (to obtain test p-value).

* logistic regression with the available features in medical scoring system.

Table S5. Sensitivity and specificity of sepsis and sepsis with high risk of death with scoring systems

Prediction of sepsis	Sensitivity	Specificity
qSOFA (>=2)	10.8%	96.4%
SIRS (>=4)	3.6%	99.4%
NEWS (>=5)	52.7%	79.2%
Prediction of sepsis with high risk of death	Sensitivity	Specificity
qSOFA (>=2)	21.2%	93.1%
SIRS (>=4)	5.1%	97.0%

Sensitivity and specificity for the prediction of sepsis (among study population, first scenario) and the prediction of sepsis with high risk of death (among patients with sepsis, second scenario), based on medical scoring system. (Data from CETAT database.)

Table S6. Sensitivity and specificity for the prediction of sepsis with scoring system

Prediction of sepsis	Sensitivity	Specificity
qSOFA (>=2)	11.4%	96.5%
SIRS (>=4)	6.1%	98.5%
NEWS (>=5)	44.4%	81.6%

Sensitivity and specificity for the prediction of sepsis based on medical scoring system. (Data from MIMIC-IV database.)

Table S7. Sensitivity and specificity of scoring systems updated with additional clinical features

Prediction of sepsis	Sensitivity	Specificity
qSOFA-based LR*	41.8%	86.4%
SIRS-based LR*	58.6%	76.0%
NEWS-based LR*	63.9%	76.4%
Prediction of sepsis with high risk of death	Sensitivity	Specificity
qSOFA-based LR*	54.9%	68.1%
SIRS-based LR*	67.0%	62.1%
NEWS-based LR*	58.7%	67.6%

Sensitivity and specificity for the prediction of sepsis (among study population, first scenario) and the prediction of sepsis with high risk of death (among patients with sepsis, second scenario), based on logistic regression with the available features in medical scoring system. (Data from CETAT database.) * logistic regression with the available features in medical scoring system.

Table S8. Sensitivity and specificity of scoring systems updated with additional clinical features

Prediction of sepsis	Sensitivity	Specificity
qSOFA-LR*	54.9%	68.1%
SIRS-LR*	67.0%	62.1%
NEWS-LR*	56.8%	73.8%

Sensitivity and specificity for the prediction of sepsis based on logistic regression with the available features in medical scoring system. (Data from MIMIC-IV database.)
* logistic regression with the available features in medical scoring system.

Table S9. The additional clinical features selected based on different criteria in CETAT database.

Characteristic	Sepsis				High-risk
	AIC	p<0.1	p<0.05	p<0.01	AIC
NE%	✓	✓			
*MONO%	✓				✓
EOS%		✓	✓	✓	
*ANC	✓	✓	✓	✓	
*ALC	✓	✓			
*Baso#	✓	✓	✓	✓	
*Eos#	✓	✓	✓	✓	
*HCT	✓				
*MCV	✓	✓	✓		
*MCHC	✓	✓			
*RDW	✓	✓	✓		
*MPV	✓	✓	✓		
*PCT					✓
*γ-GT		✓	✓		
AST					✓
*TP					✓
*ALB	✓				
*GLB		✓	✓	✓	
*A/G	✓				

	22	20	17	11	6
TBIL		✓	✓	✓	
*DBIL	✓				✓
Urea		✓	✓	✓	
CO2CP	✓	✓	✓	✓	
Potassium	✓				
Sodium	✓	✓	✓	✓	
Calcium	✓	✓	✓		
Phosphorus		✓	✓		
AG	✓				
Osmotic pressure					✓
*PT	✓				
APTT	✓	✓	✓	✓	
Fibrinogen	✓	✓	✓	✓	
The number of added variables in the baseline model	**22**	**20**	**17**	**11**	**6**

