**Appendix Table 1.** Top 100 (of 215) features considered for model inclusion. Table includes inclusion in the final model, missingness, descriptive statistics, details about units, and narrative descriptions for features included in the final model.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Rank | Feature | Included in model | Missing (%) | Mean | SD | Units (if included model | Extended description (if included model) |
| 1 | APR-DRG risk of mortality | Y | 6.2% | 2.5 | 1.1 | integer (integer 1-4) | APR-DRG risk of mortality is pulled from administrative/billing records. |
| 2 | Last Glasgow Coma Score | Y | 9.0% | 13.4 | 3.0 | integer (integer 1-15) | Glasgow Coma Scores are provided for most ICU patients and the last observed score is used. |
| 3 | APR-DRG severity of illness | Y | 1.1% | 2.6 | 1.1 | integer (integer 1-4) | APR-DRG severity of illness is pulled from administrative/billing records. |
| 4 | Last measured shock index (HR/SBP) x age | Y | 0.3% | 43.7 | 21.1 | beats per minute/mmHg x age in years | The last measured heart rate is divided by the last measured systolic blood pressure and multiplied by patient age. |
| 5 | Medicare cost weight index | Y | 0.0% | 2.4 | 2.4 | number for MSDRG | Medicare MS-DRG are pulled from administrative/billing records. The weights represent the average resources required to care for cases with a particular DRG and updated each year. |
| 6 | Last pulse oximetry | Y | 0.4% | 96.1 | 7.4 | % | Last pulse oximetry measurement |
| 7 | Last shock index (HR/SBP) | Y | 0.3% | 0.7 | 0.3 | beats per minute/mmHg | The last measured heart rate is divided by the last measured systolic blood pressure. |
| 8 | Last heart rate | Y | 0.1% | 82.4 | 20.1 | beats per minute | Last heart rate |
| 9 | Last CO2 measurement | Y | 3.3% | 24.5 | 4.7 | mmol/L | Last end tidal CO2 measurement |
| 10 | Mean pulse oximetry | Y | 0.4% | 97.0 | 2.6 | mean of % | Mean of all pulse oximetry percentages from as early as 24 hours pre-ICU admission to up to 24 hours post-ICU admission. |
| 11 | Last mechanical ventilation status (Y/N) | Y | 4.0% | 0.1 | 0.4 | 1/0 | This indicator variable is derived from a single result called 'Oxygen Delivery'. If the result contains "vent" or "mechanical ventilation, it is coded as a 1, otherwise 0. |
| 12 | Last systolic blood pressure | Y | 0.1% | 123.5 | 23.9 | mmHg | Last systolic blood pressure |
| 13 | Mean respiratory rate | Y | 0.1% | 19.2 | 3.8 | breaths per minute | Mean of all respiratory rates from as early as 24 hours pre-ICU admission to up to 24 hours post-ICU admission. |
| 14 | Mean temperature | Y | 3.8% | 98.3 | 0.8 | mean of temperature Fahrenheit | Mean of all temperature in Fahrenheit from as early as 24 hours pre-ICU admission to up to 24 hours post-ICU admission. |
| 15 | Last evidence of any oxygen therapy (Y/N) | Y | 4.0% | 0.6 | 0.5 | 1/0 | This indicator variable is derived from a single result called 'Oxygen Delivery'. If the result contains anything other than the result "Room air", it is coded as a 1, otherwise 0. |
| 16 | Change in creatinine level | Y | 3.4% | -0.1 | 0.8 | mg/dL | Last creatinine minus first creatinine |
| 17 | Last blood urea nitrogen | Y | 3.5% | 25.1 | 20.4 | mg/dL | Last blood urea nitrogen |
| Rank | Feature | Included in model | Missing (%) | Mean | SD | Units (if included model | Extended description (if included model) |
| 18 | Mean systolic blood pressure | N | 0.1% | 124.4 | 19.1 |  |  |
| 19 | Platelet count | N | 7.6% | 232.1 | 104.9 |  |  |
| 20 | Last respiratory rate | N | 0.1% | 19.1 | 6.7 |  |  |
| 21 | First Glasgow Coma Score | N | 9.0% | 13.1 | 3.3 |  |  |
| 22 | Heart rate mean | N | 0.1% | 85.0 | 16.0 |  |  |
| 23 | Mean Glasgow Coma Score | N | 9.0% | 13.2 | 2.9 |  |  |
| 24 | Aspartate aminotransferase | N | 24.3% | 78.0 | 436.6 |  |  |
| 25 | Surgical IP flag | N | 0.0% | 0.3 | 0.4 |  |  |
| 26 | Last Creatinine level | N | 3.4% | 1.6 | 1.8 |  |  |
| 27 | Time since last inpatient visit | N | 52.3% | 337.5 | 498.1 |  |  |
| 28 | Last red blood cell distribution | N | 16.5% | 15.4 | 2.6 |  |  |
| 29 | Albumin Level | N | 23.1% | 3.4 | 0.7 |  |  |
| 30 | Last white blood cell count | N | 9.0% | 11.3 | 8.1 |  |  |
| 31 | Last lymphocytes | N | 11.7% | 15.2 | 11.1 |  |  |
| 32 | Change in hematocrit | N | 1.5% | 5.3 | 173.7 |  |  |
| 33 | Last chloride level | N | 3.2% | 103.4 | 6.3 |  |  |
| 34 | Change in white blood cells | N | 6.4% | 419.5 | 2004.8 |  |  |
| 35 | Change in CO2 level | N | 1.8% | 4.7 | 223.7 |  |  |
| 36 | Systolic blood pressure x age | N | 0.0% | 2.0 | 5.5 |  |  |
