# P-POSSUM and the NELA Score overpredict mortality for laparoscopic emergency bowel surgery: an analysis of the NELA database

Authors: Mr Alexander R Darbyshire MBBS BSc MRCS[1], Dr Ina Kostakis BSc PhD[2], Mr Philip H Pucher MD PhD FRCS[1], Prof David Prytherch BSc PhD[2] and Mr Stuart J Mercer DM MA BM BCh FRCS[1].

1. Department of General Surgery, Portsmouth Hospitals University NHS Trust. Southwick Hill Road, PO6 3LY.
2. Centre for Healthcare Modelling and Informatics, University of Portsmouth. Buckingham Building, Lion Terrace, PO1 3HE.

Corresponding author: alexander.darbyshire@nhs.net. Mobile: 07980322141. Twitter handle: @AlexDarbyshire2

Keywords: risk-adjustment, emergency laparotomy, laparoscopy, P-POSSUM, NELA score.

Running title: P-POSSUM and the NELA Score overpredict mortality for emergency laparoscopy.

Manuscript word count: 2,611 words.

Conflicts of interest: there are no conflicts of interest to disclose for any of the authors.

**Abstract**

*Background*

Risk-stratification has become a key part of the care processes for patients having emergency bowel surgery. This study aimed to determine if operative approach influences risk-model performance, and risk-adjusted mortality rates in the United Kingdom.

*Methods*

A prospectively planned analysis was conducted using National Emergency Laparotomy Audit (NELA) data from December 2013 to November 2018. The risk-models investigated were P-POSSUM and the NELA Score, with model performance assessed in terms of discrimination and calibration. Risk-adjusted mortality was assessed using Standardised Mortality Ratios (SMR). Analysis was performed for the total cohort, and cases performed open, laparoscopically and converted to open. Sub-analysis was performed for cases with ≤20% predicted mortality.

*Results*

Data was available for 116 396 patients with P-POSSUM predicted mortality, and 46 935 patients with the NELA score. Both models displayed excellent discrimination with little variation between operative approaches (c-statistic: P-POSSUM 0.801-0.836; NELA Score 0.811-0.862). The NELA score was well calibrated across all deciles of risk, but P-POSSUM over-predicted risk beyond 20% mortality. Calibration plots for operative approach demonstrated that both models increasingly over-predicted mortality for laparoscopy, relative to open and converted to open surgery. SMRs calculated using both models consistently demonstrated that risk-adjusted mortality with laparoscopy was a third lower than open surgery.

*Conclusion*

Risk-adjusted mortality for emergency bowel surgery is lower for laparoscopy than open surgery, with P-POSSUM and NELA score both over-predicting mortality for laparoscopy. Operative approach should be considered in the development of future risk-models that rely on operative data.

## Introduction

The National Emergency Laparotomy Audit (NELA) has reported on the processes and results of emergency bowel surgery in England and Wales since 2013. It was started in response to high overall mortality rates and has produced modest but sustained improvements in outcomes[1]. It requires hospitals to adhere to set standards of care, notably consultant delivered surgery and improved access to intensive care post-operatively[2,3]. A cornerstone of these changes has been the pre-operative risk stratification of patients, with prioritisation of high-risk cases. The audit initially used P-POSSUM (Portsmouth-Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity) to risk stratify patients but identified that this over-predicted mortality in higher-risk cases[4]. Subsequently a new risk-model, the NELA score, was developed using data from the first two-years of the audit and is now used as its risk stratification tool[5]. Both models use a similar range of biochemical, physiological and operative variables in their regression equation, and have been externally validated[4–8].

Standardised mortality ratios (SMR) are an established method of comparing clinical outcomes. They are the ratio of observed versus expected deaths in a population, once crude mortality has been adjusted for, often using a regression model developed from important case-mix variable[9,10]. Although P-POSSUM and the NELA score have now been popularised for providing pre-operative risk predictions in emergency laparotomy, they were originally developed to provide hospital level risk-adjusted statistics such as SMRs[4,5].

Laparoscopy is routinely used for elective gastrointestinal surgery, with well recognised patient benefits[11–16]. Although its uptake in emergency surgery is limited in the UK, there has been a small increase in NELA cases attempted (16.7% to 20.2%) and completed laparoscopically (8.1% to 10.0%)[1]. However, neither risk-model considers operative approach as a variable in their algorithm. This analysis aims to assess the performance of the NELA score and P-POSSUM at generating risk-adjusted statistics for each operative approach at a population level.

## Methods

The NELA dataset is a prospectively collected national registry of all emergency bowel surgery performed in England and Wales. The recorded data items and inclusion criteria have been previously reported[1]. Anonymised data collected between 01/12/2013 to 31/11/2018 were extracted. Data analysis is permitted under the 2006 NHS Act.

Data on patient age, sex, pre-operative predicted mortality, grade of surgeon, degree of peritoneal soiling, presence of malignancy, timing of surgery, intensive care stay, post-operative length of stay and 30-day mortality were available for analysis. The risk-models evaluated in this analysis were P-POSSUM and the NELA score[4,5]. The primary outcome used was 30-day mortality. Operative approach is categorised by NELA as: open, laparoscopic, converted to open and laparoscopically assisted. Laparoscopically assisted cases were excluded from analysis because the procedure is not clearly defined, and there were insufficient cases for robust analysis (1.2% of dataset)[17]. Procedure specific sub-analysis have not been conducted as both models are general risk-models designed to be used on a range of procedures[4,5]. The increase in laparoscopy rate was only 1.6% over the 6-year study period, so a temporal analysis of model performance over time was not conducted.

*Statistical Analysis*

Model discrimination was assessed by plotting receiver operator characteristic (ROC) curves for each operative approach and calculating the c-statistic. Discrimination demonstrates how likely the model is to correctly classify a binary outcome, in this case whether a patient is dead or alive at 30-days. A c-statistic of 0.7 is considered good and 0.8 excellent; with 0.5 being no better at predicting the outcome than random chance[18]. Model calibration was assessed by dividing data evenly into deciles of predicted risk and plotting the ratio of observed vs predicted deaths for each decile. This demonstrates how close the model's predictions match observed outcomes as risk increases. Calibration plots were generated for P-POSSUM and NELA score as the total cohort and by operative approach. P-POSSUM is known to over-predict risk above 20% mortality, so a sub-analysis was conducted using only patients with predicted risk between 0-20%[5]. Linear regression lines are plotted to display trends with a line of best fit where the ratio of observed to expected deaths is exactly equal[17]. SMRs were calculated using P-POSSUM and NELA score predicted mortalities for each operative approach, and sub-study with 0-20% predicted mortality. An SMR of 1 indicates that the ratio of observed vs predicted deaths is exactly equal. If the SMR increases or decreases, then risk-adjusted mortality rates are respectively higher or lower than predicted by the model. Available demographic, operative factors and outcomes have been compared for each operative approach. Categorical variables are presented as counts (proportions) and compared with $\chi^2$ test. Continuous variables are presented as median (interquartile range) and compared with Kruskal-Wallis Test. Data analysis has been performed with R: A language and environment for statistical computing (Vienna, Austria). This article has been written in line with the STROBE guidelines (STROBE checklist included in Appendix 4)[19].

**Results**

A total of 116 396 patients were included in analysis (see Figure 1). The NELA score was introduced part-way through this dataset, with predicted mortality available for 46 935 patients[20]. Complete data was available for operative approach and 30-day mortality. Some data items displayed in Table 1 had a small amount of missing data (0.3-3.9%) but were kept in analysis as the key outcomes were available (see Appendix 1 for details).

Summary of demographic, operative and patient outcomes is presented in Table 1. Patients undergoing laparoscopic surgery were younger, had lower predicted mortality, rates of peritoneal contamination or presence of malignancy. They were more likely to be operated on by a consultant surgeon during daytime. In this unadjusted analysis, length of stay and 30-day mortality was lower for all cases started laparoscopically. The most frequently performed surgical procedures are summarised in Table 2, with adhesiolysis, right hemicolectomy and perforated peptic ulcer repair having the highest rates of laparoscopy.

Both models displayed excellent discrimination, with the NELA score outperforming P-POSSUM and a c-statistic of 0.862 compared to 0.809. ROC curves plotted for each operative approach showed only minor variation (c-statistics displayed in Table 3). This demonstrates that irrespective of operative approach both models will correctly predict whether a patient will be dead or alive at 30-days more than 80% of the time (see Appendix 2 for plots).

