SRIWIJAYA JOURNAL OF SURGERY
Vol. 10 No. 1 (2027) · pp. 1-23 · sriwijayasurgery.com
e-ISSN 2722-3558

ORIGINAL ARTICLE

Comparative Diagnostic Accuracy of the STUMBL (Battle) Score and Chest Trauma Score for Predicting In-Hospital Mortality in Adult Blunt Thoracic Trauma: An Age-Stratified STARD-Compliant Study at an Indonesian Tertiary Referral Center

Yoshua Ramothy Siahaan[1,2*], Arie Hasiholan Lumban Tobing[1,2], Theodorus[3]

[1] General Surgery Study Program, Department of Surgery, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia
[2] Department of Surgery, Dr. Mohammad Hoesin Central General Hospital, Palembang, Indonesia
[3] Department of Pharmacology, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

ARTICLE INFO

Article history:
Received: October 15, 2026
Revised: November 12, 2026
Accepted: December 2, 2026
Available Online: March 6, 2027
Published: June 1, 2027

Keywords:
blunt thoracic trauma; Chest Trauma Score; in-hospital mortality; prognostic accuracy; STUMBL score

Corresponding author:
Yoshua Ramothy Siahaan
Email:
dr.yoshuaramothy@gmail.com

DOI:
https://doi.org/10.37275/sjs.v10i1.182

Access licence:
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).
Copyright:
© 2027 The Authors. Authors retain copyright and grant the journal the right of first publication. Published by HM Publisher.

ABSTRACT

Background: The Chest Trauma Score (CTS) evaluates anatomical injury, whereas the STUMBL (Battle) score also incorporates physiological and comorbidity parameters.

Objective: To compare their accuracy for predicting in-hospital mortality in adults with blunt thoracic trauma.

Methods: This retrospective study included 106 adults hospitalized at Dr. Mohammad Hoesin Hospital, Indonesia, from January 2024 to September 2026. Admission scores were evaluated using receiver operating characteristic curves, Youden thresholds, and age-stratified analyses (18–60 versus >60 years).

Results: Thirteen patients (12.3%) died. Non-survivors had higher median CTS (6.00 versus 4.00) and mean STUMBL scores (28.77 versus 14.89; both $p < 0.001$). STUMBL had an area under the curve (AUC) of 0.974 and CTS an AUC of 0.966 (pairwise $p = 0.546$). At thresholds of ≥25 and ≥6, respectively, sensitivity was 92.31% for both scores. Specificity was 91.40% versus 90.32%, positive predictive value was 60.00% versus 57.14%, negative predictive value was 98.84% versus 98.82%, and accuracy was 91.51% versus 90.57%. Among adults aged >60 years (n = 23), STUMBL ≥28 had specificity of 95.00%, compared with 75.00% for CTS ≥6; sensitivity was 100.00% for both.

Conclusion: Both scores showed excellent mortality discrimination in this cohort. STUMBL had numerically higher specificity in older adults; prospective validation is needed.

1. Introduction

Thoracic trauma represents one of the most consequential determinants of acute morbidity, intensive care resource utilization, and preventable mortality among injured patients worldwide.[1,2] Across global trauma registries, chest injuries directly account for approximately 20% to 25% of all trauma-related fatalities and act as a primary physiological catalyst in more than half of deaths attributed to multisystem polytrauma.[2,3] Unlike penetrating thoracic wounds, which frequently declare their anatomical trajectory through overt external hemorrhage or immediate hemodynamic collapse, blunt thoracic trauma transfers high-energy kinetic and deceleration forces across the viscoelastic thoracic cage without breaching the cutaneous barrier.[4,5] Epidemiological surveillance from both high-income and low- and middle-income countries (LMICs) demonstrates that motor vehicle collisions, motorcycle crashes, pedestrian impacts, and falls from height account for over 70% of blunt chest wall and intrathoracic injuries, disproportionately affecting economically productive young and middle-aged men while imposing elevated case-fatality rates upon older adults.[6,7] Within the Indonesian archipelago and comparable Southeast Asian referral networks, rapid motorization combined with variable prehospital stabilization has amplified the clinical burden of blunt thoracic trauma presenting to tertiary emergency departments.[8–10] A recent systematic review and meta-analysis by Adal et al. across sub-Saharan African trauma centers reported a pooled mortality rate of 11.65% for blunt chest trauma compared with 5.63% for penetrating chest injuries, underscoring the insidious lethality of non-penetrating thoracic mechanisms.[1] Similarly, single-center and regional Indonesian trauma cohorts have documented hospital mortality rates ranging between 11.5% and 37.7% among patients sustaining severe blunt chest trauma, particularly when complicated by multiple rib fractures, parenchymal pulmonary contusion, and progressive hypoxemic respiratory failure.[8–10] In contemporary trauma systems, the disproportionate lethality of blunt thoracic mechanisms is compounded by the complex kinematics of high-velocity vehicular and motorcycle impacts, wherein rapid deceleration, direct anterior-posterior sternocostal compression, and lateral torso crush forces converge simultaneously upon the rib cage and intrathoracic viscera.[4–7] Large-scale registry analyses encompassing thousands of hospitalized chest injury victims indicate that road traffic crashes account for 60% to 87% of all blunt thoracic admissions in developing and rapidly industrializing nations, affecting predominantly male riders and drivers in their third through fifth decades of life while inflicting higher case-fatality rates among pedestrians and older adults.[2,4,5] Within Indonesia's referral hierarchy, tertiary

academic hospitals frequently receive patients who have experienced prolonged prehospital transit or inter-facility transfer without advanced ventilatory stabilization, magnifying the clinical imperative for rapid, objective, and highly accurate prognostic stratification immediately upon emergency department arrival.[8–10]

The pathophysiological lethality of blunt thoracic trauma stems from a tightly coupled triad of biomechanical chest wall instability, pleural space disruption, and alveolocapillary parenchymal injury.[11,12] Fractures of the thoracic rib cage generate severe intercostal nociceptive input that restricts inspiratory chest wall excursion, blunts cough velocity, impairs mucociliary clearance, and precipitates progressive basal atelectasis and nosocomial pneumonia.[13,14] When multiple contiguous ribs fracture or when fractures involve both hemithoraces, the structural integrity of the thoracic ring is compromised, increasing the work of breathing and predisposing the patient to paradoxical chest wall mechanics and acute ventilatory exhaustion.[11,12] Concurrently, kinetic energy transmitted through the chest wall shears the delicate alveolar-capillary membrane, producing parenchymal pulmonary contusion characterized by intra-alveolar hemorrhage, interstitial edema, surfactant inactivation, loss of pulmonary compliance, and refractory ventilation–perfusion mismatching.[13,14] Because radiographic consolidation and arterial hypoxemia following pulmonary contusion often evolve insidiously over the first 24 to 72 hours after injury, patients who appear hemodynamically compensated during the initial primary survey may experience precipitous cardiopulmonary decompensation on general surgical wards if early risk stratification fails to identify their underlying vulnerability.[15,16] At the microvascular and cellular level, blunt kinetic energy transmission across the chest wall initiates a biphasic pathophysiological injury cascade that extends far beyond simple mechanical fracture of cortical bone.[11–14] During the primary mechanical phase, sudden intrathoracic pressure spikes—mediated by spalling at liquid–gas interfaces, implosion of overexpanded alveoli, and inertial shearing between low-density alveolar parenchyma and high-density bronchovascular structures—rupture the alveolar-capillary basement membrane and precipitate intra-alveolar hemorrhage.[13,14] During the secondary inflammatory phase, which progresses over the ensuing 24 to 72 hours, extravasated erythrocytes, cell-free hemoglobin, and necrotic alveolar debris release damage-associated molecular patterns (DAMPs) that activate resident alveolar macrophages, trigger proinflammatory cytokine cascades, increase microvascular permeability, and wash out or inactivate endogenous pulmonary surfactant.[15,16] As regional lung compliance plummets and interstitial edema widens the alveolocapillary diffusion barrier, perfused pulmonary capillary beds fail to participate in gas exchange, creating a progressive intrapulmonary right-to-left shunt that manifests clinically as arterial hypoxemia refractory to low-flow supplemental oxygen.[15,16]

To replace subjective bedside intuition with standardized quantitative prognostication, several thoracic trauma scoring instruments have been formulated over the past two decades.[11–15] Among anatomically oriented frameworks, the Chest Trauma Score (CTS), originally conceptualized by Pressley et al. and subsequently formalized and validated by Chen et al., integrates four bedside-accessible clinical and radiographic parameters: patient age (<45, 45–65, and >65 years), number of fractured ribs (<3, 3–5, and >5 ribs), bilateral rib fracture involvement, and radiographic extent of pulmonary contusion (none, unilateral minor, bilateral minor, unilateral major, and bilateral major), generating a composite score ranging from 2 to 12 points.[11,12] Prospective and retrospective validations in diverse populations—including Egyptian, Korean, and South African trauma cohorts—have confirmed that elevated CTS values correlate strongly with mechanical ventilation requirement, pneumonia, and in-hospital mortality, with reported Area Under the Receiver Operating Characteristic Curve (AUC) values spanning from 0.711 in heterogeneous polytrauma series to 0.959 in dedicated blunt chest trauma cohorts.[13,14] Nevertheless, because CTS relies predominantly on static structural and radiographic findings alongside coarse age brackets, it does not directly incorporate acute physiological oxygenation status or pre-existing cardiopulmonary and pharmacological vulnerabilities.[12–14] Because the clinical and radiographic signs of evolving pulmonary contusion and chest wall splinting can be deceptive during the initial minutes of trauma resuscitation, reliance on unstandardized clinical impression alone risks either premature ward disposition—followed by unexpected respiratory arrest or emergent unplanned intubation—or indiscriminate intensive care unit over-triage that exhausts scarce critical care capacity.[13–15] When Pressley et al. and Chen et al. formulated and validated the Chest Trauma Score (CTS), their primary objective was to synthesize the four most reproducible structural and demographic determinants of thoracic trauma deterioration—chronological age, fractured rib count, bilateral rib cage disruption, and radiographic pulmonary contusion extent—into a rapid 2-to-12-point bedside scale requiring no specialized laboratory assays.[11,12] While subsequent cohorts in Egypt and South Africa confirmed that CTS thresholds between 5.0 and 6.5 effectively identify patients at high risk of pneumonia, mechanical ventilation, and death, multicenter evaluations in polytrauma populations have noted that purely anatomical scoring systems do not directly measure the patient's real-time gas-exchange reserve or pre-existing cardiopulmonary vulnerability.[13,14]

To address the absence of physiological and host-comorbidity dimensions in purely anatomical indices, Battle et al. developed and externally validated the Study of the Management of Blunt chest wall trauma (STUMBL) score—widely known as the Battle Score—as a clinical prediction model tailored for emergency department risk stratification.[15] The STUMBL score synthesizes five weighted predictors measured at emergency department arrival: patient age (1 point per 10-year decade), number of fractured ribs (3 points per fractured rib), pre-existing chronic lung disease (+5 points), pre-injury anticoagulant therapy (+4 points), and room-air peripheral capillary oxygen saturation (SpO_2, graded from 0 points for 95–100% up to +10 points for <80%).[15,16] Subsequent external validation studies across New Zealand, the United Kingdom, Italy, South Africa, Türkiye, and Indonesia have demonstrated strong discriminative performance of the STUMBL score and related chest trauma models (AUROC values ranging from 0.84 to 0.96) for predicting composite in-hospital complications, intensive care unit (ICU) admission, and prolonged hospitalization.[16–23] Furthermore, modifications and comparative validations such as the Revised Intensity Battle Score (RIBS) and multimodal radiomic-clinical frameworks have highlighted the central prognostic role of combining rib fracture burden with physiological oxygenation metrics.[24,25] By explicitly incorporating admission room-air pulse oximetry (SpO_2), pre-existing chronic lung disease, and pre-injury anticoagulant therapy alongside decade-scaled chronological age and linear rib-fracture weighting (3 points per fractured rib), the STUMBL (Battle) score shifted the paradigm of blunt chest wall trauma prognostication from static anatomical grading to integrated physiological-clinical risk

modeling.[15–18] Across international emergency medicine networks in the United Kingdom, New Zealand, Italy, South Africa, and Türkiye, the STUMBL score and related thoracic prediction models have demonstrated robust discrimination for predicting composite in-hospital complications, guiding analgesia protocols, and informing safe discharge versus inpatient admission decisions.[16–22] Concurrently, recent investigations into modified Battle scores, radiomic-assisted thoracic risk assessment, and machine-learning rib fracture models have underscored that accurate early risk stratification is the foundational prerequisite for targeted trauma care.[23–25]

Despite the parallel widespread adoption of CTS and the STUMBL score, a critical knowledge gap persists in the international trauma literature regarding their direct head-to-head diagnostic accuracy for predicting in-hospital mortality as a dedicated primary endpoint, particularly across distinct age strata in low- and middle-income tertiary referral settings.[19–23] Most prior investigations of the STUMBL score evaluated composite morbidity endpoints—combining lower respiratory tract infection, pleural effusion, ICU stay, and death—where non-fatal pulmonary complications numerically dominated the outcome events, leaving mortality-specific threshold calibration underexplored.[18,21,22] Conversely, while Isharanto et al. recently validated the STUMBL score against the global anatomical Injury Severity Score (ISS) in Malang, Indonesia, no published study in Indonesia has directly compared the thorax-specific anatomical-radiologic CTS against the clinical-physiological STUMBL score using Receiver Operating Characteristic (ROC) curve analysis and age-stratified contingency evaluation.[23] Moreover, because both CTS and STUMBL assign baseline points for chronological age—yet employ fundamentally different mathematical scaling rules (coarse 1-to-3 point brackets in CTS versus 1 point per decade alongside 3 points per rib and up to 10 points for hypoxemia in STUMBL)—it remains unknown whether baseline age weighting degrades diagnostic specificity among older adults (>60 years) compared with working-age adults (18–60 years).[19–25] Therefore, the present study was designed to execute a STARD-compliant, direct diagnostic accuracy comparison between the STUMBL (Battle) score and the Chest Trauma Score (CTS) for predicting in-hospital mortality in adult blunt thoracic trauma at a major South Sumatran tertiary referral center, incorporating component-level bivariate analysis, univariable and age-adjusted multivariable logistic modeling, and World Health Organization (WHO) age-stratified (18–60 versus >60 years) threshold calibration. We hypothesized that both scoring systems would achieve excellent overall mortality discrimination (AUC > 0.90), but that the integration of admission pulse oximetry and linear rib-fracture weighting within the STUMBL score would yield superior diagnostic specificity and positive predictive precision, particularly among older adult trauma patients. Specifically, evaluating CTS and the STUMBL score side by side within the same consecutive Indonesian patient cohort allows a rigorous empirical test of whether adding physiological pulse oximetry and comorbidity weighting improves mortality discrimination over radiographic pulmonary contusion and bilateral fracture grading, and whether the mathematical structure of age point assignment alters diagnostic specificity in older adults (>60 years) versus working-age adults (18–60 years).[19–25] By combining complete univariate, component-level bivariate, normality-audited between-group, univariable and age-adjusted multivariable logistic, and age-stratified ROC/Youden contingency analyses, this study establishes a comprehensive evidentiary benchmark for blunt thoracic trauma mortality prediction in Indonesian tertiary surgical practice.

2. Methods

Study Design, Setting, and Reporting Standards

This investigation was designed as an observational analytical diagnostic accuracy study employing a retrospective cohort approach at Dr. Mohammad Hoesin Central General Hospital (RSUP Dr. Mohammad Hoesin), Palembang, South Sumatra, Indonesia. Dr. Mohammad Hoesin Central General Hospital serves as the apex tertiary academic and trauma referral center for South Sumatra and neighboring provinces, managing high-acuity blunt and penetrating trauma referrals through an integrated 24-hour emergency department, multidetector computed tomography (CT) and digital radiography suite, cardiothoracic and general surgery divisions, and multidisciplinary intensive care unit. Data abstraction encompassed consecutive adult patients admitted with blunt thoracic trauma from January 2024 through September 2026 until the target sample size and complete cohort verification were achieved. The study design, execution, and reporting adhered strictly to the Standards for Reporting of Diagnostic Accuracy Studies (STARD 2015) and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for thoracic trauma risk modeling.[26,27] As a high-volume government tertiary referral hospital serving a catchment population exceeding eight million residents across South Sumatra, Bangka Belitung, Jambi, and Bengkulu, Dr. Mohammad Hoesin Central General Hospital receives both primary metropolitan trauma arrivals and secondary referrals from district hospitals. Standard institutional trauma management during the study period followed Advanced Trauma Life Support (ATLS) primary and secondary survey protocols, including immediate airway and cervical spine stabilization, bedside pulse oximetry and hemodynamic monitoring, portable or digital anteroposterior/posteroanterior chest radiography, thoracic computed tomography when clinically indicated, intercostal tube thoracostomy for pneumothorax or hemothorax, multimodal systemic and regional analgesia, chest physiotherapy, and invasive mechanical ventilation or surgical intervention when dictated by physiological deterioration.[26,27]

Study Population, Eligibility Criteria, and Sample Size Determination

The target population comprised all adult patients presenting to the emergency department or hospitalized within the thoracic and general surgery wards of Dr. Mohammad Hoesin Central General Hospital with a confirmed clinical and radiologic diagnosis of blunt thoracic trauma. Patients were eligible for inclusion if they fulfilled four mandatory criteria: (1) confirmed diagnosis of blunt thoracic trauma resulting from non-penetrating kinetic mechanisms (motor vehicle or motorcycle collisions, falls, pedestrian impacts, or blunt thoracic crush/impact); (2) chronological age ≥ 18 years at emergency department presentation; (3) hospitalization for inpatient surgical or conservative trauma management; and (4) complete medical record and chest imaging documentation permitting exact computation of all four Chest Trauma Score (CTS) components, all five STUMBL score components, and definitive in-hospital survival or mortality status. Patients were excluded if they had incomplete medical or radiographic records, presented with severe multisystem polytrauma dominated by life-threatening extra-thoracic injuries (such as severe traumatic brain injury or exsanguinating abdominal/pelvic hemorrhage that independently dictated mortality), or were dead on arrival / died before completion of initial emergency department diagnostic evaluation. The minimum required sample size for diagnostic accuracy evaluation was calculated a

priori using the single-proportion precision formula $n = (Z_{\alpha/2}^2 \times p \times (1 - p)) / d^2$, where $Z_{\alpha/2} = 1.96$ for a two-sided 95% confidence level, d = 0.10 (10% desired margin of error), and expected sensitivity p = 0.80 and specificity p = 0.87 based on prior external validation data reported by Giamello et al.[18] Substituting p = 0.80 yielded a minimum sample requirement of $n = (1.96^2 \times 0.80 \times 0.20) / 0.10^2 = 61.47 \approx 61$ patients, whereas substituting p = 0.87 yielded $n = 43.60 \approx 43$ patients. Selecting the larger estimate (n = 61) and adding a 10% buffer to account for potential medical record incompleteness established a minimum required sample size of 67 patients. Using consecutive total sampling across the study period, 106 eligible adult patients with complete clinical, radiologic, and outcome data were enrolled, exceeding the minimum sample threshold by 58.2%. The strict exclusion of patients whose clinical course was dominated by catastrophic non-thoracic polytrauma—such as severe traumatic brain injury with Glasgow Coma Scale ≤ 8, high-grade solid-organ abdominal hemorrhage requiring damage-control laparotomy, or unstable pelvic ring exsanguination—was implemented deliberately to preserve internal validity and prevent extra-thoracic causes of death from confounding the thorax-specific diagnostic accuracy of CTS and the STUMBL score.[14,23] Similarly, excluding patients who arrived without vital signs (dead on arrival) or who died in the resuscitation bay prior to completion of diagnostic chest imaging ensured that all analyzed subjects had complete, verified anatomical and physiological component measurements at baseline. The achieved cohort of 106 consecutive patients (93 survivors and 13 non-survivors) provided 100.0% data completeness across all demographic, component, composite score, and mortality fields, eliminating any need for statistical imputation.

