

Boey Score as a Predictor of Mortality in Patients with Perforated Peptic Ulcer After Laparotomy at a Private Regional Hospital: An Analytic Study at Efarina Etaham Hospital, Pematang Siantar, Indonesia, 2022–2026

Rinadi Andara*

* Department of Surgery, Efarina Etaham Pematang Siantar Hospital, North Sumatra, Indonesia
dr.rinadi.andara@gmail.com

DOI: 10.29322/IJSRP.16.10.2026.p17810

<https://dx.doi.org/10.29322/IJSRP.16.10.2026.p17810>

Paper Received Date: 4th September 2026

Paper Acceptance Date: 6th October 2026

Paper Publication Date: 18th October 2026

Abstract- Perforated peptic ulcer (PPU) remains a major surgical emergency with persistently high postoperative mortality, particularly in resource-limited private regional hospitals where simple, rapid risk-stratification tools are most needed. This retrospective analytic cohort study evaluated the Boey score (preoperative shock, major comorbidity, and delayed presentation >24 hours) as a predictor of postoperative mortality in 67 patients with PPU who underwent laparotomy at RS Efarina Etaham, Pematang Siantar, Indonesia, between November 2022 and June 2026, using total sampling. Overall mortality was 25.4%. Delayed presentation >24 hours was the most frequent component (94.0%), followed by major comorbidity (61.2%) and preoperative shock (28.4%). Mortality rose steeply with increasing Boey score, from 0% at scores 0–1 to 87.5% at score 3 (chi-square $p < 0.001$; logistic regression OR 57.42 per 1-point increase, 95% CI 9.05–364.13, $p < 0.001$). Receiver operating characteristic analysis showed outstanding discriminative ability, with an area under the curve (AUC) of 0.936 (95% CI 0.873–0.986). The Youden-optimal cut-off was Boey score ≥ 3 (sensitivity 82.4%, specificity 96.0%), while a cut-off of ≥ 2 achieved perfect sensitivity (100%) as a screening threshold. These findings confirm the Boey score as a robust, low-cost bedside prognostic tool for mortality risk stratification in PPU patients treated at private regional hospitals with limited diagnostic resources.

Index Terms- Boey score, laparotomy, mortality, perforated peptic ulcer, regional hospital.

I. INTRODUCTION

Perforated peptic ulcer (PPU) is one of the most common abdominal surgical emergencies and remains a leading cause of morbidity and mortality worldwide, despite advances in acid-suppressive therapy and *Helicobacter pylori* eradication that have reduced the overall incidence of uncomplicated peptic ulcer disease.[1][2] Reported incidence ranges from 3.8 to 14 per 100,000 population per year, with marked geographic variation driven by *H. pylori* prevalence, patterns of non-steroidal anti-inflammatory drug (NSAID) use, and lifestyle factors such as smoking and alcohol consumption.[3]

Postoperative mortality after PPU repair remains high and has changed little over recent decades, reported at 10–20% globally and exceeding 30% in elderly patients or those with delayed diagnosis and treatment.[1][2] The multicentre GRACE cohort study of 1,874 patients from multiple countries reported a 30-day mortality of 9.3%, with significant variation

between high-income and low-to-middle-income countries.[4] This persistently high and variable mortality underscores the importance of early identification of high-risk patients so that treatment can be intensified promptly.

Several prognostic scoring systems have been developed to predict outcomes in PPU, including the Boey score, the Peptic Ulcer Perforation (PULP) score, the Mannheim Peritonitis Index (MPI), the Acute Physiology and Chronic Health Evaluation (APACHE) II, and the American Society of Anesthesiologists (ASA) physical status classification.[5][6][7][8] Among these, the Boey score, introduced by Boey and colleagues in 1987, remains one of the most widely used because of its simplicity: it comprises only three clinical variables assessable immediately at presentation — preoperative shock (systolic blood pressure < 90 mmHg), major comorbidity, and delayed presentation > 24 hours from symptom onset.[9] The score has demonstrated good predictive value for postoperative mortality and morbidity across populations, with a reported area under the receiver operating

characteristic curve (AUC) of around 0.79 in validation studies from India.[10][11] The 2020 World Society of Emergency Surgery (WSES) guidelines recommend the Boey score as one of the risk-stratification tools for PPU (grade 2C recommendation), while noting the need for further validation across diverse populations.[12]

