

Original Article



Integration of CURB-65 and CT Severity Score for Prognostic Prediction in Community-acquired Pneumonia

Ökkeş Gültekin¹, Orhan Özçekici², İnan Arslantaş²¹Clinic of Pulmonology, Kemalpaşa State Hospital, İzmir, Türkiye²Clinic of Radiology, Kemalpaşa State Hospital, İzmir, Türkiye

Cite this article as: Gültekin Ö, Özçekici O, Arslantaş İ. Integration of CURB-65 and CT severity score for prognostic prediction in community-acquired pneumonia. *Thorac Res Pract.* 2026;27(Spec No 1):45-52

ABSTRACT

OBJECTIVE: This study compared confusion, urea, respiratory rate, blood pressure, age ≥ 65 years (CURB-65) and computed tomography severity score (CTSS) for predicting 30-day mortality, intensive care unit (ICU) admission, and length of hospital stay in patients with community-acquired pneumonia (CAP) and assessed whether CTSS added prognostic value to CURB-65.

MATERIAL AND METHODS: This retrospective study included 456 adults with CT-confirmed CAP. CURB-65 and CTSS were calculated. Multivariable logistic regression and receiver operating characteristic analyses were performed for 30-day mortality and ICU admission, and area under the curve (AUC) values were compared using the DeLong test.

RESULTS: For mortality prediction, the AUC values were 0.894 for CURB-65, 0.762 for mean CTSS, and 0.909 for the combined mortality model. The combined model did not significantly outperform CURB-65 ($P = 0.081$). For ICU admission, the AUC values were 0.851, 0.716, and 0.881, respectively; the combined ICU model significantly outperformed CURB-65 ($P = 0.003$). The mean CTSS was independently associated with mortality and ICU admission, but was not associated with length of hospital stay.

CONCLUSION: CURB-65 remained a strong and practical predictor of mortality and ICU admission in CAP. Mean CTSS provided additional prognostic information for mortality and ICU admission, but not for length of hospital stay. The combined mortality model did not significantly improve discrimination compared with CURB-65 alone, whereas the combined ICU model did significantly improve discrimination in predicting ICU admission.

KEYWORDS: Community-acquired pneumonia, CURB-65, CT severity score, mortality, intensive care, ROC analysis

Received: 12.05.2026

Revision Requested: 02.06.2026

Last Revision Received: 30.06.2026

Accepted: 14.07.2026

Epub: 09.10.2026

Publication Date: 21.10.2026

INTRODUCTION

Lower respiratory tract infections, particularly pneumonia, remain a major cause of morbidity and mortality worldwide despite a decline in prevalence over the past three decades.¹ Because pneumonia has a broad clinical spectrum, accurate risk stratification at diagnosis is critical for determining the site of care and predicting prognosis.² Several prognostic scoring systems have therefore been developed and are widely used in clinical practice.

Confusion, urea, respiratory rate, blood pressure, age ≥ 65 years (CURB-65) is one of the most widely used and easily applied scoring systems. It includes confusion, urea level, respiratory rate, blood pressure, and age ≥ 65 years.³ CURB-65 is commonly used for mortality prediction and is included in international guidelines.^{2,4}

However, CURB-65 has several limitations for prognostication. The score was primarily developed to predict 30-day mortality, and its age component may substantially influence the score in older patients. In younger patients, reliance on clinical findings that may appear later, such as hypotension and confusion, may lead to early misclassification of some potentially severe cases as low risk.^{5,6} Its sensitivity for predicting intensive care unit (ICU) need is also limited, and it may not always provide adequate discrimination for critical patient triage.^{7,8} The absence of factors that may influence disease

Corresponding author: Ökkeş Gültekin, MD, e-mail: o.gultekin@hotmail.com.tr



Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Turkish Thoracic Society. Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

course, such as hypoxemia and radiological extent, may further limit its prognostic performance.⁹

In recent years, thoracic computed tomography (CT) has played an increasingly important role in assessing the severity and prognosis of pneumonia. The CT severity score (CTSS), a semiquantitative assessment method, enables objective evaluation of the extent and severity of lung involvement.¹⁰ Although CT-based severity scoring was widely studied during the coronavirus disease-2019 (COVID-19) pandemic,¹¹ lobar-based assessment of parenchymal involvement may also provide clinically relevant prognostic information in non-COVID community-acquired pneumonia (CAP).¹²⁻¹⁴

Clinical scores primarily reflect the systemic physiological response and clinical severity, whereas radiological scores reflect the anatomical extent of lung involvement. Therefore, these two approaches may provide complementary information. However, studies evaluating CURB-65 together with CTSS and investigating the additional prognostic contribution of combined models remain limited.¹³

Based on this complementary framework, we hypothesized that CTSS would provide prognostic information beyond CURB-65 by capturing the radiological disease burden. Therefore, this study aimed to compare the performance of CURB-65 and CTSS in predicting 30-day mortality, ICU admission, and length of hospital stay in patients with CAP, and to evaluate the additional prognostic contribution of CTSS to CURB-65.

MATERIAL AND METHODS

This retrospective study included patients aged 18 years or older who presented to the emergency department or the pulmonology outpatient clinic with CAP and who had pneumonia-compatible findings on thoracic CT between July 2023 and July 2025. This study was approved by the Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital before data extraction and analysis (approval date: September 11, 2025; approval number: 2025/08-09), and all patient data were anonymized.

