**Title:** Predictive Approaches for Acute Dialysis Requirement in COVID-19

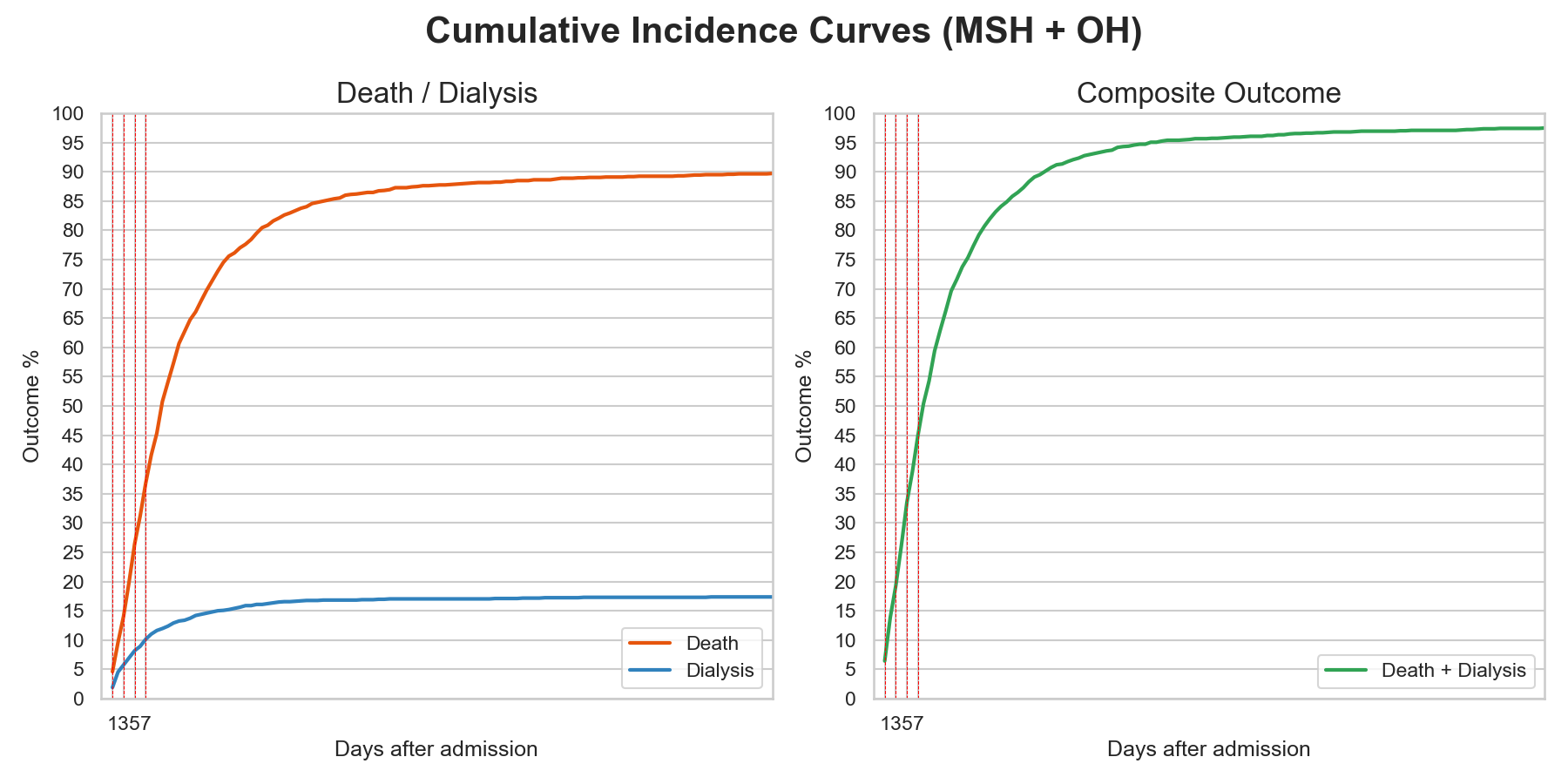
**Supplemental Figures:**

1. **Supplemental Figure 1:** Cumulative incidence plots of death, dialysis, and death and dialysis.
2. **Supplemental Figure 2.** Calibration for the Non-Imputed XGBoost Model over different time horizons. The difference for the calibration line over the model (under both isotonic and sigmoid activation) is not significantly different from the line of perfect calibration
3. **Supplemental Figure 3:** Sensitivity, specificity, and positive predictive value of XGBoost without imputation model at various prediction thresholds
4. **Supplemental Figure 4:** Ten features with highest SHAP scores in the XGBoost without imputation model at hospital day 1, 3, 5, and 7.

