e-ISSN: 2722-3558

Sriwijaya Journal of Surgery

[SJS]

https://sriwijayasurgery.com

Establishing an APACHE II Cut-off Score for Predicting Mortality in Post-Thoracotomy Patients: A Single-Center Cohort Analysis

Awan Rochaniawan[1*], Ahmat Umar[2], Erial Bahar[3]

[1]Surgery Study Program, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

[2]Department of Thoraco-Cardiovascular Surgery, Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia

[3]Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

ARTICLE INFO

Keywords:
APACHE II score
Critical care
Mortality
Thoracic surgery
Thoracotomy

***Corresponding author:**
Awan Rochaniawan

E-mail address:
awan.rochaniawan@gmail.com

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.37275/sjs.v8i2.131

A B S T R A C T

Introduction: Thoracotomy represents a significant physiological challenge with considerable mortality risk. Early, objective risk stratification in the General Intensive Care Unit (GICU) is essential for guiding clinical management. This study sought to evaluate the utility of the Acute Physiology and Chronic Health Evaluation II (APACHE II) score as a prognostic tool in a heterogeneous post-thoracotomy population at a tertiary referral center in Southeast Asia. **Methods:** A retrospective cohort study was conducted on 33 consecutive patients admitted to the GICU following thoracotomy between January and December 2024. The APACHE II score was calculated using the most deranged physiological values within the first 24 hours of admission. The primary outcome was in-hospital mortality. Statistical analyses included non-parametric tests for group comparisons, Spearman's rank correlation, and Receiver Operating Characteristic (ROC) curve analysis. A novel aspect of this study was the post-hoc stratification of the cohort by the primary surgical indication (malignancy versus non-malignancy) to explore sources of prognostic variability. **Results:** The overall mortality rate was 27.3% (9 of 33 patients). Non-survivors had a significantly higher median APACHE II score than survivors (23 vs. 8; $p < 0.001$). A strong, positive correlation was observed between the APACHE II score and mortality (Spearman's $\rho = 0.706$; $p < 0.001$). ROC analysis demonstrated excellent discriminatory performance for the overall cohort, with an Area Under the Curve (AUC) of 0.956 (95% CI: 0.891–1.000). A score of ≥12.5 was identified as the optimal cut-off, yielding a sensitivity of 88.9% and specificity of 87.5%. Analysis of the APACHE II components revealed that mortality was primarily driven by derangements in neurological (GCS), renal (Creatinine), and acid-base (pH) parameters. **Conclusion:** In this preliminary, single-center analysis, the initial 24-hour APACHE II score demonstrated potential as a powerful prognostic marker for in-hospital mortality following thoracotomy. A candidate cut-off score of ≥12.5 successfully identified a high-risk subgroup. However, given the study's significant limitations, including a small and heterogeneous sample, these findings should be interpreted as hypothesis-generating. They underscore the need for larger, prospective studies to validate this cut-off and to develop more refined prognostic models for specific subgroups of post-thoracotomy patients.

1. Introduction

Thoracotomy, the surgical creation of an opening into the thoracic cavity, remains an indispensable approach for the management of a vast array of pulmonary, mediastinal, and esophageal diseases.[1] It provides unparalleled exposure for complex oncologic resections, control of thoracic trauma, and eradication of deep-seated infections. However, this extensive access is achieved at the cost of a profound physiological insult. The procedure involves the transection of major muscle groups, retraction or resection of ribs, and often necessitates periods of single-lung ventilation.[2] This combination inflicts a significant surgical stress response, compromises

respiratory mechanics, and predisposes patients to a cascade of life-threatening postoperative complications. Consequently, thoracotomy is associated with substantial rates of morbidity and a reported in-hospital mortality ranging from 5% to as high as 15%, depending on the procedural complexity and the patient's underlying condition.[3] Given these high stakes, a large proportion of post-thoracotomy patients require admission to an intensive care setting for vigilant monitoring and advanced organ support. Within this critical care environment, the ability to accurately and objectively stratify patients based on their risk of adverse outcomes is a cornerstone of effective management. Early identification of high-risk individuals allows for the pre-emptive allocation of resources, heightened clinical vigilance, and the timely initiation of aggressive, targeted therapies. Conversely, the confident identification of low-risk patients can facilitate the de-escalation of care and optimize patient flow in resource-limited settings.[4]

For decades, clinicians have relied on prognostic scoring systems to supplement clinical judgment in this endeavor. Among the most enduring and widely implemented of these is the Acute Physiology and Chronic Health Evaluation II (APACHE II) score.[5] Developed by Knaus et al. in 1985, the APACHE II system synthesizes 12 acute physiological variables, patient age, and chronic health status into a single, numerical expression of illness severity. Its strength lies in its capacity to provide a holistic, quantitative snapshot of a patient's physiological derangement at a specific point in time. The score has been validated across a multitude of critically ill populations and serves as a global benchmark for comparing ICU performance and as an essential tool in clinical research for risk adjustment.[6] The post-thoracotomy patient represents a clinical archetype whose pathophysiology aligns remarkably well with the components of the APACHE II score.[7] The immediate postoperative period is often characterized by hypoxemia from ventilation-perfusion mismatch, hypercapnia and acidosis from splinting-induced hypoventilation, tachycardia and fever from the systemic inflammatory response, and hypotension from fluid shifts or hemorrhage. Each of these derangements is directly quantified by the APACHE II system. It is therefore logical to hypothesize that the score should serve as a potent predictor of outcome in this specific surgical cohort.[8]

