

ORIGINAL ARTICLE

Modified ABSI Versus Revised Baux for In-Hospital Burn Mortality at an Indonesian Burn Unit: A Paired Accuracy and Calibration Study

Ilham Taufan Ikramiawan^{1*}, Abda Arif², Theodorus³

¹ Department of Surgery, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

² Department of Plastic Surgery, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

³ Department of Pharmacology, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

ARTICLE INFO

Article history:

Received: April 21, 2026

Revised: May 22, 2026

Accepted: June 18, 2026

Available Online: June 25, 2026

Published: June 28, 2026

Keywords:

Burns; Revised Baux score; Abbreviated Burn Severity Index; Hospital mortality; Triage

Corresponding author:

Ilham Taufan Ikramiawan

Email:

ilhamtaufan0@gmail.com

DOI:

<https://doi.org/10.37275/sjs.v9i1.175>

Access licence:

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

Copyright:

© 2026 The Authors. Authors retain copyright and grant the journal the right of first publication. Published by HM Publisher.

ABSTRACT

Background: The Modified Abbreviated Burn Severity Index (mABSI) and the Revised Baux score (R-Baux) are bedside scores for predicting death after burn injury, but they have not been compared in the same patients in Sumatra.

Objective: To compare the discrimination, accuracy, agreement and calibration of the two scores for in-hospital mortality.

Methods: Paired retrospective diagnostic-accuracy study of 45 consecutive eligible burn admissions in 2025 at Dr. Mohammad Hoesin General Hospital, Palembang. Both scores were derived from admission data; the reference standard was in-hospital death. We estimated areas under the receiver operating characteristic curve (AUC), accuracy at Youden-optimal cutoffs, paired McNemar tests, agreement, logistic risk curves and calibration against the published mABSI risk categories. Paired analyses needing patient-level linkage are reported as ranges compatible with the observed tables.

Results: Twenty patients (44.4%) died. AUCs were similar: 0.811 (95% CI 0.690–0.932) for the mABSI, 0.808 (95% CI 0.679–0.937) for the R-Baux and 0.803 for burn size alone, with no paired difference ($p \geq 0.922$; differences >0.15 excluded). The Youden-optimal cutoff was ≥ 6 for the mABSI (sensitivity 75.0%, specificity 72.0%) and ≥ 81 for the R-Baux (70.0% and 84.0%). Deaths were 2.36 times those expected from the published mABSI categories, and at scores of 4–5 mortality was 27.8% rather than 2%. Predicted mortality reached 50% at an mABSI of 6.3 and an R-Baux of about 85.

Conclusion: Both scores ranked mortality equally well but added nothing detectable to burn size, and the published categories underestimated local risk. Locally recalibrated estimates may suit triage but remain provisional pending multicenter validation.

1. Introduction

Burn injury remains a large and unevenly distributed source of death and disability. An analysis of the Global Burden of Disease 2021 data estimated about 15 million new burn injuries in that year, with the disability burden inversely related to the socio-demographic index,¹ and projections for 195 countries anticipate 10.3 million cases in 2030 and 10.9 million in 2050, with disability-adjusted life-years rising from 11.5 to 14.1 million and the burden concentrated in South Asia and sub-Saharan Africa.² Hospital mortality differs by an order of magnitude between health systems. It was 4.3% among adults in a population cohort from Catalonia³ and 3% across a meta-analysis of 465,433 patients treated in Chinese hospitals,⁴ whereas referral centers in middle-income settings report 21.8% in Karnataka,⁵ 29% in Rawalpindi,⁶ 32.2% in the Dominican Republic⁷ and 39% in Damascus.⁸ Indonesian referral units sit at the upper end of this range: in-hospital mortality was 42.2% among 595 adults treated at the national referral burn center in Jakarta,⁹ 27.8% at Dr. Soetomo Hospital in Surabaya¹⁰ and 25% at an Indonesian emergency center.¹¹

Deaths after burn injury follow two overlapping phases, and both depend on decisions taken in the first hours after admission. Early deaths reflect burn shock and the futility of care in the largest burns; in a British series of patients with more

than 50% of total body surface area (TBSA) burned, mortality was 71.9% and 71.0% of deaths occurred within 48 hours.¹² Later deaths are driven by infection and organ failure: in a British intensive-care cohort multiple-organ failure peaked on day 5 and sepsis on day 10,¹³ sepsis affected 22.9% of older adults in a Canadian two-center cohort with a median onset on day 11,¹⁴ mortality was 40.7% among Chinese patients who developed acute kidney injury against 2.6% among those who did not,¹⁵ and gastrointestinal dysfunction carried an odds ratio of 14.07 for 90-day death.¹⁶ Inhalation injury amplifies these pathways; injury graded severe at bronchoscopy nearly triples the odds of death,¹⁷ depresses cardiac index during early resuscitation¹⁸ and doubles the risk of ventilator-associated pneumonia.¹⁹ Surgical timing matters as well: in a French prospective multicenter cohort, delaying excision beyond 48 hours in burns over 20% TBSA was associated with more than twice the odds of 90-day death,²⁰ and in burns of 70% TBSA or more a first tangential excision after day 3 was associated with a hazard ratio of 2.37.²¹ An accurate estimate of risk at admission therefore shapes the intensity of resuscitation, the timing of surgery and the allocation of scarce intensive-care beds.

Two bedside scores dominate this task. The Revised Baux score (R-Baux) adds age in years to the percentage of TBSA burned and adds 17 points for inhalation injury; it was derived

in 39,888 patients of the American National Burn Repository.²² A meta-analysis of 21 validation studies gave a summary area under the receiver operating characteristic curve (AUC) of 0.93, with weaker accuracy at the extremes of age,²³ and a meta-analysis of nine risk models found the R-Baux (pooled AUC 0.92), the original Baux score (0.92) and the FLAMES score (Fatality by Longevity, APACHE II score, Measured Extent of burn, and Sex; 0.94) statistically indistinguishable.²⁴ The Abbreviated Burn Severity Index (ABSI) sums points for age, sex, inhalation injury, full-thickness burn and TBSA; pooled across 16 studies, higher ABSI values were associated with death (odds ratio [OR] 1.72; 1.59 per point when analyzed continuously).²⁵ In 14,984 patients of the German Burn Registry, Bartels and colleagues found that sex no longer predicted death and that the age relationship was non-linear, and they derived the Modified ABSI (mABSI), which drops sex, re-weights age and keeps the published threat-to-life categories.^{26,28} In the same registry the mABSI reached an AUC of 0.928 against 0.939 for the R-Baux,²⁷ and across seven sources on five continents it fitted survival well in countries with a very high or high Human Development Index but less well than the original ABSI in middle- and low-index countries.²⁸ Full-thickness involvement, the one variable the mABSI adds to the R-Baux, was irrelevant to survival below 20% TBSA in German registry data and became important only above it.²⁹

Evidence from resource-limited settings is less consistent. In 504 intensive-care patients in Punjab, India, the mABSI was more specific and the R-Baux more sensitive, but the R-Baux over-predicted mortality.³⁰ Reported R-Baux cutoffs range from 74.5 in Uganda³¹ and 80–81.5 in West Virginia and Ethiopia^{32,33} to 105.5 in north-western Syria,³⁴ and discrimination falls where burns are large and mortality high: the R-Baux and the ABSI both had an AUC of 0.79 among Croatian patients with at least 20% TBSA³⁵ and about 0.75 among adults with 30–60% TBSA in New Delhi.³⁶ Indonesian data are limited to Java. In Surabaya the ABSI and the R-Baux had near-identical accuracy (sensitivity 82.2% and 80.8%, specificity 82.1% and 80.0%, respectively),¹⁰ a second Surabaya series found the R-Baux, the Belgian Outcome of Burn Injury score and the ABSI similar,³⁷ the ABSI performed best among four scores at an Indonesian emergency center,¹¹ and in Jakarta burn size alone discriminated death with an AUC of 0.813.⁹

Dr. Mohammad Hoesin General Hospital (RSMH) in Palembang is the referral burn unit for South Sumatra. Its patients reach the unit after referral pathways and pre-hospital

delays that differ from those of the registries in which both scores were derived, so neither the ranking ability of the scores nor the absolute risks attached to them can be assumed to transfer. To our knowledge, no study has compared the mABSI and the R-Baux in the same patients outside Java, none has tested the published mABSI risk categories in an Indonesian unit, and none has asked whether either composite score adds information beyond burn size in a cohort dominated by large flame burns in young adults. We therefore compared the two scores in the same patients at RSMH, with four aims: to estimate the discrimination and accuracy of each score for in-hospital death; to compare them directly and quantify their agreement; to test the calibration of the published mABSI risk categories and derive local risk curves; and to determine whether either score improves on burn size alone. We hypothesized that the two scores would discriminate similarly and that the mABSI, which adds burn depth, would be the more sensitive.

2. Methods

Study design and setting

This was a retrospective, single-center, paired diagnostic-accuracy study in which both index tests were applied to every patient and compared against a single reference standard. It was conducted in the Burn Unit of the Department of Surgery at RSMH, Palembang, South Sumatra, Indonesia, the tertiary teaching hospital of the Faculty of Medicine, Universitas Sriwijaya, and the referral burn center for the province. The unit provides burn-shock resuscitation, surgical excision and grafting, intensive care, nutrition support and inpatient rehabilitation within the plastic and reconstructive surgery service. The study is reported in accordance with the STARD 2015 statement for diagnostic-accuracy studies, and the observational design follows the STROBE checklist.

Participants

The medical records of all patients admitted with an acute burn injury between 1 January and 31 December 2025 were screened. Patients were eligible if they had an admission diagnosis of acute burn of any etiology, complete admission data for age, %TBSA, inhalation injury and burn depth, and a documented in-hospital outcome. Patients with concurrent major trauma were excluded because neither score was derived for combined injuries, as were records too incomplete to compute either score. Of 92 admissions, 18 were excluded for concurrent major trauma and 29 for incomplete records, leaving 45 patients for analysis, as shown in Figure 1.

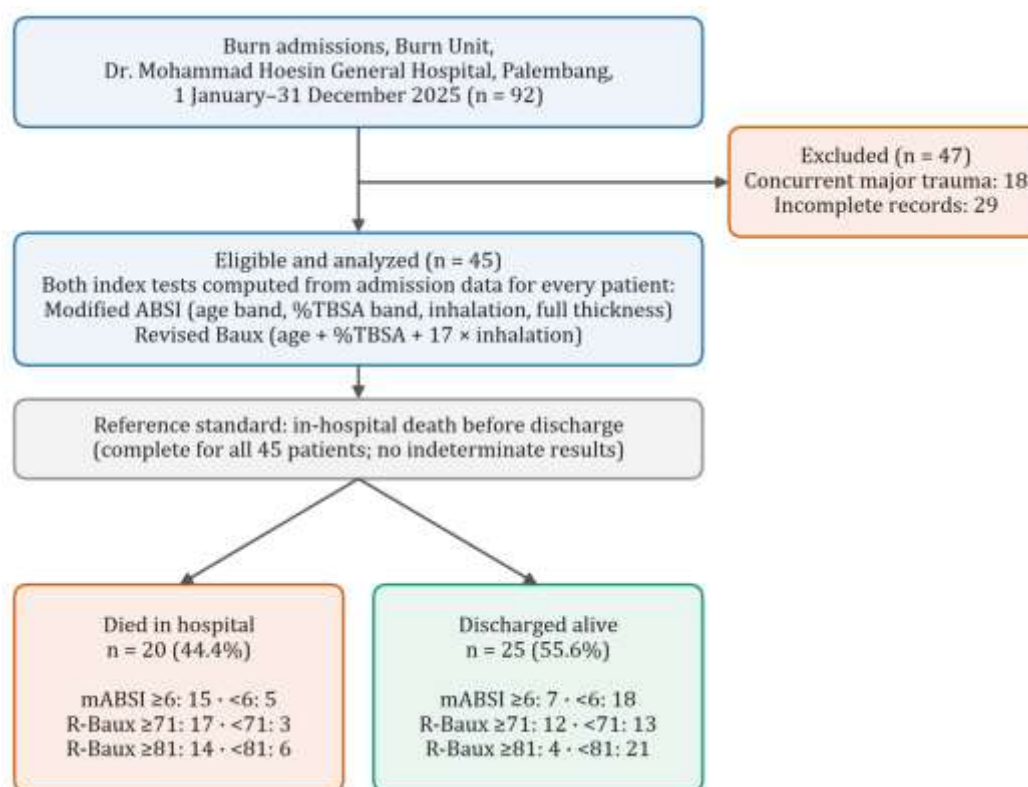


Figure 1. STARD flow of patients through the study. Both index tests were derived for every analyzed patient from admission data; the reference standard was in-hospital death. Counts below each outcome box give the classification by the dichotomized scores.

Burn care during the study period

Care followed the unit's burn protocol. Airway assessment followed an ABCDE sequence, and suspected inhalation injury — closed-space exposure, facial or oropharyngeal burns, carbonaceous sputum, hoarseness or stridor — prompted early intubation, with fiberoptic bronchoscopy when available. Fluid resuscitation used the Parkland formula (4 mL × kg × %TBSA of Ringer's lactate over 24 hours, half in the first 8 hours from injury), titrated to a urine output of 0.5 mL/kg/h in adults and adjusted for inhalation injury, electrical injury and delayed presentation. Escharotomy was performed urgently for circumferential full-thickness burns threatening perfusion or ventilation. Tangential excision with split-thickness autografting was the default operation for deep dermal and full-thickness burns, usually within the first week, with staged excision and temporary cover when donor sites were limited. Enteral nutrition targeted 25–30 kcal/kg/day and 1.5–2.0 g of protein/kg/day, silver-based dressings were changed every 24–48 hours, and patients received physiotherapy and occupational therapy.

Index tests

The mABSI was calculated as the sum of four component scores:²⁶

$$\text{mABSI} = S_{\text{age}} + S_{\text{TBSA}} + S_{\text{inh}} + S_{\text{FT}} \quad (1)$$

where S_{age} is 1 point for age 0–40 years, 2 for 41–70 years, 3 for 71–80 years, 4 for 81–90 years and 5 for 91–100 years; S_{TBSA} is 1 point per 10% band of TBSA (1–10% = 1 to 91–100% = 10); and S_{inh} and S_{FT} add 1 point each for inhalation injury and for any full-thickness burn, giving a possible range of 2–17. The published threat-to-life categories map totals of 2–3, 4–5, 6–7, 8–9, 10–11 and 12 or more to probabilities of survival of at least 99%, 98%, 80–90%, 50–70%, 20–40% and 10% or less.^{26,28} The R-Baux was calculated as:²²

$$\text{R-Baux} = \text{age} + \% \text{TBSA} + 17 \times R \quad (2)$$

where R is 1 when inhalation injury is present and 0 otherwise. Age was taken from the admission record. %TBSA was estimated with the Lund–Browder chart in patients younger than 15 years and with the rule of nines in older patients. Burn depth was assessed by the attending burn surgeon at admission and refined at the first excision, and the depth range of each burn (from its shallowest to its deepest component) was recorded; a full-thickness burn was any burn whose deepest component was full thickness. Inhalation injury was diagnosed clinically and confirmed by bronchoscopy when performed. Both scores were computed from admission documentation by the first author without reference to the discharge outcome, and a random 10% of records was abstracted a second time by the second author with complete agreement. The mABSI was recorded as an integer and the R-Baux in 10-point bands (20 or less, 21–30, and so on up to 141–150). Among patients who died, the component points summed to two more than the recorded totals (152 against 150), indicating a one- or two-point scoring discrepancy in at most two records; the recorded totals — the scores as applied — were analyzed, and the sensitivity analysis allowed explicitly for the discrepancy.

