**Name of Journal: *World Journal of Clinical Cases***

**Manuscript NO: 46762**

**Manuscript Type: ORIGINAL ARTICLE**

***Clinical Trials Study***

**Comparative analysis of APACHE-II and P-POSSUM scoring systems in predicting postoperative mortality in patients undergoing emergency laparotomy**

Nag DS *et al*. Comparative analysis of APACHE-II and P-POSSUM

**Deb Sanjay Nag, Ankur Dembla, Pratap Rudra Mahanty, Shashi Kant, Abhishek Chatterjee, Devi Prasad Samaddar, Parul Chugh**

**Deb Sanjay Nag, Pratap Rudra Mahanty, Shashi Kant, Abhishek Chatterjee, Devi Prasad Samaddar,** Department of Anaesthesiology and Critical Care, Tata Main Hospital, Jamshedpur 831001, India

**Ankur Dembla,** Department of Anaesthesiology and Critical Care, Darya Ram Hospital, Sonipat 131001, India

**Parul Chugh,** Department of Biomedical Statistics, Sir Ganga Ram Hospital, New Delhi 110060, India

**ORCID number:** Deb Sanjay Nag ([0000-0003-2200-9324](https://orcid.org/0000-0003-2200-9324)); Ankur Dembla ([0000-0001-7257-0326](http://orcid.org/0000-0001-7257-0326" \t "_blank)); Pratap Rudra Mahanty ([0000-0001-8063-8538](http://orcid.org/0000-0001-8063-8538" \t "_blank)); Shashi Kant ([0000-0002-2718-479X](http://orcid.org/0000-0002-2718-479X" \t "_blank)); Abhishek Chatterjee ([0000-0002-9570-2747](http://orcid.org/0000-0002-9570-2747" \t "_blank)); Devi Prasad Samaddar ([0000-0001-7759-970X](http://orcid.org/0000-0001-7759-970X" \t "_blank)); Parul Chugh ([0000-0003-4898-7652](http://orcid.org/0000-0003-4898-7652" \t "_blank))

**Author contributions:** Nag DS, Dembla A, Mahanty PR, and Chatterjee A planned and conducted of the study, scored the patients, and prepared and edited the manuscript; Kant S and Samaddar DP prepared manuscript, reviewed the literature, and analyzed the findings; Chugh P was involved in the study as a biomedical statistician and guided the plan of the study, determined appropriate sample size, and performed statistical analysis and interpretation.

**Institutional review board statement:** The observational study was approved by the Hospital Ethics Committee, Tata Main Hospital.

**Clinical trial registration statement:** The registration identification No. NCT02471612.

**Conflict-of-interest statement:** The authors declare that they have no conflicts of interest.

**Open-Access:** This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

**Manuscript source:** Invited manuscript

**Corresponding author: Deb Sanjay Nag, MBBS, MD, Associate Specialist, Doctor, Senior Researcher,** Department of Anaesthesiology and Critical Care, Tata Main Hospital, C Road West, Northern Town, Jamshedpur 831001, India. ds.nag@tatasteel.com

**Telephone:** +91-943-1166582

**Fax:** +91-657-2224559

**Received:** February 22, 2019

**Peer-review started:** February 24, 2019

**First decision:** June 21, 2019

**Revised:** June 28, 2019

**Accepted:** July 20, 2019

**Article in press:**

**Published online:**

**Abstract**

***BACKGROUND***

Laparotomy remains one of the commonest emergency surgical procedures. Early prognostic evaluation would aid in selecting the high-risk patients for an aggressive treatment. Awareness about risks could potentially contribute to the quality of perioperative care and optimum utilization of resources. Portsmouth modification of Physiological and operative severity for the enumeration of mortality and morbidity (P-POSSUM) and the acute physiology and chronic health evaluation II (APACHE-II) have been the most widely used scoring systems for emergency laparotomies. It is always better to have a single scoring system to predict outcomes and audit healthcare organizations.

***AIM***

To compare the ability of APACHE-II and P-POSSUM to predict postoperative morbidity and mortality in patients undergoing emergency laparotomy.

***METHODS***

All patients undergoing emergency laparotomy at the Tata Main Hospital, Jamshedpur between December 2013 and November 2014 were included in the study. In this observational study, P-POSSUM and APACHE-II scoring were done, and the outcome analysis evaluated with mortality being the primary outcome.

***RESULTS***

For P-POSSUM, at a cut off value of 63 to predict mortality using receiver operating characteristics curve analysis, the area under the curve was 0.989; and for APACHE-II, at the cut off value of 24, the area under the curve was 0.965.

***CONCLUSION***

Because the ability of APACHE-II to predict mortality was similar to P-POSSUM and APACHE-II does not need scoring for intra-operative findings and histopathology reports, APACHE-II can be used pre-operatively to assess the risk in patients undergoing emergency laparotomy. However, for audit purposes, either of the two scoring systems can be used.

**Key words:** Laparotomy; Emergencies; Acute physiology and chronic health evaluation II; Morbidity; Mortality

**© The Author(s) 2019.** Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** Portsmouth modification of Physiological and operative severity for the enumeration of mortality and morbidity (P-POSSUM) and the acute physiology and chronic health evaluation II (APACHE-II) have been the most widely used scoring systems for emergency laparotomies. To date, no study with statistically significant sample size has compared them in predicting mortality in emergency laparotomies. P-POSSUM cannot be done for patients who are managed conservatively and can be scored only when histopathology reports are available. In this study, both P-POSSUM and APACHE-II were found to be equally accurate. Therefore, APACHE-II scoring system can be used as effectively as P-POSSUM with the added advantage that it can be used in the acute stratification of the patients into risk groups even before surgery.