MONO: monocyte; ANC: absolute neutrophil count; ALC: Absolute Lymphocyte Count; Baso: Basophils; EOS: Eosinophils; HCT: Hematocrit; MCV: Mean Corpuscular Volume; MCHC: mean corpuscular hemoglobin concentration; RDW: Red Cell Distribution Width; MPV: Mean Platelet Volume; PCT: Plateletcrit; γ-GT: Gamma-Glutamyl Transferase; TP: Total Protein; ALB: Albumin; GLB: Globulin; A/G: Albumin/Globulin Ratio; DBIL: Direct Bilirubin; PT: Prothrombin Time.
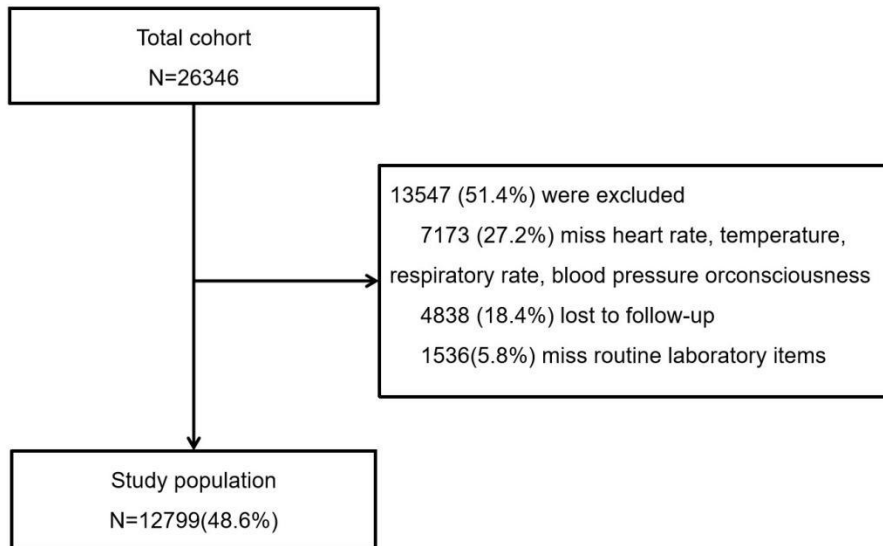
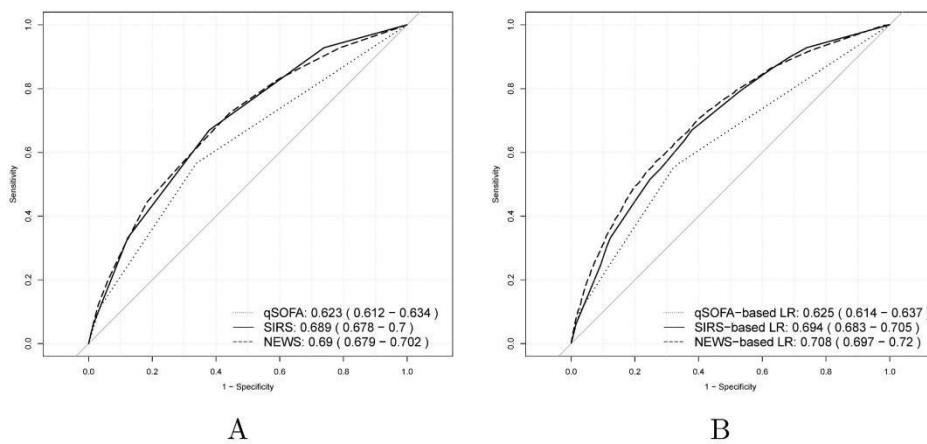**Figure S1.** Flowchart illustrating composition of study population.

Figure S2. Prediction performance (receiver operating characteristic curve, or ROC) for the prediction of sepsis. (Data from MIMIC IV database.) A. prediction based on medical scoring system; B. prediction based on logistic regression with the available features in medical scoring system. For each evaluation, the area under the ROC curve (AUC-ROC) with 95% confidence interval are provided.