Calibration plot of all cases demonstrated that P-POSSUM's predictions closely followed the line of best fit up to 20% predicted mortality, beyond which it increasingly over predicted mortality, while the NELA score's predictions matched the observed proportion of deaths well throughout all deciles (see Figure 2). P-POSSUM similarly over predicted mortality for each operative approach, but more so for laparoscopy (see Plot 1 in Figure 3). The NELA score's mortality predictions for open surgery and cases converted to open, closely matched observed mortality across all risk deciles, but it too increasingly over predicted risk for laparoscopic surgery (see Plot 2 in Figure 3).

A sub-analysis was conducted using only patients with a predicted P-POSSUM mortality of ≤20%, as its calibration showed reasonable performance up to this threshold, with the NELA score for comparison. P-POSSUM's predictions more closely followed the line of best fit for open surgery, but it continued to overpredict mortality for laparoscopic surgery and cases converted to open, as mortality increased (see Plot 3 in Figure 4). The NELA score consistently provided accurate predictions of mortality across all deciles of risk for open and converted to open surgery. However, it continued to increasingly overpredict mortality for laparoscopic surgery (see Plot 4 in Figure 4; larger images of the plots in Figure 4 are included in Appendix 2).

Standardised mortality ratios (SMR) calculated using both risk-models for each operative approach are displayed in Table 4. As expected from the calibration plots, P-POSSUM for the total cohort adjusted poorly for mortality. In the sub-analysis of ≤20% predicted mortality, where P-POSSUM is best calibrated, the SMR for laparoscopic surgery was half that of open surgery (0.81 vs 0.40); with cases converted to open in-between (0.65). The NELA score produced very similar SMRs in the total cohort and sub-analysis and was excellent at adjusting for mortality in open surgery (0.92-0.94) and cases converted to open (0.86-0.91). For laparoscopic surgery it performed poorly, overpredicting mortality by a third (0.67-0.68).

**Discussion**

This population-based analysis has demonstrated that emergency bowel surgery performed laparoscopically has a much lower risk adjusted mortality than open surgery, calculated using two widely utilised and established risk-models[5,7,8]. The NELA score and P-POSSUM demonstrated excellent discrimination irrespective of operative approach, supporting their position as effective risk-adjustment tools. The NELA score was well calibrated across all risk bands, but P-POSSUM was limited to 20% predicted mortality, beyond which it over-predicted risk. Calibration plots for each operative approach however, demonstrated that both models increasingly over-predicted mortality for laparoscopy, relative to open surgery. Similarly, standardised mortality ratios calculated with both models consistently demonstrated that risk-adjusted mortality with laparoscopy was a third lower than open surgery. This raises the significant question of why is risk-adjusted mortality lower for patients who undergo laparoscopic surgery?

It could quite reasonably be a problem with the models, methodology or the datasets. Since *Jarman's* seminal work using Hospital Episode Statistics to calculate SMRs for English hospitals, they have been plagued by controversy[9,10]. The potential for inadequate case-mix adjustment due to classification errors, leading to unreliable risk-adjusted outcomes, has been clearly demonstrated with the Dr Foster Unit algorithm[21]. While possible that the coding of data in the NELA dataset has inaccuracies, the variables used in both models such as blood tests are robust to this flaw, and the majority of them leave little room for significant interobserver error. Saying this, a recent National survey of UK Consultant Surgeons found that 18% of respondents were unaware that laparoscopic cases were reported to NELA [22]. This certainly suggests there may be an element of under-reporting of laparoscopic cases in some centres, though case ascertainment for NELA has always been good at approximately 80%[1,23].

Both models were developed on datasets of patients who've had open abdominal surgery, and one could reasonably argue that they are therefore not representative of the patients selected for laparoscopic surgery[4,5]. The cohort of patients suitable for emergency laparoscopy could have subtle differences, which if accounted for in the development dataset, would alter the coefficients. This may reduce the overall accuracy of the model but improve its predictions for emergency laparoscopy. We have observed in simple comparative analysis that patients undergoing laparoscopy were younger with lower predicted mortality, less likely to have peritoneal soiling or active malignancy and more likely to be operated on by a consultant during the daytime. While the numeric differences are small, it suggests that the population selected for successful laparoscopy have factors supportive of more positive outcomes. It is likely that there is a degree of selection bias favouring these patients. Analysis of NELA data investigating associations between pre-operative factors and operative approach identified that laparoscopy was less likely if patients had a higher predicted mortality or ASA grade, a pre-operative CT scan or recent surgery, and more likely to be performed by consultant surgeon[24]. This again implies there are differences between patients having open and laparoscopic surgery, that need to be recognised in future risk-models. P-POSSUM's poor calibration is likely due to its development on outdated data from 1993-1995, on patients having elective and emergency bowel and vascular surgery. Surgical outcomes have improved substantially since then, so the dataset is no longer representative of the modern surgical population.

Bearing this in mind most gastro-intestinal surgery whether it be laparoscopic or open remains similar in terms of its operative steps. In addition, both risk-models use similar demographic, clinical, biochemical and operative data in their algorithms. All of these variables apply to patients requiring emergency bowel surgery irrespective of the operative approach. It is hard to see how patient age, history of cardio-respiratory disease or presence of renal failure on blood tests would influence mortality predictions in patients having open but not laparoscopic surgery. While there is likely to be selection bias in which patients are chosen for laparoscopy instead of open surgery, the variables that influence mortality should still be adjusted for by the risk-model. Some models do consider operative approach as a variable, such as the Surgical Risk Score, which allows predictions to be made for a range open and laparoscopic procedures[25,26]. However, it's performance in emergency surgery has proven inferior to other new models such as POTTER, and likely due to it being developed on data from mainly elective cases[27]. Further research on this topic is clearly needed, and a detailed analysis of the variables used to derive the P-POSSUM and NELA mortality predictions for each operative approach would provide a useful insight into why they are overpredicting mortality for laparoscopy.

An alternative theory is that laparoscopy confers a survival advantage over open emergency bowel surgery, something which is not measured by the models. A study using NELA data controlled for confounding factors by matching laparoscopic to open cases (n=11,753 and n=23,506 respectively) prior to comparative analysis. It clearly demonstrated that cases started laparoscopically had significantly lower mortality (6.0 vs 9.1%, p=< 0.001), blood loss (<100 ml, 64.4 vs 52.0%, p=< 0.001), and duration of hospital stay (median 8 [IQR 5–14] vs 10 [7–18] days, p=< 0.001)[28]. This compelling evidence suggests that laparoscopy may actually be superior to an open approach. Further research is however, limited and disease specific. The LASSO randomised trial compared laparoscopic to open adhesiolysis, and while detecting reduced length of stay, was not powered to detect change in mortality[29]. The LaCeS randomised feasibility study demonstrated that emergency laparoscopic colorectal surgery is safe but was only designed to inform a future trial[30]. Other registry-based studies investigating adhesiolysis, perforated peptic ulcer repairs and colorectal resections have reported improved outcomes with laparoscopy but are restricted by their observational nature[31–33]. It does make intuitive sense that the reduced post-operative pain and physiological burden conferred by minimally invasive surgery could translate to a survival advantage, in the high-risk patients needing emergency bowel surgery[34–36]. While this study provides evidence to support this theory, the findings are not conclusive. A randomised trial comparing laparoscopic to open emergency bowel surgery would provide evidence to substantiate this theory and the data to build a representative risk-model for both operative approaches.

This study is of course limited by its observational nature and the potential for selection bias that comes with it. We believe that the use of a large national dataset and transparent methodology overcomes much of this limitation. In addition, a case-matched analysis of NELA data controlling for available confounding factors has demonstrated that mortality is reduced with emergency laparoscopy[28]. Standardised mortality ratios also have their own limitations, which we have discussed, but feel that this alone does not explain our findings. The strength of this study is that it uses a prospectively collected national dataset providing reliable population level data, with relevant and generalisable findings. We have used the two risk-models adopted or developed by the NELA team to provide risk-adjusted statistics for emergency bowel surgery and conducted a thorough assessment of their performance. We have demonstrated at a population level that emergency bowel surgery performed laparoscopically has a lower risk-adjusted mortality than open surgery[28].

While this study cannot fully account for the difference in outcomes observed, it raises a number of questions. It is clear that operative approach needs to be considered in the development or recalibration of future risk-models for emergency bowel surgery. The NELA score was recently re-calibrated with updated coefficients, but no additional exploratory analysis was performed[37]. At present, surgeons who offer laparoscopy to patients needing emergency bowel surgery cannot confidently use either model to provide accurate predictions of mortality. Furthermore, it raises the question of why risk-adjusted mortality is so much lower with laparoscopy, which merits further research.

