

ORIGINAL ARTICLE

Prognostic Utility and Exploratory Cut-off Threshold of the First 24-Hour APACHE II Score for In-Hospital Mortality After Thoracotomy: A Single-Center Retrospective Cohort Study

Awan Rochaniawan¹, Ahmat Umar², Erial Bahar³¹ Surgery Study Program, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia² Department of Thoraco-Cardiovascular Surgery, Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia³ Department of Anatomy, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

ARTICLE INFO

Article history:

Received: August 12, 2025

Revised: September 10, 2025

Accepted: October 4, 2025

Available Online: October 15, 2025

Published: November 2, 2025

Keywords:

APACHE II score; Critical care; In-hospital mortality; Risk stratification; Thoracotomy

Corresponding author:

Awan Rochaniawan

Email:

awan.rochaniawan@gmail.com

DOI:

<https://doi.org/10.37275/sjs.v8i2.159>

Access licence:

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0).

Copyright:

© 2025 The Authors. Authors retain copyright and grant the journal the right of first publication. Published by HM Publisher.

ABSTRACT

Background: Open thoracotomy imposes severe physiological stress, so objective early risk stratification is needed at General Intensive Care Unit (GICU) admission.**Objective:** To evaluate how well the first 24-hour Acute Physiology and Chronic Health Evaluation II (APACHE II) score predicts in-hospital mortality after thoracotomy, to set an exploratory cut-off, and to examine differences by surgical indication and organ component.**Methods:** This retrospective cohort study analyzed 33 consecutive adults admitted directly to the GICU of Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia, after open thoracotomy (January–December 2024). Analyses included Mann-Whitney U, Spearman's ρ , logistic regression, and Receiver Operating Characteristic (ROC) analysis with Youden's J index.**Results:** In-hospital mortality was 27.3% (9/33). Non-survivors had higher median APACHE II scores than survivors (23.0 [IQR: 19.0–27.0] vs. 8.0 [IQR: 6.75–12.0]; $U = 9.5$, $p < 0.001$). The score correlated with mortality ($\rho = 0.706$, $p < 0.001$) and remained significant on multivariable regression (aOR = 1.74; 95% CI: 1.13–2.69; $p = 0.0127$). The AUC was 0.956 (95% CI: 0.891–1.000). A cut-off of ≥ 12.5 gave 88.9% sensitivity, 87.5% specificity, 72.7% positive predictive value, and 95.5% negative predictive value. Mortality was 32.0% in malignancy and 12.5% in non-malignancy cases ($p = 0.394$). Glasgow Coma Scale (6.33 vs. 1.83 points), serum creatinine (3.44 vs. 0.96), and arterial pH (1.89 vs. 0.33) components differed between non-survivors and survivors (all $p < 0.001$).**Conclusion:** The initial 24-hour APACHE II score is a hypothesis-generating predictor of post-thoracotomy mortality at a candidate threshold of ≥ 12.5 .

1. Introduction

Open thoracotomy is the surgical creation of a formal entry into the pleural cavity through the chest wall. It remains an indispensable approach for complex pulmonary, pleural, mediastinal, and esophageal disease.¹ Video-assisted thoracoscopic surgery (VATS) and robotic platforms have expanded for early-stage parenchymal lesions. Even so, open access still gives unmatched exposure in several settings. These include locally advanced intrathoracic malignancy, dense fibrothorax, chronic pleural empyema needing radical decortication, deep descending necrotizing mediastinitis, and life-threatening penetrating or blunt thoracic trauma.^{1,2} That exposure has a price. The physiological insult is large, multisystem, and unlike that of nearly any other noncardiac operation.^{2,3} The surgeon must divide major accessory respiratory muscles, including the latissimus dorsi, serratus anterior, and intercostal musculature; the ribs are then retracted for a long time or resected, and the patient usually undergoes intraoperative one-lung ventilation (OLV).^{2,4} These mechanical and ventilatory insults trigger a strong neuroendocrine and inflammatory stress response. Chest wall compliance falls, functional residual capacity shrinks, and patients become prone to a cascade of dangerous pulmonary and extrapulmonary complications after surgery.^{3,4}

As a result, open thoracotomy brings considerable perioperative morbidity. In-hospital mortality ranges from 3% to 15% in contemporary elective series. It exceeds 20% to 30% in high-acuity cohorts that need intensive care after surgery.^{5,6}

Many patients having major open thoracic procedures therefore go straight to a General Intensive Care Unit (GICU). There they receive invasive hemodynamic monitoring, ventilator weaning, multimodal analgesia, and organ support.^{4,7} In this resource-intensive setting, objective and accurate stratification during the first postoperative hours is a foundation of care.^{7,8} Early recognition of a patient heading toward multi-organ failure lets the intensivist and thoracic team act early. They can optimize hemodynamics, protect organs, and watch more closely.^{8,9} The reverse also matters. Identifying patients likely to recover without complications allows safe de-escalation of invasive monitoring. It also allows quicker transfer to the ward and better bed allocation in tertiary centers working under high occupancy.^{4,9}

The physiological toll of an open thoracotomy begins with the incision and the spreading of the ribs.^{1,2} Standard posterolateral and anterolateral approaches divide the latissimus dorsi, serratus anterior, trapezius, or rhomboid muscles. Electrocautery strips the periosteum, and rigid metal retractors distract the intercostal space.^{1,2} This force often produces occult costovertebral or costochondral microfractures. It can also compress the intercostal nerve bundles and cause localized skeletal muscle ischemia-reperfusion injury.^{2,10} Afterward, nociceptive afferent input through the intercostal and phrenic nerves inhibits diaphragmatic contraction by reflex and causes involuntary chest wall splinting^{2,9}; forced vital capacity (FVC) and functional residual capacity (FRC) therefore usually fall by 40% to 60% during the first 24 to 48 hours. Tidal breathing shifts below the closing volume of dependent small airways.

Basal atelectasis, mucociliary stasis, and intrapulmonary shunting follow.^{2,10,11}

Intraoperative OLV adds a distinct two-lung injury pattern.^{2,10,11} The dependent ventilated lung receives the whole right ventricular output. It is ventilated inside a constrained hemithorax, compressed by the mediastinum and abdominal contents. This predisposes it to volutrauma, barotrauma, atelectrauma, and high mechanical power transfer.^{10,11} The non-dependent operative lung, meanwhile, collapses completely and undergoes hypoxic pulmonary vasoconstriction, surgical compression, and manipulation. When the bronchus is re-expanded, reactive oxygen species (ROS) surge, the endothelial glycocalyx sheds, and the alveolar-capillary barrier breaks down. Clinically this is recognized as re-expansion pulmonary edema and ischemia-reperfusion lung injury.^{2,11,12} Multicenter randomized trials and meta-analyses by Piccioni et al.,² El Tahan et al.,¹¹ and Liu et al.¹⁰ show this clearly. Postoperative pulmonary complications (PPCs) still occur in 24.6% to 30.8% of thoracic surgical patients. This holds even when contemporary lung-protective OLV is applied strictly. It uses low tidal volumes of 4–6 mL/kg predicted body weight, titrated positive end-expiratory pressure (PEEP), and recruitment maneuvers, together with goal-directed norepinephrine-supported fluid regimens.¹² Pulmonary cytokines and damage-associated molecular patterns (DAMPs) also spill into the systemic circulation. This turns a local intrathoracic insult into a generalized systemic inflammatory response syndrome (SIRS) that threatens the brain, kidneys, and cardiovascular system.^{9,13,14}

For four decades, intensive care clinicians have used composite physiological scores to complement bedside judgment and to standardize risk assessment across mixed ICU populations^{7,8}; the Acute Physiology and Chronic Health Evaluation II (APACHE II) score was first formulated by Knaus and colleagues.⁷ It is still the most widely validated severity-of-illness system in the world.^{7,8} The score combines 12 routine physiological, laboratory, and neurological measurements. Each is recorded at its most deranged value during the first 24 hours of intensive care admission. Age points and weighted chronic organ insufficiency or immunocompromised status are added. The result is an integer from 0 to 71.⁷ Preoperative thoracic risk models include the Thoracscore, the European Society Objective Score (ESOS), and the Eurolung models. They rely mainly on static demographic and anatomical variables. APACHE II instead captures the dynamic physiological aftermath of the surgical and anesthetic insult itself.^{3,9,11}

The clinical picture of the post-thoracotomy patient fits closely with the acute variables that APACHE II quantifies.^{9,11} During the first 24 hours after open thoracotomy, arterial hypoxemia is common. It stems from ventilation-perfusion mismatch, alveolar derecruitment, and pulmonary capillary endothelial injury, in both the ventilated dependent lung and the re-expanded non-dependent lung after OLV^{2,11}; incisional pain and diaphragmatic dysfunction add respiratory splinting, alveolar hypoventilation, hypercapnia, and respiratory acidosis. Fluid shifts, blood loss, and myocardial strain add to this. So does release of pro-inflammatory cytokines such as interleukin-6 (IL-6) and tumor necrosis factor- α (TNF- α). Together they show up as tachycardia, fever or hypothermia, leukocytosis, hypotension, lactic acidosis, and prerenal or intrinsic acute kidney injury (AKI).^{9,11} Every one of these perturbations carries weight in the 12 acute physiological domains of the instrument; it is therefore reasonable to hypothesize that the first 24-hour APACHE II score works as a sensitive, integrated barometer of mortality risk in post-thoracotomy GICU patients.^{8,9}

A major conceptual difficulty is that any single prognostic threshold must serve a population that is etiologically and pathophysiologically heterogeneous.^{13,15} Open thoracotomy is not one disease. It is an anatomical corridor used for very different clinical syndromes.^{1,13} A patient who has an extended lobectomy or pneumonectomy for intrathoracic malignancy arrives with a chronic substrate. It includes cancer-associated cachexia, sarcopenia, nutritional depletion, paraneoplastic inflammation, and prior cytotoxic or radiation therapy.^{13,15} A patient who has open thoracotomy and decortication for stage III fibrinous empyema or mediastinitis presents instead with an acute or subacute hyper-inflammatory septic cascade.^{1,16–18} A patient taken to emergency exploratory thoracotomy for penetrating or blunt chest trauma suffers abrupt hemorrhagic shock and direct mechanical visceral disruption.^{19,20} Pooling these subgroups into one cohort risks an averaged prognostic signal. That signal can hide real differences in baseline severity and in mortality trajectories between oncologic and non-oncologic indications.^{13,15}

International comparisons of thoracic surgical risk models show why immediate postoperative physiological scoring in the GICU is indispensable.^{3,15,21,22} The French Thoracscore and the European Society of Thoracic Surgeons (ESTS) Parsimonious Eurolung 1 and Eurolung 2 models are preoperative tools; they were derived mostly in high-volume European centers where minimally invasive VATS resection of early-stage non-small cell lung cancer (NSCLC) predominates.^{21,22} External validation in non-European or open thoracotomy populations shows marked calibration drift and weak discrimination. Pathy et al.¹⁵ tested the Thoracscore in an Indian cohort having posterolateral thoracotomy. Predicted mortality was only 0.44%, but observed mortality was 3.2%, an underprediction of more than seven-fold. Ponzholzer et al.^{21,22} reported in external validation cohorts that the Parsimonious Eurolung 1 and Eurolung 2 scores missed patient-level vulnerability such as computed-tomography-quantified sarcopenia. They also correlated weakly with postoperative morbidity and mortality after lung cancer resection. Bang et al.³ found that open thoracotomy and thoracoscopic-to-open conversion prolonged recovery and hospital stay compared with minimally invasive lobectomy. Blood loss, prolonged endotracheal intubation, and pleural drainage duration drove this difference. All preoperative scores share one limit: they cannot incorporate intraoperative physiological disturbance. The first 24-hour postoperative APACHE II score captures exactly those variables on arrival in the GICU.^{7–9}

The prognostic performance of APACHE II is well documented in general mixed medical-surgical ICUs and after emergency abdominal surgery. Its discrimination, optimal threshold, and organ-component drivers after thoracotomy in Southeast Asian tertiary settings remain insufficiently characterized.^{3,23,24} Regional features can change how established models perform. These include late-stage oncologic referral, a heavy burden of complicated pleuropulmonary infection in tuberculosis-endemic areas, and different perioperative critical care workflows^{3,15,23}; we therefore studied a consecutive cohort of post-thoracotomy patients admitted to the GICU of a major tertiary referral center in South Sumatra, Indonesia, with two analytic aims. The first was to evaluate the association and discriminatory performance of the initial 24-hour APACHE II score for in-hospital mortality. It also aimed to derive an exploratory, statistically optimal cut-off using Receiver Operating Characteristic (ROC) analysis. The second went beyond aggregate score validation. We compared malignancy with non-malignancy patients in a post-hoc etiological subgroup

analysis. We also decomposed all 12 acute physiological components of the score to identify which organ-system failures drive postoperative death.