Nag DS, Dembla A, Mahanty PR, Kant S, Chatterjee A, Samaddar DP, Chugh P. Comparative analysis of APACHE-II and P-POSSUM scoring systems in predicting postoperative mortality in patients undergoing emergency laparotomy. *World J Clin Cases* 2019; In press

**INTRODUCTION**

Laparotomy remains one of the commonest emergency surgical procedures. Even after advances in surgical skills, antimicrobial agents and postoperative care, the mortality has remained high (14.9%-19.4%)[[1](#_ENREF_1" \o "Siegel, 2012 #882),2]. Only over the last few years, various perioperative quality improvement initiatives involving early interventions, intensive postoperative care, and consultant led approaches have ensured a decrease in the average mortality rate to 11.1% in some studies[3].

Early prognostic evaluation would aid in selecting the high-risk patients for an aggressive treatment[3]. Awareness about risks could potentially contribute to the quality of perioperative care and optimum utilization of resources[4]. Regular audit and continuous improvement of clinical practice is essential to providing quality medical care[5] . The doctor is legally bound to discuss the prognosis and the possible outcomes of the available treatment modalities[6] . Estimating the risk preoperatively will help predict which patients would need aggressive treatment, which patients would need damage control surgery *versus* definitive procedure, and who would benefit from postoperative intensive care and organ support[7] .

An ideal scoring system should accurately predict outcomes, help determine who deserves more aggressive care, guide in deciding the extensiveness of surgery, and can be used broadly across emergency laparotomies for various disease pathologies[8] . The scoring system should also be capable of analyzing risk-adjusted morbidity and mortality amongst various healthcare providers[9] .

Portsmouth modification of Physiological and operative severity for the enumeration of mortality and morbidity (P-POSSUM) and the acute physiology and chronic health evaluation II (APACHE-II) have been the most widely used scoring systems for emergency laparotomies. While P-POSSUM remains the tool of choice in the United Kingdom[6] , disparities have been observed between APACHE II and P-POSSUM in their discriminatory ability to predict mortality[10].

It has been suggested that preoperative assessment of individual risk would help the treating team and the patient make shared decisions[10]. Although P-POSSUM is the commonest scoring system used for audit purposes in the United Kingdom[10] for National Emergency Laparotomy Audit, Enhanced Peri-Operative Care for High-risk patients or Emergency Laparotomy Pathway Quality improvement Care (ELPQuiC), it needs 18 data points as compared to 12 data points for APACHE-II. In addition, it needs intraoperative details such as blood loss, peritoneal contamination, and histopathology reports to suggest malignancy.

It is always better to have a single scoring system to predict outcomes and audit of healthcare organizations. Therefore, it has been suggested that studies should “update the performance (primarily the calibration) of APACHE-II and P-POSSUM” and compare its ability to predict postoperative morbidity and mortality[6,10].

**MATERIALS AND METHODS**

After approval from the institutional ethics committee, this single center prospective observational study was conducted from December 2013 to November 2014. All patients undergoing emergency laparotomy at the Tata Main Hospital, Jamshedpur, India, during this period were included in this study. All patients below 18 years, those with acute trauma, undergoing re-exploratory laparotomy, or any laparotomy for vascular surgery were excluded from the study.

All patients were scored with APACHE-II on being posted for emergency surgery. Twelve components of the Physiologic Scores of and one component of the Operative Score for P-POSSUM were scored at the time of being posted for surgery. Four of the six components of the Operative Score for P-POSSUM were done intraoperatively (category, number of procedures, blood loss, and peritoneal soiling), while one was done postoperatively on availability of histopathology reports (malignancy).

The patients were followed up for at least 30 d after discharge or death (during admission or within 30 d after discharge) by telephonic interview. While postoperative mortality was the primary outcome that was analyzed, the following secondary outcomes were also compared: (1) Length of stay (LOS); (2) need for postoperative ventilator support (any time during the postoperative period, either immediate based on the assessment of the anesthesiologist or later due to respiratory failure); (3) need for postoperative ionotropic support (inotropic support would be initiated if the patient remained hypotensive despite fluid resuscitation to maintain a mean arterial pressure ≥ 65 mmHg); (4) acute kidney injury (AKI) (diagnosed based on the Kidney Disease: Improving Global Outcomes Acute Kidney Injury Work Group (2012) guidelines[11]); (5) patients needing re-exploration; and (6) cardiac morbidity (acute myocardial infarction or arrhythmias needing treatment).

***Statistical analysis***

Receiver operating characteristics curve (ROC) was used as a statistical method to measure the diagnostic accuracy. Area under the curve (AUC) was used to measure the “size” of the prediction, and it consisted of graphically plotting “sensitivity" and the “1–specificity" relationship[12]. AUC can range from 0.5 to 1.0, and a result of 1.0 indicates a perfect discriminatory ability. An AUC value > 0.8 is considered good, a range between 0.60-0.80 is considered as moderate, and an AUC value < 0.60 is regarded as poor. The ROC curve was used to display the optimal cut-off point when sensitivity and specificity reached an optimum for both values, by which the point on the ROC curved line was closest to the upper left corner on the curve. Statistical analysis was performed using the SPSS program for Windows, version 17.0 (Chicago, IL, United States). Continuous variables are presented as mean ± standard deviation (SD) or median (min-max), and categorical variables are presented as absolute numbers and percentage. Data were checked for normality before statistical analysis. Normally distributed continuous variables were compared using the unpaired t-test, whereas the Mann-Whitney U-test was used for those variables that were not normally distributed. Categorical variables were analyzed using either the chi square test or Fisher’s exact test.