However, a critical challenge in applying any single prognostic tool to this population is its inherent heterogeneity. A "thoracotomy" is not a uniform procedure but an approach used for widely divergent pathologies. The physiological trajectory and mortality risk of a patient undergoing resection for a localized lung cancer are fundamentally different from those of a patient with a septic, multi-loculated empyema or a patient with exsanguinating thoracic trauma.[9] These subgroups—oncologic, septic, and traumatic—follow distinct pathophysiological pathways to critical illness. Combining them into a single analytical cohort risks generating an "average" prognostic signal that may not be precisely applicable to any individual subgroup. This conceptual challenge underscores the need for studies that not only validate scoring systems but also explore their performance across these clinically distinct populations. While the utility of APACHE II has been established in numerous international studies, its performance characteristics in specific, regional healthcare settings, particularly in Southeast Asia, are less well-defined. Local practice patterns, resource availability, and patient genetics can influence both the presentation of critical illness and its outcomes. Therefore, local validation studies are crucial to confirm the transportability and clinical utility of these established models. This study, conducted at a major tertiary referral hospital in Indonesia, was conceived to address this gap.[10]

The primary novelty of this investigation is its focused, dual-pronged approach. First, it represents an essential effort to validate the APACHE II score in a post-thoracotomy cohort within a specific Indonesian clinical context, a region underrepresented in the critical care literature. Second, and more importantly, it moves beyond simple validation by seeking to establish a statistically derived, clinically actionable

cut-off value using ROC curve analysis. This aims to translate a continuous risk score into a practical, dichotomous tool for frontline clinicians. However, we acknowledge the inherent heterogeneity of our cohort. Therefore, this study was designed not only to test the overall performance of the score but also to conduct a preliminary exploration of its behavior across different pathological subgroups. The primary aim of this study was to investigate the correlation between the APACHE II score, calculated within the first 24 hours of GICU admission, and in-hospital mortality among a heterogeneous cohort of post-thoracotomy patients. The co-primary aim was to determine the optimal APACHE II cut-off score that best discriminates between survivors and non-survivors. A secondary, exploratory aim was to analyze the distinct physiological profiles and outcomes of patients stratified by the primary indication for thoracotomy (malignancy vs. non-malignancy).

2. Methods

A retrospective, observational, single-center cohort study was designed and conducted. The study protocol was submitted to and approved by the Bioethics and Humanities Unit of the Faculty of Medicine, Universitas Sriwijaya, Indonesia. Institutional permission to access patient data was granted by the directorship of Dr. Mohammad Hoesin General Hospital, Palembang. The ethics committee granted a waiver of the requirement for individual patient consent due to the study's retrospective, non-interventional design and the anonymization of all patient data. The study was conducted in strict adherence to the ethical principles outlined in the Declaration of Helsinki. The study was conducted in the General Intensive Care Unit (GICU) of Dr. Mohammad Hoesin General Hospital, a large, government-funded tertiary referral and teaching hospital in South Sumatra, Indonesia. The GICU is a 20-bed, mixed medical-surgical unit managed by a team of intensivists, with 24/7 on-site coverage by critical care fellows and residents. The unit provides a full spectrum of advanced life support, including invasive mechanical ventilation and continuous renal replacement therapy. The study population comprised all consecutive adult patients (age ≥ 18 years) who underwent a thoracotomy procedure and were subsequently admitted directly to the GICU between January 1st, 2024, and December 31st, 2024. A total sampling strategy was employed to include all eligible patients within the study period, thereby maximizing the sample size and providing a complete institutional snapshot. Inclusion criteria were: 1. Age 18 years or older; 2. Admission to the GICU immediately following a thoracotomy. For the purposes of this study, a thoracotomy was defined as any surgical approach involving an incision through the chest wall into the pleural space, including posterolateral, anterolateral, and axillary incisions. Patients who were converted from a minimally invasive approach to an open thoracotomy were included; 3. A minimum GICU length of stay of 24 hours to ensure a complete dataset for the initial APACHE II calculation; 4. Availability of a complete medical record with all necessary clinical and laboratory data for APACHE II scoring. Patients were excluded if their medical records had substantial missing data for any of the 12 acute physiological variables, rendering an accurate APACHE II score calculation impossible.

A standardized data abstraction instrument was designed based on the original APACHE II scoring methodology. A single, trained member of the research team, who was not blinded to patient outcomes, systematically extracted data from the hospital's integrated electronic and paper-based medical record system. To ensure data fidelity, a second investigator independently reviewed a randomly selected 15% of the abstracted records to verify accuracy and resolve any discrepancies by consensus. The following variables were collected: Demographic and Baseline Data: Age at the time of surgery (in years), sex, and the primary indication for thoracotomy. The indication was categorized into three groups based on the final pathological or clinical diagnosis: Malignancy (including primary lung cancer, mediastinal tumors, metastatic disease), Infection (including empyema,

lung abscess, mediastinitis), and Trauma (including penetrating or blunt thoracic injury requiring surgical exploration). APACHE II Score Calculation: Data for the 12 acute physiological variables were meticulously collected. As per standard APACHE II methodology, the most deranged value recorded within the first 24 hours following GICU admission was used for scoring. "Most deranged" was defined as the value furthest from the normal physiological range, conferring the highest number of points. For the Glasgow Coma Scale (GCS), the lowest recorded score was used. In cases where patients were sedated or pharmacologically paralyzed, the GCS score recorded immediately prior to the administration of these agents was used. If no pre-sedation score was available, a score of 15 was assumed as per established guidelines, under the premise that the altered consciousness was iatrogenic and not due to underlying pathology. The final APACHE II score (range 0-71) was calculated by summing the points from the acute physiology score, the age score, and the chronic health points. Primary Outcome: The primary endpoint was all-cause, in-hospital mortality. This was recorded as a binary outcome (survived to hospital discharge vs. deceased).