| 37 | Shock index start | N | 0.1% | 45.0 | 45.1 |  |  |
| 38 | Fraction of inspired oxygen (FiO2) | N | 62.1% | 56.6 | 37.3 |  |  |
| 39 | Number of ventilation or any oxygen therapy events | N | 4.0% | 12.5 | 22.3 |  |  |
| 40 | Last glucose level | N | 2.2% | 142.4 | 62.5 |  |  |
| 41 | Change in blood urea nitrogen | N | 1.6% | 27.0 | 500.5 |  |  |
| Rank | Feature | Included in model | Missing (%) | Mean | SD | Units (if included model | Extended description (if included model) |
| 42 | Systolic blood pressure x age end | N | 0.0% | 2.0 | 5.6 |  |  |
| 43 | Religious affiliation (Y) | N | 27.9% | 0.3 | 0.5 |  |  |
| 44 | Last sodium level | N | 3.2% | 138.6 | 5.0 |  |  |
| 45 | First red blood cell count | N | 3.1% | 4.1 | 0.9 |  |  |
| 46 | Last temperature Fahrenheit | N | 3.8% | 98.3 | 0.9 |  |  |
| 47 | Troponin I level | N | 72.2% | 2.5 | 17.5 |  |  |
| 48 | Weight | N | 48.0% | 81.7 | 25.8 |  |  |
| 49 | Total hematology events | N | 28.0% | 13.5 | 35.4 |  |  |
| 50 | Age at admission | N | 0.0% | 63.7 | 17.8 |  |  |
| 51 | Relative change in neutrophils | N | 2.0% | 51.4 | 727.3 |  |  |
| 52 | Change in heart rate | N | 0.1% | -8.1 | 26.7 |  |  |
| 53 | Last diastolic blood pressure | N | 0.2% | 66.7 | 15.7 |  |  |
| 54 | Change in shock index (HR/SBP) x age | N | 0.3% | -1.3 | 22.5 |  |  |
| 55 | Change in temperature Fahrenheit | N | 3.8% | 0.0 | 1.4 |  |  |
| 56 | Anion gap | N | 26.4% | 15.4 | 6.4 |  |  |
| 57 | Chemistry events | N | 8.6% | 14.1 | 28.1 |  |  |
| 58 | Blood urea nitrogen to creatinine ratio | N | 25.3% | 19.7 | 10.8 |  |  |
| 59 | Partial pressure of oxygen | N | 55.6% | 146.1 | 122.6 |  |  |
| 60 | Respiratory rate change | N | 0.1% | -0.6 | 8.4 |  |  |
| 61 | First red blood cell distribution | N | 16.6% | 15.3 | 2.6 |  |  |
| 62 | Change in diastolic blood pressure | N | 0.2% | -8.3 | 21.6 |  |  |
| 63 | Absolute change in neutrophils | N | 2.0% | 51.3 | 714.1 |  |  |
| 64 | Time since last emergency or inpatient visit | N | 32.8% | 235.7 | 399.9 |  |  |
| 65 | Chloride level | N | 3.3% | 100.9 | 7.1 |  |  |
| Rank | Feature | Included in model | Missing (%) | Mean | SD | Units (if included model | Extended description (if included model) |
| 66 | Last glomerular filtration rate (non-African American) | N | 32.5% | 61.4 | 46.2 |  |  |
| 67 | First temperature Fahrenheit | N | 3.8% | 98.2 | 1.3 |  |  |
| 68 | Change in sodium level | N | 1.5% | 7.6 | 295.4 |  |  |
| 69 | First glucose level | N | 2.9% | 167.2 | 127.1 |  |  |
| 70 | Last calcium level | N | 3.2% | 8.4 | 0.8 |  |  |
| 71 | First white blood cell count | N | 10.1% | 11.7 | 8.9 |  |  |
| 72 | Count of heat CT scans | N | 9.6% | 0.5 | 1.9 |  |  |
| 73 | Count of previous visits | N | 0.0% | 6.4 | 13.4 |  |  |
| 74 | Absolute neutrophil count | N | 11.9% | 9.0 | 6.4 |  |  |
| 75 | Relative eosinophil level | N | 12.4% | 1.4 | 2.2 |  |  |
| 76 | Alkaline phosphatase | N | 24.2% | 102.4 | 92.2 |  |  |
| 77 | Last red blood cell count | N | 3.0% | 3.8 | 0.8 |  |  |
| 78 | Change in Glasgow coma scale | N | 9.0% | 0.3 | 2.4 |  |  |
| 79 | Number of urine tests carried out | N | 48.5% | 22.0 | 31.1 |  |  |
| 80 | Sodium level | N | 3.3% | 137.5 | 5.9 |  |  |
| 81 | Diastolic blood pressure mean | N | 0.2% | 67.6 | 11.8 |  |  |
| 82 | Last albumin level | N | 23.1% | 3.2 | 0.7 |  |  |
| 83 | Change in shock index (HR/SBP) | N | 0.3% | 0.0 | 0.6 |  |  |
| 84 | Phosphate | N | 62.2% | 3.8 | 2.5 |  |  |
| 85 | Total number of ventilation events | N | 4.0% | 2.2 | 11.6 |  |  |
| 86 | Count of CT scans | N | 9.6% | 1.5 | 4.9 |  |  |
| 87 | First potassium level | N | 3.1% | 4.2 | 0.8 |  |  |
| 88 | First glomerular filtration rate (non-African American) | N | 30.2% | 57.2 | 47.5 |  |  |
| 89 | Creatine kinase total | N | 63.2% | 168.3 | 181.4 |  |  |
| 90 | Urinalysis for bilirubin | N | 48.8% | 0.1 | 0.3 |  |  |
| Rank | Feature | Included in model | Missing (%) | Mean | SD | Units (if included model | Extended description (if included model) |
| 91 | First systolic blood pressure | N | 0.1% | 134.8 | 33.5 |  |  |
| 92 | Change in blood oxygen saturation | N | 0.4% | 0.1 | 9.2 |  |  |
| 93 | Last hematocrit | N | 2.8% | 34.1 | 6.6 |  |  |
| 94 | First pulse oximetry | N | 0.4% | 96.0 | 6.1 |  |  |
| 95 | First CO2 measurement | N | 3.3% | 24.2 | 5.5 |  |  |
| 96 | RACE (white, black, other, unknown) | N | 0.0% | NA | NA |  |  |
| 97 | Change in calcium level | N | 2.9% | 4.2 | 195.1 |  |  |
| 98 | First Glasgow Coma Score | N | 9.0% | 13.1 | 3.3 |  |  |
| 99 | Change in albumin level | N | 23.1% | -0.2 | 0.4 |  |  |
| 100 | SRG9 Payer | N | 0.0% | NA | NA |  |  |

