SUPPLEMENTARY MATERIAL

**Prospective multicentre external validation of postoperative mortality prediction tools in patients undergoing emergency laparotomy**

|  |  |
| --- | --- |
| Contents | Page |
| Table S1. Inclusion and Exclusion criteria…………………………………………………………………………….. | 2 |
| Table S2. Adjustment of the ACS-NSQIP calculator according to the surgeon’s preoperative assessment………………………………………………………………………………………………………………………….. | 3 |
| Table S3. Comparison of the distribution of important variables Hospitals with development data of the NELA and ACS-NSQIP…………………………………………………………………… | 4 |
| Figure S1. Receiver Operating Characteristic (ROC) Curve showing the discriminating performance of the models when predicting 30-day post-operative death in emergency laparotomies……………………………………………………………………………………………………………………….. | 6 |
| Figure S2. Calibration of prognostic models when predicting 30-day postoperative death….. | 7 |
| Table S4. Predictive performance measures of prognostic models for 30-day postoperative death after updating of calibration intercept and slope………………………………………………………. | 8 |
| Figure S3. Calibration of prognostic models when predicting 30-day postoperative death after updating of intercept and slope (recalibration)…………………………………………………………… | 9 |
| Figure S4. Decision curves showing the net benefit in clinical decision-making of using each prognostic model of 30-day postoperative mortality after updating of intercept and slope (recalibration)……………………………………………………………………………………………………………………… | 10 |
| Figure S5. Forest plot with hospital-specific Brier scores of the Surgeon's assessment prognostic model and overall pooled Brier score based on random-effects meta-analysis….. | 11 |
| Figure S6. Forest plot with hospital-specific Brier scores of the NELA prognostic model and overall pooled Brier score based on random-effects meta-analysis…………………………………….. | 12 |
| Figure S7. Forest plot with hospital-specific Brier scores of the P-POSSUM prognostic model and overall pooled Brier score based on random-effects meta-analysis……………………. | 13 |
| Figure S8. Forest plot with hospital-specific Brier scores of the POTTER prognostic model and overall pooled Brier score based on random-effects meta-analysis………………………………. | 14 |
| Figure S9. Forest plot with hospital-specific Brier scores of the ACS-NSQIP prognostic model and overall pooled Brier score based on random-effects meta-analysis……………………. | 15 |
| Figure S10. Forest plot with hospital-specific Brier scores of the ACS-NSQIP adjusted prognostic model and overall pooled Brier score based on random-effects meta-analysis….. | 16 |
| Table S5: Checklist for transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD)………………………………………………………………………….. | 17 |

Table S1: Inclusion and Exclusion criteria

|  |
| --- |
| **Inclusion criteria:** |
| * Age >18yrs
* Emergency laparotomy (operation simultaneously with resuscitation usually within one hour) or
* urgent (operation as soon as possible after resuscitation, within 24hrs)
* Operation in the gastrointestinal tract:
* Open or laparoscopic, or laparoscopically assisted procedures.
* Procedures involving the stomach, small or large bowel, or rectum for conditions such as
* perforation, ischaemia, abdominal abscess, bleeding or obstruction
* Wash out/evacuation of intraperitoneal abscess or haematoma
* Bowel resection/repair due to incarcerated/incisional hernias
* Bowel resection or repair due to incarcerated umbilical, inguinal or femoral hernias
* Open or laparoscopic adhesiolysis
* Laparotomy/laparoscopy with inoperable pathology
* Return to theatre for repair of a substantial dehiscence of major abdominal wound (i.e.“burst abdomen”)
* Return to theatre after any operation (including vascular, gynaecology, urology, cardiac) meeting the criteria above
* In the case of multiple procedures in the abdominopelvic cavity the patient is included if the main procedure is a general surgical one (i.e. if bowel resection happens during an open aneurysm repair it should not be included)
* Any intra-abdominal procedure not identifiable within exclusion criteria should be included.
 |
| **Exclusion criteria:** |
| * Patients under 18
* Elective operation
* Diagnostic laparoscopy or laparotomy where no other procedure is performed (NB, if no procedure is performed due to inoperable pathology, then include)
* Appendicectomy with or without drainage of localized abscess
* Cholecystectomy with or without drainage of localized abscess
* Hernia repair without bowel resection
* Minor abdominal wound revision
* Vascular surgery
* Gynaecological surgery – c-section – ruptured ectopic pregnancy
* Surgery relating to organ transplantation
 |