Main Points

- Confusion, urea, respiratory rate, blood pressure, age ≥ 65 years (CURB-65) was a strong predictor of 30-day mortality and intensive care unit (ICU) admission in patients with community-acquired pneumonia.
- Computed tomography severity score was independently associated with mortality and ICU admission, but its discriminative performance was lower than that of CURB-65.
- The combined mortality model did not significantly improve mortality discrimination over CURB-65 alone, whereas the combined ICU model significantly improved discrimination for ICU admission.
- The findings support the central role of scores that reflect systemic clinical response in the prognosis of pneumonia.

During the study period, 3,712 thoracic CT examinations were performed. The radiology database was retrospectively searched for thoracic CT reports obtained at presentation to the emergency department or pulmonology outpatient clinic that contained at least one of the predefined terms “consolidation,” “infiltration,” or “pneumonia.” This search identified 1,176 potentially eligible CT reports. Medical records were then reviewed to confirm CAP and determine eligibility according to the predefined criteria. After excluding patients with bilateral pleural effusion, severe acute respiratory syndrome coronavirus 2 positivity, absence of CT at initial presentation, repeated CT within the same clinical episode, trauma-related CT, insufficient data for CURB-65 calculation or outcome assessment, or alternative diagnoses, the final study population consisted of 456 patients. The detailed patient selection process and exclusion numbers are summarized in Figure 1.

During the study planning stage, the minimum required sample size was estimated to be 200 patients, based on previously reported 30-day mortality rates of approximately 10% for CAP, assuming 80% statistical power and a two-sided alpha level of 0.05 for logistic regression analysis. No sampling was performed; all eligible patients during the study period were included, and the final sample size exceeded the prespecified minimum requirement.

Demographic information, comorbidities, vital signs, and laboratory values were recorded, and CURB-65 scores were calculated for all patients. Radiological assessment was based exclusively on thoracic CT images obtained at presentation; chest radiographs were not evaluated. All CT examinations were independently evaluated by two radiologists, one with ≤ 5 years and the other with >10 years of radiology experience. Both radiologists were blinded to patients’ clinical and laboratory data, CURB-65 scores, ICU admission, mortality outcomes,

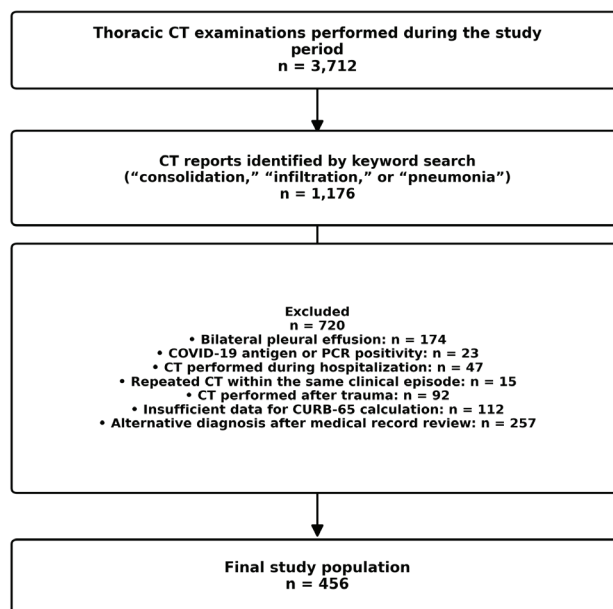


Figure 1. Flowchart showing the selection of the study population from thoracic CT examinations performed during the study period and the reasons for exclusion

CT: computed tomography, COVID-19: coronavirus disease-2019

and each other's assessments. CTSS was graded according to the percentage of parenchymal involvement in each lung lobe as follows: 0 = no involvement, 1 = <5%, 2 = 5–25%, 3 = 26–50%, 4 = 51–75%, and 5 = >75%. The total score, calculated by summing the scores of the five lobes, ranged from 0 to 25. The arithmetic mean of the scores assigned by the two radiologists was used as the final CTSS value in all statistical analyses.

For each outcome, separate multivariable logistic regression models were constructed: CURB-65-based, mean CTSS-based, and combined. The combined models included CURB-65, mean CTSS, and candidate covariates retained for multivariable analysis. For each patient, the predicted probability generated by the corresponding combined model was saved and used as the combined-model test variable in receiver operating characteristic (ROC) analyses.

Statistical Analysis

All data were analyzed using IBM SPSS Statistics for Windows, version 27.0 (IBM Corp., Armonk, NY, USA), and Python version 3.13.5 (Python Software Foundation, Wilmington, DE, USA) with NumPy 2.3.5, SciPy 1.17.0, and scikit-learn 1.8.0. Continuous variables were presented as mean $\pm$ standard deviation, and categorical variables were presented as numbers and percentages.

For binary outcomes, univariate logistic regression analysis was first performed to identify candidate predictors. Variables showing an association at $P < 0.20$ in univariate analyses were included as candidate variables in multivariable logistic regression models. Multivariable analysis results were reported as odds ratios (OR) with 95% confidence intervals (CI). The predicted probability from each combined multivariable model was used as the combined-model test variable in the ROC analysis. Non-parametric area under the curve (AUC) estimates, 95% CIs, optimal cutoffs based on the Youden index, and paired comparisons between ROC curves were calculated. AUC differences were evaluated using the DeLong method. For logistic regression models, the model fit was evaluated using the Omnibus test, -2 log likelihood, and Cox-Snell and Nagelkerke pseudo- R^2 values; calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test.

