BJA
British Journal of Anaesthesia

Development and internal validation of a novel risk adjustment model for adult patients undergoing emergency laparotomy surgery: the NELA risk model.

Journal:	*British Journal of Anaesthesia*
Manuscript ID	BJA-2018-00292-PM025.R2
Article Type:	Clinical Investigation
Date Submitted by the Author:	01-Jun-2018
Complete List of Authors:	Eugene, Natalie; National Emergency Laparotomy Audit, Royal College of Anaesthetists; Royal College of Surgeons of England, Clinical Effectivness Unit Oliver, Charles M; National Emergency Laparotomy Audit, Royal College of Anaesthetists; University College London, Division of surgical outcomes and interventional science; University College London Hospitals NHS Foundation Trust, Anaesthetics Basset, Michael; National Emergency Laparotomy Audit, Royal College of Anaesthetists; Manchester University NHS Foundation Trust, Anaesthetics Poulton, Thomas; National Emergency Laparotomy Audit, Royal College of Anaesthetists; Northern School Of Anaesthesia & Intensive Care Medicine, Anaesthetics Kuryba, Angela; National Emergency Laparotomy Audit, Royal College of Anaesthetists; Royal College of Surgeons of England, Clinical Effectivness Unit Johnston, Carolyn; National Emergency Laparotomy Audit, Royal College of Anaesthetists; St Georges Healthcare NHS Trust, Anaesthetics Anderson, Iain; National Emergency Laparotomy Audit, Royal College of Anaesthetists; Salford Royal NHS Foundation Trust, Anaesthetics Moonesinghe, Suneetha; National Emergency Laparotomy Audit, Royal College of Anaesthetists; UCL / UCLH Surgical Outcomes Research Centre, Anaesthesia and Intensive Care Grocott, Michael; National Emergency Laparotomy Audit, Royal College of Anaesthetists; University Hospital Southampton NHS Foundation Trust, Anaesthesia and Critical Care Research Unit; University of Southampton, Integrative Physiology and Critical Illness Group Murray, Dave; National Emergency Laparotomy Audit, Royal College of Anaesthetists; James Cook

	University Hospital, Anaesthetics Cromwell, David; National Emergency Laparotomy , Royal College of Anaesthetists; Royal College of Surgeons of England, Clinical Effectivness Unit; London School of Hygiene and Tropical Medicine, Department of Health Services Research & Policy Walker, Kate; National Emergency Laparotomy Audit, Royal College of Anaesthetists; Royal College of Surgeons of England, Clinical Effectivness Unit; London School of Hygiene and Tropical Medicine, Department of Health Services Research & Policy
Mesh keywords:	emergency laparotomy, postoperative mortality, risk adjustment

Subedited BJA-2018-00292-PM025.R2; accepted 20th June; 3 figures, 3 tables, 1 Supplementary file (web-based)

Development and internal validation of a novel risk adjustment model for adult patients undergoing emergency laparotomy surgery: the NELA risk model

Eugene N[1,2], Oliver CM[1,3,4], Bassett MG[1,5], Poulton TE[1,5,6], Kuryba A[1,2], Johnston C[1,7], Anderson ID[1,8], Moonesinghe SR[1,4], Grocott MP[1,9,10], Murray DM[1,11], Cromwell DA[1,2,12], Walker K [1,2,12] on behalf of the NELA collaboration

1. National Emergency Laparotomy Audit, Royal College of Anaesthetists, London, UK
2. Clinical Effectiveness Unit, Royal College of Surgeons of England, UK.
3. Division of surgery and interventional science, University College London, 74 Huntley Street, London WC1E 6AU, UK
4. UCLH Surgical Outcomes Research Centre, Department of Anaesthesia and Perioperative Medicine, University College London Hospitals NHS Foundation Trust, UK
5. Department of Applied Health Research, University College London, UK
6. Northern School of Anaesthesia and Intensive Care Medicine, Newcastle, UK
7. St George NHS Hospital, London, UK
8. Salford Royal Foundation NHS Trust, Salford, UK
9. Anaesthesia & Critical Care Research Group, NIHR Biomedical Research Centre, University Hospital Southampton, UK.
10. Integrative Physiology and Critical Illness Group, Faculty of Medicine, University of Southampton, Southampton, UK
11. Department of Anaesthesia, James Cook University Hospital, Middlesbrough, UK
12. Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine, London, UK

Keywords:

general surgery; emergency laparotomy; postoperative mortality; risk adjustment;

Corresponding author

Prof David Cromwell

Department of Health Services Research and Policy

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place, London WC1H 9SH, United Kingdom

Email: david.cromwell@lshtm.ac.uk

Telephone: + (0)20 7869 6608

Word count: 2983

ABSTRACT

Background

Among patients undergoing emergency laparotomy, 30-day postoperative mortality is around 10-15%. The risk of death among these patients, however, varies greatly due to their clinical characteristics. We developed a risk prediction model for 30-day postoperative mortality to enable better comparison of outcomes between hospitals.

Methods

We analysed data from the National Emergency Laparotomy Audit (NELA) on patients having an emergency laparotomy between December 2013 and November 2015. A prediction model was developed using multivariable logistic regression, with potential risk factors identified from existing prediction models, national guidelines and clinical experts. Continuous risk factors were transformed if necessary to reflect their non-linear relationship with 30-day mortality. The performance of the model was assessed in terms of its calibration and discrimination. Interval validation was conducted using bootstrap resampling.

Results

There were 4,458 (11.5%) deaths within 30-days among the 38,830 patients undergoing emergency laparotomy. Variables associated with death included (among others): age, blood pressure, heart rate, physiological variables, malignancy, and American Society of Anesthesiologists (ASA) physical status classification. The predicted risk of death among patients ranged from 1% to 50%. The model demonstrated excellent calibration and discrimination, with a C-statistic of 0.863 (95% CI: 0.858, 0.867). The model retained its high discrimination during internal validation, with a bootstrap derived C-statistic of 0.861.

Discussion

The NELA risk prediction model for emergency laparotomies discriminates well between low and high-risk patients and is suitable for producing risk-adjusted provider mortality statistics.