Table S2: Adjustment of the ACS-NSQIP calculator according to the surgeon’s preoperative assessment

|  |  |  |
| --- | --- | --- |
| **ACS-NSQIP predicted probability of 30-day mortality** | **Surgeon’s preoperative assessment of mortality risk** | **Chosen option of adjustment (for underestimation) on the ACS-NSQIP online calculation** |
| <0.05 | <5% | 1 – No adjustment necessary |
| <0.05 | 5 - 10% | 2 – Risk somewhat higher then estimate |
| <0.05 | 11 - 20% | 3 – Risk significantly higher than estimate |
| <0.05 | >20% | 3 – Risk significantly higher than estimate |
| 0.05 – 0.10 | <5% | 1 – No adjustment necessary |
| 0.05 – 0.10 | 5 - 10% | 1 – No adjustment necessary |
| 0.05 – 0.10 | 11 - 20% | 2 – Risk somewhat higher then estimate |
| 0.05 – 0.10 | >20% | 3 – Risk significantly higher than estimate |
| 0.10 – 0.20 | <5% | 1 – No adjustment necessary |
| 0.10 – 0.20 | 5 - 10% | 1 – No adjustment necessary |
| 0.10 – 0.20 | 11 - 20% | 1 – No adjustment necessary |
| 0.10 – 0.20 | >20% | 2 – Risk somewhat higher then estimate |
| >0.20 | Any | 1 – No adjustment necessary |

**Table S3**. Comparison of the distribution of important variables of 631 patients undergoing emergency laparotomy in 11 Greek Hospitals with development data of the NELA and ACS-NSQIP risk prediction models.

|  |  |  |  |
| --- | --- | --- | --- |
| **Characteristic** | **Greek cohort, n(%)** | **NELA, n(%)** | **ACS-NSQIP, n(%)** |
| Age group |  |  |  |
| ≤60 | 201 (32) | 13121 (34) | - |
| >60 | 430 (68) | 25709 (66) | - |
| Male Sex | 341 (54) | 18 740 (48) | 604016 (43) |
| ASA class |  |  |  |
| 1-2 | 351 (56) | 17190 (44) | 777115 (55) |
| 3 | 171 (27) | 13706 (35) | 541404 (38) |
| 4-5 | 110 (17) | 7934 (21) | 95487 (7) |
| Steroid use (1) | 55 (9) | - | 43296 (3) |
| Ascites (2) | 81 (13) | - | 8697 (1) |
| Preoperative functional status |  |  |  |
| Independent | 444 (70) | - | 1344929 (95) |
| Partially dependent | 154 (25) | - | 52500 (4) |
| Totally dependent | 33 (5) | - | 16577 (1) |
| Sepsis (3) |  |  |  |
| SIRS | 218 (34) | - | 55090 (4) |
| Sepsis | 81 (13) | - | 33725 (2) |
| Septic shock | 25 (4) | - | 8546 (1) |
| Ventilator dependent | 16 (3) | - | 10119 (1) |
| Smoking (4) | 186 (29) | - | 272322 (19) |
| Dyspnoea |  |  |  |
| Moderate exertion | 36 (6) | 6210 (16) | 110720 (8) |
| At rest | 19 (3) | 4632 (12) | 15571 (1) |
| Haemodialysis or CVVH | 10 (2) | - | 22829 (2) |
| Acute renal failure | 77 (12) | - | 7103 (1) |
| Anticipated severity of malignancy |  |  |  |
| None | 462 (73) | 29774 (77) | - |
| Primary | 78 (12) | 4496 (12) | - |
| Nodal metastasis | 19 (3) | 1655 (4) | - |
| Distant metastasis | 73 (12) | 2905 (7) |  28173 (2) (5) |
| Diabetes mellitus | 103 (16) | - | 215180 (15) |
| Urgency of operation |  |  |  |
| Expedited (>18 hours) | 71 (11) | 6405 (17) | - |
| Urgent (6-18 hours) | 179 (28) | 11735 (30) | - |
| Urgent (2-6 hours) | 206 (33) | 15051 (39) | - |
| Immediate (<2 hours) | 177 (28) | 5639 (14) | - |
| Operative severity |  |  |  |
| Major | 388 (62) | 24453 (63) | - |
| Major+ | 243 (38) | 14377 (37) | - |
| Peritoneal soiling |  |  |  |
| None | 171 (27) | 14537 (37) | - |
| Serous fluid | 234 (37) | 9992 (26) | - |
| Localised pus | 54 (9) | 4183 (11) | - |
| Free bowel content, pus, or blood | 172 (27) | 10118 (26) | - |
| Intraoperative blood loss |  |  |  |
| <100ml | 357 (56) | 18380 (47) | - |
| 101-500ml | 237 (38) | 17463 (45) | - |
| 501-999 ml | 25 (4) | 2001 (5) | - |
| >1000ml | 12 (2) | 986 (3) | - |
| 30-day mortality | 103 (16) | 4458 (12) | 18,909 (1) |

NA, not applicable; ASA, American Society of Anesthesiologists; CVVH, continuous veno-venous hemofiltration; BMI, body mass index. Dashes (-) imply that data for these variables were not available in the original model development publications.

