# SDC 3. Test Sample Performance Data.

Figure 1. Receiver Operating Characteristic Curves.

Figure 2. Precision Recall Curves

Figure 3: Calibration Plots

Table 1. Accuracy, Precision, and Negative Predictive Value at Sensitivities and Specificities of .85, .90, .95, and .99 at the lower boundary of the 95% confidence interval.

Figure 1: Receiver Operating Characteristic Curves: Test Sample
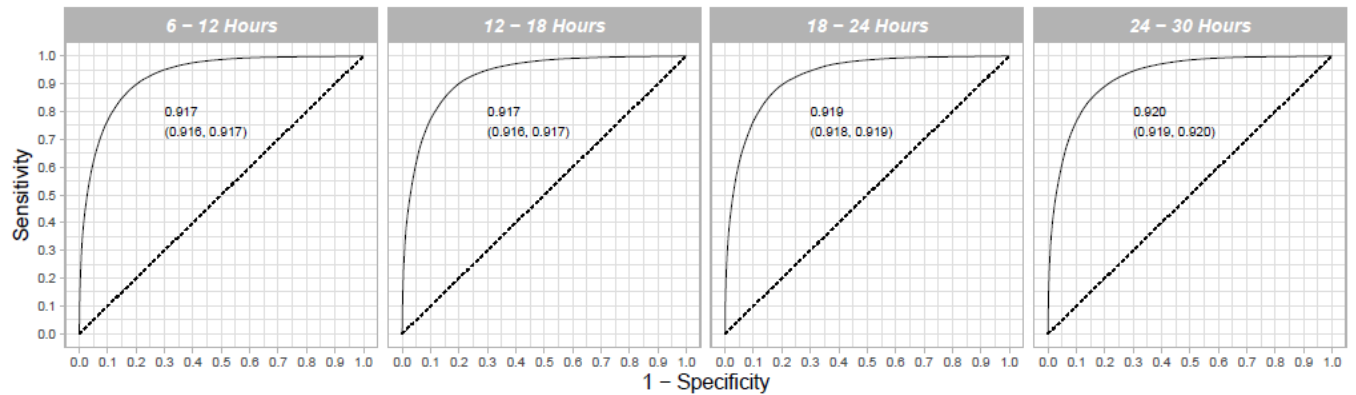
Receiver Operating Characteristic Curves. Receiver operating characteristic curves for the test sample classifying care as non-ICU or ICU for the respective future time periods.

Figure 2. Precision Recall Curve: Test Sample.
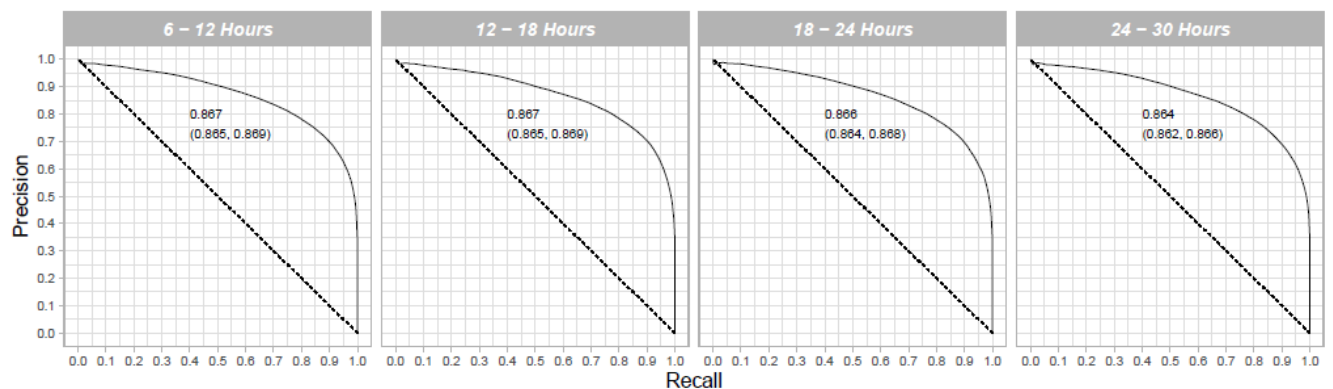
Precision Recall Curves. Precision recall curves for the test sample classifying care as routine or ICU for the respective future time periods. The area under the precision recall curves and 95% CIs are included in each panel. The areas were computed with integral approximations and the CIs were computed using a logit method.[24]

