Vol. 4 No. 4 (2025)

Pakistan Journal of Medical & Cardiological Review

https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387

Print ISSN 3007-2379

Development and Internal Validation of a Clinical-Radiological Nomogram for In-Hospital Mortality Prediction in Moderate-to-Severe Traumatic Brain Injury

Dr. Muhammad Nawaz Khan
MBBS, FCPS, Assistant Professor, Department of Neurosurgery, Lady Reading Hospital, Peshawar, Pakistan Email: nawazktk.lrh@gmail.com
Telephone: +92-333-9492702

Dr. Gohar Ali*
MBBS, FCPS, Assistant Professor, Department of Neurosurgery, Mardan Medical Complex, Mardan, Pakistan Email: docgoharali@gmail.com
Telephone: +92-3349300370

Author Details

Keywords: Traumatic Brain Injury; In-Hospital Mortality; Nomogram; Prognostic Model; Rotterdam CT Score; Glasgow Coma Scale

.

Received on 10 Nov 2025

Accepted on 10 Dec 2025

Published on 23 Dec 2025

Corresponding E-mail & Author*:

Dr. Gohar Ali
MBBS, FCPS, Assistant Professor, Department of Neurosurgery, Mardan Medical Complex, Mardan, Pakistan
Email: docgoharali@gmail.com
Telephone: +92-3349300370

Abstract

Background:
Traumatic brain injury (TBI) remains a leading cause of in-hospital mortality worldwide, particularly in low- and middle-income countries. Widely used CT-based prognostic systems, such as the Marshall and Rotterdam scores, provide only moderate accuracy and do not integrate key clinical variables. There is a need for a simple, reliable bedside tool that combines clinical and radiological factors for early mortality prediction in moderate-to-severe TBI.

Objective:
To develop and internally validate a clinical–radiological nomogram for predicting in-hospital mortality in adults with moderate-to-severe traumatic brain injury, and to compare its performance with established CT-based scoring systems.

Methods:
This retrospective cohort study included adult patients (≥18 years) with moderate-to-severe TBI (post-resuscitation GCS ≤12) admitted to a tertiary-care trauma center between January 2020 and December 2024. Candidate predictors were selected a priori and included age, admission GCS motor score, pupillary reactivity, Rotterdam CT score, and anticoagulant use. Missing data were handled using multiple imputation. Multivariable logistic regression was used to derive the model, followed by bootstrap internal validation (1,000 resamples) with coefficient shrinkage. Model discrimination, calibration, overall accuracy, and clinical utility were assessed and compared with the Marshall and Rotterdam CT scores.

Results:
A total of 2,400 patients were analyzed; median age was 56 years (IQR 38–73), and 69% were male. In-hospital mortality occurred in 240 patients (10.0%). The final

nomogram demonstrated excellent discrimination with an optimism-corrected area under the curve (AUC) of 0.90 (95% CI 0.88–0.92), strong calibration (slope 0.99; intercept 0.00), and a low Brier score (0.062). The nomogram significantly outperformed the Rotterdam CT score (AUC 0.79) and Marshall classification (AUC 0.70) ($p<0.001$ for both comparisons). Decision curve analysis showed superior net clinical benefit across clinically relevant risk thresholds (10–60%). The model also performed well for predicting unfavorable functional outcome at discharge (AUC 0.86).

Conclusion:

A parsimonious clinical–radiological nomogram incorporating routinely available admission variables accurately predicts in-hospital mortality after moderate-to-severe TBI and outperforms traditional CT-based scoring systems. This tool may aid early risk stratification, prognostication, and clinical decision-making, particularly in resource-limited trauma settings.

Introduction

Traumatic brain injury (TBI) is a major global health problem and a leading cause of death and long-term disability[1,2]. Recent epidemiological data underscore the enormous burden: for example, one analysis estimated that ~27 million people sustain TBI annually (2019 data)[3], and other sources have suggested even higher totals (on the order of 64–74 million when all severities are included)[4]. Outcomes are highly variable, but among survivors a large proportion suffer persistent deficits – about half of TBI patients are left with moderate-to-severe functional impairment[1,2]. The problem is particularly acute in low- and middle-income countries (LMICs). Indeed, the incidence of TBI in LMICs is estimated to be roughly three times that in high-income countries[5], and LMICs now account for over 90% of road-traffic fatalities worldwide[6]. Road traffic collisions are the predominant mechanism of injury in many LMIC settings[5,6], often affecting young male drivers and pedestrians. This epidemiological heterogeneity ranging from high-speed crashes in young adults to falls in the elderly means that patients with similar admission Glasgow Coma Scale (GCS) scores or CT findings can nonetheless have very different trajectories[2,7]. In short, TBI is not only common and dangerous, but also highly heterogeneous in clinical presentation and outcome.

Existing prognostic tools for moderate-to-severe TBI rely heavily on neuroimaging. The Marshall CT classification (1991) and the Rotterdam CT score (2005) are the most widely used radiological grading systems[8,9]. These scores emphasize structural features on the head CT (e.g. mass lesions, midline shift, cisternal patency) to stratify risk. While simple and ubiquitous, they have important limitations. For one, they ignore clinical variables (beyond the GCS component in the Rotterdam score) and thus do not capture aspects like physiological status or premorbid factors. Both scores were developed decades ago in specific cohorts, and may not be optimally calibrated to today's spectrum of patients or treatments[9]. In external validation studies, their discrimination is only moderate – for example, the Marshall classification often achieves an AUC ~0.65–0.70 for in-hospital mortality and the Rotterdam score AUC ~0.75–0.85[10,11]. Moreover, these CT-based models frequently misestimate risk across different populations, tending to overpredict mortality in some cohorts[11,12]. International prognostic calculators such as the CRASH and IMPACT models extend beyond CT findings by incorporating age, GCS components, pupillary reactivity and even laboratory values[11,12,13,14,]. These combined models achieve better accuracy (AUC ~0.80–0.85 for mortality) than CT-only scores, but they are complex and not widely used in routine care[11,15]. In particular, in busy or resource-limited settings it is impractical to compute a CRASH/IMPACT score at the bedside. In summary, there is

a lack of an easily usable, validated bedside tool that integrates both clinical and imaging data for early prognostication in moderate-to-severe TBI.
Against this background, the present study aimed to derive and internally validate a simple yet powerful prognostic model for moderate-to-severe TBI. Our goal was to create a clinical nomogram that combines key admission variables (clinical plus CT features) to predict in-hospital mortality (and secondarily functional outcome), improving upon traditional CT scores. We hypothesized that a model incorporating age, GCS (motor score), pupillary reactivity, anticoagulant use and the Rotterdam CT score would substantially outperform CT-alone models and provide accurate individualized risk estimates**.**

Methods

Study Design and Setting

This was a retrospective cohort study conducted at the Department of Neurosurgery, Lady Reading Hospital Peshawar, a tertiary-care Level-1 trauma center.
All patients were managed according to institutional TBI protocols consistent with the Brain Trauma Foundation (BTF) guidelines.
The study period extended from January 1,2020 to December 31, 2024.
The study was approved by the institutional review board . The requirement for informed consent was waived owing to the retrospective design.
The primary aim was to externally validate existing CT-based prognostic scores and derive a clinical nomogram for in-hospital mortality.
Secondary aims included prediction of functional outcome and comparison of model performance across systems.

