Comparative evaluation of P-POSSUM and NELA scores in predicting 30-day mortality following emergency laparotomy: A prospective observational study

Mahendra Lodha[1], Abhijeet Khoth[1], Karthik N[1], Mayank Badkur[1], Satya Prakash Meena[1], Niladri Banerjee[1], Swathi M[2]

[1]Department of Surgery, All India Institute of Medical Sciences, Jodhpur, India
[2]Department of Anesthesiology, Dr. S N Medical College, Jodhpur, India

ABSTRACT

Objective: Emergency laparotomy carries a 10-18% mortality risk, influenced by factors such as age, medical conditions, and sarcopenia. Scoring models like the Portsmouth physiological and operative severity score (P-POSSUM) and the National Emergency Laparotomy Audit (NELA) have been developed to predict outcomes and assist decision-making. Both models are widely used, but their effectiveness in predicting outcomes, particularly in the Indian context, requires further evaluation. This study aimed to compare the P-POSSUM and NELA scores in predicting 30-day mortality for patients undergoing emergency laparotomy.

Material and Methods: This single-institution prospective observational study included 238 adult patients of age ≥18 years undergoing emergency laparotomy for acute abdominal conditions, following ethical approval. P-POSSUM and NELA scores were calculated preoperatively, and their predictive accuracy was evaluated by comparing predicted versus observed mortality using sensitivity, specificity, positive and negative predictive values, and the area under the receiver operating characteristic curve.

Results: The NELA area under the curve was 0.699, while the P-POSSUM area under the curve was 0.687. NELA demonstrated higher sensitivity (73.9%) and specificity (45.6%) than P-POSSUM, which had a sensitivity of 52.2% and specificity of 27.4%. P-POSSUM and NELA scores were significantly higher in patients requiring intensive care unit admission than in those who did not.

Conclusion: Our study found that the NELA score outperforms the P-POSSUM score in predicting 30-day mortality in emergency laparotomy patients, indicating that NELA is a more reliable tool for preoperative risk stratification and clinical decision-making.

Keywords: Emergency laparotomy, P-POSSUM, NELA, mortality

Corresponding Author
Karthik N

E-mail: karthikn0912@gmail.com
ORCID ID: orcid.org/0009-0009-6711-8103

INTRODUCTION

The average incidence of mortality after emergency laparotomy varies from 10% to 18% in different studies (1,2). The mortality-related risks in emergency laparotomies are much higher than any major gastrointestinal surgeries (2). The outcomes of emergency laparotomy are impacted by several factors, including the patient's age, medical comorbidities, general condition, presence of contamination, sarcopenia, etc. We need to focus on pre-operative considerations and associated factors to estimate the survival probability of patients undergoing emergency laparotomy. Prediction scores like the National Emergency Laparotomy Audit (NELA) and the Portsmouth-physiological and operative severity score for enumeration of mortality and morbidity (P-POSSUM) aid clinicians in predicting patient outcomes and supplement decision-making (3,4).

Given the rise in emergency laparotomies, it is crucial to identify reliable risk assessment tools to recognise high-risk patients early and allocate resources appropriately. A comparative analysis in India found that the P-POSSUM score effectively predicted mortality preoperatively in emergency laparotomy cases (5). A study conducted in Sweden demonstrated that P-POSSUM scores are highly accurate in predicting mortality among geriatric patients undergoing laparotomy in emergency settings (6). A study in New Zealand concluded that the NELA score is the most predictive tool for assessing mortality risk among emergency laparotomy patients (7). A UK study found that the P-POSSUM score moderately predicts mortality in elderly patients undergoing emergency abdominal surgery (8). Few studies have shown that both

P-POSSUM and NELA scores tend to overestimate mortality in patients undergoing emergency laparotomies (9,10). While some studies found no significant differences between the two scores in estimating mortality, others found that NELA outperformed P-POSSUM in clinical practice (11-15).