Data were entered into a secure database and all statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 27.0 (IBM Corp., Armonk, NY, USA). For all inferential tests, a two-tailed p-value of < 0.05 was considered statistically significant. The cohort's characteristics were described using summary statistics. Continuous variables were presented as mean ± standard deviation (SD), while categorical variables were presented as frequencies and percentages. The distribution of the total APACHE II score was formally tested using the Shapiro-Wilk test. A p-value < 0.05 indicated a non-normal distribution, guiding the selection of non-parametric statistical methods. The Mann-Whitney U test was employed to compare the distribution of APACHE II scores between the survivor and non-survivor groups. To quantify the strength and direction of the association between the ordinal APACHE II score and the binary mortality outcome, Spearman's rank correlation coefficient (ρ) was calculated. To address the issue of cohort heterogeneity, a planned post-hoc analysis was conducted. Patients were stratified based on the primary surgical indication (Malignancy vs. Non-Malignancy, which combined the Infection and Trauma groups due to small numbers). Descriptive statistics and mortality rates were calculated for each subgroup. A component analysis was also performed, comparing the mean scores for each of the 12 physiological variables between survivors and non-survivors to identify the primary drivers of mortality. A Receiver Operating Characteristic (ROC) curve was generated for the total cohort by plotting sensitivity against (1 - specificity) across all possible APACHE II score thresholds. The Area Under the Curve (AUC) with its 95% confidence interval (CI) was calculated to measure the overall discriminatory ability of the score. The optimal cut-off point was determined using the Youden's J index (J = Sensitivity + Specificity - 1), which identifies the threshold that maximizes the difference between the true positive rate and the false positive rate. For this identified cut-off, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated with their respective 95% CIs.

3. Results

Figure 1 provides a comprehensive and schematic overview of the study population, encapsulating the fundamental demographic, diagnostic, and prognostic characteristics of the 33 post-thoracotomy patients who required admission to the General Intensive Care Unit (GICU). The cohort is of a modest size (N=33), representing a focused, single-center experience. The mean age of 52.4 years (± 14.8 years) is a clinically significant finding, placing the typical patient squarely in their middle-aged, economically productive years, rather than in an exclusively elderly population often associated with high-risk surgery. This demographic suggests that the physiological insults leading to critical illness were substantial enough to overwhelm the reserves of individuals who might otherwise be

considered to have good baseline health. Furthermore, a distinct male predominance is evident, with males accounting for 60.6% of the cohort. This observation aligns with the established epidemiology of major thoracic diseases, particularly smoking-related malignancies, which historically have a higher incidence in men. This demographic profile immediately frames the clinical problem as one affecting middle-aged men with serious thoracic pathology. The data are dominated by malignancy, which constituted the indication for surgery in an overwhelming 75.8% of patients. This single statistic fundamentally reframes the cohort: these are not merely "post-surgical" patients, but predominantly "post-oncologic surgery" patients. This distinction carries profound implications for their baseline physiological state, nutritional status, and immunological competence. Cancer and its treatments can induce a state of frailty and diminished reserve, rendering these patients exquisitely vulnerable to the major physiological stress of a thoracotomy. The remaining cohort is composed of patients with infectious (18.2%) and traumatic (6.0%) indications. While smaller in number, their inclusion highlights the heterogeneous nature of the post-thoracotomy population, encompassing patients with acute septic physiology and those recovering from severe physical injury, each with their own unique trajectory of critical illness. The in-hospital mortality rate was a sobering 27.3%, meaning more than one in every four patients admitted to the GICU after thoracotomy did not survive to discharge. This figure unequivocally establishes the high-risk nature of the cohort and underscores the urgent clinical need for accurate prognostic tools. It speaks to the severity of the critical illness that developed in these individuals, where even the advanced support of a modern GICU could not alter the fatal course for a significant portion. Juxtaposed against this is the survival of 72.7% of the patients. This majority outcome is equally important, as it demonstrates that while the risk is high, a favorable outcome is achievable for most. The clinical challenge, therefore, is to prospectively identify the 27.3% destined for a poor outcome from the 72.7% who will likely survive.

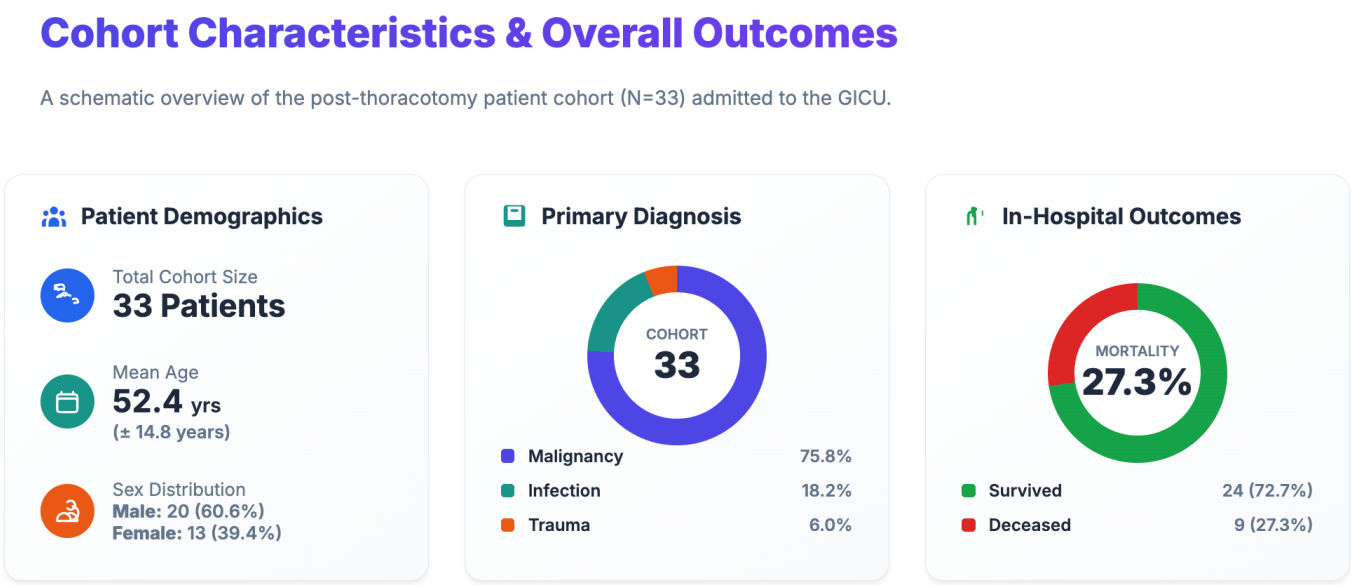

Figure 1. Cohort characteristics & overall outcomes.