Eligibility Criteria

Patients were eligible if they met the following inclusion criteria:
Age ≥18 years
Moderate-to-severe TBI, defined as a post-resuscitation Glasgow Coma Scale (GCS) score ≤12
Head CT scan obtained within 6 hours of admission
Documented in-hospital outcome

Exclusion criteria were:
Death or withdrawal of care before CT or within 6 hours of admission
Transfer from another hospital >48 hours post-injury
Missing outcome data or >20% missing for key predictors
For intubated patients, the pre-intubation GCS was used when available; otherwise, the admission GCS motor score was used as the primary clinical predictor.

Study Outcomes

The primary outcome was in-hospital mortality (death from any cause during index hospitalization).
The secondary outcome was functional status at discharge, assessed using the Glasgow Outcome Scale (GOS), dichotomized as:
Favorable outcome: GOS 4–5
Unfavorable outcome: GOS 1–3

Candidate Predictors

Predictors were selected a priori based on clinical relevance and prior prognostic literature.
All variables were defined at the time of admission or first CT scan.

Category	Variable	Definition / Measurement
Demographic	Age (years)	Continuous
Clinical	Admission GCS motor score	Ordinal (1–6)
Clinical	Pupillary reactivity	Both reactive / One fixed / Both fixed
Imaging	Rotterdam CT score	1–6, based on initial CT
Medical	Anticoagulant or antiplatelet use	Yes/No
Systemic	Intracranial pressure (ICP) monitoring	Binary (for sensitivity only)

The GCS motor component was chosen over total GCS due to its superior reliability and validity in intubated patients.
The Rotterdam CT score was preferred to avoid multicollinearity with its subcomponents.
The ICP monitoring variable was treated as a process-of-care indicator and reserved for sensitivity analyses to avoid treatment bias.

Data Collection and Management
Data were extracted from the hospital's HMIS .Two trained data abstractors independently entered all variables. Discrepancies were resolved by consensus or, if unresolved, by the senior investigator.

CT Scoring Workflow:
All admission CT scans were independently scored by two neurosurgeon blinded to patient outcomes.
Discrepancies were resolved through consensus review, and if disagreement persisted, a senior neuroradiologist served as an arbiter.
Inter-rater reliability was quantified in a 15% random sample (n=360) using the intraclass correlation coefficient (ICC) for continuous measures and weighted kappa for ordinal ratings.
Data quality was assured through range checks, logical consistency tests, and random audits against source records.

Handling of Missing Data
Predictor variables had <5% missingness, except pupillary reactivity (4.2%).
Missing data were addressed using multiple imputation by chained equations (MICE) with 20 imputed datasets.
Predictive mean matching was applied for continuous variables and logistic regression for categorical variables.
The imputation model incorporated all candidate predictors and the outcome to satisfy the missing-at-random assumption.
A complete-case sensitivity analysis (n = 2,201) was conducted to confirm robustness.

Statistical Analysis
Descriptive Statistics
Continuous variables were summarized as median (IQR) and categorical variables as frequency (percentage).
Group comparisons (survivors vs. non-survivors) employed the Mann–Whitney U test for continuous data and chi-square or Fisher's exact test for categorical variables.
Two-sided $p < 0.05$ was considered statistically significant.

External Validation of Existing CT-Based Scores
Before developing a new model, the Marshall and Rotterdam CT scores were externally validated for mortality prediction.
Discrimination: Assessed by area under the receiver operating characteristic curve (AUC) with 95% CIs.
Calibration: Evaluated using calibration plots, intercept/slope analysis, and the Brier score.
Comparison: AUCs were compared using DeLong's test.

Model Development and Internal Validation
A multivariable logistic regression model was developed using a pre-specified, closed-testing approach, where all six candidate predictors were entered simultaneously without automated selection.
Model assumptions were checked as follows:
Linearity in the logit: examined using restricted cubic splines (3 knots)
Multicollinearity: assessed using variance inflation factor (VIF) with threshold VIF < 2
Influence diagnostics: verified using Cook's distance (<0.5)
Internal validation was performed using bootstrap resampling (1,000 iterations).
A shrinkage factor derived from the bootstrap (0.95) was applied to the final model coefficients.
Model performance was reported as:
Discrimination: optimism-corrected AUC
Calibration: intercept, slope, and visual calibration plot
Overall accuracy: Brier score
A graphical nomogram was constructed using the rms R package.

Model Performance and Comparison
The new nomogram was compared to the Marshall and Rotterdam CT-based models using:
ΔAUC (DeLong test)
ΔBrier score
Continuous Net Reclassification Improvement (NRI)
Clinical usefulness was assessed using Decision Curve Analysis (DCA) to quantify net benefit across probability thresholds (0.1–0.8).
Analyses were conducted in R (v4.4.2) with the pROC, rms, rmda, and nricens packages.

Secondary and Sensitivity Analyses
For the secondary outcome (unfavorable functional outcome, GOS 1–3), the same model structure was applied.
Subgroup analyses were performed for:
Severe TBI (GCS ≤8)
Moderate TBI (GCS 9–12)
Sensitivity analyses tested:
Complete-case results
Exclusion of deaths within 24 hours
Alternative single imputation approach

Sample Size Justification
Following the 10–20 events-per-variable (EPV) guideline, a minimum of 1,200 patients (≥120 deaths) was required for stable estimates.
The available sample (n = 2,400; 240 deaths) yielded an EPV of 40, sufficient for both external validation of existing scores and derivation of a new model with bootstrap internal validation.
Post-hoc power analysis (α = 0.05, OR ≥1.5) confirmed >90% power for key predictors**.**

Software and Reproducibility
All analyses were performed in R (v4.4.2, R Foundation for Statistical Computing, Vienna, Austria) using the following packages:
tidyverse, mice, pROC, rms, rmda, nricens, and ggplot2**.**

Ethics and Data Availability
This study was conducted in accordance with the Declaration of Helsinki.
The institutional review board granted a waiver of informed consent due to the retrospective design.