In previous studies of perforative peritonitis[13], it was found that the sensitivity of APACHE-II was 87.5% at cut off value 16–20. For the sample size calculation, using a two tailed alpha value (0.05) and a beta value (0.2), 150 patients would have been sufficient to detect a significant difference of 10% between APACHE-II and P-POSSUM scoring systems in predicting postoperative mortality in patients undergoing emergency laparotomy. Thus, our sample size of 157 appears to be adequate to assess if there is any difference between the two scoring systems to predict mortality.

**RESULTS**

A total of 159 patients met the inclusion criteria. Two patients sought referral to a higher center and were lost on follow up and were excluded from the study.

Of the total 157 studied patients, 89 had perforative peritonitis, 57 had intestinal obstruction, and 11 were operated because of other reasons that included pancreatitis (4), cholecystitis (2), ruptured liver abscess (1), liver hematoma (1), rectal prolapsed (1), empyema gall bladder (1), and spontaneous hemoperitoneum because of thrombocytopenia (1).

The age of the patients ranged from 18 to 82 years. Of the 157 analyzed patients, 99 (63.1%) were male and 58 (36.9%) were females. Twenty-three (14.6%) of the total patients analyzed died, and 134 (85.4%) survived. The mean ± SD of LOS was 10.18 ± 8.24 and ranged from 1 to 70 d. Sixty-three patients (40.1%) required postoperative ventilatory support, 48 (30.6%) required perioperative ionotropic support, and 32 (20.4%) developed AKI in the postoperative period. Four out of the 157 analyzed patients required re-exploration. A total of eight patients developed postoperative cardiac morbidity.

The median age [interquartile range (IQR)] amongst the survivors was 46 (30-60) years and 60 (44-69) years for those who died in the postoperative period. The statistically significant *P* value (0.029) indicated that increasing age is associated with a higher risk of mortality. While 43.5% of the patients who died were males, 56.5% of the patients who died were females, indicating a statistically significant (*P =* 0.035) increased risk of mortality amongst female patients.

While the median APACHE-II score amongst the patients who died in the postoperative period was 31 (min-max 25-35), the median P-POSSUM Physiologic Score and Operative Score amongst them was 52 and 22, respectively (min-max 46-58 and 20-24, respectively). *P* < 0.001 signifies that higher scores are associated with statistically significant increased mortality.

For APACHE-II, the cut off value was found to be 24 to predict Mortality ROC analysis. In our studied patients, APACHE-II score of < 24 was associated with a significantly lower mortality of 17.4% as compared to an APACHE-II score of ≥ 24, which was associated with a mortality of 82.6% (*P* < 0.001) (Table 1). Using ROC, at cut off value 24, the AUC [95% confidence interval (CI)] was 0.965 (0.928–1.000). Sensitivity, specificity, positive predictive value, and negative predictive value of APACHE-II was found to be 82.6%, 98.5%, 90.5%, and 97.1%, respectively.

In comparison, for P-POSSUM the cut off value found to be 63 to predict Mortality using ROC analysis. P-POSSUM score of <63 was associated with a significantly lower mortality of 8.7% as compared to a score of ≥ 63 which was associated with a mortality of 91.3% (*P* < 0.001) (Table 2). Using ROC, at cut off value 63, AUC (95%CI) was 0.989 (0.974–1.000). Sensitivity, specificity, positive predictive value, and negative predictive value of P-POSSUM was found to be 91.3%, 99.3%, 95.5%, and 98.5% respectively.

Using Pearson's Linear Correlation Coefficient, APACHE-II showed an overall predictive value of 95.5% with an odds ratio (OR) of 1.315, 95%CI of 1.193-1.448, and a *P* < 0.001. Similarly, P-POSSUM showed an overall predictive value of 98.1% with an OR of 1.364, 95%CI of 1.193-1.559, and a *P* < 0.001. Box-plots in R (Pearson Correlation Coefficient) using APACHE-II and P-POSSUM are depicted in Figures 1 and 2 respectively.

Multivariate logistic regression model has been used to identify independent risk factors (APACHE-II and P-POSSUM) for mortality. A ROC, the graphic display between the “sensitivity" and the “1–specificity" relationship to measure diagnostic accuracy of the true positives *versus* the false positives for APACHE-II and P-POSSUM, is depicted in Figure 3. AUC was 0.965 (using a cut-off value 0f 24) for APACHE-II and 0.989 (using a cut-off value 63) for P-POSSUM. AUC can range from 0.5 to 1.0, and a result of 1.0 indicates a perfect discriminatory ability.

Although both the scores were significantly good in predicting postoperative mortality in patients undergoing emergency laparotomy, the AUC of P-POSSUM (0.989) appeared better than APACHE-II (0.965). However, on comparing the sensitivity and specificity of APACHE-II and P-POSSUM (Table 3), there appears to be no statistically significant difference between their ability to predict postoperative mortality. Except for APACHE-II's inability to predict re-exploration, both were able to predict all the secondary outcomes in a statistically significant manner (*P* < 0.001) (Table 4).