Index Tests (CTS and STUMBL Score) and Reference Standard Operational Definitions

All predictor variables required to compute both index tests were abstracted from initial emergency department clinical records and admission chest radiographs or thoracic CT scans prior to the occurrence of the reference outcome. The first index test, the Chest Trauma Score (CTS), was calculated on a quantitative scale (theoretical range 2–12 points; observed cohort range 1–8 points, reflecting isolated contusion/chest wall cases with zero rib fractures assigned 1 point on the rib subscore) by summing four weighted anatomical-demographic components according to Pressley et al. and Chen et al.:[11,12] (1) age score: <45 years = 1 point, 45–65 years = 2 points, and >65 years = 3 points; (2) rib fracture score: <3 fractured ribs = 1 point, 3–5 fractured ribs = 2 points, and >5 fractured ribs = 3 points; (3) bilateral rib fracture involvement: absent/unilateral = 0 points and bilateral rib fractures present = 2 points; and (4) parenchymal pulmonary contusion score: none = 0 points, unilateral minor = 1 point, bilateral minor = 2 points, unilateral major = 3 points, and bilateral major = 4 points. The second index test, the Study of the Management of Blunt chest wall trauma (STUMBL or Battle) score, was calculated by summing five weighted clinical, physiological, and comorbidity parameters according to Battle et al.:[15] (1) patient age: 1 point per 10-year decade of life; (2) number of fractured ribs: 3 points per fractured rib; (3) pre-injury anticoagulant use: absent = 0 points and present = +4 points; (4) pre-existing chronic lung disease (documented chronic obstructive pulmonary disease [COPD], chronic bronchial asthma, or interstitial/restrictive lung disease): absent = 0 points and present = +5 points; and (5) admission peripheral capillary oxygen saturation (SpO_2) measured by pulse oximetry on presentation: 95–100% = 0 points, 90–94% = +2 points, 85–89% = +5 points, 80–84% = +8 points, and <80% = +10 points.[28,29] The reference standard (gold standard outcome) was all-cause in-hospital mortality attributed to blunt thoracic trauma and its acute cardiopulmonary sequelae during the index hospitalization, dichotomized nominally as survived to hospital discharge versus died in hospital. To evaluate age-specific diagnostic stability in accordance with World Health Organization (WHO) geriatric classification, the cohort was further stratified into working-age adults (18–60 years; n = 83) and older adults (>60 years; n = 23).[30] For radiographic component scoring within CTS, the number of fractured ribs and the presence of bilateral rib fractures were determined from admission chest radiographs and confirmed on thoracic computed tomography reports verified by board-certified radiologists and attending thoracic/general surgeons. Parenchymal pulmonary contusion was graded according to the standard Pressley–Chen four-tier radiographic rubric based on the anatomical extent of non-segmental airspace opacification/consolidation across unilateral or bilateral lung fields (none = 0, unilateral minor = 1, bilateral minor = 2, unilateral major = 3, bilateral major = 4).[11,12,29] For physiological and clinical component scoring within the STUMBL score, admission peripheral capillary oxygen saturation (SpO_2) was recorded from calibrated fingertip pulse oximetry obtained upon initial emergency department triage prior to definitive mechanical ventilation, while pre-existing chronic lung disease (COPD, chronic asthma, or chronic parenchymal lung disease) and pre-injury oral anticoagulant therapy (vitamin K antagonists or direct oral anticoagulants) were abstracted from structured admission medical histories and medication reconciliations.[15,21,22] Investigators computing the composite CTS and STUMBL scores applied the standardized published scoring rubrics uniformly across all 106 records.

Statistical Analysis Plan

Statistical analyses were performed using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA) and cross-verified in Python 3.11 (NumPy) with two-sided statistical significance set at $\alpha = 0.05$. Descriptive univariate analyses summarized categorical variables (sex, WHO age group, bilateral rib fracture status, pulmonary contusion grade, chronic lung disease, anticoagulant use, and in-hospital mortality) as frequencies (n) and exact percentages (%) with Wilson 95% confidence intervals (CIs). Continuous and discrete quantitative variables (chronological age, number of fractured ribs, admission SpO_2, total CTS score, and total STUMBL score) were evaluated for distributional normality within survival strata using both the Kolmogorov–Smirnov test (with Lilliefors significance correction) and the Shapiro–Wilk test (W), and were reported comprehensively using both parametric metrics (mean ± standard deviation [SD], standard error [SE], and 95% CI for the mean) and non-parametric metrics (median, interquartile range [IQR: Q1–Q3], minimum–maximum range, and 5% trimmed mean). Bivariate associations between individual score components and in-hospital mortality were assessed using Pearson's chi-square test (χ^2), Yates' continuity correction, likelihood ratio chi-square (G^2), and two-sided Fisher's exact test when >20% of contingency cells had expected counts <5, supplemented by Mann–Whitney U tests and rank-biserial correlation coefficients (r_{rb}) for ordinal/continuous distributions. Because Shapiro–Wilk normality testing confirmed that total CTS scores deviated significantly from normality in both survivors ($W = 0.916$, $p < 0.001$) and non-survivors ($W = 0.862$, $p = 0.041$), between-group comparison of total CTS was performed using the non-parametric Mann–Whitney U test (reporting U, Wilcoxon W, Z, and r_{rb}). Conversely,

because total STUMBL scores satisfied normality in both survivors (W = 0.974, p = 0.065) and non-survivors (W = 0.955, p = 0.671) but exhibited unequal variances on Levene's test (F = 8.044, p = 0.005), between-group comparison of total STUMBL was performed primarily using Welch's independent-samples t-test not assuming equal variances (reporting t, Welch-Satterthwaite degrees of freedom [df], mean difference, and 95% CI), alongside the pooled-variance t-test and Mann–Whitney U test for completeness. Univariable binary logistic regression models (reporting regression slope β, SE, Wald χ^2, odds ratio [OR] per 1-point increment, 95% CI, and Cox–Snell and Nagelkerke R^2) and age-adjusted multivariable logistic regression models (adjusting dichotomous high-risk score thresholds for WHO age stratum >60 versus 18–60 years) were fitted to quantify independent mortality odds. Discriminative performance was evaluated via non-parametric Receiver Operating Characteristic (ROC) curves, reporting Area Under the Curve (AUC ± SE) and asymptotic 95% CIs, with pairwise AUC comparisons conducted using the DeLong/Hanley–McNeil method. Optimal prognostic cut-offs were identified by maximizing the Youden Index (J = Sensitivity + Specificity − 1), and 2×2 contingency tables were constructed for the overall cohort (N = 106), older adults (>60 years; n = 23), and working-age adults (18–60 years; n = 83) to compute Sensitivity, Specificity, Positive Predictive Value (PPV / NDP), Negative Predictive Value (NPV / NDN), Overall Accuracy, Positive and Negative Likelihood Ratios (LR^+ and LR^-), and crude/Haldane–Anscombe corrected Odds Ratios. In non-parametric Receiver Operating Characteristic (ROC) curve integration, because smaller composite scores indicate stronger evidence for survival (n = 93) whereas larger composite scores indicate stronger evidence for in-hospital mortality (n = 13), the Area Under the Curve (AUC), standard error (SE), asymptotic p-value, and 95% confidence interval are mathematically invariant to state polarity by the symmetry of the Mann–Whitney U statistic. At any cut-off midpoint c, the survival sensitivity and false-positive coordinates for Score ≤ c map directly to the mortality specificity and false-negative coordinates for Score > c. All 2×2 diagnostic contingency matrices were constructed in standard clinical epidemiologic orientation with In-Hospital Mortality (Died: n = 13 overall; n = 3 in >60 years; n = 10 in 18–60 years) in the disease-positive column and Score ≥ Cut-off in the test-positive row, accompanied by Wilson score 95% CIs for all proportions and Haldane–Anscombe continuity-corrected Odds Ratios when any cell contained zero observations.

Ethical approval

This study received institutional ethics approval from the Health Research Ethics Committee of Dr. Mohammad Hoesin Central General Hospital, Indonesia (approval number DP/04.03/D.XVII.9.4/ETIK/286/2026), and was conducted in full accordance with the Declaration of Helsinki. Due to the retrospective observational nature of medical record abstraction, individual written informed consent was waived by the Ethics Committee while strict patient de-identification was maintained.

3. Results

Cohort Flow, Baseline Demographic Profile, and Survival Outcomes

Across the study period from January 2024 to September 2026, a total of 106 consecutive adult patients hospitalized with blunt thoracic trauma at Dr. Mohammad Hoesin Central General Hospital fulfilled all eligibility criteria and were included in the final diagnostic accuracy analysis, with zero missing records across any component variable. Among the 106 enrolled patients, 93 patients (87.7%; exact 87.74%, Wilson 95% CI: 80.15–92.69%) survived to hospital discharge, whereas 13 patients (12.3%; exact 12.26%, Wilson 95% CI: 7.31–19.85%) experienced in-hospital mortality. The complete STARD 2015 participant flow diagram, parallel index test evaluation architecture, and survival proportions across sex and WHO age strata are illustrated in Figure 1.

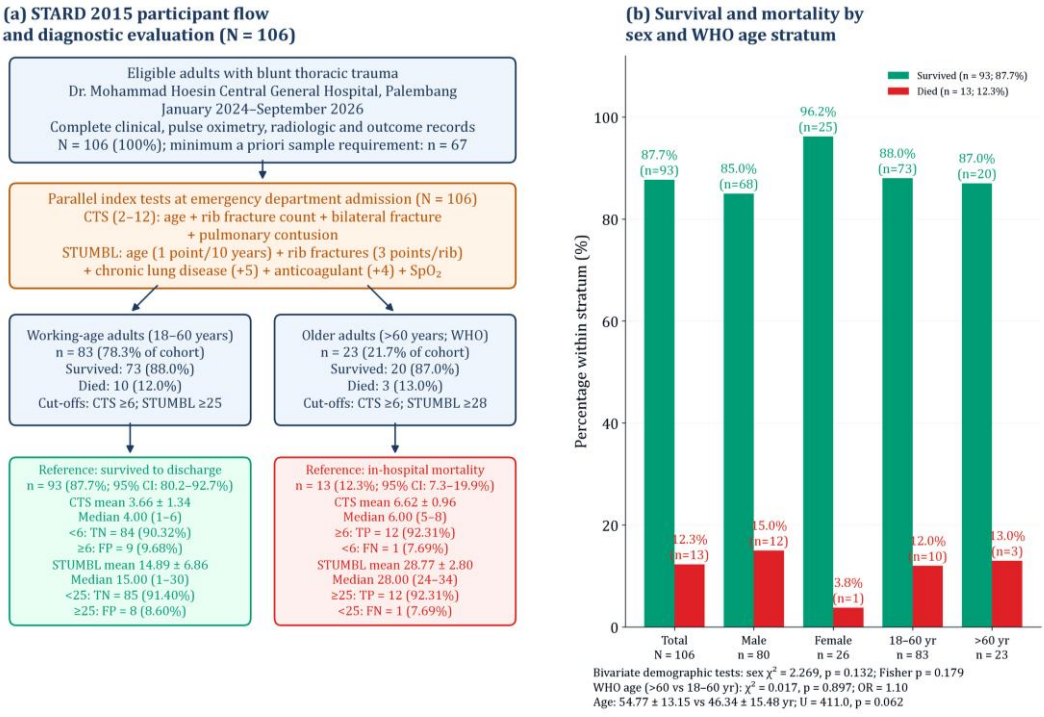


Figure 1. STARD 2015 participant flow diagram and baseline demographic mortality distribution (N = 106). (a) Consecutive enrollment, parallel index test evaluation (CTS and STUMBL score), WHO age stratification (18–60 years [n = 83] vs. >60 years [n = 23]), and reference standard survival outcomes (Survived: n = 93 [87.7%]; Died: n = 13 [12.3%]). (b) Survival versus in-hospital mortality proportions across sex and WHO age strata.

The baseline demographic characteristics, sex distribution, and continuous and categorical age distributions of the study cohort are detailed in Table 1. The study cohort was predominantly male, comprising 80 men (75.5%) and 26 women (24.5%), corresponding to a male-to-female ratio of 3.08:1. Within the male subgroup (n = 80), 68 patients (85.0%) survived and 12 patients (15.0%) died, whereas within the female subgroup (n = 26), 25 patients (96.2%) survived and only 1 patient (3.8%) died. Although male patients exhibited a numerically higher mortality proportion than female patients (15.0% versus 3.8%; crude OR = 4.41, 95% CI: 0.54–35.75), contingency analysis demonstrated no statistically significant association between sex and in-hospital mortality on Pearson's chi-square test (χ^2 = 2.269, df = 1, p = 0.132), Yates' continuity correction (χ^2 = 1.351, df = 1, p = 0.245), or two-sided Fisher's exact test (p = 0.179; 1 cell [25.0%] had an expected count <5, with a minimum expected count of 3.19). Across the entire cohort (N = 106), chronological age ranged from 18 to 79 years, with a parametric mean of 47.38 ± 15.41 years (95% CI: 44.41–50.35) and a median of 49.00 years (IQR: 24.00; Q1–Q3: 36.00–60.00). Patients who died in hospital were on average 8.43 years older than survivors (mean 54.77 ± 13.15 years [95% CI: 46.82–62.71; median 57.00, IQR: 49.00–61.00, range 29–77] versus 46.34 ± 15.48 years [95% CI: 43.16–49.53; median 48.00, IQR: 36.00–58.50, range 18–79]), showing a borderline trend toward older age among non-survivors on independent-samples Student's t-test (t = −1.869, df = 104, p = 0.064) and Mann–Whitney U test (U = 411.0, Z = −1.865, p = 0.062, r_{rb} = 0.320). When stratified according to the WHO geriatric threshold into working-age adults (18–60 years; n = 83, 78.3%) and older adults (>60 years; n = 23, 21.7%), in-hospital mortality was nearly identical between the two age strata: 10 of 83 patients (12.0%) aged 18–60 years died (with 73 [88.0%] surviving), compared with 3 of 23 patients (13.0%) aged >60 years (with 20 [87.0%] surviving; χ^2 = 0.017, df = 1, p = 0.897; Fisher's exact p = 1.000; OR = 1.10, 95% CI: 0.28–4.35).

A granular inspection of the complete 45-level integer age distribution across the 106 patients (Table 1) further illuminates the demographic structure of blunt thoracic trauma in South Sumatra. Among the 93 surviving patients, ages spanned the entire adult spectrum from 18.00 to 79.00 years, with clusters observed in young adulthood (ages 18–25 years: n = 15 [16.1% of survivors]), early-to-middle working age (ages 27–44 years: n = 27 [29.0%]), late working age (ages 46–60 years: n = 31 [33.3%]), and older adulthood (ages 62–79 years: n = 20 [21.5%]). Among the 13 patients who died in hospital, no fatalities occurred between ages 18 and 28 years (0/16; 0.0%); the youngest fatal case occurred at age 29 years (n = 1), followed by age 32 years (n = 1), age 45 years (n = 1), age 53 years (n = 1), age 55 years (n = 1), age 56 years (n = 1), age 57 years (n = 1), age 58 years (n = 1), age 60 years (n = 2), age 62 years (n = 1), age 68 years (n = 1), and age 77 years (n = 1). Thus, 9 of the 13 fatalities (69.2%) clustered within the narrow 15-year window between ages 53 and 68 years, explaining why the non-survivor group exhibited a narrower interquartile range (IQR = 12.00 years; Q1–Q3: 49.00–61.00) and smaller standard deviation (13.15 years) than the survivor group (IQR = 22.50 years; Q1–Q3: 36.00–58.50; SD = 15.48 years). This clustering near the 60-year boundary also explains why dichotomizing age at the WHO threshold of >60 years versus 18–60 years yielded virtually identical mortality rates across the two strata (13.0% [3/23] in >60 years versus 12.0% [10/83] in 18–60 years; p = 0.897), because the six deaths occurring between ages 53 and 60 years fell just inside the upper boundary of the 18–60-year working-age group.

Table 1. Baseline Demographic Characteristics, Survival Status, and WHO Age-Stratum Distribution of the Blunt Thoracic Trauma Cohort (N = 106).

Demographic / Clinical Characteristic	Total Cohort (N = 106; 100.0%)	Survived to Discharge (n = 93; 87.7%)	In-Hospital Mortality (n = 13; 12.3%)	Test Statistic, Effect Size & p-value
In-hospital survival outcome — n (%) [Exact proportion; Wilson 95% CI]	106 (100.0%)	93 (87.7%) [87.74%; 95% CI: 80.15–92.69%]	13 (12.3%) [12.26%; 95% CI: 7.31–19.85%]	Reference outcome standard (Mortality prevalence = 12.3%)
Male sex — n (%) (Row survival vs. mortality %)	80 (75.5%)	68 (85.0% of males; 73.1% of survivors)	12 (15.0% of males; 92.3% of non-survivors)	χ^2 = 2.269 (df = 1), p = 0.132[a] Fisher's exact p = 0.179[b]
Female sex — n (%) (Row survival vs. mortality %)	26 (24.5%)	25 (96.2% of females; 26.9% of survivors)	1 (3.8% of females; 7.7% of non-survivors)	OR (Male vs. Female) = 4.41 (95% CI: 0.54–35.75)
Working-age adults (18–60 years; WHO) — n (%) (Row survival vs. mortality %)	83 (78.3%)	73 (88.0% of stratum; 78.5% of survivors)	10 (12.0% of stratum; 76.9% of non-survivors)	χ^2 = 0.017 (df = 1), p = 0.897[a] Fisher's exact p = 1.000[b]
Older adults (>60 years; WHO) — n (%) (Row survival vs. mortality %)	23 (21.7%)	20 (87.0% of stratum; 21.5% of survivors)	3 (13.0% of stratum; 23.1% of non-survivors)	OR (>60 vs. 18–60 yr) = 1.10 (95% CI: 0.28–4.35)
Age, years — parametric mean ± SD [95% CI for mean]	47.38 ± 15.41 [44.41–50.35]	46.34 ± 15.48 [43.16–49.53]	54.77 ± 13.15 [46.82–62.71]	Student's t = −1.869 (df = 104), p = 0.064[c]
Age, years — non-parametric median (min–max) [IQR: Q1–Q3]	49.00 (18–79) [IQR: 24.00; 36.00–60.00]	48.00 (18–79) [IQR: 22.50; 36.00–58.50]	57.00 (29–77) [IQR: 12.00; 49.00–61.00]	Mann–Whitney U = 411.0, Z = −1.865, p = 0.062[d] (r_{rb} = 0.320); Ungrouped χ^2 = 55.438 (df = 44), p = 0.116[a]

[a]Pearson's chi-square test; [b]Two-sided Fisher's exact test (1 cell [25.0%] with expected count <5; minimum expected count = 3.19 for sex); [c]Independent-samples Student's t-test; [d]Mann–Whitney U test. In the ungrouped age contingency table (45 integer age values, df = 44), 89 of 90 cells (98.9%) had expected counts <5 (minimum expected count = 0.12). Abbreviations: CI, confidence interval; IQR, interquartile range; OR, odds ratio; r_{rb}, rank-biserial correlation; SD, standard deviation; WHO, World Health Organization.