2. Methods

Study design and setting

This was a single-center, retrospective, observational cohort study. It followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines and the principles of the Standards for Reporting of Diagnostic Accuracy Studies (STARD).^{25,26} Data came from the General Intensive Care Unit (GICU) of Dr. Mohammad Hoesin General Hospital in Palembang, South Sumatra, Indonesia. The hospital is the main government-funded academic tertiary referral center for South Sumatra province and the surrounding regional health networks. Its GICU is a 20-bed, high-acuity, mixed medical-surgical unit. Board-certified intensivists, intensive care fellows, and anesthesiology and surgical residents staff it around the clock, 24 hours a day; the unit offers comprehensive organ support: invasive lung-protective mechanical ventilation, bronchoscopy, invasive arterial and central venous monitoring, vasoactive and inotropic infusion protocols, and continuous renal replacement therapy (CRRT).

Participants and eligibility

The target population was all consecutive adults who had an open thoracotomy in the central operating theater complex. They were then transferred directly to the GICU for postoperative intensive care between January 1, 2024, and December 31, 2024. Total consecutive sampling captured every eligible patient in the 12-month window. This limits sampling selection bias within the 12-month window. It gives a complete institutional picture of high-acuity post-thoracotomy admissions. Patients were included if they met four mandatory criteria. First, they were 18 years of age or older at the index operation. Second, they were admitted directly to the GICU right after open thoracotomy. We defined this anatomically as any approach with a formal incision through the chest wall musculature and intercostal space into the pleural cavity; it covered posterolateral, anterolateral, and muscle-sparing axillary thoracotomies, and procedures begun by VATS that needed conversion to open thoracotomy.^{1,2} Third, GICU stay lasted at least 24 hours, giving a complete window for initial physiological scoring. Fourth, complete electronic and paper records held every clinical, physiological, and laboratory parameter needed to compute the APACHE II score. Patients were excluded if their records lacked essential laboratory or arterial blood gas values among the 12 acute physiological variables; patients under 18 years of age were also excluded, as were those who had median sternotomy or purely minimally invasive thoracoscopic procedures without conversion to open thoracotomy.

Variables and data collection

A standardized, pre-piloted abstraction instrument followed the original operational definitions of Knaus et al.⁷ A trained surgical investigator extracted the data from the hospital's integrated electronic medical record and the archived GICU hourly flow sheets. To protect data fidelity and limit transcription error, a second independent investigator audited a randomly selected 15% subsample of charts. By consensus, concordance was 100% across all physiological, demographic, and outcome variables.

Baseline variables abstracted for each patient were age at surgery in completed years, biological sex (male or female), and the primary surgical indication for open thoracotomy; using the definitive operative findings and the histopathological or

microbiological diagnosis, we classified the indication into three etiological categories. Malignancy covered primary bronchogenic carcinoma, malignant mediastinal neoplasms, and intrathoracic metastatic resections.^{13,27,28} Infection covered complicated stage III thoracic empyema needing decortication, ruptured or refractory lung abscess, and descending necrotizing mediastinitis.^{1,16,17} Trauma covered penetrating or blunt thoracic injury requiring urgent exploration, hemorrhage control, or parenchymal repair.^{19,20}

The APACHE II score was calculated for each patient from the most deranged value of each of the 12 acute physiological parameters during the first 24 hours after GICU arrival^{7,8}; following standard APACHE II methodology, the most deranged value was the measurement furthest from the normal reference range, and so the one earning the most points. Each parameter scores 0 to 4 points, except the Glasgow Coma Scale and serum creatinine.⁷ The 12 parameters were core body temperature (°C), mean arterial pressure (MAP, mmHg), ventricular heart rate (beats/min), respiratory rate (breaths/min), and pulmonary oxygenation; oxygenation was quantified by the alveolar-arterial oxygen tension gradient (A-aDO₂) when the fraction of inspired oxygen (FiO₂) was ≥0.50. It was quantified by arterial partial pressure of oxygen (PaO₂, mmHg) when FiO₂ was <0.50. The remaining parameters were arterial pH, serum sodium (mmol/L), and serum potassium (mmol/L). Serum creatinine (mg/dL) had its points doubled from 0–4 to 0–8 in acute renal failure. Hematocrit (%) and white blood cell count ($\times 10^9/L$) completed the laboratory set. Neurological status used the Glasgow Coma Scale (GCS), contributing 15 minus the actual GCS score, or 0 to 12 points.⁷

Sedation and neuromuscular blockade could inflate neurological severity during immediate postoperative ventilation. To prevent this, we assigned the GCS from the documented assessment made just before sedative or paralytic agents were given, or during a structured daily sedation interruption^{7,29}; sometimes no neurological deficit existed before sedation, and altered consciousness was due only to anesthetic or sedative infusion. In that case a normal GCS of 15 (0 APACHE II neurological points) was assigned, following international intensive care scoring conventions; persistent coma or stupor that was verified after sedation washout, or that stemmed from hypoxic-ischemic, septic, or metabolic encephalopathy, was scored at the lowest verified GCS.^{29–31} Age points were 0 (≤44 years), 2 (45–54 years), 3 (55–64 years), 5 (65–74 years), or 6 points (≥75 years). Chronic health points were added for documented severe preexisting organ insufficiency or immunocompromised status. The first applied to biopsy-proven cirrhosis, New York Heart Association Class IV heart failure, severe chronic restrictive or obstructive pulmonary disease, and chronic dialysis dependency. The second applied to active chemotherapy, radiotherapy, high-dose corticosteroid therapy, or advanced malignancy. These points were 2 for elective postoperative admission and 5 for emergency postoperative admission.⁷ The composite APACHE II score (theoretical range 0–71 points) was the arithmetic sum of the Acute Physiology Score (APS), age points, and chronic health points.

To limit measurement bias inherent in retrospective intensive care scoring, we enforced strict operational rules for the two highest-weighted components, the Glasgow Coma Scale (GCS) and serum creatinine^{7,14,29,32}; first, patients in the immediate post-thoracotomy GICU setting often arrive intubated and sedated with short-acting propofol, dexmedetomidine, or fentanyl infusions. Following international APACHE II conventions validated in neurological and surgical ICUs,^{7,29,31} depression of consciousness caused solely by residual general

anesthesia or continuous sedative infusion was never scored as neurological organ failure. We used instead the pre-sedation neurological status, or the best response documented during a structured sedation hold within the first 24 hours; elevated GCS component points (15 minus the observed GCS, spanning 0 to 12 points) therefore reflect genuine, persistent neurological dysfunction. Examples are perioperative hypoxic-ischemic encephalopathy, sepsis-associated encephalopathy, metabolic or uremic coma, and severe hypoactive or agitated delirium unmasked after anesthetic washout.^{29,31} Second, we followed the original APACHE II specification of Knaus et al.⁷ Creatinine points (normally 0 to 4) were doubled, to 0 to 8, whenever creatinine was elevated (≥ 1.5 mg/dL) together with acute oliguria. The same applied to an acute rise meeting Kidney Disease: Improving Global Outcomes (KDIGO) acute kidney injury criteria. This weights the severe prognostic hazard of de novo postoperative acute renal failure appropriately against stable chronic kidney disease.^{14,32–34}

Statistical analysis

The primary endpoint was all-cause in-hospital mortality. It was defined as death from any cause during the index hospitalization after open thoracotomy and recorded as a dichotomous outcome (survived to hospital discharge vs. deceased in hospital).³⁸ Analyses used the Statistical Package for the Social Sciences (SPSS) version 27.0 (IBM Corp., Armonk, NY, USA), cross-verified with Python 3.14 scientific libraries (SciPy version 1.15 and Statsmodels version 0.14). Categorical variables are summarized as exact frequencies (n) and percentages (%) summing to 100.0%. Group comparisons used Fisher's exact test, with odds ratios (OR) and 95% confidence intervals (CIs). Distributional normality of continuous variables was assessed with the Shapiro-Wilk test (W statistic). Normally distributed variables, such as age, are given as mean $\pm$ standard deviation (SD) and compared with Welch's two-sample t-test. Non-normal variables, including the composite APACHE II score, are given as median with interquartile range [IQR: 25th–75th percentile] and also as mean $\pm$ SD, to ease international meta-analytic comparison.^{8,9}

The composite APACHE II score departed significantly from normality across the cohort (Shapiro-Wilk $W = 0.8629$, $p = 0.0007$). The non-parametric Mann-Whitney U test was therefore the primary inferential test for comparing score distributions between survivors and non-survivors; it was supplemented by the standardized Z statistic, the non-parametric effect size r ($|Z|/\sqrt{N}$), and the rank-biserial correlation coefficient ($r_{rb} = 2 \times \text{AUC} - 1$). Spearman's rank correlation coefficient (ρ) quantified the graded monotonic association between the initial 24-hour APACHE II score and in-hospital mortality. Univariable and multivariable binary logistic regression gave unadjusted OR and adjusted odds ratios (aOR) with 95% CIs, the latter adjusting for surgical indication (malignancy vs. non-malignancy) and age^{23,24}; discriminatory performance was assessed with an empirical Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) with its 95% CI. The statistically optimal dichotomous cut-off was the value maximizing Youden's J index ($J = \text{Sensitivity} + \text{Specificity} - 1$).²⁴ At that threshold we built a 2×2 contingency matrix. From it we computed sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). We also computed overall diagnostic accuracy (with Wilson score 95% CIs), positive and negative likelihood ratios (LR+ and LR-), and the diagnostic odds ratio (DOR). A planned post-hoc subgroup analysis then stratified the cohort into Malignancy ($n = 25$) and Non-Malignancy ($n = 8$, combining Infection and Trauma) groups. An organ-component

decomposition compared the mean point contributions of all 12 acute physiological variables between survivors and non-survivors. For every inferential test, a two-tailed p -value < 0.05 defined statistical significance.

The study covers a specialized population. It comprises open thoracotomy admissions needing at least 24 hours of GICU care at one tertiary center over one calendar year ($N = 33$, with 9 deaths and 24 survivors). The analysis plan therefore included explicit safeguards against small-sample inferential bias.^{23,24} First, APACHE II is an integer scale (0, 1, 2, ..., 71 points), so empirical ROC coordinates were evaluated at the arithmetic midpoints between consecutive observed integer values. A Youden-optimal threshold of ≥ 12.5 points thus separates patients with an integer score of ≤ 12 (low-risk stratum) from those with an integer score of ≥ 13 (high-risk stratum).²⁴ Second, we computed Wilson score 95% confidence intervals for sensitivity, specificity, PPV, NPV, and accuracy. They cover better than Wald normal approximations in small samples and at extreme proportions. Third, the multivariable binary logistic regression model had three predictors (APACHE II score, malignancy indication, and age) for 9 mortality events, an Events-Per-Variable (EPV) ratio of 3.0. We checked model stability. For this, we compared the unadjusted log-odds coefficient of the APACHE II score ($\beta = 0.549$, $\text{OR} = 1.73$, $p = 0.0192$) with the multivariable adjusted coefficient ($\beta = 0.556$, $\text{aOR} = 1.74$, $p = 0.0127$).

Ethical approval

The study protocol was reviewed and approved by the Ethics Committee of CMHC Research Center Indonesia (Reference number: 2024/155). The Directorate of Medical and Nursing Services of Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia, authorized access to anonymized clinical records. Following national regulatory standards and the ethical principles of the Declaration of Helsinki, the Ethics Committee CMHC Research Center Indonesia formally waived individual written informed consent; the grounds were the retrospective, non-interventional design and the strict de-identification of all patient health information before statistical analysis.

3. Results

Cohort characteristics, indications, and overall outcomes

During the 12-month study period from January 1, 2024, to December 31, 2024, a total of 33 consecutive adult patients met all eligibility criteria following open thoracotomy and direct admission to the GICU of Dr. Mohammad Hoesin General Hospital ($N = 33$), with zero patients excluded due to missing physiological or outcome records. As detailed in Table 1 and illustrated in Figure 1 (Panel a), the overall cohort had a mean chronological age of 52.4 ± 14.8 years (range: 24.1–77.2 years). Biological sex distribution demonstrated a clear male predominance, with 20 male patients (60.6%) and 13 female patients (39.4%), as summarized in Table 1. When stratified by final hospital survival status (Table 1 and Figure 1a), non-survivors were numerically older than survivors (60.8 ± 14.5 years vs. 49.3 ± 13.9 years), exhibiting a clinically notable 11.5-year age difference (95% CI: -0.6 to +23.6 years) that approached but did not reach formal statistical significance on two-sample comparison ($t = 2.04$, $p = 0.060$). Sex distribution did not differ significantly between outcome groups, with males accounting for 66.7% (6/9) of non-survivors and 58.3% (14/24) of survivors ($\text{OR} = 1.43$; 95% CI: 0.28–7.26; Fisher's exact $p = 1.000$).

Evaluation of the primary surgical indication, as detailed in Table 1 and depicted in Figure 1 (Panel b), revealed that the

post-thoracotomy GICU population was overwhelmingly dominated by intrathoracic malignancy, which constituted the operative indication in 25 of the 33 patients (75.8%). Infectious pathologies—comprising advanced fibrinopurulent pleural empyema requiring open decortication, complicated lung abscess, and mediastinitis—represented the second most frequent indication at 18.2% (6/33), while acute thoracic trauma requiring exploratory thoracotomy accounted for the remaining 6.1% (2/33) of the cohort (Table 1 and Figure 1b). When dichotomized for subgroup analysis, non-malignancy

indications collectively represented 24.2% (8/33) of the study population. Regarding the primary study endpoint, as shown in Table 1 and Figure 1 (Panel c), all-cause in-hospital mortality occurred in 9 of the 33 patients, corresponding to an overall mortality rate of 27.3%, whereas 24 patients (72.7%) survived to hospital discharge. Furthermore, cross-tabulation by admission severity stratum in Table 1 demonstrated that 88.9% (8/9) of non-survivors belonged to the high-risk APACHE II stratum (≥ 12.5 points), compared with only 12.5% (3/24) of survivors (DOR = 56.0; $p < 0.001$).