The length of hospital stay was logarithmically transformed because it had a non-normal distribution. Correlations between log-transformed length of stay and prognostic scores were evaluated using Spearman correlation, and independent associations were assessed using multivariable linear regression.

To assess coefficient stability and potential overfitting in the combined models, a bootstrap robustness analysis was performed using 1,000 simple bootstrap samples. Bias-corrected and accelerated (BCa) 95% CIs and two-sided bootstrap P values were calculated for the regression coefficients.

Interobserver reliability for CTSS between the two radiologists was assessed using the intraclass correlation coefficient based on a two-way mixed-effects model, an absolute-agreement definition, and the mean of two measurements. Intraclass correlation coefficient (ICC) values were reported with 95% CIs.

A P value of < 0.05 was considered statistically significant for all tests.

RESULTS

A total of 456 patients were included in the study. Of the patients, 46.3% were female, and the mean age was 66.40 ± 17.67 years. The most common comorbidity was hypertension (40.1%). The mean CURB-65 score was 1.37 ± 1.12 , and the mean CTSS was 4.99 ± 3.29 . A total of 221 patients (48.5%) were hospitalized, 74 (16.2%) were admitted to the ICU, and the 30-day mortality rate was 11.0% (Table 1). Interobserver reliability for CTSS between the two radiologists was excellent (average-measures ICC: 0.981; 95% CI: 0.977–0.984; $P < 0.001$).

In the mortality analyses, variables showing an association at $P < 0.20$ in univariate analysis were included as candidate variables in multivariable models. To avoid conceptual overlap and overadjustment, age and the number of involved lobes were

Table 1. Demographic and clinical characteristics

Parameter	Number/percentage of cases
Sex/female	211 (46.3%)
Age, years, mean $\pm$ SD	66.40 $\pm$ 17.67
Comorbidities	
Hypertension	183 (40.1%)
Diabetes mellitus	95 (20.8%)
COPD	74 (16.2%)
Coronary artery disease	28 (6.1%)
Alzheimer's disease	93 (20.4%)
Congestive heart failure	37 (8.1%)
Atrial fibrillation	32 (7.0%)
Asthma	32 (7.0%)
Other	42 (9.2%)
CURB-65 score	
0	128 (28.1%)
1	119 (26.1%)
2	142 (31.1%)
3	48 (10.5%)
4–5	19 (4.2%)
Number of affected lobes	
1	131 (28.7%)
2	141 (30.9%)
3	103 (22.6%)
4	43 (9.4%)
5	38 (8.3%)
ICU admission	74 (16.2%)
Invasive mechanical ventilation requirement	33 (7.2%)
Mortality	50 (11.0%)

Data are presented as mean $\pm$ standard deviation for normally distributed continuous variables and as n (%) for categorical variables

COPD: chronic obstructive pulmonary disease, ICU: intensive care unit, CURB-65: confusion, urea, respiratory rate, blood pressure, age ≥ 65 years, SD: standard deviation

not included in the CURB-65-based and mean CTSS-based models, respectively. All mortality models were statistically significant and showed acceptable calibration. The Omnibus $\chi^2(6)$ values were 122.998, 81.106, and 132.205 for the CURB-65-based, mean CTSS-based, and combined mortality models, respectively; all with $P < 0.001$. The corresponding Nagelkerke R^2 values were 0.474, 0.326, and 0.504, respectively. Hosmer-Lemeshow goodness-of-fit tests indicated acceptable calibration for all three models: CURB-65-based model, $\chi^2(8) = 9.111$, $P = 0.333$; mean CTSS-based model, $\chi^2(8) = 5.080$, $P = 0.749$; and combined mortality model, $\chi^2(8) = 5.839$, $P = 0.665$.

In the CURB-65-based mortality model, CURB-65 was independently associated with 30-day mortality (OR: 5.380; 95% CI: 3.369–8.591; $P < 0.001$). In the mortality model based on mean CTSS, age (OR: 1.043; 95% CI: 1.013–1.073; $P = 0.004$), Alzheimer's disease (OR: 3.701; 95% CI: 1.690–8.105; $P = 0.001$), and mean CTSS (OR: 1.349; 95% CI: 1.221–1.490; $P < 0.001$) were independently associated with 30-day mortality. In the combined mortality model, both CURB-65 (OR: 4.496; 95% CI: 2.854–7.084; $P < 0.001$) and mean CTSS (OR: 1.206; 95% CI: 1.073–1.355; $P = 0.002$) remained independently associated with 30-day mortality, whereas asthma, Alzheimer's disease, congestive heart failure, and pleural effusion were not significantly associated with 30-day mortality (Table 2).

In the bootstrap robustness analysis of the combined mortality model, CURB-65 (bootstrap $P = 0.001$; BCa 95% CI for B, 1.005–2.395) and mean CTSS (bootstrap $P = 0.002$; BCa 95% CI

for B, 0.064–0.327) retained statistically significant coefficients. The BCa CIs for asthma, Alzheimer's disease, congestive heart failure, and pleural effusion included zero. The asthma estimate was unstable across bootstrap samples, consistent with sparse-data effects.