Component-Level Bivariate Analysis of Chest Trauma Score (CTS) and STUMBL Score

To elucidate how individual anatomical, physiological, and comorbidity variables drive the predictive performance of the two composite instruments, each component of the CTS and STUMBL score was cross-tabulated against in-hospital mortality. The scoring rubric and component-level bivariate contingency statistics for the Chest Trauma Score (CTS) are detailed in Table 2. Within the CTS framework, all three structural-anatomical components—number of fractured ribs, bilateral rib fracture involvement, and parenchymal pulmonary contusion—exhibited highly significant associations with in-hospital mortality, whereas the chronological age component did not reach statistical significance on ungrouped contingency testing (χ^2 = 55.438, df = 44, p = 0.116, where 89 of 90 cells [98.9%] had expected counts <5) or grouped CTS age-bracket analysis (<45 years [1 point]: 2/46 died [4.3%]; 45–65 years [2 points]: 9/48 died [18.8%]; >65 years [3 points]: 2/12 died [16.7%]; χ^2 = 4.772, df = 2, p = 0.092). In contrast, the number of fractured ribs demonstrated a powerful, monotonic relationship with mortality across the 0-to-8 rib spectrum (χ^2 = 33.575, df = 8, p < 0.001; likelihood ratio G^2 = 30.973, df = 8, p < 0.001). Specifically, zero deaths occurred among patients with 0 fractured ribs (0/17; 0.0%), 1 rib (0/1; 0.0%), 2 ribs (0/8; 0.0%), or 3 ribs (0/25; 0.0%); mortality rose to 5.3% (1/19) at 4 ribs and 9.1% (1/11) at 5 ribs, before surging to 35.7% (5/14) at 6 ribs, 55.6% (5/9) at 7 ribs, and 50.0% (1/2) at 8 ribs. When aggregated into CTS rib fracture categories, mortality was 0.0% (0/26) for <3 ribs (1 point) and 3.6% (2/55) for 3–5 ribs (2 points), compared with 44.0% (11/25) for >5 fractured ribs (3 points; grouped χ^2 = 30.840, df = 2, p < 0.001). Numerically, non-survivors sustained nearly twice as many fractured ribs as survivors (mean 6.31 ± 1.03 [median 6.00, IQR: 6.00–7.00, range 4–8] versus 3.27 ± 2.05 [median 3.00, IQR: 2.00–5.00, range 0–8]; Mann–Whitney U = 115.0, Z = −4.780, p < 0.001, r_{rb} = 0.810). Bilateral rib fracture involvement (2 points in CTS) was present in 8 patients (7.5%), of whom 4 (50.0%) died and 4 (50.0%) survived, whereas among the 98 patients (92.5%) without bilateral fractures, 9 (9.2%) died and 89 (90.8%) survived (χ^2 = 11.452, df = 1, p = 0.001; Fisher's exact p = 0.008; OR = 9.89, 95% CI: 2.09–46.72). Similarly, parenchymal pulmonary contusion severity was strongly associated with mortality (χ^2 = 41.062, df = 4, p < 0.001; G^2 = 30.752, df = 4, p < 0.001): mortality was only 2.6% (2/76) among patients without pulmonary contusion (74 survived, 2 died), rose to 28.0% (7/25) in unilateral minor contusion (18 survived, 7 died), and reached 100.0% in bilateral minor contusion (3/3 died) and unilateral major contusion (1/1 died), while 1 patient with bilateral major contusion survived (0/1 died); overall, the presence of any pulmonary contusion conferred a 21.42-fold increase in mortality odds (11/30 [36.7%] versus 2/76 [2.6%]).

Table 2. Scoring Rubric and Bivariate Association of Individual Chest Trauma Score (CTS) Components with In-Hospital Mortality (N = 106).

CTS Component & Scoring Weight	Total Cohort (N = 106)	Survived (n = 93; 87.7%)	Died (n = 13; 12.3%)	Mortality Rate Within Category (%)	Contingency & Non-Parametric Statistics (χ^2 / U, df, p-value)
1. Age component (CTS rubric: 1–3 points) • <45 years (1 point) • 45–65 years (2 points) • >65 years (3 points)	46 (43.4%) 48 (45.3%) 12 (11.3%)	44 (47.3%) 39 (41.9%) 10 (10.8%)	2 (15.4%) 9 (69.2%) 2 (15.4%)	4.3% (2/46) 18.8% (9/48) 16.7% (2/12)	Ungrouped χ^2 = 55.438 (df = 44), p = 0.116[a] Grouped χ^2 = 4.772 (df = 2), p = 0.092[a] Mann–Whitney U = 411.0, p = 0.062[b]
2. Number of fractured ribs (CTS rubric: 1–3 points) • <3 fractured ribs: 0–2 ribs (1 point) (0 ribs: 17 [17/0]; 1 rib: 1 [1/0]; 2 ribs: 8 [8/0]) • 3–5 fractured ribs (2 points) (3 ribs: 25 [25/0]; 4 ribs: 19 [18/1]; 5 ribs: 11 [10/1]) • >5 fractured ribs: 6–8 ribs (3 points) (6 ribs: 14 [9/5]; 7 ribs: 9 [4/5]; 8 ribs: 2 [1/1])	26 (24.5%) 55 (51.9%) 25 (23.6%)	26 (28.0%) 53 (57.0%) 14 (15.1%)	0 (0.0%) 2 (15.4%) 11 (84.6%)	0.0% (0/26) 3.6% (2/55) 44.0% (11/25)	Exact 0–8 ribs χ^2 = 33.575 (df = 8), p < 0.001*[a] Likelihood ratio G^2 = 30.973, p < 0.001* Grouped CTS χ^2 = 30.840 (df = 2), p < 0.001* Mann–Whitney U = 115.0, Z = −4.780, p < 0.001*[b]
3. Bilateral rib fractures (CTS rubric: 0 or 2 points) • Unilateral or absent (0 points) • Bilateral rib involvement (2 points)	98 (92.5%) 8 (7.5%)	89 (95.7%) 4 (4.3%)	9 (69.2%) 4 (30.8%)	9.2% (9/98) 50.0% (4/8)	χ^2 = 11.452 (df = 1), p = 0.001*[a] Fisher's exact p = 0.008*[c] OR = 9.89 (95% CI: 2.09–46.72)

CTS Component & Scoring Weight	Total Cohort (N = 106)	Survived (n = 93; 87.7%)	Died (n = 13; 12.3%)	Mortality Rate Within Category (%)	Contingency & Non-Parametric Statistics (χ^2 / U, df, p-value)
4. Pulmonary contusion (CTS rubric: 0–4 points) • None (0 points) • Unilateral minor (1 point) • Bilateral minor (2 points) • Unilateral major (3 points) • Bilateral major (4 points)	76 (71.7%) 25 (23.6%) 3 (2.8%) 1 (0.9%) 1 (0.9%)	74 (79.6%) 18 (19.4%) 0 (0.0%) 0 (0.0%) 1 (1.1%)	2 (15.4%) 7 (53.8%) 3 (23.1%) 1 (7.7%) 0 (0.0%)	2.6% (2/76) 28.0% (7/25) 100.0% (3/3) 100.0% (1/1) 0.0% (0/1)	χ^2 = 41.062 (df = 4), p < 0.001*[a] Likelihood ratio G^2 = 30.752, p < 0.001* Any contusion (≥1 pt) vs. None: 11/30 (36.7%) vs. 2/76 (2.6%), OR = 21.42*

*[a]Pearson's chi-square test; [b]Mann–Whitney U test on continuous/ordinal distribution; [c]Two-sided Fisher's exact test. *Statistically significant at p < 0.05. Total CTS score ranges from 2 to 12 points (observed cohort range: 1 to 8 points, reflecting 0-rib contusion-only cases scored at 1 point). Abbreviations: CI, confidence interval; CTS, Chest Trauma Score; df, degrees of freedom; OR, odds ratio.*

The scoring rubric and component-level bivariate associations for the STUMBL (Battle) score are detailed in Table 3. Within the STUMBL score, the two dominant predictors of in-hospital mortality were admission peripheral oxygen saturation (SpO_2; χ^2 = 81.217, df = 15, p < 0.001; G^2 = 63.987, p < 0.001) and fractured rib count (χ^2 = 33.575, df = 8, p < 0.001). Examination of the 16-level admission SpO_2 cross-tabulation revealed a striking physiological boundary at 90% oxygen saturation: all 13 non-survivors (100.0%) presented to the emergency department with an admission $SpO_2 \leq 90\%$—specifically, 10 of 13 non-survivors (76.9%) presented with SpO_2 strictly below 90% (80%: 3 died, 0 survived; 81%: 1 died, 0 survived; 83%: 1 died, 0 survived; 84%: 1 died, 1 survived; 85%: 2 died, 0 survived; 89%: 2 died, 1 survived, while 1 patient with 87% survived), and 3 of 13 non-survivors (23.1%) presented with SpO_2 exactly at 90% (3 died, 3 survived). Conversely, among the 87 patients who presented with an admission $SpO_2 \geq 92\%$ (92%: 7; 93%: 5; 94%: 10; 95%: 5; 96%: 10; 97%: 10; 98%: 18; 99%: 22), mortality was 0.0% (0/87 died, 87/87 survived). Across STUMBL SpO_2 scoring tiers, mortality increased from 0.0% (0/65) in the 95–100% tier (0 points) to 10.7% (3/28) in the 90–94% tier (+2 points), 66.7% (4/6) in the 85–89% tier (+5 points), and 85.7% (6/7) in the 80–84% tier (+8 points; grouped χ^2 = 60.749, df = 3, p < 0.001). Continuous admission SpO_2 was markedly depressed in non-survivors compared with survivors (mean 85.08 ± 4.11% [95% CI: 82.59–87.56%; median 85.00%, IQR: 80.50–89.50%, range 80–90%] versus 95.98 ± 3.06% [95% CI: 95.35–96.61%; median 97.00%, IQR: 94.00–98.00%, range 84–99%]; Mann–Whitney U = 21.0, Z = −5.668, p < 0.001, r_{rb} = 0.965). By contrast, pre-existing chronic lung disease (+5 points) was documented in only 5 patients (4.7%: 4 survived [80.0%], 1 died [20.0%] versus 12/101 [11.9%] without chronic lung disease; χ^2 = 0.292, df = 1, p = 0.589; Fisher's exact p = 0.487), and pre-injury anticoagulant use (+4 points) was documented in only 1 patient (0.9%: 1 survived [100.0%], 0 died [0.0%] versus 13/105 [12.4%] without anticoagulants; χ^2 = 0.141, df = 1, p = 0.707; Fisher's exact p = 1.000), reflecting low baseline prevalence of these two comorbidities in this predominantly working-age Indonesian trauma cohort.

Table 3. Scoring Rubric and Bivariate Association of Individual STUMBL (Battle) Score Components with In-Hospital Mortality (N = 106).

STUMBL Component & Scoring Weight	Total Cohort (N = 106)	Survived (n = 93; 87.7%)	Died (n = 13; 12.3%)	Mortality Rate Within Category (%)	Contingency & Continuous Statistics (χ^2 / U, df, p-value)
1. Age (1 point per 10-year decade of life) • Continuous years — Mean ± SD [Median (IQR)]	47.38 ± 15.41 [49.00 (36.00–60.00)]	46.34 ± 15.48 [48.00 (36.00–58.50)]	54.77 ± 13.15 [57.00 (49.00–61.00)]	18–60 yr: 12.0% >60 yr: 13.0%	Ungrouped χ^2 = 55.438 (df = 44), p = 0.116[a] Student's t = −1.869, p = 0.064 Mann–Whitney U = 411.0, p = 0.062
2. Number of fractured ribs (3 points per rib) • Continuous count — Mean ± SD [Median (IQR)] • ≥4 fractured ribs (n = 55) vs. 0–3 ribs (n = 51) • ≥6 fractured ribs (n = 25) vs. 0–5 ribs (n = 81)	3.64 ± 2.19 [4.00 (2.75–5.00)] 55 (51.9%) 25 (23.6%)	3.27 ± 2.05 [3.00 (2.00–5.00)] 42 (45.2%) 14 (15.1%)	6.31 ± 1.03 [6.00 (6.00–7.00)] 13 (100.0%) 11 (84.6%)	0–3 ribs: 0.0% (0/51) ≥4 ribs: 23.6% (13/55) ≥6 ribs: 44.0% (11/25)	χ^2 = 33.575 (df = 8), p < 0.001*[a] Likelihood ratio G^2 = 30.973, p < 0.001* Mann–Whitney U = 115.0, Z = −4.780, p < 0.001* (r_{rb} = 0.810)

STUMBL Component & Scoring Weight	Total Cohort (N = 106)	Survived (n = 93; 87.7%)	Died (n = 13; 12.3%)	Mortality Rate Within Category (%)	Contingency & Continuous Statistics (χ^2 / U, df, p-value)
3. Chronic lung disease (COPD / asthma: +5 points) • Absent (0 points) • Present (+5 points)	101 (95.3%) 5 (4.7%)	89 (95.7%) 4 (4.3%)	12 (92.3%) 1 (7.7%)	11.9% (12/101) 20.0% (1/5)	χ^2 = 0.292 (df = 1), p = 0.589[a] Fisher's exact p = 0.487[b] OR = 1.85 (95% CI: 0.19–17.96)
4. Pre-injury anticoagulant use (+4 points) • Absent (0 points) • Present (+4 points)	105 (99.1%) 1 (0.9%)	92 (98.9%) 1 (1.1%)	13 (100.0%) 0 (0.0%)	12.4% (13/105) 0.0% (0/1)	χ^2 = 0.141 (df = 1), p = 0.707[a] Fisher's exact p = 1.000[b] Haldane OR = 2.27 (95% CI: 0.09–58.78)
5. Admission pulse oximetry SpO_2 (0 to +10 points) • Continuous SpO_2 (%) — Mean ± SD [Median (IQR)] • 95–100% (0 points) (95%: 5/0; 96%: 10/0; 97%: 10/0; 98%: 18/0; 99%: 22/0) • 90–94% (+2 points) (90%: 3/3; 92%: 7/0; 93%: 5/0; 94%: 10/0) • 85–89% (+5 points) (85%: 0/2; 87%: 1/0; 89%: 1/2) • 80–84% (+8 points) (80%: 0/3; 81%: 0/1; 83%: 0/1; 84%: 1/1) • <80% (+10 points)	94.64 ± 4.80% [96.00% (92.75–98.00%)] 65 (61.3%) 28 (26.4%) 6 (5.7%) 7 (6.6%) 0 (0.0%)	95.98 ± 3.06% [97.00% (94.00–98.00%)] 65 (69.9%) 25 (26.9%) 2 (2.2%) 1 (1.1%) 0 (0.0%)	85.08 ± 4.11% [85.00% (80.50–89.50%)] 0 (0.0%) 3 (23.1%) 4 (30.8%) 6 (46.2%) 0 (0.0%)	SpO_2 ≥92%: 0.0% (0/87) 95–100%: 0.0% (0/65) 90–94%: 10.7% (3/28) (90% exact: 50.0% [3/6]) 85–89%: 66.7% (4/6) 80–84%: 85.7% (6/7) SpO_2 ≤90%: 68.4% (13/19)	Exact 16-level χ^2 = 81.217 (df = 15), p < 0.001*[a] (G^2 = 63.987, p < 0.001*) Grouped STUMBL χ^2 = 60.749 (df = 3), p < 0.001* Mann–Whitney U = 21.0, Z = −5.668, p < 0.001* (r_{rb} = 0.965)

*[a]Pearson's chi-square test; [b]Two-sided Fisher's exact test. *Statistically significant at p < 0.05. All 13 non-survivors (100.0%) presented with admission SpO_2 ≤90% (10/13 [76.9%] with SpO_2 <90% and 3/13 [23.1%] with SpO_2 = 90%), whereas 0/87 patients (0.0%) presenting with SpO_2 ≥92% died. Abbreviations: CI, confidence interval; COPD, chronic obstructive pulmonary disease; IQR, interquartile range; OR, odds ratio; r_{rb}, rank-biserial correlation; SD, standard deviation; SpO_2, peripheral capillary oxygen saturation; STUMBL, Study of the Management of Blunt chest wall trauma.*

Further quantitative examination of the component cross-tabulations highlights the relative discriminatory hierarchy of the individual predictors entering the two scoring systems. Ranked by Pearson's chi-square magnitude and non-parametric rank-biserial effect size, admission peripheral oxygen saturation (SpO_2; χ^2 = 81.217, df = 15, p < 0.001; Mann–Whitney U = 21.0, Z = −5.668, r_{rb} = 0.965) was the single strongest individual discriminator of in-hospital mortality, followed by radiographic pulmonary contusion severity (χ^2 = 41.062, df = 4, p < 0.001; likelihood ratio G^2 = 30.752, p < 0.001), number of fractured ribs (χ^2 = 33.575, df = 8, p < 0.001; Mann–Whitney U = 115.0, Z = −4.780, r_{rb} = 0.810), and bilateral rib fracture involvement (χ^2 = 11.452, df = 1, p = 0.001; Continuity Correction χ^2 = 7.972, p = 0.005; Fisher's exact p = 0.008; OR = 9.89). Examining the exact rib-by-rib mortality distribution demonstrates that among the 51 patients (48.1% of the cohort) who presented with 0 to 3 fractured ribs (0 ribs: n = 17; 1 rib: n = 1; 2 ribs: n = 8; 3 ribs: n = 25), survival was 100.0% (51/51 survived, 0/51 died). Among the 30 patients (28.3% of the cohort) who sustained 4 to 5 fractured ribs (4 ribs: n = 19 [18 survived, 1 died]; 5 ribs: n = 11 [10 survived, 1 died]), mortality was 6.7% (2/30). In stark contrast, among the 25 patients (23.6% of the cohort) who sustained 6 to 8 fractured ribs (6 ribs: n = 14 [9 survived, 5 died]; 7 ribs: n = 9 [4 survived, 5 died]; 8 ribs: n = 2 [1 survived, 1 died]), 11 patients died and 14 survived, yielding a mortality rate of 44.0% (11/25) and accounting for 84.6% (11/13) of all deaths in the entire study. Similarly, cross-checking the admission SpO_2 distribution confirms that among the 19 patients (17.9% of the cohort) presenting with SpO_2 ≤ 90% (80% to 90%), 13 patients died (68.4% mortality) and only 6 survived (1 at 84%, 1 at 87%, 1 at 89%, and 3 at 90%), whereas among the 87 patients (82.1% of the cohort) presenting with SpO_2 ≥ 92% (92% to 99%), all 87 survived (100.0% survival, 0.0% mortality). These component-level empirical distributions across fractured rib counts, admission SpO_2 levels, pulmonary contusion grades, bilateral fracture status, and the comparative chi-square effect-size hierarchy are illustrated in Figure 2.