**Author Contributions**

Project initiation: SM and AD. Planning of statistical analysis: AD, IK, DP and PP. Cleaning and analysis of data: AD, IK and SM. Interpretation of results: AD, IK, DP, PP and SM. Drafting of initial version of paper: AD and IK. Revision of paper: AD, IK, DP, PP and SM.

**Conflicts of Interest**

There are no conflicts of interest to disclose.

**Funding**

This project has not received any research funding.

**References**


1. The NELA Project Team. *Sixth Patient Report of the National Emergency Laparotomy Audit*. 1–51 https://www.nela.org.uk/reports (2020).

2. Oliver, C. M. *et al.* Organisational factors and mortality after an emergency laparotomy: multilevel analysis of 39 903 National Emergency Laparotomy Audit patients. *BJA* **121**, 1346–1356 (2018).

3. Aggarwal, G. *et al.* Evaluation of the Collaborative Use of an Evidence-Based Care Bundle in Emergency Laparotomy. *JAMA Surgery* **154**, e190145-10 (2019).

4. Prytherch *et al.* POSSUM and Portsmouth POSSUM for predicting mortality. *British Journal of Surgery* **85**, 1217–1220 (1998).

5. Eugene, N. *et al.* Development and internal validation of a novel risk adjustment model for adult patients undergoing emergency laparotomy surgery: the National Emergency Laparotomy Audit risk model. *British Journal of Anaesthesia* **121**, 739–748 (2018).

6. Moonesinghe, S. R., Mythen, M. G., Das, P., Rowan, K. M. & Grocott, M. P. W. Risk Stratification Tools for Predicting Morbidity and Mortality in Adult Patients Undergoing Major Surgery. *Anesthesiology* **119**, 959–981 (2013).

7. Oliver, C. M., Walker, E., Giannaris, S., Grocott, M. P. W. & Moonesinghe, S. R. Risk assessment tools validated for patients undergoing emergency laparotomy: a systematic review. *BJA* **115**, 849–860 (2015).

8. Boyd-Carson, H. *et al.* The association of pre-operative anaemia with morbidity and mortality after emergency laparotomy. *Anaesthesia* **75**, 904–912 (2020).

9. Jarman, B. *et al.* Explaining differences in English hospital death rates using routinely collected data. *BMJ* **318**, 1515–1520 (1999).

10. Iezzoni, L. I. The risks of risk adjustment. *Jama* **278**, 1600–1607 (1997).

11. Guillou, P. J. *et al.* Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): multicentre, randomised controlled trial. *Lancet (London, England)* **365**, 1718–1726 (2005).

12. Jayne, D. G. *et al.* Five-year follow-up of the Medical Research Council CLASICC trial of laparoscopically assisted versus open surgery for colorectal cancer. *British Journal of Surgery* **97**, 1638–1645 (2010).

13. Biere, S. S. A. Y. *et al.* Minimally invasive versus open oesophagectomy for patients with oesophageal cancer: a multicentre, open-label, randomised controlled trial. *Lancet (London, England)* **379**, 1887–1892 (2012).

14. Pas, M. H. van der *et al.* Laparoscopic versus open surgery for rectal cancer (COLOR II): short-term outcomes of a randomised, phase 3 trial. *The Lancet Oncology* **14**, 210–218 (2013).

15. Bonjer, H. J. *et al.* A randomized trial of laparoscopic versus open surgery for rectal cancer. *New England Journal of Medicine* **372**, 1324–1332 (2015).

16. Mariette, C. *et al.* Hybrid Minimally Invasive Esophagectomy for Esophageal Cancer. *New England Journal of Medicine* **380**, 152–162 (2019).

17. Calster, B. V. *et al.* Calibration: the Achilles heel of predictive analytics. *BMC medicine* **17**, 230–7 (2019).

18. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* **44**, 837–845 (1988).

19. Elm, E. von *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of clinical epidemiology* **61**, 344–349 (2008).

20. The NELA Project Team. *Third Patient Report of the National Emergency Laparotomy Audit (NELA)*. 1–190 https://www.nela.org.uk/downloads/The%20Third%20Patient%20Report%20of%20the%20National%20Emergency%20Laparotomy%20Audit%202017%20-%20Full%20Patient%20Report.pdf (2017).

21. Mohammed, M. A. *et al.* Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ* **338**, b780-8 (2009).

22. Heywood, N. *et al.* The laparoscopy in emergency general surgery (LEGS) study: a questionnaire survey of UK practice. *Ann Royal Coll Surg Engl* **103**, 120–129 (2021).

23. The NELA Project Team. *The First Patient Report of the National Emergency Laparotomy Audit*. 1–151 https://www.nela.org.uk/All-Patient-Reports (2015).

24. Badrick, E. *et al.* Laparoscopy in emergency general surgery (The LEGS Study): NELA database analysis. *British Journal of Surgery* (2020).

25. Bilimoria, K. Y. B. M. M. *et al.* Development and Evaluation of the Universal ACS NSQIP Surgical Risk Calculator: A Decision Aid and Informed Consent Tool for Patients and Surgeons. *Journal of the American College of Surgeons* **217**, 833-842.e3 (2013).

26. Liu, Y., Cohen, M. E., Hall, B. L., Ko, C. Y. & Bilimoria, K. Y. Evaluation and Enhancement of Calibration in the American College of Surgeons NSQIP Surgical Risk Calculator. *Journal of the American College of Surgeons* **223**, 231–239 (2016).

27. Bertsimas, D., Dunn, J., Velmahos, G. C. & Kaafarani, H. M. A. Surgical Risk Is Not Linear: Derivation and Validation of a Novel, User-friendly, and Machine-learning-based Predictive OpTimal Trees in Emergency Surgery Risk (POTTER) Calculator. *Annals of Surgery* **268**, 574–583 (2018).

28. Pucher, P. H., Mackenzie, H., Tucker, V. & Mercer, S. J. A national propensity score-matched analysis of emergency laparoscopic versus open abdominal surgery. *British Journal of Surgery* 1–7 (2021) doi:10.1093/bjs/znab048.

29. Sallinen, V. *et al.* Laparoscopic versus open adhesiolysis for adhesive small bowel obstruction (LASSO): an international, multicentre, randomised, open-label trial. *The Lancet Gastroenterology & Hepatology* **4**, 278–286 (2019).

30. Harji, D. P. *et al.* Laparoscopic versus open colorectal surgery in the acute setting (LaCeS trial): a multicentre randomized feasibility trial. *British Journal of Surgery* **107**, 1595–1604 (2020).

31. Søreide, K., Thorsen, K. & Søreide, J. A. Strategies to improve the outcome of emergency surgery for perforated peptic ulcer. *British Journal of Surgery* **101**, e51-64 (2014).

32. Moghadamyeghaneh, Z., Talus, H., Ballantyne, G., Stamos, M. J. & Pigazzi, A. Short-term outcomes of laparoscopic approach to colonic obstruction for colon cancer. *Surgical Endoscopy* 1–11 (2020) doi:10.1007/s00464-020-07743-w.

33. Darbyshire, A. R., Kostakis, I., Pucher, P. H., Toh, S. & Mercer, S. J. The impact of laparoscopy on emergency surgery for adhesional small bowel obstruction: prospective single centre cohort study. *Annals of the Royal College of Surgeons of England* 1–8 (2021) doi:10.1308/rcsann.2020.7079.

34. Shibata, J. *et al.* Surgical stress response after colorectal resection: a comparison of robotic, laparoscopic, and open surgery. *Techniques in Coloproctology* **19**, 275–280 (2015).

35. Crippa, J., Mari, G. M., Miranda, A., Costanzi, A. T. M. & Maggioni, D. Surgical Stress Response and Enhanced Recovery after Laparoscopic Surgery - A systematic review. *Chirurgia* **113**, 455–10 (2018).

36. Foss, N. B. & Kehlet, H. Challenges in optimising recovery after emergency laparotomy. *Anaesthesia* **75**, A4876-7 (2020).

37. Martin, P. *NELA risk adjustment model: recalibration of predictor coefficients*. 1–6 https://data.nela.org.uk/getattachment/Support/NELA-Risk-Adjustment-Model/NELA-(2020)-Recalibrated-risk-model-coefficients.pdf.aspx?lang=en-GB (2020).