Figure 2 provides a multi-faceted and compelling statistical narrative, visually articulating the profound association between the initial 24-hour APACHE II score and the ultimate outcome of in-hospital mortality. The data reveal a profound chasm in physiological derangement between the two groups. The survivor group (n=24) presented with a median APACHE II score of 8, a value indicative of a moderate but manageable level of critical illness. In stark contrast, the non-survivor group (n=9) exhibited a median score of 23, a value nearly three times higher that signifies a state of severe, multi-system physiological distress. This is not a subtle statistical variation; it is a clear quantitative signal of a fundamentally different clinical trajectory present from the earliest moments of GICU admission. The high degree of statistical significance, confirmed by the Mann-Whitney U Test ($p < 0.001$), robustly underscores that this vast difference in illness severity is highly unlikely to be a product of random chance. It reports a Spearman's rank correlation coefficient (ρ) of 0.706. This strong, positive value provides a more nuanced understanding than a simple comparison of medians. It indicates that not only are the scores different between groups, but there is a consistent, directional, and predictable association: as the APACHE II score increases, the likelihood of mortality systematically increases with it. This dose-response-like relationship is crucial, as it validates the score's ability to function as a continuous barometer of risk across its entire range. The infographic rightly concludes that this strong correlation signifies that higher APACHE II scores are significantly associated with an increased risk of death, solidifying the score's role as a reliable prognostic marker. This section details the results of the Receiver Operating Characteristic (ROC) curve analysis, which assesses the score's overall ability to discriminate between survivors and non-survivors. The Area Under the Curve (AUC) of 0.956 is an exceptional result, falling into the category of "excellent" diagnostic accuracy. This high value indicates that if one were to randomly select a patient who survived and one who died, the APACHE II score would correctly identify the sicker patient over 95% of the time. The analysis further identifies an "Optimal Cut-off" score of ≥12.5, providing a clear, evidence-based threshold for risk stratification. The performance metrics of this cut-off are particularly illuminating. A Sensitivity of 88.9% demonstrates that the test is highly effective at identifying patients at risk, correctly flagging nearly nine out of every ten patients who will ultimately die. The Specificity of 87.5% is similarly strong, indicating the test is also proficient at correctly identifying those who will survive. However, the true clinical power of this tool is revealed in its predictive values. The positive predictive value (PPV) of 72.7% suggests that while a high score is a serious warning, a significant portion of patients in this high-risk category can still be salvaged with intensive care. Conversely, the exceptionally high negative predictive value (NPV) of 95.5% is perhaps the most powerful metric for clinical decision-making. It provides a high degree of confidence that a patient with an APACHE II score below 12.5 has an excellent prognosis and is very likely to survive. This metric offers a strong evidence base for clinicians to de-escalate care, allocate resources effectively, and provide reassuring information to patients' families.

Figure 3 transitions from the broad prognostic validation presented in the preceding figures to a more granular and clinically insightful exploratory analysis. It stratifies the patient population into two clinically relevant subgroups: the "Malignancy Group" (n=25), comprising patients undergoing thoracotomy for oncologic indications, and the "Non-Malignancy Group" (n=8), which includes patients with infectious or traumatic pathologies. The comparison between these groups is stark and illuminating. Patients in the Malignancy Group entered the GICU with a substantially higher baseline level of illness severity, as evidenced by a median APACHE II score of 11. This is in sharp contrast to the Non-Malignancy Group, which presented with a much lower median score of 7.5. This divergence in initial severity scores is directly mirrored in the clinical outcomes.