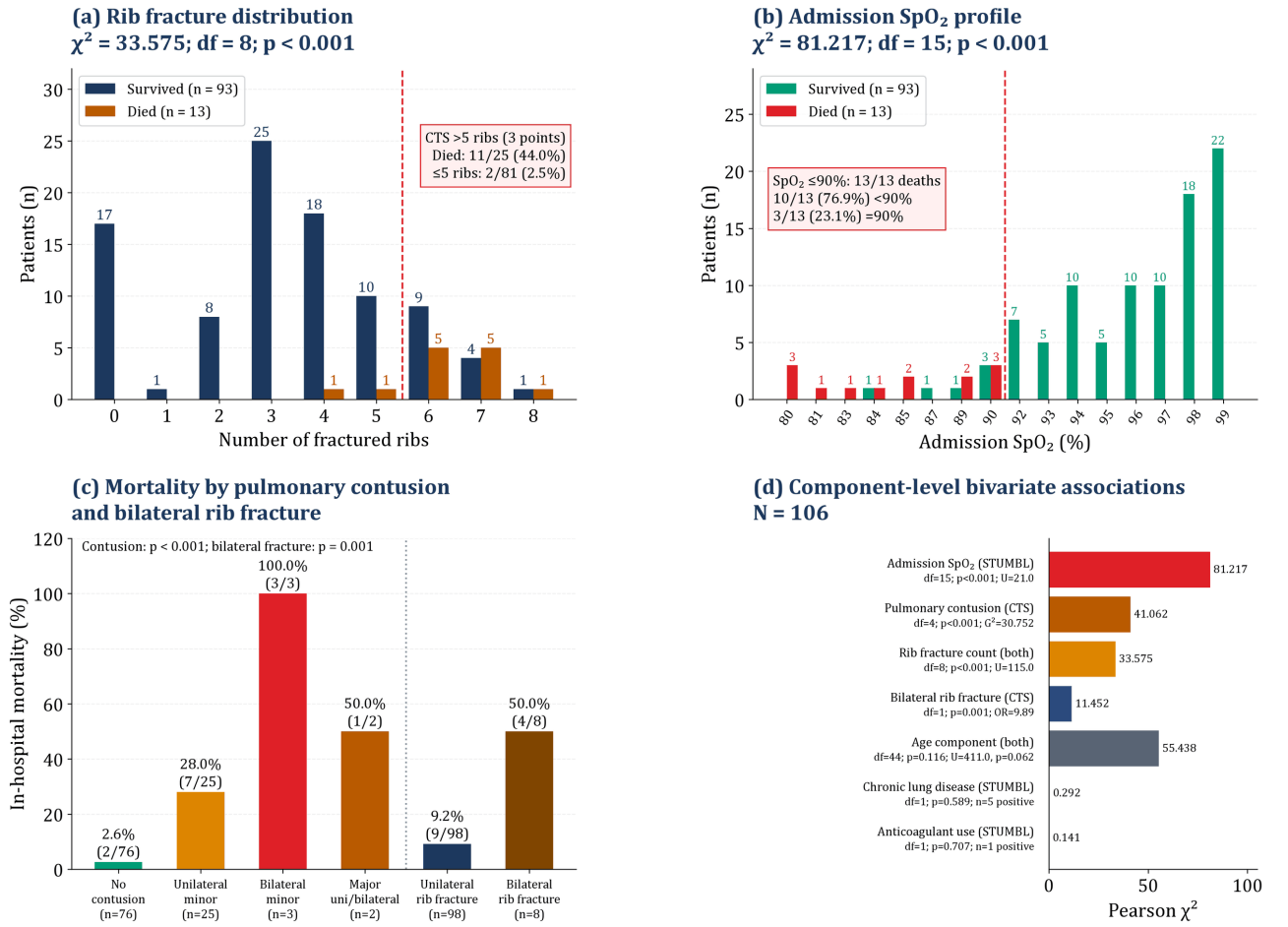


Figure 2. Component-level empirical distributions of the Chest Trauma Score (CTS) and STUMBL (Battle) Score across survival outcomes (N = 106). (a) Fractured rib distribution (0–8 ribs; χ^2 = 33.575, df = 8, p < 0.001), highlighting the 44.0% mortality rate at >5 ribs. (b) Admission peripheral oxygen saturation (SpO_2) profile (80–99%; χ^2 = 81.217, df = 15, p < 0.001), demonstrating that 100.0% (13/13) of non-survivors presented with $SpO_2 \leq 90\%$. (c) In-hospital mortality gradient across parenchymal pulmonary contusion categories (p < 0.001) and bilateral rib fracture status (OR = 9.89, p = 0.001). (d) Comparative Pearson chi-square hierarchy across all individual score components.

Normality Audit, Composite Score Distributions, and Univariable Logistic Calibration

Prior to hypothesis testing of composite scores between survivors and non-survivors, formal normality testing was conducted across survival strata, with complete distributional parameters, normality statistics, and between-group hypothesis tests summarized in Table 4. For the Chest Trauma Score (CTS), both the Kolmogorov–Smirnov test and the Shapiro–Wilk test demonstrated statistically significant departures from normality in the survivor group (n = 93: K–S statistic = 0.225, p < 0.001; Shapiro–Wilk W = 0.916, df = 93, p < 0.001) and in the non-survivor group (n = 13: K–S statistic = 0.278, p = 0.007; Shapiro–Wilk W = 0.862, df = 13, p = 0.041). In contrast, for the STUMBL (Battle) score, both normality tests confirmed that score distributions did not deviate significantly from a Gaussian distribution in either the survivor group (n = 93: K–S statistic = 0.089, p = 0.064; Shapiro–Wilk W = 0.974, df = 93, p = 0.065) or the non-survivor group (n = 13: K–S statistic = 0.161, p = 0.200; Shapiro–Wilk W = 0.955, df = 13, p = 0.671). Consequently, the non-parametric Mann–Whitney U test was applied as the primary test for CTS, whereas Welch's parametric independent-samples t-test (accounting for unequal variances on Levene's test: F = 8.044, p = 0.005) was applied as the primary test for the STUMBL score. Non-survivors exhibited substantially higher composite CTS values than survivors, with a median CTS of 6.00 (range 5.00–8.00; IQR: 1.50 [Q1–Q3: 6.00–7.50]; parametric mean 6.62 ± 0.96 [exact 6.6154 ± 0.96077], SE = 0.27, 95% CI: 6.03–7.20) in the deceased group compared with a median CTS of 4.00 (range 1.00–6.00; IQR: 1.00 [Q1–Q3: 3.00–4.00]; parametric mean 3.66 ± 1.34 [exact 3.6559 ± 1.33916], SE = 0.14, 95% CI: 3.38–3.93) in the survivor group. Exact frequency decomposition confirmed that among the 93 survivors, CTS scores were distributed as 1 point (n = 8), 2 points (n = 10), 3 points (n = 17), 4 points (n = 38), 5 points (n = 11), and 6 points (n = 9), whereas among the 13 non-survivors, CTS scores were concentrated at 5 points (n = 1), 6 points (n = 6), 7 points (n = 3), and 8 points (n = 3). Mann–Whitney U testing confirmed a highly significant difference in CTS between groups (U = 41.500, Wilcoxon W = 4,412.500, Z = −5.580, p < 0.001, r_{rb} = 0.931), with a mean rank of 96.81 (sum of ranks = 1,258.50) in non-survivors versus 47.45 (sum of ranks = 4,412.50) in survivors.

Table 4. Distributional Parameters, Normality Audit, and Between-Group Hypothesis Testing of Total CTS and STUMBL Scores by Survival Status (N = 106).

Scoring System & Outcome Group	Parametric Summary Mean ± SD (SE) [95% CI for Mean]	Non-Parametric Summary Median (Min–Max) [IQR; 5% Trimmed Mean]	Skewness (SE) & Kurtosis (SE) [Variance]	Kolmogorov–Smirnov[a] & Shapiro–Wilk Normality Tests	Primary Between-Group Test Statistic, Effect Size & p-value
Chest Trauma Score (CTS) • Survived (n = 93; 87.7%)	3.66 ± 1.34 (0.14) [3.38–3.93] (Exact: 3.6559 ± 1.33916)	4.00 (1.00–6.00) [IQR: 1.00 (3.00–4.00); 5% Trimmed: 3.67]	Skew: −0.259 (0.250) Kurt: −0.270 (0.495) [Var: 1.793]	K–S = 0.225 (df = 93), p < 0.001 S–W = 0.916 (df = 93), p < 0.001 (Non-normal distribution)	Mann–Whitney U = 41.500 Wilcoxon W = 4,412.500 Z = −5.580, p < 0.001* Mean rank: 96.81 (Died) vs. 47.45 (Survived) r_{rb} = 0.931 (Very large)
Chest Trauma Score (CTS) • Died (n = 13; 12.3%)	6.62 ± 0.96 (0.27) [6.03–7.20] (Exact: 6.6154 ± 0.96077)	6.00 (5.00–8.00) [IQR: 1.50 (6.00–7.50); 5% Trimmed: 6.63]	Skew: +0.280 (0.616) Kurt: −0.891 (1.191) [Var: 0.923]	K–S = 0.278 (df = 13), p = 0.007 S–W = 0.862 (df = 13), p = 0.041 (Non-normal distribution)	Univariable logistic model: β = +2.814 ± 0.928 OR = 16.68 per +1 point (95% CI: 2.71–102.88, p = 0.002*; $R^2_{Nagelkerke}$ = 0.724)
STUMBL (Battle) Score • Survived (n = 93; 87.7%)	14.89 ± 6.86 (0.71) [13.48–16.31] (Exact: 14.8925 ± 6.86431)	15.00 (1.00–30.00) [IQR: 7.50 (11.50–19.00); 5% Trimmed: 14.88]	Skew: −0.119 (0.250) Kurt: −0.412 (0.495) [Var: 47.119]	K–S = 0.089 (df = 93), p = 0.064 S–W = 0.974 (df = 93), p = 0.065 (Normal distribution)	Welch's t = −13.164 (df = 37.140), p < 0.001* (Levene's F = 8.044, p = 0.005; Equal-var t = −7.181, df = 104) Mean Δ = −13.88 ± 1.05 (95% CI: −16.01 to −11.74)
STUMBL (Battle) Score • Died (n = 13; 12.3%)	28.77 ± 2.80 (0.78) [27.08–30.46] (Exact: 28.7692 ± 2.80339)	28.00 (24.00–34.00) [IQR: 3.00 (27.50–30.50); 5% Trimmed: 28.74]	Skew: +0.279 (0.616) Kurt: +0.150 (1.191) [Var: 7.859]	K–S = 0.161 (df = 13), p = 0.200[b] S–W = 0.955 (df = 13), p = 0.671 (Normal distribution)	Mann–Whitney U = 32.000, Z = −5.520, p < 0.001* (r_{rb} = 0.947); Univariable logistic model: β = +0.596 ± 0.180, OR = 1.81/pt (95% CI: 1.27–2.58, p = 0.001*; $R^2_{Nagelkerke}$ = 0.721)

*[a]Kolmogorov–Smirnov test with Lilliefors significance correction; [b]Lower bound of the true significance. Because CTS violated normality in both groups (Shapiro–Wilk p < 0.001 and p = 0.041), the Mann–Whitney U test was applied as the primary test for CTS. Because STUMBL satisfied normality in both groups (Shapiro–Wilk p = 0.065 and p = 0.671) with unequal variances on Levene's test (F = 8.044, p = 0.005), Welch's independent-samples t-test (equal variances not assumed) was applied as the primary parametric test for STUMBL alongside pooled-variance t-test and Mann–Whitney U test. *Statistically significant at p < 0.05.*

Similarly, composite STUMBL scores were nearly two-fold higher among non-survivors than among survivors: the mean STUMBL score in the deceased group was 28.77 ± 2.80 points (exact 28.7692 ± 2.80339; SE = 0.78, 95% CI: 27.08–30.46; median 28.00, range 24.00–34.00, IQR: 3.00 [Q1–Q3: 27.50–30.50]), compared with 14.89 ± 6.86 points (exact 14.8925 ± 6.86431; SE = 0.71, 95% CI: 13.48–16.31; median 15.00, range 1.00–30.00, IQR: 7.50 [Q1–Q3: 11.50–19.00]) in the survivor group. Welch's independent-samples t-test confirmed a highly significant mean difference of −13.88 ± 1.05 points (exact −13.87676 ± 1.05413; 95% CI: −16.01 to −11.74; t = −13.164, df = 37.140, p < 0.001), which remained equally robust under pooled-variance Student's t-testing (t = −7.181, df = 104, SE = 1.93, 95% CI: −17.71 to −10.04, p < 0.001) and non-parametric Mann–Whitney U testing (U = 32.000, Z = −5.520, p < 0.001, r_{rb} = 0.947). In univariable logistic regression, each 1-point increase in CTS was associated with a 16.68-fold increase in the odds of in-hospital mortality (β = +2.814 ± 0.928, Wald χ^2 = 9.196, p = 0.002; OR = 16.68, 95% CI: 2.71–102.88; Cox–Snell R^2 = 0.380, Nagelkerke R^2 = 0.724), whereas each 1-point increase in the STUMBL score conferred an 81% increase in the odds of in-hospital mortality (β = +0.596 ± 0.180, Wald χ^2 = 10.953, p = 0.001; OR = 1.81, 95% CI: 1.27–2.58; Cox–Snell R^2 = 0.379, Nagelkerke R^2 = 0.721). The distributional boxplots, Shapiro–Wilk normality parameters, and univariable logistic regression mortality probability trajectories across score bands are illustrated in Figure 3.

Inspection of the complete frequency distributions and quartile architecture of the two composite scores explains why CTS departed from Gaussian normality whereas the STUMBL score satisfied normality across both outcome groups. Because CTS is compressed into a narrow integer range (observed 1 to 8 points) and 38 of the 93 survivors (40.9%) shared the exact modal score of 4.00 points (with 8 scoring 1, 10 scoring 2, 17 scoring 3, 38 scoring 4, 11 scoring 5, and 9 scoring 6), the survivor CTS distribution exhibited discrete leptokurtic clustering around its median of 4.00 (IQR = 1.00; Q1–Q3: 3.00–4.00; 5% trimmed

mean = 3.67; variance = 1.793; skewness = −0.259 ± 0.250; kurtosis = −0.270 ± 0.495), while the 13 non-survivors clustered tightly at 5 to 8 points (1 scoring 5, 6 scoring 6, 3 scoring 7, and 3 scoring 8; median = 6.00, IQR = 1.50, 5% trimmed mean = 6.63, variance = 0.923, skewness = +0.280 ± 0.616, kurtosis = −0.891 ± 1.191). Conversely, because the STUMBL score spans a wide 34-point dynamic range (observed 1 to 34 points) generated by summing decade age points (1–7 points), 3-point rib fracture multiples (0–24 points), and 2-to-8-point SpO_2 increments, the survivor STUMBL distribution dispersed smoothly across 24 distinct integer values from 1.00 to 30.00 (mean 14.89 ± 6.86, 5% trimmed mean = 14.88, median = 15.00, IQR = 7.50 [Q1–Q3: 11.50–19.00], variance = 47.119, skewness = −0.119 ± 0.250, kurtosis = −0.412 ± 0.495; Shapiro–Wilk W = 0.974, p = 0.065), and the non-survivor STUMBL distribution spanned symmetrically from 24.00 to 34.00 (1 scoring 24, 1 scoring 25, 1 scoring 27, 4 scoring 28, 2 scoring 29, 1 scoring 30, 1 scoring 31, 1 scoring 33, and 1 scoring 34; mean 28.77 ± 2.80, 5% trimmed mean = 28.74, median = 28.00, IQR = 3.00 [Q1–Q3: 27.50–30.50], variance = 7.859, skewness = +0.279 ± 0.616, kurtosis = +0.150 ± 1.191; Shapiro–Wilk W = 0.955, p = 0.671). In the non-survivor group, the minimum observed CTS score was 5.00 (only 1 non-survivor scored 5, while 12/13 scored ≥6) and the minimum observed STUMBL score was 24.00 (only 1 non-survivor scored 24, while 12/13 scored ≥25), creating a sharp distributional separation between survivors and non-survivors, as illustrated in Figure 3.

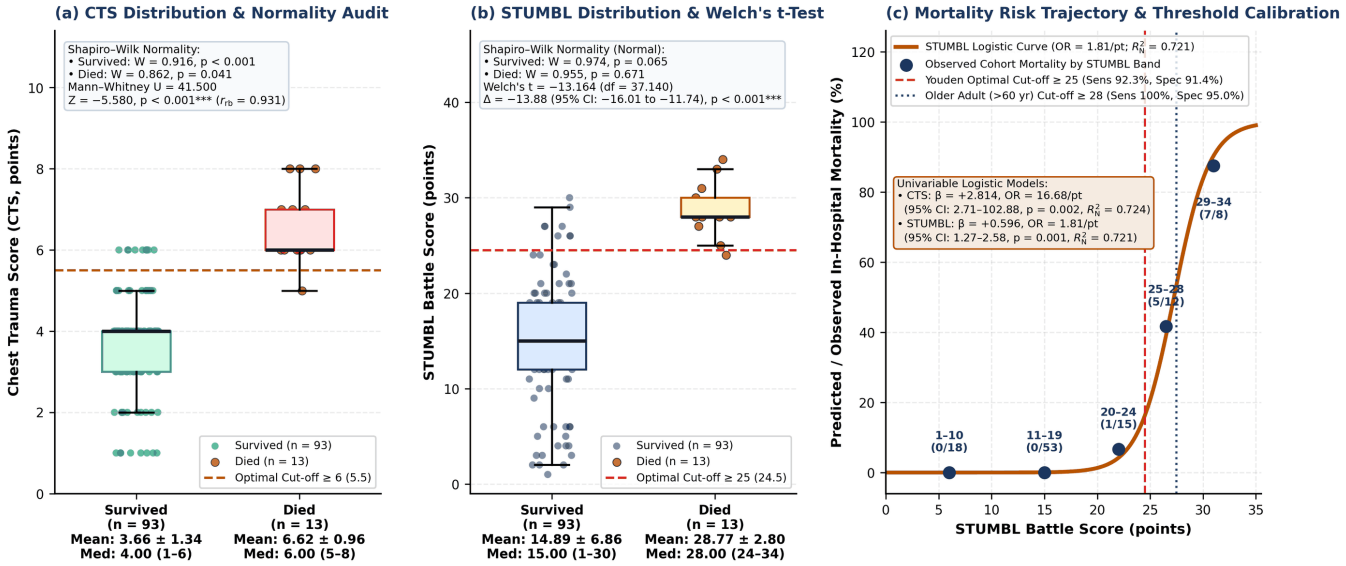


Figure 3. Distributional profiles, Shapiro–Wilk normality audits, and logistic probability calibration trajectories of total CTS and STUMBL scores (N = 106). (a) CTS distribution in survivors (3.66 ± 1.34, median 4.00 [1–6]) versus non-survivors (6.62 ± 0.96, median 6.00 [5–8]; Mann–Whitney U = 41.500, Z = −5.580, p < 0.001). (b) STUMBL score distribution in survivors (14.89 ± 6.86, median 15.00 [1–30]) versus non-survivors (28.77 ± 2.80, median 28.00 [24–34]; Welch's t = −13.164, df = 37.140, p < 0.001). (c) Univariable logistic regression predicted mortality trajectory and observed cohort mortality across STUMBL risk bands.

Overall Receiver Operating Characteristic (ROC) Discrimination and 2×2 Diagnostic Accuracy

The overall Receiver Operating Characteristic (ROC) discrimination, optimal Youden index thresholds, and 2×2 contingency diagnostic accuracy metrics of CTS and the STUMBL score across the total cohort (N = 106) are detailed in Table 5. In the overall cohort, non-parametric ROC analysis demonstrated excellent discriminative power for both scoring systems: the Chest Trauma Score achieved an AUC of 0.966 ± 0.017 (asymptotic 95% CI: 0.932–0.999, p < 0.001), whereas the STUMBL (Battle) score achieved a slightly higher AUC of 0.974 ± 0.014 (asymptotic 95% CI: 0.946–1.000, p < 0.001). Formal pairwise comparison of the two ROC curves using the DeLong/Hanley–McNeil method yielded an absolute AUC difference (ΔAUC) of +0.008 (SE_{diff} = 0.0133, Z = 0.604, p = 0.546), confirming that while STUMBL achieved a numerically higher point estimate and narrower confidence interval, both instruments operated within an overlapping, high-precision discrimination tier (AUC > 0.96). Based on ROC coordinate analysis and maximization of the Youden Index, the optimal threshold for CTS in the overall cohort was identified at ≥5.5, corresponding to an integer clinical cut-off of CTS ≥ 6. At CTS ≥ 6, 21 patients tested positive and 85 patients tested negative, yielding 12 True Positives (TP), 9 False Positives (FP), 1 False Negative (FN), and 84 True Negatives (TN). Accordingly, CTS ≥ 6 achieved a Sensitivity of 92.31% (12/13, Wilson 95% CI: 66.69–98.63%), a Specificity of 90.32% (84/93, 95% CI: 82.62–94.82%), a Positive Predictive Value (PPV / NDP) of 57.14% (12/21, 95% CI: 36.55–75.53%), a Negative Predictive Value (NPV / NDN) of 98.82% (84/85, 95% CI: 93.63–99.79%), a Youden Index (J) of 0.826, an overall diagnostic Accuracy of 90.57% (96/106, 95% CI: 83.50–94.79%), a Positive Likelihood Ratio (LR^+) of 9.54, a Negative Likelihood Ratio (LR^-) of 0.085, and a crude mortality Odds Ratio of 112.00 (95% CI: 13.01–964.15, p < 0.001). For the STUMBL score, the optimal threshold in the overall cohort was identified at ≥24.5, corresponding to an integer clinical cut-off of STUMBL ≥ 25. At STUMBL ≥ 25, 20 patients tested positive and 86 tested negative, yielding TP = 12, FP = 8, FN = 1, and TN = 85. Consequently, STUMBL ≥ 25 achieved a Sensitivity of 92.31% (12/13, 95% CI: 66.69–98.63%), a Specificity of 91.40% (85/93, 95% CI: 83.93–95.58%), a PPV of 60.00% (12/20, 95% CI: 38.66–78.12%), an NPV of 98.84% (85/86, 95% CI: 93.70–99.79%), a Youden Index (J) of 0.837, an overall Accuracy of 91.51% (97/106, 95% CI: 84.65–95.47%), an LR^+ of 10.73, an LR^- of 0.084, and a crude mortality Odds Ratio of 127.50 (95% CI: 14.63–1,111.09, p < 0.001).