**DISCUSSION**

Emergency laparotomy “describes an exploratory procedure for which the clinical presentation, underlying pathology, anatomical site of surgery, and perioperative management vary considerably”[1]. The mere fact that over 400 different surgical procedures have been described as a part of emergency laparotomy reflect the diversity in pathology[1]. Often there is little time to optimize these patients, resulting in significant adverse outcomes. The unadjusted 30-d postoperative mortality rate was 14.6% at our hospital. A study published in 2011 from a 650-bed general hospital (Royal United Hospitals, Bath) serving a population of over half a million reported a 30-d mortality of 16.9% amongst 124 patients undergoing emergency laparotomy[2]. Like their study, we also excluded emergency vascular surgery, re-exploration, and simple appendectomy[2]. Similarly, the Emergency Laparotomy Network[1] covering 35 NHS hospitals reported a 30-d mortality of 14.9% amongst 1853 patients who underwent emergency laparotomy. Similar incidence of mortality after emergency laparotomy of 20.2%[14] and 17%[15] were reported in 2017. Without adjusting for age, patient comorbidity, surgical presentation, and complexity of the involved pathology, we cannot be certain whether our 30-d postoperative mortality (14.6%) represents equivalent or better quality of care in comparison to that provided in European countries (14.9%-20.2%)[1,2]. However there is increased understanding that standardization of care and quality improvement bundles can improve morbidity and mortality after emergency surgery[7]. The male preponderance in our study group (63.1%) and statistically significant increased mortality amongst females (56.5% as compared to 43.5% in males) was in stark contrast to the UK Emergency Laparotomy Network observations[1]. However, there is some evidence supporting our observation. Similar studies in India have shown a male preponderance for patients undergoing emergency laparotomy (69.5%)[16]. Certain scoring systems, like the Mannheim Peritonitis Index, assign a higher risk for the female patients[17], a risk validated by our study also.

While no mortality was observed in any of our patients who were less than 20 years of age, it increased from 11.11% in the 21-40 year age group to 13.33% in the 41-60 year age group, 23.68% in the 61-80 year age group, and 33.33% amongst those above 80 years of age. Amongst the patients analyzed by Emergency Laparotomy Network, the risk of mortality increased by approximately 4% for each additional 10 years of age[1]. Increasing age has been identified as an independent risk factor, and increase in mortality with age has been observed in most studies, thus validating the inclusion of age as risk factor[2,18,19].

In our study, the LOS (± SD) was 10.18 (± 8.24) d. This was similar to the observations by the Emergency Laparotomy Network in whom the median [IQR (range)] postoperative length of stay for all patients was 11 d [6–21 (0–216)][1]. Although 30-d mortality after implementation of the ELPQuiC bundle indicated a reduction in the risk of death (14% to 10.5%), it had no bearing on the LOS, which remained at its median value of 11 d both before and after ELPQuiC[7]. A number of factors, including the survival of patients who would not previously have survived surgery and the availability of suitable discharge facilities, may explain the lack of reduction of LOS even with improved quality of care. Similar LOS of a median [IQR (range)] of 13 [8–24 (1–176)] d following emergency laparotomy has been reported by other studies as well[20]. While higher scores of APACHE-II a P-POSSUM do indicate some correlation with the LOS, the degree of correlation expressed by the Spearman Rank Correlation Coefficient is relatively small, 0.322 for APACHE-II and 0.374 for P-POSSUM.

In our studied patients, APACHE-II score of < 24 was associated with a significantly lower mortality of 17.4%, as compared to a score of ≥ 24 which was associated with a mortality of 82.6%. At cut off value 24, the AUC (95%CI) was 0.965 (0.928-1.000). While all studies have so far shown the ability of APACHE-II scores to predict mortality and similar AUC has been reported in other studies as well for patients undergoing emergency laparotomy either for varied causes[21] or for perforative peritonitis[22], our study has demonstrated the strongest correlation to date with AUC of 0.965 (as compared to 0.74-0.86 in other studies)[21,22].

In our study, P-POSSUM at cut off value of 63 to predict mortality using ROC analysis, a score of < 63 was associated with a significantly lower mortality of 8.7% as compared to a P-POSSUM score of ≥ 63 which was associated with a mortality of 91.3% (*P* < 0.001). Using ROC, at cut off value 63, AUC (95%CI) was 0.989 (0.974–1.000). While our observations are similar to other studies[9,16,23,24], which demonstrates the ability of the P-POSSUM to predict mortality, our AUC of 0.989 at the cut off value of 63 shows a fairly high degree of accuracy of P-POSSUM. Studies have shown that P-POSSUM is a poor predictor in trauma[9], possibly resulting in our study showing a higher predictive ability of mortality, as we had excluded such cases.

While some studies have tried comparing APACHE-II and P-POSSUM across all surgeries[25], others have used it for specific pathologies[26,27]. To date, no study with statistically significant sample size has compared APACHE-II and P-POSSUM in its ability to predict mortality in patients undergoing emergency laparotomy. Our study can potentially fill in the present void in published literature comparing APACHE-II and P-POSSUM in predicting mortality in patients undergoing emergency laparotomy.

ELPQuiC[7] has used P-POSSUM as a scoring system to assess the impact of introduction of quality improvement bundles, but our study shows that either of the scoring systems (APACHE-II or P-POSSUM) can be used as a tool for surgical audit and the impact of quality improvement initiatives on hospital mortality.