A study highlighted that while P-POSSUM and APACHE-II are often used to predict mortality in emergency laparotomy patients, no scoring system currently provides highly accurate or easily calculable risk predictions (16). Due to the rising number of emergency laparotomies in India, both P-POSSUM and NELA scoring models are widely used. However, their validity in predicting mortality and morbidity in emergency laparotomy patients, particularly in the Indian population, still requires further evaluation. Most of the studies are retrospective, and there is a lack of well-designed prospective observational studies in the Indian population that establish the effectiveness of both scoring models (17). This study compares the NELA and P-POSSUM scoring systems in estimating thirty-day mortality for patients undergoing emergency laparotomies.

MATERIAL and METHODS

Study Design and Patients

This single-centre, prospective observational study was conducted at our tertiary care hospital from July 2022 to January 2024. The Institutional Ethics Committee approved the study All India Institute of Medical Sciences, Jodhpur (IEC/2022/4135). All patients who underwent emergency laparotomy in the department of general surgery were enrolled. Eligibility criteria included adults aged 18 years or older who had undergone emergency laparotomy for any acute abdominal aetiology through a midline incision of 5 cm or longer. Patients undergoing trauma laparotomies were excluded.

Study Procedure and Outcomes

All the patients admitted to our department of surgery underwent comprehensive medical evaluations as a part of standard practice. Through convenience sampling, patients who met the inclusion criteria were selected and given detailed explanations of the study format using a patient information sheet. Written informed consent was taken from the patients willing to participate. The NELA and P-POSSUM mortality risk scores were calculated for each patient in the preoperative room based on the respective scoring algorithms. The primary objective was to compare the effectiveness of both scores in predicting thirty-day mortality. The predictive accuracy of both mortality risk scores was assessed by comparing predicted mortality rates with observed mortality rates using metrics such as specificity, sensitivity, area under the receiver operating characteristic curve, and positive and negative predictive values. Secondary objectives included assessing the length of

postoperative hospital stay and redo surgeries within 30 days. Demographic details, comorbid conditions, and substance use information were collected. All the patients were followed up until discharge.

Sample Size

Using the mean and standard deviation values from Lai et al.'s (15) study -16.3±21.4 for P-POSSUM and 9.8±12.7 for NELA- the sample size was calculated through OpenEpi software. The required sample size was 238, accounting for 80% power, a 95% confidence interval, and a 10% contingency.

Statistical Analysis

Data were analysed using the IBM Statistical Package for Social Sciences version 29.0. Descriptive data were reported for each variable. Data for continuous variables were expressed as mean or median, and compared using the Student's t-test or Mann-Whitney U test, depending on the distribution. The analysis focused on evaluating discrimination and calibration for each risk prediction tool selected. Discrimination of a risk prediction tool refers to its ability or inability to correctly classify patients with or without mortality following an emergency laparotomy, as determined by the area under the curve (AUC)- receiver operating characteristic (ROC) curve. The AUC provides a quantitative assessment of the discrimination of a risk prediction tool, enabling comparison between different tools. The observed thirty-day mortality rate was compared with the predicted thirty-day mortality rate for each risk prediction tool using the chi-square test. A p-value greater than 0.05 indicated that the expected and observed mortality rates were similar, suggesting good calibration of the risk prediction tool.

RESULTS

A total of 238 patients were enrolled during the study period. It was observed that the patients' mean age was 50.2 years. Almost 45% of cases in the study group were aged between 36 and 55. The mean postoperative hospital stay was 9.94 days. Notably, 85.7% of cases in the study group did not need postoperative intensive care unit (ICU). Of the 34 cases requiring postoperative ICU care, 21 patients were admitted for less than three days, while 5 were admitted for more than a week. The patient characteristics are detailed in Table 1.