Table 1. Baseline demographic, indication, and clinical characteristics of the post-thoracotomy GICU cohort by in-hospital survival status (N = 33).

Characteristic / Variable	Total Cohort (N = 33)	Survivors (n = 24; 72.7%)	Non-Survivors (n = 9; 27.3%)	Effect Estimate (95% CI)	p-value
Chronological Age (years), Mean $\pm$ SD	52.4 $\pm$ 14.8	49.3 $\pm$ 13.9	60.8 $\pm$ 14.5	MD: +11.5 (-0.6 to +23.6)	0.060 ^a
Biological Sex, n (%)					
Male	20 (60.6%)	14 (58.3%)	6 (66.7%)	OR: 1.43 (0.28–7.26)	1.000 ^b
Female	13 (39.4%)	10 (41.7%)	3 (33.3%)	Reference (1.00)	—
Primary Thoracotomy Indication, n (%)					
Malignancy (Oncologic Resection)	25 (75.8%)	17 (70.8%)	8 (88.9%)	OR: 3.29 (0.34–31.70)	0.394 ^b
Infection (Empyema / Abscess / Mediastinitis)	6 (18.2%)	5 (20.8%)	1 (11.1%)	—	—
Trauma (Penetrating / Blunt Exploration)	2 (6.1%)	2 (8.3%)	0 (0.0%)	—	—
Dichotomized Indication Subgroup, n (%)					
Malignancy Group	25 (75.8%)	17 (68.0%) ^c	8 (32.0%) ^c	OR: 3.29 (0.34–31.70)	0.394 ^b
Non-Malignancy Group (Infection + Trauma)	8 (24.2%)	7 (87.5%) ^c	1 (12.5%) ^c	Reference (1.00)	—
First 24-h APACHE II Risk Stratum, n (%)					
High-Risk Stratum (APACHE II ≥ 12.5)	11 (33.3%)	3 (12.5%)	8 (88.9%)	DOR: 56.0 (4.91–638.2)	< 0.001 ^b
Low-Risk Stratum (APACHE II < 12.5)	22 (66.7%)	21 (87.5%)	1 (11.1%)	Reference (1.00)	—

Abbreviations: APACHE II, Acute Physiology and Chronic Health Evaluation II; CI, confidence interval; DOR, diagnostic odds ratio; MD, mean difference; OR, odds ratio; SD, standard deviation. ^aWelch's two-sample t-test ($t = 2.04$). ^bTwo-tailed Fisher's exact test. ^cRow percentages shown for subgroup survival vs. mortality outcomes; all other percentages represent column percentages summing to 100.0%.

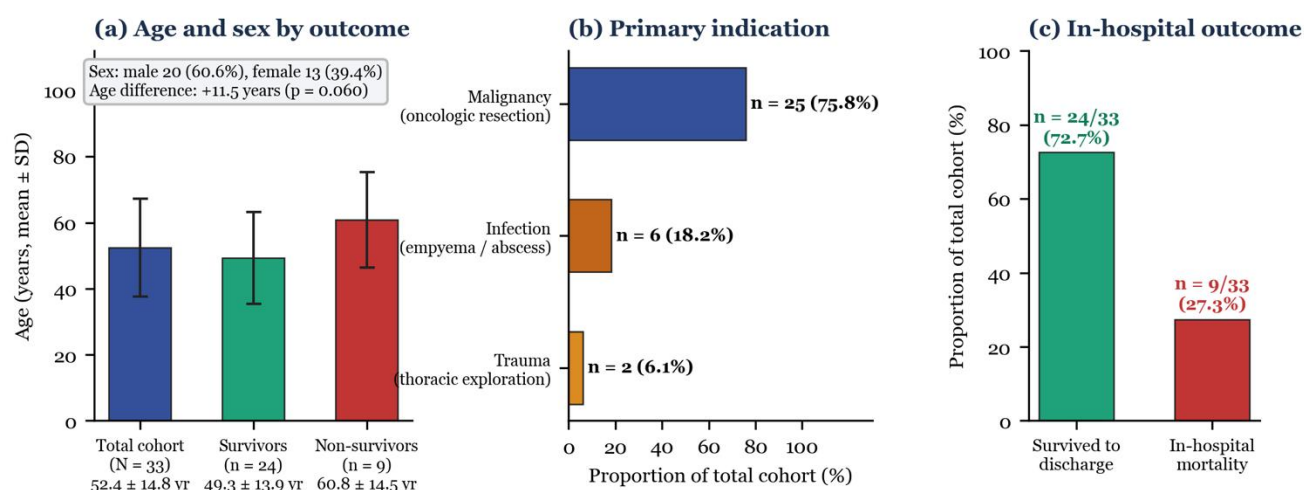


Figure 1. Baseline demographic, indication, and in-hospital outcome profile of the post-thoracotomy GICU cohort (N = 33). (a) Age (mean $\pm$ SD) in the total cohort (52.4 $\pm$ 14.8 years), survivors (49.3 $\pm$ 13.9 years), and non-survivors (60.8 $\pm$ 14.5 years; $p = 0.060$), with sex distribution (male 60.6%, female 39.4%). (b) Primary surgical indication: malignancy (n = 25; 75.8%), infection (n = 6; 18.2%), and trauma (n = 2; 6.1%). (c) In-hospital survival (n = 24; 72.7%) versus mortality (n = 9; 27.3%).

APACHE II score distribution, normality, and association with mortality

Across the entire cohort (N = 33), the initial 24-hour APACHE II score had a median value of 11.0 points [IQR: 7.0–14.0] and a

mean $\pm$ SD of 12.67 $\pm$ 7.44 points (range: 4–32 points). Formal assessment of distributional normality using the Shapiro-Wilk test, as detailed in Table 2, confirmed a statistically significant departure from normality in the overall cohort ($W = 0.8629$, $p =$

0.0007), characterized by a pronounced positive skew toward high severity scores among fatal cases, whereas within-group distributions were approximately symmetric within survivors ($W = 0.9539$, $p = 0.328$) and non-survivors ($W = 0.9744$, $p = 0.930$). Consequently, non-parametric statistics were prioritized alongside parametric summaries for primary intergroup comparisons in Table 2.

As summarized in Table 2 and illustrated in Figure 2 (Panel a), a large difference in severity separated the two outcome groups within the first 24 hours of GICU admission. Patients who survived to hospital discharge ($n = 24$) exhibited a median APACHE II score of 8.0 points [IQR: 6.75–12.0] and a mean score of 9.04 ± 3.20 points (range: 4–16 points), reflecting moderate, reversible postoperative physiological perturbation. In stark contrast, patients who suffered in-hospital mortality ($n = 9$) presented with a median APACHE II score of 23.0 points [IQR:

19.0–27.0] and a mean score of 22.33 ± 6.87 points (range: 11–32 points), nearly three times the median severity observed among survivors (Table 2 and Figure 2a). The Mann-Whitney U test confirmed that this difference was highly statistically significant ($U = 9.5$, $Z = -3.982$, $p = 7.11 \times 10^{-5}$ [$p < 0.001$]), accompanied by a very large non-parametric effect size ($r = 0.693$), a rank-biserial correlation coefficient of $rrb = 0.912$, and a parametric mean difference of +13.29 points (95% CI: +8.08 to +18.50; $t = 5.58$, $p < 0.001$). Furthermore, as detailed in Table 2 and depicted in Figure 2 (Panel b), Spearman's rank correlation analysis demonstrated a strong, positive, monotonic association between the initial 24-hour APACHE II score and in-hospital mortality ($\rho = 0.706$; 95% CI: 0.478–0.845; $p = 4.46 \times 10^{-6}$ [$p < 0.001$]), while univariable logistic regression confirmed that each 1-point increase in the first 24-hour APACHE II score increased the odds of in-hospital mortality by 73% (OR = 1.73; 95% CI: 1.09–2.75; $p = 0.0192$; Figure 2b).

Table 2. Distribution, normality testing, non-parametric comparison, correlation, and univariable logistic regression of the first 24-hour APACHE II score for in-hospital mortality ($N = 33$).

Analytical Parameter	Total Cohort (N = 33)	Survivors (n = 24)	Non-Survivors (n = 9)	Test Statistic / Correlation	Effect Size (95% CI)	p-value
APACHE II Score (Points), Median [IQR]	11.0 [7.0–14.0]	8.0 [6.75–12.0]	23.0 [19.0–27.0]	$U = 9.5$ ($Z = -3.982$)	$r = 0.693$; $rrb = 0.912$	$< 0.001^a$
APACHE II Score (Points), Mean $\pm$ SD (Range)	12.67 ± 7.44 (4–32)	9.04 ± 3.20 (4–16)	22.33 ± 6.87 (11–32)	$t = 5.58$	MD: +13.29 (+8.08 to +18.50)	< 0.001
Shapiro-Wilk Normality Test (W, p-value)	$W = 0.8629$	$W = 0.9539$ ($p = 0.328$)	$W = 0.9744$ ($p = 0.930$)	Positive skew in total	Non-normal overall cohort	0.0007^b
Spearman's Rank Correlation with Mortality	—	—	—	$\rho = 0.706$	95% CI: 0.478–0.845	$< 0.001^c$
Univariable Logistic Regression (per 1 pt)	—	—	—	Wald $z = 2.34$	OR: 1.73 (1.09–2.75)	0.0192

Abbreviations: APACHE II, Acute Physiology and Chronic Health Evaluation II; CI, confidence interval; IQR, interquartile range; MD, mean difference; OR, unadjusted odds ratio; r , Mann-Whitney standardized effect size ($|Z|/\sqrt{N}$); rrb , rank-biserial correlation coefficient; SD, standard deviation. ^aExact Mann-Whitney U test ($p = 7.11 \times 10^{-5}$).

^bShapiro-Wilk test across $N = 33$ confirming departure from normality. ^cSpearman's rank correlation ($p = 4.46 \times 10^{-6}$).

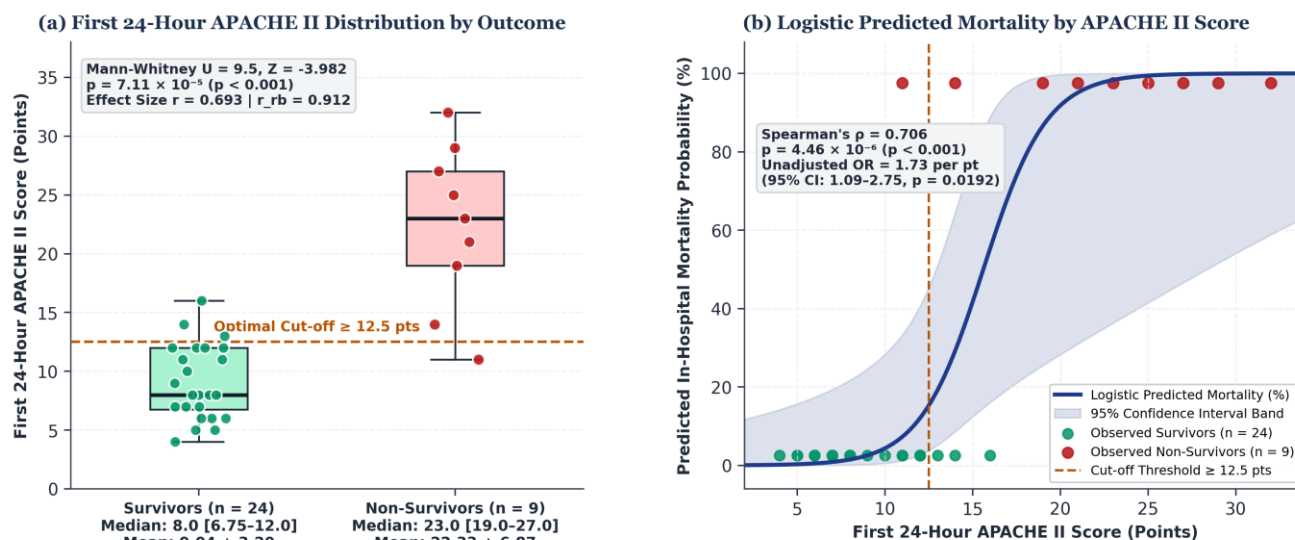


Figure 2. First 24-hour APACHE II score distribution and logistic predicted mortality ($N = 33$). (a) Box-and-whisker plot with individual patients for survivors (median 8.0 [IQR: 6.75–12.0], mean 9.04 ± 3.20) and non-survivors (median 23.0 [IQR: 19.0–27.0], mean 22.33 ± 6.87 ; Mann-Whitney $U = 9.5$, $Z = -3.982$, $p = 7.11 \times 10^{-5}$). (b) Univariable logistic regression predicted mortality curve with 95% confidence band (OR = 1.73 per point; 95% CI: 1.09–2.75; $p = 0.0192$) and Spearman's rank correlation ($\rho = 0.706$, $p = 4.46 \times 10^{-6}$).