Reference standard

The reference standard was in-hospital death from any cause before discharge. Patients transferred elsewhere were classified by their final RSMH disposition, and deaths within 24 hours of admission were retained, because excluding them would remove the patients in whom early stratification is most consequential; the number of such patients, the timing of death and the length of stay were not recorded, so the interval between the index tests (admission data) and the reference standard was the whole hospital stay. The outcome was complete for all analyzed patients, and there were no indeterminate index-test results.

Sample size

All eligible admissions in the study year were included. The planning calculation at protocol stage assumed a sensitivity of 97% estimated with a precision of $\pm 5\%$:

$$n = \frac{z^2 \cdot P(1-P)}{\Delta^2} = \frac{1.96^2 \times 0.97 \times 0.03}{0.05^2} \approx 45 \quad (3)$$

where z is the standard normal deviate, P the anticipated sensitivity and Δ the precision. Because this formula gives the number of deaths, not of patients, needed for the sensitivity estimate, it would require about 101 patients at the observed mortality; the achieved precision is therefore reported for every estimate, and a sample-size analysis for future studies is given in the Discussion.

Statistical analysis

Analyses were performed in Python 3.14 with NumPy 2.4, SciPy 1.18 (including the HiGHS integer-programming solver) and matplotlib 3.10, using purpose-written exact routines. Categorical variables are reported as counts and percentages with Wilson score intervals, which retain close-to-nominal coverage in small samples:

$$CI = \frac{x + \frac{z^2}{2} \pm z \sqrt{\frac{x(n-x)}{n} + \frac{z^2}{4}}}{n + z^2} \quad (4)$$

where x is the number of events and n the denominator. Associations with death were tested with Fisher's exact test for 2×2 tables and the Fisher–Freeman–Halton exact test for larger tables, with odds ratios (OR, Woolf 95% CI, Haldane correction for empty cells), Newcombe risk differences and Cramér's V ; trends across ordered categories used the exact Cochran–Armitage test. Holm's method adjusted the seven baseline comparisons. Score distributions were compared with the exact permutation Mann–Whitney test with mid-ranks, and the rank-biserial correlation was reported as the effect size.

Discrimination was summarized by the empirical AUC with the DeLong variance, a logit-transformed interval and a bootstrap percentile interval (20,000 resamples). For every possible cutoff we computed sensitivity (Sn), specificity (Sp), predictive values, likelihood ratios with log-method intervals, accuracy and the Youden index:

$$J = Sn + Sp - 1, \quad LR+ = \frac{Sn}{1-Sp}, \quad LR- = \frac{1-Sn}{Sp} \quad (5)$$

Two pairs of cutoffs were analyzed: the dichotomized classification used in the unit (mABSI ≥ 6 and R-Baux ≥ 71 , the 71–80 band and above) and the Youden-optimal cutoffs. The R-Baux threshold of 71 is a band boundary rather than a validated clinical cutoff; it was analyzed because the unit classified patients jointly by both scores at these thresholds. The stability of the Youden-optimal cutoffs was examined by repeating the selection in each of 20,000 bootstrap resamples, and the optimism of the accuracy at the selected cutoffs was estimated with Harrell's bootstrap procedure, in which the cutoff selected in each resample is applied back to the original data. Because the R-Baux and %TBSA were recorded in bands, the attainable AUC of the raw scores was bounded by resolving every within-band tie for and against the score. Paired sensitivities and specificities were compared with exact McNemar tests and Newcombe's hybrid-score interval for paired proportions, and agreement between the dichotomized scores with Cohen's κ (Fleiss–Cohen–Everitt large-sample standard error), the prevalence- and bias-adjusted κ (PABAK) and positive and negative percent agreement. The exact McNemar test was the pre-specified paired test; Newcombe intervals are reported as estimates of the paired differences. Paired AUCs were compared with the DeLong test:

$$\text{Var}(A_1 - A_2) = \frac{S_{10}}{m} + \frac{S_{01}}{n} \quad (6)$$

where A_1 and A_2 are the two AUCs, m and n the numbers of deaths and survivors, and $S_{10} = S_{10,11} + S_{10,22} - 2S_{10,12}$ and $S_{01} = S_{01,11} + S_{01,22} - 2S_{01,12}$ combine the variances (S_{11} , S_{22}) and covariance (S_{12}) of the DeLong structural components of the two scores among deaths (10) and survivors (01). Death was modeled on each score with logistic regression to obtain local risk curves:

$$\text{logit}(\text{Pr}(\text{death})) = \beta_0 + \beta_1 x, \quad \text{LD}_{50} = -\frac{\beta_0}{\beta_1} \quad (7)$$

where x is the score and LD_{50} the score at which predicted mortality reaches 50% (delta-method CI). Model performance was described by the likelihood-ratio χ^2 , Nagelkerke R^2 and the Brier score. Calibration was assessed with a Hosmer–Lemeshow statistic over the observed score levels with a parametric-bootstrap p value (20,000 replicates), which has little power with 20 deaths and was not taken as evidence of good fit, and with the optimism-corrected calibration slope from Harrell's bootstrap (2,000 resamples). Firth-penalized estimates were checked for small-sample bias, and predicted mortality with 95% CIs was tabulated at each score level. Calibration of the published mABSI categories was assessed with one-sided exact binomial tests of observed deaths against the upper bound of each category's expected mortality, an overall observed/expected ratio with an exact Poisson interval, stratum-specific likelihood ratios, and logistic recalibration of the published probabilities (category midpoints), which gives a calibration-in-the-large intercept (slope fixed at 1) and a calibration slope. Clinical utility was assessed with decision curves:

$$\text{Net benefit} = \frac{TP}{N} - \frac{FP}{N} \times \frac{p_t}{1-p_t} \quad (8)$$

where TP and FP are the true and false positives at threshold probability p_t among N patients. Mortality and AUCs were compared descriptively with published cohorts, reporting the differences and, for mortality, Cohen's h . All tests were two-sided unless stated, with $\alpha = 0.05$, and p values are reported to three decimal places. Holm adjustment was applied to the baseline comparisons; comparisons of secondary accuracy indices are exploratory and were not adjusted for multiplicity.

Sensitivity analysis over data sets consistent with the observations

The records provided the full cross-classification of each score, of %TBSA and of every component with the outcome, and the joint dichotomized classification of the two scores; because no patient was positive on the mABSI and negative on the R-Baux, this joint classification is fully determined within each outcome group. The data did not record which mABSI value and which R-Baux band belonged to the same patient beyond that classification, nor the cross-classification of the components with one another. These quantities are needed for the paired DeLong test, for paired comparisons at other cutoffs and for multivariable models. No patient-level value was imputed. Instead, an admissible data set was defined as any assignment of age band, %TBSA band, inhalation injury, full thickness, recorded mABSI and R-Baux band to the patients of each outcome group that reproduces every observed marginal distribution, the observed dichotomized classification and the mABSI arithmetic (allowing the two-point recording discrepancy among deaths), and in which each patient's R-Baux band is attainable from that patient's age band, %TBSA band and inhalation status. Linear quantities — the DeLong covariance and the counts behind paired tests — were minimized and maximized exactly by integer linear

programming. Model coefficients were explored over 3,000 randomly oriented admissible data sets plus 24 constructed extremes, each fitted by Firth-penalized logistic regression without discarding any fit; extremes are reported as the most extreme values found. A conclusion holding in every admissible data set was treated as robust, and a range spanning the null as a finding that the data cannot identify the effect.

Ethical approval

The study protocol received an ethical exemption from the Health Research Ethics Committee of Dr. Mohammad Hoesin General Hospital Palembang (No. DP.04.03/D.XVIII.06.08/ETIK/033/2026) on 30 January 2026, before data abstraction began in February 2026. Individual informed consent was waived because only de-identified existing medical records were used. The study was conducted in accordance with the Declaration of Helsinki, and each record was assigned a study identifier before analysis.

3. Results

Participants

Of the 92 burn admissions, 45 patients met the eligibility criteria, as shown in Figure 1. Twenty died in hospital, giving a mortality of 44.4% (95% CI 30.9–58.8). Most patients were male (35/45, 77.8%), and the mean age was 40.8 ± 19.6 years (range 1–74); 23 patients (51.1%) were aged 0–40 years, 21 (46.7%) 41–70 years and one (2.2%) 71–80 years. Flame burns accounted for 29 admissions (64.4%), electrical burns for 9 (20.0%), scalds for 5 (11.1%) and chemical burns for 2 (4.4%). Inhalation injury was documented in 6 patients (13.3%) and a full-thickness component in 16 (35.6%). The characteristics of patients who died and of those discharged alive, together with the mortality within each characteristic, are detailed in Table 1.

Table 1. Characteristics of the 45 patients with burns and in-hospital mortality by characteristic

Characteristic	Died (<i>n</i> = 20)	Discharged alive (<i>n</i> = 25)	Total (<i>n</i> = 45)	Mortality, % (deaths/ <i>n</i>)	Exact <i>p</i> (Holm)	Effect size
All patients	20	25	45	44.4 (20/45)		
Male sex, <i>n</i> (%)	15 (75.0)	20 (80.0)	35 (77.8)	42.9 (15/35) vs 50.0 (5/10) female	0.731 (1.000)	OR 0.75 (0.18–3.07)
Age band (mABSI points), <i>n</i> (%)					0.459 (1.000); trend 1.000	<i>V</i> = 0.19
0–40 years (1 point)	11 (55.0)	12 (48.0)	23 (51.1)	47.8 (11/23)		
41–70 years (2 points)	8 (40.0)	13 (52.0)	21 (46.7)	38.1 (8/21)		
71–80 years (3 points)	1 (5.0)	0 (0.0)	1 (2.2)	100.0 (1/1)		
Etiology, <i>n</i> (%)					0.031 (0.155)	<i>V</i> = 0.43
Flame	17 (85.0)	12 (48.0)	29 (64.4)	58.6 (17/29)		
Electrical	3 (15.0)	6 (24.0)	9 (20.0)	33.3 (3/9)		
Scald	0 (0.0)	5 (20.0)	5 (11.1)	0.0 (0/5)		
Chemical	0 (0.0)	2 (8.0)	2 (4.4)	0.0 (0/2)		
Flame vs all other mechanisms	17 vs 3	12 vs 13	29 vs 16	58.6 (17/29) vs 18.8 (3/16)	0.013	OR 6.14 (1.43–26.35)
Inhalation injury, <i>n</i> (%)	4 (20.0)	2 (8.0)	6 (13.3)	66.7 (4/6) vs 41.0 (16/39)	0.383 (1.000)	OR 2.88 (0.47–17.63)
Full-thickness component, <i>n</i> (%)	8 (40.0)	8 (32.0)	16 (35.6)	50.0 (8/16) vs 41.4 (12/29)	0.755 (1.000)	OR 1.42 (0.42–4.83)
%TBSA >30%, <i>n</i> (%)	16 (80.0)	7 (28.0)	23 (51.1)	69.6 (16/23) vs 18.2 (4/22)	0.001 (0.006)	OR 10.29 (2.53–41.75)
%TBSA band (10 bands)	—	—	—	see next table	0.015 (0.087); trend <0.001	<i>V</i> = 0.65

Overall in-hospital mortality 20/45 = 44.4% (Wilson 95% CI 30.9–58.8). Mortality column: deaths/patients in the category; for binary characteristics, present vs absent. Exact *p*: Fisher's exact test (2 × 2) or Fisher–Freeman–Halton exact test (2 × *k*), two-sided; Holm-adjusted *p* across the seven characteristic tests in parentheses; trend: exact Cochran–Armitage test across ordered categories. OR, odds ratio of death (Woolf 95% CI); *V*, Cramér's *V*. Mean age of the whole cohort 40.8 ± 19.6 years (range 1–74); age by outcome was available only in the score's age bands.

Neither sex (exact *p* = 0.731) nor age band (*p* = 0.459) was associated with death; mortality was 47.8% at 0–40 years and 38.1% at 41–70 years, with no trend across age bands (exact trend *p* = 1.000). Etiology was associated with death (*p* = 0.031; Cramér's *V* = 0.43): 17 of 29 patients with flame burns died (58.6%) against 3 of 16 with other mechanisms (18.8%; OR 6.14 (1.43–26.35), *p* = 0.013), and no patient with a scald or chemical burn died. Inhalation injury (4/6 vs 16/39 died; OR 2.88 (0.47–17.63), *p* = 0.383) and full-thickness involvement (8/16 vs 12/29; OR 1.42 (0.42–4.83), *p* = 0.755) were not associated with death. Apart from burn size (%TBSA >30%; Holm-adjusted *p* = 0.006), none of the baseline differences remained significant after Holm adjustment.

Mortality rose steeply with burn size, as shown in Figure 2. No patient with burns of 10% TBSA or less died, mortality was 25.0% at 11–20% and 62.5% at 31–40%, and all five patients with burns above 70% died (exact Cochran–Armitage trend *p* < 0.001). Burns of more than 30% TBSA carried an OR of 10.29 (95% CI 2.53–41.75) for death (16/23 vs 4/22). In an exploratory analysis, mortality also differed between recorded depth ranges (*p* = 0.012), as detailed in Table 2: 7 of 9 patients whose full-thickness burns were recorded together with superficial components died, whereas none of the 6 whose full-thickness burns were recorded with mid-dermal components did (*p* = 0.007).

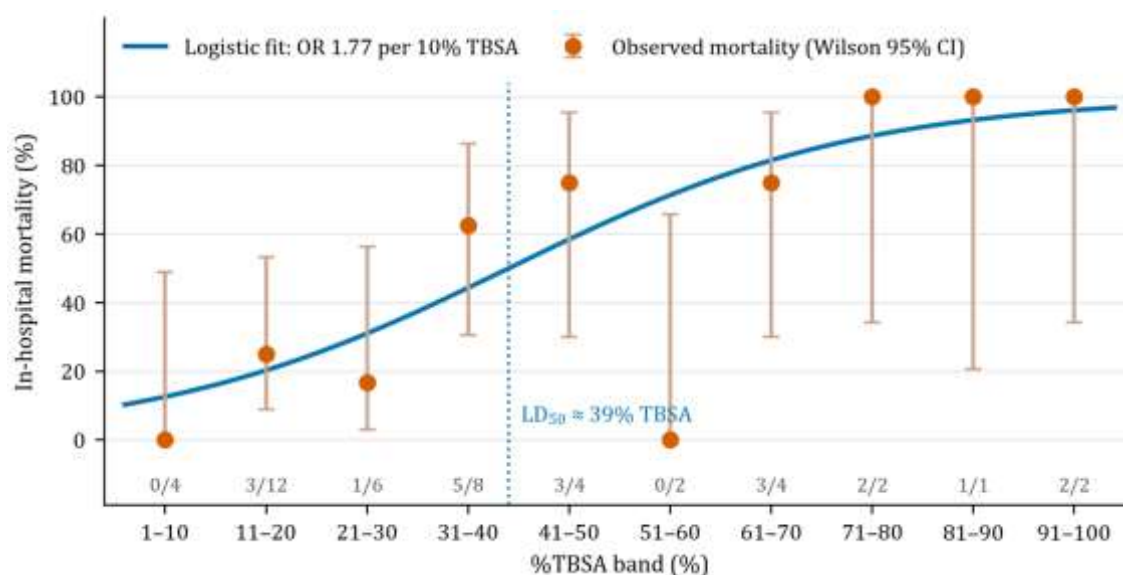


Figure 2. In-hospital mortality by %TBSA band (points: observed proportion with Wilson 95% CI; numbers: deaths/patients) with the fitted logistic curve. The dotted line marks the burn size with a predicted mortality of 50% (LD_{50}).