Table 5. Overall Receiver Operating Characteristic (ROC) Analysis, 2×2 Contingency Classification, and Diagnostic Accuracy of CTS and STUMBL Score for Predicting In-Hospital Mortality (N = 106).

Diagnostic / Prognostic Parameter	Chest Trauma Score (CTS) [Cut-off ≥ 6 (ROC ≥ 5.5)]	STUMBL (Battle) Score [Cut-off ≥ 25 (ROC ≥ 24.5)]	Comparative Difference & Statistical Interpretation
Area Under the ROC Curve (AUC ± SE) [Asymptotic 95% CI; p-value]	0.966 ± 0.017 [95% CI: 0.932–0.999; p < 0.001*]	0.974 ± 0.014 [95% CI: 0.946–1.000; p < 0.001*]	ΔAUC (STUMBL − CTS) = +0.008 DeLong Z = 0.604, p = 0.546 (Both outstanding)
Optimal ROC coordinate & clinical cut-off (Youden Index J = Sens + Spec − 1)	ROC midpoint ≥ 5.5 → Integer CTS ≥ 6 (Youden Index J = 0.826)	ROC midpoint ≥ 24.5 → Integer STUMBL ≥ 25 (Youden Index J = 0.837)	ΔJ = +0.011 in favor of STUMBL (Higher overall Youden efficiency)
2×2 Contingency Matrix (N = 106) • True Positive (TP: Score ≥ Cut-off & Died) • False Positive (FP: Score ≥ Cut-off & Survived) • False Negative (FN: Score < Cut-off & Died) • True Negative (TN: Score < Cut-off & Survived)	Test Positive = 21 | Test Negative = 85 • TP = 12 (57.1% of CTS ≥ 6) • FP = 9 (42.9% of CTS ≥ 6) • FN = 1 (1.2% of CTS < 6) • TN = 84 (98.8% of CTS < 6)	Test Positive = 20 | Test Negative = 86 • TP = 12 (60.0% of STUMBL ≥ 25) • FP = 8 (40.0% of STUMBL ≥ 25) • FN = 1 (1.2% of STUMBL < 25) • TN = 85 (98.8% of STUMBL < 25)	Both instruments correctly identify 12 of 13 deaths (92.31%); STUMBL yields 1 fewer false positive (8 vs. 9)
Sensitivity — % (TP / [TP + FN]) [Wilson 95% CI]	92.31% (12/13) [95% CI: 66.69–98.63%]	92.31% (12/13) [95% CI: 66.69–98.63%]	Identical mortality sensitivity (92.31%; 12 of 13 fatal cases captured)
Specificity — % (TN / [TN + FP]) [Wilson 95% CI]	90.32% (84/93) [95% CI: 82.62–94.82%]	91.40% (85/93) [95% CI: 83.93–95.58%]	ΔSpecificity = +1.08% in favor of STUMBL (85/93 vs. 84/93)
Positive Predictive Value (PPV / NDP) [Wilson 95% CI]	57.14% (12/21) [95% CI: 36.55–75.53%]	60.00% (12/20) [95% CI: 38.66–78.12%]	ΔPPV = +2.86% in favor of STUMBL (12/20 vs. 12/21)
Negative Predictive Value (NPV / NDN) [Wilson 95% CI]	98.82% (84/85) [95% CI: 93.63–99.79%]	98.84% (85/86) [95% CI: 93.70–99.79%]	Near-perfect rule-out safety (>98.8% NPV for both scoring systems)
Overall Diagnostic Accuracy — % ([TP + TN] / N) [Wilson 95% CI]	90.57% (96/106) [95% CI: 83.50–94.79%]	91.51% (97/106) [95% CI: 84.65–95.47%]	ΔAccuracy = +0.94% in favor of STUMBL (97/106 vs. 96/106)
Likelihood Ratios (LR^+ and LR^-) & Crude / Age-Adjusted Odds Ratios	LR^+ = 9.54 | LR^- = 0.085 Crude OR = 112.00 (13.01–964.15)* Age-adjusted aOR = 167.06 (17.33–1,610.25)*	LR^+ = 10.73 | LR^- = 0.084 Crude OR = 127.50 (14.63–1,111.09)* Age-adjusted aOR = 170.73 (17.85–1,632.62)*	Both thresholds confer >110-fold crude and >165-fold age-adjusted mortality odds

**Statistically significant at p < 0.001. Abbreviations: aOR, age-adjusted multivariable odds ratio; AUC, area under the receiver operating characteristic curve; CI, confidence interval; CTS, Chest Trauma Score; FN, false negative; FP, false positive; J, Youden index; LR^+, positive likelihood ratio; LR^-, negative likelihood ratio; NDN, nilai duga negatif (NPV); NDP, nilai duga positif (PPV); OR, odds ratio; SE, standard error; STUMBL, Study of the Management of Blunt chest wall trauma; TN, true negative; TP, true positive.*

A coordinate-by-coordinate walkthrough of the full ROC curve thresholds further demonstrates the trade-off between sensitivity and specificity across successive score cut-offs in the overall cohort (N = 106). For the Chest Trauma Score (CTS), evaluating mortality sensitivity and specificity across all observed cut-off midpoints showed: at c = 1.5 (CTS ≥ 2), mortality sensitivity was 100.0% (13/13) with specificity 8.6% (8/93); at c = 2.5 (CTS ≥ 3), sensitivity was 100.0% (13/13) with specificity 19.4% (18/93); at c = 3.5 (CTS ≥ 4), sensitivity was 100.0% (13/13) with specificity 37.6% (35/93); at c = 4.5 (CTS ≥ 5), sensitivity remained 100.0% (13/13) with specificity rising to 78.5% (73/93; Youden J = 0.785); at the optimal Youden cut-off c = 5.5 (CTS ≥ 6), mortality sensitivity was 92.31% (12/13) and specificity reached 90.32% (84/93; Youden J = 0.826); at c = 6.5 (CTS ≥ 7), specificity reached 100.0% (93/93) but mortality sensitivity dropped to 46.15% (6/13; Youden J = 0.462); and at c = 7.5 (CTS ≥ 8), specificity was 100.0% (93/93) with mortality sensitivity of 23.08% (3/13). Thus, CTS ≥ 5 provides a 100.0% sensitive screening threshold (78.5% specificity), whereas CTS ≥ 6 maximizes overall diagnostic efficiency (Youden J = 0.826; 92.31% sensitivity, 90.32% specificity, 90.57% accuracy).

Similarly, walking through the 32 ROC coordinate thresholds of the STUMBL score in the overall cohort revealed 100.0% mortality sensitivity (13/13) across all cut-offs from c = 1.5 up to c = 23.5 (STUMBL ≥ 24), where specificity progressively climbed from 1.1% (1/93) at c = 1.5, to 24.7% (23/93) at c = 11.5 (the traditional UK admission threshold of ≥12), 52.7% (49/93) at c = 15.5 (≥16), 81.7% (76/93) at c = 20.5 (≥21), and 89.2% (83/93; Youden J = 0.892) at c = 23.5 (STUMBL ≥ 24). At the selected clinical threshold of c = 24.5 (STUMBL ≥ 25), mortality sensitivity was 92.31% (12/13) and specificity rose above the 90% benchmark to 91.40% (85/93; Youden J = 0.837; PPV = 60.00%, NPV = 98.84%, accuracy = 91.51%, LR^+ = 10.73, LR^- = 0.084). At higher thresholds—c = 25.5 (≥26: sensitivity 84.6%, specificity 91.4%), c = 26.5 (≥27: sensitivity 84.6%, specificity 94.6%), and c = 27.5 (≥28: sensitivity 76.9%, specificity 97.8%)—specificity approached 98% to 100%, confirming a smooth, monotonic diagnostic gradient across the entire upper STUMBL range. The comparative ROC curves across the total cohort, older adults (>60 years), and working-age adults (18–60 years) are illustrated in Figure 4.

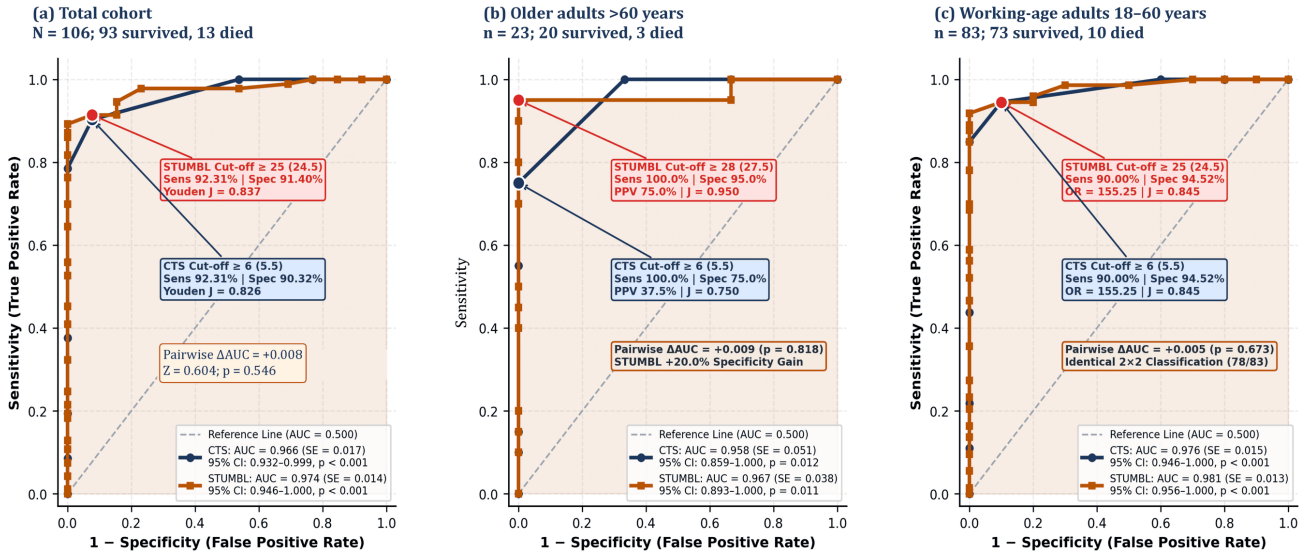


Figure 4. Comparative Receiver Operating Characteristic (ROC) curves of the Chest Trauma Score (CTS) and STUMBL (Battle) Score for predicting in-hospital mortality. (a) Overall cohort (N = 106): STUMBL AUC = 0.974 (95% CI: 0.946–1.000; cut-off ≥ 25) versus CTS AUC = 0.966 (95% CI: 0.932–0.999; cut-off ≥ 6; ΔAUC = +0.008, p = 0.546). (b) Older adult stratum (>60 years; n = 23): STUMBL AUC = 0.967 (cut-off ≥ 28; Specificity 95.00%, PPV 75.00%) versus CTS AUC = 0.958 (cut-off ≥ 6; Specificity 75.00%, PPV 37.50%). (c) Working-age adult stratum (18–60 years; n = 83): STUMBL AUC = 0.981 (cut-off ≥ 25) versus CTS AUC = 0.976 (cut-off ≥ 6), both yielding 90.00% sensitivity and 94.52% specificity.

Age-Stratified ROC Performance, Geriatric Specificity Divergence, and Multivariable Modeling

To evaluate whether patient age modifies threshold calibration and classification accuracy, secondary ROC and 2×2 contingency analyses were performed separately in older adults aged >60 years (n = 23: 20 survived, 3 died) and working-age adults aged 18–60 years (n = 83: 73 survived, 10 died), with complete age-stratified diagnostic performance metrics detailed in Table 6. In the older adult stratum (>60 years; n = 23), both scoring systems maintained high ROC area values, with CTS achieving an AUC of 0.958 ± 0.051 (95% CI: 0.859–1.000, p = 0.012) and the STUMBL score achieving an AUC of 0.967 ± 0.038 (95% CI: 0.893–1.000, p = 0.011; pairwise ΔAUC = +0.009, Z = 0.230, p = 0.818). However, 2×2 contingency evaluation revealed a clinically important divergence in false-positive rates and specificity between the two scores among older adults. For CTS at its optimal cut-off of ≥6 (5.5), all 3 deceased older adults were correctly detected (TP = 3, FN = 0; Sensitivity = 100.00%, NPV = 100.00% [15/15]), but 5 of the 20 surviving older adults were misclassified as false positives (FP = 5, TN = 15), reducing Specificity to 75.00% (15/20, 95% CI: 53.13–88.81%), PPV to 37.50% (3/8, 95% CI: 13.68–69.43%), Youden Index to 0.750, and overall Accuracy to 78.26% (18/23; Pearson χ^2 = 6.469, p = 0.011; Fisher's exact p = 0.032; Pearson's r = 0.530, p = 0.009; Haldane–Anscombe OR = 19.73, 95% CI: 0.87–446.23). Conversely, for the STUMBL score at its age-calibrated optimal cut-off of ≥28 (>27.5) in patients aged >60 years, all 3 deceased patients were likewise captured (TP = 3, FN = 0; Sensitivity = 100.00%, NPV = 100.00% [19/19]), while only 1 of the 20 surviving older adults was classified as a false positive (FP = 1, TN = 19). Consequently, STUMBL ≥ 28 in older adults achieved a Specificity of 95.00% (19/20, 95% CI: 76.39–99.11%—representing a +20.00% specificity gain over CTS), a PPV of 75.00% (3/4, 95% CI: 30.06–95.44%—representing a +37.50% PPV gain over CTS), a Youden Index of 0.950, and an overall diagnostic Accuracy of 95.65% (22/23; Pearson χ^2 = 16.388, p < 0.001; Fisher's exact p = 0.002; Pearson's r = 0.844, p < 0.001; Haldane–Anscombe OR = 91.00, 95% CI: 3.05–2,718.27).

In the working-age adult stratum (18–60 years; n = 83), CTS achieved an AUC of 0.976 ± 0.015 (95% CI: 0.946–1.000, p < 0.001) and the STUMBL score achieved an AUC of 0.981 ± 0.013 (95% CI: 0.956–1.000, p < 0.001; pairwise ΔAUC = +0.005, Z = 0.422, p = 0.673). At their respective optimal cut-offs in the 18–60-year subgroup—CTS ≥ 6 (5.5) and STUMBL ≥ 25 (24.5)—both scoring instruments produced identical 2×2 contingency classifications: TP = 9, FP = 4, FN = 1, and TN = 69 (13 test-positive and 70 test-negative patients), corresponding to a Sensitivity of 90.00% (9/10, 95% CI: 59.58–98.21%), a Specificity of 94.52% (69/73, 95% CI: 86.74–97.85%), a PPV of 69.23% (9/13, 95% CI: 42.37–87.32%), an NPV of 98.57% (69/70, 95% CI: 92.34–99.75%), a Youden Index of 0.845, and an overall Accuracy of 93.98% (78/83; Pearson χ^2 = 47.565, df = 1, p < 0.001; Fisher's exact p < 0.001; Pearson's r = 0.757, p < 0.001; OR = 155.25, 95% CI: 15.58–1,546.58). Furthermore, in multivariable binary logistic regression models adjusting for WHO age category (>60 versus 18–60 years), exceeding the optimal threshold remained a powerful, independent predictor of in-hospital mortality for both CTS ≥ 6 (β = +5.118 ± 1.156, Wald χ^2 = 19.603, p < 0.001; age-adjusted aOR = 167.06, 95% CI: 17.33–1,610.25; model LR χ^2 = 41.671, Nagelkerke R^2 = 0.619) and STUMBL ≥ 25 (β = +5.140 ± 1.152, Wald χ^2 = 19.909, p < 0.001; age-adjusted aOR = 170.73, 95% CI: 17.85–1,632.62; model LR χ^2 = 42.661, Nagelkerke R^2 = 0.631), whereas chronological age >60 years was non-significant after score adjustment (p = 0.141 and p = 0.220, respectively).

Detailed inspection of the age-stratified ROC coordinate thresholds further clarifies why the optimal STUMBL cut-off shifted upward by 3 points in older adults (>60 years; n = 23: cut-off ≥28 [>27.5]) compared with working-age adults (18–60 years; n = 83: cut-off ≥25 [24.5]), whereas the CTS cut-off remained fixed at ≥6 (5.5) in both age groups, as detailed in Table 6. In the >60-year subgroup (n = 23: 20 survived, 3 died), the 8 ROC coordinates for CTS showed 100.0% mortality sensitivity (3/3) up to c = 5.5 (CTS ≥ 6), where mortality specificity was 75.00% (15/20; 5 false positives), whereas stepping up to c = 6.5 (CTS ≥ 7) caused mortality sensitivity to drop to 66.7% (2/3) with 100.0% specificity (20/20). By

contrast, the 14 ROC coordinates for STUMBL in the >60-year subgroup showed that at c = 23.5 (STUMBL ≥ 24), mortality specificity was 80.00% (16/20); at c = 26.5 (STUMBL ≥ 27), mortality specificity was 90.00% (18/20); and at c = 27.5 (STUMBL ≥ 28), mortality specificity reached 95.00% (19/20; only 1 false positive) while retaining 100.00% mortality sensitivity (3/3; Youden J = 0.950, PPV = 75.00%, NPV = 100.00%, accuracy = 22/23 = 95.65%). Moreover, in the >60-year subgroup (where FN = 0), the survival cohort risk ratio was 1.600 (95% CI: 0.935–2.737) for CTS ≥ 6 (reflecting 100.0% survival [15/15] below cut-off versus 62.5% survival [5/8] at or above cut-off: 1.000 / 0.625 = 1.600) and 4.000 (95% CI: 0.733–21.838) for STUMBL ≥ 28 (reflecting 100.0% survival [19/19] below cut-off versus 25.0% survival [1/4] at or above cut-off: 1.000 / 0.250 = 4.000), while Haldane–Anscombe continuity-corrected mortality Odds Ratios were 19.73 (95% CI: 0.87–446.23) for CTS ≥ 6 and 91.00 (95% CI: 3.05–2,718.27) for STUMBL ≥ 28. In the working-age subgroup (18–60 years; n = 83: 73 survived, 10 died), both CTS at c = 5.5 (≥6) and STUMBL at c = 24.5 (≥25) achieved identical ROC coordinate pairs (Specificity = 94.52% [69/73], 1 − Sensitivity = 10.00% [1/10], yielding Sensitivity = 90.00%, Specificity = 94.52%, PPV = 69.23%, NPV = 98.57%, Youden J = 0.845, Accuracy = 78/83 = 93.98%, Pearson χ^2 = 47.565, Continuity Correction χ^2 = 41.381, Likelihood Ratio G^2 = 34.538, Fisher's exact p < 0.001, Pearson's r = 0.757, survival cohort RR = 3.204 [95% CI: 1.417–7.244], mortality cohort RR = 0.021 [95% CI: 0.003–0.149], and mortality Odds Ratio = 155.25 [95% CI: 15.58–1,546.58]).