In the present study, both scoring systems were found to be accurate in predicting the mortality of patients, with patients having higher scores having a higher mortality. APACHE-II scores correlate well with mortality and are effective in the prediction of outcome. It considers the acute physiology of the patient and can be completed before surgery. Therefore, it is very useful in the acute stratification of the patients into risk groups and in predicting which patients can be considered for more extensive procedures. However, the APACHE-II score does not consider the etiology of peritonitis or the nature of peritoneal contamination, which has an important bearing on the outcome. In comparison, the P-POSSUM system appears to be of value as the physiologic status is assessed just before the operation or more accurately after full resuscitation and also takes the operative findings into consideration.

However, the P-POSSUM model also has its limitations. First of all, it does not include the patients who are managed conservatively and those who have refused or been denied surgery due to the significant associated risk of mortality. Secondly, while recording the operative variables such as estimated blood loss or peritoneal contamination the surgeon’s eye may be biased. And finally, the scores are not complete until the histopathology reports are available and may significantly delay the scoring and assessing of the risk. Possibly, that is the reason why in our study APACHE-II, being a physiologic score, was a poor indicator of the need for a re-exploration surgery (Spearman Rank Correlation Coefficient of 0.112) (*P* = 0.112). P-POSSUM is possibly a better predictor of the need for re-exploration (Spearman Rank Correlation Coefficient of 0.178) as it includes the intra-operative finding with a *P =* 0.026. This indicated that although P-POSSUM has some correlation with possible need re-exploration as compared to APACHE-II (which had no correlation), the correlation was quite low.

Higher APACHE-II and P-POSSUM correlated well with our secondary outcomes like the postoperative need for inotropic support or ventilatory support or AKI. Such patients who need postoperative organ support are best managed in a critical care setup. Ability of APACHE-II to predict these sicker patients (without relying on the intraoperative or histology findings as for P-POSSUM) could allow us to plan better, optimize and utilize such scarce resources.

Because the ability of APACHE-II to predict mortality is similar to P-POSSUM, and the fact that APACHE-II does not need scoring for intra-operative findings and histopathology reports, APACHE-II can be used pre-operatively to assess the risk in patients undergoing emergency laparotomy. However, for audit purposes, either of the two scoring systems can be used.

**ARTICLE HIGHLIGHTS**

***Research background***

Various scoring systems have been used historically to predict outcomes in patients who are at increased risk of morbidity and mortality during their hospital stay. Emergency laparotomy, despite being one of the commonest surgical procedures, continued to have reasonably high postoperative mortality. Doctors are legally bound to discuss with their patients and relatives the potential risk of complications and adverse outcomes. A robust scoring system enables us to quantify the risk and serves as a tool to measure risk-based outcomes and enable audit of clinical results and impact of improvement initiatives.

***Research motivation***

Portsmouth modification of Physiological and operative severity for the enumeration of mortality and morbidity (P-POSSUM) and the acute physiology and chronic health evaluation II (APACHE-II) have been the most widely used scoring systems for emergency laparotomies. P-POSSUM remains the tool of choice in the United Kingdom. However, it is subject to observational bias while quantifying intraoperative blood loss and peritoneal contamination. It is always better that we have a single scoring system to predict outcomes and audit healthcare organizations. Besides, delay in histopathology reports would delay the P-POSSUM score of the patient, and patients managed conservatively or refused surgery could not be scored. In these circumstances, the APACHE-II score had the advantage of being available in the pre-operative period itself. However, to date no study with statistically significant sample size has compared P-POSSUM and APACHE-II in their ability to predict mortality in emergency laparotomies. This study aims to bridge this gap and assess if APACHE-II can be used as a single scoring system to predict outcomes and for audit of outcomes across healthcare organizations.

***Research objectives***

The study was conducted to compare the predictability of APACHE-II and P-POSSUM scoring systems on postoperative mortality and to see any correlation between these scoring systems and length of stay, requirement of postoperative ventilatory support, inotropic support, development of acute kidney injury (AKI), cardiac morbidity, and need for re-exploration. While the study showed that both APACHE-II and P-POSSUM can equally predict mortality, it also demonstrated comparability in predicting increased length of stay and need for postoperative ventilatory support, higher incidence of AKI, and increased risk of cardiac morbidity. However, P-POSSUM was a better predictor of the need for re-exploration as compared to APACHE-II. The study was successful in demonstrating that both APACHE-II and P-POSSUM can be interchangeably used not only for postoperative mortality but also for effectively predicting morbidity. With the advantage that the APACHE-II scoring can be done preoperatively, the study justifies the fact that APACHE-II can be the single scoring system to predict outcomes and audit healthcare organizations for emergency laparotomies.

***Research methods***

All patients undergoing emergency laparotomy at Tata Main Hospital (Jamshedpur, India) form December 2013 to November 2014 were included in the study. All patients were scored with APACHE-II and P-POSSUM scoring systems. Receiver operating characteristics curve (ROC) was used as a statistical method to measure the diagnostic accuracy. Area under the curve (AUC) was used to measure the “size” of the prediction, and it consisted of graphically plotting “sensitivity” and the “1–specificity” relationship. The ROC curve was used to display the optimal cut-off point when sensitivity and specificity reached an optimum for both values, by which the point on the ROC curved line was closest to the upper left corner on the curve.