Editor's Key Points

- Valid and reliable risk prediction models can guide clinical practice and better inform bench-marking
- Some perioperative risk factors are modifiable, or at least alert clinical teams to the need for higher levels of care for high-risk patients
- This NELA risk model is recommended for healthcare quality evaluations for patients undergoing emergency laparomy

INTRODUCTION

Each year, approximately 33,000 patients undergo emergency laparotomy surgery in the UK[1]. Patients requiring an emergency laparotomy present with various conditions (such as perforation, ischaemia, abdominal abscess, bleeding or obstruction), and have an urgent need for clinical assessment to ensure appropriate perioperative management[2-3] . As emergency laparotomy is a common procedure with high postoperative mortality, there is potential to prevent a substantial number of deaths by benchmarking the performance of providers. But, without risk adjustment, hospital outcomes might not be comparable, and benchmarking may create unwelcome incentives including an aversion to selecting high-risk patients for surgery[4-7].

Various models are available to estimate the short-term risk of death after emergency bowel surgery, including: the Portsmouth Physiological and Operative Severity Score for the enumeration of Mortality and Morbidity (P-POSSUM) model[8-12], the Biochemistry and Haematology Outcome Model (BHOM)[13]; the Surgical Outcome Risk Tool (SORT)[14] as well as others[3,15-24] . Systematic reviews[25,26] of such models have identified substantial limitations in their design because they were often derived using small, single site studies, and/or were restricted to specific populations. This makes it difficult to draw general conclusions about their performance.

In response to the limitations of pre-existing prediction models, we undertook to develop a new model for calculating the risk-adjusted 30-day mortality of care providers performing emergency laparotomy using data on over 38,000 patients from the UK National Emergency Laparotomy Audit (NELA)[27]. The resulting model was intended for use in producing risk adjusted postoperative mortality of hospitals and/or clinical teams, and thereby support benchmarking and quality improvement.

METHODS

We used data submitted to NELA from 186 National Health Service (NHS) hospitals in England and Health Boards in Wales between 1 December 2013 and 30 November 2015. NELA was commissioned by the Healthcare Quality Improvement Partnership (HQIP) and funded by NHS England and the Welsh government, and this study was undertaken as part of the work by the Audit to evaluate the outcomes after emergency laparotomy achieved by English and Welsh NHS hospitals. Patients were eligible for inclusion in NELA if their emergency procedure involved the stomach, small or large bowel, or rectum for conditions such as perforation, ischaemia, abdominal abscess, bleeding or obstruction (see appendix A2). Procedures for appendicitis, vascular, trauma or obstetric emergencies were outside the scope of the Audit. Data collection was approved by the Confidentiality Advisory Group under section 251 of the NHS Act 2006.

The participating NHS hospitals in England and Wales submitted data on 43,566 patients. This represented approximately 70% of patients recorded in Hospital Episode Statistics[27] as having an eligible emergency laparotomy during the two year period. Patient records with missing values for one or more risk factors were removed (n=4,736), leaving 38,830 patients with complete data for inclusion in the analysis (Figure A1 in the web-supplement).

To derive 30-day all-cause postoperative mortality, patient records were linked to the Office for National Statistics (ONS) death register. For NELA patients that could not be linked to an ONS record (63 cases, 0.1%), the study used their 30-day (in-hospital) mortality available within the NELA dataset. This was considered acceptable because, among patients with dates of death in both the NELA and ONS datasets, the dates were the same for 98.6% of patients.

Selection and definition of risk factors

Potential risk factors were identified from previous reviews of existing prediction models[25], from national guidelines and from consulting with clinical experts. Decisions about their inclusion into the risk model was based on the following criteria[28] – that the risk factors: (1) were routinely measured in clinical practice, (2) were beyond the control of the provider, (3) reflect patient risk immediately before surgery and (4) were completely recorded or likely to be missing at random in the dataset.

The candidate risk factors are listed in Table 1. The factors cover basic patient demographics, pre-operative laboratory tests (creatinine, potassium, haemoglobin, white blood count (WBC) and urea), and other clinical measurements such as heart rate, systolic blood pressure, the Glasgow coma score (GCS), the American Society of Anesthesiologists physical status classification (ASA grade) and the NCEPOD (National Confidential Enquiry into Patient Outcome and Death) urgency scale[29]. Cardiac and respiratory signs were measured using the criteria defined by Copeland et al. when developing the POSSUM score[8].

*insert Table 1 here

A patient's degree of peritoneal soiling, severity of operative procedure, blood loss during surgery, and severity of malignancy were measured at the end of surgery. For patients missing these intraoperative values, we used values estimated by the clinician as part of pre-operative risk assessment, which will have been based on the surgical disgnosis and anticipated surgical findings. The proportion of patients missing intraoperative data was 0.4% for peritoneal soiling, 0.3% for operative severity, 1.0% for blood loss, and 0.4% for severity of malignancy.

Model development and statistical analysis

A multivariable logistic regression model was developed on all patients with complete data using a stepwise backward elimination process with the initial model including all variables.

The model included a random intercept term for each hospital to account for any lack of independence in the data due to patients being clustered within hospitals. All analysis was carried out on Stata® version 14 (StataCorp LP, College Station, Texas, USA).

In developing the model, it was necessary to reclassify some categorical variables because there were few patients with particular values. The variables with combined values included respiratory history, ASA grade and GCS score (see Table 1 for groups). We also found the distributions of urea and creatinine values to be highly skewed and these variables were therefore log-transformed. All continuous physiological risk factors except haemoglobin had extreme values at one or both ends of their distribution. Consequently, the distributions were winsorised at the 1st and/or 99th percentile (see web-supplement A3 for the limits).

For some continuous risk factors, their relationship with mortality was linear. When this was not the case, it was possible to capture the non-linear relationship using a linear plus quadratic term. However, this proved inadequate for serum sodium concentration, and appropriate form was determined using a fractional polynomial[30]. This process identified the equation: $sodium^3 + sodium^3 \times log[sodium]$ as a good fit for the data and indicated that mortality increased outside the range 135-145 mmol/L. Figure 1 describes these non-linear relationships.