It was observed that 23 cases died within 30 days of surgery, while 215 cases survived. It was observed that the P-POSSUM score in cases needing ICU admission was 15.15, while it was 11.00 in cases not needing ICU admission. The NELA score was 12.41 in cases requiring ICU admission and 5.18 in cases not requiring ICU. the scores were significantly higher in cases ICU admission.

It was observed that the P-POSSUM score in cases that did not survive was 17.61, while it was 10.97 in cases that survived.

Additionally, the NELA score was 13.61 in cases that did not survive, while it was 5.4 in cases that survived. The difference in P-POSSUM and NELA scores between cases categorized by mortality was statistically significant. A comparison of both scores in predicting the 30-day mortality and postoperative ICU care is shown in Table 2.

In our study on predicting 30-day mortality for patients undergoing laparotomy in emergency settings, we analysed the effectiveness of both scores using ROC curve analysis. It was observed that NELA AUC was 0.699, while P-POSSUM AUC was 0.687. AUC NELA (73.9%) indicates significantly higher sensitivity than P-POSSUM (52.2%). NELA (45.6%) also has higher specificity than P-POSSUM (27.4%). A comparison of both scores in predicting thirty-day mortality based on ROC curve analysis is shown in Table 3 and Figure 1.

DISCUSSION

Our study compared the effectiveness of NELA and P-POSSUM scores in estimating thirty-day mortality for patients undergoing an emergency laparotomy. In our study, mean patient age was 50.2±18.3 years, with nearly half aged 36-55 years. Comparable studies by Rinisha et al. (14) and Hunter Emergency Laparotomy Collaborator Group (18) reported mean ages of 66.0±17 years and 45.48±15.75 years, respectively. Contrary to our findings, Naidoo et al. (19) reported a mean age of 38.2 years for non-trauma emergency laparotomy patients. In Singapore, Lai et al. (15) found a higher mean age of 65.9 years ±14.7, likely due to differences in demographics, ethnicity, and patient severity. With 64% male patients in our study, this aligns with findings from Naidoo et al. (19) and Lai et al. (15), reporting a higher proportion of male patients. In contrast to our findings, Sharma et al. (20) in Birmingham, Barghash et al. (11), and Rinisha et al. (14) in Karnataka reported more female patients, with 78.4% female patients (11,20). This difference may reflect variations in study locations and populations. In our study, ICU-admitted cases had significantly higher scores. P-POSSUM averaged 15.15 (vs. 11.00 for non-ICU cases), and NELA averaged 12.41 (vs. 5.18 for non-ICU cases). This indicates that both scores effectively identified patients requiring ICU admission. Of 34 ICU cases, 21 stayed under 3 days, 8 stayed 4-7 days, and 5 stayed more than a week. Rinisha et al. (14) reported a mean postoperative ICU stay of 1.5±0.3 days for emergency laparotomy patients. In our study, 9.7% (23 cases) died during a 30-day follow-up. Hunter Emergency Laparotomy Collaborator Group (18) reported 10.5% mortality within 30 days for emergency laparotomy patients. NELA predicted 25.4% deaths (52 cases) in high-risk patients, compared to 18.8% (46 cases) with P-POSSUM.

Our study found that P-POSSUM and NELA scores were significantly higher in patients who did not survive, suggesting these scores may effectively differentiate mortality risk in emergency laparotomy cases. Lai et al. (15) found that NELA and P-POSSUM over-predicted mortality, with NELA demonstrating superior performance compared to P-POSSUM. In the Rinisha et al. (14) study, the prediction of mortality using NELA scores was found to correspond better to observed mortality data than the P-POSSUM scores at 30 and 60-day mortality (14). In a retrospective study, Darbyshire et al. (9) concluded that the NELA prediction score was better-calibrated than P-POSSUM, which over-estimated the mortality risk of more than 20% among emergency laparotomy patients. Both the scoring systems