***Research results***

Out of a total of 159 patients who met the inclusion criteria, only 157 could be included in the study. For APACHE-II, the cut off value was found to be 24 for predicting mortality by ROC analysis. In comparison, for P-POSSUM, the cut off value was found to be 63 to predict mortality using ROC analysis. Multivariate logistic regression model was used to identify independent risk factors for mortality. A ROC, the graphic display between the “sensitivity"and the “1–specificity" relationship to measure diagnostic accuracy of the true positives *versus* the false positives for APACHE-II and P-POSSUM, depicted that AUC was 0.965 (using a cut-off value of 24) for APACHE-II and 0.989 (using a cut-off value 63) for P-POSSUM. Both the scores were significantly good in predicting postoperative mortality in patients undergoing emergency laparotomy and on comparing the sensitivity and specificity of APACHE-II and P-POSSUM, there appears to be no statistically significant difference between their ability to predict postoperative mortality. Except for APACHE-II's inability to predict re-exploration, both can predict all the secondary outcomes in a statistically significant manner.

***Research conclusions***

This is possibly the first adequately powered study with alpha value (0.05) and a beta value (0.2) and statistically significant sample size that has compared P-POSSUM and APACHE-II in predicting mortality in emergency laparotomies. P-POSSUM above 63 and APACHE-II above 24 not only indicates higher risk, it also increases the risk of postoperative morbidity. However, APACHE-II, being a physiologic score, was a poor indicator of the need for a re-exploration after laparotomy. P-POSSUM is a significantly better predictor of the possibility of re-exploration. While P-POSSUM continues to be the most commonly used scoring system for audit purposes, risk-based outcome comparisons across hospitals and impact of quality improvement initiatives using APACHE-II would ensure that a single scoring system can be used not only for individual patient’s risk assessment and prognostication but also used interchangeably with P-POSSUM for audit purposes as well.

***Research perspectives***

This study demonstrates that compared to the more widely used P-POSSUM, which needs 18 data points, APACHE-II needs only 12 data points, is easily available for risk assessment in the preoperative period, and does not need subjective assessments (intraoperative blood loss or peritoneal contamination) or wait for histopathology reports. While this study was an adequately powered single center study, future research should focus on multi-center trials to strengthen the findings of our study.

**REFERENCES**

1 **Saunders DI**, Murray D, Pichel AC, Varley S, Peden CJ; UK Emergency Laparotomy Network. Variations in mortality after emergency laparotomy: the first report of the UK Emergency Laparotomy Network. *Br J Anaesth* 2012; **109**: 368-375 [PMID: 22728205 DOI: 10.1093/bja/aes165]

2 **Clarke A**, Murdoch H, Thomas MJ, Cook TM, Peden CJ. Mortality and postoperative care after emergency laparotomy. *Eur J Anaesthesiol* 2011; **28**: 16-19 [PMID: 20829702 DOI: 10.1097/EJA.0b013e32833f5389]

3 **Hussain A**, Mahmood F, Teng C, Jafferbhoy S, Luke D, Tsiamis A. Patient outcome of emergency laparotomy improved with increasing "number of surgeons on-call" in a university hospital: Audit loop. *Ann Med Surg* (Lond) 2017; **23**: 21-24 [PMID: 29021897 DOI: 10.1016/j.amsu.2017.09.013]

4 **Neary WD**, Heather BP, Earnshaw JJ. The Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM). *Br J Surg* 2003; **90**: 157-165 [PMID: 12555290 DOI: 10.1002/bjs.4041]

5 **Mercer S**, Guha A, Ramesh V. The P-POSSUM scoring systems for predicting the mortality of neurosurgical patients undergoing craniotomy: Further validation of usefulness and application across healthcare systems. *Indian J Anaesth* 2013; **57**: 587-591 [PMID: 24403619 DOI: 10.4103/0019-5049.123332]

6 **Nag DS**. Assessing the risk: Scoring systems for outcome prediction in emergency laparotomies. *Biomedicine* (Taipei) 2015; **5**: 20 [PMID: 26615537 DOI: 10.7603/s40681-015-0020-y]

7 **Huddart S**, Peden CJ, Swart M, McCormick B, Dickinson M, Mohammed MA, Quiney N; ELPQuiC Collaborator Group; ELPQuiC Collaborator Group. Use of a pathway quality improvement care bundle to reduce mortality after emergency laparotomy. *Br J Surg* 2015; **102**: 57-66 [PMID: 25384994 DOI: 10.1002/bjs.9658]

8 **Rix TE**, Bates T. Pre-operative risk scores for the prediction of outcome in elderly people who require emergency surgery. *World J Emerg Surg* 2007; **2**: 16 [PMID: 17550623 DOI: 10.1186/1749-7922-2-16]

9 **Chieng TH**, Roslan AC, Chuah JA. Risk-adjusted analysis of patients undergoing laparotomy using POSSUM and P-POSSUM score in Queen Elizabeth Hospital, Sabah. *Med J Malaysia* 2010; **65**: 286-290 [PMID: 21901947]

10 **Oliver CM**, Walker E, Giannaris S, Grocott MP, Moonesinghe SR. Risk assessment tools validated for patients undergoing emergency laparotomy: a systematic review. *Br J Anaesth* 2015; **115**: 849-860 [PMID: 26537629 DOI: 10.1093/bja/aev350]

11 **Kellum JA**, Lameire N; KDIGO AKI Guideline Work Group. Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). *Crit Care* 2013; **17**: 204 [PMID: 23394211 DOI: 10.1186/cc11454]