Figure 3. Calibration Plot. Test Sample.
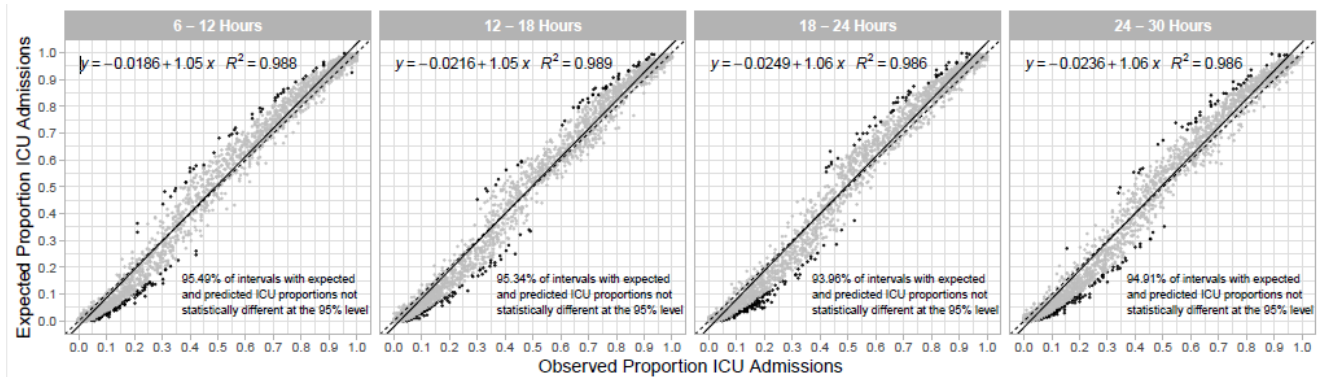
Calibration Plot. The y axis shows the expected proportion of ICU care areas for the time periods based on the risk intervals and the x-axis shows the observed proportion of ICU care areas in the time periods. The line of identity is the dashed line. The limits of the risk intervals were initially defined by 3001 equally spaced empirical quantiles. Some intervals were combined to have a larger number of risk values but ensuring the intervals were smaller than 0.005 resulting in 99% of the intervals containing at least 97 time periods and 2,967, 2,939, 2,912, and 2,925 risk intervals in the four models. Within each interval we compute the average expected risk of ICU admission and the observed risk of ICU admission. The circles indicate the observed and the expected proportions of ICU 6-hour time periods over ascending Criticality Index intervals. The black circles indicate the observed/expected proportions whose 95% confidence intervals do not cross the line of identity.   A linear regression is reported in each panel with their respective R^2 and the fitted mean is represented with the solid line. Calibration was accomplished for each model by using their output for B-spline polynomials as covariates in a linear logistic regression with outcome ICU/Regular care for the respective future time. This calibration method is similar to the Platt scaling method for support vector machines.[20]

Table 1. Accuracy, Precision, and Negative Predictive Value (NPV) at Sensitivities and Specificities of .85, .90, .95, and .99 at the lower boundary of the 95% confidence interval.