In Indonesia, data on the predictive performance of the Boey score remain limited and are largely confined to tertiary referral hospitals in major cities. A study at RSUD Dr. Soetomo Surabaya involving 65 PPU patients reported an overall mortality of 53.8%, with preoperative shock as the dominant predictive factor, but did not report an overall AUC for the Boey score.[13] Data on Boey score performance in private regional hospitals — which typically face greater diagnostic resource constraints and more frequent delays in referral — remain scarce, even though risk stratification based on simple clinical data is arguably most valuable in exactly this setting.

Efarina Etaham Pematang Siantar Hospital is a private regional hospital in North Sumatra that manages a substantial caseload of PPU, predominantly among patients with limited healthcare access and a tendency toward delayed presentation. Preliminary descriptive data from 67 post-laparotomy PPU patients at this hospital (2022-2026) showed an overall mortality of 25.4%, within the range reported in the literature, but the relationship with the Boey score and its individual components had not previously been formally analyzed. This study therefore aimed to analyze the ability of the Boey score and its components to predict postoperative mortality in PPU patients undergoing laparotomy at Efarina Etaham Pematang Siantar Hospital.

II. MATERIALS AND METHODS

A. Study Design and Setting

This was a retrospective analytic observational (cohort) study conducted at the Department of Surgery, Efarina Etaham Pematang Siantar Hospital, North Sumatra, Indonesia. Data were collected in August-September 2026 from the electronic medical records of patients with perforated peptic ulcer (PPU) who underwent laparotomy between November 2022 and June 2026, following institutional ethical clearance.

B. Population and Sample

The study population comprised all patients diagnosed with PPU who underwent laparotomy at the study site during the study period. A total sampling method was used: all patients meeting the inclusion criteria and none of the exclusion criteria were enrolled, yielding a final sample of 67 patients.

Inclusion criteria: age ≥ 18 years with PPU diagnosed on clinical grounds (acute abdominal pain with peritonism), radiological findings (free subdiaphragmatic air and/or pneumoperitoneum on CT), and confirmed intraoperatively; and complete medical-record data for preoperative hemodynamic status, comorbidity, symptom onset, and outcome.

Exclusion criteria: age < 18 years; perforation due to abdominal trauma; perforation due to gastric/duodenal malignancy; incomplete data for any Boey score component; and discharge against medical advice before the final outcome (survival/death) could be determined.

C. Variables

The independent variables were the Boey score and its three components: preoperative shock (systolic blood pressure < 90 mmHg at emergency triage), major comorbidity (a clinically significant systemic comorbid illness identified preoperatively, e.g. diabetes, hypertension, chronic obstructive pulmonary disease, renal or hepatic impairment), and delayed presentation (symptom-to-operation interval ≥ 24 hours). Each component was scored as present (1 point) or absent (0 points), giving a total Boey score of 0-3. Postoperative complications (e.g. anastomotic/patch leakage) were not included in the comorbidity component, since the Boey score by definition reflects only the patient's preoperative condition. The dependent variable was postoperative mortality (in-hospital death vs. survival to discharge).

D. Statistical Analysis

Analyses were performed in Python (SciPy, scikit-learn, statsmodels). Baseline characteristics and component distributions were summarized descriptively. The association between categorical Boey score and mortality was tested using Pearson's chi-square test, with Fisher's exact test applied where expected cell counts were < 5 . Binary logistic regression was used to estimate the odds ratio (OR) and 95% confidence interval (CI) for mortality per 1-point increase in Boey score. Discriminative ability was assessed by receiver operating characteristic (ROC) curve analysis, with the area under the curve (AUC) and its 95% CI estimated via bootstrap resampling (5,000 replicates). The optimal cut-off was determined using the Youden index (sensitivity + specificity - 1), and diagnostic performance (sensitivity, specificity, positive and negative predictive value) was reported at each candidate cut-off. A two-sided $p < 0.05$ was considered statistically significant.