Table 2. In-hospital mortality by %TBSA band and by recorded burn-depth range

Stratum	Died / n	Mortality, % (Wilson 95% CI)	Stratum	Died / n	Mortality, % (Wilson 95% CI)
%TBSA 1–10	0/4	0.0 (0.0–49.0)	Epidermal–mid-dermal	0/1	0.0 (0.0–79.3)
%TBSA 11–20	3/12	25.0 (8.9–53.2)	Superficial–mid-dermal	6/18	33.3 (16.3–56.3)
%TBSA 21–30	1/6	16.7 (3.0–56.4)	Superficial–deep-dermal	6/10	60.0 (31.3–83.2)
%TBSA 31–40	5/8	62.5 (30.6–86.3)	Superficial–full-thickness	7/9	77.8 (45.3–93.7)
%TBSA 41–50	3/4	75.0 (30.1–95.4)	Mid-dermal–full-thickness	0/6	0.0 (0.0–39.0)
%TBSA 51–60	0/2	0.0 (0.0–65.8)	Deep-dermal–full-thickness	1/1	100.0 (20.7–100.0)
%TBSA 61–70	3/4	75.0 (30.1–95.4)			
%TBSA 71–80	2/2	100.0 (34.2–100.0)			
%TBSA 81–90	1/1	100.0 (20.7–100.0)			
%TBSA 91–100	2/2	100.0 (34.2–100.0)			

%TBSA band: exact Fisher–Freeman–Halton $p = 0.015$; exact Cochran–Armitage trend $p < 0.001$. Depth range (shallowest–deepest component recorded at admission or first excision): exact $p = 0.012$. Full-thickness involvement is present in the last three depth categories. Full-thickness burns recorded with superficial components vs with mid-dermal components: 7/9 vs 0/6 deaths, exact $p = 0.007$ (exploratory).

Score distributions

Both scores were higher in patients who died, as summarized in Table 3 and shown in Figure 3. The median mABSI was 6 (interquartile range [IQR] 5–10) in patients who died and 5 (IQR 4–6) in survivors (exact Mann–Whitney $p < 0.001$; rank-biserial

$r = 0.62$); the observed range was 2–13. The median R-Baux band was 91–100 (IQR 71–80 to 111–120) in patients who died and 61–70 (IQR 51–60 to 71–80) in survivors ($p < 0.001$; $r = 0.62$). The distributions overlapped: 5 of the 20 deaths occurred at an mABSI of 5 or less, and 7 of the 25 survivors scored 6 or more.

Table 3. Distribution of the two scores and of %TBSA by outcome, and their discrimination of in-hospital death

Measure	Modified ABSI (points)	Revised Baux (10-point band)	%TBSA band alone
Died: median (IQR); range	6 (5–10); 4–13	91–100 (71–80 to 111–120); 51–60 to 141–150	41–50 (31–40 to 61–70); 11–20 to 91–100
Discharged alive: median (IQR); range	5 (4–6); 2–9	61–70 (51–60 to 71–80); ≤20 to 111–120	11–20 (11–20 to 31–40); 1–10 to 61–70
Died vs alive: mean ± SD	7.50 ± 2.70 vs 4.84 ± 1.68	band code 9.20 ± 2.71 vs 6.00 ± 2.43	—
Exact Mann–Whitney p	<0.001	<0.001	—
Rank-biserial r ($= 2 \times \text{AUC} - 1$)	0.62	0.62	0.61
AUC	0.811	0.808	0.803
DeLong 95% CI	0.690–0.932	0.679–0.937	0.675–0.931
Logit-transformed 95% CI	0.661–0.904	0.647–0.906	0.645–0.901
Bootstrap percentile 95% CI (20,000)	0.680–0.917	0.672–0.921	—
Attainable raw-score AUC (band ties resolved)	— (integer score)	0.764–0.852	0.750–0.856
p vs AUC 0.5	<0.001	<0.001	<0.001
Youden-optimal cutoff (bootstrap selection, %)	≥6 (54); ≥5 (34)	≥81 (82); ≥71 (8)	—
Maximum Youden J (bootstrap 95% CI)	0.47 (0.30–0.72)	0.54 (0.32–0.77)	—

The Revised Baux score was recorded in 10-point bands (≤20, 21–30, ..., 141–150); its band code (1 code = 10 points) was used for rank statistics, AUC and the mean ± SD row. IQR, interquartile range. AUC, area under the receiver-operating-characteristic curve (ties counted as one half).

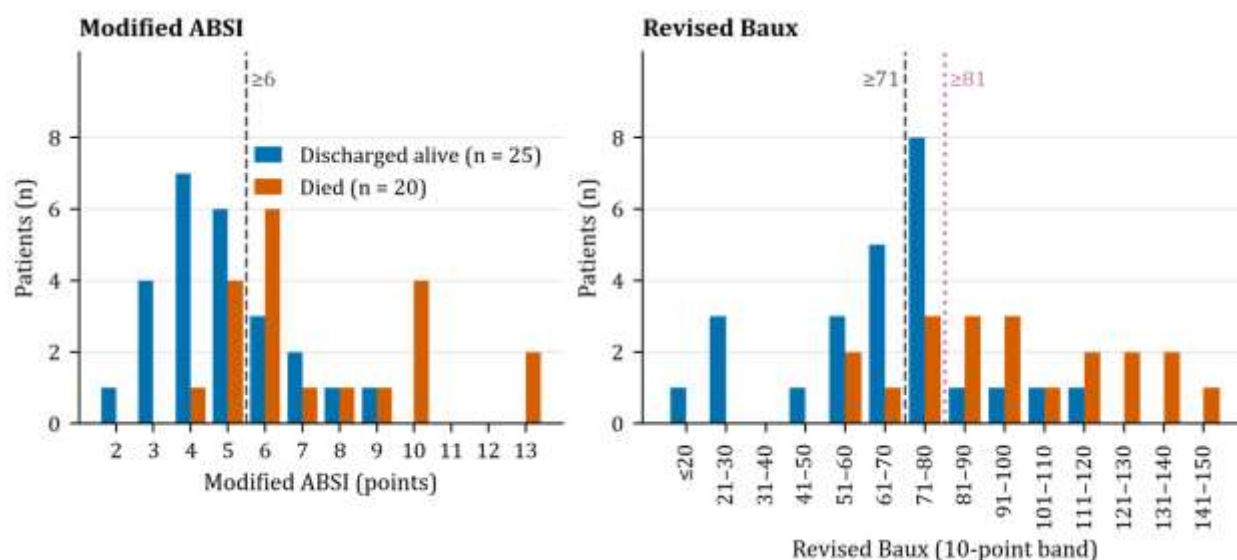


Figure 3. Distribution of the Modified ABSI (left) and the Revised Baux score in 10-point bands (right) among patients who died and those discharged alive. Dashed lines mark the cutoffs used for the dichotomized classification (≥ 6 and ≥ 71); the dotted line marks the Youden-optimal Revised Baux cutoff (≥ 81).

Discrimination and accuracy

The empirical ROC curves of the two scores and of burn size alone were close throughout their range, as shown in Figure 4. The AUC was 0.811 (DeLong 95% CI 0.690–0.932) for the mABSI, 0.808 (0.679–0.937) for the R-Baux and 0.803 (0.675–0.931) for the %TBSA band alone; each differed from 0.5 ($p <$

0.001 for all three). Logit-transformed and bootstrap intervals, listed in Table 3, gave the same conclusions. Because the R-Baux and burn size were recorded in bands, their raw-score AUCs could lie anywhere between 0.764 and 0.852 and between 0.750 and 0.856, respectively, depending on the order of patients within bands; the mABSI is an integer score and its ties are genuine.

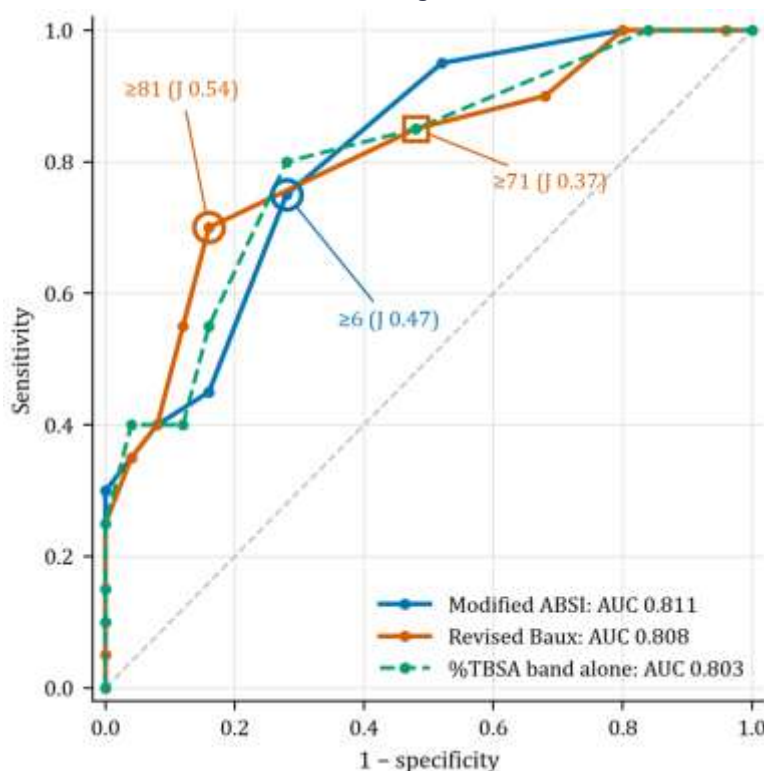


Figure 4. Empirical receiver-operating-characteristic curves for the Modified ABSI, the Revised Baux score and %TBSA band alone. Circles mark the Youden-optimal cutoffs; the square marks the Revised Baux cutoff of ≥ 71 .

The Youden-optimal cutoff was ≥ 6 for the mABSI, selected in 54% of bootstrap resamples (≥ 5 in 34%), and ≥ 81 for the R-Baux, selected in 82%. Accuracy at these and neighboring cutoffs is detailed in Table 4. At mABSI ≥ 6 , sensitivity was 75.0% (95% CI 53.1–88.8), specificity 72.0% (52.4–85.7), the positive likelihood ratio 2.68 (1.36–5.27) and the negative likelihood ratio 0.35 (0.16–0.77), with $J = 0.47$. At R-Baux ≥ 81 , sensitivity was 70.0% (48.1–85.5), specificity 84.0% (65.3–93.6), the

positive likelihood ratio 4.37 (1.70–11.23) and the negative likelihood ratio 0.36 (0.18–0.71), with $J = 0.54$. The lower R-Baux cutoff of ≥ 71 raised sensitivity to 85.0% at the cost of a specificity of 52.0% ($J = 0.37$), and an mABSI of ≥ 5 identified 19 of the 20 deaths (95.0%) with a specificity of 48.0%. After bootstrap correction for optimism, the Youden index was 0.40 for mABSI ≥ 6 (sensitivity 70.9%, specificity 68.7%) and 0.50 for R-Baux ≥ 81 (sensitivity 67.7%, specificity 82.2%).

Table 4. Diagnostic accuracy of the Modified ABSI and the Revised Baux score for in-hospital death at clinically relevant cutoffs

Score and cutoff	TP/FP/FN/TN	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	LR+ (95% CI)	LR- (95% CI)	J
Modified ABSI ≥ 5	19/13/1/12	95.0 (76.4–99.1)	48.0 (30.0–66.5)	59.4 (42.3–74.5)	92.3 (66.7–98.6)	1.83 (1.24–2.70)	0.10 (0.01–0.73)	0.43
Modified ABSI $\geq 6^*$	15/7/5/18	75.0 (53.1–88.8)	72.0 (52.4–85.7)	68.2 (47.3–83.6)	78.3 (58.1–90.3)	2.68 (1.36–5.27)	0.35 (0.16–0.77)	0.47
Modified ABSI ≥ 7	9/4/11/21	45.0 (25.8–65.8)	84.0 (65.3–93.6)	69.2 (42.4–87.3)	65.6 (48.3–79.6)	2.81 (1.01–7.80)	0.65 (0.43–1.01)	0.29
Revised Baux $\geq 71^\dagger$	17/12/3/13	85.0 (64.0–94.8)	52.0 (33.5–70.0)	58.6 (40.7–74.5)	81.3 (57.0–93.4)	1.77 (1.13–2.77)	0.29 (0.10–0.87)	0.37
Revised Baux $\geq 81^*$	14/4/6/21	70.0 (48.1–85.5)	84.0 (65.3–93.6)	77.8 (54.8–91.0)	77.8 (59.2–89.4)	4.37 (1.70–11.23)	0.36 (0.18–0.71)	0.54
Revised Baux ≥ 91	11/3/9/22	55.0 (34.2–74.2)	88.0 (70.0–95.8)	78.6 (52.4–92.4)	71.0 (53.4–83.9)	4.58 (1.48–14.23)	0.51 (0.31–0.85)	0.43

* Youden-optimal cutoff. † Cutoff of the unit's dichotomized Revised Baux classification (71–80 band and above). Death is the positive state (prevalence 20/45). TP, true positives; FP, false positives; FN, false negatives; TN, true negatives; PPV/NPV, positive/negative predictive value; LR+/LR-, positive/negative likelihood ratio (log-method 95% CI); J, Youden index. Proportion CIs by the Wilson score method. Bootstrap optimism-corrected values at the Youden-optimal cutoffs (Harrell, 20,000 resamples): mABSI ≥ 6 , J = 0.40 (sensitivity 70.9%, specificity 68.7%); R-Baux ≥ 81 , J = 0.50 (sensitivity 67.7%, specificity 82.2%).

Paired comparison and agreement

The paired classification of all 45 patients by the two dichotomized scores is detailed in Table 5. No patient was classified as high risk by the mABSI (≥ 6) and low risk by the R-Baux (< 71); the 7 discordant patients were all R-Baux-positive and mABSI-negative, and 2 of them died. The R-Baux classification was therefore the more sensitive (85.0% vs 75.0%; paired difference –10.0 percentage points, 95% CI –26.2

to 4.5; exact McNemar $p = 0.500$) and the mABSI classification the more specific (72.0% vs 52.0%; +20.0 points, 4.2 to 34.2; exact McNemar $p = 0.063$). Agreement was substantial ($\kappa = 0.69$, 95% CI 0.49–0.89; PABAK 0.69; overall agreement 84.4%). Because mABSI-positive patients were a subset of R-Baux-positive patients, a rule treating a patient as high risk when either score was positive reproduced the R-Baux classification exactly, and a rule requiring both reproduced the mABSI classification.