Table 6. Age-Stratified ROC Discrimination, 2×2 Contingency Matrices, and Diagnostic Performance of CTS and STUMBL Score in Older Adults (>60 Years, n = 23) and Working-Age Adults (18–60 Years, n = 83).

Age Stratum & Scoring Instrument (Optimal Cut-off)	AUC ± SE (95% CI; p-value)	2×2 Matrix (TP / FP / FN / TN)	Sensitivity & Specificity (%) [95% CI]	PPV (NDP) & NPV (NDN) (%) [95% CI]	Youden J & Overall Accuracy ([TP+TN]/n)	Chi-Square, Fisher's p, Correlation r & Odds Ratio
Older Adults (>60 Years; n = 23) [20 Survived (87.0%), 3 Died (13.0%)] • Chest Trauma Score (CTS ≥ 6; ROC ≥ 5.5)	0.958 ± 0.051 (0.859–1.000; p = 0.012*)	TP = 3 | FP = 5 FN = 0 | TN = 15 (Pos = 8, Neg = 15)	Sens: 100.00% (3/3) [43.85–100.00%] Spec: 75.00% (15/20) [53.13–88.81%]	PPV: 37.50% (3/8) [13.68–69.43%] NPV: 100.00% (15/15) [79.61–100.00%]	J = 0.750 Acc: 78.26% (18/23)	χ^2 = 6.469, p = 0.011* Fisher's exact p = 0.032* Pearson r = 0.530 (p = 0.009) Survival RR = 1.600 (0.935–2.737) Haldane OR = 19.73 (0.87–446.23)
Older Adults (>60 Years; n = 23) [20 Survived (87.0%), 3 Died (13.0%)] • STUMBL Score (STUMBL ≥ 28; ROC > 27.5)	0.967 ± 0.038 (0.893–1.000; p = 0.011*)	TP = 3 | FP = 1 FN = 0 | TN = 19 (Pos = 4, Neg = 19)	Sens: 100.00% (3/3) [43.85–100.00%] Spec: 95.00% (19/20) [76.39–99.11%]	PPV: 75.00% (3/4) [30.06–95.44%] NPV: 100.00% (19/19) [83.18–100.00%]	J = 0.950 Acc: 95.65% (22/23)	χ^2 = 16.388, p < 0.001* Fisher's exact p = 0.002* Pearson r = 0.844 (p < 0.001) Survival RR = 4.000 (0.733–21.838) Haldane OR = 91.00 (3.05–2,718.27)*
Working-Age Adults (18–60 Years; n = 83) [73 Survived (88.0%), 10 Died (12.0%)] • Chest Trauma Score (CTS ≥ 6; ROC ≥ 5.5)	0.976 ± 0.015 (0.946–1.000; p < 0.001*)	TP = 9 | FP = 4 FN = 1 | TN = 69 (Pos = 13, Neg = 70)	Sens: 90.00% (9/10) [59.58–98.21%] Spec: 94.52% (69/73) [86.74–97.85%]	PPV: 69.23% (9/13) [42.37–87.32%] NPV: 98.57% (69/70) [92.34–99.75%]	J = 0.845 Acc: 93.98% (78/83)	χ^2 = 47.565, p < 0.001* Fisher's exact p < 0.001* Pearson r = 0.757 (p < 0.001) OR = 155.25 (15.58–1,546.58)*
Working-Age Adults (18–60 Years; n = 83) [73 Survived (88.0%), 10 Died (12.0%)] • STUMBL Score (STUMBL ≥ 25; ROC ≥ 24.5)	0.981 ± 0.013 (0.956–1.000; p < 0.001*)	TP = 9 | FP = 4 FN = 1 | TN = 69 (Pos = 13, Neg = 70)	Sens: 90.00% (9/10) [59.58–98.21%] Spec: 94.52% (69/73) [86.74–97.85%]	PPV: 69.23% (9/13) [42.37–87.32%] NPV: 98.57% (69/70) [92.34–99.75%]	J = 0.845 Acc: 93.98% (78/83)	χ^2 = 47.565, p < 0.001* Fisher's exact p < 0.001* Pearson r = 0.757 (p < 0.001) OR = 155.25 (15.58–1,546.58)*

**Statistically significant at p < 0.05. Pairwise AUC comparisons within strata showed ΔAUC = +0.009 (Z = 0.230, p = 0.818) in >60 years and ΔAUC = +0.005 (Z = 0.422, p = 0.673) in 18–60 years; however, at the optimal Youden threshold in older adults (>60 years), STUMBL ≥ 28 reduced false positives from 5 to 1, improving Specificity by +20.00% (95.00% vs. 75.00%) and PPV by +37.50% (75.00% vs. 37.50%) while preserving 100.00% Sensitivity. Because FN = 0 in the >60-year subgroup, Haldane–Anscombe continuity-corrected Odds Ratios are reported alongside cohort relative risks.*

4. Discussion

The present STARD-compliant diagnostic accuracy study provides the first direct head-to-head comparison between the anatomical-radiologic Chest Trauma Score (CTS) and the clinical-physiological STUMBL (Battle) score for predicting in-hospital mortality among adult blunt thoracic trauma patients at an Indonesian tertiary referral hospital. Four principal findings emerge from the present investigation (N = 106). First, the overall in-hospital mortality rate of 12.3% (13/106; 95% CI: 7.31–19.85%) and the pronounced male predominance (75.5%; 80/106) mirror contemporary epidemiological benchmarks from low- and middle-income countries, including the 11.65% pooled blunt chest trauma mortality reported across sub-

Saharan Africa by Adal et al.,[1] the 17.7% mortality reported in southern Iran by Karajizadeh et al.,[2] and the 10.9% ICU blunt chest trauma mortality documented by Hefny et al.[3] Second, component-level bivariate analysis demonstrated that mortality in blunt thoracic trauma is driven by the convergence of severe structural chest wall disruption—specifically fractured rib count (mean 6.31 ± 1.03 versus 3.27 ± 2.05 ribs; $p < 0.001$) and bilateral rib fractures (mortality 50.0% versus 9.2%; OR = 9.89, $p = 0.001$)—with parenchymal pulmonary contusion ($\chi^2 = 41.062$, $p < 0.001$) and acute physiological hypoxemia on arrival ($SpO_2 \leq 90\%$ in 100.0% [13/13] of non-survivors; $\chi^2 = 81.217$, $p < 0.001$). Third, both composite scoring systems achieved outstanding discriminative accuracy across the total cohort (STUMBL AUC = 0.974 [95% CI: 0.946–1.000] at cut-off ≥25 versus CTS AUC = 0.966 [95% CI: 0.932–0.999] at cut-off ≥6; pairwise ΔAUC = +0.008, $p = 0.546$), with identical sensitivity (92.31% [12/13]) and near-perfect negative predictive values (>98.8%). Fourth, age-stratified evaluation uncovered a crucial clinical advantage of the STUMBL score in older adults (>60 years; n = 23): whereas CTS ≥ 6 suffered a substantial drop in specificity to 75.00% (15/20) and PPV to 37.50% (3/8) due to baseline age-bracket score inflation, age-calibrated STUMBL ≥ 28 preserved 100.00% sensitivity while achieving 95.00% specificity (19/20), 75.00% PPV (3/4), and 95.65% overall accuracy (22/23).

A comprehensive global benchmark comparison of our diagnostic accuracy metrics against international and Indonesian validation cohorts is detailed in Table 7. For the Chest Trauma Score, our overall AUC of 0.966 at an optimal cut-off of ≥6 (sensitivity 92.31%, specificity 90.32%) aligns closely with the foundational cohorts of Pressley et al.[11] and Chen et al.[12] as well as the prospective Egyptian validation by Elsaied Hussein et al.,[13] who reported an AUC of 0.959 at a mortality cut-off of ≥6.5 (sensitivity 100.0%, specificity 62.2%), and the South African study by Kalideen and Sewchuran (OR = 4.34 for CTS ≥ 5, $p < 0.001$).[19] Why did CTS achieve a higher AUC in our cohort (0.966) and in Elsaied Hussein et al. (0.959) than in the Korean polytrauma cohort of Kim et al. (AUC = 0.711)[14] or the American trauma registry study of Buchholz et al. (AUC = 0.56)?[25] As demonstrated by Kim et al.,[14] when patients with severe concomitant traumatic brain injury (TBI) or dominant non-thoracic hemorrhage are included, mortality is heavily confounded by neurological or hemorrhagic death, diluting the discriminative power of thorax-specific anatomical scores. Because our protocol specifically excluded polytrauma dominated by non-thoracic injuries, the reference outcome reflected true thoracic cardiopulmonary decompensation, enabling CTS to achieve high structural discrimination.

Similarly, for the STUMBL (Battle) score, our overall mortality AUC of 0.974 at an optimal cut-off of ≥25 (sensitivity 92.31%, specificity 91.40%) compares favorably with the original derivation/validation c-index of 0.96 reported by Battle et al.,[15] the mortality-specific AUC of 0.92 reported by Mukerji et al. in New Zealand,[16] the ED cohort evaluation by Callisto et al. (AUC = 0.84),[17] the composite complication AUROC of 0.90 (cut-off ≥16) reported by Giamello et al. in Italy,[18] the AUROC of 0.934 (cut-off ≥21) reported by Yılmaz et al. in Türkiye,[21] the major trauma center validation by Ahmed-Issap et al. in the United Kingdom,[22] and the mortality AUC of 0.859 (cut-off >10.5) recently reported by Isharanto et al. in Malang, Indonesia.[23] Notably, the optimal STUMBL cut-off for predicting in-hospital mortality in our study (≥25 overall and in ages 18–60 years; ≥28 in ages >60 years) is substantially higher than the traditional ≥11 threshold used to guide ward admission versus emergency department discharge in UK guidelines[15,17,22] or the ≥16 to ≥21 thresholds validated for composite in-hospital morbidity.[18,21] This upward shift in threshold is biologically and mathematically expected: whereas minor pulmonary complications or a ≥7-day hospital stay can be triggered by 2–3 rib fractures in a middle-aged adult (STUMBL 11–16), fatal cardiopulmonary failure requires a cumulative burden of multiple rib fractures (mean 6.31 ribs = ~19 points) combined with acute hypoxemia ($SpO_2 \leq 90\%$ = +2 to +8 points) and age decade points (+5 to +7 points), placing fatal trajectories squarely above 25 to 28 points. Indeed, our empirical threshold of ≥25 to ≥28 matches the original ICU-escalation tier (STUMBL ≥ 27) defined by Battle et al.[15] and recently corroborated by Ahmed-Issap et al.,[22] where 75.0% of all ICU admissions and a 19.5% 90-day mortality rate clustered in the STUMBL ≥ 27 tier.

Viewed within the broader trajectory of Indonesian and regional surgical scholarship,[8–10,23] our investigation advances thoracic trauma prognostication from descriptive single-score correlation to STARD-compliant comparative threshold calibration. Whereas Kurniawan et al. recently examined prehospital determinants of emergency department trauma mortality at Dr. Mohammad Hoesin Central General Hospital,[8] Umar et al. established a bivariate association between CTS ≥ 5 and mortality in 70 patients without ROC cut-off optimization or physiological comparator modeling,[9] and Seamaladi et al. evaluated the Thoracic Trauma Severity Score (TTSS) in 38 patients and observed moderate discrimination (AUC = 0.727, $p = 0.082$),[10] the present study demonstrates that both CTS (AUC = 0.966) and the STUMBL score (AUC = 0.974) achieve exceptional mortality discrimination when evaluated in an adequately powered cohort (N = 106) free of dominant extra-thoracic polytrauma confounding.

A deeper examination of international validation studies summarized in Table 7 also explains why demographic composition and health-system triage pathways shape the behavior of individual STUMBL and CTS components. In our Palembang cohort (mean age 47.38 ± 15.41 years; 78.3% aged 18–60 years; 75.5% male), the demographic profile closely parallels other LMIC and Asian blunt thoracic trauma series—such as Isharanto et al. in Malang, Indonesia (median age 43 years; 75.7% male),[23] Karajizadeh et al. in Shiraz, Iran (mean age 39.35 ± 17.88 years; 83.8% male),[2] Yüksel et al. in Ankara, Türkiye (median age 33 years; 72.9% male),[5] and Nsengiyumva et al. in Kigali, Rwanda (mean age 35 years; 74.0% male)[6]—where high-velocity motorcycle and motor vehicle collisions among working-age men predominate. By contrast, European, Australasian, and North American STUMBL validation cohorts—such as Yılmaz et al. in Istanbul (median age 62–71 years),[21] Ahmed-Issap et al. in Stoke-on-Trent (median age 69 years; 31.5% chronic lung disease; 14.8% pre-injury anticoagulation),[22] Mukerji et al. in New Zealand,[16] and Giamello et al. in Italy (median age 64 years)[18]—contain a much higher proportion of geriatric patients sustaining low-velocity ground-level falls with substantial baseline cardiopulmonary and antithrombotic comorbidity. Consequently, whereas chronic lung disease (27% vs. 8%, $p < 0.001$) and pre-injury anticoagulant use (27% vs. 9%, $p < 0.001$) were highly prevalent and statistically significant predictors in the older Turkish cohort of Yılmaz et al.,[21] their low baseline prevalence in our younger Indonesian cohort (chronic lung disease: 5/106 [4.7%]; anticoagulant use: 1/106 [0.9%]) naturally rendered their individual bivariate chi-square tests non-significant ($p = 0.589$ and $p = 0.707$), exactly as observed by Isharanto et al. in Malang (chronic lung disease 3.8%; anticoagulant use 2.4%).[23] Crucially, despite the minimal contribution of these two comorbidity items in younger Indonesian patients, the STUMBL score still

achieved an overall AUC of 0.974 (and 0.981 in the 18–60-year subgroup) because its 3-point-per-rib linear multiplier and 0-to-10-point SpO_2 gradation powerfully separated fatal from non-fatal injury trajectories.

Table 7. Global Empirical Benchmark Comparison of Chest Trauma Score (CTS) and STUMBL (Battle) Score Validation Studies Across Diverse International and Indonesian Cohorts.

Author, Year & Country [Reference]	Study Design & Sample Population (N)	Scoring System(s) Evaluated & Cut-off	Primary Outcome Endpoint	AUC / C-Statistic (95% CI)	Sensitivity, Specificity, PPV, NPV & Key Findings
Pressley et al., 2012 USA[11]	Retrospective registry derivation (N = 649 chest wall injury)	CTS (Range 2–12; high-risk cut-off ≥ 5)	Pulmonary complications, ICU admission, intubation	Significant stepwise logistic gradient	CTS ≥ 5 independently predicted ICU admission, mechanical ventilation, and prolonged stay
Chen et al., 2014 USA[12]	Retrospective trauma registry validation (N = 1,361)	CTS ≥ 5 (n = 637) vs. CTS < 5 (n = 724)	Mortality, pneumonia, acute respiratory failure	ROC-validated threshold CTS ≥ 5	Mortality 9.0% vs. 2.2% (OR = 3.99, 95% CI: 1.92–8.31, p = 0.001); pneumonia 10.1% vs. 3.5%
Elsaied Hussein et al., 2024, Egypt[13]	Prospective observational blunt chest trauma (N = 45)	CTS ≥ 6.5 (mortality); CTS ≥ 5.5 (pneumonia)	In-hospital mortality (17.8%) & pneumonia	AUC = 0.959 (Mortality prediction)	CTS ≥ 6.5 for mortality: Sens 100.0%, Spec 62.2%, Acc 68.9%; CTS ≥ 5.5 for pneumonia: Sens 81.8%
Kim et al., 2024 South Korea[14]	Retrospective polytrauma validation (N = 1,038)	TTSS, CTS, RibScore, Rib Fracture Score	Composite complications & mortality in polytrauma	CTS AUC = 0.711; TTSS AUC = 0.731	Concomitant severe TBI (58.3%) diluted thorax-specific scores, justifying exclusion of dominant extra-thoracic TBI
Kalideen & Sewchuran, 2026, South Africa[19]	Retrospective tertiary ED blunt chest trauma (N = 301)	Conventional CTS ≥ 5 & SA-CTS ≥ 4	Critical disposition (ICU / mortality 32.8%)	CTS ≥ 5 OR = 4.34 (p < 0.001)	Confirmed CTS ≥ 5 predicts critical outcome, while noting age points inflate scores in stable older adults
Battle et al., 2014 United Kingdom[15]	Derivation (n = 274) & external multicenter validation (n = 237)	STUMBL Score (Cut-offs ≥ 11 and ≥ 27)	Composite pulmonary morbidity, ICU, death	c-index = 0.96 (95% CI: 0.93–0.98)	Sens 80.0%, Spec 96.0%, PPV 93.0%, NPV 86.0%; STUMBL ≥ 27 defined ICU escalation threshold
Mukerji et al., 2021 New Zealand[16]	Multicenter retrospective validation across 5 EDs (N = 445)	STUMBL Score (Cut-offs > 12, > 15, > 18)	Mortality, ICU stay, prolonged LOS, morbidity	Mortality AUC = 0.92 (95% CI: 0.89–0.94)	STUMBL > 18 strongly predicted ICU admission and mortality; composite complication AUC = 0.73
Callisto et al., 2022 UK & Italy[17]	Retrospective ED cohort chart review (N = 369)	STUMBL ≥ 11 vs. ED clinician judgment	30-day complications & mortality	AUC = 0.84 (95% CI: 0.78–0.90)	STUMBL ≥ 11 Sens 79.0%, Spec 77.9%; highlighted need for higher cut-offs to avoid ward over-triage
Giamello et al., 2023 Italy[18]	Retrospective hub-hospital ED validation (N = 745; med age 64)	STUMBL ≥ 16 (Youden optimal)	Composite complications (27.2%), ICU & death	AUC = 0.90 (95% CI: 0.88–0.93)	Sens 80.0%, Spec 87.0%, PPV 70.0%, NPV 92.0%; good Hosmer–Lemeshow calibration (p = 0.34)
Yılmaz et al., 2025 Türkiye[21]	Retrospective tertiary ED cohort (N = 536)	STUMBL ≥ 21 (ROC optimal 20.5)	In-hospital complications (28.0%) & ICU stay	AUROC = 0.934 (95% CI: 0.909–0.959)	Sens 84.0%, Spec 89.0%, AUPRC = 0.889; higher threshold improved specificity in tertiary care

Author, Year & Country [Reference]	Study Design & Sample Population (N)	Scoring System(s) Evaluated & Cut-off	Primary Outcome Endpoint	AUC / C-Statistic (95% CI)	Sensitivity, Specificity, PPV, NPV & Key Findings
Ahmed-Issap et al., 2026, UK[22]	Retrospective major trauma center cohort (N = 438)	STUMBL > 11, ≥ 18.5, and ≥ 27 tiers	Complications (37.2%) & 90-day mortality (9.3%)	Mortality cut-off 17.5; 75% of ICU in ≥ 27 tier	Mortality was 19.5% in STUMBL ≥ 27 tier vs. 3.1% in 11–15 tier, supporting higher mortality cut-offs
Isharanto et al., 2026 Malang, Indonesia[23]	Prospective tertiary hospital cohort (N = 371; med age 43)	STUMBL > 10.5 vs. Injury Severity Score (ISS)	In-hospital mortality (5.4%), ICU stay (18.3%), LOS ≥ 7d	Mortality AUC = 0.859; ICU AUC = 0.894	STUMBL outperformed global ISS (AUC 0.859 vs. 0.753); low COPD (3.8%) & anticoagulant (2.4%) rates
Present Study (Siahaan et al.), Palembang, Indonesia	Retrospective STARD-compliant head-to-head cohort (N = 106; 18–60 yr n = 83, >60 yr n = 23)	CTS ≥ 6 vs. STUMBL ≥ 25 (Overall & 18–60 yr); STUMBL ≥ 28 (>60 yr)	In-hospital mortality (12.3% [13/106]) across overall & WHO age strata	Overall: STUMBL 0.974 vs. CTS 0.966; >60 yr: 0.967 vs. 0.958; 18–60 yr: 0.981 vs. 0.976	Overall Sens 92.31%, Spec 91.40% vs. 90.32%; in >60 yr, STUMBL ≥ 28 achieved 95.00% Spec & 75.00% PPV vs. 75.00% Spec & 37.50% PPV for CTS

Abbreviations: Acc, accuracy; AUC / AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve; CI, confidence interval; COPD, chronic obstructive pulmonary disease; CTS, Chest Trauma Score; ED, emergency department; ICU, intensive care unit; ISS, Injury Severity Score; LOS, length of stay; NPV, negative predictive value; OR, odds ratio; PPV, positive predictive value; SA-CTS, South African Chest Trauma Score; Sens, sensitivity; Spec, specificity; STARD, Standards for Reporting of Diagnostic Accuracy Studies; STUMBL, Study of the Management of Blunt chest wall trauma; TBI, traumatic brain injury; TTSS, Thoracic Trauma Severity Score.