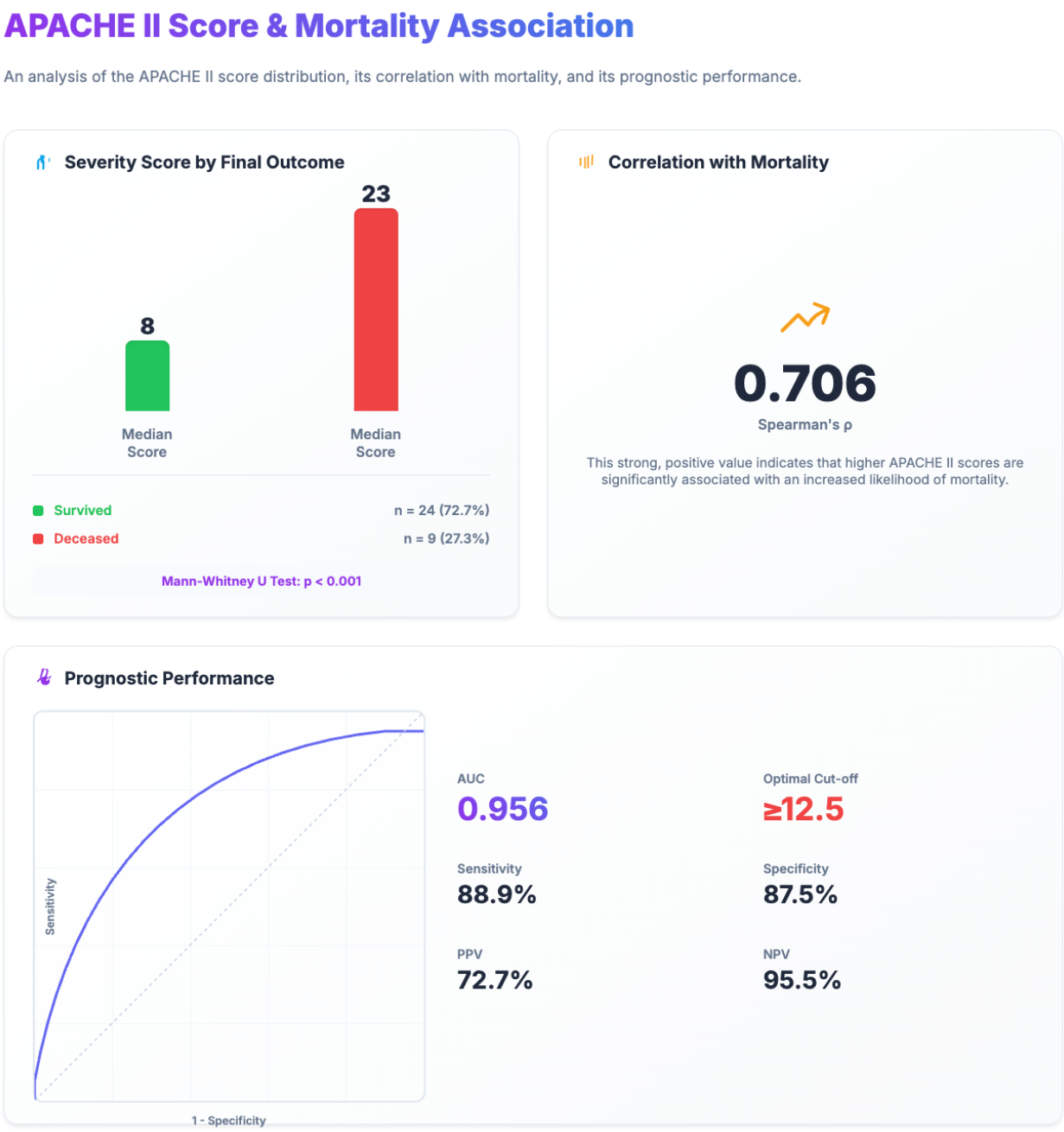

Figure 2. APACHE II score & mortality association.

The mortality rate in the Malignancy Group was a staggering 32.0%, a figure that is more than two and a half times higher than the 12.5% mortality rate observed in the Non-Malignancy Group. This finding is of profound clinical significance. It strongly suggests that the primary diagnosis is not merely a background detail but is, in fact, a powerful independent modulator of both the initial physiological insult and the ultimate risk of death. Patients with cancer often approach major surgery with a diminished physiological reserve due to factors like cachexia, malnutrition, and underlying immunosuppression. This pre-existing vulnerability likely makes them less resilient to the profound stress of a thoracotomy,

leading to more severe postoperative derangements and a higher probability of a fatal outcome. This panel powerfully argues against a one-size-fits-all approach to prognostication, highlighting that the context of why a thoracotomy was performed is a crucial variable in accurately assessing a patient's risk. It employs a sophisticated horizontal "tornado plot" to compare the mean scores of the most critical APACHE II components between survivors and non-survivors. This analysis moves beyond the total score to pinpoint the specific organ systems that fail in fatal cases. The results are unequivocal. The three components with the most dramatic and visually striking disparity between the groups were the Glasgow Coma Scale (GCS), Serum Creatinine, and Arterial pH. The mean GCS score component in non-survivors was 6.33, more than three times higher than the 1.83 seen in survivors. This identifies neurological compromise—whether from sepsis-associated encephalopathy, perioperative hypoxia, or metabolic derangement—as a primary harbinger of mortality. Similarly, the mean Creatinine score component was 3.44 in non-survivors versus a mere 0.96 in survivors, indicating that the development of acute kidney injury is a pivotal event in the cascade toward a fatal outcome. Finally, the mean Arterial pH score component of 1.89 in non-survivors, compared to 0.33 in survivors, signals that severe, uncompensated acidosis, a marker of profound circulatory shock and cellular dysfunction, is a near-terminal finding. While other parameters like Heart Rate and PaO_2 were also worse in non-survivors, their differences were far less pronounced. Figure 3 provides a crucial layer of depth to the study's narrative. It first demonstrates that the risk following thoracotomy is not uniform but is significantly elevated in patients with malignancy. It then dissects this risk, revealing that the fatal pathway in this cohort is not typically one of simple respiratory failure, but a systemic, multi-organ collapse characterized by a devastating triad of encephalopathy, acute kidney injury, and profound metabolic acidosis. This component analysis provides clinicians with a clear "signature of mortality," directing their vigilance and therapeutic efforts toward protecting

Figure 3. Exploratory subgroup & component analysis.

4. Discussion

This single-center, retrospective study was undertaken to evaluate the prognostic utility of the APACHE II score in a cohort of critically ill patients following thoracotomy. The investigation yielded several key findings: the APACHE II score demonstrated a powerful correlation with in-hospital mortality; the score exhibited excellent discriminatory capacity as measured by the AUC; and a candidate cut-off score of ≥12.5 was identified that effectively stratified patients. However, these results, while statistically robust within our dataset, must be interpreted with significant caution. The study's primary value lies not in providing a definitive, generalizable answer, but in generating a hypothesis and offering a detailed clinical and pathophysiological exploration of risk in a unique and understudied patient population. The core strength of this study is its ability to move beyond a simple statistical correlation and begin to dissect the "why" behind the prognostic power of the APACHE II score.[11] While non-survivors were demonstrably sicker across nearly all parameters, mortality was not driven by a uniform decline. Instead, it was overwhelmingly associated with catastrophic failure in three specific systems: neurological, renal, and metabolic (acid-base balance). The most striking contributor to mortality risk was neurological dysfunction, as evidenced by the dramatic difference in the GCS component score (mean of 6.33 in non-survivors vs. 1.83 in survivors). A low GCS in a post-surgical patient is a profoundly ominous sign, representing the final common pathway of numerous insults.[12] It can result from direct cerebral hypoperfusion during intraoperative hypotension, sustained postoperative hypoxia, the neurotoxic effects of sepsis-associated encephalopathy, or severe metabolic derangements. It reflects a state where the patient's homeostatic mechanisms have failed to protect the most vital organ.