**Appendix Table 2**. XGboost was used to generate 100 different machine learning models from 100 bootstrap samples of training data. Frequency is the number of times a feature was included in the final model out of 100. Average rank in the average feature ranking of the feature across the 100 models.

|  |  |  |
| --- | --- | --- |
| Feature | Frequency | Avg. Rank |
| APR-DRG risk of mortality (integer 1-4) | 100 | 1 |
| Last Glasgow Coma Scale (integer 1-15) | 100 | 2.02 |
| APR-DRG severity of illness (integer 1-4) | 100 | 3.13 |
| Last measured shock index (HR/SBP) x age | 100 | 4.87 |
| Last shock index (HR/SBP) | 100 | 5.84 |
| Last pulse oximetry | 100 | 6.61 |
| Last heart rate | 100 | 6.65 |
| Medicare cost weight index | 100 | 8.11 |
| Mean pulse oximetry | 100 | 9.96 |
| Last CO2 measurement | 100 | 10.17 |
| Last mechanical ventilation status (Y/N) | 100 | 11.74 |
| Last systolic blood pressure | 100 | 12.04 |
| Mean respiratory rate | 100 | 12.22 |
| Mean temperature Fahrenheit | 100 | 14.6 |
| Last evidence of any oxygen therapy (Y/N) | 100 | 16.9 |
| Change in creatinine level | 100 | 17.88 |
| Last blood urea nitrogen | 99 | 26.73 |