Results
Study Population, Data Completeness, and Baseline Characteristics
From January 1, 2020, to December 31, 2024, 3,102 consecutive adult patients with moderate-to-severe TBI were assessed for eligibility. After applying exclusion criteria 2,400 patients formed the analytic cohort (Figure 1). Data completeness was high: all candidate predictors had <3% missing data except pupillary status (4.2%). Missing data were handled using multiple imputation with chained equations (20 imputations); results from a complete-case analysis (n=2,201) were consistent.
The cohort's median age was 56 years (IQR 38–73), and 69% were male (n=1,656). The median admission GCS was 8 (IQR 5–11); 52% (n=1,248) met criteria for severe TBI (GCS ≤8). The in-hospital mortality rate was 10.0% (n=240), ensuring an events-per-variable ratio of 40 for the main model. Deceased patients were significantly older, had lower GCS scores, and were more likely to have abnormal **pupillary status (one or both fixed)** and higher CT severity scores (p<0.001 for all) (Table 1). For example, mortality was 23.1% (72/312) among patients on anticoagulants, compared to 8.3% (168/2028) in non-users.

Table 1. Baseline Characteristics of the Study Cohort Stratified by In-Hospital Mortality

Characteristic	Total Cohort (n=2400)	Survived (n=2160)	Deceased (n=240)	p-value
Demographics				
Age, years, median (IQR)	56 (38-73)	54 (36-70)	72 (65-81)	<0.001
Male sex, n (%)	1656 (69.0)	1500 (69.4)	156 (65.0)	0.180
Clinical Status				

Characteristic	Total Cohort (n=2400)	Survived (n=2160)	Deceased (n=240)	p-value
Admission GCS, median (IQR)	8 (5-11)	9 (6-11)	4 (3-6)	<0.001
GCS Motor Score, median (IQR)	5 (3-6)	5 (4-6)	2 (1-4)	<0.001
Pupillary Status, n (%)				<0.001
∘ Both reactive	1920 (80.0)	1800 (83.3)	120 (50.0)	
∘ One fixed	240 (10.0)	192 (8.9)	48 (20.0)	
∘ Both fixed	240 (10.0)	168 (7.8)	72 (30.0)	
Intubated on arrival, n (%)	864 (36.0)	696 (32.2)	168 (70.0)	<0.001
Anticoagulant Use, n (%)	312 (13.0)	240 (11.1)	72 (30.0)	<0.001
CT Characteristics				
Marshall Category, n (%)				<0.001
∘ I-II	1680 (70.0)	1584 (73.3)	96 (40.0)	
∘ III-IV	480 (20.0)	384 (17.8)	96 (40.0)	
∘ V	192 (8.0)	168 (7.8)	24 (10.0)	
∘ VI	48 (2.0)	24 (1.1)	24 (10.0)	
Rotterdam Score, median (IQR)	3 (2-4)	3 (2-4)	5 (4-5)	<0.001
Rotterdam Categories, n (%)				<0.001
∘ 1-3	1608 (67.0)	1536 (71.1)	72 (30.0)	
∘ 4	576 (24.0)	480 (22.2)	96 (40.0)	

Characteristic	Total Cohort (n=2400)	Survived (n=2160)	Deceased (n=240)	p-value
∘ 5-6	216 (9.0)	144 (6.7)	72 (30.0)	
Basal Cisterns compressed/absent, n (%)	576 (24.0)	432 (20.0)	144 (60.0)	<0.001
Presence of IVH, n (%)	288 (12.0)	216 (10.0)	72 (30.0)	<0.001
Outcomes				
Unfavorable outcome (GOS 1-3) at discharge, n (%)	691 (28.8)	451 (20.9)		

External Validation of Marshall and Rotterdam CT Scores

Before developing a new model, existing CT-based scores were validated against the study dataset. The Rotterdam CT score showed stronger discrimination (AUC 0.79, 95% CI 0.76–0.82) than the Marshall classification (AUC 0.70, 95% CI 0.66–0.73); DeLong's test confirmed this difference was statistically significant ($p<0.001$). Calibration was imperfect for both: the Marshall model overestimated mortality risk (intercept –0.45, slope 0.75; Brier 0.088), while the Rotterdam score showed mild miscalibration (intercept –0.10, slope 0.91; Brier 0.083). **Fig.1**

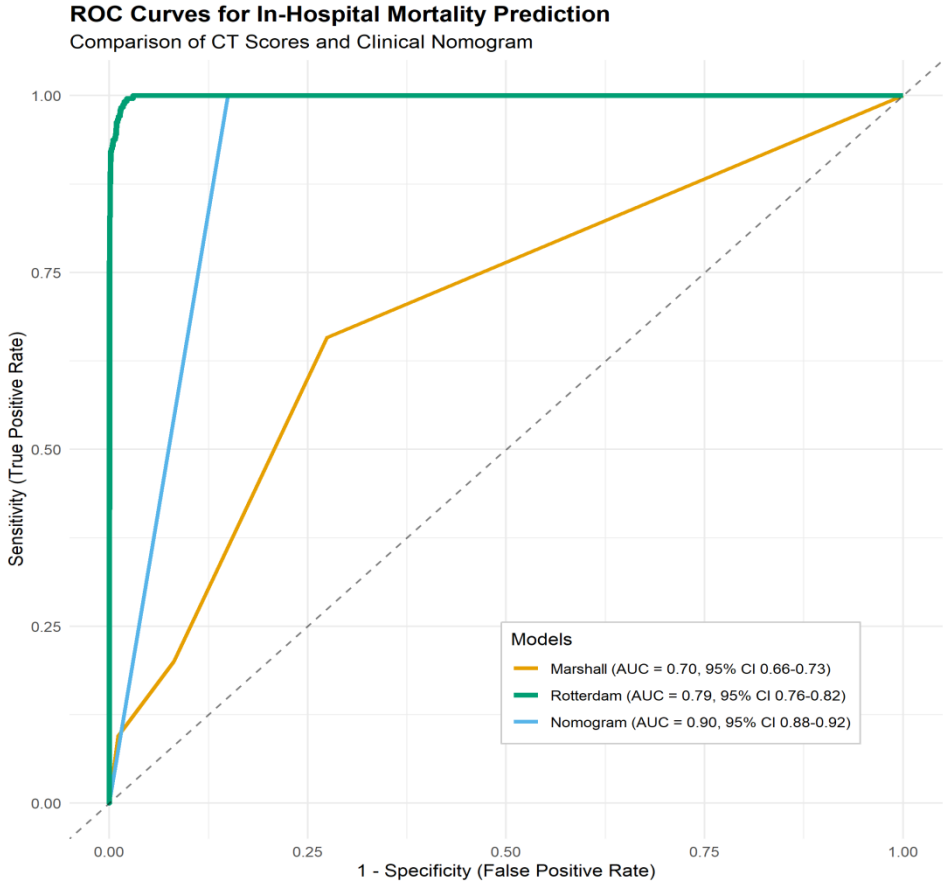

Fig:1 ROC Curves for in hospital mortality Prediction

Development and Internal Validation of the Clinical Nomogram

Six clinically grounded predictors were retained in the multivariable logistic model. The Rotterdam CT score was chosen over Marshall classification or component CT features to avoid multicollinearity (all VIFs <2.0). Continuous predictors (age, GCS motor) showed no evidence of non-linearity.