Table 1. Characteristics of patients (n=238)		
Patient characteristics		**Total (%)**
Mean age in years ± SD		50.2±18.3
Sex (%)	Male	152 (63.9%)
	Female	86 (36.1%)
Median operative duration in minutes (IQR)		180 (160-192)
Diagnosis (%)	Perforation peritonitis	105 (44.1%)
	Acute intestinal obstruction	116 (48.7%)
	Miscellaneous	17 (7.1%)
Mean duration of postoperative hospital stay in days ± SD		9.9±5.0
Duration of ICU stay (34)	1-3 days	21 (8.8%)
	4-7 days	8 (3.4%)
	>7 days	5 (2.1%)
Redo surgeries within 30 days (%)	Needed	22 (9.2%)
	Not needed	216 (90.8%)
SD: Standard deviation, IQR: Interquartile range, ICU: Intensive care unit		

| **Table 2.** Comparison of P-POSSUM and NELA scores in predicting 30-day mortality and postoperative ICU care					
Outcome		**P-POSSUM (mean ± SD)**	***p-value**	**NELA (mean ± SD)**	***p-value**
30-day mortality	Death (n=23)	17.6±13.3	0.005	13.6±15.5	<0.01
	Recovered (n=215)	11.0±10.4		5.4±6.3	
Postoperative ICU care	Needed (n=34)	15.2±12.9	0.039	12.4±8.8	<0.01
	Not needed (n=204)	11.0±10.4		5.2±7.4	
*: p-value is calculated using the independent t-test, P-POSSUM: Portsmouth physiological and operative severity score, NELA: National Emergency Laparotomy Audit, SD: Standard deviation, ICU: Intensive care unit					

Table 3. Comparative analysis of P-POSSUM and NELA scores in predicting 30-day mortality based on ROC curve analysis

ROC curve in the prediction of mortality	P-POSSUM	NELA
Ideal cut-off value	13.30	3.77
AUC	0.687	0.699
Sensitivity	52.2%	73.9%
Specificity	27.4%	45.6%

AUC: Area under curve, P-POSSUM: Portsmouth physiological and operative severity score, NELA: National Emergency Laparotomy Audit, ROC: Receiver operating characteristic
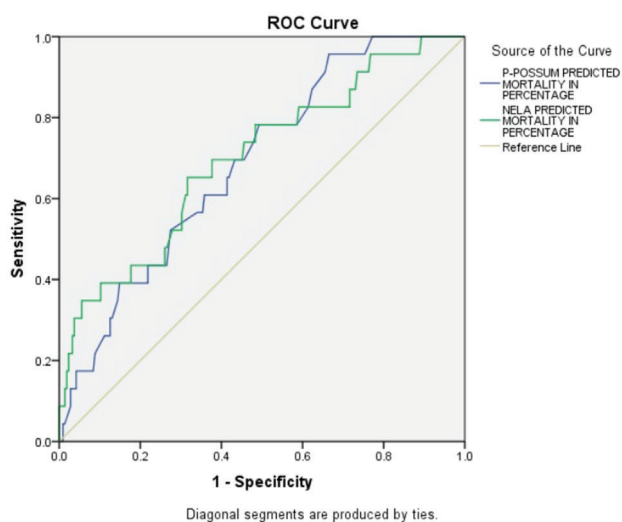
Figure 1. Receiver operating characteristics curve for P-POSSUM and NELA in the mortality prediction in emergency laparotomy cases.

P-POSSUM: Portsmouth physiological and operative severity score, NELA: National Emergency Laparotomy Audit, ROC: Receiver operating characteristic

showed good discrimination with slight variation between operative approaches, over-predicting mortality for laparoscopy (9). Thahir et al. (12) also reported that P-POSSUM over-predicted the risk of mortality, while NELA underestimated the same risk. Alabbasy et al. (21) found 30-day and 90-day mortality rates of 10.3% and 13.1%, respectively, among 670 patients undergoing emergency laparotomy, with AUCs of 0.774 preoperatively for the NELA score and 0.763 for the P-POSSUM score. Their findings, corroborated by Barghash et al. (11), indicated no statistically significant difference in mortality prediction between the two scoring models.