Table 5. Paired comparison of the dichotomized Modified ABSI and Revised Baux classifications in the same 45 patients

Comparison	Died (n = 20)	Discharged alive (n = 25)	Paired statistic
Both positive (mABSI ≥ 6 and R-Baux ≥ 71)	15	7	
R-Baux ≥ 71 only	2	5	
mABSI ≥ 6 only	0	0	
Both negative	3	13	
Sensitivity, mABSI ≥ 6 vs R-Baux ≥ 71	75.0% vs 85.0%	—	Δ –10.0 points (–26.2 to 4.5); McNemar exact $p = 0.500$
Specificity, mABSI ≥ 6 vs R-Baux ≥ 71	—	72.0% vs 52.0%	Δ +20.0 points (4.2 to 34.2); McNemar exact $p = 0.063$
Agreement of the two dichotomized scores (all 45)			$\kappa = 0.69$ (0.49–0.89), $p < 0.001$; PABAK 0.69; overall agreement 38/45 (84.4%)
Positive / negative percent agreement			86.3% / 82.1%
Youden-optimal pair (mABSI ≥ 6 vs R-Baux ≥ 81): sensitivity	75.0% vs 70.0%	—	McNemar exact $p = 1.000$ in every admissible data set
Youden-optimal pair: specificity	—	72.0% vs 84.0%	McNemar exact $p = 0.250$ –0.508 across admissible data sets
Either-positive rule (mABSI ≥ 6 or R-Baux ≥ 81)	15–17 detected	7–10 false positives	Attainable range
Both-positive rule (mABSI ≥ 6 and R-Baux ≥ 81)	12–14 detected	1–4 false positives	Attainable range

Rows 1–4 are observed: no patient was positive on the Modified ABSI (≥ 6) and negative on the Revised Baux (≥ 71), so the classification of every patient by both scores and outcome is fully determined. Δ , paired difference (mABSI minus R-Baux) with Newcombe's hybrid-score 95% CI for paired proportions. κ , Cohen's kappa; PABAK, prevalence- and bias-adjusted kappa. The last four rows depend on how patients pair up across the 81–90 band and are therefore given as the range attainable over every admissible data set.

At the Youden-optimal pair (mABSI ≥ 6 vs R-Baux ≥ 81), the sensitivities (75.0% vs 70.0%) and specificities (72.0% vs 84.0%) did not differ in any admissible data set (McNemar $p = 1.000$ and 0.250–0.508). An either-positive rule at these cutoffs would have detected 15–17 of the 20 deaths with 7–10 false positives among the 25 survivors, and a both-positive rule 12–14 deaths with 1–4 false positives.

Risk curves

Observed mortality at each score level and the fitted logistic curves are shown in Figure 5, and the model statistics are

detailed in Table 6. Each additional mABSI point multiplied the odds of death by 1.84 (1.29–2.98), each 10 R-Baux points by 1.69 (1.26–2.55) and each 10% band of TBSA by 1.77 (1.28–2.71). Predicted mortality reached 50% at an mABSI of 6.3 points (95% CI 5.1–7.5), an R-Baux of about 85 (71–99) and a burn size of about 39% TBSA (27–52%). The three models fitted similarly: Nagelkerke R^2 0.371, 0.386 and 0.359; Brier scores 0.180, 0.175 and 0.180 against 0.247 for the overall mortality alone; and Hosmer–Lemeshow p values (low power) of 0.472, 0.743 and 0.330.

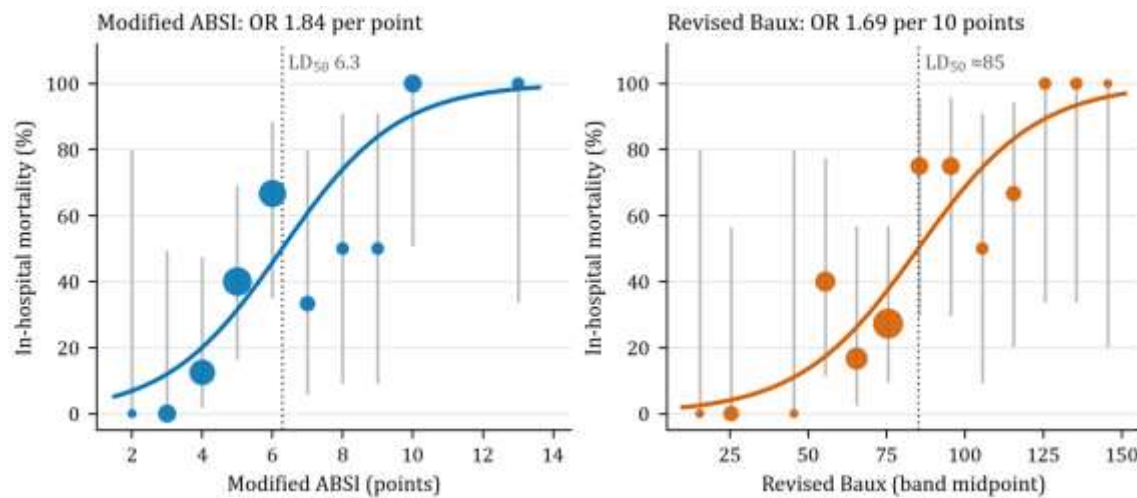


Figure 5. Observed mortality at each score level (marker size proportional to the number of patients; bars, Wilson 95% CI) and fitted logistic risk curves for the Modified ABSI (left) and the Revised Baux score (right). Dotted lines mark the score with 50% predicted mortality (LD_{50}).

Table 6. Single-predictor logistic models of in-hospital death fitted to the observed score–outcome data, with decision-curve net benefit

Model (observed data)	OR per unit (profile 95% CI)	LR χ^2 (p)	Nagelkerke R^2	LD_{50} (95% CI)	Brier score	Hosmer–Lemeshow p	Firth OR	Net benefit at 20%	Net benefit at 30%
Modified ABSI (per point)	1.84 (1.29–2.98)	14.58 (<0.001)	0.371	6.3 (5.1–7.5)	0.180	0.472	1.72	0.333	0.298
Revised Baux (per 10 points)	1.69 (1.26–2.55)	15.32 (<0.001)	0.386	85 (71–99)	0.175	0.743	1.61	0.306	0.263
%TBSA (per 10% band)	1.77 (1.28–2.71)	14.03 (<0.001)	0.359	39% (27%–52%)	0.180	0.330	1.68	0.328	0.263
mABSI, published categories	—	—	—	—	—	—	—	0.167	0.159
Treat all as high risk	—	—	—	—	—	—	—	0.306	0.206

LD_{50} , score at which predicted mortality is 50% (delta-method CI); Revised Baux and %TBSA converted from band codes to points/percentages at band midpoints. Brier score of the null model (overall mortality) 0.247. Hosmer–Lemeshow: Pearson statistic over observed score levels with a parametric-bootstrap (20,000) p ; the test has little power with 20 deaths. Optimism-corrected calibration slope (Harrell bootstrap, 2,000 resamples): 1.00, 1.00 and 1.02. Net benefit = $TP/N - FP/N \times p_i / (1 - p_i)$ at threshold probabilities p_i of 20% and 30% (apparent performance).

Predicted mortality at each observed score level is listed in Table 7. An mABSI of 4 corresponded to a predicted mortality of 20.0% (95% CI 8.6–40.1), 5 to 31.5% (95% CI 17.8–49.4), 6 to

45.8% (95% CI 29.4–63.2) and 8 to 74.0% (95% CI 47.0–90.1), whereas the published categories assign 2% to scores of 4–5 and 10–20% to scores of 6–7.

Table 7. Locally predicted in-hospital mortality at each score level (single-predictor logistic models fitted in this cohort)

mABSI	Died / n	Predicted mortality, % (95% CI)	R-Baux band	Died / n	Predicted mortality, % (95% CI)
2	0/1	6.9 (1.4–28.1)	≤20	0/1	2.5 (0.2–21.0)
3	0/4	12.0 (3.6–33.4)	21–30	0/3	4.2 (0.6–24.6)
4	1/8	20.0 (8.6–40.1)	31–40	—	6.8 (1.3–28.7)
5	4/10	31.5 (17.8–49.4)	41–50	0/1	11.0 (3.0–33.4)
6	6/9	45.8 (29.4–63.2)	51–60	2/5	17.4 (6.5–39.0)
7	1/3	60.8 (39.3–78.8)	61–70	1/6	26.3 (12.9–46.0)
8	1/2	74.0 (47.0–90.1)	71–80	3/11	37.6 (22.6–55.4)
9	1/2	83.9 (53.5–96.0)	81–90	3/4	50.5 (33.2–67.8)
10	4/4	90.6 (59.3–98.4)	91–100	3/4	63.4 (42.0–80.5)
11	—	94.6 (64.6–99.4)	101–110	1/2	74.5 (49.2–89.8)
12	—	97.0 (69.4–99.8)	111–120	2/3	83.2 (55.3–95.2)
13	2/2	98.3 (73.7–99.9)	121–130	2/2	89.4 (60.7–97.9)
			131–140	2/2	93.4 (65.6–99.1)
			141–150	1/1	96.0 (70.0–99.6)

In-sample estimates from 45 patients and 20 deaths (Wald 95% CI on the logit scale); levels without patients are model extrapolations. For comparison, the published Modified ABSI categories assign mortality of ≤1% to scores 2–3, 2% to 4–5, 10–20% to 6–7, 30–50% to 8–9, 60–80% to 10–11 and ≥90% to ≥12.

Calibration of the published risk categories

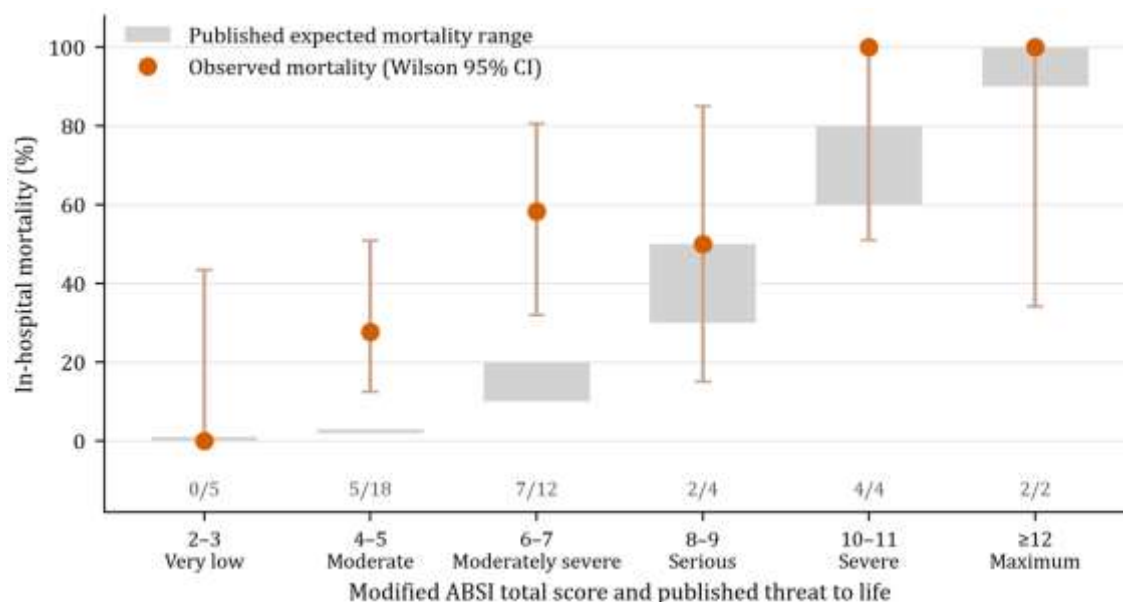
Observed mortality exceeded the published mABSI estimates, as detailed in Table 8 and shown in Figure 6. In the 'moderate' category (scores 4–5), for which the published mortality is 2%, 5 of 18 patients died (27.8%, 95% CI 12.5–50.9; exact $p < 0.001$); in the 'moderately severe' category (6–7; published 10–20%), 7 of 12 died (58.3%, 32.0–80.7; $p = 0.004$). Overall, 20 deaths were

observed against 8.5 expected from the category midpoints (observed/expected 2.36, 95% CI 1.44–3.64) and against 10.0 expected even from the upper bounds of the published ranges (Poisson $p = 0.003$). The higher categories were too small to test individually. Logistic recalibration of the published probabilities gave a calibration-in-the-large intercept of 2.26 (95% CI 1.47–3.05), corresponding to about 9.6-fold higher odds of death than predicted, and a calibration slope of 0.69 (0.31–1.18).

Table 8. Observed in-hospital mortality against the published Modified ABSI threat-to-life categories

mABSI total	Threat to life	Published mortality	Died / n	Observed mortality, % (95% CI)	Exact <i>p</i> vs published upper bound	Stratum LR (95% CI)
2-3	Very low	≤1%	0/5	0.0 (0.0-43.4)	—	0.11 (0.01-1.92)‡
4-5	Moderate	2%	5/18	27.8 (12.5-50.9)	<0.001	0.48 (0.21-1.12)
6-7	Moderately severe	10-20%	7/12	58.3 (32.0-80.7)	0.004	1.75 (0.65-4.69)
8-9	Serious	30-50%	2/4	50.0 (15.0-85.0)	0.688	1.25 (0.19-8.11)
10-11	Severe	60-80%	4/4	100.0 (51.0-100.0)	0.410	11.14 (0.64-195.47)‡
≥12	Maximum	≥90%	2/2	100.0 (34.2-100.0)	—	6.19 (0.31-122.05)‡

Published mortality = 1 - published probability of survival for each total-score band. Exact one-sided binomial test of the observed deaths against the upper bound of the published range. Observed/expected deaths using category midpoints: 20/8.5 = 2.36 (95% CI 1.44-3.64); against the upper bounds, 20 observed vs 10.0 expected (Poisson *p* = 0.003). Stratum LR, stratum-specific likelihood ratio = $P(\text{stratum} | \text{death}) / P(\text{stratum} | \text{survival})$; ‡ 0.5 added to empty cells. Logistic recalibration of the published probabilities (category midpoints): calibration-in-the-large intercept 2.26 (odds ratio 9.6) (95% CI 1.47-3.05); calibration slope 0.69 (0.31-1.18).

**Figure 6.** Observed in-hospital mortality (Wilson 95% CI) in each published Modified ABSI threat-to-life category against the published expected mortality range (gray bars).

Sensitivity analysis: paired comparisons and adjusted effects

Across every admissible data set, the AUC difference between the scores (0.003) was far from significant, as detailed in Table 9 and shown in Figure 7: the paired standard error ranged from

0.031 to 0.074, the *p* value from 0.922 to 0.968, and even the widest 95% CI (-0.142 to 0.148) excluded differences larger than 0.15. Neither score improved on burn size alone in any admissible data set (mABSI minus %TBSA 0.008, *p* 0.736-0.897; R-Baux minus %TBSA 0.005, *p* 0.785-0.958).