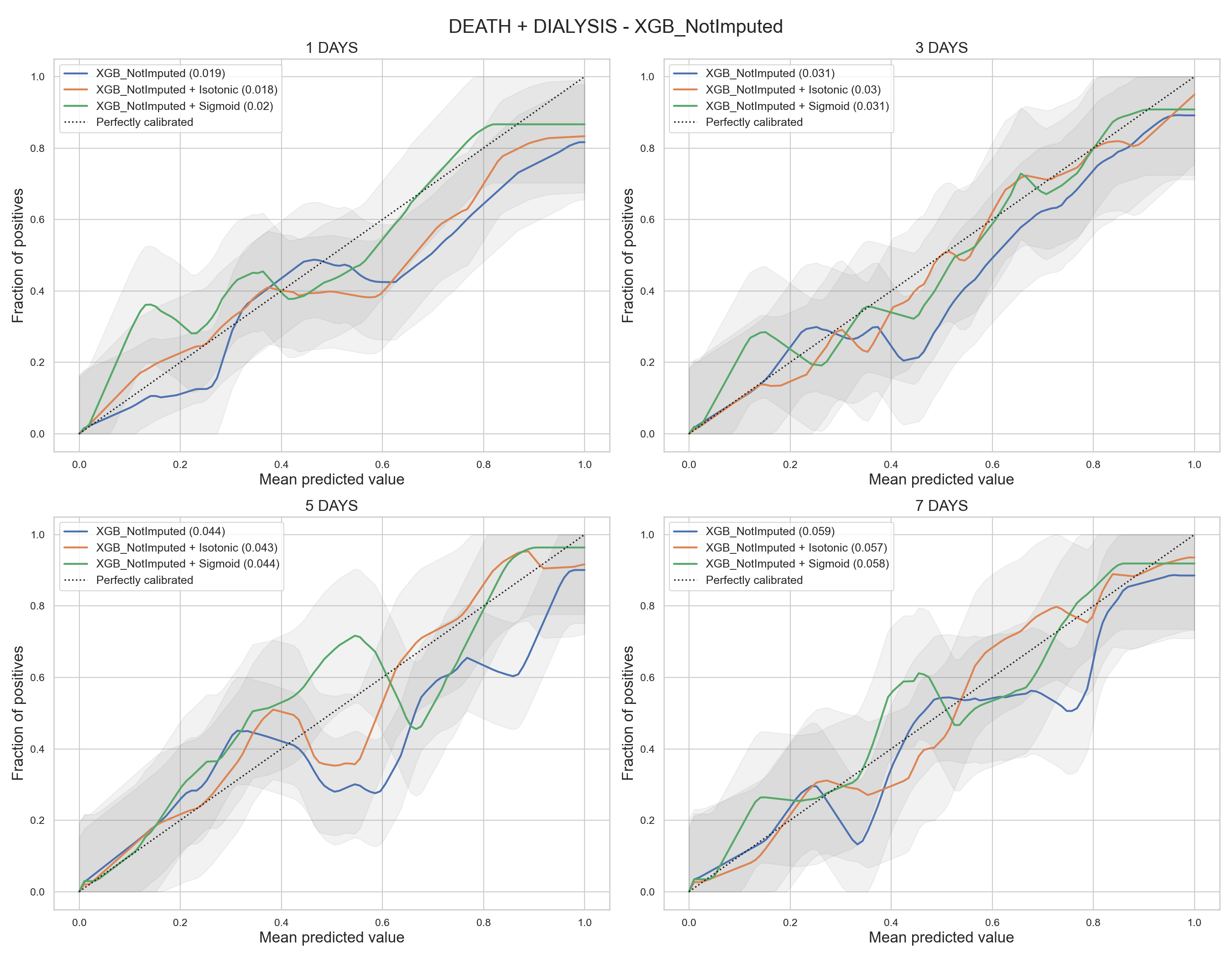
**Supplemental Tables:**

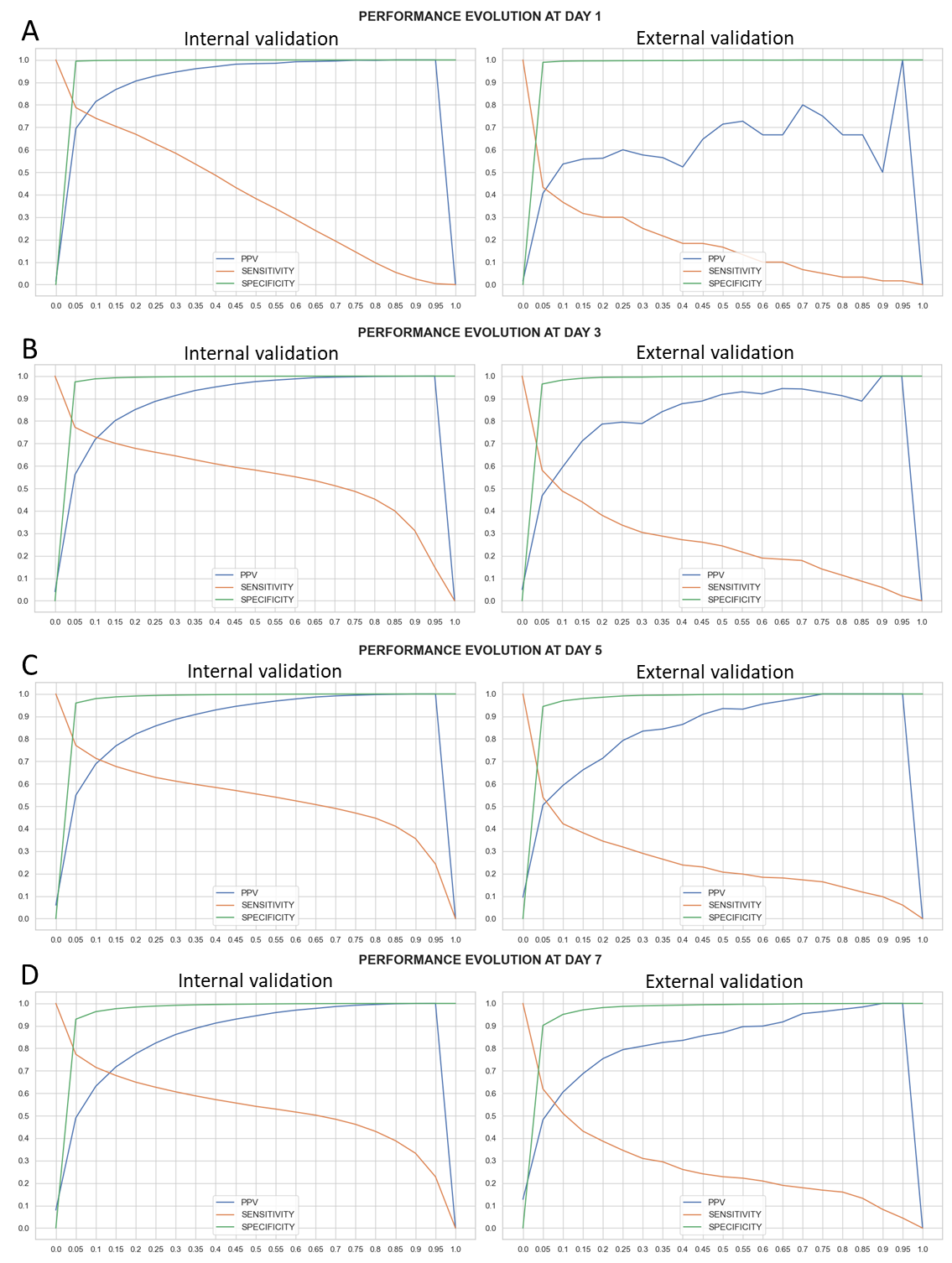
1. **Supplemental Table 1:** Number of death and dialysis in the MSH and OH cohort at prediction time windows.
2. **Supplemental Table 2.** Model Performance in Patients Admitted to the Intensive Care Unit in the External Validation Set
3. **Supplemental Table 3:** Performance Metrics of Different Models in Internal and External validation.