12 **Søreide K**, Kørner H, Søreide JA. Diagnostic accuracy and receiver-operating characteristics curve analysis in surgical research and decision making. *Ann Surg* 2011; **253**: 27-34 [PMID: 21294285 DOI: 10.1097/sla.0b013e318204a892]

13 **Kulkarni SV**, Naik AS, Subramanian N Jr. APACHE-II scoring system in perforative peritonitis. *Am J Surg* 2007; **194**: 549-552 [PMID: 17826077 DOI: 10.1016/j.amjsurg.2007.01.031]

14 **Tengberg LT**, Cihoric M, Foss NB, Bay-Nielsen M, Gögenur I, Henriksen R, Jensen TK, Tolstrup MB, Nielsen LB. Complications after emergency laparotomy beyond the immediate postoperative period - a retrospective, observational cohort study of 1139 patients. *Anaesthesia* 2017; **72**: 309-316 [PMID: 27809332 DOI: 10.1111/anae.13721]

15 **Tolstrup MB**, Watt SK, Gögenur I. Morbidity and mortality rates after emergency abdominal surgery: an analysis of 4346 patients scheduled for emergency laparotomy or laparoscopy. *Langenbecks Arch Surg* 2017; **402**: 615-623 [PMID: 27502400 DOI: 10.1007/s00423-016-1493-1]

16 **Kumar P**, Rodrigues GS. Comparison of POSSUM and P-POSSUM for risk-adjusted audit of patients undergoing emergency laparotomy. *Ulus Travma Acil Cerrahi Derg* 2009; **15**: 19-22 [PMID: 19130334]

17 **Malik AA**, Wani KA, Dar LA, Wani MA, Wani RA, Parray FQ. Mannheim Peritonitis Index and APACHE II--prediction of outcome in patients with peritonitis. *Ulus Travma Acil Cerrahi Derg* 2010; **16**: 27-32 [PMID: 20209392]

18 **Al-Temimi MH**, Griffee M, Enniss TM, Preston R, Vargo D, Overton S, Kimball E, Barton R, Nirula R. When is death inevitable after emergency laparotomy? Analysis of the American College of Surgeons National Surgical Quality Improvement Program database. *J Am Coll Surg* 2012; **215**: 503-511 [PMID: 22789546 DOI: 10.1016/j.jamcollsurg.2012.06.004]

19 **Vester-Andersen M**, Lundstrøm LH, Møller MH, Waldau T, Rosenberg J, Møller AM; Danish Anaesthesia Database. Mortality and postoperative care pathways after emergency gastrointestinal surgery in 2904 patients: a population-based cohort study. *Br J Anaesth* 2014; **112**: 860-870 [PMID: 24520008 DOI: 10.1093/bja/aet487]

20 **Shapter SL**, Paul MJ, White SM. Incidence and estimated annual cost of emergency laparotomy in England: is there a major funding shortfall?. *Anaesthesia* 2012; **67**: 474-478 [PMID: 22493955 DOI: 10.1111/j.1365-2044.2011.07046.x]

21 **van Ruler O**, Kiewiet JJ, Boer KR, Lamme B, Gouma DJ, Boermeester MA, Reitsma JB. Failure of available scoring systems to predict ongoing infection in patients with abdominal sepsis after their initial emergency laparotomy. *BMC Surg* 2011; **11**: 38 [PMID: 22196238 DOI: 10.1186/1471-2482-11-38]

22 **Delibegovic S**, Markovic D, Hodzic S. APACHE II scoring system is superior in the prediction of the outcome in critically ill patients with perforative peritonitis. *Med Arh* 2011; **65**: 82-85 [PMID: 21585179]

23 **Mohil RS**, Bhatnagar D, Bahadur L, Rajneesh, Dev DK, Magan M. POSSUM and P-POSSUM for risk-adjusted audit of patients undergoing emergency laparotomy. *Br J Surg* 2004; **91**: 500-503 [PMID: 15048756 DOI: 10.1002/bjs.4465]

24 **Brooks MJ**, Sutton R, Sarin S. Comparison of Surgical Risk Score, POSSUM and p-POSSUM in higher-risk surgical patients. *Br J Surg* 2005; **92**: 1288-1292 [PMID: 15981213 DOI: 10.1002/bjs.5058]

25 **Yurtlu DA**, Aksun M, Ayvat P, Karahan N, Koroglu L, Aran GÖ. Comparison of Risk Scoring Systems to Predict the Outcome in ASA-PS V Patients Undergoing Surgery: A Retrospective Cohort Study. *Medicine* (Baltimore) 2016; **95**: e3238 [PMID: 27043696 DOI: 10.1097/MD.0000000000003238]

26 **Hong S**, Wang S, Xu G, Liu J. Evaluation of the POSSUM, p-POSSUM, o-POSSUM, and APACHE II scoring systems in predicting postoperative mortality and morbidity in gastric cancer patients. *Asian J Surg* 2017; **40**: 89-94 [PMID: 26420667 DOI: 10.1016/j.asjsur.2015.07.004]

27 **Crea N**, Di Fabio F, Pata G, Nascimbeni R. APACHE II, POSSUM, and ASA scores and the risk of perioperative complications in patients with colorectal disease. *Ann Ital Chir* 2009; **80**: 177-181 [PMID: 20131533]

**P-Reviewer:** Yan SL **S-Editor:** Cui LJ **L-Editor:** Filipodia **E-Editor:**

**Specialty type:** Medicine, research and experimental

**Country of origin:** India

**Peer-review report classification**

Grade A (Excellent): 0

Grade B (Very good): 0

Grade C (Good): C

Grade D (Fair): 0

Grade E (Poor): 0

**Table 1 Discriminating ability of APACHE-II**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **APACHE-II** | **Survived** | | **Mortality** | | ***P* value** |
| **Frequency** | **%** | **Frequency** | **%** |
| < 24 | 132 | 98.5 | 4 | 17.4 | < 0.001 |
| ≥ 24 | 2 | 1.5 | 19 | 82.6 |
| Total | 134 | 100 | 23 | 100 |

APACHE-II: Acute physiology and chronic health evaluation II.