**Tables and Figures**

Table 1: summary of demographic, operative factors and patient outcomes for each operative approach.

| | Open[*] (n=99 414) | Laparoscopic[*] (n=9 089) | Converted[*] (n=7 893) | p value[†] |
|---|---|---|---|---|
| Age category (years) | | | | |
| 18 to 50 | 18 255 (18.4) | 3030 (33.3) | 2476 (31.4) | <0.001 |
| 51 to 70 | 34 698 (34.9) | 3024 (33.3) | 2717 (34.4) | |
| 71 to 80 | 26 068 (26.2) | 1694 (18.6) | 1571 (19.9) | |
| >81 | 20 393 (20.5) | 1341 (14.8) | 1129 (14.3) | |
| Sex | | | | |
| Female | 51 243 (51.5) | 4786 (52.7) | 4022 (51.0) | 0.064 |
| Male | 48 171 (48.5) | 4303 (47.3) | 3871 (49.0) | |
| Pre-operative predicted mortality (%) | | | | |
| P-POSSUM | 7.6 (2.9-23.7) | 3.5 (1.7-8.4) | 4.1 (1.8-11.9) | <0.001 |
| NELA Score | 5.0 (1.5-14.3) | 1.6 (0.5-5.5) | 2.2 (0.6-7.4) | <0.001 |
| Grade of surgeon | | | | |
| Consultant | 87 732 (88.2) | 8279 (91.1) | 7087 (89.8) | <0.001 |
| Post-CCT/SAS | 3927 (4.0) | 269 (3.0) | 269 (3.4) | |
| Registrar | 7755 (7.8) | 541 (6.0) | 537 (6.8) | |
| Peritoneal soiling | | | | |
| None | 36 212 (36.6) | 4352 (48.2) | 2305 (29.3) | <0.001 |
| Serous fluid | 27 327 (27.6) | 1923 (21.3) | 1902 (24.2) | |
| Localised pus | 9898 (10.0) | 1087 (12.0) | 1188 (15.1) | |
| Free contamination | 25 593 (25.8) | 1672 (18.5) | 2471 (31.4) | |
| Active malignancy | | | | |
| None | 76 171 (76.9) | 7252 (80.3) | 6478 (82.3) | <0.001 |
| Primary disease | 11 308 (11.4) | 889 (9.8) | 755 (9.6) | |
| Nodal metastases | 4361 (4.4) | 333 (3.7) | 261 (3.3) | |
| Distant metastases | 7217 (7.3) | 562 (6.2) | 373 (4.7) | |
| Time of surgery | | | | |
| 08:00-17:00 | 57 519 (60.2) | 6257 (71.6) | 4896 (64.3) | <0.001 |
| 17:00-08:00 | 38 010 (39.8) | 2479 (28.4) | 2719 (35.7) | |
| Intensive care length of stay (days) | | | | |
| | 0 (0-3) | 0 (0) | 0.0 (0-2) | <0.001 |
| Post-operative length of stay (days) | | | | |
| | 11 (7-20) | 7 (4-12) | 9 (6-16) | <0.001 |
| 30-day mortality | | | | |
| | 11 056 (11.1) | 311 (3.4) | 500 (6.3) | <0.001 |

[*] Data are displayed as counts (%), except pre-operative predicted mortality and post-operative length of stay which are median (IQR).

[†] Comparative statistics performed with $\chi^2$ test for categorical data and Kruskal-Wallis test for continuous data.

Table 2: summary of the most common surgical procedures by operative approach.

| Procedure | Open[‡] | Laparoscopic[‡] | Converted[‡§] |
|---|---|---|---|
| Adhesiolysis | 16 336 (84.1) | 1907 (9.8) | 1174 (6.0) |
| Small bowel resection | 16 998 (90.4) | 340 (1.8) | 1468 (7.8) |
| Right Colectomy | 12 735 (84.8) | 957 (6.4) | 1319 (8.8) |
| Hartmann's procedure | 13 509 (92.1) | 323 (2.2) | 830 (5.7) |
| Perforated peptic ulcer repair | 4882 (77.0) | 977 (15.4) | 478 (7.5) |
| Subtotal Colectomy | 5302 (85.1) | 721 (11.6) | 207 (3.3) |
| Stoma formation | 4243 (79.7) | 825 (15.5) | 257 (4.8) |
| Left Colectomy/Anterior Resection | 3425 (87.8) | 247 (6.3) | 227 (5.8) |
| Drainage of abscess | 2188 (72.9) | 443 (14.8) | 369 (12.3) |
| Washout only | 1837 (65.8) | 726 (26.0) | 227 (8.1) |



[‡] Data are displayed as counts (%).

[§] Converted to open.

Table 3: c-statistics for P-POSSUM and NELA score by operative approach.

| | P-POSSUM (95% CI)** | NELA Score (95% CI)** |
|---|---|---|
| All Cases | 0.809 (0.805-0.813) | 0.862 (0.857-0.867) |
| Open | 0.801 (0.797-0.805) | 0.854 (0.849-0.860) |
| Laparoscopic | 0.836 (0.815-0.858) | 0.811 (0.857-0.906) |
| Converted to Open | 0.811 (0.793-0.831) | 0.855 (0.866-0.906) |

** 95% confidence intervals.

Table 4: standardised mortality ratios calculated using both risk-models by operative approach.

| Risk-model | Operative Approach | n=[††] | SMR[‡‡] (95% CI)[§§] |
|---|---|---|---|
| P-POSSUM | Open | 99 550 | 0.61 (0.60-0.62) |
| | Laparoscopic | 9095 | 0.39 (0.35-0.44) |
| | Converted to Open | 7900 | 0.56 (0.51-0.61) |
| P-POSSUM predicted mortality ≤20% | Open | 71 552 | 0.81 (0.79-0.84) |
| | Laparoscopic | 8065 | 0.40 (0.34-0.47) |
| | Converted to Open | 6625 | 0.65 (0.57-0.74) |
| NELA Score | Open | 39 175 | 0.94 (0.91-0.97) |
| | Laparoscopic | 4205 | 0.68 (0.57-0.79) |
| | Converted to Open | 3603 | 0.91 (0.79-1.03) |
| NELA score predicted mortality ≤20% | Open | 32 073 | 0.92 (0.87-0.96) |
| | Laparoscopic | 3956 | 0.67 (0.53-0.81) |
| | Converted to Open | 3278 | 0.86 (0.70-1.03) |



[††] Number of cases in each group.
[‡‡] Standardised Mortality Ratio.
[§§] 95% confidence intervals.

Figure 1: study flow chart.

Figure 2: calibration plot for P-POSSUM and NELA Score.

Figure 3: calibration plots by operative approach 1. P-POSSUM; 2. NELA Score; 3. P-POSSUM ≤20% predicted mortality; 4. NELA Score ≤20% predicted mortality. The continuous bold black line in each plot represents the line of best fit where observed mortality exactly matches predicted mortality.

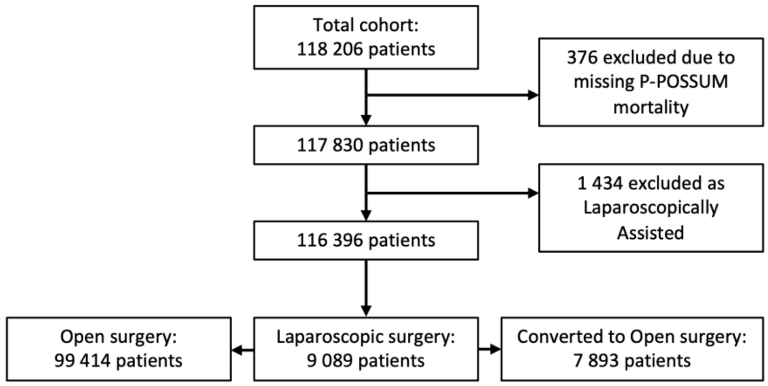



Figure 1: study flow chart.

793x458mm (72 x 72 DPI)

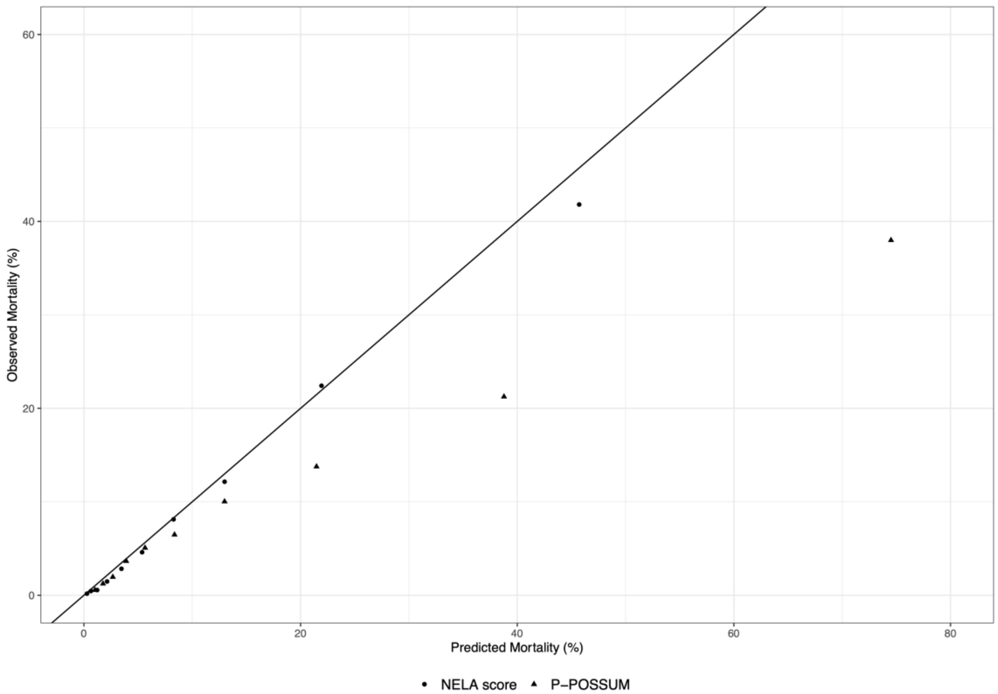



Figure 2: calibration plot for P-POSSUM and NELA Score.