**Supplemental Figure 1:** Cumulative incidence plots of death, dialysis, and death and dialysis.

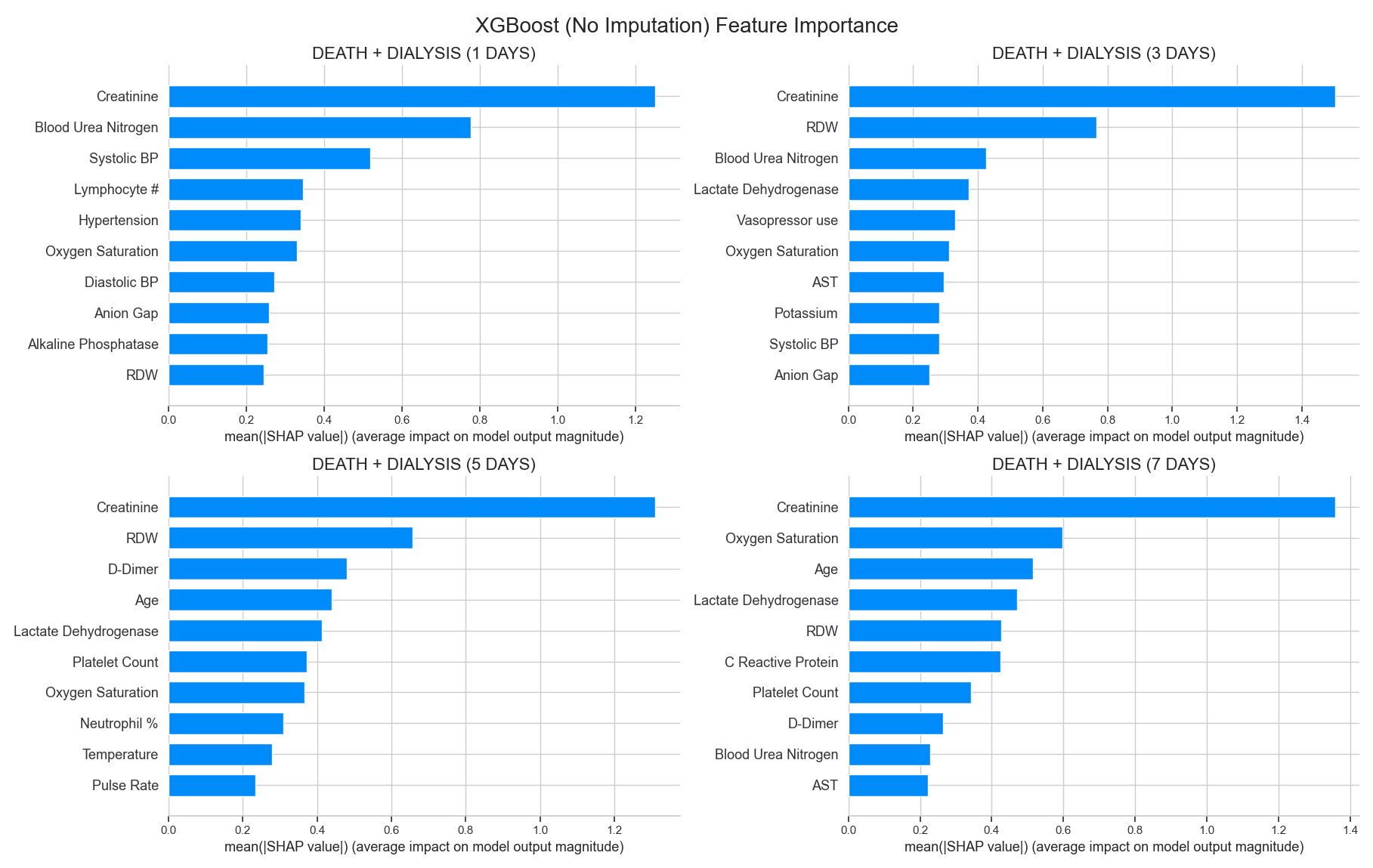


**Supplemental Figure 2.** Calibration for the Non-Imputed XGBoost Model over different time horizons. The difference for the calibration line over the model (under both isotonic and sigmoid activation) is not significantly different from the line of perfect calibration



**Supplemental Figure 3:** Sensitivity, specificity, and positive predictive value of XGBoost non-imputed model at various prediction thresholds for prediction at A) Day 1; B) Day 3; C) Day 5, and D) Day 7. Left panels are results from the internal validation cohort and right panels are results from the external validation cohort. X-axis is the range of model prediction and Y-axis is values for sensitivity, specificity, and positive predictive value (PPV). Green line is sensitivity, orange line is specificity, and blue line is PPV.

**Supplemental Figure 4:** Ten features with highest SHAP scores in the XGBoost without imputation model at hospital day 1, 3, 5, and 7.



**Supplemental Table 1:** Number of death and dialysis in the MSH and OH cohort at prediction time windows.

|  |  |  |  |
| --- | --- | --- | --- |
| **MSH (N=2442)** | | | |
| **Time Period** | **Death** | **Dialysis** | **Total Outcomes** |
| 1 Day | 17 | 18 | 35 |
| 3 Days | 47 | 54 | 101 |
| 5 Days | 75 | 72 | 147 |
| 7 Days | 107 | 91 | 198 |

|  |  |  |  |
| --- | --- | --- | --- |
| **OH (N=3651)** | | | |
| **Time Period** | **Death** | **Dialysis** | **Total Outcomes** |
| 1 Day | 50 | 10 | 60 |
| 3 Days | 153 | 31 | 184 |
| 5 Days | 299 | 49 | 348 |
| 7 Days | 409 | 59 | 468 |

**Supplemental Table 2:** Model Performance in Patients Admitted to the Intensive Care Unit in the External Validation Set