Notes.

1. Steroid use for chronic condition
2. Ascites within 30 days prior to surgery
3. Systemic sepsis within 48h from surgery
4. Smoking within 12 months from surgery
5. Disseminated cancer

Figure S1. Receiver Operating Characteristic (ROC) Curve showing the discriminating performance of the models when predicting 30-day post-operative death in emergency laparotomies



Note: For NELA, P-POSSUM, POTTER, ACS-NSQIP and ACS-NSQIP-adjusted, the ROC curves were plotted by calculating the sensitivity and specificity for all values ranging from 0 to 1, to construct a curve, and for the surgeon’s prediction, they were calculated for each of the four categories, and the points were combined to form the curve.

Figure S2. Calibration of prognostic models when predicting 30-day postoperative death.



## Table S4. Predictive performance measures of prognostic models for 30-day postoperative death after updating of calibration intercept and slope

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Prognostic model** | **N** | **Overall fit** | **Discrimination** | **Calibration** | **Clinical utility** |
| Brier scaled % (95% CI) | AUC (95% CI) | DeLong p-value | E:O ratio | CITL (95% CI) | Slope (95% CI) | HL-GOF p-value | NB, 5% | NB, 10% | NB, 20% |
| Surgeon | 610 | 10.6 ( 1.3, 18.7) | 0.79 (0.75, 0.82) | Ref. | 0.91 | 0.16 (-0.09, 0.41) | 0.74 (0.57, 0.91) | <0.001 | 0.12 | 0.10 | 0.04 |
| NELA RCM | 623 | 22.2 (13.4, 29.5) | 0.85 (0.82, 0.88) | 0.005 | 1.00 | -0.00 (-0.24, 0.24) | 1.00 (0.79, 1.21) | 0.560 | 0.13 | 0.11 | 0.08 |
| P-POSSUM RCM | 622 | 10.9 ( 4.5, 16.8) | 0.79 (0.75, 0.82) | 0.868 | 1.00 | -0.00 (-0.23, 0.23) | 1.00 (0.75, 1.25) | 0.199 | 0.12 | 0.10 | 0.06 |
| POTTER RCM | 486 | 20.9 (10.9, 29.1) | 0.84 (0.81, 0.87) | 0.081 | 0.93 | -0.00 (-0.28, 0.28) | 1.00 (0.74, 1.26) | 0.853 | 0.12 | 0.10 | 0.06 |
| ACS-NSQIP RCM | 618 | 24.3 (16.4, 31.6) | 0.84 (0.81, 0.87) | 0.030 | 1.01 | 0.00 (-0.24, 0.24) | 1.00 (0.79, 1.21) | 0.806 | 0.13 | 0.10 | 0.08 |

**Note.** RCM, re-calibrated model; N, number of patients in the analysis; CI, confidence interval; AUC, area under the curve; E:O, ratio of expected and observed events; CITL, calibration-in-the-large; Slope, calibration slope; HL-GOF, Hosmer-Lemeshow goodness-of-fit test; NB, net benefit (calculated at decision thresholds 5%, 10% and 20%).

**Figure S3**.Calibration of prognostic models when predicting 30-day postoperative death after updating of intercept and slope (recalibration). The blue line is a smoothed locally weighted regression (lowess) line that shows the agreement between predicted probabilities and observed proportions of 30-day mortality. The dashed diagonal line indicates perfect calibration. The circled points represent mean risks in decile groups of predicted probabilities, with vertical lines representing 95% confidence intervals. The spike plot on the x-axis summarises the density of patients in the range of predicted risks of 30-day death. RCM, re-calibrated model; E:O, ratio of expected and observed deaths; CITL, calibration-in-the-large; Slope, calibration slope; AUC, area under the curve; GOF goodness-of-fit.



## Figure S4. Decision curves showing the net benefit in clinical decision-making of using each prognostic model of 30-day postoperative mortality after updating of intercept and slope (recalibration).



## Figure S5. Forest plot with hospital-specific Brier scores of the Surgeon's assessment prognostic model and overall pooled Brier score based on random-effects meta-analysis.



**Note.** The overall meta-analyzed scaled Brier score is represented by the diamond centered on its estimated value with the diamond width corresponding to the length of the confidence interval. The green whiskers extending from the overall diamond represent the prediction interval, which provides a plausible range for the scaled Brier score in a future, new study.

## Figure S6. Forest plot with hospital-specific Brier scores of the NELA prognostic model and overall pooled Brier score based on random-effects meta-analysis.