Table 9. Sensitivity analysis over every patient-level data set consistent with the observed data: paired AUC comparisons and Firth multivariable models

Quantity	Attainable range over admissible data sets	Data sets with <i>p</i> < 0.05	Verdict
ΔAUC, mABSI - R-Baux (paired DeLong)	0.003 fixed; SE 0.031-0.074; 95% CI at widest -0.142 to 0.148; <i>p</i> 0.922-0.968	None	No difference (robust); differences >0.15 excluded
ΔAUC, mABSI - %TBSA band	0.008 fixed; 95% CI at widest -0.113 to 0.129; <i>p</i> 0.736-0.897	None	No added discrimination (robust)
ΔAUC, R-Baux - %TBSA band	0.005 fixed; 95% CI at widest -0.182 to 0.192; <i>p</i> 0.785-0.958	None	No added discrimination (robust)
Joint model: mABSI OR per point R-Baux	1.10-1.43 (central 95%: 1.21-1.34); <i>p</i> 0.043-0.859	0.0%	Not identifiable
Joint model: R-Baux OR per 10 points mABSI	1.26-1.44 (central 95%: 1.29-1.37); <i>p</i> 0.029-0.413	0.0%	Not identifiable
Joint model AUC	0.796-0.852	—	≤0.04 above the better single score
Component model: %TBSA OR per 10%	1.62-3.42; maximum <i>p</i> 0.001	100.0%	Robust independent predictor
Component model: age band OR	0.29-17.37	0.1%	Not identifiable
Component model: inhalation injury OR	0.08-112.06	7.2%	Not identifiable
Component model: full-thickness OR	0.04-17.93	0.3%	Not identifiable

Paired AUC rows: exact minimum and maximum of the DeLong covariance by integer linear programming. Model rows: 3,000 admissible data sets sampled at random vertices plus 24 constructed extremes (most extreme values found, a lower bound on the attainable range); 'Data sets with *p* < 0.05' refers to the random sample. Firth-penalized logistic regression, penalized likelihood-ratio *p*; no fit was discarded. Component model: age band, %TBSA band (per band), inhalation injury and full thickness entered together. Sample size for future studies: detecting an AUC difference of 0.05 with 80% power would need about 132 patients at the narrowest and 772 at the widest attainable paired SE; estimating a sensitivity of 75% within ±10 or ±5 percentage points would need about 162 or 648 patients at the observed mortality.

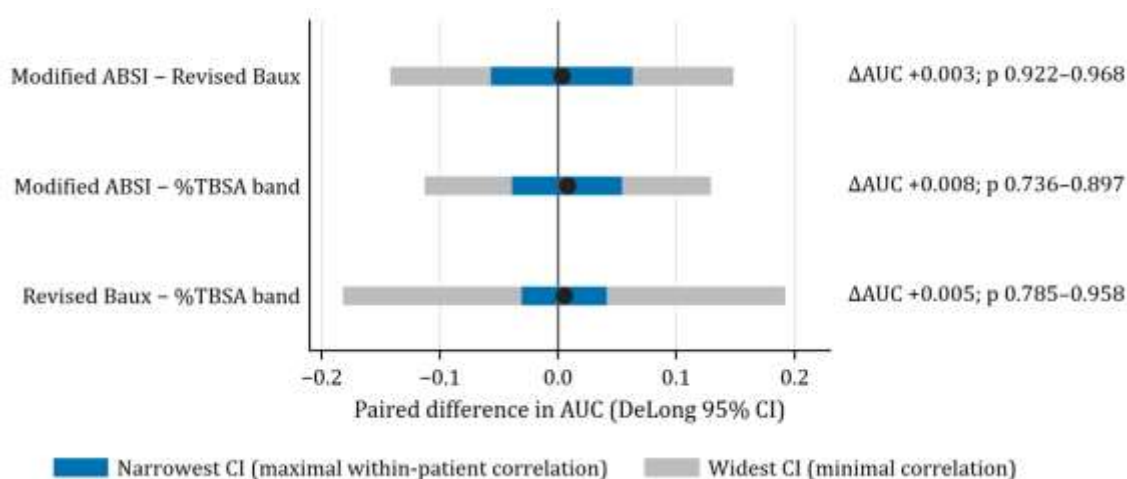


Figure 7. Paired differences in AUC across every admissible patient-level data set. Point: the difference (fixed by the observed data); blue bar: the narrowest attainable DeLong 95% CI; gray bar: the widest.

In models containing both scores, the adjusted OR per mABSI point ranged from 1.10 to 1.43 and that per 10 R-Baux points from 1.26 to 1.44; neither reached significance in any of the randomly sampled data sets, although constructed extremes reached $p = 0.043$ and 0.029 , respectively, so the independent contribution of either score cannot be identified from these data.

The combined model's AUC ranged from 0.796 to 0.852. In the four-component model, %TBSA was the only robust predictor (adjusted OR per 10% band 1.62–3.42; significant in every sampled data set), whereas the attainable ORs for age band (0.29–17.37), inhalation injury (0.08–112.06) and full thickness (0.04–17.93) all spanned 1, as shown in Figure 8.

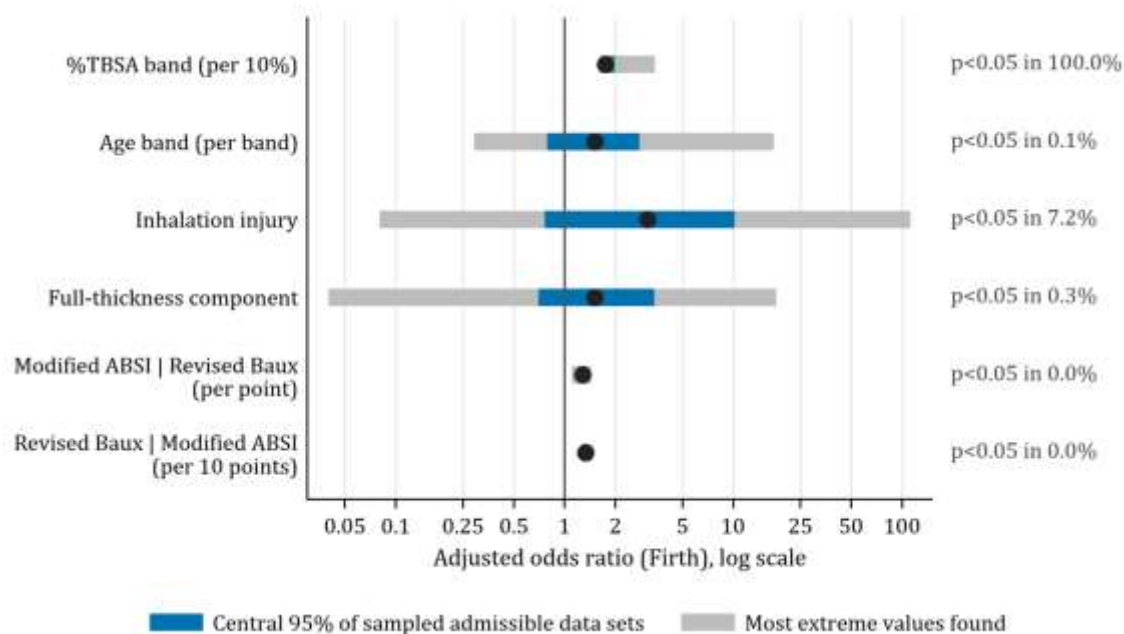


Figure 8. Adjusted odds ratios of death across admissible data sets: four-component model (upper four rows) and joint model of the two scores (lower two rows). Blue/green bar, central 95% of sampled data sets; gray bar, most extreme values found; dot, median; right, percentage of sampled data sets with $p < 0.05$.

Clinical utility

The decision curves of the three locally fitted models were close to one another and exceeded the treat-all strategy at threshold probabilities of death above about 20%, as shown in Figure 9. At a threshold of 30%, the net benefit was 0.298 for the mABSI,

0.263 for the R-Baux and 0.263 for burn size alone, against 0.206 for treating every patient as high risk. Using the published mABSI category probabilities instead of local risks lowered the net benefit to 0.159 at the same threshold and to 0.167 at 20%, because patients in the 'moderate' category were classified as low risk although more than a quarter of them died.

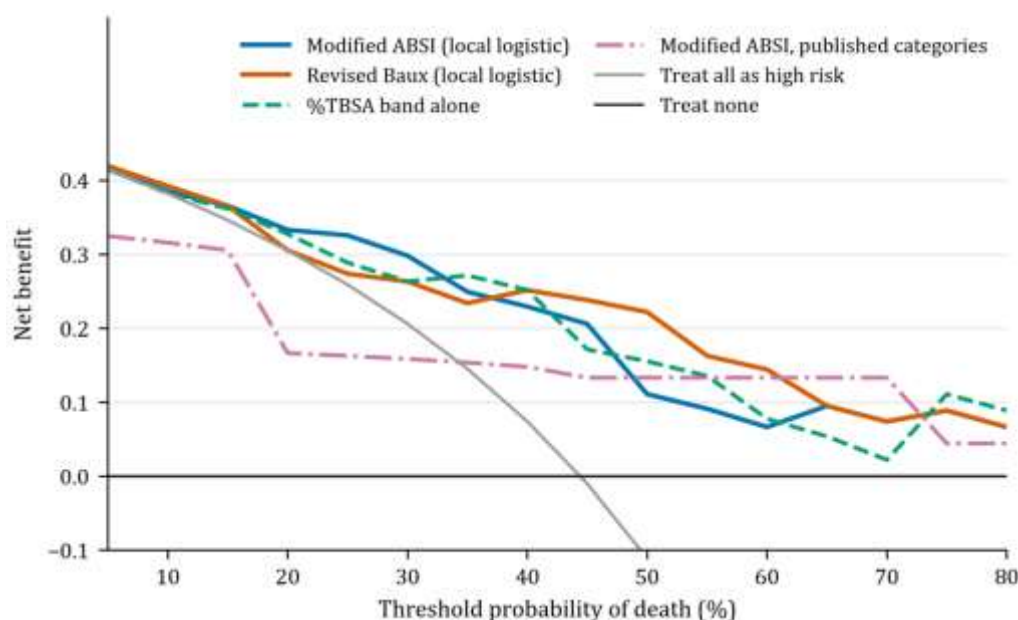


Figure 9. Decision curves: net benefit of classifying patients as high risk by each model across threshold probabilities of death, compared with treating all or none. Local models were fitted in this cohort (apparent performance); the pink dash-dot curve uses the published Modified ABSI category probabilities without refitting.

Comparison with published cohorts

Mortality in this cohort was similar to that at the national referral burn center in Jakarta (42.2%)⁹ but higher than in Surabaya (27.8%)¹⁰ and at an Indonesian emergency center (25.0%),¹¹ as detailed in Table 10. The R-Baux AUC was lower than the meta-analytic summary of 0.93²³ and the German registry value of 0.939,²⁷ but similar to cohorts with large burns

or high mortality, such as Zagreb (0.79)³⁵ and a Polish intensive-care unit (0.86).³⁸ The mABSI AUC was lower than the German registry value of 0.928, and the AUC of burn size alone matched the 0.813 reported in Jakarta, as shown in Figure 10. The comparator cohorts differed from ours in mortality and burn size, so these AUC differences describe case mix as much as score performance.

Table 10. Discrimination and mortality in this cohort compared with published cohorts

Comparator (setting)	Measure	Comparator value	This study	Difference
Discrimination (AUC)				
Edgar et al. ²³ : meta-analysis of 21 studies	R-Baux AUC	0.930 (0.910–0.950)	0.808	–0.122
Mrad et al. ²⁴ : meta-analysis of 34 studies	R-Baux AUC	0.920 (0.900–0.930)	0.808	–0.112
Bagheri et al. ²⁷ : German Burn Registry, <i>n</i> = 7,276	R-Baux AUC	0.939 (0.930–0.948)	0.808	–0.131
Rahimpour et al. ³² : West Virginia, USA, <i>n</i> = 1,104	R-Baux AUC	0.927 (0.892–0.962)	0.808	–0.119
Belayneh et al. ³³ : Addis Ababa, Ethiopia, <i>n</i> = 149	R-Baux AUC	0.923 (0.850–0.996)	0.808	–0.115
Baraka et al. ³¹ : Uganda, 3 hospitals, <i>n</i> = 101	R-Baux AUC	0.943	0.808	–0.135
Tomaka et al. ³⁸ : Polish burn ICU, <i>n</i> = 144	R-Baux AUC	0.860	0.808	–0.052
Goder et al. ³⁹ : Older adults, Toronto, <i>n</i> = 109	R-Baux AUC	0.839	0.808	–0.031
Costa et al. ⁴⁰ : Ventilated burn patients, Portugal	R-Baux AUC	0.834	0.808	–0.026
Skunca et al. ³⁵ : TBSA ≥20%, Zagreb, <i>n</i> = 149	R-Baux AUC	0.790	0.808	+0.018
Sharma et al. ³⁶ : TBSA 30–60%, New Delhi	R-Baux AUC	0.746 (0.671–0.821)	0.808	+0.062
Bagheri et al. ²⁷ : German Burn Registry, <i>n</i> = 7,276	Modified ABSI AUC	0.928 (0.918–0.939)	0.811	–0.117
Wardhana et al. ⁹ : Jakarta, Indonesia, <i>n</i> = 595	%TBSA AUC	0.813	0.803	–0.010
In-hospital mortality				
Wardhana et al. ⁹ : Jakarta national referral burn center, adults	In-hospital mortality	42.2% (251/595)	44.4% (20/45)	<i>h</i> = +0.05
Iustitiani et al. ¹⁰ : Dr. Soetomo Hospital, Surabaya	In-hospital mortality	27.8% (73/263)	44.4% (20/45)	<i>h</i> = +0.35
Herlianita et al. ¹¹ : Indonesian emergency center	In-hospital mortality	25.0% (18/72)	44.4% (20/45)	<i>h</i> = +0.41
Skunca et al. ³⁵ : TBSA ≥20%, Zagreb	In-hospital mortality	49.0% (73/149)	44.4% (20/45)	<i>h</i> = –0.09
Tomaka et al. ³⁸ : Burn ICU, Poland	In-hospital mortality	43.1% (62/144)	44.4% (20/45)	<i>h</i> = +0.03

Mortality: published counts; *h*, Cohen's *h*. Comparator values are taken from the published abstracts; the R-Baux AUC of Mrad et al. was pooled from 22 of the 34 studies. Comparator cohorts differ from this cohort in mortality and burn size; AUC differences therefore reflect case mix as well as score performance.