1236x874mm (72 x 72 DPI)

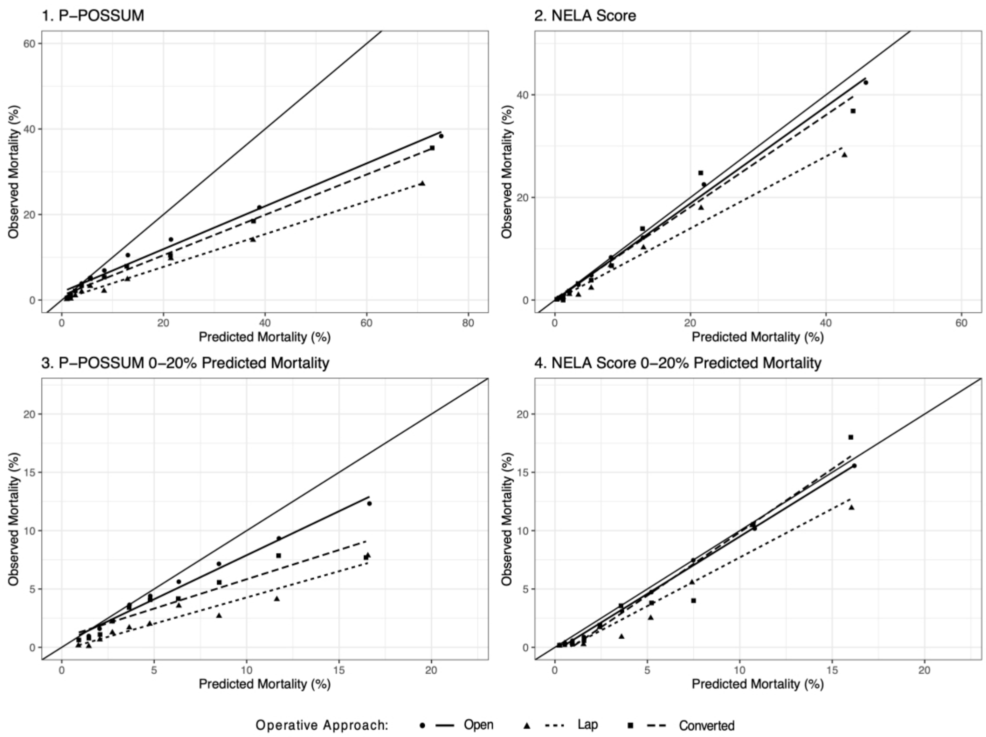
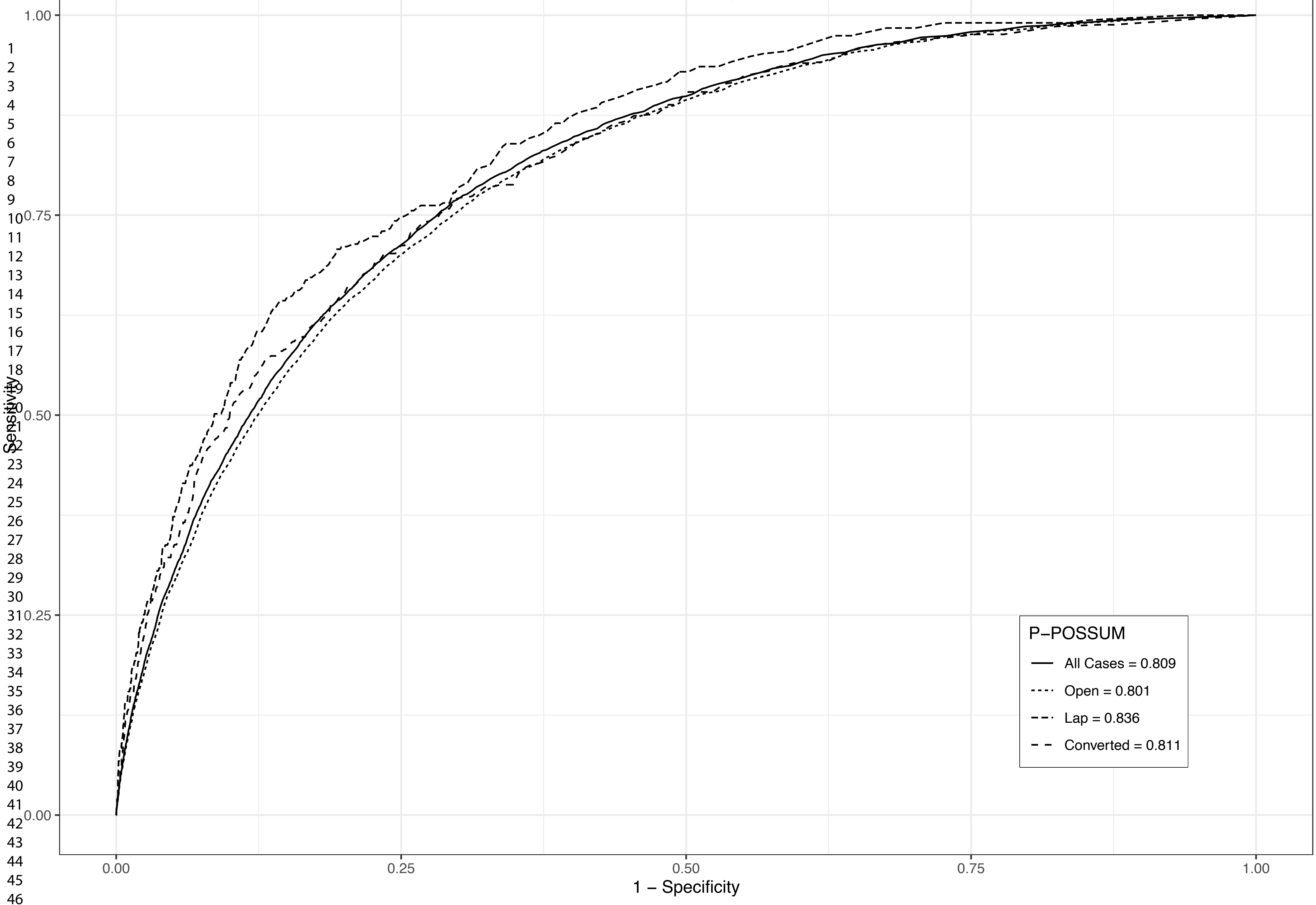