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cohort** | **OUTCOME** | **CLASSIFIER** | **ACCURACY** | **AUROC** | **AUPRC** | **F1-Score** | **Sensitivity** | **Specificity** | **PPV** | **NPV** |
| **ICU Patients**  **(OH)** | **Dialysis or Death at Day 1** | **LASSO** | 0.9 | 0.66 | 0.31 | 0.33 | 0.42 | 0.93 | 0.28 | 0.96 |
| **Logistic**  **Regression** | 0.58 | 0.68 | 0.19 | 0.17 | 0.74 | 0.57 | 0.10 | 0.97 |
| **Random**  **Forest** | 0.93 | 0.82 | 0.42 | 0.49 | 0.58 | 0.95 | 0.42 | 0.97 |
| **XGB**  **Imputed** | 0.85 | 0.81 | 0.4 | 0.33 | 0.63 | 0.86 | 0.23 | 0.97 |
| **XGB**  **Not Imputed** | 0.78 | 0.95 | 0.64 | 0.35 | 1.00 | 0.76 | 0.21 | 1.00 |
| **Dialysis or Death at Day 3** | **LASSO** | 0.81 | 0.82 | 0.62 | 0.5 | 0.71 | 0.83 | 0.39 | 0.95 |
| **Logistic**  **Regression** | 0.64 | 0.73 | 0.34 | 0.37 | 0.79 | 0.62 | 0.24 | 0.95 |
| **Random**  **Forest** | 0.84 | 0.84 | 0.64 | 0.54 | 0.69 | 0.87 | 0.45 | 0.95 |
| **XGB**  **Imputed** | 0.78 | 0.86 | 0.62 | 0.48 | 0.79 | 0.77 | 0.35 | 0.96 |
| **XGB**  **Not Imputed** | 0.9 | 0.93 | 0.79 | 0.67 | 0.81 | 0.91 | 0.58 | 0.97 |
| **Dialysis or Death at Day 5** | **LASSO** | 0.83 | 0.84 | 0.68 | 0.62 | 0.69 | 0.86 | 0.57 | 0.92 |
| **Logistic**  **Regression** | 0.72 | 0.75 | 0.44 | 0.52 | 0.74 | 0.72 | 0.41 | 0.91 |
| **Random**  **Forest** | 0.83 | 0.84 | 0.68 | 0.63 | 0.72 | 0.85 | 0.56 | 0.92 |
| **XGB**  **Imputed** | 0.83 | 0.84 | 0.68 | 0.63 | 0.71 | 0.86 | 0.56 | 0.92 |
| **XGB**  **Not Imputed** | 0.85 | 0.9 | 0.8 | 0.71 | 0.85 | 0.86 | 0.60 | 0.96 |
| **Dialysis or Death at Day 7** | **LASSO** | 0.79 | 0.86 | 0.75 | 0.65 | 0.79 | 0.79 | 0.55 | 0.92 |
| **Logistic**  **Regression** | 0.71 | 0.76 | 0.5 | 0.57 | 0.78 | 0.69 | 0.45 | 0.91 |
| **Random**  **Forest** | 0.8 | 0.84 | 0.72 | 0.64 | 0.71 | 0.83 | 0.57 | 0.90 |
| **XGB**  **Imputed** | 0.84 | 0.85 | 0.73 | 0.69 | 0.75 | 0.87 | 0.64 | 0.92 |
| **XGB**  **Not Imputed** | 0.84 | 0.87 | 0.82 | 0.71 | 0.78 | 0.86 | 0.65 | 0.92 |