The second pillar of this mortality signature was acute kidney injury (AKI), captured by the serum creatinine component (mean of 3.44 vs. 0.96).[13] The kidneys are exquisitely sensitive to the insults of major surgery. The combination of the systemic inflammatory response, potential nephrotoxic agents, and periods of hypovolemia or hypotension creates a perfect storm for acute tubular necrosis. The development of AKI is not merely a biomarker of illness; it is an engine of further decline.[14] It leads to volume overload, intractable metabolic acidosis, electrolyte emergencies, and uremic encephalopathy, thereby perpetuating a vicious cycle of multi-organ failure. Finally, the development of severe metabolic acidosis, reflected in the arterial pH component (mean of 1.89 vs. 0.33), was a key feature of non-survivors. While some of this may be a residual respiratory acidosis from hypoventilation, a significant metabolic component likely represents systemic tissue hypoperfusion and the generation of lactic acid. It is a direct measure of shock at the cellular level. When a patient can no longer buffer this acid load, it signals a terminal decline in circulatory integrity. This triad—encephalopathy, renal failure, and acidosis—paints a vivid picture. Mortality in our post-thoracotomy cohort was not typically a simple failure of the lungs, but a systemic, multi-organ collapse where the brain and kidneys were the critical failing points. This insight is clinically actionable, suggesting that management priorities for high-risk patients should focus aggressively on neuro-protection (optimizing cerebral perfusion pressure), reno-protection (maintaining MAP, avoiding nephrotoxins), and the early recognition and reversal of shock.[15]

Our study identified a cut-off score of ≥12.5 with an impressive AUC of 0.956. While statistically sound within our cohort, the application of a single prognostic threshold to a pathologically heterogeneous population is fraught with conceptual difficulties. Our subgroup analysis (Table 4) begins to reveal this complexity. The mortality rate in the malignancy group was 32.0%, more than double the 12.5% rate in the non-malignancy group. These are, in effect, two different diseases. The patient with cancer often enters surgery in a state of chronic frailty, with diminished physiological reserve due to cachexia and potential

immunosuppression.[16] Their postoperative course may be one of slow decline. In contrast, a patient with a septic empyema presents with an acute, hyper-inflammatory state, and their risk is front-loaded. Therefore, the 12.5 cut-off is a statistical composite. It may be too low for the septic patient (where scores are expected to be high) and perhaps too high for the elective oncology patient. The exceptionally high AUC may, in part, be an artifact of the model's ability to easily distinguish between very sick septic patients and less sick elective patients, a distinction that is already clinically apparent.[17] The true test of a scoring system is its ability to discriminate risk within these more homogenous subgroups. Our study is underpowered to perform this analysis definitively, but our findings strongly suggest that a one-size-fits-all approach to prognostication is suboptimal.[18] The 12.5 cut-off should not be interpreted as a universal constant, but rather as an institutional benchmark that can trigger a "clinical pause"—a moment for the team to reassess a patient whose physiological derangement has crossed a significant threshold—while acknowledging that the true risk is modified by the underlying diagnosis.

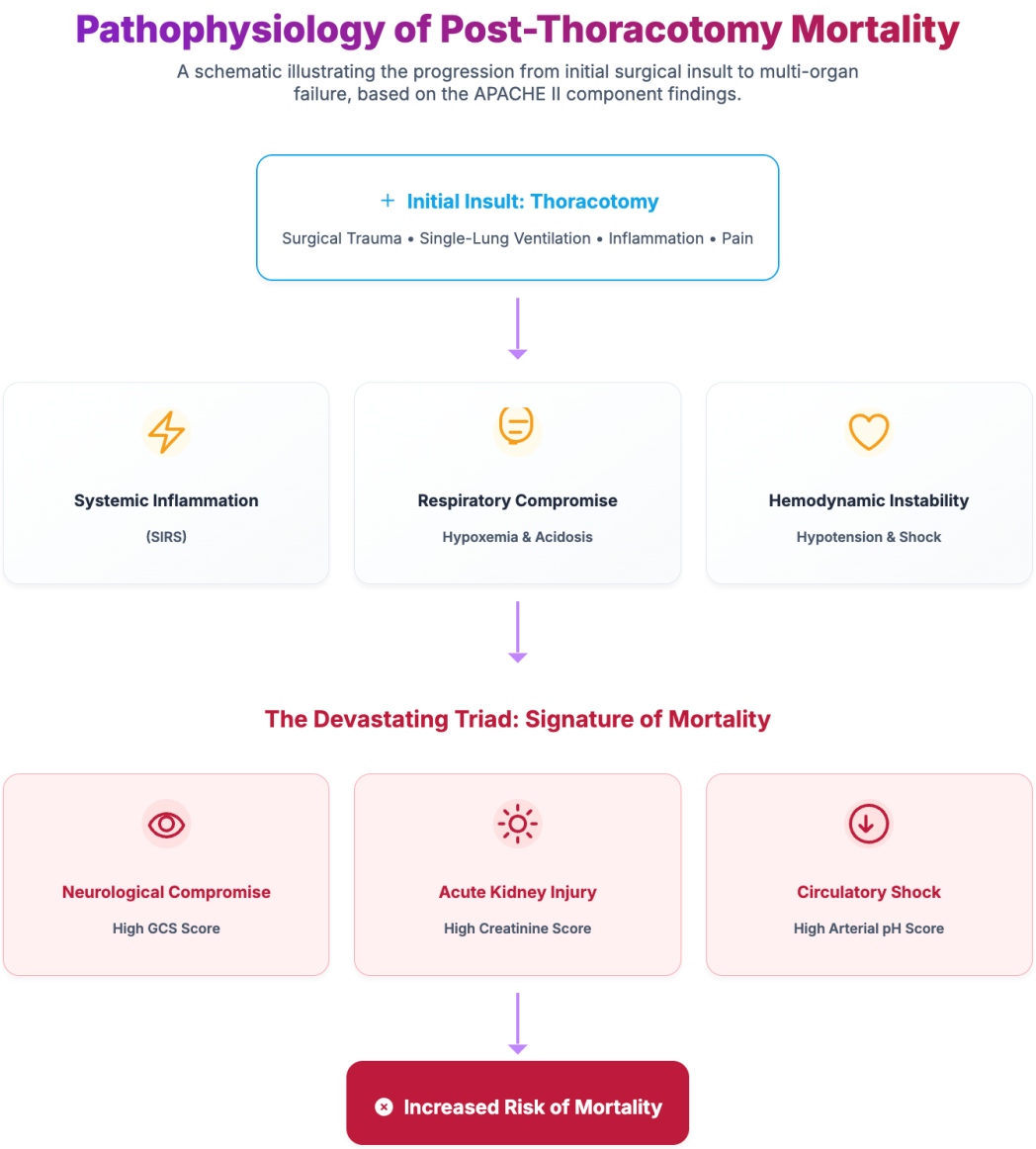

Figure 4. Pathophysiology of post-thoracotomy mortality.