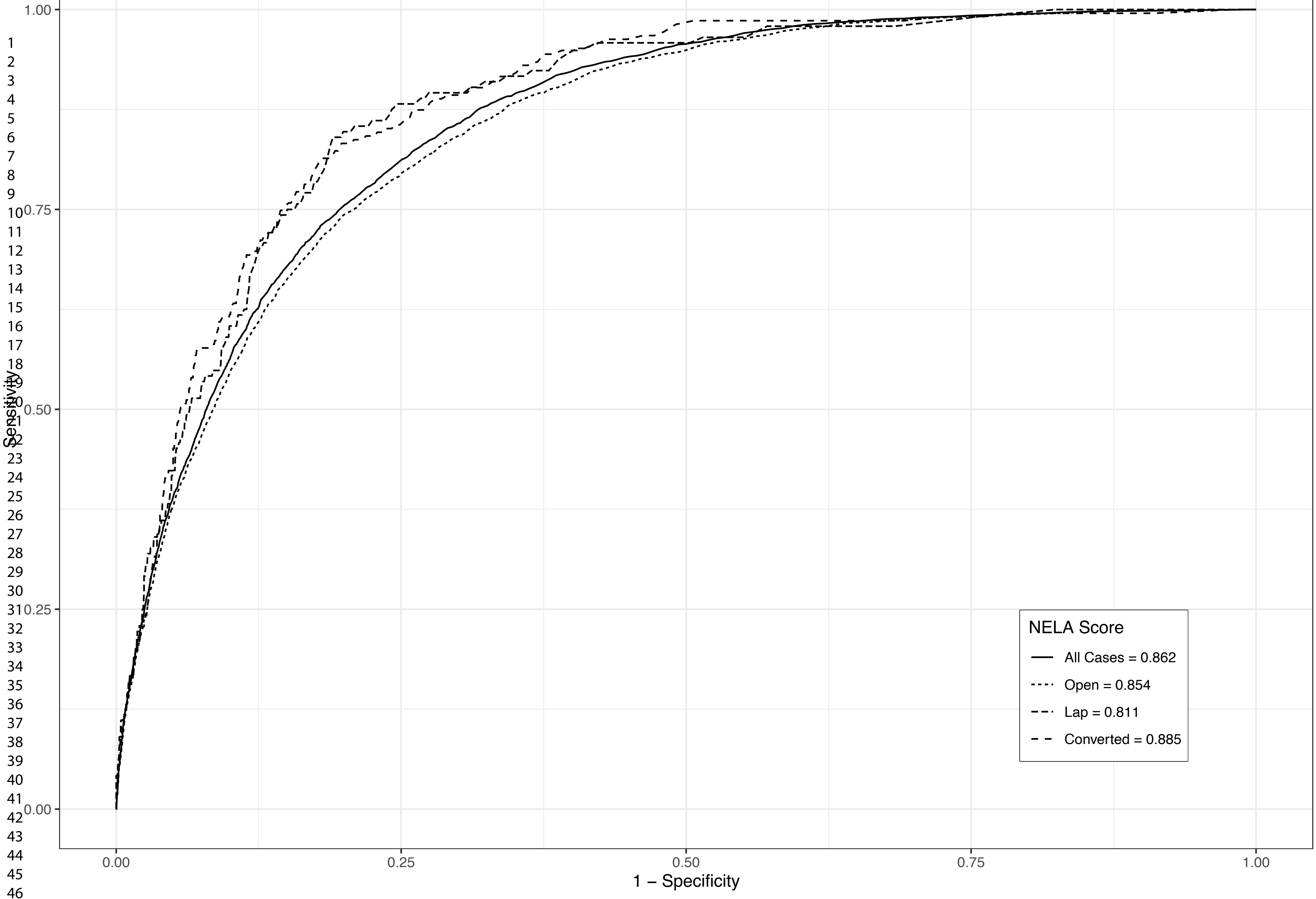

Figure 3: calibration plots by operative approach 1. P-POSSUM; 2. NELA Score; 3. P-POSSUM ≤20% predicted mortality; 4. NELA Score ≤20% predicted mortality. The continuous bold black line in each plot represents the line of best fit where observed mortality exactly matches predicted mortality.

1238x943mm (72 x 72 DPI)

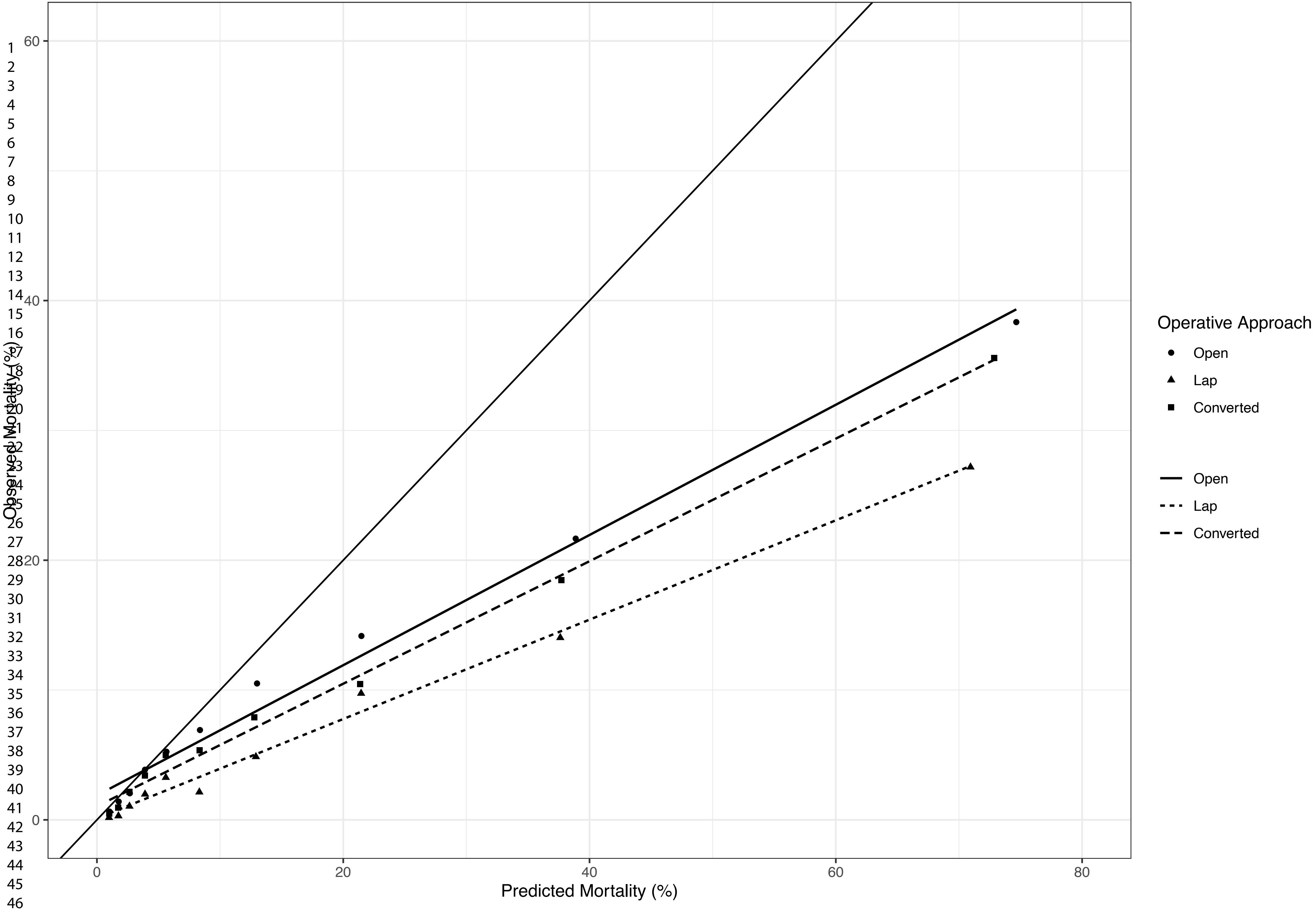
**Appendix 1. Missing Data Items**

| Data Item | Number Missing | Proportion Missing (%) |
|---|---|---|
| P-POSSUM predicted mortality | 376 | 0.3 |
| Procedure performed | 2064 | 1.7 |
| Peritoneal soiling | 492 | 0.4 |
| Active malignancy | 467 | 0.4 |
| Intensive care length of stay | 335 | 0.3 |
| Post-operative length of stay | 1376 | 1.2 |
| Time of surgery | 4615 | 3.9 |