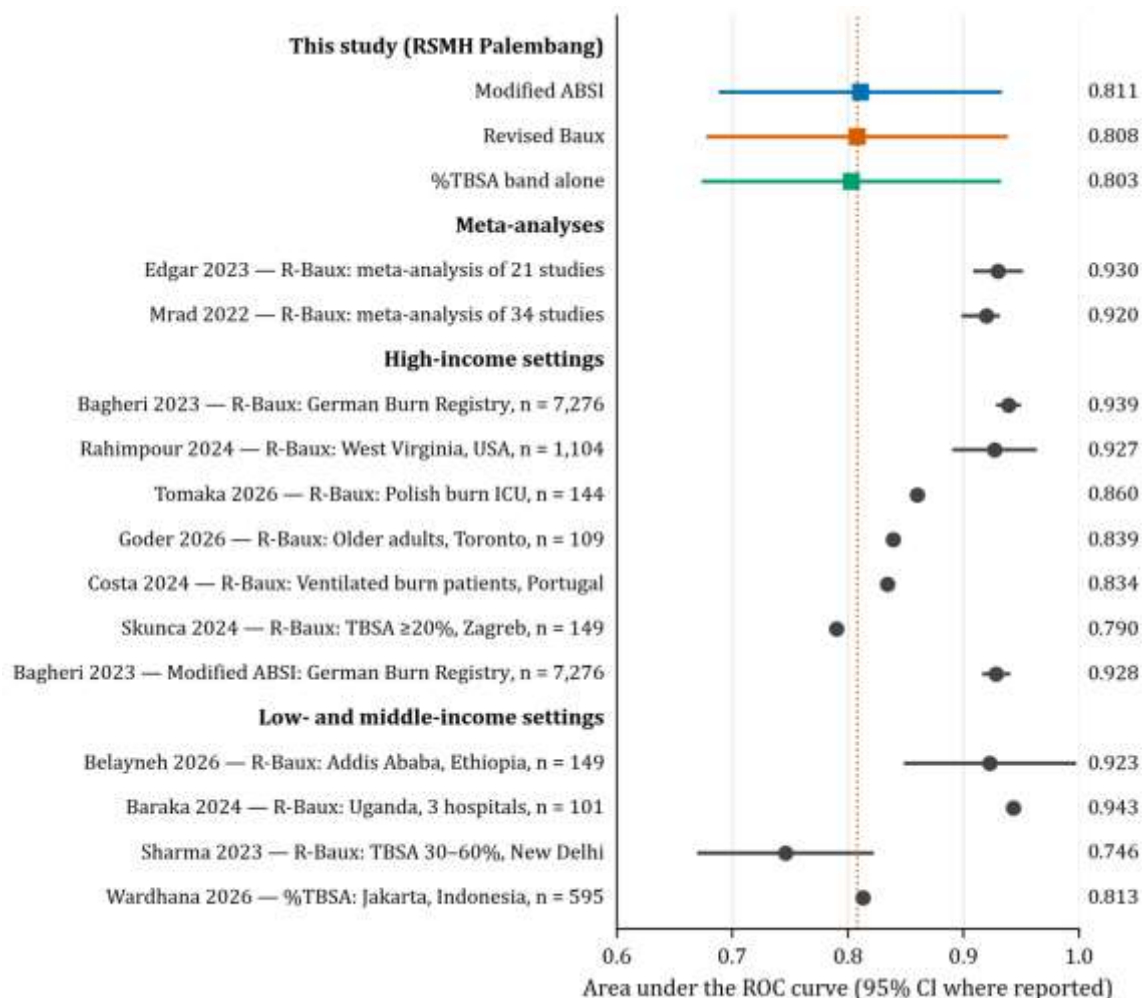


Figure 10. Areas under the ROC curve in this cohort (squares) against published meta-analyses and cohorts grouped by setting (circles; horizontal lines, 95% CI where reported). The dotted line marks Revised Baux AUC.

4. Discussion

In this paired comparison of 45 patients treated in the RSMH burn unit, the mABSI and the R-Baux ranked the risk of in-hospital death equally well, with AUCs of 0.811 and 0.808, and neither ranked it detectably better than burn size alone (0.803); these conclusions held in every patient-level data set consistent with the observations, which excluded AUC differences larger than about 0.15. At their Youden-optimal cutoffs the mABSI (≥ 6) and the R-Baux (≥ 81) offered different trade-offs between sensitivity and specificity, but neither was demonstrably more accurate, and contrary to our hypothesis the mABSI classification used in the unit was the more specific rather than the more sensitive. The most consequential finding concerns absolute risk rather than ranking: in-hospital mortality was more than twice that implied by the published mABSI risk categories, and the score values at which half the patients died were lower than in the registries in which both scores were developed.

Discrimination in context

The AUCs observed here lie below the values reported in high-income registries and meta-analyses but within the range reported for severely injured cohorts. The summary R-Baux AUC across 21 studies was 0.93,²³ German registry values were 0.939 for the R-Baux and 0.928 for the mABSI,²⁷ and single-center series from West Virginia (0.927),³² Jordan (0.918),⁴¹ Kuwait (above 0.90)⁴² and British Columbia (0.95 for the Baux index)⁴³ were similar. Values close to ours come from cohorts resembling ours in severity: 0.79 for both the R-Baux and the

ABSI in Croatian patients with at least 20% TBSA and 49% mortality,³⁵ 0.75 for both in Indian adults with 30–60% TBSA,³⁶ 0.86 in a Polish intensive-care unit with 43.1% mortality,³⁸ 0.834 in mechanically ventilated Portuguese patients⁴⁰ and 0.839 in older Canadian adults.³⁹ This pattern is expected from the statistics of discrimination. When mortality is low, as in registries in which most patients have small burns, the score separates a large group of survivors with trivial injuries from a small group of non-survivors with massive burns, and the AUC is driven by these easy comparisons. In a referral cohort in which half the patients have burns over 30% TBSA, the informative comparisons are among patients with intermediate scores, where survival depends on events after admission, and the AUC falls. Lower discrimination here therefore reflects case mix more than a defect in either score.

Head-to-head comparisons of the two scores elsewhere have reached similar conclusions. In Punjab, the mABSI was more specific and the R-Baux more sensitive,³⁰ the same trade-off as at the dichotomized cutoffs in our unit. Surabaya data showed near-identical accuracy for the ABSI and the R-Baux,¹⁰ a second Surabaya series found no meaningful difference among the R-Baux, the Belgian Outcome of Burn Injury score and the ABSI,³⁷ and multi-score comparisons in Germany⁴⁴ and Croatia³⁵ found differences among the leading scores small or non-significant, whereas in Poland the R-Baux discriminated better than the ABSI (AUC 0.86 vs 0.78)³⁸ and in Brazil, where the R-Baux was not tested, burn-specific indices outperformed general intensive-care scores.⁴⁵ Only studies with several hundred deaths can separate scores whose AUCs differ by a few

hundredths. In the present cohort the paired comparison excluded differences larger than 0.15 in every admissible data set, which is informative against the claim that either score is grossly superior but cannot exclude the small differences reported in large registries.

Why burn size carried the discrimination

The finding that neither composite score added discrimination to %TBSA alone follows from the composition of the cohort. Age, the second pillar of both scores, was not associated with death: mortality was 47.8% in patients aged 40 years or younger and 38.1% in those aged 41–70, only one patient was older than 70, and the attainable adjusted OR for age spanned 1. Both scores were derived in populations in which age contributes strongly — the R-Baux in an American repository and the mABSI in a

German registry whose age scale was rebuilt because mortality rises non-linearly with age^{22,26} — and their accuracy falls at the extremes of age.²³ Where patients are older, age-related information adds substantially: frailty raised the odds of death threefold after adjustment for the R-Baux in a British cohort,⁴⁶ frailty and ambulation separated survivors from non-survivors among Turkish patients aged 65 years or more,⁴⁷ and the R-Baux under-predicted mortality in older Canadians.³⁹ In a young cohort such as ours the age component adds little variance, and the scores collapse towards burn size, as summarized in Figure 11. Inhalation injury, the third shared component, was documented in only six patients, too few to contribute; in a Beijing intensive-care cohort, by contrast, 69.1% of patients had inhalation injury.⁴⁸

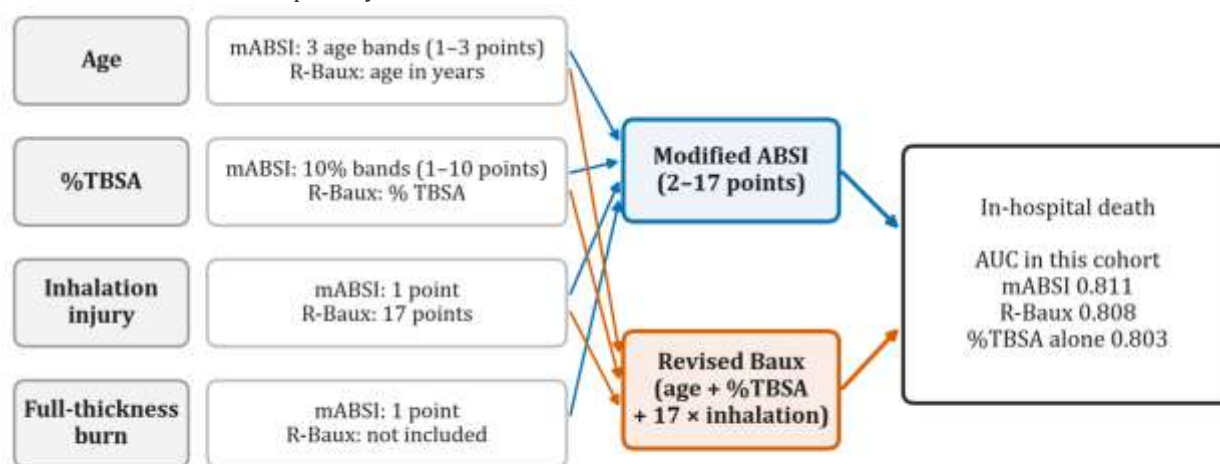


Figure 11. How the two scores weight the same admission variables, and the AUCs observed for each score and for burn size alone. The downstream pathway shown at the bottom was not measured in this study.

Burn size itself behaved as in other referral populations. The 50%-mortality burn size of about 39% TBSA in our unit is close to the lethal area 50 of 37% reported in a mixed adult and pediatric cohort in Bloemfontein, South Africa,⁴⁹ 42.6% for men with partial-thickness burns in Damascus⁸ and 34% among French patients older than 65 years with burns of at least 10% TBSA,⁵⁰ and below the 48.0% reported in the Dominican Republic.⁷ In Jakarta, burn size discriminated death with an AUC of 0.813 and an optimal cutoff of about 33.8%, burns of 31–40% carried an adjusted OR of 8.27 and burns over 40% an OR of 22.3,⁹ mirroring our OR of 10.29 for burns over 30%. In British Columbia the lethal score 50 (LS50, based on the Baux score) improved over four decades,⁴³ which shows that these thresholds are properties of a care system as much as of the injury.

Cutoffs, paired classification and combined rules

The Youden-optimal R-Baux cutoff of ≥ 81 agrees closely with the cutoffs of 81.5 in Addis Ababa³³ and 80 in West Virginia,³² and lies between the 74.5 reported in Uganda³¹ and the 85 reported in South India,⁵¹ whereas a cutoff above 100 (105.5) was reported in north-western Syria, where mortality was lower (18.0%).³⁴ In Saudi Arabia no patient with an R-Baux of 110 or more survived and none below 36 died.⁵² For the original ABSI, an Indian prospective series observed 100% mortality above 11,⁵³ and a South Indian series an optimal cutoff of 7.⁵¹ The cutoffs obtained here should be regarded as local and provisional: the mABSI optimum moved to ≥ 5 in a third of bootstrap resamples, and the R-Baux optimum, although more stable, rests on 20 deaths.

The structure of the paired classification has a practical implication. Because every patient classified as high risk by the mABSI was also high risk by the R-Baux at the dichotomization used in the unit, the two scores were nested at these thresholds rather than complementary — a property of the thresholds chosen, not of the scores: combining them added nothing, and the choice between them was a choice between sensitivity (R-Baux ≥ 71 : 85.0%) and specificity (mABSI ≥ 6 : 72.0%). At the Youden-optimal cutoffs, an either-positive rule could detect at most two more deaths than mABSI ≥ 6 (17 vs 15) at the cost of up to three more false positives (10 vs 7), or at most three more deaths than R-Baux ≥ 81 at the cost of up to six more false positives (10 vs 4), among 25 survivors. For a unit with few intensive-care beds, a single score at a deliberately chosen threshold is simpler and loses nothing.

Calibration and the excess of deaths

The excess of observed over expected deaths is the clearest result of the study and the one with the most direct consequence for counseling. More than a quarter of patients in the 'moderate' mABSI category died against a published 2%, and more than half of those in the 'moderately severe' category died against 10–20%. The direction agrees with the multinational validation in which the modified score fitted high-income countries but the original ABSI, which predicts higher mortality at the same score, fitted middle- and low-income countries better,²⁸ and with the under-prediction by the R-Baux in older Canadians (observed/expected 1.30).³⁹ It is opposite to the German registry experience, in which patients with ABSI values of 12 or more without full-thickness burns survived in 46% of cases

against less than 10% predicted.²⁹ Risk estimates exported from registries therefore need recalibration before they are quoted to families.

Several features of burn care in the region plausibly contribute to the excess. Sepsis caused 56.4% of deaths in a resource-limited unit in Damascus⁸ and 40.1% in the Dominican registry,⁷ infection-related organ failure accounts for most late deaths in intensive care,^{13,14} and acute kidney injury and gastrointestinal dysfunction mark patients at very high risk.^{15,16} Organ-dysfunction scores track these trajectories only partly in burn patients with infection.⁵⁴ Timing of excision differs between settings: delayed excision was associated with higher mortality in a French prospective cohort²⁰ and in burns of 70% TBSA or more,²¹ completing excision within a week reduced bacteremia in Japanese patients with large flame burns,⁵⁵ and early excision reduced six-month mortality in older French patients with a mean age of 83 years.⁵⁶ Yet the benefit is context-dependent: an early-excision priority system was associated with higher mortality than a conservative approach in a resource-restricted South African unit,⁵⁷ a North Indian randomized trial of early excision shortened hospital stay without changing mortality,⁵⁸ and the mortality of a Karnataka center fell from 23.9% to 7.9% over five years as early tangential excision and allografts were introduced.⁵ Pre-hospital delay, transfer after partial management elsewhere and the availability of intensive-care beds could not be measured here, but all of them operate after the score is calculated and push observed mortality above registry expectations.

Selection could have moved the observed/expected ratio in either direction. If the 29 records excluded as incomplete belonged mainly to patients who died early, before full documentation, the true excess would be larger; if they belonged mainly to short stays with minor burns, the cohort would be enriched for severe injury and the excess partly artifactual. The second explanation cannot account for the excess within score categories, however, because calibration compares patients with the same score: 27.8% mortality at mABSI 4–5 is not a product of enrichment for higher scores. Some contributors to the excess are likely to apply to other Indonesian provincial units — referral after initial management elsewhere, travel time across a large province, limited intensive-care capacity and a case mix dominated by flame injury — whereas others, such as the unit's excision practice and staffing during the study year, are specific to RSMH. Recalibration should therefore be repeated in each unit, or in a pooled registry with unit-level terms, rather than transferred from this study.

The local risk curves offer a practical substitute. The mABSI value with 50% predicted mortality was 6.3 points in our unit, whereas the published categories place 50% mortality at the boundary between the 8–9 (30–50%) and 10–11 (60–80%) categories; the corresponding R-Baux value of about 85 is far below the mean R-Baux of 130 in a British series of massive burns¹² and the value of 110 above which no Saudi patient survived⁵² and close to the mean R-Baux of 80.6 among patients who died in Bloemfontein.⁴⁹ Until multicenter Indonesian data are available, these local estimates, presented with their uncertainty, may inform the counseling of families at RSMH better than the published category probabilities, although they are in-sample estimates that require external validation. The decision-curve analysis quantifies the cost of not doing so: at a 20% threshold, triage based on the published categories yielded about half the net benefit of the locally calibrated score.