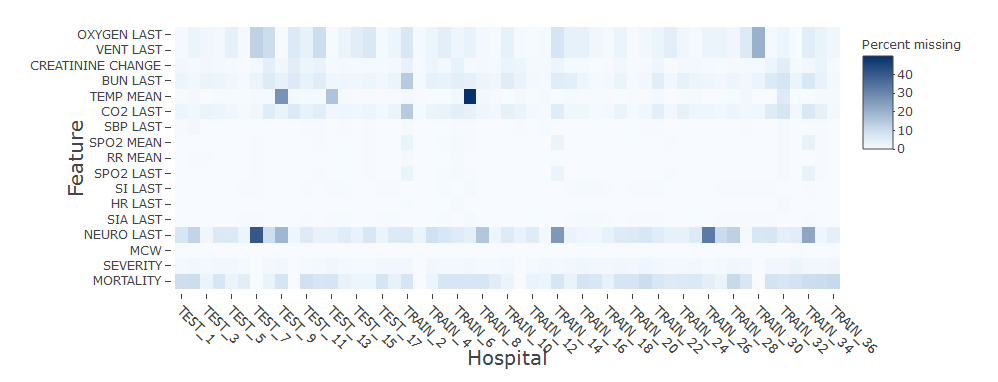
**Appendix Table 3**. The RIPD score has an alternate version (RIPD\_reduced.rds) that does not require any administrative/DRG terms. This version of the score uses only 14 features removing the APR-DRG risk of mortality, the APR-DRG severity of illness, and the Medicare cost weight index.

|  |  |  |  |
| --- | --- | --- | --- |
| Feature | Mean (SD) | Missingness | Relative influence |
| Last Glasgow Coma Score (integer 1-15) | 13.4 (3) | 8.98% | 30.54% |
| Last shock index | 0.7 (0.3) | 0.26% | 12.51% |
| Last measured shock index (HR/SBP) x age | 43.7 (21.1) | 0.26% | 8.59% |
| Last blood urea nitrogen | 25.1 (20.4) | 3.54% | 5.62% |
| Last mechanical ventilation status (Y/N) | 14.4% (Y) | 3.96% | 5.18% |
| Last systolic blood pressure | 123.5 (23.9) | 0.15% | 5.08% |
| Mean respiratory rate | 19.2 (3.8) | 0.10% | 4.88% |
| Last pulse oximetry | 96.1 (7.4) | 0.44% | 4.80% |
| Last evidence of any oxygen therapy (Y/N) | 59.2% (Y) | 3.96% | 4.74% |
| Last CO2 measurement | 24.5 (4.7) | 3.26% | 4.54% |
| Mean pulse oximetry | 97 (2.6) | 0.44% | 4.09% |
| Last heart rate | 82.4 (20.1) | 0.08% | 3.84% |
| Mean temperature Fahrenheit | 98.3 (0.8) | 3.75% | 3.22% |
| Change in creatinine level | -0.14 (0..83) | 3.59% | 2.37% |

**Appendix Table 4**. RIPD\_reduced crude and standardized mortality ratios, accuracy, and discrimination in training set, validation set. The RIPD\_reduced score has no DRG-related codes and uses only the clinical variables shown above in Appendix Table 3.