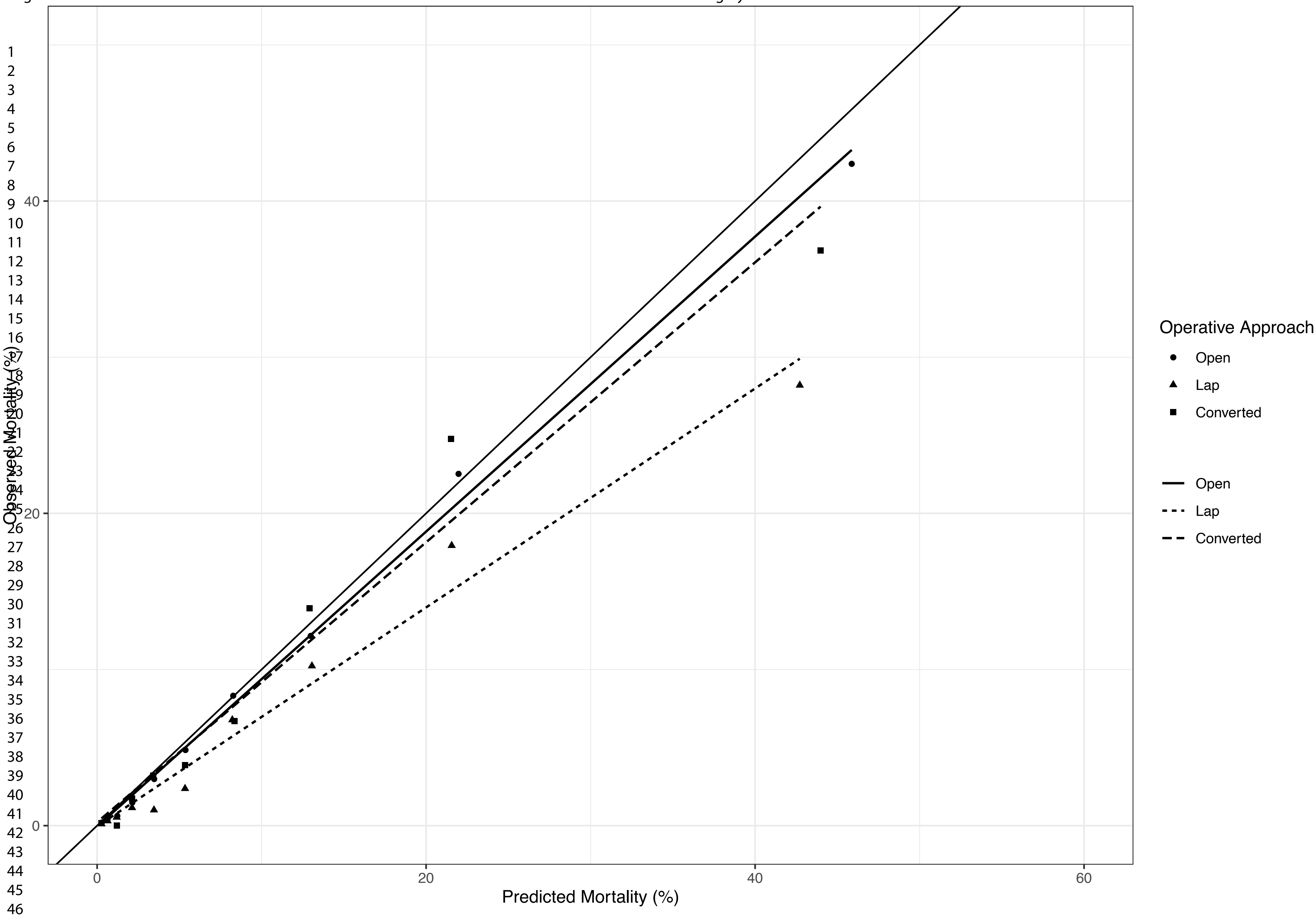

1.00
0.75
0.50
0.25
0.00
Sensitivity
0.00
0.25
0.50
0.75
1.00
1 − Specificity
P−POSSUM
All Cases = 0.809
Open = 0.801
Lap = 0.836
Converted = 0.811


1.00
0.75
0.50
0.25
0.00
Sensitivity
0.00
0.25
0.50
0.75
1.00
1 − Specificity
NELA Score
All Cases = 0.862
Open = 0.854
Lap = 0.811
Converted = 0.885

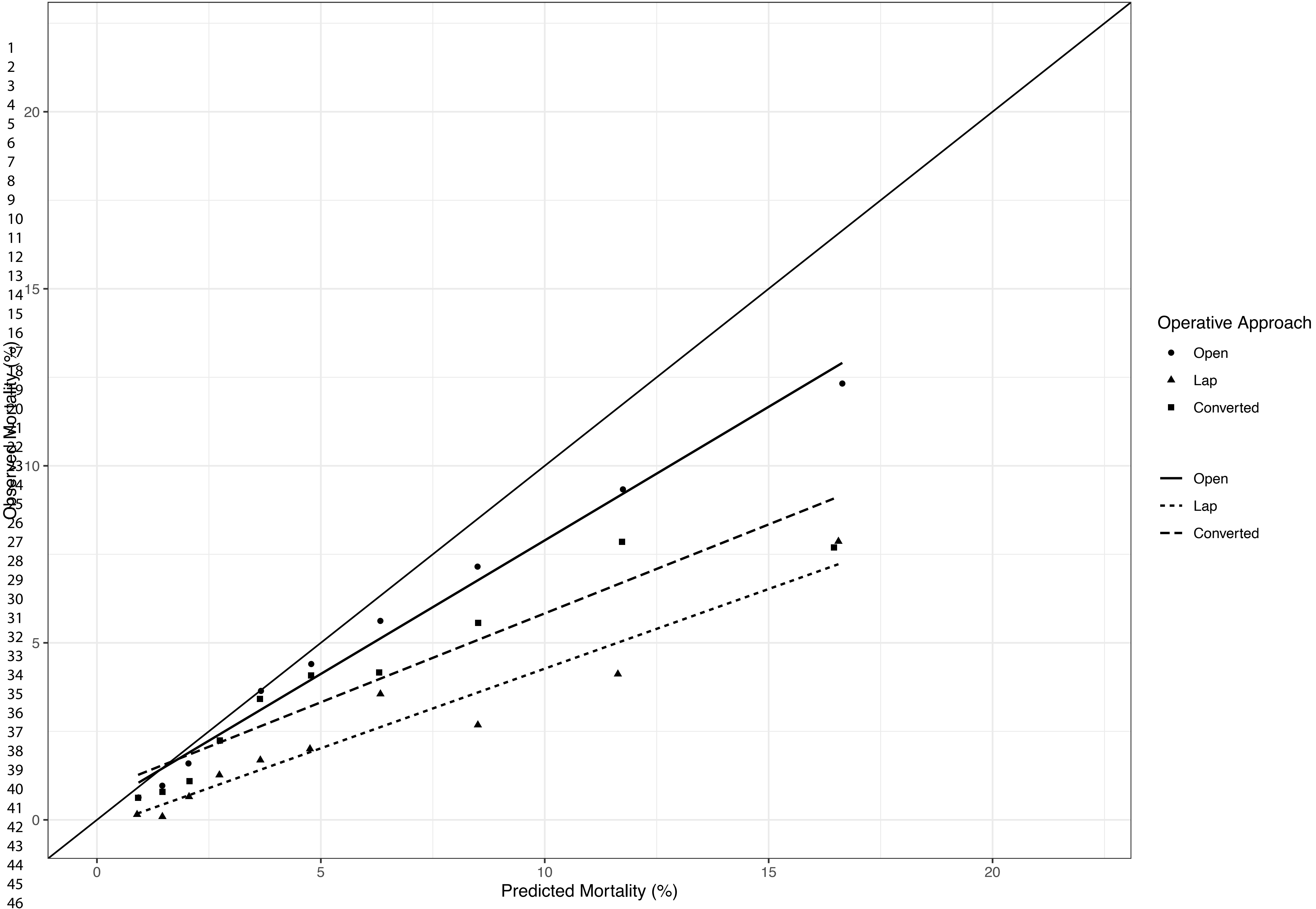
1. P–POSSUM


Observed Mortality (%)
60
40
20
0
0
20
40
60
80
Predicted Mortality (%)
Operative Approach
Open
Lap
Converted
Open
Lap
Converted