Score components

Full-thickness involvement, the variable that distinguishes the mABSI from the R-Baux, was not associated with death in our cohort, and its adjusted effect could not be identified. Larger cohorts consistently find it predictive: full-thickness burns were independent predictors in a German intensive-care cohort,⁴⁴ carried an adjusted hazard ratio of 2.6 in Addis Ababa intensive-care units⁵⁹ and an adjusted OR of 4.33 in the Dominican registry,⁷ and deep burns were associated with death in a two-center Ethiopian cohort.⁶⁰ Third-degree burns were an independent predictor in a Florida cohort,⁶¹ and third-degree area predicted death in Beijing.⁴⁸ The German registry analysis offers the most plausible explanation for the null result here: full thickness mattered only above 20% TBSA,²⁹ where burn size already places most patients in the high-risk range. The recorded depth ranges point to a second explanation. Patients whose full-thickness burns were recorded together with superficial components died far more often than those whose full-thickness burns were recorded with mid-dermal components, which suggests that the depth range reflects the extent and pattern of a burn — large mixed-depth flame burns as opposed to localized deep contact or electrical burns — more than depth alone. Because this finding is exploratory, it should be confirmed in data that record the percentage of TBSA at each depth.

Flame injury was the only etiology associated with death, a pattern shared with Chinese meta-analytic data⁴ and a Turkish geriatric series in which flame burns accounted for 87.5% of deaths,⁴⁷ and consistent with the predominance of flame and explosion injuries (79.4% of admissions) at the Jakarta referral center.⁹ Neither score was designed for electrical injury, in which the visible %TBSA under-represents deep tissue damage, or for chemical injury; together these accounted for a quarter of our admissions. Their inclusion probably lowered discrimination slightly, because most such patients had small burns and survived — which the scores predict correctly — while their occasional deaths are driven by mechanisms the scores do not capture. The study data did not cross-tabulate the scores by etiology, so an analysis restricted to flame and scald burns was not possible; future registries should report score performance separately for electrical injury. Inhalation injury showed the expected direction (66.7% vs 41.0% mortality) but involved only six patients. In larger data sets it carried a pooled OR of 6.67 in a Chinese meta-analysis,⁴ an OR of 8.48 in Pakistan⁶ and an adjusted OR of 6.53 in Ethiopia,⁶⁰ although a Brazilian intensive-care cohort found similar mortality with and without it.¹⁹ Sex was not associated with death here, whereas female sex carried higher mortality in Kuwait,⁴² Pakistan⁶ and Ethiopia³³ and a higher risk of bacteremia in Switzerland;⁶² the mABSI's exclusion of sex therefore deserves testing in larger Indonesian cohorts. Body weight, which modifies burn mortality in a large electronic health record database,⁶³ was not recorded.

Beyond the two scores

More complex models discriminate better in development data but rarely transfer. Gradient-boosted trees reached an AUC of 0.966 in Isfahan, but sensitivity fell from 78.1% to 68.3% on temporal validation;⁶⁴ random forests reached 0.882 in an Iranian cohort;⁶⁵ LightGBM models using the lactate dehydrogenase-to-lymphocyte ratio 0.916 in Shanghai⁶⁶ and models for burns of 50% TBSA or more 0.981;⁶⁷ and a CatBoost model with APACHE IV outperformed the ABSI and the R-Baux in a public data set of Korean patients.⁶⁸ In a Swedish intensive-care cohort, however, machine-learning models did not clearly outperform the Baux score.⁶⁹ Among conventional scores, the BUrn Mortality Prediction (BUMP) score (AUC 0.947) improved

slightly on the R-Baux in German registry data,²⁷ the Bogenhausen ABSI (BABSI) reached 0.963 in a German single-center cohort,⁷⁰ the ABSI reached 0.909 over 20 years in Vienna,⁷¹ a Nanchang logistic model 0.868 against a best R-Baux performance in patients aged 60 or more,⁷² and a simple model of age, TBSA and inhalation injury performed as well as the composite scores in Poland.³⁸ The ABSI predicted mortality as well as a general intensive-care score in São Paulo,⁷³ and ABSI-based estimates were closest to observed mortality in a Turkish burn center.⁷⁴ Laboratory variables, comorbidities and frailty add information,^{46,61,71} but they are not available at the bedside on arrival in every Indonesian referral unit, and admission scores that need only four variables keep their place for initial triage. Risk-adjusted outcome monitoring with the R-Baux has been sustained for a decade in a regional burn center,⁷⁵ and a nationwide Japanese analysis found no association between hospital volume and mortality,⁷⁶ suggesting that benchmarking quality with a locally calibrated score is feasible even in smaller units.

Implications for practice

Three implications follow for RSMH and similar Indonesian units. First, either score can be used to rank patients at admission, and the choice can rest on familiarity and the clinical purpose: an R-Baux of 71 or more identifies 85% of patients who will die at the cost of classifying 12 of 25 survivors as high risk, which suits a unit with spare intensive-care capacity, and an mABSI of 6 or more or an R-Baux of 81 or more identifies a smaller group with positive likelihood ratios of 2.68 and 4.37, respectively. Second, the published mABSI probabilities should not be quoted to families; a patient with an mABSI of 4–5 faced about a one-in-four risk of death in this unit rather than one in fifty, and the local estimates listed in Table 7, presented with their confidence intervals, may be a better basis for counseling, recognizing that they are in-sample estimates from 20 deaths that require external validation. Third, because burn size carried nearly all the discrimination, the accuracy of the %TBSA estimate at admission is the most important input to either score. Training in TBSA estimation, documentation of the percentage burned at each depth and recording of the raw R-Baux value rather than a band are inexpensive steps that would improve both triage and future validation.

Research agenda

Four studies would resolve the questions left open. First, a prospective multicenter registry of burn admissions in Sumatra and other Indonesian provinces should record raw age, raw %TBSA, the percentage of TBSA at each depth, inhalation injury confirmed by bronchoscopy, time from injury to admission and the treatment pathway, so that both scores can be recalibrated and compared on identical patients. Second, detecting an AUC difference of 0.05 between the scores with 80% power would require about 132 patients if the scores are as correlated as the most correlated admissible data set and about 772 if they are as weakly correlated as the least; estimating a sensitivity of 75% to within ± 10 percentage points requires about 162 patients at the mortality observed here, and to within ± 5 points about 648. Third, a stepped implementation study should test whether a locally calibrated score linked to an admission pathway — intensive-care referral, early excision planning and structured family meetings — changes the timing of excision and mortality. Fourth, cost and capacity analyses should establish which threshold of predicted risk is appropriate for intensive-care admission in a unit with fixed bed numbers, using decision curves estimated in the registry data.

Strengths and limitations

The study has several strengths. Both index tests were applied to every patient against an objective reference standard with no missing outcomes, all single-score analyses were computed exactly from fully observed tables, the paired analyses were bounded without imputing data, and the calibration of the published risk categories was tested directly rather than assumed. To our knowledge, it is also the first comparison of the two scores in the same patients outside Java.

The limitations are substantial. This was a retrospective single-center study with 20 deaths, so every estimate is imprecise, the cutoffs are provisional and differences between the scores smaller than about 0.15 in AUC cannot be excluded. More than half of the admissions were excluded, 29 of them for incomplete records; incompleteness may be related to severity and outcome, and spectrum bias cannot be ruled out. The R-Baux was available only in 10-point bands, so its raw-score AUC could lie anywhere between 0.764 and 0.852, published cutoffs could not be replicated exactly, and raw age was available only as a cohort summary. Accuracy at the Youden-optimal cutoffs is optimistic because the cutoffs were selected in the same data, although the optimism-corrected Youden indices remained 0.40 and 0.50. The outcomes of the 29 patients excluded for incomplete records are unknown. Burn depth and inhalation injury were assessed clinically and bronchoscopy was not performed in every patient. The joint distribution of the two scores and of the components was not recorded, so paired and adjusted analyses are reported as ranges, and the adjusted effects of age, inhalation injury and full thickness could not be identified. Finally, time from injury, comorbidity, cause of death and the timing of excision were not available, so the reasons for the excess mortality remain hypotheses.

5. Conclusion

In consecutive burn admissions to the RSMH burn unit in Palembang, the Modified ABSI and the Revised Baux score ranked the risk of in-hospital death equally well (AUC 0.811 and 0.808), and neither ranked it detectably better than burn size alone, with differences larger than about 0.15 in AUC excluded. The Youden-optimal cutoffs were an mABSI of 6 or more (sensitivity 75.0%, specificity 72.0%) and an R-Baux of 81 or more (70.0% and 84.0%); at the dichotomization used in the unit the R-Baux was the more sensitive and the mABSI the more specific classification, and combining them added nothing. Mortality was 2.36 times that implied by the published mABSI risk categories (calibration-in-the-large intercept 2.26), and half of the patients died at an mABSI of about 6.3 and an R-Baux of about 85. Indonesian burn units should consider locally calibrated risk estimates, once validated, for triage and family counseling, record raw scores and depth-specific burn size, and contribute to a multicenter registry large enough to recalibrate both scores and to test whether risk-guided admission pathways improve survival.

6. Declarations

Ethics Approval and Consent to Participate

The study protocol received an ethical exemption from the Health Research Ethics Committee of Dr. Mohammad Hoesin General Hospital Palembang (No. DP.04.03/D.XVIII.06.08/ETIK/033/2026; 30 January 2026). Individual informed consent was waived because only de-identified existing medical records were used. The study was conducted in accordance with the Declaration of Helsinki.

Consent for Publication

Not applicable; the article contains no individual patient data, images or other identifying information.

Availability of Data and Materials

All aggregate data analyzed are contained in the tables of this article. The de-identified patient-level dataset is held by the Burn Unit, Dr. Mohammad Hoesin General Hospital Palembang, and is available from the corresponding author on reasonable request, subject to approval by the hospital's ethics committee. The analysis code is available from the corresponding author.

Artificial Intelligence Use Statement

The authors did not use artificial intelligence tools in preparing this article.

Author Contribution Statement

I.T.I.: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft. A.A.: Conceptualization, Supervision, Validation, Writing – review & editing. T.: Methodology, Formal analysis, Supervision, Writing – review & editing. All authors read and approved the final manuscript.

Funding Statement

This research received no specific grant 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 competing interests.

Acknowledgements

The authors thank the nursing and medical-records staff of the Burn Unit, Dr. Mohammad Hoesin General Hospital Palembang.

7. References

- Shao J, Zhu Y, Xie M, et al. Assessing the global burns burden in the context of the WHO Prevention and Care Strategy: insights from the Global Burden of Disease Study 2021. *Nurs Health Sci.* 2026;28(2):e70329. doi: 10.1111/nhs.70329.
- Hauc SC, Rodriguez VIG, Atayde JEP, et al. Projecting the incidence of burn injuries and its associated economic burden for 195 countries. *Burns.* 2026;52(8):108066. doi: 10.1016/j.burns.2026.108066.
- Abarca L, Guilabert P, Martin N, et al. Epidemiology and mortality in patients hospitalized for burns in Catalonia, Spain. *Sci Rep.* 2023;13(1):14364. doi: 10.1038/s41598-023-40198-2.
- Li P, Zhang Q, Li D. Epidemiology and prognosis of burn injuries in China: a meta-analysis. *Eur J Med Res.* 2025;30(1):581. doi: 10.1186/s40001-025-02741-x.
- Patel P, Rao T, Thomas J, et al. Epidemiology and outcomes of burn injuries: a retrospective study at a tertiary care center. *Ann Plast Surg.* 2026;97(3S):S238-44. doi: 10.1097/sap.0000000000004799.
- Khalid S, Butala S, Alam Z, et al. Burn epidemiology and outcome disparities: a study of patients hospitalized with burns in Pakistan. *Adv Skin Wound Care.* 2026. Epub ahead of print. doi: 10.1097/asw.0000000000000490.
- Cueva-Ramírez JE, Valdez R, Tejada R, et al. Expanded burn registry: epidemiological profile, clinical variables, and multivariate analysis of outcomes in the Dominican Republic (2020–2024). *Burns.* 2026;52(8):108079. doi: 10.1016/j.burns.2026.108079.
- Domat L, Azar A, Alkheder A, et al. Epidemiology, management, and alarmingly high mortality of burn patients: a retrospective analysis from a resource-limited tertiary centre in Syria. *Int Wound J.* 2026;23(4):e70910. doi: 10.1111/iwj.70910.
- Wardhana A, Leksono TP, Sanjaya IS, et al. Association of total body surface area on effective surgical treatment and mortality in adult burn patients: a 5-year retrospective analysis in an Indonesian national referral hospital. *Turk J Emerg Med.* 2026;26(2):94-101. doi: 10.4103/tjem.tjem_246_25.
- Iustitiati MAA, Nata'atmadja BS. Comparison of Abbreviated Burn Severity Index (ABSI) score with R-Baux score as a predictor of mortality in burn patients. *J Rekonstr Estet.* 2022;7(2):43-50. doi: 10.20473/jre.v7i2.41216.
- Herlianita R, Purwanto E, Wahyuningsih I, et al. Clinical outcome and comparison of burn injury scoring systems in burn patient in Indonesia. *Afr J Emerg Med.* 2021;11(3):331-4. doi: 10.1016/j.afjem.2021.04.005.
- Walker H, Chipp E. Short and long term mortality following massive burn injury (>50% TBSA) in a UK burns centre. *Burns.* 2022;48(8):1845-9. doi: 10.1016/j.burns.2022.02.006.
- McIver WJ, Mullhi R, Chipp E, et al. Multi-organ failure in adult patients admitted to intensive care with severe burns—a retrospective cohort study. *Eur Burn J.* 2026;7(3):41. doi: 10.3390/ebj7030041.
- Tedesco DJ, Hutter MF, Khalaf F, et al. Sepsis in burn care: incidence and outcomes. *Mil Med Res.* 2025;12(1):55. doi: 10.1186/s40779-025-00643-x.
- Chen X, Zhou J, Wang Y, et al. The prevalence, risk factors, and in-hospital mortality of acute kidney injury according to KDIGO criteria in burn patients. *Ren Fail.* 2025;47(1):2564376. doi: 10.1080/0886022x.2025.2564376.
- He QL, Gao SW, Qin Y, et al. Gastrointestinal dysfunction is associated with mortality in severe burn patients: a 10-year retrospective observational study from South China. *Mil Med Res.* 2022;9(1):49. doi: 10.1186/s40779-022-00403-1.
- Virdis E, Fara G, Paliogiannis P, et al. The prognostic role of inhalation injury in burn patients assessed by flexible bronchoscopy: a systematic review and meta-analysis. *Front Med (Lausanne).* 2026;13:1808915. doi: 10.3389/fmed.2026.1808915.
- Fuciños LC, Rotaache AA, Cabezas EF, et al. Cardiopulmonary hemodynamic alterations during early resuscitation in burn patients with inhalation injury: do they impact prognosis? *J Burn Care Res.* 2026;47(5):1549-57. doi: 10.1093/jbcr/irag096.
- Normann TC, Bellaver P, Limeira BI, et al. Association of inhalation injury with ventilator-associated pneumonia and clinical outcomes in critically ill patients with burns: a retrospective study. *Can J Respir Ther.* 2026;62:64-73. doi: 10.29390/001c.159978.
- Legrand M, Hoffmann C, Roquilly A, et al. Early surgery and outcomes in severe burns: a prospective multicenter observational study. *Ann Surg.* 2026. Epub ahead of print. doi: 10.1097/sla.0000000000007201.
- Huang R, Yao Y, Xie S, et al. Optimal timing and approaches for first tangential excision in patients with ≥70% total body surface area burns: insights from a retrospective cohort of a mass casualty incident. *Ann Surg.* 2026;283(6):959-71. doi: 10.1097/sla.0000000000006626.
- Osler T, Glance LG, Hosmer DW. Simplified estimates of the probability of death after burn injuries: extending and updating the Baux score. *J Trauma.* 2010;68(3):690-7. doi: 10.1097/ta.0b013e3181c453b3.
- Edgar MC, Bond SM, Jiang SH, et al. The revised Baux score as a predictor of burn mortality: a systematic review and meta-analysis. *J Burn Care Res.* 2023;44(6):1278-88. doi: 10.1093/jbcr/irad075.
- Mrad MA, Al Qurashi AA, Shah Mardan QNM, et al. Risk models to predict mortality in burn patients: a systematic review and meta-analysis. *Plast Reconstr Surg Glob Open.* 2022;10(12):e4694. doi: 10.1097/gox.00000000000004694.
- Deng G, Tang L, Yang Q, et al. Association of the abbreviated burn severity index with mortality in severely burned patients: a meta-analysis. *PLoS One.* 2025;20(2):e0319199. doi: 10.1371/journal.pone.0319199.
- Bartels P, Thamm OC, Elrod J, et al. The ABSI is dead, long live the ABSI - reliable prediction of survival in burns with a modified Abbreviated Burn Severity Index. *Burns.* 2020;46(6):1272-9. doi: 10.1016/j.burns.2020.05.003.
- Bagheri M, Fuchs PC, Lefering R, et al. The Burn Mortality Prediction (BUMP) score - an improved mortality prediction score