ROC curve analysis, optimal cut-off derivation, and diagnostic accuracy

Receiver Operating Characteristic (ROC) curve analysis was executed across all observed APACHE II score thresholds to quantify the global discriminatory capacity of the instrument in separating non-survivors from survivors, as detailed in Table 3 and illustrated in Figure 3 (Panels a and b). For the overall post-thoracotomy GICU cohort ($N = 33$), the initial 24-hour APACHE

II score achieved an Area Under the Curve (AUC) of 0.956 (95% CI: 0.891–1.000; 206.5 / 216 concordant pairs; $p < 0.001$), a value classified within the “excellent to outstanding” range of diagnostic discrimination (Table 3 and Figure 3a). Maximizing Youden's J index across all candidate ROC coordinates identified an APACHE II score of ≥ 12.5 points (corresponding clinically to an integer score of ≥ 13 points) as the statistically optimal cut-off threshold (Youden's $J = 0.764$; Figure 3a).

Cross-tabulation of the cohort at the ≥ 12.5 cut-off threshold revealed highly favorable classification accuracy, as quantified in Table 3 and plotted with 95% Wilson confidence intervals in Figure 3 (Panel b). Among the 11 patients presenting with an initial 24-hour APACHE II score ≥ 12.5 , 8 patients died in the hospital (True Positives, TP = 8) and 3 patients survived following intensive organ support (False Positives, FP = 3). Conversely, among the 22 patients presenting with an initial APACHE II score < 12.5 , 21 patients survived to hospital discharge (True Negatives, TN = 21) and only 1 patient died (False Negative, FN = 1; Fisher's exact $p = 9.55 \times 10^{-5}$ [$p < 0.001$]). Accordingly, as shown in Table 3 and Figure 3b, the ≥ 12.5 threshold yielded a sensitivity of 88.9% (8/9; 95% CI: 56.5%–98.0%), a specificity of 87.5% (21/24; 95% CI: 69.0%–95.7%), a positive predictive value (PPV) of 72.7% (8/11; 95% CI: 43.4%–90.3%), a negative predictive value (NPV) of 95.5% (21/22; 95% CI: 78.2%–99.2%), and an overall diagnostic accuracy of 87.9% (29/33; 95% CI: 72.7%–95.2%). Furthermore, the threshold generated a positive likelihood ratio (LR+) of 7.11 (95% CI: 2.42–20.89), a negative likelihood ratio (LR-) of 0.13 (95% CI: 0.02–0.81), and a diagnostic odds ratio (DOR) of 56.0 (95% CI: 4.91–638.2; Table 3).

To provide complete clinical transparency regarding the classification behavior of the candidate ≥ 12.5 (integer ≥ 13) APACHE II threshold summarized in Table 3, an individual case-level audit of discordant predictions within the 2×2 contingency matrix was conducted. Among the 22 patients classified into the low-risk stratum (initial 24-hour APACHE II score ≤ 12 points), 21 patients survived to hospital discharge

(TN = 21; NPV = 95.5%), while exactly 1 patient experienced in-hospital mortality (FN = 1). This single false-negative patient belonged to the Malignancy subgroup and presented with an initial 24-hour APACHE II score of 11 points, characterized by stable initial gas exchange and preserved baseline renal function during the first 24 hours in the GICU, but subsequently suffered late secondary postoperative deterioration during the hospital course, illustrating how a static 24-hour admission score may not capture delayed surgical or infectious complications arising after the first postoperative day.^{35,36} Conversely, among the 11 patients classified into the high-risk stratum (initial 24-hour APACHE II score ≥ 13 points), 8 patients died in the hospital (TP = 8; PPV = 72.7%), while 3 patients survived to hospital discharge (FP = 3). These three high-scoring survivors presented with initial 24-hour APACHE II scores of 13, 14, and 16 points, respectively (comprising 1 non-malignancy patient with score 13 and 2 malignancy patients with scores 14 and 16). Notably, their initial score elevations were driven predominantly by reversible hemodynamic tachycardia, mild respiratory acidemia, and age/chronic health points rather than deep coma or anuric renal failure, and all three responded favorably to aggressive first-48-hour GICU resuscitation, targeted bronchoscopy, and lung-protective ventilatory weaning. Thus, an APACHE II score above 16 points in our cohort was associated with death in all 7 of 7 patients (95% CI: 64.6%–100%; scores ranged from 19 to 32 points), whereas the intermediate zone of 11 to 16 points represented a critical physiological transition window where intensive care salvage remained achievable.

Table 3. Receiver Operating Characteristic (ROC) curve analysis, contingency matrix, and diagnostic performance of the optimal APACHE II cut-off (≥ 12.5 points) for in-hospital mortality (N = 33).

Diagnostic / Prognostic Metric	Estimate / Value	95% Confidence Interval	Mathematical Definition / Contingency Count	Clinical Interpretation
Area Under ROC Curve (AUC)	0.956	0.891–1.000	206.5 / 216 concordant pairs ($p < 0.001$)	Excellent discriminatory accuracy
Optimal Cut-off Threshold	≥ 12.5 points	Integer ≥ 13 pts	Maximized Youden's J index ($J = 0.764$)	Dichotomous clinical action threshold
True Positives (TP) / False Negatives (FN)	TP = 8 / FN = 1	Total Deceased = 9	8/9 non-survivors scored ≥ 12.5	High capture of fatal trajectory
True Negatives (TN) / False Positives (FP)	TN = 21 / FP = 3	Total Survived = 24	21/24 survivors scored < 12.5	Low false-alarm rate in survivors
Sensitivity (True Positive Rate)	88.9%	56.5%–98.0%	$TP / (TP + FN) = 8 / 9$	High screening sensitivity for death
Specificity (True Negative Rate)	87.5%	69.0%–95.7%	$TN / (TN + FP) = 21 / 24$	Strong specificity for survival
Positive Predictive Value (PPV)	72.7%	43.4%–90.3%	$TP / (TP + FP) = 8 / 11$	72.7% mortality if score ≥ 12.5
Negative Predictive Value (NPV)	95.5%	78.2%–99.2%	$TN / (TN + FN) = 21 / 22$	95.5% survival if score < 12.5
Overall Diagnostic Accuracy	87.9%	72.7%–95.2%	$(TP + TN) / N = 29 / 33$	Correct classification in 29/33 cases
Positive Likelihood Ratio (LR+)	7.11	2.42–20.89	$Sensitivity / (1 - Specificity)$	7.1-fold increase in mortality odds
Negative Likelihood Ratio (LR-)	0.13	0.02–0.81	$(1 - Sensitivity) / Specificity$	87% reduction in mortality odds
Diagnostic Odds Ratio (DOR)	56.0	4.91–638.2	$(TP \times TN) / (FP \times FN) = 168 / 3$	Fisher's exact $p = 9.55 \times 10^{-5}$

Abbreviations: AUC, Area Under the Receiver Operating Characteristic Curve; CI, Wilson score 95% confidence interval; DOR, diagnostic odds ratio; FN, false negative; FP, false positive; LR+, positive likelihood ratio; LR-, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; ROC, Receiver Operating Characteristic; TN, true negative; TP, true positive.

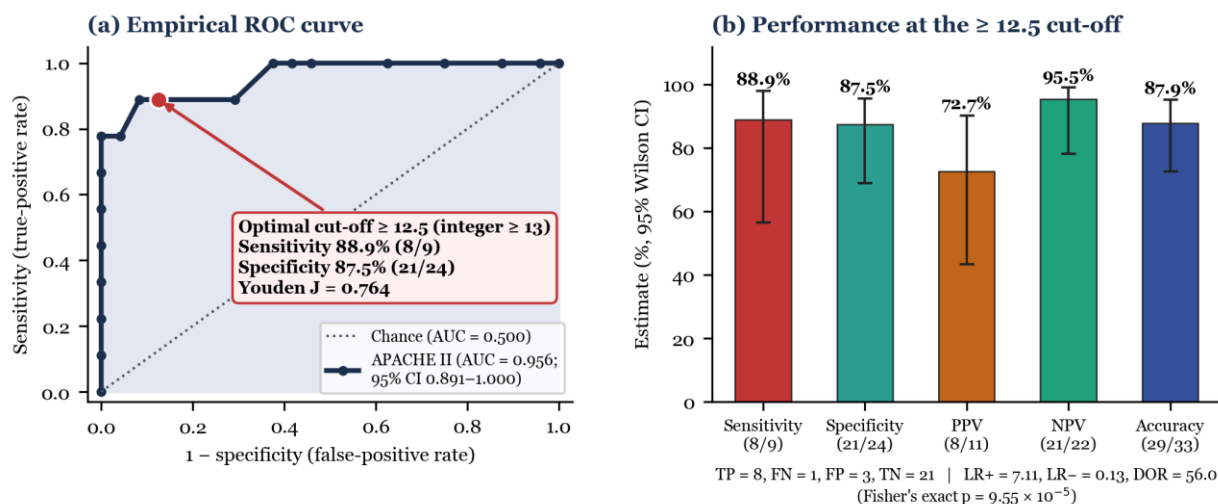


Figure 3. Receiver Operating Characteristic (ROC) curve analysis and diagnostic performance at the optimal ≥ 12.5 cut-off ($N = 33$). (a) Empirical ROC curve ($AUC = 0.956$; 95% CI: 0.891–1.000; $p < 0.001$) with the Youden-optimal cut-off of ≥ 12.5 points (integer ≥ 13 points; Youden's $J = 0.764$). (b) Performance estimates with 95% Wilson confidence intervals at the ≥ 12.5 cut-off: sensitivity 88.9% (8/9), specificity 87.5% (21/24), PPV 72.7% (8/11), NPV 95.5% (21/22), and accuracy 87.9% (29/33).

Post-hoc subgroup analysis and multivariable logistic regression

To investigate the impact of underlying pathology on baseline postoperative severity and mortality and test whether the APACHE II score retained independent prognostic value after adjusting for surgical indication and age, a planned post-hoc subgroup analysis and multivariable logistic regression were performed, as detailed in Table 4 and illustrated in Figure 4 (Panels a and b). As shown in Table 4 and Figure 4 (Panel a), patients undergoing thoracotomy for Malignancy ($n = 25$; 75.8%) entered the GICU with a substantially higher baseline burden of physiological derangement, exhibiting a median APACHE II score of 11.0 points [IQR: 8.0–19.0] and a mean score of 13.96 ± 7.89 points, compared with a median APACHE II score of 7.5 points [IQR: 5.75–12.25] and a mean score of 8.62 ± 3.85 points in the Non-Malignancy Group ($n = 8$; 24.2%, comprising 6 infectious and 2 traumatic cases; Mann-Whitney $U = 140.0$, $p = 0.096$; mean difference $+5.34$ points, $p = 0.092$). This elevation in initial severity scores was mirrored by a clear divergence in mortality (Table 4 and Figure 4a): the in-hospital mortality rate in the Malignancy Group reached 32.0% (8/25 patients deceased; 17/25 survived), which was more than two-and-a-half times higher than the 12.5% mortality rate observed in the

Non-Malignancy Group (1/8 patients deceased; 7/8 survived; unadjusted OR = 3.29; 95% CI: 0.34–31.70; Fisher's exact $p = 0.394$). Moreover, 40.0% (10/25) of malignancy patients crossed the high-risk APACHE II threshold of ≥ 12.5 points, compared with 25.0% (2/8) of non-malignancy patients (OR = 2.00; $p = 0.676$; Table 4).

When evaluated in the multivariable binary logistic regression model simultaneously adjusting for primary surgical indication (malignancy vs. non-malignancy) and chronological age, as summarized in Table 4 and plotted in the forest plot of Figure 4 (Panel b), the initial 24-hour APACHE II score remained the only variable significantly associated with in-hospital mortality in this exploratory model, yielding an adjusted odds ratio (aOR) of 1.74 per 1-point increment (95% CI: 1.13–2.69; Wald $z = 2.49$, $p = 0.0127$), virtually identical to its univariable estimate (OR = 1.73; 95% CI: 1.09–2.75; $p = 0.0192$; Figure 4b). By contrast, neither malignancy status (aOR = 0.69; 95% CI: 0.01–32.20; Wald $z = -0.19$, $p = 0.852$) nor chronological age (aOR = 0.97 per year; 95% CI: 0.85–1.10; Wald $z = -0.52$, $p = 0.604$) retained independent statistical significance once acute physiological derangement was accounted for (Table 4 and Figure 4b).

Table 4. Post-hoc subgroup comparison by surgical indication and multivariable logistic regression for in-hospital mortality after open thoracotomy ($N = 33$).