*insert Figure 1 here

After the backward elimination variable selection process, we investigated whether the effect of some risk factors on postoperative mortality differed between levels of ASA grade. In discussions prior to the model development, this was considered clinically plausible for: age, respiratory history, cardiac signs, GCS and presence of malignancy. An interaction between systolic BP and age was also considered. The strength of these interactions was examined using non-parametric resampling with 100 bootstrap samples, and the model included those interactions which had a P value<0.01 in at least 90% of bootstrap samples.

Assessment of model performance

The performance of the model was assessed in terms of its calibration and discrimination. Calibration describes the level of agreement between the predicted and observed risks. The predicted and observed mortality was compared in deciles of predicted risk. Discrimination indicates the ability of a model to distinguish between patients with a lower and higher risk of postoperative mortality and was evaluated using the C-statistic (area under the receiver-operator charactistic curve)[32].

Internal validation and comparison with other models

Internal validation was performed using bootstrap resampling. This process involved re-fitting the model to a series of 100 random samples drawn from the dataset, and produced an overall C-statistic from all samples. The process adjusts the C-statistic for over-optimism that may arise when a model is validated with the data used to build the model[31].

The calibration and discrimination of the NELA model were compared to five models identified in the literature: P-POSSUM, CR-POSSUM, SORT, IRCS and BHOM. To ensure a fair comparison, the five models were re-calibrated to reflect the overall mortality rate in the NELA dataset, whilst retaining the weight assigned to each risk factor in the model. Re-calibration involved, first, calculating the predicted log odds of death for each patient in the NELA dataset using the published model equation. A logistic regression model was then fitted to the predicted log odds, together with an intercept term. The estimated intercept was then added to the predicted log odds to obtain a re-calibrated value.

RESULTS

Overall, 4,458 (11.5%) of the 38,830 patients undergoing emergency laparotomy died within 30 days of their surgery. There was a small difference in annual mortality across the data

collection period, being 12.0% in year 1 and 11.1% in year 2. Mortality in the analysed data was slightly higher than mortality among patients missing at least one risk factor (10.4%). Mortality among all patients (including incomplete cases) was 11.4%.

Model fitting

All potential risk factors in Table 1 were included in the final model, except haemoglobin. Interactions between ASA grade and age, and between ASA grade and respiratory signs were also included (see Figure 1, and Figure A2 in the web-supplement, for plots of relationships); both had met the selection criterion (P value <0.01) in 100% of the bootstrap samples. The heuristic shrinkage factor was estimated to be 0.992, suggesting that there is little chance for overfitting within the NELA dataset.

Table 2 reports the adjusted odds ratios for 30 day mortality for each risk factor in the model, with the effect of ASA grade reported by age and respiratory signs. As the effect of a continuous risk factor on mortality is not easily expressed when the relationship is non-linear, Table 2 shows the odds ratios for selected values of the continuous factors. The model equation is described in web-supplement A1.

*insert Table 2 here

Assessment of model performance

In the development dataset, the model demonstrated excellent discrimination, with a C-statistic of 0.863 (95% CI: 0.858, 0.867). It also had very good calibration across all levels of risk (Figure 2), the difference between the observed and predicted risk being no larger than 3% in any decile. The calibration plot also highlights the considerable heterogeneity in risk faced by patients undergoing emergency laparotomy. In the top two deciles, the observed 30-day mortality rates were 28% and 48%, respectively; in the bottom two deciles, the rates were 0.1% and 0.3%, respectively. During internal validation, the NELA model retained its

excellent discrimination across the bootstrap samples, returning an overall C-statistic of 0.861, which was close to the value from development dataset.

Comparison with other models

Table 3 reports the discrimination of the NELA model with five other models, both in terms of that achieved in their original development datasets, as well as that achieved in the NELA dataset. The P-POSSUM and SORT models both had a C-statistic of 0.81 in the NELA dataset. The BHOM had the poorest discrimination, with a C-statistic of less than 0.6.

*insert Table 3 here

The calibration plots for the NELA, P-POSSUM, CR-POSSUM, SORT, BHOM & IRCS models are shown in Figure 3. The SORT and P-POSSUM models predicted a similar range of patient risk as the NELA model, hence their relatively high discrimination. Indeed, the top deciles of risk for the BHOM and IRCS models only extended to around 30%. Calibration within the deciles of predicted risk was found to be poorer than the NELA model for all of these models except SORT. P-POSSUM and CR-POSSUM were both observed to under-predict risk in patients with moderate to high risk and to over-predict risk for patients in the highest decile. The BHOM and IRCS models both showed a lack of calibration throughout the deciles of predicted risk.

*insert Figure 3 here

DISCUSSION

The NELA risk model of postoperative mortality after emergency laparotomy was developed to support provider benchmarking by enabling the production of risk-adjusted 30-day mortality rates. It incorporated risk factors that are routinely collected in clinical practice and was derived using a large, comtemporary population-based dataset. The model had very good calibration and excellent discrimination in the development dataset, and retained its

performance during internal validation. In contrast with other models, it also avoids categorising the continuous risk factors and allows for non-linear relationships. Interactions between key risk factors were included when supported by robust evidence. We recommend this model be used to adjust short-term postoperative mortality rates when comparing hospitals and/or clinical teams that undertake emergency laparotomy.

Comparison with previous studies

Various risk models are available to estimate short-term postoperative mortality in patients undergoing an emergency laparotomy. A systematic review[25] which reviewed research published before April 2013 found that the largest previous study to develop a prediction model included 37,553 patients across 142 sites in the USA[21]. The model was based on the ACS NSQIP (The American College of Surgeons National Surgical Quality Improvement Program) and was developed using statistical approaches that may produce a suboptimal model, such as the categorisation of continuous variables and automated variable selection methods. The final model included 37 risk factors and its internal discrimination was high (C-statistic = 0.87). However, there has yet to be an external validation of the model to show whether it retain this level of performance in other situations. We were not able to compare the NSQIP model to the NELA model as it required many risk factors that are not collected by NELA, such as body mass index (BMI) and smoking status.