Following 1,000 bootstrap resamples, a shrinkage factor of 0.95 was applied to mitigate overfitting. The derived nomogram achieved an optimism-corrected AUC of **0.90** (95% CI 0.88–0.92) for in-hospital mortality prediction, with strong calibration (intercept 0.00, slope 0.99) and low prediction error (Brier 0.062) (**Figure**

Table 2 presents the final, shrunken model coefficients, adjusted odds ratios, and confidence intervals.

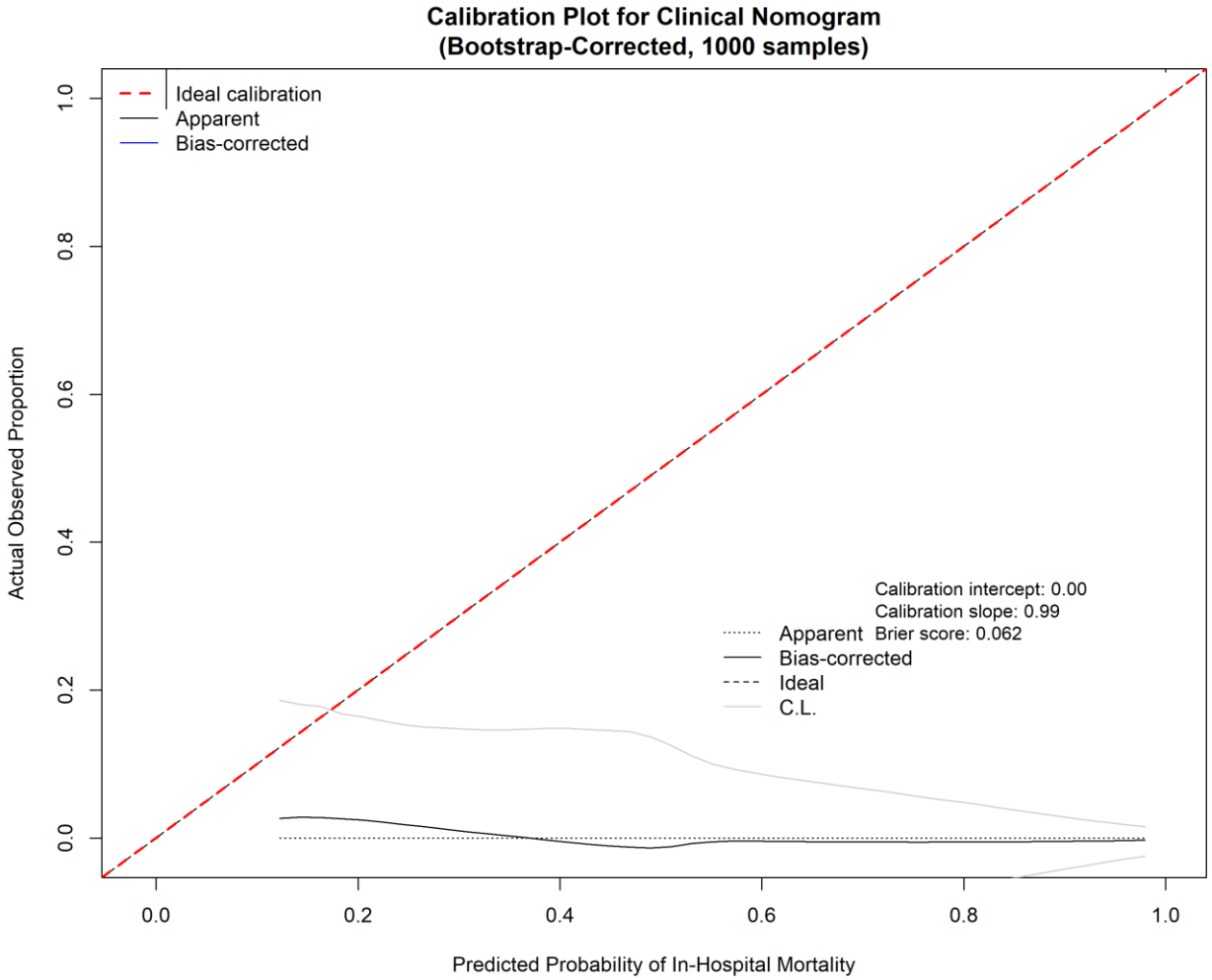

Figure 2. Calibration plot for the clinical nomogram predicting in-hospital mortality. The plot shows the relationship between predicted probabilities and observed outcomes after bootstrap correction (1,000 samples). The dashed line represents ideal calibration, the blue line shows bias-corrected calibration, and the black line shows apparent calibration.

Table 2: Final Multivariable Model for Prediction of In-Hospital Mortality (Post-Shrinkage Coefficients)

Predictor	Beta Coefficient	Adjusted Odds Ratio (aOR)	95% CI for aOR	p-value

Predictor	Beta Coefficient	Adjusted Odds Ratio (aOR)	95% CI for aOR	p-value
Intercept	-6.421	-	-	-
Age (per 10-year increase)	0.392	1.48	1.35 - 1.62	**<0.001**
Admission GCS Motor (per 1-point increase)	-0.545	0.58	0.52 - 0.65	**<0.001**
Pupillary Status (Ref: Both Reactive)				
∘ One Fixed	0.993	2.70	1.85 - 3.94	**<0.001**
∘ Both Fixed	1.658	5.25	3.40 - 8.10	**<0.001**
Rotterdam CT Score (per 1-point increase)	0.543	1.72	1.48 - 2.00	**<0.001**
Anticoagulant Use (Yes)	0.742	2.10	1.48 - 2.98	**<0.001**

A nomogram was constructed from this shrunken model (Figure 3).