In our study, we found, using the ROC curve, that the NELA model is more specific and sensitive, in the 30-day mortality analysis. The AUC of 0.873 in NELA revealed its better predictive value than the P-POSSUM score (AUC 0.544) in predicting thirty-day mortality in a study conducted by Rinisha et al. (14). In contrast to our findings, Lai et al. (15) found that the area under the receiver operating characteristics curve was similar for the NELA (0.86)

and the P-POSSUM score models (0.84). Linganathan et al. (22) in 2024 found that the NELA scoring method had lesser accuracy in predicting 30-day mortality among emergency laparotomy patients aged above 80 years; they found that the ROC graph analysis of NELA showed that the AUC was 0.78 in the age group of above 80 years and 0.89 in the age group of below 80 years, however, the score was not well-calibrated. This difference was due to different inclusion criteria and age groups. Overall, the findings of this retrospective study noted that the NELA tool performs better, supporting other findings in the literature.

The NELA and P-POSSUM scores have demonstrated discrimination, irrespective of the pre- or post-operative approach, to be preferred as one of the most effective risk-adjustment tools. The NELA method was well-adjusted and calibrated across all risk bands. However, the P-POSSUM method had limited predicted mortality, beyond which over-predicted risk in comparison. However, both scores concerning open surgery were found to be overestimating the mortality among the patients associated with emergency laparotomies.

The study's prospective design strengthened its ability to establish temporal relationships, reduce recall bias, and provide a reliable assessment of associations between exposures and outcomes. NELA and P-POSSUM scores show moderate predictive value (AUC under 0.7).

Study Limitations

This may suggest that while NELA performs better, both scores have limitations. The operative scores rely on subjective assessments, like peritoneal contamination, and do not account for operative duration, the time of presentation to the healthcare facility, and operative approach. The low specificity in both scores suggests they may overestimate risk in the population. This overestimation aligns with the findings of other studies and could be relevant for discussions on their use in Indian settings. The results align with previous research, showing that while both scores are predictive, NELA may be more useful in assessing 30-day mortality risk in emergency laparotomy patients. Further discussion on the clinical implications of these findings and the limitations of the scores in the Indian context could add valuable insight.

CONCLUSION

Our study demonstrated that the NELA score outperforms the P-POSSUM score in estimating thirty-day mortality for emergency laparotomy patients. NELA's superior accuracy suggests it may be a more reliable tool for preoperative risk stratification and clinical decision-making in this high-risk patient population.

Informed Consent: Written informed consent was taken from the patients willing to participate.

REFERENCES

1. Ahmed A, Azim A. Emergency laparotomies: Causes, pathophysiology, and outcomes. Indian J Crit Care Med. 2020;24(Suppl 4):S183-S189.

2. Barazanchi AWH, Xia W, MacFater W, Bhat S, MacFater H, Taneja A, et al. Risk factors for mortality after emergency laparotomy: scoping systematic review. ANZ J Surg. 2020;90:1895-1902.

3. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. Br J Surg. 1998;85:1217-1220.

4. National Emergency Laparotomy Audit (NELA) | The Royal College of Anaesthetists [Internet]. Available from: https://rcoa.ac.uk/research/research-projects/national-emergency-laparotomy-audit-nela.

5. Roy VK, Kumar M, Sinha S, Gopal K, Roy SS. Comparative analysis of P-POSSUM score at admission and pre-operatively in predicting postoperative mortality in patients undergoing emergency laparotomy. South Asian J Case Rep Rev. 2022;9:30-34.

6. Cao Y, Bass GA, Ahl R, Pourlotfi A, Geijer H, Montgomery S, et al. The statistical importance of P-POSSUM scores for predicting mortality after emergency laparotomy in geriatric patients. BMC Med Inform Decis Making. 2020;20:86.