The comparison of the NELA model to the five other predication models found that it outperformed them all in terms of discrimination. In addition, all models except SORT were observed to have worse calibration. This might reflect the fact that we were only able to evaluate the NELA model using the data on which it was developed, although its performance changed little during internal validation. However, Table 3 demonstrates that, during external validation, performance tends to decrease. Consequently, an external validation of the NELA model would be desirable. Another reason for the poorer performance of the five published models could be their development in smaller patient

cohorts. SORT[14] was developed on the largest cohort (16,788 patients), but only a fraction of these patients had an emergency laparotomy. Among the five models, P-POSSUM is the most widely used tool for risk assessment in clinical practice in emergency laparotomy, but its original equation has proven to be poorly calibrated in contemporary populations, particularly in higher risk patients. However, after re-calibration, we found that it still performed reasonably well.

Implications

The NELA model was developed to enable the production of risk-adjusted hospital-level postoperative mortality. It was not designed for use within a clinical setting to predict individual patient risk. The variables selected into the model reflect this design aim and therefore risk factors which could improve the prediction of individual patient risk may not be included. A model to predict individual patient risk should be based only on information available before surgery, and the perioperative variables used in this risk-adjustment model were, in some case, only available postoperatively.

Several features of this model are worth highlighting in relation to the association between mortality risk and individual risk factors. First, there were non-linear relationships between mortality and several continuous risk factors. U-shaped relationships were identified for potassium and creatinine, demonstrating that mortality is higher in those patients outside the normal range, a finding consistent with previously published analyses [33]. Second, we observed that the association between mortality and some other risk factors differed by ASA grade. For example, we found that the impact of a high ASA grade was particularly marked in younger patients, whereas older patients were at a relatively high risk of death across all ASA categories. This suggests that it is worth investigating whether particular patient groups might be helped by individualised care including augmented pathways and levels of support and a shared approach to decision-making.

Strengths and limitations of the study

The main strength of this study is the large sample size from a national population. With case ascertainment at 65% and 70% for year 1 and 2 of the audit respectively [27], we are confident that the dataset is representative of patients within England and Wales that underwent an emergency laparotomy in an NHS hospital, especially as the mortality in patients not captured in NELA was similar to that of the patients captured (results not shown).

Oher strengths include (1) the linkage of NELA records with ONS mortality data which allowed us to reliably capture all deaths (in or out of hospital) and so have complete follow-up, and (2) the richness of this dataset due to the large number of routinely collected clinical variables and the small proportion of missing data items. This enabled the model to include risk factors that were not measured in previous studies.

One limitation of this study is the development of the risk adjustment model excluding patients with missing data on 1 or more risk factors. The distribution of known values across the risk factors were similar in paients with complete and incomplete data, which suggested the data were missing at random, but we noted that postoperative mortality was a little lower in the patients with missing data. However, only 11% of patients were missing any risk factors and excluding these patients had a minimal effect on the overall mortality. Another limitation is the definitions used for some comorbidities. These were chosen based on the definitions used in the initial description of the POSSUM model in 1991. Alternative methods exist for how some comorbidities are classified or described (e.g. the New York Heart Association (NYHA) classification for heart failure). Comparison of how comorbidities are defined, and consideration of how these might be updated, could add improved discrimination to future models and could be considered in future iterations of NELA and other observational studies of major surgery.

Conclusion

Emergency laparotomy is associated with a high rate of mortality and morbidity, and comparative benchmarking has the potential to greatly improve outcomes for patients. The NELA model has demonstrated excellent performance in predicting short-term postoperative mortality and will enable fair comparisons to be made between providers of emergency laparotomy. We expect the NELA model to retain its performance when it is applied to data collected in other settings because it was developed in a large, population-based dataset with a robust process of model development (eg, almost all decisions about the model building decided a priori). The performance of the model is therefore likely to compare very favourably with other models when validated using external data.

Contributors

Initiated the project: NE, CMO, MGB, AK, AID, SRM, MPG, DMM, DAC, KW; planning of statistical analysis: NE, CMO, MGB, AID, SRM, MPG, DMM, DAC, KW; cleaning and analysis of data: NE, KW; Interpretation of results: NE, CMO, MGB, TEP, AK, JC, IDA, SRM, MPG, DMM, DAC, KW; Drafted initial paper: NE, KW; Revised paper: NE, CMO, MGB, TEP, AK, JC, IDA, SRM, MPG, DMM, DAC, KW. KW is guarantor.

Acknowledgements

We are grateful to the NHS hospital staff across England and Wales who submitted data to the Audit since 2013.

The NELA Project Team during the period of data collection also comprised: Mr Martin Cripps, Mr Paul Cripps, Mrs Emma Davies, Ms Sharon Drake, Dr Mike Galsworthy, Mr James Goodwin, Dr Tom Salih, Mr Jose Lourtie, Mr Dimitri Papadimitriou, Dr Carol Peden.

Source of funding

This study was undertaken as part of the work by the National Emergency Laparotomy Audit to evaluate the clinical outcomes achieved by English and Welsh NHS hospitals. The Audit is commissioned by the Healthcare Quality Improvement Partnership (HQIP) as part of the National Clinical Audit Programme (NCAPOP). The funder has no responsibility or liability for the accuracy, currency, reliability and/or correctness of this publication, and did not play a role in the analysis and interpretation of data, writing of the manuscript, or the decision to submit

the manuscript for publication.

Ethics approval

The study is exempt from UK National Research Ethics Committee approval as it involved data collected for the purposes of clinical audit. ONS date of death were made available by

Conflict of interests

IDA is the Vice-President of the Association of surgeons of Great Britain and Ireland.