E. Ethical Considerations

This study used anonymized medical-record data and was conducted as part of an internal quality-of-care evaluation at RS Efarina Etaham Pematang Siantar, with minimal risk to subjects.

III. RESULTS

A. Baseline Characteristics

Sixty-seven post-laparotomy PPU patients met the inclusion criteria. Mean age was 57.8 ± 14.3 years (range 21-92), with 30 patients (44.8%) aged ≥ 60 years. Most patients were male (58; 86.6%) versus female (9; 13.4%). Seventeen patients (25.4%) died during hospitalization and 50 (74.6%) survived to discharge (Table 1).

Table 1. Baseline characteristics of study subjects (n=67)

Characteristic	n	%
Age (years), mean $\pm$ SD	57.8 $\pm$ 14.3	–
Age ≥ 60 years	30	44.8
Male sex	58	86.6
In-hospital mortality	17	25.4

B. Boey Score Component and Total Score Distribution

Delayed presentation >24 hours was the most frequent component (63; 94.0%), followed by major comorbidity (41; 61.2%) and preoperative shock (19; 28.4%) (Table 2). Boey score 0 was rare (2; 3.0%); scores 1 and 2 predominated (23 and 26 patients), and 16 patients (23.9%) had the maximum score of 3.

Table 2. Boey score components and total score (n=67)

Component / Score	n	%
Preoperative shock	19	28.4
Major comorbidity	41	61.2
Delay >24 hours	63	94.0
Total score 0	2	3.0
Total score 1	23	34.3
Total score 2	26	38.8
Total score 3	16	23.9

C. Association Between Boey Score and Mortality

Mortality increased sharply with increasing Boey score, from 0% at scores 0-1 to 11.5% at score 2 and 87.5% at score 3 (Table 3).

Table 3. Boey score vs. mortality (n=67)

Score	Died	Survived	Mortality (%)
0	0	2	0.0
1	0	23	0.0
2	3	23	11.5
3	14	2	87.5
Total	17	50	25.4

Pearson's chi-square test showed a highly significant association between Boey score category and mortality ($\chi^2=43.74$, $df=3$, $p<0.001$). Because the expected count for score 0 was <5 , Fisher's exact test was also applied: low (0-1) vs. high (2-3) Boey score was significantly associated with mortality ($p<0.001$), as was score 3 vs. <3 (OR=112.0, $p<0.001$). Logistic regression treating Boey score as a continuous variable showed that each 1-point increase was associated with a 57.4-fold increase in mortality risk (OR=57.42, 95% CI 9.05-364.13, $p<0.001$), confirming a strong dose-response relationship. The very wide confidence interval partly reflects quasi-complete separation arising from the absence of deaths at scores 0-1 in this sample, and the OR should be interpreted as indicating direction and strength of association rather than a precise point estimate.

D. Discriminative Ability: ROC Curve and AUC

The Boey score showed outstanding discriminative ability for mortality, with an AUC of 0.936 (bootstrap 95% CI 0.873-

0.986) (Fig. 1). By convention, an AUC of 0.5 indicates no discrimination, 0.7-0.8 is acceptable, 0.8-0.9 is excellent, and >0.9 is outstanding; the AUC observed here therefore reflects a highly reliable ability of the Boey score to distinguish patients at high risk of death from those at low risk in this population.

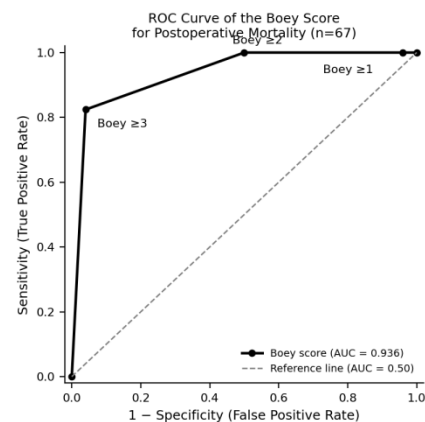


Fig. 1. ROC curve of the Boey score for postoperative mortality (n=67).