In ROC analysis for mortality prediction, the AUC values were 0.894 for CURB-65 (95% CI: 0.852–0.936; $P < 0.001$), 0.762 for mean CTSS (95% CI: 0.689–0.834; $P < 0.001$), and 0.909 for the combined mortality model (95% CI: 0.870–0.949; $P < 0.001$) (Figure 2). In paired DeLong comparisons, the combined mortality model did not significantly outperform CURB-65 (AUC difference: 0.015; 95% CI: -0.002 to 0.033; $P = 0.081$). CURB-65 significantly outperformed mean CTSS (AUC difference: 0.132; 95% CI: 0.059–0.206; $P = 0.0004$), and the combined mortality model significantly outperformed mean CTSS (AUC difference: 0.148; 95% CI: 0.085–0.210; $P < 0.001$).

At the Youden-optimal thresholds, CURB-65 ≥ 3 showed a sensitivity of 72.0% and specificity of 92.4% for mortality prediction. For mean CTSS, the optimal threshold was ≥ 8 , with a sensitivity of 52.0% and specificity of 86.7%. For the combined mortality model, the optimal predicted-probability threshold was ≥ 0.20768 , with a sensitivity of 74.0% and specificity of 92.1%. Thus, the combined mortality model provided only a limited numerical improvement over CURB-65, but did not show statistically significant superiority.

Table 2. Independent predictors of 30-day mortality

Model	Variable	B	S.E.	Wald	P	OR	95% CI lower	95% CI upper
Model 1 CURB-65	Number of lobes	0.152	0.165	0.854	0.355	1.165	0.843	1.609
	Asthma	-0.704	1.178	0.357	0.550	0.495	0.049	4.978
	Alzheimer's disease	0.586	0.415	1.995	0.158	1.797	0.797	4.053
	CHF	0.561	0.515	1.189	0.275	1.753	0.639	4.808
	CURB-65	1.683	0.239	49.644	<0.001*	5.380	3.369	8.591
	Effusion	0.337	0.418	0.651	0.420	1.400	0.618	3.174
Model 2 Mean CTSS	Age	0.042	0.015	8.165	0.004*	1.043	1.013	1.073
	CHF	0.591	0.503	1.378	0.240	1.805	0.673	4.841
	Alzheimer's disease	1.309	0.400	10.711	0.001*	3.701	1.690	8.105
	Effusion	0.344	0.375	0.842	0.359	1.411	0.676	2.943
	Asthma	-0.828	1.084	0.583	0.445	0.437	0.052	3.660
	Mean CTSS	0.299	0.051	34.775	<0.001*	1.349	1.221	1.490
Model 3 CURB-65 + mean CTSS	Asthma	-0.607	1.181	0.264	0.607	0.545	0.054	5.516
	Alzheimer's disease	0.719	0.430	2.801	0.094	2.053	0.884	4.765
	CHF	0.596	0.528	1.272	0.259	1.814	0.644	5.106
	Effusion	0.313	0.422	0.552	0.457	1.368	0.599	3.127
	CURB-65	1.503	0.232	41.994	<0.001*	4.496	2.854	7.084
	Mean CTSS	0.187	0.059	9.934	0.002*	1.206	1.073	1.355

Multivariable logistic regression analysis

* $P < 0.05$

CHF: congestive heart failure, CTSS: computed tomography severity score, CI: confidence interval, OR: odds ratio, CURB-65: confusion, urea, respiratory rate, blood pressure, age ≥ 65 years, S.E.: standard error

In the ICU admission analyses, variables showing an association at $P < 0.20$ in univariate analysis were included as candidates in multivariable models. To avoid conceptual overlap and overadjustment, age was not included in the CURB-65-based model, and the number of involved lobes was not included in

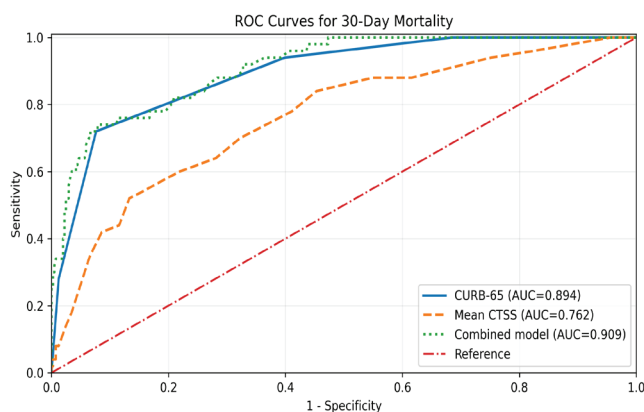


Figure 2. Comparison of ROC curves for CURB-65, mean CTSS, and the combined mortality model in predicting 30-day mortality. ROC curves for CURB-65, mean CTSS, and the combined mortality model based on predicted probability; AUCs were 0.894, 0.762, and 0.909, respectively

ROC: receiver operating characteristic, CURB-65: confusion, urea, respiratory rate, blood pressure, age ≥ 65 years, CTSS: computed tomography severity score, AUC: area under the curve

the mean CTSS-based model. All ICU admission models were statistically significant and showed acceptable calibration. The Omnibus $\chi^2(6)$ values were 135.121 for the CURB-65-based model, 83.792 for the mean CTSS-based model, and 139.623 for the combined ICU model, all with $P < 0.001$. The corresponding Nagelkerke R^2 values were 0.436, 0.285, and 0.449, respectively. Hosmer–Lemeshow goodness-of-fit tests indicated acceptable calibration for all three models: the CURB-65-based model ($\chi^2(8) = 8.236$, $P = 0.411$); the mean CTSS-based model ($\chi^2(8) = 10.482$, $P = 0.233$); and the combined ICU model ($\chi^2(8) = 8.549$, $P = 0.382$).