DMM's Trust is reimbursed for the time commitment to NELA and has received PAs as the chair of the NELA Project Team since 2017, and NELA National Clinical Lead 2012-2017

MPG received PAs as the chair of the NELA Project Team. MPG is the National Specialty Lead for Anaesthesia, Perioperative Medicine and Pain within the UK National Institute of Heath Research Clinical Research Network, an elected council member of the Royal College of Anaesthetists and President of the Critical Care Medicine section of the Royal Society of Medicine. MPWG serves on the board of ERAS UK, Oxygen Control Systems Ltd, the Evidence Based Perioperative Medicine (EBPOM) social enterprise as well as the medical advisory board of Sphere Medical Ltd and the international advisory board of the American Society of Enhanced Recovery (ASER). MG has received honoraria for speaking and/or travel expenses from Edwards Lifesciences, Fresenius-Kabi, BOC Medical (Linde Group), Ely-Lilly Critical Care, and Cortex GmBH. MG is executive chair of the Xtreme-Everest Oxygen Research Consortium.is a medical adviser for Sphere Medical Ltd and Director of Oxygen Control Systems Ltd and received an honorarium and travel expenses from Edwards Lifesciences in 2016.

SRM is the Associate National Clinical Director for elective care, NHS England.

References

1. RCS & ASGBI, 2013. https://www.rcseng.ac.uk/news/emergency-general-surgery-rcs-and-asgbi-position-statement#.V4ycLU0UWUl *(accessed 21/06/2017).*
2. Saunders DI, Murray D, Pichel AC, Varley S, Peden CJ, members of the UK Emergency Laparotomy Network. Variations in mortality after emergency laparotomy: the first report of the UK Emergency Laparotomy Network. British journal of anaesthesia. 2012; 109(3):368-75.
3. Green G, Shaikh I, Fernandes R, Wegstapel H. Emergency laparotomy in octogenarians: A 5-year study of morbidity and mortality. World journal of gastrointestinal surgery. 2013; 5(7):216.
4. Moonesinghe SR. Individualised surgical outcomes: please look the other way. 2013
5. Burns EM, Pettengell C, Athanasiou T, Darzi A. Understanding the strengths and weaknesses of public reporting of surgeon-specific outcome data. Health Affairs. 2016;35(3):415-21.
6. Alderson D, Cromwell D. Publication of surgeon-specific outcomes. British Journal of Surgery. 2014; 101(11):1335-7.
7. Werner RM, Asch DA. The unintended consequences of publicly reporting quality information. JAMA 2005; 293(10):1239-44.
8. Copeland GP, Jones D, Walters MP. POSSUM: a scoring system for surgical audit. British Journal of Surgery. 1991; 78(3):355-60.
9. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. British Journal of Surgery. 1998; 85(9):1217-20.
10. Tekkis PP, Prytherch DR, Kocher HM, Senapati A, Poloniecki JD, Stamatakis JD, Windsor AC. Development of a dedicated risk-adjustment scoring system for

colorectal surgery (colorectal POSSUM). British Journal of Surgery. 2004; 91(9):1174-82.

11. Sreeharsha H, Sp R, Sreekar H, Reddy R. Efficacy of POSSUM score in predicting the outcome in patients undergoing emergency laparotomy. Polish Journal of Surgery. 2014; 86(4):159-65.
12. Leung E, McArdle K, Wong LS. Risk-adjusted scoring systems in colorectal surgery. International Journal of Surgery. 2011; 9(2):130-5.
13. Prytherch DR, Sirl JS, Weaver PC, Schmidt P, Higgins B, Sutton GL. Towards a national clinical minimum data set for general surgery. British Journal of Surgery. 2003; 90(10):1300-5.
14. Protopapa KL, Simpson JC, Smith NC, Moonesinghe SR. Development and validation of the surgical outcome risk tool (SORT). British Journal of Surgery. 2014; 101(13):1774-83.
15. Sluis FJ, Espin E, Vallribera F, Bock GH, Hoekstra HJ, Leeuwen BL, Engel AF. Predicting postoperative mortality after colorectal surgery: a novel clinical model. Colorectal Disease. 2014; 16(8):631-9.
16. Zerbib P, Kulick JF, Lebuffe G, Khoury-Helou A, Plenier I, Chambon JP. Emergency major abdominal surgery in patients over 85 years of age. World Journal of Surgery. 2005; 29(7):820-5.
17. Skala K, Gervaz P, Buchs N, Inan I, Secic M, Mugnier-Konrad B, Morel P. Risk factors for mortality–morbidity after emergency–urgent colorectal surgery. International journal of colorectal disease. 2009; 24(3):311-6.
18. Garcea G, Ganga R, Neal CP, Ong SL, Dennison AR, Berry DP. Preoperative early warning scores can predict in-hospital mortality and critical care admission following emergency surgery. Journal of Surgical Research. 2010; 159(2):729-34.
19. Aslar AK, Özdemir S, Mahmoudi H, Kuzu MA. Analysis of 230 cases of emergent surgery for obstructing colon cancer—lessons learned. Journal of Gastrointestinal Surgery. 2011; 15(1):110-9.

20. Iversen LH, Bülow S, Christensen IJ, Laurberg S, Harling H. Postoperative medical complications are the main cause of early death after emergency surgery for colonic cancer. British Journal of Surgery. 2008; 95(8):1012-9.

21. 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. Journal of the American College of Surgeons. 2012; 215(4):503-11.

22. Kwok AC, Lipsitz SR, Bader AM, Gawande AA. Are targeted preoperative risk prediction tools more powerful? A test of models for emergency colon surgery in the very elderly. Journal of the American College of Surgeons. 2011; 213(2):220-5.

23. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Critical care medicine. 1985; 13(10):818-29.

24. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, Cohen ME. 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. 2013; 217(5):833-42.

25. Oliver CM, Walker E, Giannaris S, Grocott MP, Moonesinghe SR. Risk assessment tools validated for patients undergoing emergency laparotomy: a systematic review. BJA: British Journal of Anaesthesia. 2015; 115(6):849-60.

26. Stonelake S, Thomson P, Suggett N. Identification of the high risk emergency surgical patient: Which risk prediction model should be used? Annals of Medicine and Surgery. 2015; 4(3):240-7.