- based on data of the German burn registry. *Burns*. 2023;49(1):110-9. doi: 10.1016/j.burns.2022.02.007.
28. Elrod J, Mohr C, Branski L, et al. Country economic status is strongly associated with burn survival - validation of the (modified) ABSI. *Int J Equity Health*. 2025;24(1):5. doi: 10.1186/s12939-024-02353-7.
 29. Gorenflo J, Reichert B, German Burn Registry, et al. The prognostic relevance of full-thickness burns on ABSI. *Burns*. 2024;50(9):107197. doi: 10.1016/j.burns.2024.07.001.
 30. Shah S, Verma R, Mittal RK, et al. Better among the two for burn mortality prediction in developing nations: revised Baux or modified Abbreviated Burn Severity Index? *Int J Appl Basic Med Res*. 2024;14(1):7-11. doi: 10.4103/ijabmr.ijabmr_350_23.
 31. Baraka SM, Kiswezi A, Olasinde AA, et al. Role of the revised Baux score in predicting mortality among burn patients in an African low-income country: a multicentre prospective cohort. *Ann Med Surg (Lond)*. 2024;86(8):4364-7. doi: 10.1097/ms9.0000000000001774.
 32. Rahimpour A, Fox N, Kahley G, et al. Burn mortality in an Appalachian referral center: an examination of mortality prediction scores in a 13-year retrospective study. *Cureus*. 2024;16(6):e62912. doi: 10.7759/cureus.62912.
 33. Belayneh ES, Abebe MW, Fikremariam M, et al. Accuracy of the revised Baux score for predicting in-hospital mortality of patients with burns: a retrospective cohort study from Ethiopia. *J Plast Reconstr Aesthet Surg*. 2026;119:375-81. doi: 10.1016/j.bjps.2026.06.024.
 34. Karaca B, Çelik B, Emem MK. Evaluation of clinical outcomes and comparison of prediction models in the burn population hospitalized from the emergency department: can burn mortality scores be used in a post-conflict area such as northwest Syria? *Ulus Travma Acil Cerrahi Derg*. 2023;29(3):409-18. doi: 10.14744/tjtes.2023.17731.
 35. Skunca A, Mesic A, Zagorac D, et al. Predicting mortality in severe burns: a comparison of four mortality prediction scores and the role of organizational changes in the Croatian Burn Center. *Eur Burn J*. 2024;5(4):410-7. doi: 10.3390/ebj5040036.
 36. Sharma S, Tandon R. External validation of three burn-specific mortality prediction models in adult burn patients at a tertiary care hospital in India. *Indian J Plast Surg*. 2023;56(1):39-43. doi: 10.1055/s-0043-1760825.
 37. Primabudi ARI, Budi AS, Saputro ID. Comparison of revised Baux, Belgian Outcome of Burn Injury, and Abbreviated Burn Severity Index scores as model predictors of mortality in burn patients: a retrospective study. *J Wound Manag Res*. 2025;21(2):79-84. doi: 10.22467/jwrmr.2024.03174.
 38. Tomaka P, Rzońca P, Rzońca M, et al. Prognostic value of mortality scoring systems in patients with severe burns: identifying key predictors of mortality and comparative analysis between survivors and non-survivors. *Med Sci Monit*. 2026;32:e951713. doi: 10.12659/msm.951713.
 39. Goder N, Meghdadi Y, Thibedeau M, et al. Recalibrating the revised Baux score: admission and 48-hour landmark survival models in contemporary older adult patients with burn injury. *Burns*. 2026;52(8):108141. doi: 10.1016/j.burns.2026.108141.
 40. Costa S, Ferros C, Reigota A, et al. Prognostic scores for mortality in invasive mechanically ventilated burn patients. *Cureus*. 2024;16(6):e63502. doi: 10.7759/cureus.63502.
 41. Alshdowh H, Alshraideh JA, Al Qadire M, et al. Burn injuries in Jordan: a 5-year retrospective analysis of presentation, management and hospital mortality. *Int Wound J*. 2026;23(2):e70839. doi: 10.1111/iwj.70839.
 42. Hassan Z, Burhamah W, Alabdulmuhsen S, et al. The analysis and accuracy of mortality prediction scores in burn patients admitted to the intensive care burn unit (ICBU). *Ann Med Surg (Lond)*. 2021;65:102249. doi: 10.1016/j.amsu.2021.102249.
 43. Zhang JX, Ahmed SN, Pangli H, et al. Predicting and estimating burn outcomes: an institutional analysis of 4622 cases. *J Burn Care Res*. 2022;43(6):1426-33. doi: 10.1093/jbcr/irac051.
 44. Obed D, Salim M, Dastagir N, et al. Comparative analysis of composite mortality prediction scores in intensive care burn patients. *Int J Environ Res Public Health*. 2022;19(19):12321. doi: 10.3390/ijerph191912321.
 45. de Carvalho VF, Paggiaro AO, Goldner A, et al. Retrospective evaluation of the accuracy of five different severity scores to predict the mortality in burns patients. *J Burn Care Res*. 2024;45(5):1175-82. doi: 10.1093/jbcr/irab057.
 46. Smith OA, Tridente A, Sen SK. Can we predict mortality, length of stay, and time taken to heal at the point of admission in burn patients? *Burns*. 2026;52(1):107779. doi: 10.1016/j.burns.2025.107779.
 47. Özlü Ö, Başaran A. Elderly burns: Clinical frailty scale and functional ambulation classification in predicting prognosis. *Ulus Travma Acil Cerrahi Derg*. 2022;28(6):812-7. doi: 10.14744/tjtes.2022.49400.
 48. Wang C, Dou Z, Qin F, et al. Epidemiology and risk prediction of patients with severe burns admitted to a burn intensive care unit in a burn center in Beijing: a 5-year retrospective study. *Heliyon*. 2022;8(12):e12572. doi: 10.1016/j.heliyon.2022.e12572.
 49. Mathonsi K, Arko-Cobbah E. Outcomes of burns patients in a developing country: a single centre's experience. *S Afr Med J*. 2023;113(10):49-53. doi: 10.7196/samj.2023.v113i10.1029.
 50. Schmitt A, Lherm M, Brennac C, et al. Epidemiology and prognostic factors in older adults with severe burns ($\geq 10\%$ TBSA): a 13-year retrospective study from a French national burn center. *Burns*. 2026;52(8):108104. doi: 10.1016/j.burns.2026.108104.
 51. Shetty N, Mashalkar NS, Ellur SR, et al. Sensitivity and specificity of abbreviated burn severity index and revised Baux score in mortality prediction in burn patients admitted to a tertiary care center in South India. *Indian J Burns*. 2022;30(1):78-83. doi: 10.4103/ijb.ijb_15_22.
 52. Alshammari SM, Almarzouq S, Alghamdi AA, et al. Mortality and survival analysis of burn patients admitted in a critical care burn unit, Saudi Arabia. *Saudi J Med Med Sci*. 2022;10(3):216-20. doi: 10.4103/sjms.sjms_618_21.
 53. Usmani A, Pipal DK, Bagla H, et al. Prediction of mortality in acute thermal burn patients using the Abbreviated Burn Severity Index score: a single-center experience. *Cureus*. 2022;14(6):e26161. doi: 10.7759/cureus.26161.
 54. Gimenez FMP, Cardoso LTQ, Kerbaui G, et al. Analysis of the SOFA score, quick-SOFA, and SIRS criteria in burn patients with infection. *Rev Bras Enferm*. 2025;78(3):e20240111. doi: 10.1590/0034-7167-2024-0111.
 55. Kaita Y, Matsuda T, Yamaguchi Y. The significance of completing burn excision within one week after injury. *Burns*. 2025;51(5):107506. doi: 10.1016/j.burns.2025.107506.
 56. Cavadore N, Thuau F, Cavadore T, et al. Impact of surgical excision timing on the survival in geriatric burn patients: a retrospective study. *J Burn Care Res*. 2026;47(1):266-72. doi: 10.1093/jbcr/iraf153.
 57. Allorto N, Bishop DG. Selective early excision versus a conservative surgical approach. Management of deep burn injury in a resource-restricted setting: a prospective, observational study. *World J Surg*. 2026;50(3):523-8. doi: 10.1002/wjs.70262.
 58. Kakani A, Rao N, Mago V, et al. Impact of early excision in acute burns: a randomized controlled trial from a tertiary centre in North India. *Burns*. 2026;52(9):108200. doi: 10.1016/j.burns.2026.108200.
 59. Bayu HT, Fekede MS, Asrat WA, et al. Survival status and predictors of mortality among burn patients admitted to intensive care units at burn centres in Addis Ababa, Ethiopia: a retrospective follow-up study. *Int J Emerg Med*. 2026;19(1):75. doi: 10.1186/s12245-026-01188-z.
 60. Bayou RK, Demmissie MB, Bayou BK, et al. Predictors of mortality in burn patients at selected tertiary public hospitals in Addis Ababa, Ethiopia: a two-year retrospective study. *Eur Burn J*. 2026;7(2):28. doi: 10.3390/ebj7020028.
 61. Nehila T, Mikhael M, Arora S, et al. Revised Baux score identifies a new risk factor for mortality: history of diabetes. *Adv Skin Wound Care*. 2026;39(6):304-7. doi: 10.1097/asw.0000000000000471.
 62. Schweizer NJM, Scheier TC, Buehler PK, et al. Impact of sex on the development of bacteremia in critically ill burn patients: a

- retrospective cohort study. *Burns*. 2026;52(2):107845. doi: 10.1016/j.burns.2025.107845.
63. Pu SY, Huang YL, Pu CM, et al. The role of body weight in burn mortality. *Burns*. 2026;52(4):107950. doi: 10.1016/j.burns.2026.107950.
 64. Sabet Kouhanjani Y, Sattari M, Ehteshami A. From data to decisions: predicting inpatient burn mortality with advanced classification models. *PLoS One*. 2026;21(1):e0338564. doi: 10.1371/journal.pone.0338564.
 65. Motamedi M, Moallemkolaei NM, Hesamirostami M, et al. Simultaneous prediction of early and delayed mortality in burn patients: a comparative machine learning analysis of feature importance in a single-center retrospective study. *BMC Med Inform Decis Mak*. 2025;26(1):7. doi: 10.1186/s12911-025-03311-1.
 66. Sun N, Fan H, Zhu L, et al. Development of a LightGBM-based survival prediction model for major burn patients using LDH/LYM ratio as the core predictor. *Burns*. 2026;52(6):108022. doi: 10.1016/j.burns.2026.108022.
 67. Wang SQ, Qiu K, Zheng QR, et al. Development and validation of web-based, interpretable predictive models for sepsis and mortality in extensive burns. *Front Cell Infect Microbiol*. 2025;15:1586087. doi: 10.3389/fcimb.2025.1586087.
 68. Ren H, Xu Y. Interpretative machine learning for predicting 60-day mortality in burn patients with suspected infection. *World J Emerg Med*. 2026;17(3):244-9. doi: 10.5847/wjem.j.1920-8642.2026.046.
 69. Fransén J, Lundin J, Fredén F, et al. A proof-of-concept study on mortality prediction with machine learning algorithms using burn intensive care data. *Scars Burn Heal*. 2022;8:20595131211066585. doi: 10.1177/20595131211066585.
 70. Rein S, Schmiechen J, Gille J, et al. Prognostic scoring systems for burns: a comparative analysis of their predictive accuracies for mortality in burn patients. *Eur Burn J*. 2026;7(1):18. doi: 10.3390/ebj7010018.
 71. Christ A, Staud CJ, Krotka P, et al. Revalidating the prognostic relevance of the Abbreviated Burn Severity Index (ABSI): a twenty-year experience examining the performance of the ABSI score in consideration of progression and advantages of burn treatments from a single center in Vienna. *J Plast Reconstr Aesthet Surg*. 2024;94:160-8. doi: 10.1016/j.bjps.2024.04.041.
 72. Chen H, Wu X, Zou L, et al. A comparative study of the predictive value of four models for death in patients with severe burns. *Burns*. 2024;50(3):550-60. doi: 10.1016/j.burns.2023.10.019.
 73. Melo FDL, Gragnani A, de Oliveira AF, et al. Predicting mortality for critically ill burns patients, using the Abbreviated Burn Severity Index and Simplified Acute Physiology Score 3. *Injury*. 2022;53(2):453-6. doi: 10.1016/j.injury.2021.11.027.
 74. Yazıcı H, Uçar AD, Namdaroğlu O, et al. Mortality prediction models for severe burn patients: which one is the best? *Ulus Travma Acil Cerrahi Derg*. 2022;28(6):790-5. doi: 10.14744/tjtes.2021.29540.
 75. Eriksson S, Martin N, Smailes S, et al. Cumulative sum (CUSUM) analysis: a 10 year experience of outcome monitoring in a regional burn centre. *Burns*. 2024;50(8):1999-2006. doi: 10.1016/j.burns.2024.07.035.
 76. Watanabe T, Watanabe A, Iwagami M, et al. The association between hospital volume and in-hospital mortality in severe burn patients: a nationwide study using the Japanese Burn Registry. *Burns*. 2026;52(5):107959. doi: 10.1016/j.burns.2026.107959.