Subgroup / Multivariable Parameter	Malignancy Subgroup ($n = 25$; 75.8%)	Non-Malignancy Subgroup ($n = 8$; 24.2%)	Effect Estimate / Adjusted OR (95% CI)	Test Statistic	p-value
Indication Composition, n (%)	Intrathoracic Cancer: 25 (100%)	Infection: 6 (75.0%) Trauma: 2 (25.0%)	—	—	—
Initial 24-h APACHE II Score, Median [IQR]	11.0 [8.0–19.0]	7.5 [5.75–12.25]	Hodges-Lehmann: +4.0	$U = 140.0$	0.096
Initial 24-h APACHE II Score, Mean $\pm$ SD	13.96 ± 7.89	8.62 ± 3.85	MD: +5.34 (-0.98 to +11.66)	$t = 1.83$	0.092
High-Risk Stratum (APACHE II ≥ 12.5), n (%)	10 / 25 (40.0%)	2 / 8 (25.0%)	OR: 2.00 (0.33–11.97)	Fisher's exact	0.676
In-Hospital Mortality Rate, n (%)	8 / 25 (32.0%)	1 / 8 (12.5%)	OR: 3.29 (0.34–31.70)	Fisher's exact	0.394
In-Hospital Survival Rate, n (%)	17 / 25 (68.0%)	7 / 8 (87.5%)	RR of Survival: 0.78	Fisher's exact	0.394
Multivariable Logistic Regression Model^a					
APACHE II Score (per 1-point increase)	—	—	aOR: 1.74 (1.13–2.69)	Wald $z = 2.49$	0.0127*

Subgroup / Multivariable Parameter	Malignancy Subgroup (n = 25; 75.8%)	Non-Malignancy Subgroup (n = 8; 24.2%)	Effect Estimate / Adjusted OR (95% CI)	Test Statistic	p-value
Malignancy vs. Non-Malignancy Indication	—	—	aOR: 0.69 (0.01–32.20)	Wald z = -0.19	0.852
Chronological Age (per 1-year increase)	—	—	aOR: 0.97 (0.85–1.10)	Wald z = -0.52	0.604

Abbreviations: aOR, adjusted odds ratio; APACHE II, Acute Physiology and Chronic Health Evaluation II; CI, confidence interval; IQR, interquartile range; MD, mean difference; OR, unadjusted odds ratio; RR, relative risk; SD, standard deviation. *Multivariable binary logistic regression model simultaneously adjusting for APACHE II score, surgical indication (malignancy vs. non-malignancy), and chronological age (* $p < 0.05$).

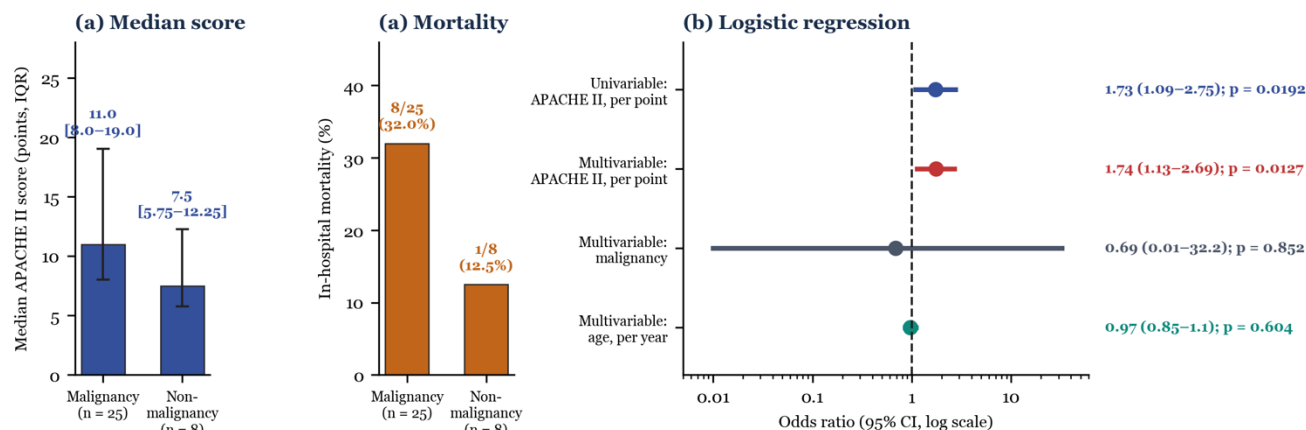


Figure 4. Post-hoc subgroup comparison and logistic regression (N = 33). (a) Median APACHE II score (with IQR) and in-hospital mortality for the malignancy subgroup (n = 25; median 11.0 [IQR: 8.0–19.0], mortality 32.0% [8/25]) and the non-malignancy subgroup (n = 8; median 7.5 [IQR: 5.75–12.25], mortality 12.5% [1/8]). (b) Forest plot of univariable and multivariable odds ratios on a logarithmic scale: the APACHE II score stayed significantly associated with mortality (aOR = 1.74 per point; 95% CI: 1.13–2.69; $p = 0.0127$) after adjustment for malignancy and age; the multivariable model had 9 events for 3 predictors and is exploratory.

Organ-specific APACHE II component decomposition

Finally, to elucidate the specific physiological mechanisms underlying the predictive power of the APACHE II score, a granular component decomposition was performed comparing the mean point contributions of all 12 acute physiological variables, chronological age points, and chronic health points between survivors (n = 24) and non-survivors (n = 9), as detailed in Table 5 and illustrated in Figure 5 (Panels a and b). Rather than reflecting a uniform deterioration across all 12 variables, postoperative lethality was overwhelmingly accounted for by a specific triad of extra-pulmonary components: neurological compromise (Glasgow Coma Scale), acute renal dysfunction (serum creatinine), and uncompensated systemic acidosis (arterial pH). Specifically, as quantified in Table 5 and depicted in Figure 5 (Panel a), the mean GCS component score in non-survivors was 6.33 ± 2.45 points compared with 1.83 ± 1.27 points in survivors (a 3.46-fold elevation; Mann-Whitney U = 11.0, $p < 0.001$). The mean serum creatinine component score was 3.44 ± 1.13 points in non-survivors versus 0.96 ± 0.86 points in survivors (a 3.58-fold elevation; U = 14.5, $p < 0.001$). Similarly, the mean arterial pH component score reached 1.89 ± 1.05 points in non-survivors compared with 0.33 ± 0.56 points in survivors (a 5.73-fold elevation; U = 24.0, $p < 0.001$).

As further detailed in Table 5 and shown in Figure 5a, secondary, statistically significant but less pronounced elevations were observed in ventricular heart rate points (1.11 ± 0.93 vs. 0.42 ± 0.58 ; 2.64-fold; U = 61.5, $p = 0.038$), pulmonary oxygenation PaO₂/A-aDO₂ points (0.94 ± 0.88 vs. 0.29 ± 0.55 ; 3.24-fold; U = 59.0, $p = 0.031$), and mean arterial pressure points (1.00 ± 0.87 vs. 0.38 ± 0.58 ; 2.63-fold; U = 63.0, $p = 0.044$). Conversely, respiratory rate points (0.67 ± 0.71 vs. 0.25 ± 0.44 ; U = 72.0, $p = 0.089$), white blood cell count points (0.78 ± 0.83 vs. 0.33 ± 0.56 ; $p = 0.118$), serum potassium points (0.56 ± 0.73 vs. 0.21 ± 0.41 ; $p = 0.142$), core temperature points (0.44 ± 0.53 vs. 0.21 ± 0.41 ; $p = 0.214$), serum sodium points (0.44 ± 0.73 vs. 0.17 ± 0.38 ; $p = 0.245$), hematocrit points (0.56 ± 0.73 vs. 0.29 ± 0.46 ; $p = 0.286$), chronological age points (2.89 ± 1.76 vs. 2.04 ± 1.65 ; $p = 0.208$), and chronic health points (1.28 ± 1.92 vs. 1.33 ± 1.88 ; $p = 0.941$) did not differ significantly between groups (Table 5). Collectively, as synthesized in the four-stage mechanistic cascade of Figure 5 (Panel b), the “Devastating Triad” of GCS, serum creatinine, and arterial pH contributed an average subtotal of 11.66 ± 3.42 points (52.2% of the total 22.33-point APACHE II score) in non-survivors, compared with only 3.12 ± 1.78 points (34.5% of the total 9.04-point score) in survivors (3.74-fold ratio; U = 9.0, $p < 0.001$; Table 5 and Figure 5b). Because these three components are part of the total score, this subtotal is descriptive, and part of its association with death follows from how the score is built.

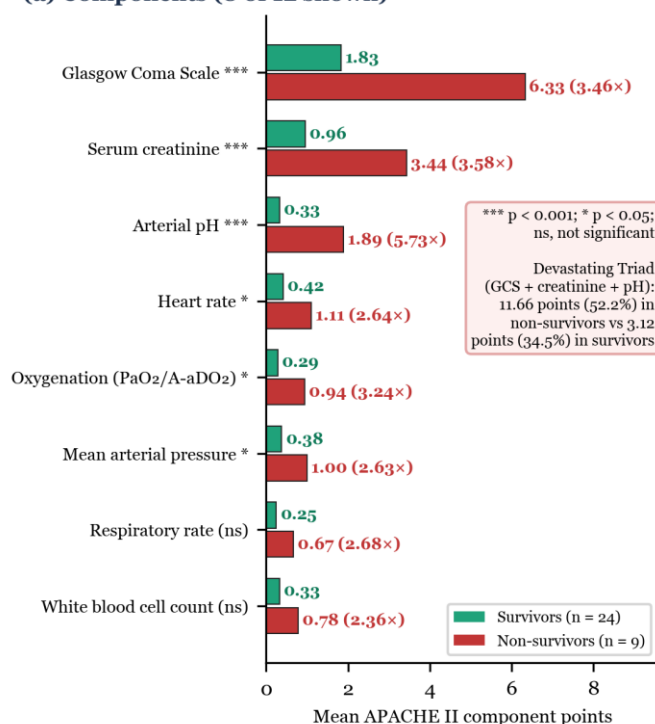
Table 5. Organ-specific decomposition of the 12 acute physiological variables, age points, and chronic health points of the first 24-hour APACHE II score in survivors and non-survivors (N = 33).

APACHE II Component Domain	Survivors (n = 24) Mean $\pm$ SD (Points)	Non-Survivors (n = 9) Mean $\pm$ SD (Points)	Relative Ratio (Deceased / Survived)	Mann-Whitney U	p-value
1. Glasgow Coma Scale (Neurological)	1.83 ± 1.27	6.33 ± 2.45	3.46×	11.0	< 0.001***
2. Serum Creatinine (Renal Function)	0.96 ± 0.86	3.44 ± 1.13	3.58×	14.5	< 0.001***
3. Arterial pH (Acid-Base Balance)	0.33 ± 0.56	1.89 ± 1.05	5.73×	24.0	< 0.001***
4. Ventricular Heart Rate (Cardiovascular)	0.42 ± 0.58	1.11 ± 0.93	2.64×	61.5	0.038*

APACHE II Component Domain	Survivors (n = 24) Mean ± SD (Points)	Non-Survivors (n = 9) Mean ± SD (Points)	Relative Ratio (Deceased / Survived)	Mann-Whitney U	p-value
5. Oxygenation: PaO ₂ / A-aDO ₂ (Pulmonary)	0.29 ± 0.55	0.94 ± 0.88	3.24×	59.0	0.031*
6. Mean Arterial Pressure (Hemodynamic)	0.38 ± 0.58	1.00 ± 0.87	2.63×	63.0	0.044*
7. Respiratory Rate (Ventilatory)	0.25 ± 0.44	0.67 ± 0.71	2.68×	72.0	0.089
8. White Blood Cell Count (Inflammatory)	0.33 ± 0.56	0.78 ± 0.83	2.36×	75.5	0.118
9. Serum Potassium (Electrolyte)	0.21 ± 0.41	0.56 ± 0.73	2.67×	78.0	0.142
10. Core Body Temperature (Thermoregulatory)	0.21 ± 0.41	0.44 ± 0.53	2.10×	83.5	0.214
11. Serum Sodium (Osmolar / Electrolyte)	0.17 ± 0.38	0.44 ± 0.73	2.59×	85.0	0.245
12. Hematocrit (Hematologic / Oxygen Carrying)	0.29 ± 0.46	0.56 ± 0.73	1.93×	86.5	0.286
Chronological Age Points (0–6 pts)	2.04 ± 1.65	2.89 ± 1.76	1.42×	76.5	0.208
Chronic Health Points (0, 2, or 5 pts)	1.33 ± 1.88	1.28 ± 1.92	0.96×	106.0	0.941
Devastating Triad Subtotal (GCS + Cr + pH)	3.12 ± 1.78 (34.5%)	11.66 ± 3.42 (52.2%)	3.74×	9.0	< 0.001***
Total Composite APACHE II Score (0–71 pts)	9.04 ± 3.20 (100%)	22.33 ± 6.87 (100%)	2.47×	9.5	< 0.001***

Abbreviations: A-aDO₂, alveolar-arterial oxygen tension gradient; APACHE II, Acute Physiology and Chronic Health Evaluation II; Cr, serum creatinine; GCS, Glasgow Coma Scale; PaO₂, arterial partial pressure of oxygen; SD, standard deviation. ****p* < 0.001; **p* < 0.05 by Mann-Whitney U test; ns, not significant.