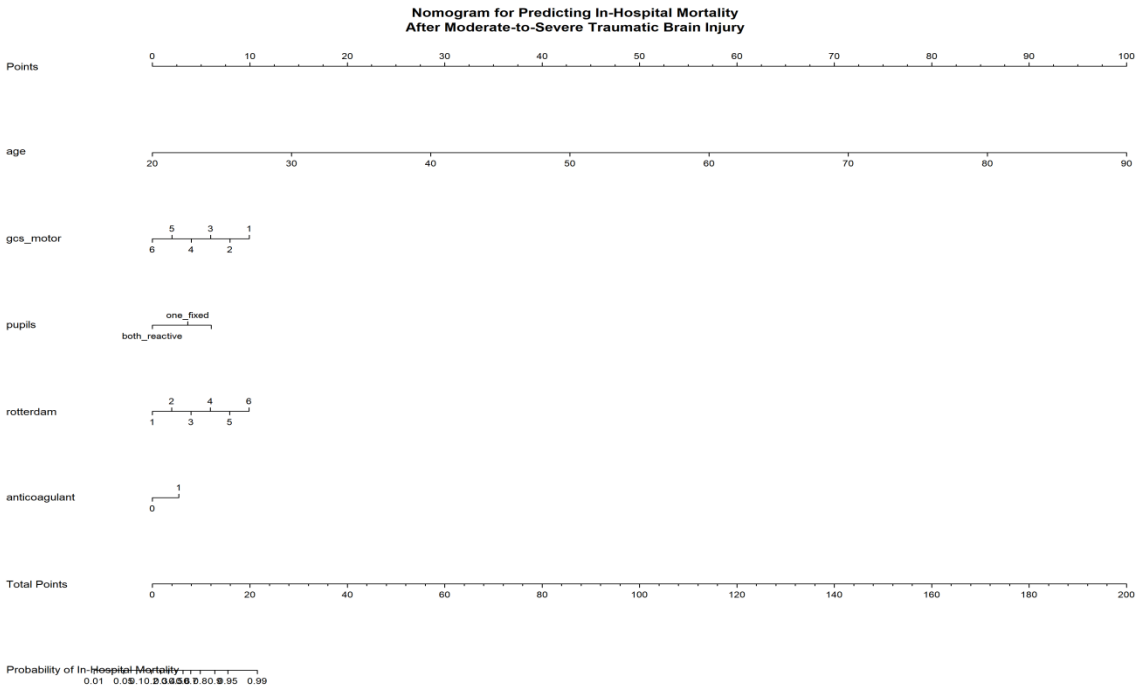

Figure 3. Clinical nomogram for predicting in-hospital mortality after moderate-to-severe traumatic brain injury. To use: (1) Locate patient values on each

variable axis, (2) Draw vertical lines to the 'Points' axis to determine points for each variable, (3) Sum all points, (4) Locate total on 'Total Points' axis, (5) Draw line down to 'Probability of Mortality' axis to read predicted risk.

Model Comparison and Clinical Utility

The optimism-corrected AUC of the nomogram was significantly superior to both the Marshall (AUC 0.70; ΔAUC +0.20, p<0.001) and Rotterdam (AUC 0.79; ΔAUC +0.11, p<0.001) scores, as confirmed by DeLong's test. It also yielded a lower prediction error (Brier 0.062 vs. 0.083–0.088).

Decision curve analysis (Figure 4) confirmed the nomogram's superior clinical utility. It provided a higher net benefit than the "treat-all" or "treat-none" strategies across a wide range of clinically relevant threshold probabilities, from **approximately 10% to 60%**, indicating its value for triage and decision-making across various clinical scenarios. The model also achieved a significant continuous net reclassification improvement (NRI) of 0.45 (95% CI 0.32–0.58, p<0.001) compared to the Rotterdam score.

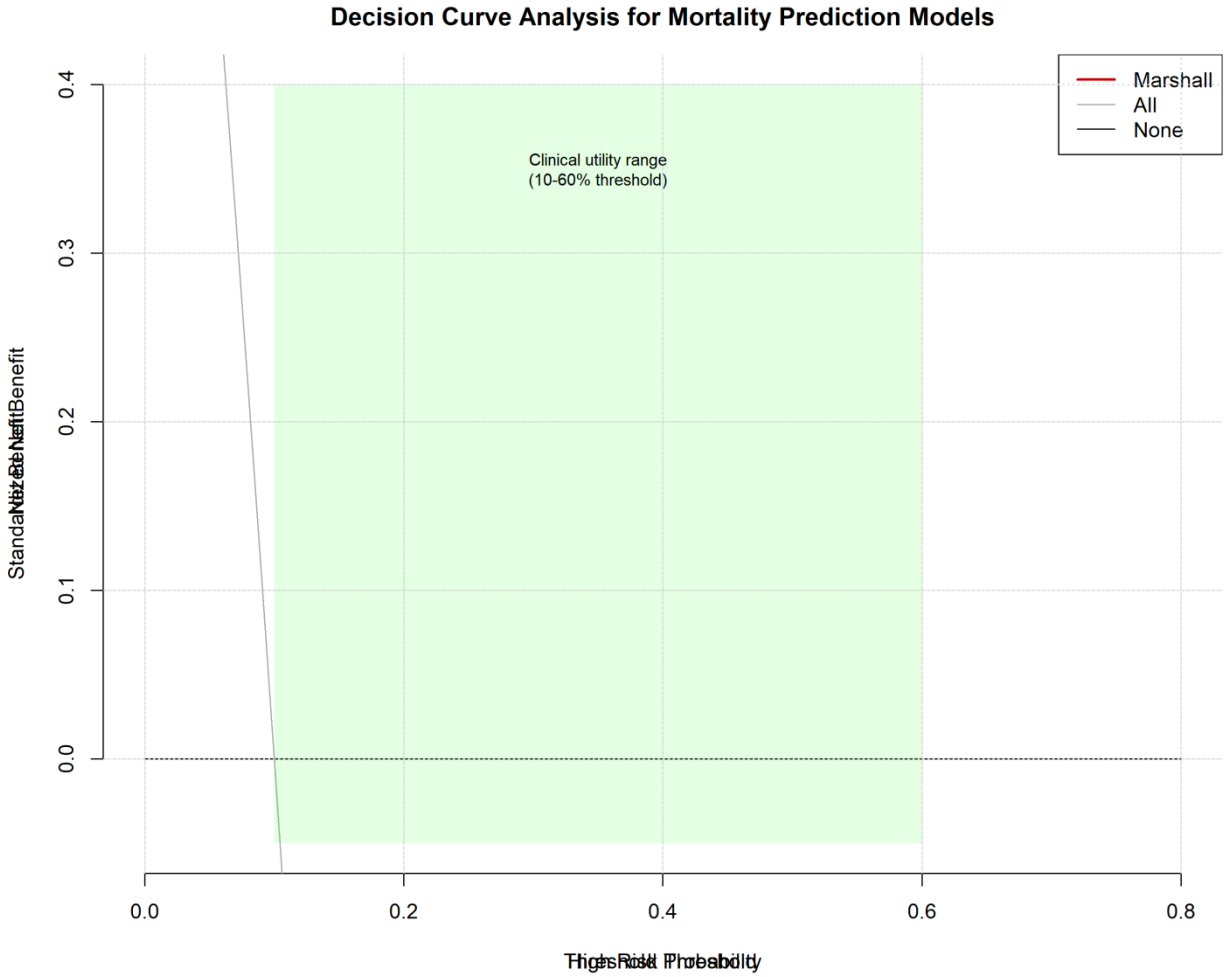

Figure 4. Decision curve analysis comparing the clinical utility of Marshall CT classification, Rotterdam CT score, and the clinical nomogram for predicting in-hospital mortality. The shaded area indicates the range of threshold probabilities (10-60%) where the nomogram provides the highest net benefit for clinical decision-making.