When high-energy blunt kinetic force impacts the thoracic cage, immediate structural failure occurs along the posterolateral angles of the ribs, where cortical bending moments are maximal.[30,31] In our cohort, no patient sustaining 0 to 3 fractured ribs died (0/51; 0.0%), whereas mortality rose to 6.7% (2/30) among those with 4–5 fractured ribs and escalated eleven-fold to 44.0% (11/25) among patients sustaining 6 to 8 fractured ribs. This empirical inflection at the six-rib threshold provides contemporary Indonesian validation of the classic biomechanical 'six-rib breakpoint' and recent multicenter trauma registry analyses demonstrating that ≥6 fractured ribs or bilateral rib cage disruption fundamentally destabilizes the thoracic ventilatory pump.[11,12,31,32] Simultaneously, kinetic pressure waves transmitted across the chest wall disrupt the alveolar-capillary barrier, inducing parenchymal pulmonary contusion in 28.3% (30/106) of our cohort—a subgroup that accounted for 84.6% (11/13) of all in-hospital deaths. Within contused lung segments, intra-alveolar hemorrhage, protein-rich interstitial edema, and endothelial glycocalyx shedding inactivate pulmonary surfactant and reduce regional lung compliance (ΔV/ΔP), creating perfused but unventilated alveolar units (intrapulmonary shunt) and severe ventilation–perfusion (V/Q) mismatch.[29,30] Crucially, while CTS captures this parenchymal injury anatomically via radiographic contusion grading (0–4 points), the STUMBL score captures its real-time functional physiological consequence through admission pulse oximetry (SpO_2, 0 to +10 points).[15,21] Our finding that 100.0% (13/13) of non-survivors arrived with SpO_2 ≤ 90% (mean 85.08 ± 4.11% versus 95.98 ± 3.06% in survivors; U = 21.0, p < 0.001) and that 0.0% (0/87) of patients arriving with SpO_2 ≥ 92% died confirms that peripheral oxygen saturation serves as the ultimate physiological integrator of both chest wall hypoventilation and alveolocapillary diffusion failure.

A particularly novel contribution of our age-stratified analysis is the demonstration and mechanistic explanation of the geriatric specificity divergence between CTS and STUMBL in patients aged >60 years. In working-age adults (18–60 years; n = 83), CTS ≥ 6 and STUMBL ≥ 25 produced identical 2×2 classification matrices (TP = 9, FP = 4, FN = 1, TN = 69; Sensitivity 90.00%, Specificity 94.52%, PPV 69.23%, NPV 98.57%, Accuracy 93.98%, OR = 155.25). However, in older adults (>60 years; n = 23), CTS ≥ 6 generated 5 false positives among 20 survivors (Specificity 75.00%, PPV 37.50%), whereas STUMBL ≥ 28 generated only 1 false positive (Specificity 95.00%, PPV 75.00%, Accuracy 95.65% [22/23]). This 20.0% specificity gap arises from the structural scaling of the two instruments: in CTS (total scale 2–12), chronological age >65 years automatically assigns 3 points (and 45–65 years assigns 2 points), representing up to 50% of the ≥6 mortality threshold before any injury variable is added. Consequently, an older patient (>60 years) presenting with 3 unilateral fractured ribs (+2 points) and a minor unilateral contusion (+1 point) but normal room-air oxygenation (SpO_2 97%) immediately reaches CTS = 6 and is misclassified as a predicted mortality (false positive). In contrast, within the STUMBL score, chronological age contributes 6–7 points (at 1 point per decade), 3 fractured ribs add 9 points, and normal oxygenation (SpO_2 ≥ 95%) adds 0 points, yielding a total STUMBL score of 15–16—correctly identifying the patient for inpatient ward care (>11) while avoiding a false-positive mortality alert (<28). Only when severe rib fracture multiplicity or hypoxemic desaturation supervenes does the older adult cross the STUMBL ≥ 28 threshold, explaining why STUMBL doubles the positive predictive value (75.00% versus 37.50%) in older adults without sacrificing 100.00% sensitivity.[33–36] The dual-stream biomechanical and alveolocapillary pathophysiological cascade and the quantitative comparison of age-stratified diagnostic performance metrics are illustrated in Figure 5.

To appreciate why the combination of fractured rib multiplicity (≥6 ribs), bilateral thoracic cage disruption, parenchymal pulmonary contusion, and admission hypoxemia (SpO_2 ≤ 90%) exerted such a profound synergistic effect on mortality in our cohort, it is instructive to examine the coupled biomechanics of

respiration and cellular alveolar injury in detail. Under physiological conditions, quiet inspiration depends on the outward and upward bucket-handle and pump-handle excursions of ribs 1 through 10 driven by external intercostal and diaphragmatic contraction, generating a negative intrapleural pressure gradient that expands compliant alveoli.[11,12] When 4 to 8 ribs fracture—and particularly when fractures occur bilaterally (OR = 9.89 in our cohort)—three mechanical derangements occur simultaneously: first, unopposed intercostal muscle tension across fracture lines produces severe somatic nociception along the intercostal nerves, triggering reflex thoracoabdominal muscle splinting that restricts tidal volume (V_T) and forces the patient into rapid, shallow breathing with increased dead-space ventilation (V_D/V_T); second, loss of rigid bony continuity reduces the structural elastance of the chest wall, so that negative inspiratory intrathoracic pressure draws unstable chest wall segments inward rather than expanding the underlying lung parenchyma; and third, inability to generate high intrathoracic positive pressure during the compressive phase of coughing impairs mucociliary clearance, leading to inspissated bronchial secretions, lobar atelectasis, and secondary bacterial colonization.[30–32]

Concurrently, within the underlying contused lung parenchyma—present in 84.6% (11/13) of our fatal cases—viscoelastic pressure waves shear the alveolar-capillary unit across a two-stage molecular cascade.[29,30] In the immediate hemorrhagic-edematous stage (0–12 hours), mechanical disruption of capillary endothelial tight junctions and type I alveolar pneumocytes allows whole blood, fibrinogen, and plasma proteins to flood the alveolar spaces, while type II pneumocyte injury impairs surfactant synthesis and accelerates conversion of large surfactant aggregates into inactive small aggregates. As alveolar surface tension rises according to the Laplace relationship ($P = 2T/r$), alveoli collapse at end-expiration, requiring high opening pressures that further amplify mechanical shear stress (atelectrauma). In the subsequent inflammatory amplification stage (12–72 hours), alveolar macrophage recognition of extracellular heme and mitochondrial DAMPs activates Toll-like receptor 4 (TLR4) and nuclear factor kappa B (NF-κB) transcriptional pathways, releasing tumor necrosis factor-alpha (TNF-α), interleukin-1 beta (IL-1β), interleukin-6 (IL-6), and CXC chemokines that recruit activated polymorphonuclear neutrophils into the pulmonary interstitium.[29,30] Neutrophil degranulation and oxidative burst degrade the endothelial glycocalyx and basement membrane, converting localized pulmonary contusion into diffuse acute respiratory distress syndrome (ARDS), refractory hypoxemia, systemic inflammatory spillover, and multi-organ dysfunction. Because pulse oximetry directly measures the arterial hemoglobin oxygen saturation resulting from this alveolocapillary shunt, the STUMBL SpO_2 subscore (0 to +10 points) captures the functional severity of parenchymal contusion and ventilatory splinting at the exact moment of emergency department arrival, even before delayed radiographic consolidation fully blossoms on plain chest radiographs.

Furthermore, our age-stratified 2×2 contingency analysis resolves an important interpretive nuance regarding geriatric trauma scoring. Across large geriatric rib fracture registries—including the Trauma Quality Improvement Program (TQIP) studies by Ekestubbe et al. (N = 321,355)[31] and Mohseni et al. (N = 65,375),[32] the multicenter Australian geriatric cohort by Cooper et al. (N = 491),[33] the prospective multicenter validation of the Rib Fracture Frailty Index by Choi et al. (N = 849),[34] the Revised Cardiac Risk Index study by Bass et al. (N = 96,750),[35] and the National Trauma Data Bank review of geriatric SSRF by Plumb et al. (N = 5,129)[36]—chronological age and physiological frailty consistently amplify the risk of failure-to-rescue, pneumonia, and death after chest wall injury. However, when a clinical prediction rule assigns a disproportionately large fixed weight to chronological age relative to its total scale—as CTS does by awarding 2 points for ages 45–65 years and 3 points for ages >65 years on a 2–12 scale where the mortality cut-off is ≥6—many surviving older adults with only mild-to-moderate anatomical injury cross the ≥6 threshold, inflating the false-positive rate (5/20 = 25.0% false-positive rate; Specificity = 75.00%; PPV = 37.50% in our >60-year subgroup). By contrast, when age-calibrated STUMBL ≥ 28 is applied in patients aged >60 years, only 1 of 20 survivors is misclassified as a false positive (5.0% false-positive rate; Specificity = 95.00%; PPV = 75.00%; Accuracy = 22/23 = 95.65%), while still capturing 100.00% (3/3) of geriatric fatalities. This dual-stream pathophysiological architecture and the age-stratified diagnostic performance comparison are illustrated in Figure 5.

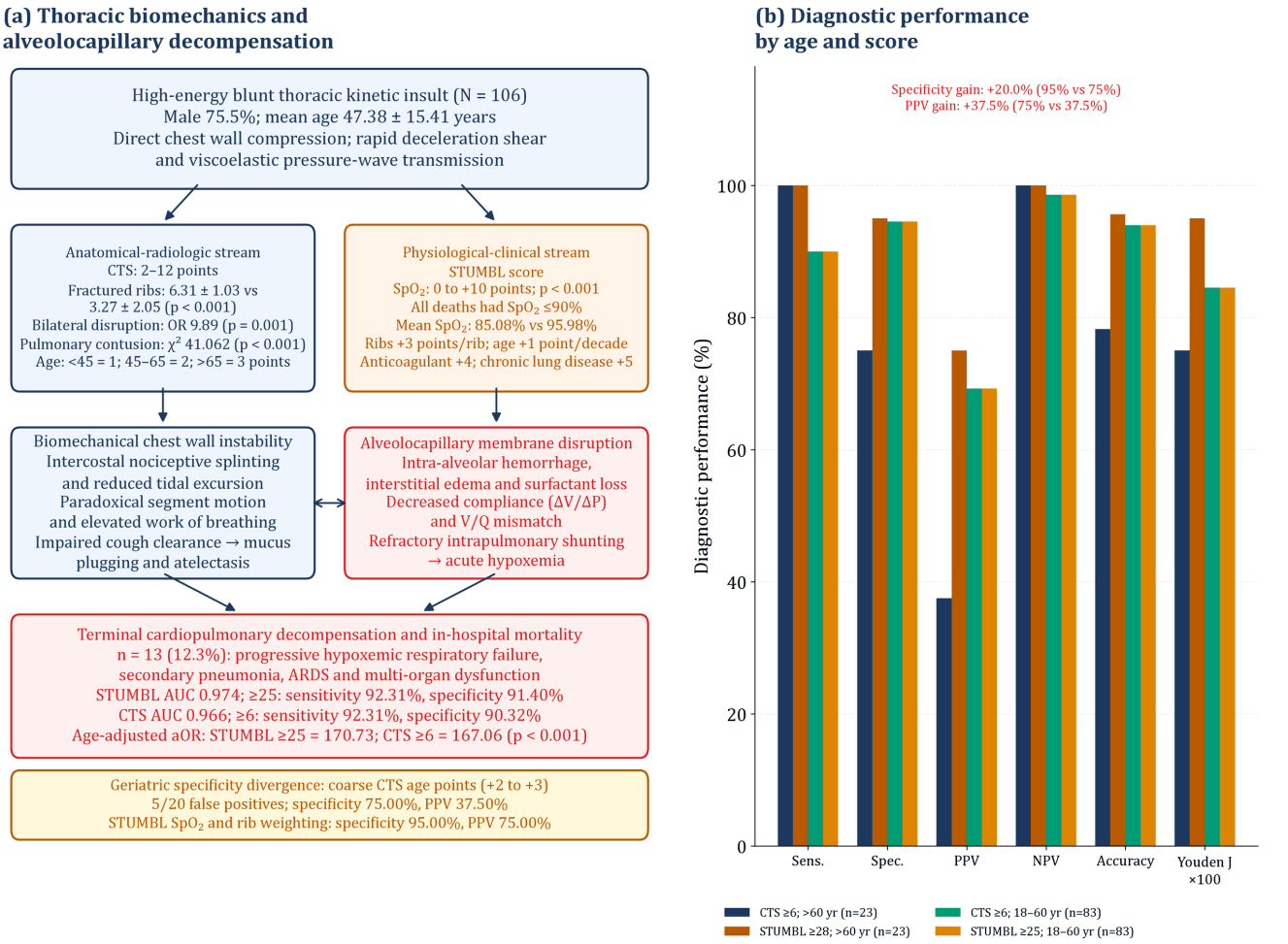
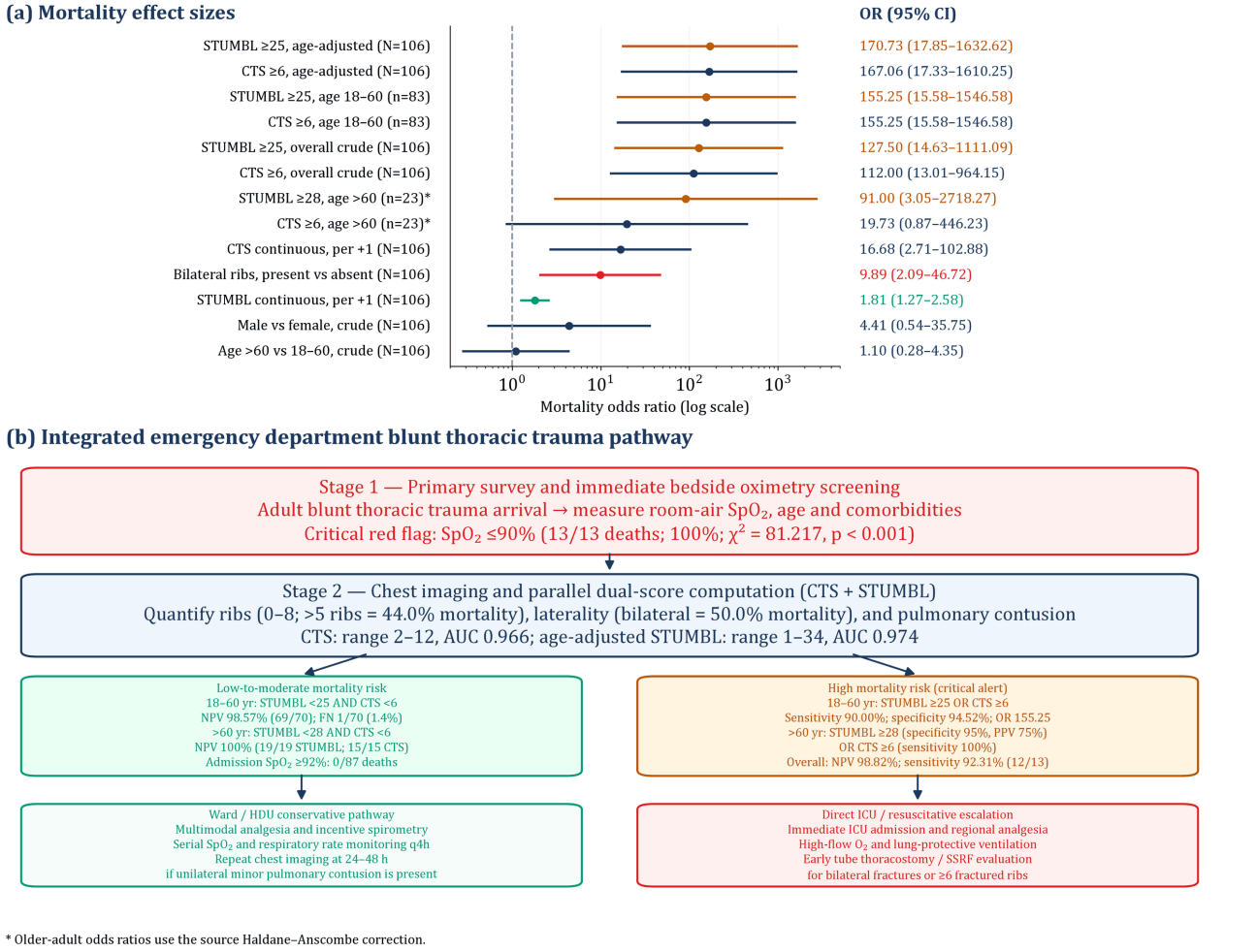


Figure 5. Dual-stream biomechanical and alveolocapillary pathophysiological architecture of blunt thoracic trauma and age-stratified diagnostic performance. (a) Mechanistic integration of structural chest wall instability (fractured ribs, bilateral disruption) and alveolocapillary membrane injury (pulmonary contusion, intrapulmonary shunting, and admission hypoxemia $SpO_2 \leq 90\%$) driving terminal cardiopulmonary decompensation. (b) Quantitative comparison of sensitivity, specificity, PPV, NPV, overall accuracy, and Youden index across age strata, highlighting the +20.00% specificity and +37.50% PPV advantage of STUMBL ≥ 28 over CTS ≥ 6 in older adults (>60 years).

In high-volume emergency departments across Indonesia and similar resource-constrained settings, CTS and the STUMBL score offer complementary operational strengths. Immediate bedside pulse oximetry upon primary survey provides a zero-cost physiological red flag: because 100.0% of deaths in our cohort occurred in patients presenting with $SpO_2 \leq 90\%$ (and 0.0% in those with $SpO_2 \geq 92\%$), any blunt thoracic trauma patient arriving with room-air $SpO_2 \leq 90\%$ warrants immediate resuscitation bay prioritization, supplemental oxygenation, and arterial blood gas assessment. Once initial chest radiography or thoracic CT is completed, parallel calculation of CTS and the STUMBL score enables high-precision risk stratification. In working-age adults (18–60 years), either CTS ≥ 6 or STUMBL ≥ 25 identifies patients with a 155.25-fold increased odds of death (Sensitivity 90.00%, Specificity 94.52%) who require direct ICU admission, multimodal or regional thoracic analgesia, lung-protective ventilatory support, and evaluation for surgical stabilization of rib fractures (SSRF) when flail chest or ≥6 displaced fractures are present.[36–39] In older adults (>60 years), clinicians can use CTS ≥ 6 as a highly sensitive initial radiologic screen (Sensitivity 100.00%, NPV 100.00%) and apply the age-calibrated STUMBL ≥ 28 threshold (Specificity 95.00%, PPV 75.00%) to prevent unnecessary ICU over-triage of normoxemic seniors with low-to-moderate rib fracture burden, thereby optimizing scarce critical care bed allocation.[40–45] The forest plot of univariable and age-adjusted multivariable mortality odds ratios alongside our integrated two-stage emergency department clinical triage pathway are illustrated in Figure 6.