Sensitivities					
Prediction Time Period	Sensitivity (1)	Accuracy	Precision	Specificity	NPV
>6 -12 Hours	0.852 (0.850, 0.854)	0.848 (0.847, 0.85)	0.744 (0.741, 0.746)	0.847 (0.845, 0.848)	0.916 (0.915, 0.918)
>6 -12 Hours	0.902 (0.900, 0.904)	0.831 (0.830, 0.832)	0.696 (0.694, 0.698)	0.794 (0.792, 0.796)	0.939 (0.938, 0.940)
>6 -12 Hours	0.952 (0.950, 0.953)	0.785 (0.783, 0.786)	0.622 (0.619, 0.624)	0.697 (0.696, 0.699)	0.965 (0.964, 0.966)
>6 -12 Hours	0.991 (0.990, 0.991	0.637 (0.990, 0.991)	0.486 (0.484, 0.488)	0.452 (0.449, 0.454)	0.990 (0.989, 0.990)
>12 - 18 Hours	0.852 (0.850, 0.854)	0.850 (0.849, 0.851)	0.746 (0.744, 0.749)	0.848 0.799	0.917 (0.915, 0.918)
>12 - 18 Hours	0.902 (0.900, 0.904)	0.834 (0.833, 0.835)	0.701 (0.698, 0.703)	0.799 (0.797, 0.802)	0.940 (0.939, 0.941)
>12 - 18 Hours	0.951 (0.950, 0.953)	0.782 (0.781, 0.784)	0.619 (0.617, 0.621)	0.694 (0.692, 0.696)	0.965 (0.964, 0.966)
>12 - 18 Hours	0.991 (0.990, 0.991)	0.615 (0.614, 0.617)	0.471 (0.469, 0.473)	0.419 (0.417, 0.421)	0.989 0.988, 0.989)
>18 – 24 Hours	0.853 (0.850, 0.855)	0.848 (0.846, 0.849)	0.742 (0.74, 0.744)	0.845 (0.844, 0.847)	0.916 (0.915, 0.918)
>18 – 24 Hours	0.902 (0.900, 0.904)	0.829 (0.828, 0.831)	0.693 (0.691, 0.696)	0.791 (0.790, 0.793)	0.939 (0.938, 0.940)
>18 – 24 Hours	0.952 (0.950, 0.953)	0.777 (0.775, 0.778)	0.612 (0.610, 0.615)	0.685 (0.683, 0.687)	0.964 (0.963, 0.965)
>18 – 24 Hours	0.991 (0.991, 0.992)	0.618 (0.617, 0.620)	0.473 (0.471, 0.475)	0.423 (0.421, 0.425)	0.989 (0.989, 0.990)
>24 - 30 Hours	0.852 (0.850, 0.854)	0.846 (0.844, 0.847)	0.739 (0.736, 0.741)	0.842 (0.736, 0.741)	0.916 (0.915, 0.917)
>24 - 30 Hours	0.902 (0.901, 0.904)	0.825 (0.824, 0.826)	0.686 (0.684, 0.689)	0.784 (0.783, 0.786)	0.939 (0.938, 0.940)
>24 - 30 Hours	0.952 (0.950, 0.953)	0.775 (0.774, 0.777)	0.611 (0.609, 0.613)	0.683 (0.681, 0.685)	0.964 (0.963, 0.965)
>24 - 30 Hours	0.991 (0.991, 0.992)	0.623 (0.622, 0.625)	0.476 (0.474, 0.478)	0.431 (0.429, 0.433)	0.989 (0.989, 0.990)
Specificities					
Prediction Time Period	Specificity (1)	Accuracy	Precision	Sensitivity	NPV
>6 -12 Hours	0.852 (0.850, 0.853)	0.85 (0.848, 0.851)	0.749 (0.746, 0.751)	0.846 (0.844, 0.848)	0.914 (0.913, 0.915)
>6 -12 Hours	0.902 (0.900, 0.903)	0.855 (0.853, 0.856)	0.802 (0.800, 0.805)	0.765 (0.763, 0.767)	0.88 (0.879, 0.881)
>6 -12 Hours	0.951 (0.950, 0.952)	0.836 (0.835, 0.837)	0.868 (0.866, 0.870)	0.615 (0.612, 0.618)	0.825 (0.824, 0.827)
>6 -12 Hours	0.991 (0.990, 0.991)	0.762 (0.761, 0.763)	0.947 (0.945, 0.949)	0.325 (0.322, 0.327)	0.737 (0.736, 0.739)