In the CURB-65-based model for ICU admission, CURB-65 (OR: 4.022; 95% CI: 2.762–5.855; $P < 0.001$), Alzheimer's disease (OR: 2.717; 95% CI: 1.347–5.478; $P = 0.005$), and pleural effusion (OR: 2.091; 95% CI: 1.066–4.100; $P = 0.032$) were independently associated with ICU admission. In the model based on mean CTSS, Alzheimer's disease (OR: 5.761; 95% CI: 2.893–11.469; $P < 0.001$), pleural effusion (OR: 2.049; 95% CI: 1.109–3.788; $P = 0.022$), age (OR: 1.024; 95% CI: 1.002–1.046; $P = 0.031$), and mean CTSS (OR: 1.239; 95% CI: 1.138–1.347; $P < 0.001$) were independently associated with ICU admission. In the combined ICU model, Alzheimer's disease (OR: 3.062; 95% CI: 1.495–6.271; $P = 0.002$), pleural effusion (OR: 2.080; 95% CI: 1.058–4.088; $P = 0.034$), CURB-65 (OR: 3.504; 95% CI: 2.424–5.065; $P < 0.001$), and mean CTSS (OR: 1.113; 95% CI: 1.009–1.227; $P = 0.033$) were independently associated with ICU admission (Table 3).

Table 3. Independent predictors of ICU admission

Model	Variable	B	S.E.	Wald	P	OR	95% CI lower	95% CI upper
Model 1 CURB-65	Asthma	−0.360	0.876	0.169	0.681	0.697	0.125	3.885
	Alzheimer's disease	0.999	0.358	7.803	0.005*	2.717	1.347	5.478
	CHF	−0.070	0.472	0.022	0.882	0.932	0.369	2.352
	Effusion	0.737	0.344	4.605	0.032*	2.091	1.066	4.100
	Number of lobes	0.009	0.137	0.004	0.947	1.009	0.772	1.319
	CURB-65	1.392	0.192	52.732	<0.001*	4.022	2.762	5.855
Model 2 Mean CTSS	Asthma	−0.576	0.807	0.509	0.476	0.562	0.116	2.735
	Alzheimer's disease	1.751	0.351	24.840	<0.001*	5.761	2.893	11.469
	CHF	0.214	0.457	0.219	0.640	1.238	0.505	3.035
	Effusion	0.718	0.313	5.243	0.022*	2.049	1.109	3.788
	Age	0.024	0.011	4.627	0.031*	1.024	1.002	1.046
	Mean CTSS	0.214	0.043	24.782	<0.001*	1.239	1.138	1.347
Model 3 CURB-65 + Mean CTSS	Asthma	−0.335	0.871	0.148	0.700	0.715	0.130	3.944
	Alzheimer's disease	1.119	0.366	9.366	0.002*	3.062	1.495	6.271
	CHF	−0.060	0.480	0.016	0.901	0.942	0.367	2.414
	Effusion	0.732	0.345	4.508	0.034*	2.080	1.058	4.088
	CURB-65	1.254	0.188	44.483	<0.001*	3.504	2.424	5.065
	Mean CTSS	0.107	0.050	4.569	0.033*	1.113	1.009	1.227

Multivariable logistic regression analysis

* $P < 0.05$

CHF: congestive heart failure, CTSS: computed tomography severity score, CI: confidence interval, CURB-65: confusion, urea, respiratory rate, blood pressure, age ≥ 65 years, S.E.: standard error, ICU: intensive care unit

In the bootstrap robustness analysis of the combined ICU model, Alzheimer’s disease, pleural effusion, and CURB-65 retained statistically significant coefficients, whereas the mean CTSS coefficient was borderline significant (bootstrap $P = 0.051$; BCa 95% CI for B, -0.012 to 0.219). Asthma and congestive heart failure were not significantly associated.

In the ROC analysis for ICU admission, the AUC values were 0.851 for CURB-65 (95% CI: 0.800–0.901; $P < 0.001$), 0.716 for mean CTSS (95% CI: 0.653–0.779; $P < 0.001$), and 0.881 for the combined ICU model based on the predicted probability (95% CI: 0.836–0.926; $P < 0.001$) (Figure 3). In paired DeLong comparisons, the combined ICU model significantly outperformed CURB-65 (AUC difference: 0.030; 95% CI: 0.010–0.050; $P = 0.0028$). Both CURB-65 and the combined ICU model significantly outperformed the mean CTSS: AUC difference: 0.135; 95% CI: 0.062–0.208; $P = 0.0003$ for CURB-65, and AUC difference: 0.165; 95% CI: 0.099–0.231; $P < 0.001$ for the combined ICU model.