Diagnostic performance at each candidate cut-off is shown in Table 4. The Youden-optimal cut-off was Boey score ≥ 3 (Youden index 0.784), with sensitivity 82.4%, specificity 96.0%, positive predictive value 87.5%, and negative predictive value 94.1%. A lower threshold of ≥ 2 achieved perfect sensitivity (100%) with moderate specificity (50.0%), and may be clinically useful as an early screening threshold to avoid missing any patient at risk of death, at the cost of more false-positive high-risk classifications.

Table 4. Diagnostic performance of the Boey score at candidate cut-offs

Cut-off	Sens. (%)	Spec. (%)	Youden
≥ 1	100.0	4.0	0.040
≥ 2	100.0	50.0	0.500
≥ 3	82.4	96.0	0.784

E. Comparison with Previously Reported AUC Values

The AUC observed in this study (0.936) was numerically higher than AUC values reported in prior Boey score validation studies (Table 5): 0.793 in Agarwal et al.[10] (n=180, India) and 0.792 in Kadam[11] (India). Possible explanations for this difference, including sample size, population homogeneity, and a ceiling effect in the delay component, are discussed in Section IV.

Table 5. AUC of the Boey score: this study vs. prior validation studies

Study	n	AUC (95% CI)
This study (2026)	67	0.936 (0.873-0.986)
Agarwal et al.	180	0.793
Kadam	–	0.792

IV. DISCUSSION

The mean age (57.8 $\pm$ 14.3 years) and male predominance (86.6%) observed in this study are consistent with the

epidemiological pattern of PPU reported internationally and in Indonesia, where middle-aged to older men appear to carry a higher risk, plausibly reflecting cumulative exposure to smoking, alcohol use, and NSAID consumption.[3] A similar demographic pattern was reported in Korea[14] and in a retrospective Indonesian study in Bali.[15] This cross-national consistency in baseline risk profile stands in contrast to the wide variation in mortality that accompanies it depending on healthcare-facility context.

Overall mortality in this study (25.4%) fell within the globally reported range (10-20%) but toward its upper limit, and was somewhat higher than the 9.3% 30-day mortality reported in the multicentre GRACE cohort, which included many high-income countries with more rapid access to care.[4] It was, however, considerably lower than the 53.8% mortality reported at RSUD Dr. Soetomo Surabaya, a tertiary referral center, in a similarly sized cohort (n=65) using a comparable Boey-component assessment methodology.[13] This marked difference between two Indonesian studies (25.4% vs. 53.8%) is plausibly explained by differences in referral pattern: RS Efarina Etaham, as a primary-contact private regional hospital, tends to manage patients presenting directly from the surrounding population, whereas a tertiary referral center typically receives more advanced, pre-referred cases.

A striking finding of this study was the very high proportion of delayed presentation >24 hours (94.0%), far exceeding that of the other two components. This is consistent with well-documented healthcare-access challenges in populations served by private regional hospitals, where patients often seek care only after symptoms have worsened, or experience delays in referral from primary care facilities.[13] This proportion was also higher than the approximately 60-70% reported by Agarwal et al.[10] and by Kadam[11] in India, possibly reflecting differences in referral-system density between urban Indian populations and the semi-rural North Sumatran catchment area of this hospital.

Because almost all patients were positive for the delay component, this component contributed limited discriminative variance to the total Boey score in this population — a ceiling effect. Variation in total score was therefore driven predominantly by preoperative shock (28.4%) and major comorbidity (61.2%), consistent with the Surabaya study, which also identified preoperative shock as the most dominant predictive component.[13] This differs from the original Boey cohort in Singapore, where the three components were more evenly distributed.[9] Such differences in component distribution across populations likely contribute to between-study variation in AUC, discussed further below.

The dose-response relationship between Boey score and mortality observed here — from 0% at scores 0-1 to 87.5% at score 3 — closely mirrors the pattern reported in the original Boey cohort and in subsequent validation studies across countries and decades.[9][10] This consistency across populations as different as 1980s Singapore and present-day Indonesia and India supports the construct validity of the Boey score: the cumulative burden of physiological risk factors

(shock, comorbidity, delay) predictably worsens outcome irrespective of geographic or temporal context.