>12 - 18 Hours	0.852 (0.850, 0.853)	0.851 (0.849, 0.852)	0.749 (0.747, 0.752)	0.849 (0.846, 0.851)	0.915 (0.914, 0.916)
>12 - 18 Hours	0.901 (0.900, 0.903)	0.856 (0.855, 0.857)	0.803 (0.801, 0.806)	0.770 (0.767, 0.772)	0.882 (0.881, 0.884)
>12 - 18 Hours	0.951 (0.950, 0.952)	0.835 (0.834, 0.837)	0.867 (0.865, 0.870)	0.614 (0.612, 0.617)	0.825 (0.824, 0.827)
>12 - 18 Hours	0.990 (0.990, 0.991)	0.762 (0.760, 0.763)	0.946 (0.944, 0.949)	0.324 (0.321, 0.326)	0.737 (0.735, 0.739)
>18 – 24 Hours	0.852 (0.850, 0.853)	0.849 (0.848, 0.851)	0.749 (0.746, 0.751)	0.845 (0.842, 0.847)	0.913 (0.912, 0.914)
>18 – 24 Hours	0.901 (0.900, 0.903)	0.854 (0.900, 0.903)	0.802 (0.799, 0.804)	0.763 (0.761, 0.766)	0.879 (0.878, 0.881)
>18 – 24 Hours	0.951 (0.950, 0.952)	0.835 (0.834, 0.836)	0.867 (0.865, 0.87)	0.613 (0.610, 0.616)	0.867 (0.865, 0.87)
>18 – 24 Hours	0.990 (0.990, 0.991)	0.761 (0.759, 0.762)	0.946 (0.944, 0.948)	0.321 (0.318, 0.324)	0.736 (0.735, 0.738)
>24 - 30 Hours	0.852 (0.850, 0.853)	0.848 (0.847, 0.849)	0.748 (0.745, 0.750)	0.842 (0.840, 0.844)	0.912 (0.910, 0.913)
>24 - 30 Hours	0.901 (0.900, 0.903)	0.853 (0.852, 0.855)	0.802 (0.799, 0.804)	0.762 (0.759, 0.764)	0.879 (0.877, 0.880)
>24 - 30 Hours	0.951 (0.950, 0.952)	0.833 (0.832, 0.834)	0.867 (0.864, 0.869)	0.608 (0.605, 0.610)	0.823 (0.821, 0.824)
>24 - 30 Hours	0.990 (0.990, 0.991)	0.763 (0.761, 0.764)	0.947 (0.945, 0.949)	0.327 (0.324, 0.330)	0.738 (0.736, 0.739)

(1) Lower bound of the 95% CI
(2) Sensitivity = true positive rate = ICU care time periods.
(3) Specificity = true negative rate = non-ICU care time periods.
(4) Precision = positive predictive value = true positives (cared for in the ICU)/[true positive + false positive]. Number needed to evaluate = 1/precision.

Legend.   The identification of true positives (ICU care time periods) is most relevant to identifying those patients expected to transfer to the ICU from non-ICU care areas or remain in the ICU. The identification of true negatives (non-ICU care time periods) is most relevant to identifying those patients not expected to transfer to the ICU or transfer from the ICU to non-ICU care areas.   The data shown are the estimates and 95% confidence intervals (CI).

Legend: The identification of true positives is most relevant to identifying those patients expected to transfer to the ICU from non-ICU care areas or remain in the ICU. The identification of true negatives is most relevant to identifying those patients not expected to transfer to the ICU. The data shown are the estimates and 95% confidence intervals (CI). Positives = time periods when patients in ICU care areas. Negatives = time periods when patient in non-ICU care areas.