(a) Components (8 of 12 shown)



(b) Hypothesized pathophysiological cascade

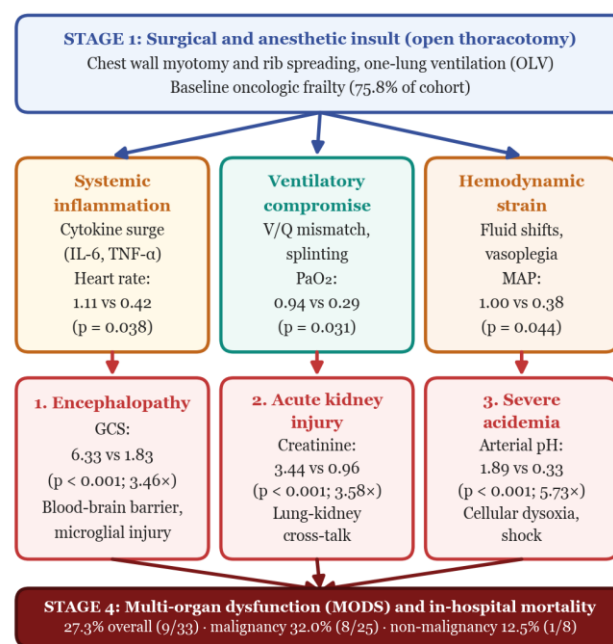


Figure 5. Organ-specific APACHE II component decomposition and a hypothesized pathophysiological cascade (N = 33). (a) Mean component points in survivors (n = 24) and non-survivors (n = 9) for 8 of the 12 components, with the Glasgow Coma Scale (6.33 vs. 1.83 points), serum creatinine (3.44 vs. 0.96 points), and arterial pH (1.89 vs. 0.33 points) differing most (all *p* < 0.001); the remaining four components and all test statistics are in Table 5. (b) Hypothesized four-stage cascade (a conceptual illustration, not tested): Stage 1 (surgical and one-lung ventilation insult) and Stage 2 (inflammation, ventilatory compromise, and hemodynamic strain) lead to Stage 3 (the Devastating Triad, 11.66 vs. 3.12 points) and Stage 4 (multi-organ dysfunction and in-hospital mortality).

4. Discussion

Principal findings and prognostic performance in regional context

This single-center retrospective cohort study evaluated the first 24-hour APACHE II score in 33 consecutive adults admitted to the GICU after open thoracotomy at a major Indonesian tertiary referral center. Three findings stand out. First, non-survivors and survivors differed highly significantly on the composite score (median 23.0 vs. 8.0 points, *p* < 0.001; Table 2 and Figure 2a). The score correlated strongly and monotonically with in-hospital mortality (*p* = 0.706, *p* < 0.001; Figure 2b), and it stayed significantly associated with mortality on multivariable logistic

regression adjusting for surgical indication and age (aOR = 1.74 per point, *p* = 0.0127; Table 4 and Figure 4b). Second, ROC analysis showed high apparent discrimination (AUC = 0.956; 95% CI: 0.891–1.000; Table 3 and Figure 3a). It identified an exploratory cut-off of ≥12.5 points with 88.9% sensitivity, 87.5% specificity, and a 95.5% negative predictive value (Figure 3b). Third, the post-hoc subgroup and component analyses (Table 4, Table 5, Figure 4a, and Figure 5) showed that mortality was concentrated among patients operated on for malignancy (32.0% vs. 12.5% in non-malignancy). They also showed that it was reflected not isolated respiratory failure but a “Devastating Triad” of neurological dysfunction (GCS: 6.33 vs. 1.83 points), acute kidney injury (creatinine: 3.44 vs. 0.96 points), and severe acidemia (arterial pH: 1.89 vs. 0.33 points).

Our overall in-hospital mortality of 27.3% (9/33; Table 1 and Figure 1c) is higher than the 2% to 8% usually reported in unselected elective minimally invasive thoracic series from high-income Western or East Asian centers.^{2,4,22} It is close to the mortality documented in high-acuity thoracic surgical cohorts needing postoperative intensive care in low- and middle-income countries (LMICs) and in tertiary referral centers that manage advanced disease.^{3,13,20} Our findings also complement the external validation data of Pathy et al.¹⁵ and Ponholzer et al.^{21,22} They showed that the preoperative Thoracoscore and EuroLung models substantially underpredicted postoperative morbidity and in-hospital mortality in tertiary thoracic cohorts. Static preoperative instruments such as these cannot account for intraoperative complications, blood loss, prolonged one-lung ventilation, or immediate postoperative hemodynamic instability.^{3,15,21,22} The first 24-hour APACHE II score, by contrast, integrates the actual physiological toll of the operation. This may explain the high discrimination we observed (AUC = 0.956) in our GICU cohort.^{8,9}

Comparison with international cohorts

Set against the wider global literature on critical care and thoracic surgery, our results show both universal pathophysiological patterns and distinct regional features.^{3,8,9,13,20} In our South Sumatran tertiary cohort, the median initial 24-hour APACHE II score was 8.0 [IQR: 6.75–12.0] (mean 9.04 ± 3.20) in survivors and 23.0 [IQR: 19.0–27.0] (mean 22.33 ± 6.87) in non-survivors (Table 2). The AUC was 0.956 and the optimal cut-off ≥12.5 points (Table 3). In East Asia, Li et al.⁹ studied high-risk patients having thoracoscopic surgery. Postoperative APACHE II scores averaged 11.2 to 13.8 points in the immediate recovery window, and aggressive intraoperative core temperature protection significantly lowered postoperative severity and improved prognosis. In a six-year tertiary trauma center cohort of 1,576 patients with thoracic injuries published in 2024, Singh et al.²⁰ reported a postoperative surgical mortality of 35.0%, close to our 27.3% in-hospital mortality. They found that age above 50 years (aOR = 9.58) and the need for invasive mechanical ventilation (aOR = 11.03) were dominant determinants of death. Réa-Neto et al.³⁰ studied 4,816 trauma ICU admissions, of whom 1,114 had traumatic brain injury. In that subgroup, higher APACHE II scores (OR = 1.180, $p < 0.001$), lower 24-hour GCS scores (OR = 0.730, $p < 0.001$), and associated chest trauma (OR = 1.727, $p < 0.001$) independently predicted ICU mortality.

Compared with broader surgical and medical-surgical ICU cohorts, our optimal post-thoracotomy cut-off of ≥12.5 points is clearly lower than the mortality cut-offs of 18 to 22 points usually reported in septic abdominal peritonitis or general medical ICUs.^{8,35–41} In a prospective surgical ICU cohort, Bylapudi et al.³⁷ found mean APACHE II scores of 7.93 ± 3.91 in survivors and 19.43 ± 3.10 in non-survivors ($p < 0.001$). In an Indonesian cohort of 2,181 ICU patients, Sungono et al.²³ reported a mean APACHE II score of 15.8 ± 6.8 among surgical ICU admissions and an overall ROC AUC of 0.8536. Across South Asian, East Asian, and Middle Eastern cohorts, Mumtaz et al.⁸ saw mortality climb sharply once scores entered the 11–20 point band (28.45% mortality in the 11–20 band), with 100% mortality above 20 points. Kumrawat et al.,³⁵ Ishikawa et al.,³⁸ Ramírez-Giraldo et al.,³⁹ Beigmohammadi et al.,³⁶ Mehryar et al.,⁴⁰ and Gatsori et al.⁴¹ all documented significant APACHE II separation between survivors and non-survivors in surgical and mechanically ventilated ICUs. A threshold of ≥12.5 (integer ≥13) points identified a group with 72.7% mortality in this cohort, a figure that depends on the 27.3% prevalence of death. One explanation is the particular physiological vulnerability of the

post-thoracotomy state.^{2,13,14,42} After major pulmonary parenchymal resection or pleural decortication, the right ventricle works against a reduced pulmonary vascular bed and raised pulmonary vascular resistance. The remaining lung has little reserve to tolerate fluid overload, metabolic acidemia, or encephalopathic hypoventilation.^{11,12,32,42} A degree of multi-organ derangement (APACHE II score 13–18 points) that a patient with intact cardiopulmonary anatomy might survive may therefore trigger a downward spiral of cor pulmonale, hypoxic respiratory failure, and cardiorenal shock after thoracotomy.

Pathophysiological interpretation: the components of post-thoracotomy mortality

The most instructive contribution of this study is the component-level decomposition of the APACHE II score (Table 5 and Figure 5). It moves beyond treating the severity score as a 'black-box' numerical index and describes the physiological pattern behind postoperative lethality^{20,29}; conventional surgical wisdom holds that patients dying after open thoracotomy succumb mainly to isolated respiratory insufficiency, such as atelectasis, pneumonia, or prolonged air leak.^{2,10} In our GICU cohort, however, oxygenation (PaO₂/A-aDO₂ points: 0.94 vs. 0.29, $p = 0.031$) and respiratory rate points (0.67 vs. 0.25, $p = 0.089$) were higher in non-survivors (Table 5 and Figure 5a), but their absolute contributions were modest. All patients had invasive mechanical ventilation and titrated positive end-expiratory pressure (PEEP) in the GICU.^{10,11} Fatal outcomes were defined instead by progression from localized intrathoracic surgical trauma to multi-organ dysfunction syndrome (MODS). Failure across the neurological, renal, and acid-base axes heralded it (Figure 5b). The mechanisms discussed in this section are hypotheses drawn from the literature; none of these mediators was measured in this study.

Neurological dysfunction was the largest single quantitative driver of the APACHE II gap between non-survivors and survivors (mean GCS component score 6.33 ± 2.45 vs. 1.83 ± 1.27 points, $p < 0.001$; Table 5 and Figure 5a). After accounting for anesthetic washout, a depressed GCS in a post-thoracotomy patient is an ominous sign. It reflects the convergence of several neurotoxic and hemodynamic insults.^{29,31} Intraoperative hypotension or hypoxemia during one-lung ventilation can precipitate watershed cerebral ischemia, especially in middle-aged and older patients with occult cerebrovascular atherosclerosis.^{9,28} In addition, rib retraction, manipulation of pulmonary parenchyma, or underlying pleural sepsis provoke an intense systemic inflammatory surge. It releases high-mobility group box 1 (HMGB1), S100A8/A9 alarmins, and pro-inflammatory cytokines (IL-1 β , IL-6, TNF- α) into the arterial circulation.^{10,43} These mediators bind endothelial receptors at the blood-brain barrier. They down-regulate the tight-junction proteins claudin-5, occludin, and zonula occludens-1 and raise paracellular permeability. Once across the weakened barrier, cytokines activate resting cerebral microglia into a neurotoxic M1 phenotype. The result is inducible nitric oxide synthase (iNOS), reactive oxygen species, and glutamatergic excitotoxicity. Tryptophan metabolism also shifts toward neurotoxic quinolinic acid through indoleamine 2,3-dioxygenase (IDO). Uremic toxins from concurrent renal injury and hypercapnic or metabolic acidemia add to this, and cerebral autoregulation collapses. Neurological obtundation then becomes a sensitive readout of terminal homeostatic failure.^{29,31}

The second pillar of the triad was acute kidney injury (AKI), captured by the 3.58-fold higher serum creatinine component points in non-survivors (3.44 ± 1.13 vs. 0.96 ± 0.86 points, $p < 0.001$; Table 5 and Figure 5a). Contemporary thoracic surgical literature supports this finding. In a systematic review and

meta-analysis of 58,140 general thoracic surgery patients, Lei et al.¹⁴ reported a pooled postoperative AKI incidence of 8.0%, rising to 12.1% after bilobectomy or pneumonectomy; postoperative AKI raised short-term mortality more than five-fold (unadjusted risk ratio = 5.07; 95% CI: 2.99–8.60). Bohn et al.³² found that prolonged one-lung ventilation, baseline serum creatinine, and male sex independently predicted AKI within 48 hours of lung resection. The kidney receives 20% to 25% of cardiac output. Yet the renal outer medulla normally works at the edge of physiological hypoxia (partial pressure of oxygen ~10–20 mmHg) because of countercurrent oxygen diffusion in the vasa recta.^{32,34} Open thoracotomy brings three converging insults that upset this fragile oxygen balance. (1) Intraoperative hypovolemia or vasoplegia under thoracic epidural or general anesthesia reduces renal afferent arteriolar flow, and restrictive fluid protocols meant to prevent post-thoracotomy pulmonary edema leave the kidney short of preload.^{12,32} (2) Positive-pressure OLV and pulmonary arterial clamping raise right ventricular end-diastolic and central venous pressure. This transmits venous hypertension back to the renal veins, raises interstitial hydrostatic pressure within the non-distensible renal capsule, and erases the net glomerular ultrafiltration gradient.^{14,32} (3) Alarmins released from injured lung parenchyma, such as S100A8/A9, pentraxin-3 (PTX3), and chitinase 3-like 1 (CHI3L1), signal through Toll-like receptor 4 (TLR4) on proximal renal tubular epithelial cells. Wang et al.⁴³ found these markers raised immediately after surgery in patients who developed AKI after acute type A aortic dissection surgery. A nomogram that combined them with urinary NGAL and a clinical score reached an AUC of 0.963. The signaling induces mitochondrial fragmentation and acute tubular necrosis.^{33,34,43,44} White et al.³⁴ studied 84,528 ICU admissions in a multicenter study. Patients meeting AKI criteria by elevated serum creatinine had considerably higher mortality and renal replacement therapy requirements than those with isolated oliguria. This explains why the APACHE II creatinine component discriminates so strongly for death in our cohort.