27. *NELA project team.* First patient report of the National Emergency Laparotomy Audit. *London: The Royal College of Anaesthetists, 2017.* http://www.nela.org.uk/reports *(accessed 21/06/2017).*

28. Walker K, Finan PJ, van der Meulen JH. Model for risk adjustment of postoperative mortality in patients with colorectal cancer. British Journal of Surgery. 2015; 102(3):269-80.

29. *NCEPOD, 2004 http://www.ncepod.org.uk/classification.html (accessed 21/06/2017).*

30. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. Statistics in Medicine. 2007; 26(30):5512-28.

31. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology. 2001; 54(8):774-81.

32. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988 :837-45.

33. Mohammed MA, Rudge G, Wood G, Smith G, Nangalia V, Prytherch D, Holder R, Briggs J. Which is more useful in predicting hospital mortality-dichotomised blood test results or actual test values? a retrospective study in two hospitals. PloS one. 2012; 7(10):e46860.

Table 1: Patient characteristics and associated unadjusted 30 day mortality rates

Risk factor	Number (%)	Mortality	Risk factor	Number (%)	Mortality
Age (years)*			Gender		
18-39	4,116 (11)	2.3	Male	18,740 (48)	11.6
40-49	3,697 (9)	3.1	Female	20,090 (52)	11.4
50-59	5,308 (14)	6.0			
60-69	8,074 (21)	9.9	Year of NELA audit		
70-79	9,795 (25)	15.3	Year 1 (1st Dec 2013- 30th Nov 2014)	16,897 (43)	12.0
80-89	6,885 (18)	20.2	Year 2 (1st Dec 2014- 30th Nov 2015)	21,933 (57)	11.1
90+	955 (2)	24.2			
Preoperative					
ECG*			Haemoglobin (g/l)*		
No abnormalities	31,013 (80)	8.4	Low (male<130/ female<115)	16,129 (42)	14.3
AF rate 60-90	1,613 (4)	18.2	Normal (male 130-180/ female 115-165)	21,793 (56)	9.2
AF rate >90/ any other abnormal rhythm	6,204 (16)	25.0	High (male>180/female>165)	908 (2)	14.9
Cardiac signs*			Urea (mmol/l)*		
No failure	28,358 (73)	8.2	Low (<2.5)	1,647 (4)	4.4
Diuretic, digoxin, antihypertensive therapy	8,115 (21)	17.6	Normal (2.5-7.8)	22,826 (59)	6.5
Peripheral oedema, warfarin therapy or CXR	1,913 (5)	28.9	High (>7.8)	14,357 (37)	20.2
Raised jugular venous pressure or CXR	492 (2)	33.9			
Systolic BP (mmHg)*			White Blood Cell ($x10^9$/l)*		
Low (<90)	1,747 (5)	35.2	Low (<3.6)	1,290 (3)	21.7
Normal (90-120)	15,302 (39)	13.8	Normal (3.6-11.0)	17,903 (46)	9.7
High (>120)	21,781 (56)	8.0	High (>11.0)	19,637 (51)	12.5
Pulse (bpm)*			Creatinine (umol/l)*		
Low (<60)	854 (2)	6.6	Low (male <59/ female <45)	4,079 (11)	10.1
Normal (60-100)	27,602 (71)	8.9	Normal (male 59-104 / female 45-84)	23,004 (59)	6.7
High (>100)	10,374 (27)	18.9	High (male >104/ female >84)	11,747 (30)	21.3

Risk factor	Number (%)	Mortality	Risk factor	Number (%)	Mortality
Respiratory history*			Sodium (mmol/l)*		
No dyspnoea	27,988 (72)	7.4	Low (<133)	6,490 (17)	16.3
Dyspnoea on exertion or CXR	6,210 (16)	17.7	Normal (133-146)	31,783 (82)	10.2
Dyspnoea limiting exertion & at rest	4,632 (12)	28.2	High (>146)	557 (1)	28.6
Glasgow Coma Score*			Potassium (mmol/l)*		
Minor (13-15)	37,705 (97)	10.4	Low (<3.5)	4,363 (11)	13.4
Moderate (9-12)	332 (1)	43.4	Normal (3.5-5.3)	32,910 (85)	10.3
Severe (3-8)	793 (2)	48.3	High (>5.3)	1,557 (4)	30.6
ASA Score*			Urgency of surgery*		
1 & 2 (None or mild systemic disease)	17,190 (44)	2.6	Expedited (>18hrs)	6,405 (17)	6.8
3 (Severe disease, not life-threatening)	13,706 (35)	9.9	Urgent (6-18hrs)	11,735 (30)	6.9
4 (Severe, life-threatening disease)	7,123 (19)	30.8	Urgent (2-6hrs)	15,051 (39)	11.6
5 (Moribund patient)	811 (2)	58.8	Immediate (<2hrs)	5,639 (14)	26.0
Number of operations within this Admission *					
1	33,584 (87)	11.3			
2	4,815 (12)	11.9			
>2	431 (1)	18.1			
Peri-operative					
Operative severity*			Intra-operative blood loss *		
Major	24,453 (63)	9.6	<100 ml	18,380 (47)	9.7
Major+	14,377 (37)	14.7	101-500 ml	17,463 (45)	12.4
			501-999 ml	2,001 (5)	15.6
			≥1000 ml	986 (3)	19.8
Peritoneal Soiling*			Severity of malignancy*		
None	14,537 (37)	8.2	None	29,774 (77)	10.9
Serous fluid	9,992 (26)	11.9	Primary only	4,496 (12)	9.8
Localised pus	4,183 (11)	7.4	Nodal metastases	1,655 (4)	11.9
Free bowel content, pus or blood	10,118 (26)	17.5	Distant metastases	2,905 (7)	20.2