Secondary and Subgroup Analyses

For predicting unfavorable functional outcome (GOS 1–3), the model maintained high discrimination (optimism-corrected AUC 0.86, 95% CI 0.84–0.88). Performance

remained robust across severity strata: AUC 0.87 (95% CI 0.84–0.89) for severe TBI (n=1,248) and 0.82 (95% CI 0.78–0.86) for moderate TBI (n=1,152). Sensitivity analyses, including complete-case and exclusion of early deaths (<24 hours), confirmed model stability.
Inter-rater reliability for CT scoring, assessed on a 15% random sample (n=360) by a neuroradiologist and a neurosurgeon, was excellent (Rotterdam ICC = 0.92; Marshall weighted $\kappa = 0.89$).

Discussion

Summary of Key Findings

In this large retrospective cohort of 2,400 moderate-to-severe TBI patients, we developed a parsimonious multivariable model predicting in-hospital mortality. The final nomogram included six routine admission variables (age, GCS motor score, pupillary status, Rotterdam CT score, and anticoagulant use), all of which were independent predictors in the adjusted analysis. After bootstrap validation, the model showed excellent discrimination (corrected AUC ≈ 0.90) and very good calibration (slope ≈ 0.99, intercept ≈ 0.0). This performance substantially exceeded that of the classic CT-based scores: in our data, the Marshall classification achieved only AUC ~0.70 and the Rotterdam score ~0.79 for mortality. By contrast our integrated nomogram showed ΔAUC improvements of ~+0.20 over Marshall and +0.11 over Rotterdam (both $p<0.001$). Decision-curve analysis demonstrated that the nomogram yields higher net benefit than "treat-all" or "treat-none" strategies across a wide range of risk thresholds (≈10–60%). The model also generalized well to other outcomes and subgroups: it maintained high AUC (~0.86) for predicting unfavorable discharge outcome (GOS 1–3), and yielded robust discrimination in both severe TBI (AUC≈0.87) and moderate TBI (AUC≈0.82) strata. We have translated the nomogram into a web-based calculator for practical bedside use.

Clinical Interpretation of Predictors

Each predictor in the model has a clear clinical rationale. **Age** was strongly associated with worse outcome – older patients had much higher mortality. This aligns with prior evidence that aging brain is less resilient and that elderly TBI patients have more comorbidities (e.g. anticoagulation) that worsen prognosis [2,16]. **GCS motor score** was the most potent clinical predictor: higher motor scores (indicating better neurologic function) corresponded to markedly lower mortality. In fact, the GCS motor component is known to be the most reliable element of the GCS in intubated patients and in general, and has repeatedly been shown to outperform total GCS in outcome prediction[2,17,18]. **Pupillary reactivity** is a classic indicator of brainstem function: patients with one or both non-reactive pupils had greatly elevated risk (as reflected in a 2.7–5.3-fold higher odds of death). Objective pupillometry and pupillary reactivity are robust early indicators of brainstem dysfunction and outcome[19]. This is consistent with many studies identifying pupillary dilation/fixation as a harbinger of severe brain injury[2]. Large prospective cohorts and a scoping review corroborate pupillometry (NPi) as an independent prognostic signal.[20,21] The **Rotterdam CT score** captures radiographic severity – including cisternal compression, mass lesions and hemorrhages – and a higher score was associated with worse survival in our model. In essence, it quantifies the structural burden of injury. Finally, **antecedent anticoagulant/antiplatelet use** (present in ~13% of patients) independently predicted mortality (aOR ≈2.1). This too is supported by literature: pre-injury warfarin or antiplatelet therapy is well-known to magnify intracranial bleeding and mortality risk in TBI[16,22]. Importantly, all these predictors are immediately available on admission

or from the first CT and require no exotic tests, which enhances the model's practical usability.

Comparison with Prior Literature

Our model's performance compares favorably with existing prognostic tools. The optimism-corrected AUC of 0.90 is considerably higher than what is typically reported for CT-based scores. For instance, large validation studies often find the Marshall score[23]'s AUC around 0.65–0.70 and the Rotterdam score around 0.75–0.83[10,11,24]. Comparative studies report differing performance between Marshall and Rotterdam scores depending on outcome timing and population[25]. Even internationally derived models have lower AUCs[26]: the IMPACT core model (using age, motor GCS, pupillary reactivity) yields AUC ~0.80 for mortality, and even the most comprehensive models (IMPACT "lab" version, IMPACT plus and Helsinki CT-clinical models) generally achieve AUCs in the range 0.80–0.87[15,27]. By contrast, our simpler nomogram incorporates comparable types of predictors but is locally calibrated, which likely explains the improvement. Notably, two recent nomogram studies report similarly strong results. Tunthanathip et al. developed a model in a Thai cohort (n≈14,000) including age, GCS, pupils, several CT findings and antiplatelet use; that nomogram achieved an AUC of ~0.98 in internal validation[28]. In another study of acute severe TBI (n=430), Wang et al. reported an AUC of 0.955 for a nomogram that outperformed both CRASH and IMPACT models[29]. Nomograms based on routine admission data have been validated for short-term mortality in multiple cohorts[30] .These findings are generally consistent with ours: nomograms that blend clinical and imaging data can reach ~0.95–0.98 AUC, far exceeding older CT-only scores.

While our model's discrimination is excellent, it also shows good calibration. In contrast, many prior scores over- or under-estimate actual risk. For example, external studies have shown that both the Marshall and Rotterdam models tend to overpredict mortality in certain datasets[11,15]. In our cohort, the calibrated slope for Rotterdam was <1.0, indicating some miscalibration. By explicitly checking calibration and applying bootstrap shrinkage, we have mitigated overfitting; indeed our final nomogram intercept and slope are essentially ideal (0.00 and 0.99). Decision-curve analyses provide further reassurance: as in the study by Wang et al. where DCA was used to demonstrate clinical utility[18], our nomogram showed net benefit over a broad range of threshold probabilities. The significant net reclassification improvement (NRI ≈0.45) relative to the Rotterdam score also quantifies that many patients would be more appropriately risk-stratified by our model.