**Supplemental Table 3.** Additional Performance Metrics of Different Models in Internal and External validation.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cohort** | **OUTCOME** | **CLASSIFIER** | **ACCURACY** | **F1-Score** | **Sensitivity** | **Specificity** | **PPV** | **NPV** |
| **Internal Validation**  **(MSH only)** | **Dialysis or Death at Day 1** | **LASSO** | 0.89 (0.88 - 0.9) | 0.28 (0.27 - 0.3) | 0.92 (0.91 - 0.93) | 0.89 (0.88 - 0.9) | 0.19 (0.17 - 0.2) | 1.0 (1.0 - 1.0) |
| **Logistic**  **Regression** | 0.79 (0.79 - 0.8) | 0.13 (0.12 - 0.13) | 0.88 (0.87 - 0.89) | 0.79 (0.79 - 0.8) | 0.07 (0.07 - 0.07) | 1.0 (1.0 - 1.0) |
| **Random**  **Forest** | 0.88 (0.87 - 0.89) | 0.25 (0.23 - 0.26) | 0.87 (0.87 - 0.88) | 0.88 (0.87 - 0.89) | 0.16 (0.15 - 0.18) | 1.0 (1.0 - 1.0) |
| **XGB**  **Imputed** | 0.9 (0.89 - 0.9) | 0.29 (0.28 - 0.31) | 0.94 (0.94 - 0.95) | 0.9 (0.89 - 0.9) | 0.19 (0.18 - 0.21) | 1.0 (1.0 - 1.0) |
| **XGB**  **Not Imputed** | 0.96 (0.96 - 0.96) | 0.54 (0.52 - 0.56) | 0.95 (0.94 - 0.95) | 0.96 (0.96 - 0.97) | 0.42 (0.4 - 0.44) | 1.0 (1.0 - 1.0) |
| **Dialysis or Death at Day 3** | **LASSO** | 0.87 (0.87 - 0.88) | 0.38 (0.38 - 0.39) | 0.9 (0.89 - 0.9) | 0.87 (0.87 - 0.87) | 0.25 (0.24 - 0.26) | 1.0 (0.99 - 1.0) |
| **Logistic**  **Regression** | 0.79 (0.79 - 0.8) | 0.26 (0.25 - 0.26) | 0.85 (0.85 - 0.86) | 0.79 (0.78 - 0.79) | 0.15 (0.15 - 0.16) | 0.99 (0.99 - 0.99) |
| **Random**  **Forest** | 0.87 (0.87 - 0.87) | 0.38 (0.37 - 0.39) | 0.88 (0.87 - 0.88) | 0.87 (0.86 - 0.87) | 0.26 (0.25 - 0.27) | 0.99 (0.99 - 0.99) |
| **XGB**  **Imputed** | 0.86 (0.86 - 0.87) | 0.38 (0.37 - 0.39) | 0.91 (0.9 - 0.91) | 0.86 (0.86 - 0.87) | 0.25 (0.24 - 0.26) | 1.0 (1.0 - 1.0) |
| **XGB**  **Not Imputed** | 0.91 (0.91 - 0.92) | 0.5 (0.49 - 0.51) | 0.9 (0.9 - 0.91) | 0.91 (0.91 - 0.92) | 0.37 (0.35 - 0.38) | 1.0 (1.0 - 1.0) |
| **Dialysis or Death at Day 5** | **LASSO** | 0.86 (0.86 - 0.87) | 0.45 (0.44 - 0.45) | 0.9 (0.89 - 0.9) | 0.86 (0.86 - 0.86) | 0.3 (0.3 - 0.31) | 0.99 (0.99 - 0.99) |
| **Logistic**  **Regression** | 0.77 (0.77 - 0.78) | 0.31 (0.31 - 0.32) | 0.85 (0.84 - 0.86) | 0.77 (0.76 - 0.77) | 0.19 (0.19 - 0.2) | 0.99 (0.99 - 0.99) |
| **Random**  **Forest** | 0.85 (0.85 - 0.86) | 0.43 (0.42 - 0.44) | 0.88 (0.87 - 0.88) | 0.85 (0.85 - 0.86) | 0.29 (0.28 - 0.3) | 0.99 (0.99 - 0.99) |
| **XGB**  **Imputed** | 0.87 (0.87 - 0.88) | 0.47 (0.46 - 0.48) | 0.87 (0.87 - 0.88) | 0.87 (0.87 - 0.88) | 0.33 (0.32 - 0.34) | 0.99 (0.99 - 0.99) |