*Chi^2 p value <0.05

Table 2: Odds ratios (OR) for the variables included in the NELA model

Risk factor	OR	95% CI	Risk factor	OR	95% CI	Risk factor	OR	95% CI
NELA year1	1		ECG (no abnormalities)	1		WBC 5x10^9/l	1.06	1.01 to 1.12
NELA year2	0.96	0.90 to 1.03	ECG (AF rate 60-90)	1.22	1.06 to 1.41	WBC 10 x10^9/l	1	
Male	1		ECG (AF rate >90 or abnl)	1.21	1.11 to 1.31	WBC 20 x10^9/l	1.02	0.97 to 1.08
Female	1.04	0.97 to 1.12	Peritoneal soiling (none)	1		WBC 30 x10^9/l	1.26	1.16 to 1.38
Blood loss <100ml	1		Peritoneal soiling (serous fluid)	1.20	1.09 to 1.31	WBC 40 x10^9/l	1.89	1.59 to 2.24
Blood loss (101-500ml)	1.02	0.94 to 1.10	Peritoneal soiling (localised pus)	1.01	0.87 to 1.16	Urea 2 mmol/l	0.58	0.47 to 0.70
Blood loss (501-999ml)	1.04	0.89 to 1.20	Peritoneal soiling (free bowel)	1.41	1.28 to 1.55	Urea 5 mmol/l	0.80	0.76 to 0.85
Blood loss (>1,000ml)	0.85	0.70 to 1.04	Surgical urgency (>18hrs)	1		Urea 10 mmol/l	1	
No cardiac failure	1		Surgical urgency (6-18hrs)	0.91	0.80 to 1.04	Urea 20 mmol/l	1.21	1.13 to 1.29
Antihypertensive therapy	1.07	0.98 to 1.16	Surgical urgency (2-6hrs)	1.04	0.91 to 1.18	Urea 30 mmol/l	1.33	1.17 to 1.52
Borderline cardiomegaly	1.33	1.17 to 1.51	Surgical urgency (<2hrs)	1.58	1.37 to 1.82	Creatinine 40umol/l	1.16	1.03 to 1.31
Cardiomegaly	1.22	0.99 to 1.52				Creatinine 70umol/l	1.02	0.95 to 1.09
Glasgow score (13-15)	1		Sodium 125mmol/l	1.53	1.37 to 1.71	Creatinine 100umol/l	1	
Glasgow score (9-12)	1.85	1.44 to 2.38	Sodium 130 mmol/l	1.38	1.26 to 1.51	Creatinine 150umol/l	1.04	1.01 to 1.08
Glasgow score (3-8)	2.44	2.06 to 2.90	Sodium 140 mmol/l	1		Potassium 3mmol/l	1.36	1.23 to 1.51
Malignancy (none)	1		Sodium 150 mmol/l	2.99	2.20 to 4.07	Potassium 3.5 mmol/l	1.11	1.06 to 1.15
Malignancy (primary)	1.10	0.98 to 1.24	Systolic BP 80	1.75	1.57 to 1.94	Potassium 4 mmol/l	1	
Malignancy (nodal mets)	1.54	1.30 to 1.82	Systolic BP 100mmHg	1.26	1.22 to 1.32	Potassium 4.5 mmol/l	1.01	0.98 to 1.04
Malignancy (distant)	3.16	2.83 to 3.54	Systolic BP 120 mmHg	1		Potassium 5 mmol/l	1.14	1.07 to 1.21
Number procedures (1)	1		Systolic BP 150 mmHg	0.83	0.79 to 0.87	Pulse 60bpm	0.64	0.56 to 0.72
Number procedures (2)	0.78	0.70 to 0.87	Systolic BP 180 mmHg	0.83	0.72 to 0.96	Pulse 70bpm	0.76	0.71 to 0.81
Number procedures (>2)	0.75	0.56 to 0.99				Pulse 90bpm	1	
Operative severity (Major)	1					Pulse 120bpm	1.30	1.23 to 1.38
Operative (Major+)	1.17	1.09 to 1.26				Pulse 140bpm	1.40	1.22 to 1.61

Table 2: Odds ratios (OR) for the variables included in the NELA model (continued)

Model Interaction terms

	ASA 1 or 2		ASA 3		ASA 4		ASA 5	
Risk factor	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
ASA (no respiratory history and age 70)	1		2.52	2.12 to 3.00	6.28	5.25 to 7.51	12.45	9.21 to 16.83
Age 50	0.48	0.42 to 0.54	0.59	0.52 to 0.67	0.70	0.62 to 0.80	0.77	0.68 to 0.88
Age 60	0.80	0.78 to 0.82	0.86	0.83 to 0.88	0.89	0.87 to 0.92	0.91	0.88 to 0.93
Age 70 (ref)	1		1		1		1	
Age 80	2.73	2.39 to 3.13	1.95	1.71 to 2.23	1.74	1.52 to 1.99	1.81	1.58 to 2.07
Age 90	5.59	4.23 to 7.38	3.08	2.33 to 4.07	2.68	2.03 to 3.54	3.09	2.34 to 4.08
No respiratory history (ref)	1		1		1		1	
Mild dyspnoea	1.97	1.53 to 2.53	1.37	1.20 to 1.56	1.22	1.06 to 1.39	1.03	0.70 to 1.52
limiting & at rest	3.73	2.51 to 5.53	1.90	1.63 to 2.20	1.48	1.31 to 1.68	1.33	0.95 to 1.86

Table 3: Discrimination of the NELA risk model compared to published models

Model	C-Statistic in development study / other external validation	Sample size in original study	C- statistic within NELA dataset
NELA	Not applicable	N/A	0.863 (0.858, 0.867)
P-POSSUM	External validation: 0.90 [35]	10,000	0.808 (0.802, 0.815)
CR-POSSUM	Internal validation: 0.898 [10]	2,691	0.771 (0.765, 0.778)
SORT	Internal validation: 0.91 [14]	5,569	0.814 (0.808, 0.821)
IRCS	External validation: 0.83 [15]	1,252	0.695 (0.687, 0.702)
BHOM	External validation: 0.841[35]	12,259	0.578 (0.569, 0.587)

Figure 1: Model fit for continuous risk factors that had a non-linear relationship with 30-day postoperative mortality.