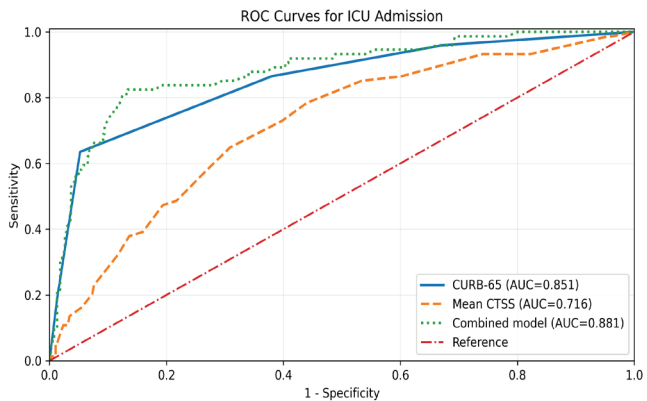


Figure 3. Comparison of ROC curves for CURB-65, mean CTSS, and the combined ICU model in predicting ICU admission. ROC curves for CURB-65, mean CTSS, and the combined ICU model based on predicted probability; AUCs were 0.851, 0.716, and 0.881, respectively

ROC: receiver operating characteristic, CURB-65: confusion, urea, respiratory rate, blood pressure, age ≥ 65 years, CTSS: computed tomography severity score, ICU: intensive care unit, AUC: area under the curve

At the Youden-optimal thresholds, CURB-65 ≥ 3 showed a sensitivity of 63.5% and specificity of 94.8% for ICU admission. For mean CTSS, the optimal threshold was ≥ 4.5 , with a sensitivity of 78.4% and specificity of 56.0%. For the combined ICU model, the optimal predicted-probability threshold was ≥ 0.20878 , with a sensitivity of 82.4% and specificity of 86.6%. A summary of ROC analyses, including AUC values, optimal cut-offs, sensitivity, specificity, Youden indices, and DeLong comparisons, is presented in Table 4.

Among hospitalized patients ($n = 221$), the length of hospital stay was not normally distributed; therefore, a logarithmic transformation was applied before analysis. In a Spearman correlation analysis, log-transformed length of stay showed a weak positive correlation with CURB-65 ($\rho = 0.230$; $P < 0.001$), whereas its correlation with CTSS was not statistically significant ($\rho = 0.110$; $P = 0.103$). In multivariable linear regression analysis, CURB-65 ($B = 0.096$; 95% CI: 0.023–0.170; $P = 0.010$), pleural effusion ($B = 0.269$; 95% CI: 0.082–0.456; $P = 0.005$), and congestive heart failure ($B = -0.295$; 95% CI: -0.570 to -0.019 ; $P = 0.036$) were independently associated with length of hospital stay. In contrast, CTSS was not independently associated with length of hospital stay ($P = 0.797$).

DISCUSSION

In this study, CURB-65 and CTSS were compared in their ability to predict 30-day mortality, ICU admission, and length of hospital stay in patients with CAP, and the performance of combined multivariable models was evaluated. CURB-65 showed strong discrimination for both mortality and ICU admission. Mean CTSS provided additional prognostic information for mortality and ICU admission in conventional multivariable models; however, its standalone discriminative performance was lower than that of CURB-65. The combined mortality model did not significantly improve discrimination over CURB-65 alone; however, the combined ICU model did significantly improve discrimination in predicting ICU admission.

In this study, the discriminatory power of CURB-65 for predicting mortality was high. This finding is consistent with previous studies reporting CURB-65 as a reliable tool for assessing pneumonia severity.^{3,15} However, CURB-65 may

Table 4. Summary of ROC analyses for 30-day mortality and ICU admission

Outcome	Model	AUC	95% CI	Optimal cut-off	Sensitivity (%)	Specificity (%)	Youden J
30-day mortality	CURB-65	0.894	0.852–0.936	≥ 3	72.0	92.4	0.644
	Mean CTSS	0.762	0.689–0.834	≥ 8	52.0	86.7	0.387
	Combined mortality model	0.909	0.870–0.949	≥ 0.20768	74.0	92.1	0.661
ICU admission	CURB-65	0.851	0.800–0.901	≥ 3	63.5	94.8	0.583
	Mean CTSS	0.716	0.653–0.779	≥ 4.5	78.4	56.0	0.344
	Combined ICU model	0.881	0.836–0.926	≥ 0.20878	82.4	86.6	0.691

The combined models were based on predicted probabilities derived from the corresponding multivariable logistic regression models. For CURB-65, the ROC-derived threshold of >2 corresponds to a clinical cut-off of ≥ 3 because CURB-65 is an integer-based score. In DeLong comparisons for 30-day mortality, the combined mortality model did not significantly outperform CURB-65 (AUC difference = 0.015; 95% CI: -0.002 to 0.033 ; $P = 0.081$), whereas CURB-65 significantly outperformed mean CTSS ($P = 0.0004$) and the combined mortality model significantly outperformed mean CTSS ($P < 0.001$). For ICU admission, the combined ICU model significantly outperformed CURB-65 (AUC difference = 0.030; 95% CI: 0.010–0.050; $P = 0.0028$), CURB-65 significantly outperformed mean CTSS ($P = 0.0003$), and the combined ICU model significantly outperformed mean CTSS ($P < 0.001$)

AUC: area under the curve, CI: confidence interval, CTSS: computed tomography severity score, ICU: intensive care unit, ROC: receiver operating characteristic, CURB-65: confusion, urea, respiratory rate, blood pressure, age ≥ 65 years

be insufficient in some patient groups because it includes a limited number of variables and does not directly incorporate hypoxemia, radiological extent, or broader markers of acute physiological reserve.^{5,6} Because CURB-65 includes age and parameters reflecting systemic physiological deterioration, its prognostic performance may be particularly strong in cohorts characterized by older age and greater clinical severity. Therefore, the high discriminative performance observed in our study may be partly related to the clinical characteristics and CT-selected nature of the patient population.