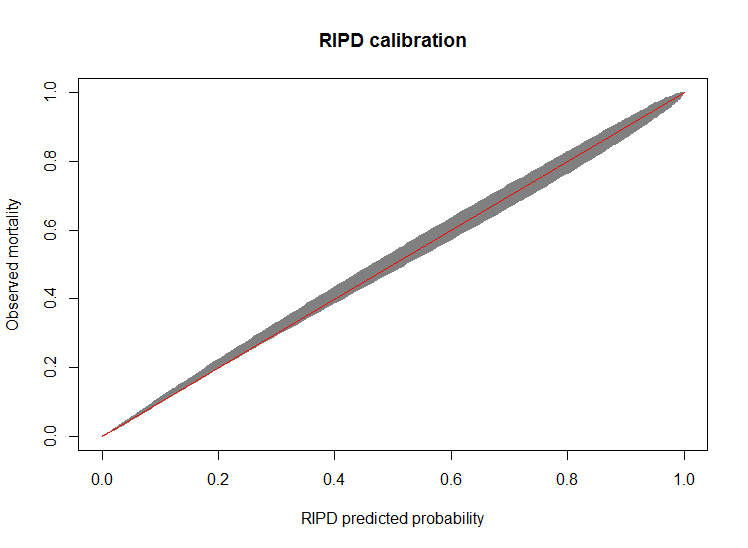
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Dataset** | **Sample size** | **Deaths** | **Mortality rate** | **SMR\***  **(95% CI)** | **Adjusted Brier Score\*\*** | **AUC** |
| Training set (82 ICUs) | 146,982 | 13,725 | 9.3% | 1.01 (0.99-1.03) | 48.8% | 0.921 |
| Validation set (49 ICUs) | 90,191 | 8,168 | 9.1% | 1.02 (1.00-1.05) | 48.6% | 0.914 |

**Appendix Figure 1**. Figure showing prevalence of missing data by feature and hospital. This figure highlights the heterogeneity of data availability across our system of hospitals. The heterogeneity of data availability supports our assertion that the RIPD algorithm performs well even in contexts where the underlying data availability varies.



OXYGEN LAST = Last evidence of any oxygen therapy; VENT LAST = Last mechanical ventilation status; CREATININE CHANGE = Last creatinine; BUN LAST = Last blood urea nitrogen; TEMP MEAN = Mean temperature Fahrenheit; CO2 LAST = Last CO2 measurement ; SBP LAST = Last systolic blood pressure; SPO2 MEAN = Mean pulse oximetry; RR MEAN = Mean respiratory rate; SPO2 LAST = Last pulse oximetry; SI LAST = Last shock index (HR/SBP); HR LAST = Last heart rate; SIA LAST = Last measured shock index (HR/SBP) x age; NEURO LAST = Last Glasgow Coma Scale; MCW = Medicare cost weight index; SEVERITY = APR-DRG severity of illness; MORTALITY = APR-DRG risk of mortality; .

**Appendix Figure 2.** Calibration curve with 95% confidence intervals, based on 100 bootstrapped samples of 500 patients from each of the 49 ICUs in the validation set. Red line represents perfect agreement between observed and predicted probabilities. Gray area represents 95% confidence area for predicted probabilities. Gray area inclusive of unity suggests good agreement between observed and predicted probabilities.



**Supplementary methods.**

*Adjusted Brier score*

The adjusted Brier score indicates the accuracy gained from using a model (i.e., the RIPD score) in excess of the accuracy of using average mortality. It has been used in other publications on ICU risk adjustment. (1,2) In the case of our validation hospitals, the model-free average mortality rate for each patient is 9.1%. Observed mortality rate in the presence of a model would give a positive Brier score if it better predicts the mortality observed in the subgroup, whether it is an APR-DRG group such as sepsis which may have a higher base mortality rate of 24% or a lower rate such as diabetes at 0.1%. Negative adjusted Brier scores would indicate a model performs worse for such a group than expected by the base model. Positive scores would indicate better performance than prediction based on simple averages.

Raw Brier score is calculated as follows:

Null Brier score = incidence of group x (1 – incidence)

Adjusted Brier score = (Null Brier Score – Raw Brier Score) / Null Brier Score

*Hyper-parameter tuning*

In order to identify optimum hyper-parameters, we used the R package, Caret (3). We altered our step size shrinkage (learning rate) from 0.5, 0.1, 0.05, 0.01, 0.005, and 0.001 and used trees with the following depths 2,3,4,5,6, and 10. Area under the curve was evaluated with each combination to maximize this value with each combination of hyper-parameters until the optimum pair (step size, 0.1; and tree depth, 4) were identified