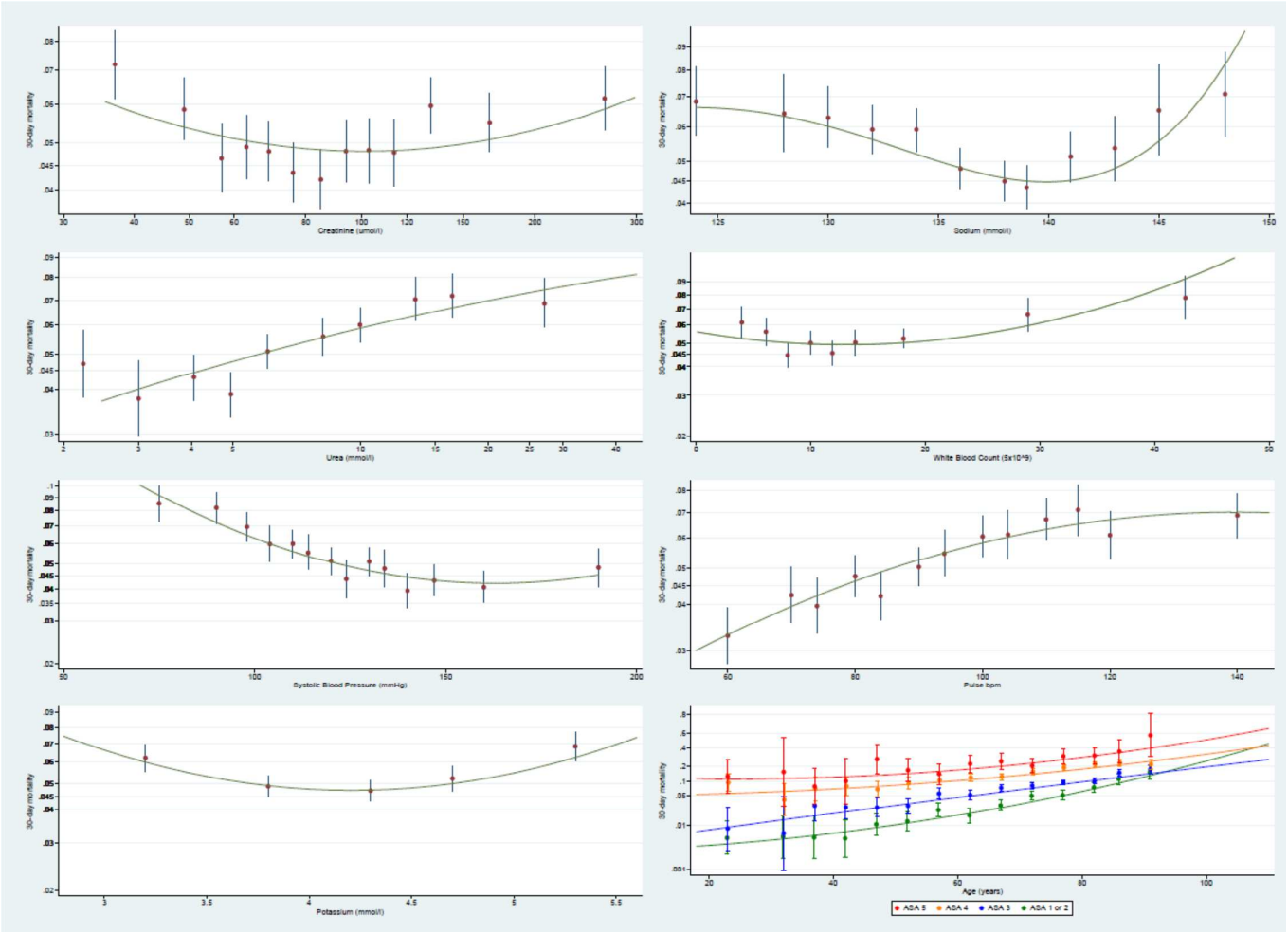

Figure 2: Calibration plot comparing the observed 30-day mortality against the predicted mortality in deciles of predicted risk from the NELA risk adjustment model

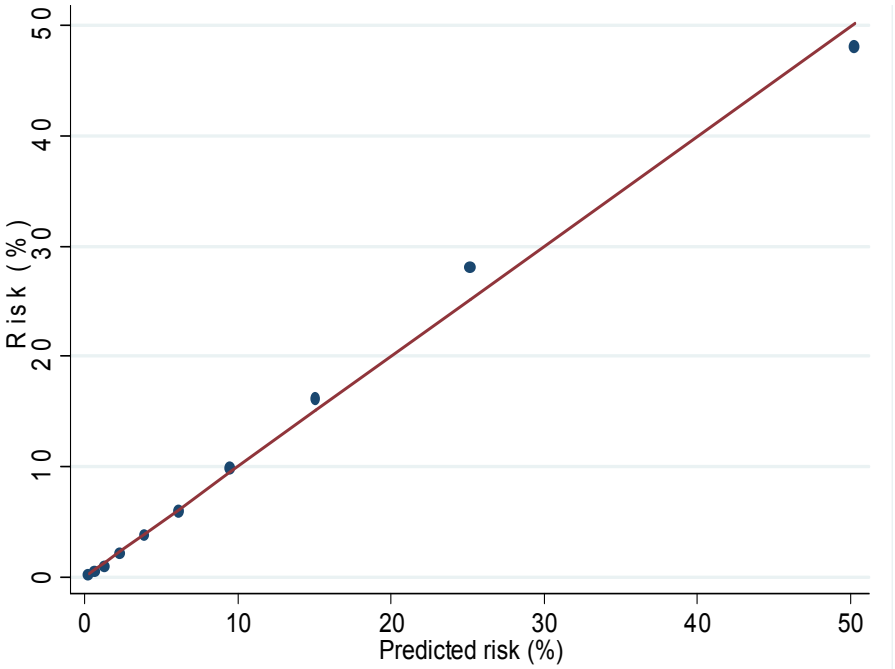

Figure 3: Calibration plot comparing the observed 30-day mortality against that predicted from the various models in deciles of predicted risk for P-POSSUM, CR-POSSUM, SORT, BHOM, IRCS and NELA risk models.