Although much of the available evidence on CT-based severity scoring originates from the COVID-19 era,¹⁰ CT-based quantification of parenchymal involvement may also be relevant in non-COVID CAP.¹²⁻¹⁴ Nevertheless, disease-specific differences in radiological patterns, clinical courses, and prognoses should be considered when extrapolating COVID-19 CTSS findings to non-COVID CAP. In the present study, CTSS was evaluated as a semiquantitative marker of anatomical disease burden rather than as a replacement for clinical severity assessment. The excellent interobserver reliability in the present study supports the reproducibility of this semiquantitative approach and is consistent with previous validation studies of CT scoring systems.^{16,17}

Although the AUC of the combined mortality model was numerically highest, the increase was not statistically significant compared with that of CURB-65. At the Youden-optimal thresholds, the combined mortality model provided slightly higher sensitivity and had specificity similar to that of CURB-65, whereas the mean CTSS alone provided lower sensitivity and specificity. These findings suggest that radiological disease burden may complement clinical severity assessment, but it does not meaningfully replace or substantially outperform CURB-65 for mortality prediction in this cohort.

For ICU admission, CURB-65 retained strong discriminative performance, but the combined ICU model achieved a statistically significant increase in AUC. The combined model also showed a more balanced sensitivity-specificity profile than either CURB-65 or mean CTSS alone. This finding suggests that integrating radiological extent with the clinical score and selected clinical covariates may improve discrimination in ICU triage, particularly in patients who are already undergoing thoracic CT for diagnostic clarification. However, this potential role should be interpreted cautiously because external validation is required before routine clinical application.¹⁸

The bootstrap findings further qualified the ICU results. CURB-65, Alzheimer's disease, and pleural effusion retained stable associations with ICU admission, whereas the mean CTSS coefficient was borderline in the bootstrap analysis. Thus, the incremental ICU discrimination provided by the combined model should not be attributed solely to a robust CTSS effect. Rather, it likely reflects the combined contribution of clinical severity, radiological extent, and selected clinical covariates.

Length-of-stay analyses also supported the limited contribution of CTSS. Log-transformed length of stay showed a weak positive correlation with CURB-65, whereas its correlation with CTSS was not statistically significant. In multivariable

linear regression analysis, CURB-65, pleural effusion, and congestive heart failure were independently associated with length of stay, whereas CTSS was not independently associated. These findings suggest that hospital stay may be more closely related to systemic clinical status and comorbidity burden than to radiological extent alone.¹⁹

From a practical perspective, CTSS should not be considered a routine replacement for bedside clinical scores because CT availability, cost, radiation exposure, and the need for radiological expertise may limit its universal use. Its potential role may be more appropriate in selected patients who are already undergoing CT for diagnostic clarification, or in those whose ICU triage remains uncertain after initial clinical assessment.

The relatively high AUC values observed for CURB-65 should be interpreted in the context of a CT-selected cohort rather than the broader CAP population. Although the mean CURB-65 score indicated a low-to-intermediate clinical risk profile, 45.8% of patients had a CURB-65 score ≥ 2 , nearly half of the cohort was hospitalized, and a substantial proportion required ICU admission. Because thoracic CT is more likely to be performed in patients with greater clinical concern, atypical presentation, diagnostic uncertainty, or relevant comorbidities, milder CAP cases managed without CT may have been underrepresented. This spectrum effect may have increased the apparent discriminative performance of CURB-65 and may limit the direct generalizability of the findings to all patients with CAP.²⁰

One of the strengths of this study is that CURB-65 and CTSS were evaluated both separately and in predicted-probability models based on logistic regression. This approach allowed comparison of the independent and combined contributions of clinical and radiological variables to prognostic prediction in the same study population. In addition, CTSS was assessed independently by two blinded radiologists, and interobserver reliability was excellent.

Study Limitations

This study has several limitations. Its retrospective and single-center design limits generalizability. In addition, the CT-selected design may have shifted the patient spectrum toward more clinically complex or severe cases and contributed to overestimation of model performance. Patients with CAP who did not undergo thoracic CT were not included; therefore, the findings may not be directly applicable to milder CAP cases managed without CT. Although a bootstrap robustness analysis was performed to assess coefficient stability, potential overfitting cannot be completely excluded because of the number of candidate predictors relative to outcome events. The bootstrap procedure evaluated the stability of regression coefficients but did not provide optimism-corrected AUC estimates, calibration slopes, or Brier scores; therefore, it should not be interpreted as a complete internal validation of model performance. In addition, sparse events in patients with asthma resulted in unstable bootstrap estimates. External validation in an independent cohort remains necessary.

CONCLUSION

In patients with CAP, CURB-65 was a strong predictor of 30-day mortality and ICU admission. Mean CTSS was independently associated with 30-day mortality and ICU admission in conventional multivariable analyses, but showed lower standalone discrimination than CURB-65 and was not independently associated with length of hospital stay. The combined mortality model did not significantly improve discrimination over CURB-65 alone, whereas the combined ICU model significantly improved discrimination for ICU admission. These findings suggest that radiological extent may provide complementary prognostic information in selected CT-assessed patients, particularly when ICU triage is being considered.