The AUC observed in this study (0.936) exceeded that reported by Agarwal et al. (0.793, n=180)[10] and by Kadam (0.792)[11]. Several factors may explain this difference. First, the smaller sample size in this study (n=67) widens the confidence interval around the AUC estimate and increases the potential for sampling-related overestimation, reflected in the relatively wide bootstrap CI (0.873-0.986) obtained here. Second, the ceiling effect in the delay component meant that score variation in this population was effectively driven by the combination of shock and comorbidity, which happened to be strongly associated with mortality in this specific cohort — a pattern that may not generalize to populations with more even component distribution. Third, a single-center, relatively homogeneous clinical population can yield higher apparent discrimination than a more heterogeneous multicentre population, analogous to the optimism typically seen in internal versus external validation of clinical prediction models. Notwithstanding these absolute differences, the consistent finding across all comparator studies is that the Boey score achieves acceptable-to-outstanding discrimination (AUC >0.7) across diverse populations and eras, supporting its robustness as a bedside prognostic tool, particularly where diagnostic resources are limited.

Based on the Youden index, a cut-off of Boey score ≥ 3 offered the best balance of sensitivity (82.4%) and specificity (96.0%), with a high positive predictive value (87.5%), making it an appropriate threshold for identifying patients at very high mortality risk who require intensive postoperative care and candid prognostic communication with families — consistent with cut-offs identified in the original Boey study and in Agarwal et al.[9][10] A lower threshold of ≥ 2 , with perfect sensitivity (100%) but reduced specificity (50.0%), may be more appropriate as an early screening threshold in the emergency department, before additional clinical information becomes available, prioritizing sensitivity over specificity in line with the tiered risk-stratification approach endorsed by the WSES guidelines[12] and consistent with the emphasis on tiered risk stratification in the Surviving Sepsis Campaign guidelines for intra-abdominal sepsis such as PPU.[16]

Limitations

This study has several limitations. First, its retrospective, single-center design with a relatively small sample (n=67) limits the precision of statistical estimates, particularly for the Boey score 0 stratum (n=2). Second, the ceiling effect in the delay component limited the ability to assess its independent contribution to overall discrimination. Third, as an internal validation study rather than an external validation in a separate population, the reported AUC may be subject to optimism bias and requires confirmation in larger, multicentre studies before broad generalization.

V. CONCLUSION

In this cohort of 67 post-laparotomy PPU patients at a private regional hospital in Indonesia, delayed presentation >24 hours was the most frequent Boey score component (94.0%), followed by major comorbidity (61.2%) and preoperative shock (28.4%), with the total Boey score distribution dominated by scores 1 and 2. The Boey score was strongly and significantly associated with postoperative mortality (chi-square $\chi^2=43.74$, $p<0.001$; Fisher's exact $p<0.001$), with a clear dose-response gradient from 0% mortality at scores 0-1 to 87.5% at score 3. The score showed outstanding discriminative ability (AUC 0.936, 95% CI 0.873-0.986), exceeding values reported in prior international validation studies, with a Youden-optimal cut-off of ≥ 3 (sensitivity 82.4%, specificity 96.0%) and a high-sensitivity screening threshold of ≥ 2 (sensitivity 100%). These findings support routine use of the Boey score as a simple, low-cost bedside risk-stratification tool for PPU patients undergoing laparotomy in private regional hospitals with limited diagnostic resources, while larger multicentre studies are needed to externally validate these findings and to explore combined models incorporating simple laboratory parameters to address the ceiling effect observed in the delay component.