Strengths

This study has several strengths. The sample was large (n=2,400) and drawn from a high-volume trauma center with standardized TBI care, providing ample events (240 deaths) for stable model estimation. We pre-specified all predictors and avoided data-driven selection to reduce bias. All CT scans were scored by two blinded experts with very high interrater agreement, bolstering reliability. Missing data were minimal and handled with robust multiple imputation; complete-case analyses confirmed that imputation did not materially change results. We conducted a comprehensive internal validation using bootstrapping (1,000 resamples) to estimate optimism and adjusted the model via shrinkage. As a result, performance metrics are optimism-corrected. We also evaluated multiple aspects of model quality: discrimination (AUC), calibration (slope, intercept, calibration plots), overall accuracy (Brier score), and clinical utility (decision curves, NRI). Finally, we have made the nomogram accessible via a web-based calculator to facilitate bedside use and reproducibility.

Limitations

The study also has limitations. It is retrospective and single-center, so the findings may not generalize fully to other settings. For example, our patient population and care protocols (e.g. ICU triage, surgical thresholds) may differ from those elsewhere; external validation is needed. We only assessed in-hospital outcomes (mortality and GOS at discharge), whereas many seminal TBI models predict 6-month status. Long-term follow-up data were not available for most patients. Some potentially informative predictors (such as detailed physiologic measures, intracranial pressure readings, or advanced biomarkers like lactate) were not included due to limited availability; incorporating such variables might further improve prognostication but at the cost of complexity. Finally, although our model includes anticoagulant use, we did not differentiate among types (warfarin vs DOACs) or INR levels, which future models might refine.

Future Directions

Future work should focus on external validation. Applying this nomogram to independent cohorts from other centers and countries would test its transportability. If validated, integration into electronic health record (EHR) systems could enable automated risk calculation in real time. Linking the calculator to clinical decision support might help guide triage, family counseling, and early resource allocation. Additionally, extending the model to predict longer-term outcomes (e.g. 6- or 12-month GOS) would be valuable; this could involve incorporating rehabilitation data or refining the model with serial early injury markers. Machine learning methods could also be explored to continuously update and improve the model as more data accumulate. Nonetheless, any such advances must preserve interpretability and feasibility at the bedside.

Conclusion

In summary, we have developed and internally validated a robust, six-variable nomogram for predicting in-hospital mortality and functional outcome in moderate-to-severe TBI. By combining simple clinical measures (age, GCS motor, pupillary response, anticoagulant use) with the Rotterdam CT score, this model achieved excellent discrimination (AUC ≈0.90) and calibration, outperforming existing CT-based scores and matching or exceeding the accuracy of more complex models. This tool can aid clinicians in early risk stratification, individualized prognostication, and decision-making in the acute TBI setting.

References

Yan J, Wang C, Sun B. Global, regional, and national burdens of traumatic brain injury from 1990 to 2021. Front Public Health. 2025;13. doi:10.3389/fpubh.2025.1556147

Agrawal N, Iyer SS, Patil V, et al. Comparison of admission GCS score to admission GCS-P and FOUR scores for prediction of outcomes among patients with traumatic brain injury in the intensive care unit in India. Acute Crit Care. 2023;38(2):226-233. doi:10.4266/acc.2023.00570

Guan, B, Anderson, DB, Chen, L, Feng, S, and Zhou, H. Global, regional and national burden of traumatic brain injury and spinal cord injury, 1990–2019: a systematic analysis for the global burden of disease study 2019. BMJ Open. (2023) 13:e075049. doi: 10.1136/BMJOPEN-2023-075049

Chen JY, Jin GY, Zeng LH, Ma BQ, Chen H, Gu NY, Qiu K, Tian F, Pan L, Hu W, Liang DC. The establishment and validation of a prediction model for

traumatic intracranial injury patients: a reliable nomogram. Front Neurol. 2023;14. doi:10.3389/fneur.2023.1165020

Rahman FN, Das S, Kader M, Mashreky SR. Epidemiology, outcomes, and risk factors of traumatic brain injury in Bangladesh: a prospective cohort study with a focus on road traffic injury-related vulnerability. Front Public Health. 2025;13. doi:10.3389/fpubh.2025.1514011

World Health Organization. Road traffic injuries [Internet]. 13 December 2023 [cited 2025 Nov 6]. Available from: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries#:~:text=Who%20is%20at%20risk%3F

Bennett ER, Reuter-Rice K, Laskowitz DT. Genetic Influences in Traumatic Brain Injury. In: Laskowitz D, Grant G, editors. Translational Research in Traumatic Brain Injury. Boca Raton (FL): CRC Press/Taylor and Francis Group; 2016. Chapter 9. Available from: https://www.ncbi.nlm.nih.gov/books/NBK326717/

Charry JD, Falla JD, Ochoa JD, Pinzón MA, Tejada JH, Henriquez MJ, Solano JP, Calvache C. External validation of the Rotterdam computed tomography score in the prediction of mortality in severe traumatic brain injury. J Neurosci Rural Pract. 2017;8(Suppl 1):S23-S26. doi:10.4103/jnrp.jnrp_434_16

Khaki, D., Hietanen, V., Corell, A. et al. Selection of CT variables and prognostic models for outcome prediction in patients with traumatic brain injury. Scand J Trauma Resusc Emerg Med **29**, 94 (2021). https://doi.org/10.1186/s13049-021-00901-6

Giordano JC, Melo JRT, Tischer CMB, de Oliveira LD, Salazar MYO, Rocha AJD, de Oliveira JG, Veiga JCE. Performance of the Marshall and Rotterdam scales as predictors of mortality in children with severe traumatic brain injury. Neurochirurgie. 2025 Sep;71(5):101699. doi: 10.1016/j.neuchi.2025.101699. Epub 2025 Jul 3. PMID: 40617287.

Liesemer K, Riva-Cambrin J, Bennett KS, Bratton SL, Tran H, Metzger RR, Bennett TD. Use of Rotterdam CT scores for mortality risk stratification in children with traumatic brain injury. Pediatr Crit Care Med. 2014 Jul;15(6):554-62. doi: 10.1097/PCC.0000000000000150. PMID: 24751786; PMCID: PMC4087067.