Ethics

Ethics Committee Approval: The approval was obtained from University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital before data extraction and analysis (approval date: September 11, 2025; approval number: 2025/08-09).

Informed Consent: Patient consent was not obtained because of the retrospective design and use of anonymized data.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ö.G., Concept: Ö.G., Design: Ö.G., Data Collection or Processing: Ö.G., O.Ö., İ.A., Analysis or Interpretation: Ö.G., O.Ö., İ.A., Literature Search: Ö.G., Writing: Ö.G., O.Ö., İ.A.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

- GBD 2023 Lower Respiratory Infections and Antimicrobial Resistance Collaborators. Global burden of lower respiratory infections and aetiologies, 1990-2023: a systematic analysis for the Global Burden of Disease Study 2023. *Lancet Infect Dis.* 2026;26(4):343-361. Erratum in: *Lancet Infect Dis.* 2026;26(4):e214. Erratum in: *Lancet Infect Dis.* 2026;26(7):e281. [\[Crossref\]](#)
- Metlay JP, Waterer GW, Long AC, et al. Diagnosis and treatment of adults with community-acquired pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. *Am J Respir Crit Care Med.* 2019;200(7):e45-e67. [\[Crossref\]](#)
- Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax.* 2003;58(5):377-382. [\[Crossref\]](#)
- Lim WS, Smith DL, Wise MP, Welham SA; British Thoracic Society. British Thoracic Society community acquired pneumonia guideline and the NICE pneumonia guideline: how they fit together. *Thorax.* 2015;70(7):698-700. [\[Crossref\]](#)
- Chalmers JD, Singanayagam A, Hill AT. Predicting the need for mechanical ventilation and/or inotropic support for young adults admitted to the hospital with community-acquired pneumonia. *Clin Infect Dis.* 2008;47(12):1571-1574. [\[Crossref\]](#)
- Zhang ZX, Yong Y, Tan WC, Shen L, Ng HS, Fong KY. Prognostic factors for mortality due to pneumonia among adults from different age groups in Singapore and mortality predictions based on PSI and CURB-65. *Singapore Med J.* 2018;59(4):190-198. [\[Crossref\]](#)
- Marti C, Garin N, Groscurin O, Poncet A, Combescure C, Carballo S, Perrier A. Prediction of severe community-acquired pneumonia: a systematic review and meta-analysis. *Crit Care.* 2012;16(4):R141. [\[Crossref\]](#)
- Ilg A, Moskowitz A, Konanki V, et al. Performance of the CURB-65 score in predicting critical care interventions in patients admitted with community-acquired pneumonia. *Ann Emerg Med.* 2019;74(1):60-68. [\[Crossref\]](#)
- Liu JL, Xu F, Zhou H, et al. Expanded CURB-65: a new score system predicts severity of community-acquired pneumonia with superior efficiency. *Sci Rep.* 2016;6:22911. Erratum in: *Sci Rep.* 2018;8:47005. [\[Crossref\]](#)
- Francone M, Iafate F, Masci GM, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. *Eur Radiol.* 2020;30(12):6808-6817. [\[Crossref\]](#)
- Li K, Wu J, Wu F, et al. The clinical and chest CT features associated with severe and critical COVID-19 pneumonia. *Invest Radiol.* 2020;55(6):327-331. [\[Crossref\]](#)
- Fan H, Xu X, Dai W, et al. Semiquantitative chest computed tomography scoring system to estimate severity in pediatric community-acquired pneumonia. *Front Pediatr.* 2025;13:1556349. [\[Crossref\]](#)
- Unal Cetin E, Kurtkulagi O, Kamis F, et al. Advancing ICU mortality prediction in community-acquired pneumonia: combining fibrinogen-to-albumin ratio, CT severity score, PSI, and CURB-65. *Biomol Biomed.* 2025;25(9):2083-2091. [\[Crossref\]](#)
- Nemoto M, Nakashima K, Noma S, et al. Prognostic value of chest computed tomography in community-acquired pneumonia patients. *ERJ Open Res.* 2020;6(4):00079-2020. [\[Crossref\]](#)
- Chalmers JD, Singanayagam A, Akram AR, et al. Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis. *Thorax.* 2010;65(10):878-883. [\[Crossref\]](#)
- Andrew D, Shyam K, Cicilet S, Johny J. Assessment of interobserver reliability and predictive values of CT semiquantitative and severity scores in COVID lung disease. *Egypt J Radiol Nucl Med.* 2021;52(1):150. [\[Crossref\]](#)
- Mruk B, Plucińska D, Walecki J, Półtorak-Szymczak G, Sklinda K. Chest computed tomography (CT) severity scales in COVID-19 disease: a validation study. *Med Sci Monit.* 2021;27:e31283. [\[Crossref\]](#)
- Efthimiou O, Seo M, Chalkou K, Debray T, Egger M, Salanti G. Developing clinical prediction models: a step-by-step guide. *BMJ.* 2024;386:e078276. [\[Crossref\]](#)
- Suter-Widmer I, Christ-Crain M, Zimmerli W, Albrich W, Mueller B, Schuetz P; ProHOSP Study Group. Predictors for length of hospital stay in patients with community-acquired pneumonia: results from a Swiss multicenter study. *BMC Pulm Med.* 2012;12:21. [\[Crossref\]](#)
- Usher-Smith JA, Sharp SJ, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. *BMJ.* 2016;353:i3139. [\[Crossref\]](#)