**Table 2 Discriminating ability of P-POSSUM**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **P-POSSUM** | **Survived** | | **Mortality** | | ***P* value** |
| **Frequency** | **%** | **Frequency** | **%** |
| < 63 | 133 | 99.3 | 2 | 8.7 | < 0.001 |
| > 63 | 1 | 0.7 | 21 | 91.3 |
| Total | 134 | 100 | 23 | 100 |

P-POSSUM: Physiological and operative severity for the enumeration of mortality and morbidity.

**Table 3 Sensitivity and specificity of APACHE-II and P-POSSUM**

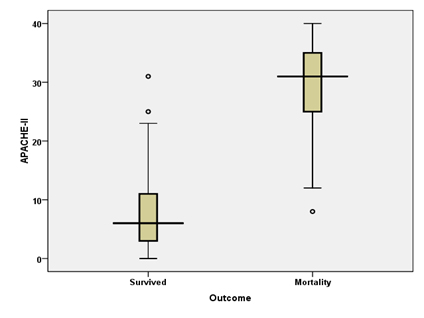
|  |  |  |  |
| --- | --- | --- | --- |
|  | **APACHE-II** | **P-POSSUM** | ***P* value** |
| Sensitivity | 82.6% | 91.3% | 0.665 |
| Specificity | 98.5% | 99.3% | 1.000 |

APACHE-II: Acute physiology and chronic health evaluation II; P-POSSUM: Physiological and operative severity for the enumeration of mortality and morbidity.

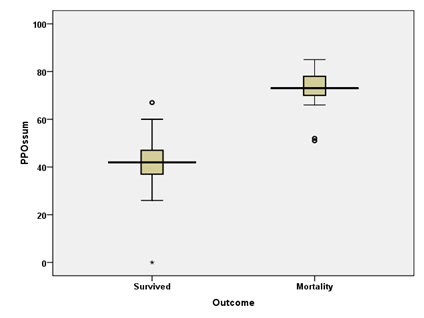
**Table 4 Discriminating ability of APACHE-II and P-POSSUM in predicting the secondary outcomes**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **APACHE-II** | **P-POSSUM** |
| LOS | *R* | 0.322 | 0.374 |
| *P* value | < 0.001 | < 0.001 |
| Ventilatory support | *R* | 0.554 | 0.572 |
| *P* value | < 0.001 | < 0.001 |
| Ionotropic support | *R* | 0.573 | 0.544 |
| *P* value | < 0.001 | < 0.001 |
| Re-exploration | *R* | 0.112 | 0.178 |
| *P* value | 0.161 | 0.026 |
| AMI or arrhythmia | *R* | 0.507 | 0.518 |
| P value | < 0.001 | < 0.001 |
| AKI | *R* | 0.507 | 0.518 |
| *P* value | < 0.001 | <0.001 |

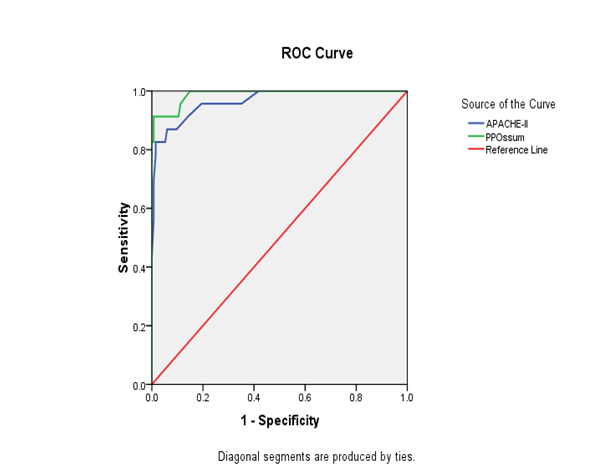
APACHE-II: Acute physiology and chronic health evaluation II; P-POSSUM: Physiological and operative severity for the enumeration of mortality and morbidity; LOS: Length of stay; AKI: Acute kidney injury; AMI: Acute myocardial infarction.

****

**Figure 1 Box-plots in Pearson correlation coefficient using APACHE-II.** APACHE-II: Acute physiology and chronic health evaluation II.



**Figure 2 Box-plots in Pearson correlation coefficient using P-POSSUM.** P-POSSUM: Physiological and operative severity for the enumeration of mortality and morbidity.

**Figure 3 Receiver operating characteristics curve for APACHE-II and P-POSSUM using the Multivariate logistic regression model.** APACHE-II: Acute physiology and chronic health evaluation II; P-POSSUM: Physiological and operative severity for the enumeration of mortality and morbidity; ROC: Receiver operating characteristics curve.