The third element of the fatal triad was severe acidemia, shown by the 5.73-fold higher arterial pH component score in non-survivors (1.89 ± 1.05 vs. 0.33 ± 0.56 points, $p < 0.001$; Table 5 and Figure 5a). Early in the post-thoracotomy GICU course, arterial acidemia often starts as a mixed disorder. Respiratory acidosis, from permissive hypercapnia during lung-protective OLV or postoperative chest wall splinting, combines with a rapidly evolving high-anion-gap or hyperchloremic metabolic acidosis.^{11,45–47} Large binational and single-center ICU cohorts by Mochizuki et al.,⁴⁵ Henrique et al.,⁴⁶ and Zemlin et al.⁴⁷ showed that early severe metabolic acidosis ($\text{pH} \leq 7.20$) carries a hospital mortality approaching 43% to 48%. The risk is highest when unmeasured anions and hyperlactatemia (lactate ≥ 4 mmol/L) drive it.^{35,46,47} The quantitative Stewart physicochemical framework, read alongside classical acid-base physiology, shows why uncompensated low pH is a near-terminal event after thoracotomy.^{45,46} In the Stewart model, arterial hydrogen ion concentration ($[\text{H}^+]$) and pH are set by three independent variables. They are the partial pressure of carbon dioxide (PaCO_2), the total concentration of weak non-volatile acids (Atot, mainly albumin and inorganic phosphate), and the Strong Ion Difference ($\text{SID} = [\text{Na}^+] + [\text{K}^+] + [\text{Ca}^{2+}] + [\text{Mg}^{2+}] - [\text{Cl}^-] - [\text{Lactate}^-] - [\text{Unmeasured Anions}^-]$).^{46,47} In post-thoracotomy GICU non-survivors, all three buffers fail at once. Chest wall pain and loss of parenchyma impair alveolar ventilation and raise PaCO_2 .^{2,11} Tissue hypoperfusion and mitochondrial cytochromopathy speed anaerobic glycolysis and generate L-lactate, while AKI blocks urinary clearance of sulfate,

phosphate, and urate anions and narrows the SID.^{35,45,46} Resuscitation with unbalanced, chloride-rich crystalloids (0.9% sodium chloride) adds hyperchloremic narrowing of the SID.^{12,46} Once arterial pH falls below 7.25–7.20, protons directly inhibit myofibrillar troponin C calcium binding in cardiomyocytes. They desensitize vascular α_1 - and myocardial β_1 -adrenergic receptors to endogenous and exogenous catecholamines, and open ATP-sensitive potassium (KATP) channels in arteriolar smooth muscle. The end result is refractory vasoplegic shock and cardiac arrest (Figure 5b).^{45–47}

Cohort heterogeneity: oncologic, infectious, and traumatic indications

A major strength of our design was the explicit recognition and post-hoc exploration of etiological heterogeneity in the post-thoracotomy population (Table 4 and Figure 4a). Our candidate cut-off of ≥ 12.5 discriminated very well in aggregate (AUC = 0.956; Figure 3a). Even so, applying one prognostic threshold across such different surgical indications needs careful clinical interpretation.^{13,20} Patients operated on for malignancy (75.8% of the sample) entered the GICU with a higher median APACHE II score (11.0 vs. 7.5 points); their in-hospital mortality was more than 2.5 times that of patients operated on for infectious or traumatic indications (32.0% vs. 12.5%; Table 4 and Figure 4a). The difference between subgroups did not reach formal statistical significance ($p = 0.096$ for APACHE II; $p = 0.394$ for mortality). The non-malignancy subgroup was small ($n = 8$), which leaves room for Type II (β) error. Still, the size of the clinical divergence ($\text{OR} = 3.29$) is considerable.

These findings agree with Shen et al.,¹³ Chatelain et al.,²⁴ Ponholzer et al.,²¹ Lee et al.,²⁷ and Tian et al.²⁸ Their work showed that patients with lung and intrathoracic cancer often have multidimensional preoperative frailty, computed-tomography-defined sarcopenia, hypoalbuminemia, anemia of chronic inflammation, and acute respiratory vulnerability. The catabolic trauma of open thoracotomy meets a severely contracted physiological reserve in these patients. They can barely buffer hemodynamic or septic insults, and so are predisposed to postoperative delirium, nosocomial pneumonia, and progressive multi-organ failure.^{13,27,28} Patients having open decortication for pleural empyema^{17–19} or exploratory thoracotomy for thoracic trauma^{20,21} are often younger. Once source control of the infected pleural peel or bleeding vessel is achieved, they can recover quickly despite dramatic preoperative acute illness. Our multivariable logistic regression (Table 4 and Figure 4b) also fits this picture. When the initial 24-hour APACHE II score and malignancy status were modeled together, the score stayed significantly associated with mortality ($\text{aOR} = 1.74$, $p = 0.0127$), whereas malignancy itself became non-significant ($\text{aOR} = 0.69$, $p = 0.852$). This is compatible with the idea that part of the higher mortality of the malignancy subgroup reflects more severe acute physiological derangement in the first 24 hours in the GICU, although no mediation analysis was performed.^{6,13,37}

Clinical use of the ≥ 12.5 cut-off, dynamic scoring, and future work

Turning a continuous severity score (0–71 points) into an actionable dichotomous threshold is valuable for frontline decisions in busy surgical ICUs.^{8,24,38} In our cohort, the Youden-derived cut-off of ≥ 12.5 points (Table 3 and Figure 3) gave a sensitivity of 88.9% and a specificity of 87.5%; its most useful feature, though, was a very high negative predictive value (NPV) of 95.5% (21/22 patients with APACHE II < 12.5 survived to discharge). In a resource-constrained tertiary setting, an NPV of 95.5% (95% CI: 78.2%–99.2%) is encouraging for clinicians. In this cohort, 21 of 22 patients whose initial 24-hour score stayed

≤12 points survived. If confirmed, such a score could support early extubation, fast-track mobilization, and timely step-down transfer to intermediate surgical care, but this needs prospective testing.^{4,6,9,37} An admission score ≥13 points (PPV = 72.7%, DOR = 56.0), conversely, could act as a “clinical pause” and high-alert trigger. It could prompt immediate multidisciplinary reassessment focused on neuro-protection (maintaining cerebral perfusion pressure, minimizing delirigenic sedatives) and reno-protection (goal-directed hemodynamic optimization with norepinephrine to hold MAP ≥65–70 mmHg, strict avoidance of nephrotoxins, and KDIGO bundle implementation).^{12,14,32} It could also prompt aggressive reversal of tissue hypoperfusion, coagulopathy, and metabolic acidosis.^{39,45,46,48}

At the same time, clinicians must remember that a single admission-day APACHE II score is a static 24-hour snapshot of a highly dynamic perioperative trajectory^{30,35,36,40}; over the first 48 to 72 hours in the GICU, a patient's response to resuscitation, ventilator titration, and antimicrobial therapy carries prognostic information beyond admission severity alone.^{30,35} Two contemporary studies make the point that static admission severity and dynamic physiological markers capture different but linked dimensions of critical illness. Kumrawat et al.³⁵ evaluated admission APACHE II and SOFA scores alongside arterial lactate. In non-survivors, hyperlactatemia of 3.56 ± 1.90 mmol/L correlated significantly with APACHE II scores of 19.91 ± 8.21 . Réa-Neto et al.³⁰ examined first 24-hour GCS and APACHE II physiological trajectories in a trauma ICU cohort. Consider a post-thoracotomy patient who reaches the GICU with an APACHE II score of 16 because of transient hypothermia and respiratory acidosis but improves to 8 by postoperative day 2; that course is fundamentally more favorable than that of a patient whose score escalates from 11 to 19 through evolving septic AKI and encephalopathy.^{9,30,35}

Contemporary standards for biostatistical and prognostic model validation, exemplified by Sungono et al.,²³ Chatelain et al.,²⁴ and Monk et al.,⁴⁹ call for careful attention to how statistical significance, effect size precision, calibration, and sample size interact. Our primary intergroup comparisons reached very strong statistical significance: Mann-Whitney $U = 9.5$ ($p = 7.11 \times 10^{-5}$), Spearman's $\rho = 0.706$ ($p = 4.46 \times 10^{-6}$), and ROC AUC = 0.956 (Table 2 and Table 3). The total cohort of $N = 33$, with 9 deaths, still imposes structural limits on multivariable modeling and subgroup inference. In our multivariable logistic regression (Table 4 and Figure 4b), the Events-Per-Variable (EPV) ratio was 3.0 (9 events across 3 predictors); the adjusted odds ratio for the continuous APACHE II score stayed stable (aOR = 1.74; 95% CI: 1.13–2.69; $p = 0.0127$). The binary malignancy covariate, however, had a wide confidence interval (aOR = 0.69; 95% CI: 0.01–32.20; $p = 0.852$); future multicenter Indonesian registries should therefore use Firth's penalized maximum likelihood logistic regression⁵⁰ and bootstrap optimism-corrected AUC validation in larger samples ($N \geq 150$ –300). Such work could derive pathology-specific risk thresholds for oncologic lung resection, infectious decortication, and trauma thoracotomy.

Strengths and limitations

This study has several notable strengths. It addresses an underrepresented Southeast Asian tertiary thoracic surgical population, its 24-hour physiological and survival follow-up data were 100% complete, its statistical verification was rigorous (non-parametric effect sizes and multivariable regression), and it offers a novel component-level dissection of the APACHE II score. Several important limitations also deserve open discussion. First, the sample was modest ($N = 33$, with 9

mortality events), reflecting a single-center experience limited to open thoracotomy admissions over one calendar year; this limits statistical power for subgroup comparisons and leaves wide 95% confidence intervals around sensitivity (56.5%–98.0%) and PPV (43.4%–90.3%). The high AUC of 0.956 may also reflect small-sample optimism, so external validation is needed before clinical generalization^{23,24,49,50}; second, the retrospective observational design depends on the accuracy of routine clinical documentation and is open to information bias, especially in assessing pre-sedation GCS scores. Strict scoring protocols and dual-investigator chart verification (15% random audit) were used to mitigate this risk. Third, important intraoperative variables were not included in the regression models and remain unmeasured potential confounders^{2,10,32,35}; they include the exact duration of one-lung ventilation, intraoperative blood loss, fluid balance, the anatomic extent of pulmonary resection, histopathologic tumor stage, and serial arterial lactate trajectories. The cut-off was also derived and assessed in the same sample, only patients staying at least 24 hours were included, the non-malignancy subgroup had 8 patients, and neither calibration nor external validation was assessed. Our findings should therefore be read as preliminary and hypothesis-generating.

5. Conclusion

This single-center retrospective cohort study took place at an Indonesian tertiary referral hospital. The initial 24-hour APACHE II score showed prognostic utility for in-hospital mortality after open thoracotomy in this derivation sample. Non-survivors had a nearly three-fold higher median score than survivors (23.0 vs. 8.0 points; $p < 0.001$). Correlation with mortality was strong ($\rho = 0.706$; $p < 0.001$). The score also remained significantly associated with mortality on multivariable logistic regression (aOR = 1.74 per point; $p = 0.0127$), and apparent overall discrimination was high (AUC = 0.956; 95% CI: 0.891–1.000); an exploratory cut-off of ≥12.5 points stratified patients with 88.9% sensitivity, 87.5% specificity, and a 95.5% negative predictive value. That gives a candidate benchmark. It can identify low-risk patients suitable for care de-escalation and high-risk patients who need intensive resuscitation. Subgroup and component analyses showed that mortality was concentrated among patients with underlying malignancy (32.0% vs. 12.5% in non-malignancy). Three components accounted for most of the difference: neurological compromise (GCS), acute kidney injury (serum creatinine), and uncompensated metabolic acidosis (arterial pH). The sample is small ($N = 33$), and the design is single-center and retrospective. These findings are hypothesis-generating. Larger, prospective, multicenter studies should validate the threshold across diverse thoracic surgical populations, adding dynamic serial scoring (Δ APACHE II), intraoperative ventilatory and hemodynamic parameters, and pathology-specific recalibration.

6. Declarations

Ethics Approval and Consent to Participate

The study protocol was reviewed and approved by the Ethics Committee CMHC Research Center Indonesia (Reference number: 2024/155), and institutional research permission was granted by the Directorate of Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia. The requirement for individual written informed consent was formally waived by the Ethics Committee CMHC Indonesia because of the retrospective, non-interventional nature of the study and complete anonymization of patient data before analysis. All procedures were conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki and its subsequent amendments.

Consent for Publication

Not applicable. All clinical, laboratory, and demographic data were fully de-identified and aggregated, and no individual patient-identifiable information or clinical photographs are included in the manuscript.

Availability of Data and Materials

The de-identified master analytical dataset and statistical computation scripts supporting the conclusions of this article are available from the corresponding author upon reasonable academic request, subject to institutional data-sharing policies.

Artificial Intelligence Use Statement

None. No artificial intelligence tools were used in preparing this manuscript.

Author Contribution Statement

A.R.: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. A.U.: Conceptualization, Supervision, Resources, Validation, Writing – review & editing. E.B.: Methodology, Validation, Formal analysis, Supervision, Writing – review & editing. All authors read and approved the final version of the manuscript for publication.