**Note.** The overall meta-analyzed scaled Brier score is represented by the diamond centered on its estimated value with the diamond width corresponding to the length of the confidence interval. The green whiskers extending from the overall diamond represent the prediction interval, which provides a plausible range for the scaled Brier score in a future, new study.

## Figure S7. Forest plot with hospital-specific Brier scores of the P-POSSUM prognostic model and overall pooled Brier score based on random-effects meta-analysis.



**Note.** The overall meta-analyzed scaled Brier score is represented by the diamond centered on its estimated value with the diamond width corresponding to the length of the confidence interval. The green whiskers extending from the overall diamond represent the prediction interval, which provides a plausible range for the scaled Brier score in a future, new study.

## Figure S8. Forest plot with hospital-specific Brier scores of the POTTER prognostic model and overall pooled Brier score based on random-effects meta-analysis.



**Note.** The overall meta-analyzed scaled Brier score is represented by the diamond centered on its estimated value with the diamond width corresponding to the length of the confidence interval. The green whiskers extending from the overall diamond represent the prediction interval, which provides a plausible range for the scaled Brier score in a future, new study.

## Figure S9. Forest plot with hospital-specific Brier scores of the ACS-NSQIP prognostic model and overall pooled Brier score based on random-effects meta-analysis.



**Note.** The overall meta-analyzed scaled Brier score is represented by the diamond centered on its estimated value with the diamond width corresponding to the length of the confidence interval. The green whiskers extending from the overall diamond represent the prediction interval, which provides a plausible range for the scaled Brier score in a future, new study.

## Figure S10. Forest plot with hospital-specific Brier scores of the ACS-NSQIP adjusted prognostic model and overall pooled Brier score based on random-effects meta-analysis.



**Note.** The overall meta-analyzed scaled Brier score is represented by the diamond centered on its estimated value with the diamond width corresponding to the length of the confidence interval. The green whiskers extending from the overall diamond represent the prediction interval, which provides a plausible range for the scaled Brier score in a future, new study.

## Table S5: Checklist for transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD)

|  |  |  |  |
| --- | --- | --- | --- |
| **Section/Topic** | **IteIIItem** | **Checklist Item** | **Page** |
| **Title and abstract** |
| Title | 1 | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | Title Page file, 1 |
| Abstract | 2 | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | Abstract file, 1-2 |
| **Introduction** |
| Background and objectives | 3a | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | Manuscript file, 1-2 |
| 3b | Specify the objectives, including whether the study describes the development or validation of the model or both. | 2 |
| **Methods** |
| Source of data | 4a | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | 2, 6-7 |
| 4b | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.  | 2 |
| Participants | 5a | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | 2 |
| 5b | Describe eligibility criteria for participants.  | 3, and Supplementary Table S1 |
| 5c | Give details of treatments received, if relevant.  | 2-3 |
| Outcome | 6a | Clearly define the outcome that is predicted by the prediction model, including how and when assessed.  | 3 |
| 6b | Report any actions to blind assessment of the outcome to be predicted.  | 2-3 |
| Predictors | 7a | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | 2-3, 7 |
| 7b | Report any actions to blind assessment of predictors for the outcome and other predictors.  | Not Applicable |
| Sample size | 8 | Explain how the study size was arrived at. | 3-4 |
| Missing data | 9 | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.  | 4 |
| Statistical analysis methods | 10c | For validation, describe how the predictions were calculated.  | 3-4 |
| 10d | Specify all measures used to assess model performance and, if relevant, to compare multiple models.  | 4-6 |
| 10e | Describe any model updating (e.g., recalibration) arising from the validation, if done. | 6 |
| Risk groups | 11 | Provide details on how risk groups were created, if done.  | Not done |
| Development vs. validation | 12 | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.  | 6-7 and Supplementary Table S3 |
| **Results** |
| Participants | 13a | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.  | 7 |
| 13b | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.  | 7, Tables 1 and 2 |
| 13c | For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).  | 8, Supplementary Table S3 |
| Model performance | 16 | Report performance measures (with CIs) for the prediction model. | 8-10, Table 3, Figures 1-3, Supplementary Figure S1 – S10 |
| Model-updating | 17 | If done, report the results from any model updating (i.e., model specification, model performance). | 9-10, Supplementary Figure S3, Table S3 and Figure S4 |
| **Discussion** |
| Limitations | 18 | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).  | 13-14 |
| Interpretation | 19a | For validation, discuss the results with reference to performance in the development data, and any other validation data.  | 10-13 |
| 19b | Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.  | 10-13 |
| Implications | 20 | Discuss the potential clinical use of the model and implications for future research.  | 14 |
| **Other information** |
| Supplementary information | 21 | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.  | 14 |
| Funding | 22 | Give the source of funding and the role of the funders for the present study.  | 15 |