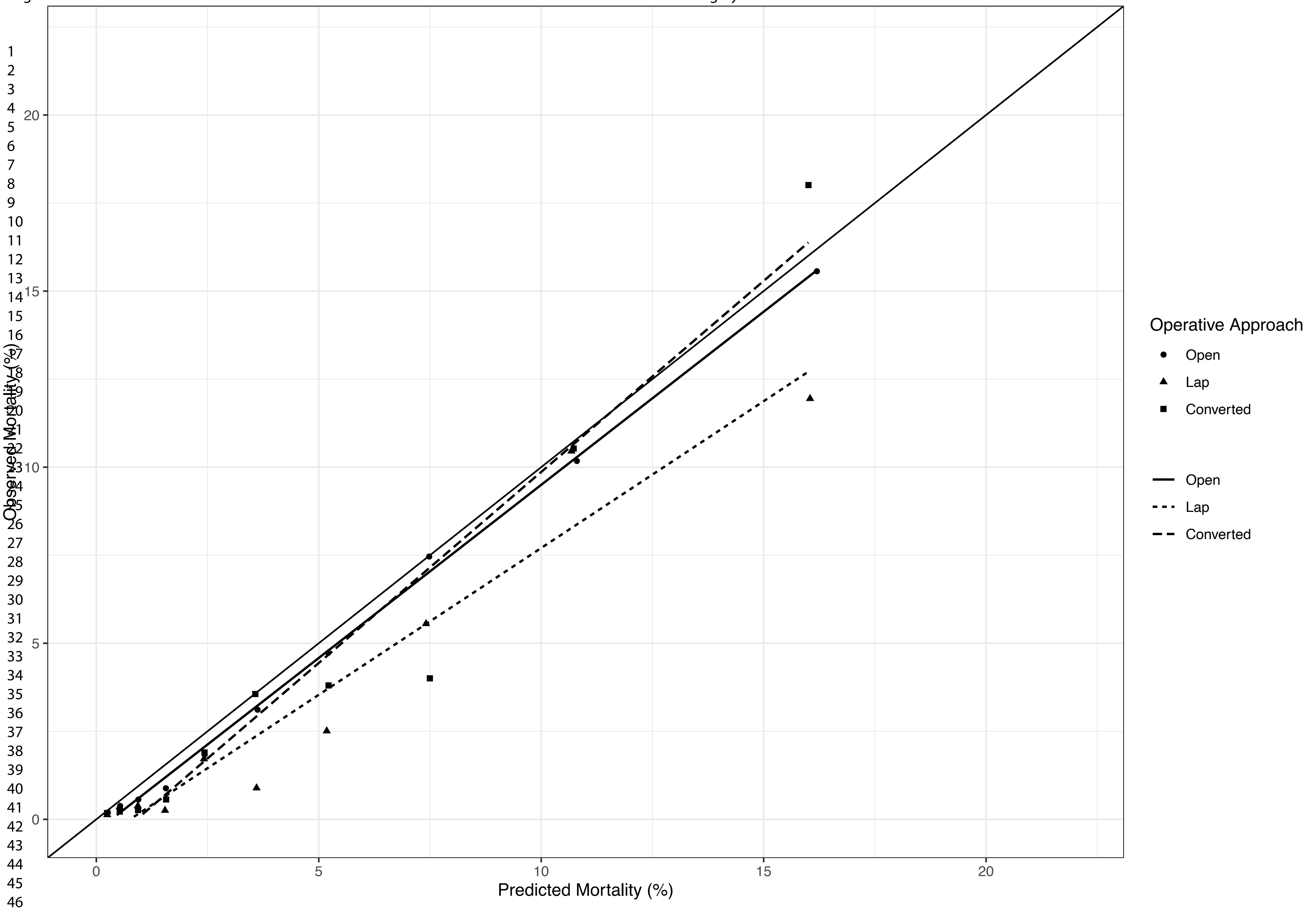
2. NELA Score


Observed Mortality (%)
40
20
0
0
20
40
60
Predicted Mortality (%)
Operative Approach
Open
Lap
Converted
Open
Lap
Converted

3. P–POSSUM 0–20% Predicted Mortality



4. NELA Score 0–20% Predicted Mortality



STROBE Statement—checklist of items that should be included in reports of observational studies

| | Item No. | Recommendation | Page No. | Relevant text from manuscript |
|---|---|---|---|---|
| **Title and abstract** | 1 | (*a*) Indicate the study's design with a commonly used term in the title or the abstract | 1 | *Methods*<br>A prospectively planned analysis was conducted using National Emergency Laparotomy Audit (NELA) data from December 2013 to November 2018. The risk-models investigated were P-POSSUM and the NELA Score, with model performance assessed in terms of discrimination and calibration. Risk-adjusted mortality was assessed using Standardised Mortality Ratios (SMR). |
| | | (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | 1-2 | Abstract |
| **Introduction** | | | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported | 3 | Introduction |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | 3 | Introduction (paragraph 3)<br>This analysis aims to assess the performance of the NELA score and P-POSSUM at generating risk-adjusted statistics for each operative approach at a population level. |
| **Methods** | | | | |
| Study design | 4 | Present key elements of study design early in the paper | 3-4 | Methods (paragraph 1)<br>The NELA dataset is a prospectively collected national registry of all emergency bowel surgery performed in England and Wales. The recorded data items and inclusion criteria have been previously reported[1]. Anonymised data collected between 01/12/2013 to 31/11/2018 were extracted.<br><br>(Paragraph 2)<br>The risk-models evaluated in this analysis were P-POSSUM and the NELA score[4,5]. The primary outcome used was 30-day mortality. |

| | | | | |
|---|---|---|---|---|
| | | | | (Paragraph 3)<br>Model discrimination was assessed by plotting receiver operator characteristic (ROC) curves for each operative approach and calculating the c-statistic.<br><br>Calibration plots were generated for P-POSSUM and NELA score as the total cohort and by operative approach. |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | 3 | Methods (paragraph 1)<br>The NELA dataset is a prospectively collected national registry of all emergency bowel surgery performed in England and Wales. The recorded data items and inclusion criteria have been previously reported[1]. Anonymised data collected between 01/12/2013 to 31/11/2018 were extracted. |
| Participants | 6 | (*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up | 3 | Methods (paragraph 1)<br>The NELA dataset is a prospectively collected national registry of all emergency bowel surgery performed in England and Wales. The recorded data items and inclusion criteria have been previously reported[1]. |
| | | | | |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | 3-4 | Methods (paragraph 2)<br>Data on patient age, sex, pre-operative predicted mortality, grade of surgeon, degree of peritoneal soiling, presence of malignancy, timing of surgery, intensive care stay, post-operative length of stay and 30-day mortality were available for analysis. The risk-models evaluated in this analysis were P-POSSUM and the NELA score[4,5]. The primary outcome used was 30-day mortality. Operative approach is categorised by NELA as: open, laparoscopic, converted to open and laparoscopically assisted.<br><br>Procedure specific sub-analysis have not been conducted as both models are general risk-models designed to be used on a range of |

| | | | | |
|---|---|---|---|---|
| | | | | procedures[4,5]. The increase in laparoscopy rate was only 1.6% over the 6-year study period, so a temporal analysis of model performance over time was not conducted. |
| Data sources/ measurement | 8* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | 3 | Methods (paragraph 1)<br>The NELA dataset is a prospectively collected national registry of all emergency bowel surgery performed in England and Wales. The recorded data items and inclusion criteria have been previously reported[1]. |
| Bias | 9 | Describe any efforts to address potential sources of bias | 6-7 | See paragraphs 2, 3 and 4 of the Discussion. |
| Study size | 10 | Explain how the study size was arrived at | 4-5 | Results (paragraph 1)<br>A total of 116 396 patients were included in analysis (see Figure 1). The NELA score was introduced part-way through this dataset, with predicted mortality available for 46 935 patients[20]. Complete data was available for operative approach and 30-day mortality. Some data items displayed in Table 1 had a small amount of missing data (0.3-3.9%) but were kept in analysis as the key outcomes were available (see Appendix 1 for details). |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | 4 | Methods (paragraph 3)<br>Continuous variables are presented as median (interquartile range) and compared with Kruskal-Wallis Test. |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | 4 | Methods (paragraph 3) |
| | | (*b*) Describe any methods used to examine subgroups and interactions | | Methods (paragraph 3) |
| | | (*c*) Explain how missing data were addressed | 4-5 | Results (paragraph 1)<br>Some data items displayed in Table 1 had a small amount of missing data (0.3-3.9%) but were kept in analysis as the key outcomes were available (see Appendix 1 for details). |
| | | (*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed | | NA |

| | | | | |
|---|---|---|---|---|
| | | (*e*) Describe any sensitivity analyses | | NA |
| **Results** | | | | |
| Participants | 13* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | 4-5 | Results (paragraph 1) |
| | | (b) Give reasons for non-participation at each stage | 4-5 | Results (paragraph 1) |
| | | (c) Consider use of a flow diagram | 4-5 | See Figure 1. |
| Descriptive data | 14* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | 5 | Results (paragraph 2) Summary of demographic, operative and patient outcomes is presented in Table 1. |
| | | (b) Indicate number of participants with missing data for each variable of interest | 4-5 | Results (paragraph 1) and see Appendix 1. |
| | | (c) *Cohort study*—Summarise follow-up time (eg, average and total amount) | 4 | Methods (paragraph 1 and 2) |
| Outcome data | 15* | *Cohort study*—Report numbers of outcome events or summary measures over time | 4-5 | Results (paragraph 1) |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | 5-6 | Results (paragraph 2, 3, 4, 5 and 6) |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | 5 | Results (paragraph 4 and 5) |
| **Discussion** | | | | |
| Key results | 18 | Summarise key results with reference to study objectives | 6 | Discussion (paragraph 1) |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | 6-8 | Discussion all paragraphs |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, | 6-8 | Discussion all paragraphs |

| | | | | |
|---|---|---|---|---|
| | | limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | | |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | 7-8 | Discussion (paragraphs 5, 6 and 7) |
| **Other information** | | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | | This study did not receive any funding. |