REFERENCES

- [1] K. Søreide, K. Thorsen, E. M. Harrison, J. Bingener, M. H. Møller, F. A. Cassivi, and J. A. Søreide, "Perforated peptic ulcer," *Lancet*, vol. 386, no. 10000, pp. 1288–1298, 2015.
- [2] K. T. Chung and V. G. Shelat, "Perforated peptic ulcer – an update," *World J. Gastrointest. Surg.*, vol. 9, no. 1, pp. 1–12, 2017.
- [3] K. Thorsen, J. A. Søreide, J. T. Kvaløy, T. Glomsaker, and K. Søreide, "Epidemiology of perforated peptic ulcer: age- and gender-adjusted analysis of incidence and mortality," *World J. Gastroenterol.*, vol. 19, no. 3, pp. 347–354, 2013.
- [4] GRACE Study Collaborative, "Global 30-day morbidity and mortality of surgery for perforated peptic ulcer: GRACE study," *Surg. Endosc.*, 2024, doi: 10.1007/s00464-024-10832-4.
- [5] M. H. Møller, M. C. Engebjerg, S. Adamsen, J. Bendix, and R. W. Thomsen, "The Peptic Ulcer Perforation (PULP) score: a predictor of mortality following peptic ulcer perforation. A cohort study," *Acta Anaesthesiol. Scand.*, vol. 56, no. 5, pp. 655–662, 2012.
- [6] M. M. Linder and H. Wacha, "The Mannheim peritonitis index: an instrument for the intraoperative prognosis of peritonitis" [in German], *Chirurg*, vol. 58, no. 2, pp. 84–92, 1987.
- [7] W. A. Knaus, E. A. Draper, D. P. Wagner, and J. E. Zimmerman, "APACHE II: a severity of disease classification system," *Crit. Care Med.*, vol. 13, no. 10, pp. 818–829, 1985.
- [8] M. Daabiss, "American Society of Anaesthesiologists physical status classification," *Indian J. Anaesth.*, vol. 55, no. 2, pp. 111–115, 2011.
- [9] J. Boey, S. K. Y. Choi, A. Poon, and T. T. Alagaratnam, "Risk stratification in perforated duodenal ulcers: a prospective validation of predictive factors," *Ann. Surg.*, vol. 205, no. 1, pp. 22–26, 1987.
- [10] A. Agarwal, S. Jain, L. N. Meena, S. A. Jain, and L. Agarwal, "Validation of Boey's score in predicting morbidity and mortality in peptic perforation peritonitis in northwestern India," *Trop. Gastroenterol.*, vol. 33, no. 4, 2012, doi: 10.7869/tg.300.
- [11] P. R. Kadam, "The utility of the Boey scoring system in predicting postoperative morbidity and mortality in patients with perforative peritonitis," *Int. Surg. J.*, vol. 13, no. 6, pp. 997–1003, 2026.
- [12] A. Tarasconi, F. Coccolini, W. L. Biffl, M. Tomasoni, L. Ansaloni, E. Picetti, et al., "Perforated and bleeding peptic ulcer: WSES guidelines," *World J. Emerg. Surg.*, vol. 15, art. 3, 2020.
- [13] S. H. N. Saputra, V. S. Budipramana, and M. D. Wibowo, "The most dominant predictability factor of Boey score on the level of mortality in perforation gaster Boey score 1 and Boey score 2 patients in RSUD Dr Soetomo Surabaya," *Biosci. Med. J. Biomed. Transl. Res.*, vol. 5, no. 10, pp. 926–930, 2021.
- [14] Y. J. Yang, C. S. Bang, S. P. Shin, T. Y. Park, K. T. Suk, and G. H. Baik, et al., "Clinical characteristics of peptic ulcer perforation in Korea," *World J. Gastroenterol.*, vol. 23, no. 14, pp. 2566–2574, 2017.
- [15] E. Kurniawan and E. Danardono, "Risk factors and outcomes of perforated peptic ulcer patients: a retrospective study," *Bali Med. J.*, vol. 12, no. 1, pp. 45–52, 2023.
- [16] L. Evans, A. Rhodes, W. Alhazzani, M. Antonelli, C. M. Coopersmith, C. French, et al., "Executive summary: Surviving Sepsis Campaign: international guidelines for the management of sepsis and septic shock 2021," *Crit. Care Med.*, vol. 49, no. 11, pp. 1974–1982, 2021.

AUTHORS

First Author – Rinadi Andara, M.Ked.(Surg), Sp.B, Department of Surgery, Efarina Etaham Pematang Siantar Hospital, North Sumatra, Indonesia, dr.rinadi.andara@gmail.com.