Figure 4 serves as the conceptual capstone of this investigation, providing a powerful and elegant schematic that visually synthesizes the study's key findings into a coherent pathophysiological narrative. It translates complex clinical and statistical observations into an intuitive, sequential flowchart that illustrates the inexorable progression from the initial surgical insult to the final, fatal outcome of multi-organ failure. The figure is not merely a summary of data points; it is a graphical representation of the very mechanism of mortality as revealed by the APACHE II component analysis, providing a clear and memorable framework for understanding why and how post-thoracotomy patients succumb to their critical illness. The flowchart begins at the apex with Stage 1: The Initial Insult. This stage represents the inciting event—the thoracotomy procedure itself. It is correctly identified as a multi-faceted trauma, encompassing not only the direct surgical injury to tissues but also the profound physiological stress of single-lung ventilation, the systemic release of inflammatory mediators, and the significant postoperative pain that characterizes this surgical approach. This initial stage acts as the primary trigger, setting in motion a cascade of systemic responses that form the basis of the patient's critical illness and are directly measured by the APACHE II scoring system. From this single trigger, the flowchart progresses downward to Stage 2: Physiological Derangements, which visually depicts the immediate, systemic consequences of the surgical insult. This stage is intelligently subdivided into three core pillars of postoperative instability, each representing a distinct but interconnected physiological axis. The first, Systemic Inflammation (SIRS), acknowledges the host's response to tissue injury, a state characterized by fever, tachycardia, and leukocytosis. The second, Respiratory Compromise, highlights the most immediate organ-specific consequence of thoracotomy, leading to the critical derangements of hypoxemia and acidosis. The third, Hemodynamic Instability, captures the circulatory consequences, including hypotension and the potential progression to a state of shock. This stage effectively illustrates how the localized surgical event rapidly evolves into a systemic crisis, with derangements that are precisely the variables quantified by the APACHE II score. The crucial turning point in the schematic is Stage 3: The Devastating Triad. This is the conceptual heart of the figure and represents the study's most important and novel insight, directly derived from the component analysis. It posits that while the initial derangements are widespread, the pathway to mortality is not a generalized decline but is instead characterized by the catastrophic failure of three specific, interrelated systems. This "Signature of Mortality" is what distinguishes non-survivors from those who can withstand the initial physiological storm. The triad consists of: Neurological Compromise, as evidenced by a high GCS component score; Acute Kidney Injury, directly measured by a high Creatinine component score; and Circulatory Shock, reflected by a high Arterial pH component score. By visually isolating these three elements in a distinct, high-alert color, the figure powerfully argues that the concurrent failure of the brain, the kidneys, and the circulatory system represents the point of no return. This is no longer just a state of SIRS or respiratory distress; it is a state of advanced, irreversible multi-organ dysfunction syndrome (MODS). Finally, the cascade culminates in Stage 4: Increased Risk of Mortality. This terminal stage is the logical and inevitable consequence of the devastating triad. The failure of these three critical homeostatic systems leads to a state that is incompatible with life, thus explaining the dramatically increased mortality observed in patients whose APACHE II scores reflected these specific derangements. Figure 4 provides a masterful visual narrative. It begins with a single, well-defined insult and logically progresses through the subsequent waves of physiological derangement. Its most significant contribution is the clear visualization of "The Devastating Triad," which crystallizes the core findings of the component analysis into an easily understandable and clinically relevant concept. The

schematic serves as a powerful educational tool, perfectly illustrating how the abstract numerical values of the APACHE II score are, in fact, direct readouts of a deadly and predictable pathophysiological cascade. It provides clinicians with not just a score, but a mechanistic understanding of the pathway to mortality, thereby guiding their focus toward the critical tasks of protecting the brain, preserving.[19]

A fundamental limitation of this study, and indeed of any study relying on a single, initial severity score, is that it provides only a static snapshot of a highly dynamic process. The first 24 hours in the GICU are a period of intense therapeutic activity and physiological flux. A patient's initial APACHE II score, heavily influenced by the immediate aftermath of anesthesia and surgery, may not reflect their true physiological trajectory. The real prognostic power in critical illness often lies in the patient's response to therapy. Contemporary research in prognostication has increasingly shifted focus from static, admission-day scores to dynamic scoring, or the analysis of the "delta-APACHE" over the first 48 to 72 hours. A patient who arrives with a high score of 20 but, with aggressive resuscitation, improves to a score of 12 by day three has a far better prognosis than a patient who arrives with a score of 10 that escalates to 18. The former demonstrates physiological resilience and response to treatment, while the latter signals a deteriorating, non-responsive state. Our study design cannot capture this crucial temporal element. Therefore, while the initial score is a useful starting point for risk stratification, it should be viewed as exactly that—a starting point. The true art of intensive care lies in continuous reassessment, and future prognostic models should aim to incorporate this dynamic element.[20]

This study's findings must be framed by its significant limitations. The primary limitation is the small sample size (N=33). This severely restricts the statistical power of our analyses and leads to wide confidence intervals for our performance metrics, particularly for sensitivity and PPV. This imprecision means that the true values could be substantially different in a larger population. The exceptionally high AUC could be a result of overfitting, where the model perfectly describes the idiosyncratic features of this small dataset but would perform less well on a new set of patients. Consequently, our findings must be considered preliminary and hypothesis-generating. Second, the retrospective design introduces potential for information bias from inaccuracies or omissions in the medical record. While we employed a rigorous data abstraction protocol, we cannot eliminate this possibility. Third, as a single-center study, our results are subject to the unique patient demographics, case-mix, and practice patterns of our institution, which may limit their generalizability. Finally, our analysis could not account for numerous important unmeasured confounders, such as intraoperative variables (surgical duration, blood loss), anesthetic techniques, and specific postoperative care protocols (pain management, fluid strategies), all of which can profoundly impact outcomes.