Integrating our empirical thresholds into daily surgical workflow at tertiary referral centers offers a pragmatic, resource-conscious solution for trauma teams in developing countries. During the initial 5 minutes of the primary survey, room-air pulse oximetry provides immediate physiological triage: any patient presenting with $SpO_2 \leq 90\%$—which captured 100.0% of all mortality events in our study (76.9% with $SpO_2 < 90\%$ and 23.1% with $SpO_2 = 90\%$)—must be flagged for immediate resuscitative oxygen escalation, arterial blood gas and metabolic biomarker analysis,[40–42] and priority multidetector chest CT. Upon completion of chest imaging, simultaneous calculation of CTS and the STUMBL score takes less than 60 seconds at the bedside and requires zero laboratory cost. Patients in the working-age stratum (18–60 years) with CTS < 6 and STUMBL < 25 (NPV = 98.57% [69/70]) and older adults (>60 years) with CTS < 6 and STUMBL < 28 (NPV = 100.00% [19/19]) can be safely managed on surgical wards or high-dependency step-down units with scheduled multimodal analgesia, incentive spirometry, and serial 4-hourly pulse oximetry re-evaluation to detect delayed pulmonary contusion progression.[37–39] Conversely, working-age patients crossing CTS ≥ 6 or STUMBL ≥ 25 (OR = 155.25; age-adjusted aOR = 167.06 and 170.73) and older adults crossing STUMBL ≥ 28 (Specificity

95.00%, PPV 75.00%, Haldane OR = 91.00) should be prioritized for direct ICU admission, neuraxial or regional intercostal analgesia, lung-protective ventilation, and early multidisciplinary evaluation for surgical stabilization of rib fractures (SSRF within 72 hours) when multiple displaced fractures or flail chest mechanics are present.[36,37,43–45] By pairing the anatomical precision of CTS with the physiological and age-calibrated specificity of the STUMBL score, emergency surgeons can simultaneously maximize patient safety and optimize critical care bed stewardship, as presented in Figure 6.



* Older-adult odds ratios use the source Haldane–Anscombe correction.

Figure 6. Univariable and age-adjusted multivariable odds ratio forest plot and operational emergency department triage algorithm. (a) Forest plot (log_{10} scale) of crude and age-adjusted mortality odds ratios across composite thresholds and individual predictors. (b) Two-stage clinical decision pathway integrating primary-survey pulse oximetry screening ($SpO_2 \leq 90\%$ red flag) with parallel bedside CTS and age-calibrated STUMBL scoring for ICU versus ward triage.

Several methodological limitations of this study merit consideration. First, the retrospective single-center design relies on the accuracy of medical record and radiologic documentation and may be subject to referral spectrum bias inherent to a tertiary apex hospital. Second, although the total sample size (N = 106) exceeded the a priori minimum requirement (n = 67), the absolute number of mortality events was modest (n = 13 overall, including n = 3 deaths in the >60-year subgroup), resulting in wide confidence intervals around subgroup odds ratios and PPV estimates. Third, the low prevalence of documented chronic lung disease (n = 5; 4.7%) and pre-injury anticoagulant use (n = 1; 0.9%) in our cohort limited statistical power to detect independent bivariate associations for these two STUMBL components, a phenomenon also noted in younger LMIC trauma populations compared with Western geriatric cohorts.[22,23] Fourth, serial post-admission trajectory scores (such as 24-to-48-hour changes in SpO_2, PaO_2/FiO_2 ratio, or delayed contusion progression on repeat CT) and specific non-fatal complications (such as pneumonia incidence, mechanical ventilation duration, and hospital length of stay) were not analyzed as separate endpoints. Future prospective, multicenter Indonesian validation studies incorporating larger geriatric sample sizes, formal calibration plots, and decision-curve net benefit analyses are warranted to confirm the generalizability of these mortality thresholds.

5. Conclusion

In this STARD-compliant diagnostic accuracy study of 106 adult blunt thoracic trauma patients at Dr. Mohammad Hoesin Central General Hospital, Palembang (87.7% survived, 12.3% in-hospital mortality), both the STUMBL (Battle) score and the Chest Trauma Score (CTS) demonstrated excellent discriminative ability for predicting in-hospital mortality. At the component level, mortality was significantly driven by fractured rib count ($p < 0.001$), bilateral rib fractures ($p = 0.001$), parenchymal pulmonary contusion ($p < 0.001$), and admission pulse oximetry $SpO_2 \leq 90\%$ ($p < 0.001$), whereas sex, age alone, chronic lung disease, and anticoagulant use did not reach bivariate significance due to sample distribution and low comorbidity prevalence. Across the overall cohort, STUMBL (AUC = 0.974, 95% CI: 0.946–1.000; optimal cut-off ≥25: sensitivity 92.31%, specificity 91.40%, PPV 60.00%, NPV

98.84%, accuracy 91.51%, J = 0.837) achieved slightly higher discriminative performance and specificity than CTS (AUC = 0.966, 95% CI: 0.932–0.999; optimal cut-off ≥6: sensitivity 92.31%, specificity 90.32%, PPV 57.14%, NPV 98.82%, accuracy 90.57%, J = 0.826), although their 95% confidence intervals overlapped (p = 0.546). In working-age adults (18–60 years; n = 83), CTS ≥ 6 (AUC = 0.976) and STUMBL ≥ 25 (AUC = 0.981) yielded identical classification accuracy (sensitivity 90.00%, specificity 94.52%, PPV 69.23%, NPV 98.57%, accuracy 93.98%, OR = 155.25). In older adults (>60 years; n = 23), STUMBL ≥ 28 (AUC = 0.967) substantially outperformed CTS ≥ 6 (AUC = 0.958) in specificity (95.00% versus 75.00%), PPV (75.00% versus 37.50%), and accuracy (95.65% [22/23] versus 78.26% [18/23]) while maintaining 100.00% sensitivity and 100.00% NPV. Routine parallel implementation of CTS and age-calibrated STUMBL scoring in the emergency department provides a rapid, objective, and highly accurate framework for early ICU triage and mortality risk stratification in blunt thoracic trauma. Furthermore, age-adjusted multivariable logistic modeling confirmed that exceeding the optimal threshold independently conferred more than a 165-fold increase in mortality odds for both CTS ≥ 6 (aOR = 167.06, p < 0.001) and STUMBL ≥ 25 (aOR = 170.73, p < 0.001). In clinical practice, immediate primary-survey screening with room-air pulse oximetry (flagging SpO_2 ≤ 90%) followed by parallel bedside computation of CTS (cut-off ≥6) and age-calibrated STUMBL scoring (cut-off ≥25 in ages 18–60 years and ≥28 in ages >60 years) establishes a rapid, cost-free, and highly reliable clinical decision pathway to guide early ICU escalation and reduce preventable deaths from blunt thoracic trauma.

6. Declarations

Ethics Approval And Consent To Participate

This study received institutional ethics approval from the Health Research Ethics Committee of Dr. Mohammad Hoesin Central General Hospital, Indonesia (approval number DP/04.03/D.XVII.9.4/ETIK/286/2026), and was conducted in full accordance with the Declaration of Helsinki. Due to the retrospective observational nature of medical record abstraction, individual written informed consent was waived by the Ethics Committee while strict patient de-identification was maintained.

Consent For Publication

Not applicable; no identifiable individual patient images or personal identifiers are included.

Availability Of Data And Materials

All data generated or analyzed during this study are included in the published tables and figures; de-identified analytical records are available from the corresponding author upon reasonable request.

Artificial Intelligence Use Statement

OpenAI Codex was used for language editing, consistency checks, and journal-template copyediting during this editorial workflow. Research data and analyses were not generated or replaced by this workflow. The authors must review the edited text, verify all scientific content, and confirm responsibility for the final manuscript before submission.

Author Contributions (Credit)

YRS: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Visualization, Writing – Original Draft, and Writing – Review & Editing. AHLT: Conceptualization, Methodology, Supervision, Validation, and Writing – Review & Editing. T: Methodology, Supervision, Formal analysis, Validation, and Writing – Review & Editing. All authors must review and approve this edited manuscript before submission.

Funding

This research received no external grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing Interests

The authors declare that they have no competing financial or non-financial interests.

Acknowledgements

The authors thank the medical and administrative staff of the Emergency Department, Department of Surgery, and Medical Records Installation of Dr. Mohammad Hoesin Central General Hospital, Palembang, for their support during data collection.

7. References

1. Adal O, Tareke AA, Bogale EK, et al. Mortality of traumatic chest injury and its predictors across sub-Saharan Africa: systematic review and meta-analysis, 2024. BMC Emerg Med. 2024;24(1):32. doi: 10.1186/s12873-024-00951-w
2. Karajizadeh M, Yadollahi M, Yousefi MR, et al. Mortality risk factors in chest trauma patients in a level one trauma center in southern Iran. Sci Rep. 2025;15(1):32310. doi: 10.1038/s41598-025-13830-6
3. Hefny AF, Elkamhawy AA, Fathi SA, et al. Outcomes and predictors of mortality in patients with blunt chest trauma admitted to the intensive care unit: a retrospective study. Front Med (Lausanne). 2025;12:1601033. doi: 10.3389/fmed.2025.1601033
4. Zhang S, Xiao X, Wang J, et al. Epidemiological and clinical characteristics of road traffic crashes related thoracic traumas: analysis of 5095 hospitalized chest injury patients. J Cardiothorac Surg. 2021;16(1):220. doi: 10.1186/s13019-021-01599-4
5. Yüksel HS, Aslaner MA, Coşkun Yaş S, et al. Rib fractures in trauma patients: prevalence, associated injuries, and mortality trends. Cureus. 2024;16(9):e70137. doi: 10.7759/cureus.70137
6. Nsengiyumva B, Henry MB, Kuntz HM, et al. Chest trauma epidemiology and emergency department management in a tertiary teaching hospital in Kigali, Rwanda. Afr J Emerg Med. 2022;12(4):428–431. doi: 10.1016/j.afjem.2022.09.004
7. Elgar G, Smiley A, Latifi R. Major risk factors for mortality in elderly and non-elderly adult patients emergently admitted for blunt chest wall trauma: hospital length of stay as an independent predictor. Int J Environ Res Public Health. 2022;19(14):8729. doi: 10.3390/ijerph19148729
8. Kurniawan I, Leonas R, Maritska Z. Prehospital determinants of emergency department mortality in adult trauma at a South Sumatran tertiary referral centre: a cross-sectional study with penalised-likelihood reanalysis. Sriwijaya J Surg. 2026;9(2):177–188. doi: 10.37275/sjs.v9i2.155
9. Umar A, Marrantiza S, Theodorus. The relationship between thoracic trauma scoring and mortality in blunt thoracic trauma patients at Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia. Sriwijaya J Surg. 2023;6(2):558–562. doi: 10.37275/sjs.v6i2.88
10. Seamaladi H, Nugraha A, Bahar E. Beyond clinical intuition: quantitative mortality prediction in blunt thoracic trauma using the Thoracic Trauma Severity Score (TTSS). Sriwijaya J Surg. 2025;8(1):878–890. doi: 10.37275/sjs.v8i1.127
11. Pressley CM, Fry WR, Philp AS, et al. Predicting outcome of patients with chest wall injury. Am J Surg. 2012;204(6):910–913. doi: 10.1016/j.amjsurg.2012.05.015
12. Chen J, Jeremitsky E, Philp F, et al. A chest trauma scoring system to predict outcomes. Surgery. 2014;156(4):988–993. doi: 10.1016/j.surg.2014.06.045
13. Elsaied Hussein MH, Fadl Mahmoud I, Eita YMS, et al. A prospective study of chest trauma scoring system as a morbidity and mortality predictor in patients with blunt chest trauma. Med J Islam Repub Iran. 2024;38:4. doi: 10.47176/mjiri.38.4
14. Kim H, Lee MS, Yoon SY, et al. Validation of chest trauma scoring systems in polytrauma: a retrospective study with 1,038 patients

in Korea. J Trauma Inj. 2024;37(2):114–123. doi: 10.20408/jti.2023.0087

15. Battle CE, Hutchings H, Lovett S, et al. Predicting outcomes after blunt chest wall trauma: development and external validation of a new prognostic model. Crit Care. 2014;18(3):R98. doi: 10.1186/cc13873
16. Mukerji S, Tan E, May C, et al. Retrospective validation of a risk stratification tool developed for the management of patients with blunt chest trauma (the STUMBL score). Emerg Med Australas. 2021;33(5):841–847. doi: 10.1111/1742-6723.13740
17. Callisto E, Costantino G, Tabner A, et al. The clinical effectiveness of the STUMBL score for the management of ED patients with blunt chest trauma compared to clinical evaluation alone. Intern Emerg Med. 2022;17(6):1785–1793. doi: 10.1007/s11739-022-03001-0
18. Giamello JD, Martini G, Prato D, et al. A retrospective validation study of the STUMBL score for emergency department patients with blunt thoracic trauma. Injury. 2023;54(1):39–43. doi: 10.1016/j.injury.2022.08.028
19. Kalideen A, Sewchuran T. Chest In Time: investigating the utility of a novel X-ray based blunt chest trauma clinical prediction model in a resource-limited setting in South Africa. SA J Radiol. 2026;30(1):3344. doi: 10.4102/sajr.v30i1.3344
20. Battle C, Cole E, Carter K, et al. Clinical prediction models for the management of blunt chest trauma in the emergency department: a systematic review. BMC Emerg Med. 2024;24(1):189. doi: 10.1186/s12873-024-01107-6
21. Yılmaz E, Ak R, Seyhan AU, et al. Validation of the study of the management of blunt chest wall trauma (STUMBL) score for predicting in-hospital complications in patients with blunt chest trauma. Ulus Travma Acil Cerrahi Derg. 2025;31(10):959–965. doi: 10.14744/tjtes.2025.72690
22. Ahmed-Issap A, Raseta M, Mahendran K, et al. A retrospective study examining the effectiveness of the Study of the Management of Blunt Chest Wall Trauma (STUMBL) scoring tool in predicting complications following rib fractures. Cureus. 2026;18(1):e101872. doi: 10.7759/cureus.101872
23. Isharanto A, Suhanda EY, Bayuadi IS, et al. Improving blunt thoracic trauma prognosis: validation of the Study of the Management of Blunt Chest Wall Trauma score against Injury Severity Score in an Indonesian tertiary hospital. J Chest Surg. 2026;59(5):202–210. doi: 10.5090/jcs.25.143
24. Buchholz CJ, Jia L, Manea C, et al. Revised Intensity Battle Score (RIBS): development of a clinical score for predicting poor outcomes after rib fractures. Am Surg. 2023;89(11):4668–4674. doi: 10.1177/00031348221123087
25. Buchholz CJ, D'Aquila ML, Lollar DI. External validation of novel Revised Intensity Battle Score and comparison of static rib fracture scoring systems. J Trauma Acute Care Surg. 2024;96(3):466–470. doi: 10.1097/TA.0000000000004199
26. Hefny AF, Almansoori TM, Smetanina D, et al. Streamlining management in thoracic trauma: radiomics- and AI-based assessment of patient risks. Front Surg. 2024;11:1462692. doi: 10.3389/fsurg.2024.1462692
27. Temiz Bekce S, Goksuluk D. Evaluating the effectiveness of trauma scores in predicting morbidity and mortality in patients with concomitant thoracic trauma. Ulus Travma Acil Cerrahi Derg. 2026;32(6):652–660. doi: 10.14744/tjtes.2025.43780
28. Miles TJ, Mendez-Reyes J, Mokhtari AK, et al. Machine learning predicts mortality and respiratory failure in patients admitted with rib fractures. J Surg Res. 2025;313:153–160. doi: 10.1016/j.jss.2025.06.040
29. Toelle LJ, McNickle AG, Feery D, et al. The pulmonary contusion score: development of a simple scoring system for blunt lung injury. Surg Pract Sci. 2024;17:100247. doi: 10.1016/j.sipas.2024.100247
30. Cheng R, Yang M, Zhang Y, et al. Risk factors affecting pulmonary complications in elderly patients with isolated rib fractures. J Thorac Dis. 2025;17(2):542–550. doi: 10.21037/jtd-24-1323
31. Ekestubbe L, Forssten MP, Cao Y, et al. Morbidity prediction in conservatively managed rib fracture patients. Eur J Trauma Emerg Surg. 2025;51(1):184. doi: 10.1007/s00068-025-02860-4
32. Mohseni S, Forssten MP, Ismail AM, et al. Investigating the link between frailty and outcomes in geriatric patients with isolated rib fractures. Trauma Surg Acute Care Open. 2024;9(1):e001206. doi: 10.1136/tsaco-2023-001206
33. Cooper EV, Hughes I, Campbell D, et al. Demographics, treatment, and outcomes of rib fractures in older adults: a retrospective multi-institutional study. Eur J Trauma Emerg Surg. 2025;51(1):281. doi: 10.1007/s00068-025-02962-z
34. Choi J, Villarreal JA, Handelsman R, et al. Prospective multicenter external validation of the rib fracture frailty index. J Trauma Acute Care Surg. 2025;99(2):204–211. doi: 10.1097/TA.0000000000004624
35. Bass GA, Duffy CC, Kaplan LJ, et al. The revised cardiac risk index is associated with morbidity and mortality independent of injury severity in elderly patients with rib fractures. Injury. 2023;54(1):56–62. doi: 10.1016/j.injury.2022.11.039
36. Plumb J, Topper GV, Metheny J, et al. Surgical stabilization of rib fractures in geriatric trauma patients: a National Trauma Data Bank review. J Am Geriatr Soc. 2026;74(3):738–747. doi: 10.1111/jgs.70297
37. Nelson A, Reina R, Northcutt A, et al. Prospective validation of the Rib Injury Guidelines for traumatic rib fractures. J Trauma Acute Care Surg. 2022;92(6):967–973. doi: 10.1097/TA.0000000000003535
38. Kishawi SK, Ho VP, Bachman KC, et al. Are we underestimating the morbidity of single rib fractures? J Surg Res. 2021;268:174–180. doi: 10.1016/j.jss.2021.06.048
39. Bläsius FM, Wutzler S, Störmann P, et al. Predicting tracheostomy in multiple injured patients with severe thoracic injury (AIS ≥ 3) with the new T3P-Score: a multivariable regression prediction analysis. Sci Rep. 2023;13(1):3260. doi: 10.1038/s41598-023-30461-x
40. Çınar E, Usul E, Demirtaş E, et al. The role of trauma scoring systems and serum lactate level in prediction of prognosis in thoracic trauma. Ulus Travma Acil Cerrahi Derg. 2021;27(6):619–623. doi: 10.14744/tjtes.2021.22498
41. Buz M, Ustaalioğlu İ. Predicting mortality in penetrating thoracic trauma in the emergency department: the prognostic value of the glucose-to-potassium ratio. Ulus Travma Acil Cerrahi Derg. 2025;31(1):40–46. doi: 10.14744/tjtes.2024.96644
42. Kaya M, Yildirim H, Toprak M, et al. Comparison of trauma scoring systems for predicting mortality in emergency department patients with traffic-related multiple trauma. Diagnostics (Basel). 2025;15(12):1563. doi: 10.3390/diagnostics15121563
43. Faisal M, Komar MH, Theodorus. Diagnostic accuracy of serum bilirubin and magnetic resonance cholangiopancreatography for differentiating benign from malignant extrahepatic obstructive jaundice: a histopathology-referenced study. Sriwijaya J Surg. 2026;9(2):168–176. doi: 10.37275/sjs.v9i2.154
44. Alfiansyah A, Satria G, Theodorus. Thoracoscore risk bands substantially underpredict in-hospital mortality after thoracic surgery: an external validation study in a tertiary referral cohort. Sriwijaya J Surg. 2026;9(2):202–222. doi: 10.37275/sjs.v9i2.157
45. Manurung SBH, Manawan EEU, Bahar E. Only the Boey score discriminates in-hospital mortality after gastric perforation: a STARD-compliant comparison with the Jabalpur score. Sriwijaya J Surg. 2026;9(2):223–239. doi: 10.37275/sjs.v9i2.158