Funding Statement

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

Conflict of Interest / Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the intensivist physicians, thoracic and cardiovascular surgeons, nursing staff, and medical record officers of the General Intensive Care Unit (GICU) and the Department of Thoraco-Cardiovascular Surgery at Dr. Mohammad Hoesin General Hospital, Palembang, and the Faculty of Medicine, Universitas Sriwijaya, for their administrative and clinical support during this investigation.

7. References

- Semmelmann A, Baar W, Haude H, et al. Risk factors for postoperative pulmonary complications leading to increased morbidity and mortality in patients undergoing thoracic surgery for pleural empyema. *J Cardiothorac Vasc Anesth*. 2023;37(9):1659-67. doi:10.1053/j.jvca.2023.04.017
- Piccioni F, Langiano N, Bignami E, et al. One-lung ventilation and postoperative pulmonary complications after major lung resection surgery: a multicenter randomized controlled trial. *J Cardiothorac Vasc Anesth*. 2023;37(12):2561-71. doi:10.1053/j.jvca.2023.04.029
- Bang HT, Vy TT, Tap NV. Length of postoperative hospital stay and related factors after lobectomy for lung cancer: a pre-enhanced recovery after surgery (ERAS) single center assessment. *Cureus*. 2024;16(2):e54724. doi:10.7759/cureus.54724
- Fryer ML, Palleiko BA, Emmerick I, et al. Enhanced recovery after surgery pathway leads to decreased length of stay for patients undergoing minimally invasive lung resection. *J Thorac Dis*. 2024;16(2):1324-37. doi:10.21037/jtd-23-1500
- Skrzypczak P, Kasprzyk M, Kamiński M, et al. Key risk factors for mortality after pneumonectomy for lung cancer: insights from a large single-center cohort study. *J Thorac Dis*. 2025;17(7):4536-49. doi:10.21037/jtd-2024-2204
- Shelley B, McCall P, Glass A, et al. Outcome following unplanned critical care admission after lung resection. *JTCVS Open*. 2022;9:281-90. doi:10.1016/j.jxon.2022.01.018
- Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. *Crit Care Med*. 1985;13(10):818-29. doi:10.1097/00003246-198510000-00009
- Mumtaz H, Ejaz MK, Tayyab M, et al. APACHE scoring as an indicator of mortality rate in ICU patients: a cohort study. *Ann Med Surg (Lond)*. 2023;85(3):416-21. doi:10.1097/MS9.0000000000000264
- Li W, Hu Z, Liu J, et al. The effect of aggressive management of intraoperative body temperature on postoperative APACHE II score and prognosis in high-risk patients undergoing thoracoscopic surgery. *J Thorac Dis*. 2022;14(9):3429-37. doi:10.21037/jtd-22-873
- Liu HM, Zhang GW, Yu H, et al. Association between mechanical power during one-lung ventilation and pulmonary complications after thoracoscopic lung resection surgery: a prospective observational study. *BMC Anesthesiol*. 2024;24(1):176. doi:10.1186/s12871-024-02562-1
- El Tahan MR, Samara E, Marczin N, et al. Impact of lower tidal volumes during one-lung ventilation: a 2022 update of the meta-analysis of randomized controlled trials. *J Cardiothorac Vasc Anesth*. 2023;37(10):1983-92. doi:10.1053/j.jvca.2023.04.018
- Gao Y, Ji D, Fang Q, et al. Effect of low-dose norepinephrine combined with goal-directed fluid therapy on postoperative pulmonary complications in lung surgery: a prospective randomized controlled trial. *J Clin Anesth*. 2024;99:111645. doi:10.1016/j.jclinane.2024.111645
- Shen CI, Yang SY, Chiu HY, et al. Prognostic factors for advanced lung cancer patients with do-not-intubate order in intensive care unit: a retrospective study. *BMC Pulm Med*. 2022;22(1):245. doi:10.1186/s12890-022-02042-7
- Lei SH, Guo GF, Yan T, et al. Acute kidney injury after general thoracic surgery: a systematic review and meta-analysis. *J Surg Res*. 2023;287:72-81. doi:10.1016/j.jss.2023.01.011
- Pathy A, Kar P, Gopinath R, et al. Thoracscore: does it predict mortality in the Indian scenario? A retrospective study. *Indian J Anaesth*. 2022;66(Suppl 5):S257-63. doi:10.4103/ijaija_24_22
- Freudenberger DC, Scheese D, Wolfe LG, et al. More isn't always better: antibiotic duration after surgical decortication in pleural empyema. *J Thorac Dis*. 2024;16(6):3873-81. doi:10.21037/jtd-24-61
- Lin CM, Chen YL, Cheng YF, et al. Optimal timing for video assisted thoracic surgery decortication for improved survival in chronic empyema. *Sci Rep*. 2024;14(1):24548. doi:10.1038/s41598-024-75569-w
- Reichert M, Pösentrup B, Hecker A, et al. Thoracotomy versus video-assisted thoracoscopic surgery (VATS) in stage III empyema: an analysis of 217 consecutive patients. *Surg Endosc*. 2018;32(6):2664-75. doi:10.1007/s00464-017-5961-7
- Khang TT. Management of cardiac trauma and penetrating cardiac injuries with severe hemorrhagic shock: a 5-year experience. *J Trauma Inj*. 2024;37(4):268-75. doi:10.20408/jti.2024.0063
- Singh AK, Prasad G, Mishra P. The impact of thoracic trauma on morbidity and outcomes: a six-year experience from a tertiary care level 1 center. *Cureus*. 2024;16(11):e73580. doi:10.7759/cureus.73580
- Ponholzer F, Groemer G, Ng C, et al. Sarcopenia and mediastinal adipose tissue as a prognostic marker for short- and long-term outcomes after primary surgical treatment for lung cancer. *Cancers (Basel)*. 2023;15(23):5666. doi:10.3390/cancers15235666
- Ponholzer F, Chorazy K, Ng C, et al. External validation of risk prediction scores in patients undergoing anatomic video-assisted thoracoscopic resection. *Surg Endosc*. 2023;37(4):2789-99. doi:10.1007/s00464-022-09786-7
- Sungono V, Hariyanto H, Soesilo TEB, et al. Cohort study of the APACHE II score and mortality for different types of intensive care unit patients. *Postgrad Med J*. 2022;98(1166):914-8. doi:10.1136/postgradmedj-2021-140376
- Chatelain E, Simon M, Hernu R, et al. Factors associated with short- and long-term outcomes in lung cancer patients requiring unplanned invasive mechanical ventilation. *Med Intensiv (Engl Ed)*. 2024;48(1):37-45. doi:10.1016/j.medine.2023.07.014
- von Elm E, Altman DG, Egger M, et al. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-9. doi:10.1016/j.jclinepi.2007.11.008

26. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. doi:10.1136/bmj.h5527
27. Lee ACH, Lee SM, Ferguson MK. Frailty is associated with adverse postoperative outcomes after lung cancer resection. *JTO Clin Res Rep*. 2022;3(11):100414. doi:10.1016/j.jtocrr.2022.100414
28. Tian JY, Hao XY, Cao FY, et al. Preoperative frailty assessment predicts postoperative mortality, delirium and pneumonia in elderly lung cancer patients: a retrospective cohort study. *Ann Surg Oncol*. 2023;30(12):7442-51. doi:10.1245/s10434-023-13696-w
29. Shahi S, Paneru H, Ojha R, et al. SOFA and APACHE II scoring systems for predicting outcome of neurological patients admitted in a tertiary hospital intensive care unit. *Ann Med Surg (Lond)*. 2024;86(4):1895-900. doi:10.1097/MS9.0000000000001734
30. Réa-Neto Á, da Silva Júnior ED, Hassler G, et al. Epidemiological and clinical characteristics predictive of ICU mortality of patients with traumatic brain injury treated at a trauma referral hospital – a cohort study. *BMC Neurol*. 2023;23(1):101. doi:10.1186/s12883-023-03145-2
31. Tasargol O. Factors affecting the mortality rate in non-COVID-19 intensive care unit patients during the COVID-19 pandemic in Cyprus: a retrospective cohort study. *Cureus*. 2023;15(10):e47610. doi:10.7759/cureus.47610
32. Bohn E, Srinathan S, Adu-Quaye J, et al. Predictors of acute kidney injury after lung resection surgery: a retrospective case-control study. *Can J Anaesth*. 2023;70(12):1901-8. doi:10.1007/s12630-023-02602-3
33. Song MJ, Jang Y, Legrand M, et al. Epidemiology of sepsis-associated acute kidney injury in critically ill patients: a multicenter, prospective, observational cohort study in South Korea. *Crit Care*. 2024;28(1):383. doi:10.1186/s13054-024-05167-9
34. White KC, Serpa-Neto A, Hurford R, et al. Sepsis-associated acute kidney injury in the intensive care unit: incidence, patient characteristics, timing, trajectory, treatment, and associated outcomes. A multicenter, observational study. *Intensive Care Med*. 2023;49(9):1079-89. doi:10.1007/s00134-023-07138-0
35. Kumrawat A, Gupta S, Dhillon HS, et al. Prognostic value of admission lactate levels in critically ill patients: a comparative study with Sequential Organ Failure Assessment and Acute Physiology and Chronic Health Evaluation II scores. *Cureus*. 2024;16(10):e71372. doi:10.7759/cureus.71372
36. Beigmohammadi MT, Amoozadeh L, Rezaei Motlagh F, et al. Mortality predictive value of APACHE II and SOFA scores in COVID-19 patients in the intensive care unit. *Can Respir J*. 2022;2022:5129314. doi:10.1155/2022/5129314
37. Bylapudi SK, Nanjan S, Ramasamy S, et al. Role of Acute Physiology, Age, and Chronic Health Evaluation (APACHE) II score in predicting outcomes of peritonitis due to hollow viscous perforation: a prospective observational study. *Cureus*. 2021;13(12):e20155. doi:10.7759/cureus.20155
38. Ishikawa D, Takehara Y, Takata A, et al. Combination of dirty mass volume and APACHE II score predicts mortality in patients with colorectal perforation. *World J Emerg Surg*. 2021;16(1):17. doi:10.1186/s13017-021-00359-y
39. Ramírez-Giraldo C, Isaza-Restrepo A, García-Peralta JC, et al. Surgical mortality in patients in extremis: futility in emergency abdominal surgery. *BMC Surg*. 2023;23(1):21. doi:10.1186/s12893-022-01897-1
40. Mehryar HR, Yarahmadi P, Anzali BC. Mortality predictive value of APACHE II scores in COVID-19 patients in the intensive care unit: a cross-sectional study. *Ann Med Surg (Lond)*. 2023;85(6):2464-8. doi:10.1097/MS9.0000000000000641
41. Gatsori P, Intas G, Katsoulas TA, et al. Arterial stiffness as a predictor of postoperative intensive care unit (ICU) mortality: an observational study. *Cureus*. 2026;18(5):e109812. doi:10.7759/cureus.109812
42. Elrakhawwy HM, Alassal MA, Shaalan AM, et al. Impact of major pulmonary resections on right ventricular function: early postoperative changes. *Heart Surg Forum*. 2018;21(1):E009-17. doi:10.1532/hsf.1864
43. Wang Z, Xu J, Zhang Y, et al. Prediction of acute kidney injury incidence following acute type A aortic dissection surgery with novel biomarkers: a prospective observational study. *BMC Med*. 2023;21(1):503. doi:10.1186/s12916-023-03215-9
44. Pickkers P, Angus DC, Bass K, et al. Phase-3 trial of recombinant human alkaline phosphatase for patients with sepsis-associated acute kidney injury (REVIVAL). *Intensive Care Med*. 2024;50(1):68-78. doi:10.1007/s00134-023-07271-w
45. Mochizuki K, Fujii T, Paul E, et al. Early metabolic acidosis in critically ill patients: a binational multicentre study. *Crit Care Resusc*. 2021;23(1):67-75. doi:10.51893/2021.1.OA6
46. Henrique LR, Souza MB, El Kadri RM, et al. Prognosis of critically ill patients with extreme acidosis: a retrospective study. *J Crit Care*. 2023;78:154381. doi:10.1016/j.jccr.2023.154381
47. Zemlin AE, Sigwadhi LN, Wiese OJ, et al. The association between acid-base status and clinical outcome in critically ill COVID-19 patients admitted to intensive care unit with an emphasis on high anion gap metabolic acidosis. *Ann Clin Biochem*. 2023;60(2):86-91. doi:10.1177/00045632221134687
48. Jansen TC, van Bommel J, Schoonderbeek FJ, et al. Early lactate-guided therapy in intensive care unit patients: a multicenter, open-label, randomized controlled trial. *Am J Respir Crit Care Med*. 2010;182(6):752-61. doi:10.1164/rccm.200912-19180C
49. Monk M, Torres J, Vickery K, et al. A comparison of ICU mortality scoring systems applied to COVID-19. *Cureus*. 2023;15(2):e35423. doi:10.7759/cureus.35423
50. Pühr R, Heinze G, Nold M, et al. Firth's logistic regression with rare events: accurate effect estimates and predictions? *Stat Med*. 2017;36(14):2302-17. doi:10.1002/sim.7273