5. Conclusion

In this preliminary, single-center cohort study, the APACHE II score, calculated within the first 24 hours of GICU admission, emerged as a statistically powerful predictor of in-hospital mortality for a heterogeneous population of post-thoracotomy patients. An analysis of the score's components suggests that mortality was principally driven by a cascade of neurological, renal, and metabolic failures. We identified a candidate APACHE II cut-off score of ≥12.5, which demonstrated excellent discriminatory capacity within our cohort and, most notably, a very high negative predictive value, making it potentially useful for identifying patients with a high likelihood of survival. However, given the profound limitations of our small and heterogeneous sample, these findings must be interpreted with extreme caution. They do not represent a definitive validation but rather a compelling proof-of-concept and a call to action. The results strongly support the need for larger, prospective, multi-center studies to validate this cut-off score and, more importantly, to develop and test

more nuanced prognostic models that account for the distinct pathophysiological trajectories of different surgical subgroups. Until such work is done, the APACHE II score should be used as one of several tools to augment, but not replace, the thoughtful clinical judgment that remains the cornerstone of managing these complex and critically ill surgical patients.

6. References

1. Situ O, Badejo O, Gwaram U. The predictive role of APACHE II score in patients with peritonitis at National Hospital Abuja. Br J Surg. 2024; 111(Suppl_6).
2. Rasslan R, de Oliveira Ferreira F, Parra JAP, da Costa Ferreira Novo F, Menegozzo CAM, Alves PHF, et al. Damage control surgery in non-traumatic abdominal emergencies: prognostic value of SOFA, APACHE II, and Mannheim peritonitis index. Updates Surg. 2025; 77(1): 255–62.
3. Modi JV, Sheth J. Observational study of comparison of BISAP score with Ranson's score and APACHE II scores in assessing severity and prognosis in patients of acute pancreatitis. Int Surg J. 2018; 6(1): 178.
4. Kumaraswamy S. A comparative study between APACHE II and Ranson scoring systems inpredicting the severity of acute pancreatitis. New Ind J Surg. 2020; 11(2): 126–9.
5. Singh S. A randomized trial of comparing single layer continuous versus double layer interrupted intestinal anastomosis techniques in a group of risk-stratified patients using APACHE-II scoring system. New Ind J Surg. 2020; 11(3): 411–8.
6. Laczynski DJ, Gallop J, Lyden SP, Rowse J, Smolock CJ, Kirksey L, et al. Benchmarking a center of excellence in vascular surgery: Using APACHE II to validate outcomes in a tertiary care institute. J Vasc Surg. 2020; 72(3): e313.
7. Sarma H, Kudari A. A study on the trends and patterns of intestinal obstruction and surgical outcome based on APACHE score II in a tertiary care centre. Int Surg J. 2021; 8(3): 856.
8. Ishikawa D, Takehara Y, Takata A, Takamura K, Sato H. Combination of dirty mass volume and APACHE II score predicts mortality in patients with colorectal perforation. World J Emerg Surg. 2021; 16(1): 17.
9. Singh S, Shankar R, Bansal M, Jain N. Comparison of Intra-abdominal pressure monitoring vs APACHE-II and Ranson's scoring in predicting the outcome of Acute Pancreatitis. Int J Surg Med. 2022; 1.
10. Garg A, Naskar D, Gupta N, Yelamanchi R, Gupta AK. Comparison of red cell distribution width and neutrophil lymphocyte ratio with APACHE II score in predicting severity of acute pancreatitis. Indian J Surg. 2022; 84(S2): 492–9.
11. Gupta AK, Yadav E, Gupta N, Yelamanchi R, Bansal LK, Durga CK. Role of intra-abdominal pressure as a prognostic marker and its correlation with APACHE II and SOFA scores in patients of acute pancreatitis. Indian J Surg. 2022; 84(S2): 485–91.
12. Mumtaz H, Ejaz MK, Tayyab M, Vohra LI, Sapkota S, Hasan 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.
13. 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.
14. Shahi S, Paneru H, Ojha R, Karn R, Rajbhandari R, Gajurel BP. 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.
15. Bujakovic T, Delibegovic S. The possibility of using the APACHE II score in prognosis of the

development of postoperative complications in the elderly with hip fractures. Eur J Orthop Surg Traumatol. 2025; 35(1): 193.

16. Subedi RP, Kumar N, Karn S, Arunkumar V, Raj N, Maheshwari P, et al. Prognostic value of the combination of serial APACHE II with serum lactate for predicting post-operative mortality in gastrointestinal perforation peritonitis: a prospective cohort study. BMC Surg. 2025; 25(1): 374.
17. Gallo E, Gaudard P, Provenchère S, Souab F, Schwab A, Bedague D, et al. Use of aprotinin versus tranexamic acid in cardiac surgery patients with high-risk for excessive bleeding (APACHE) trial: a multicentre retrospective comparative non-randomized historical study. Eur J Cardiothorac Surg. 2024; 65(2).
18. Ng WT, Ling L, Joynt GM, Chan KM. An audit of mortality by using ECMO specific scores and APACHE II scoring system in patients receiving extracorporeal membrane oxygenation in a tertiary intensive care unit in Hong Kong. J Thorac Dis. 2019; 11(2): 445–55.
19. Beigmohammadi MT, Amoozadeh L, Rezaei Motlagh F, Rahimi M, Maghsoudloo M, Jafarnejad B, 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.
20. Li W, Hu Z, Liu J, Yuan Y, Li K. 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.