7. Barazanchi A, Bhat S, Palmer-Neels K, Macfater WS, Xia W, Zeng I, et al. Evaluating and improving current risk prediction tools in emergency laparotomy. J Trauma Acute Care Surg. 2020;89:382-387.

8. Sharrock AE, McLachlan J, Chambers R, Bailey IS, Kirkby-Bott J. Emergency abdominal surgery in the elderly: Can we predict mortality? World J Surg. 2016;41:402-409.

9. Darbyshire AR, Kostakis I, Pucher PH, Prytherch D, Mercer SJ. P-POSSUM and the NELA score overpredict mortality for laparoscopic emergency bowel surgery: An analysis of the NELA database. World J Surg. 2022;46:552-560.

10. Echara ML, Singh A, Sharma G. Risk-adjusted analysis of patients undergoing emergency laparotomy using POSSUM and P-POSSUM score: A prospective study. Niger J Surg. 2019;25:45-51.

11. Barghash M, Iskandar A, Fawzy SI, Effiom D, Huck C, Hajibandeh S, et al. Predictive performance of NELA versus P-POSSUM mortality scores: Are we underestimating the risk of mortality following emergency laparotomy? Cureus. 2022;14:e32859.

12. Thahir A, Pinto-Lopes R, Madenlidou S, Daby L, Halahakoon C. Mortality risk scoring in emergency general surgery: Are we using the best tool? J Perioper Pract. 2021;31:153-158.

13. Eliezer DD, Holmes M, Sullivan G, Gani J, Pockney P, Gould T, et al. High-risk emergency laparotomy in Australia: Comparing NELA, P-POSSUM, and ACS-NSQIP calculators. J Surg Res. 2020;246:300-304.

14. Rinisha A, Pujar KA, Kalaivavani V, Dandu S. Comparison of P-POSSUM score and NELA score as prime predictors of motality in emergency laparotomies at a tertiary care centre: A retrospective cohort study. IJARS. 2024;13:1-4.

15. Lai CPT, Goo TT, Ong MW, Prakash PS, Lim WW, Drakeford PA. A comparison of the P-POSSUM and NELA risk score for patients undergoing emergency laparotomy in Singapore. World J Surg. 2021;45:2439-2446.

16. Nag DS. Assessing the risk: Scoring systems for outcome prediction in emergency laparotomies. Biomedicine (Taipei). 2015;5:20.

17. Hajibandeh S, Waterman J, Miller B, Johnson B, Higgi A, et al. Hajibandeh index versus NELA score in predicting mortality following emergency laparotomy: A retrospective cohort study. Int J Surg. 2022;102:106645.

18. Hunter Emergency Laparotomy Collaborator Group; Hunter Emergency Laparotomy Collaborator Group. High-risk emergency laparotomy in Australia: Comparing NELA, P-POSSUM, and ACS-NSQIP calculators. J Surg Res. 2020;246:300-304.

19. Naidoo R, Faurie M, Oosthuizen G, Hardcastle T. Comparative outcome analysis of trauma and non-trauma emergency laparotomy using a modified NELA tool format. S Afr J Surg. 2021;59:12-19.

20. Sharma S, Alderman J, Parekh D, Thickett D, Patel JM. Post-operative complications following emergency laparotomy are common and associated with increased late mortality - a retrospective multi-centre study. Journal of Surgery and Research. 2022;5:24-34.

21. Alabbasy MM, Elsisy AAE, Mahmoud A, Alhanafy SS. Comparison between P-POSSUM and NELA risk score for patients undergoing emergency laparotomy in Egyptian patients. BMC Surg. 2023;23:286.

22. Linganathan S, Saji AP, Hughes I, Mitra K, Hajibandeh S, Hajibandeh S. O25: Accuracy of NELA score in predicting the risk of mortality following emergency laparotomy in patients aged above and below 80: A cohort study. BJS. 2024;111(Suppl 2).