| **XGB**  **Not Imputed** | 0.91 (0.91 - 0.91) | 0.56 (0.55 - 0.56) | 0.87 (0.87 - 0.88) | 0.91 (0.91 - 0.92) | 0.42 (0.41 - 0.43) | 0.99 (0.99 - 0.99) |
| **Dialysis or Death at Day 7** | **LASSO** | 0.88 (0.87 - 0.88) | 0.53 (0.53 - 0.54) | 0.84 (0.84 - 0.85) | 0.88 (0.88 - 0.88) | 0.4 (0.39 - 0.41) | 0.98 (0.98 - 0.99) |
| **Logistic**  **Regression** | 0.77 (0.77 - 0.78) | 0.37 (0.37 - 0.37) | 0.81 (0.8 - 0.81) | 0.77 (0.77 - 0.77) | 0.24 (0.24 - 0.24) | 0.98 (0.98 - 0.98) |
| **Random**  **Forest** | 0.86 (0.86 - 0.86) | 0.5 (0.5 - 0.51) | 0.85 (0.84 - 0.85) | 0.86 (0.86 - 0.86) | 0.36 (0.36 - 0.37) | 0.99 (0.98 - 0.99) |
| **XGB**  **Imputed** | 0.87 (0.87 - 0.88) | 0.53 (0.52 - 0.54) | 0.84 (0.84 - 0.85) | 0.88 (0.87 - 0.88) | 0.4 (0.39 - 0.41) | 0.98 (0.98 - 0.98) |
| **XGB**  **Not Imputed** | 0.89 (0.89 - 0.9) | 0.57 (0.57 - 0.58) | 0.84 (0.84 - 0.85) | 0.9 (0.89 - 0.9) | 0.45 (0.44 - 0.46) | 0.99 (0.98 - 0.99) |
| **External Validation**  **(OH)** | **Dialysis or Death at Day 1** | **LASSO** | 0.79 | 0.09 | 0.67 | 0.79 | 0.05 | 0.99 |
| **Logistic**  **Regression** | 0.7 | 0.07 | 0.73 | 0.70 | 0.04 | 0.99 |
| **Random**  **Forest** | 0.89 | 0.17 | 0.72 | 0.89 | 0.10 | 0.99 |
| **XGB**  **Imputed** | 0.79 | 0.12 | 0.87 | 0.79 | 0.06 | 1.00 |
| **XGB**  **Not Imputed** | 0.87 | 0.18 | 0.87 | 0.87 | 0.10 | 1.00 |
| **Dialysis or Death at Day 3** | **LASSO** | 0.85 | 0.34 | 0.74 | 0.86 | 0.22 | 0.98 |
| **Logistic**  **Regression** | 0.67 | 0.2 | 0.82 | 0.66 | 0.11 | 0.99 |
| **Random**  **Forest** | 0.76 | 0.26 | 0.80 | 0.76 | 0.15 | 0.99 |
| **XGB**  **Imputed** | 0.78 | 0.28 | 0.84 | 0.78 | 0.17 | 0.99 |
| **XGB**  **Not Imputed** | 0.79 | 0.31 | 0.90 | 0.79 | 0.18 | 0.99 |
| **Dialysis or Death at Day 5** | **LASSO** | 0.75 | 0.4 | 0.87 | 0.74 | 0.26 | 0.98 |
| **Logistic**  **Regression** | 0.79 | 0.37 | 0.67 | 0.80 | 0.26 | 0.96 |
| **Random**  **Forest** | 0.77 | 0.39 | 0.79 | 0.76 | 0.26 | 0.97 |
| **XGB**  **Imputed** | 0.8 | 0.43 | 0.79 | 0.80 | 0.29 | 0.97 |
| **XGB**  **Not Imputed** | 0.79 | 0.43 | 0.84 | 0.79 | 0.29 | 0.98 |
| **Dialysis or Death at Day 7** | **LASSO** | 0.81 | 0.51 | 0.80 | 0.81 | 0.38 | 0.97 |
| **Logistic**  **Regression** | 0.78 | 0.45 | 0.68 | 0.80 | 0.33 | 0.94 |
| **Random**  **Forest** | 0.74 | 0.44 | 0.78 | 0.74 | 0.31 | 0.96 |
| **XGB**  **Imputed** | 0.78 | 0.48 | 0.79 | 0.78 | 0.34 | 0.96 |
| **XGB**  **Not Imputed** | 0.8 | 0.51 | 0.82 | 0.79 | 0.37 | 0.97 |