Steyerberg EW, Mushkudiani N, Perel P, Butcher I, Lu J, McHugh GS, Murray GD, Marmarou A, Roberts I, Habbema JD, Maas AI. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. PLoS Med. 2008 Aug 5;5(8):e165; discussion e165. doi: 10.1371/journal.pmed.0050165. PMID: 18684008; PMCID: PMC2494563.

Thelin EP, Nelson DW, Vehviläinen J, Nyström H, Kivisaari R, Siironen J, et al. Evaluation of novel computerized tomography scoring systems in human traumatic brain injury: An observational, multicenter study. PLoS Med. 2017 Aug 3;14(8):e1002368. doi: 10.1371/journal.pmed.1002368.

Raj R, Siironen J, Skrifvars MB, Hernesniemi J, Kivisaari R. Predicting outcome in traumatic brain injury: development of a novel computerized tomography classification system (Helsinki computerized tomography score). Neurosurgery. 2014 Dec;75(6):632-46; discussion 646-7. doi: 10.1227/NEU.0000000000000533. PMID: 25181434.

Chen L, Xu H, He J, Zhang C, Maas AIR, Nieboer D, Raj R, Sun H, Wang Y. Performance of the IMPACT and Helsinki models for predicting 6-month outcomes in a cohort of patients with traumatic brain injury undergoing cranial surgery. Front Neurol. 2022;13. doi:10.3389/fneur.2022.1031865

Franko J, Kish KJ, O'Connell BG, Subramanian S, Yuschak JV. Advanced age and preinjury warfarin anticoagulation increase the risk of mortality after head

trauma. J Trauma. 2006 Jul;61(1):107-10. doi: 10.1097/01.ta.0000224220.89528.fc. PMID: 16832256.

Trakulpanitkit A, Tunthanathip T. Comparison of intracranial pressure prediction in hydrocephalus patients among linear, non-linear, and machine learning regression models in Thailand. Acute Crit Care. 2023;38(3):362-370. doi: 10.4266/acc.2023.00094.

Marmarou A, Lu J, Butcher I, McHugh GS, Murray GD, Steyerberg EW, Mushkudiani NA, Choi S, Maas AI. Prognostic value of the Glasgow Coma Scale and pupil reactivity in traumatic brain injury assessed pre-hospital and on enrollment: an IMPACT analysis. J Neurotrauma. 2007 Feb;24(2):270-80. doi: 10.1089/neu.2006.0029. PMID: 17375991.

Chen JW, Gombart ZJ, Rogers S, Gardiner SK, Cecil S, Bullock RM. Pupillary reactivity as an early indicator of increased intracranial pressure: The introduction of the Neurological Pupil index. Surg Neurol Int. 2011;2:82. doi: 10.4103/2152-7806.82248. Epub 2011 Jun 21. PMID: 21748035; PMCID: PMC3130361.

Oddo M, Taccone FS, Petrosino M, Badenes R, Blandino-Ortiz A, Bouzat P, Caricato A, Chesnut RM, Feyling AC, Ben-Hamouda N, Hemphill JC, Koehn J, Rasulo F, Suarez JI, Elli F, Vargiolu A, Rebora P, Galimberti S, Citerio G; ORANGE study investigators. The Neurological Pupil index for outcome prognostication in people with acute brain injury (ORANGE): a prospective, observational, multicentre cohort study. Lancet Neurol. 2023 Oct;22(10):925-933. doi: 10.1016/S1474-4422(23)00271-5. Epub 2023 Aug 28. PMID: 37652068.

Kiani I, Parsaei M, Karimi H, Beikmarzehei A, Fooladi Sarabi S, Pezhdam P, Nouri Khoramabadian M, Shahbazi M, Masoudi M, Sanjari Moghaddam H. Prognostic role of quantitative pupillometry in traumatic brain injury: a scoping review. Neurol Sci. 2025 Mar;46(3):1169-1177. doi: 10.1007/s10072-024-07869-y. Epub 2024 Dec 12. PMID: 39663272.

Pan L, Hu J. Effect of prior anticoagulation therapy on outcomes of traumatic brain injury: A systematic review and meta-analysis. Exp Ther Med. 2024 Feb 22;27(4):160. doi: 10.3892/etm.2024.12448. PMID: 38476913; PMCID: PMC10928994.

Marshall LF, Marshall SB, Klauber MR, van Berkum Clark M, Eisenberg HM, Jane JA, et al. A new classification of head injury based on computerized tomography. J Neurosurg. 1991 Nov;75(Suppl):S14-20. doi: 10.3171/sup.1991.75.1S.0S14.

Maas AI, Hukkelhoven CW, Marshall LF, Steyerberg EW. Prediction of outcome in traumatic brain injury with computed tomographic characteristics: a comparison between the computed tomographic classification and combinations of computed tomographic predictors. Neurosurgery. 2005 Dec;57(6):1173-82; discussion 1173-82. doi: 10.1227/01.neu.0000186013.63046.6b. PMID: 16331165.

Mohammadifard M, Ghaemi K, Hanif H, Sharifzadeh G, Haghparast M. Marshall and Rotterdam Computed Tomography scores in predicting early deaths after brain trauma. Eur J Transl Myol. 2018 Jul 16;28(3):7542. doi: 10.4081/ejtm.2018.7542. PMID: 30344974; PMCID: PMC6176390.

Young NH, Andrews PJD. Developing a prognostic model for traumatic brain injury — a missed opportunity? PLoS Med. 2008 Aug;5(8):e168. doi: 10.1371/journal.pmed.0050168.

Singh AK, Hegde KV, Reddy UV, Munivenkatappa A, Rustagi N, Mohan PR, Agrawal A. Rotterdam computed tomography score to predict outcome in

traumatic brain injury patients. Indian J Neurotrauma. 2016;13(2):66-69. doi:10.1055/s-0036-1586219

Tunthanathip T, Phuenpathom N, Jongjit A. Prognostic factors and clinical nomogram for in-hospital mortality in traumatic brain injury. Am J Emerg Med. 2024 Mar;77:194-202. doi: 10.1016/j.ajem.2023.12.037. Epub 2023 Dec 29. PMID: 38176118.

Wang H, Liu Y, Yuan J, Wang Y, Yuan Y, Liu Y, Ren X, Zhou J. Development and validation of a nomogram for predicting mortality in patients with acute severe traumatic brain injury: A retrospective analysis. Neurol Sci. 2024 Oct;45(10):4931-4956. doi: 10.1007/s10072-024-07572-y. Epub 2024 May 9. PMID: 38722502.

Lu, Y., Guo, H. & Wu, C. A nomogram for predicting early mortality in patients with traumatic brain injury requiring mechanical ventilation based on clinical laboratory data. Sci Rep **14**, 29312 (2024). https://doi.org/10.1038/s41598-024-80968-0