



2025 Risk Model Recalibration

PC4

September 15, 2025 | Samantha Ryan | Data Scientist



Contents

Overview	3
Historic Implementation	4
2023 Recalibration	4
Severity-Adjusted POMV Model (2023)	4
2025 Recalibration	5
Requirements	5
Methods	5
Training Models	5
Evaluation Methods	6
Move to Production	6
Model Evaluation	7
Medical Risk Adjusted Metrics	7
CICU Postoperative Quality Risk Adjusted Metrics	9
Surgical Program Adjusted Metrics	11
Conclusion	17



Overview

This report outlines PC4's risk adjusted models in the context of a regular recalibration. The last recalibration started in 2022 and completed in April 2023. The goal of this project is to recalibrate (retrain) all models with PC4's most current data. PC4's risk adjusted models requiring recalibration are outlined below.

Risk Adjusted Models		
Report Name	Report Section	Variable Name
FTR	Medical Risk Adjusted Metrics	flg_cicu_ftr
FTR-major	Medical Risk Adjusted Metrics	flg_cicu_ftr_major
Medical Cardiac Arrest	Medical Risk Adjusted Metrics	flg_med_any_carrest
CICU Medical Mortality	Medical Risk Adjusted Metrics	flg_med_cicudeath
Postoperative Severity-Adjusted CICU LOS	CICU Postoperative Quality Risk Adjusted Metrics	val_post_sx_cicu_los_prepost
Postoperative Cardiac Arrest	CICU Postoperative Quality Risk Adjusted Metrics	flg_surg_post2h_carrest
CICU Surgical Mortality	CICU Postoperative Quality Risk Adjusted Metrics	flg_cicudeathyn
Severity-Adjusted POMV	CICU Postoperative Quality Risk Adjusted Metrics	val_cicu_pomv_hrs
Postoperative Duration of Mechanical Ventilation	Surgical Program Risk Adjusted Metrics	val_post_sx_vent_los_le21
All Surgical Cardiac Arrest	Surgical Program Risk Adjusted Metrics	flg_surg_any_carrest
Postop CICU LOS	Surgical Program Risk Adjusted Metrics	val_post_sx_cicu_los
Postoperative Infection	Surgical Program Risk Adjusted Metrics	flg_sx_cmp_infection
Postoperative Major Complication	Surgical Program Risk Adjusted Metrics	flg_sx_cmp_major
Postoperative ECMO Rate	Surgical Program Risk Adjusted Metrics	flg_sx_cmp_mcs_ecmo
Postoperative Overall Complication	Surgical Program Risk Adjusted Metrics	flg_sx_cmp_overall
Postoperative Unplanned Reintervention (cath or surgery) v1-v2	Surgical Program Risk Adjusted Metrics	flg_sx_cmp_reopunplan
Postoperative Seizure	Surgical Program Risk Adjusted Metrics	flg_sx_cmp_seizure
Postoperative Unplanned Reintervention (cath or surgery) v3	Surgical Program Risk Adjusted Metrics	flg_sx_cmp_unplan_any_v3
Failure-to-Rescue (FTR)	Surgical Program Risk Adjusted Metrics	flg_sx_ftr



Failure-to-Rescue After Major Complication (FTR-major)	Surgical Program Risk Adjusted Metrics	flg_sx_ftr_major
Surgical In-Hospital Mortality	Surgical Program Risk Adjusted Metrics	flg_sx_hosp_mortality
Postoperative Stroke or Intracranial Hemorrhage	Surgical Program Risk Adjusted Metrics	flg_sx_stroke_hem
Postoperative Hospital LOS	Surgical Program Risk Adjusted Metrics	val_post_sx_hospital_los

Historic Implementation

2023 Recalibration

The last recalibration for PC4 risk models occurred late 2022 and early 2023. The reference table containing model coefficients that is currently in place to generate predicted (expected) values was first updated in prod on October 13, 2022, and contains the recalibrated coefficients. Some model coefficients were updated on an ad hoc basis from support case resolutions in September 2023¹. It is assumed that the reference table was updated between October 2022 and September 2023 (based on modeling code file modification dates), but exact commits or tickets have not been traced. Coefficients are sourced from files which output as part of the modeling process.

[Historic Project Reference](#)

Supporting documentation from PC4 is in registry UI under Information > [Quality Metrics Descriptions](#)

Severity-Adjusted POMV Model (2023)

The Severity-Adjusted POMV Model was built as a separate project², outside of the 2023 Recalibration. Coefficients are generated as part of the modeling process, and imported into AMx model processor as a reference table to generate predicted values, last updated August 8, 2023.

[Historic Project Reference](#)

[Supporting documentation](#) about the model was provided by PC4

¹ SF 5362

² SF 2899, 3251, 3861



2025 Recalibration

Requirements

There were no major changes to any model inputs, except for the following:

- Addition of **val_statscore_hosp** to the Postoperative Duration of Mechanical Ventilation (val_post_sx_vent_los_le21) model as a candidate/selected independent variable
- Inclusion of the Severity-Adjusted POMV (val_cicu_pomv_hrs) model in recalibration.

Methods

Training Models

Typically, models are trained using one larger dataset (training data) and evaluated on another, smaller, dataset (test data). Training and testing data should be representative of each other, having roughly the same distribution of observed values, independent variables, etc. while also being independent of one another. For example, for training patient level data, models should not be tested and evaluated using the same patients. This ensures models are accurately evaluated for performance and can perform well with new data. In historic implementation of this project, models were recalibrated (re-trained) using the entire corpus of available data in production, and OE ratios and model coefficients were examined to determine acceptable model performance. To improve upon this process, these models were trained on a subset of the PC4 corpus; restricting records to only those with hospitalization admission date before Jan 01, 2025. The remainder of the corpus (records with hospitalization admission date on/after Jan 01, 2025) was used to evaluate performance of the current models (baseline) and compared to the performance of the 2025 recalibrated models.

A summary of this project was presented to the greater PC4 team in August and September 2025 (including PC4 leadership, statisticians, and data specialists). It was suggested by the PC4 team that data should be restricted based on hospital admission *and* discharge dates. Models did not necessitate re-training, as encounters with admit date in 2025 were not included in training to start, and there is no risk of data overlapping in evaluation. Models were evaluated on data where: hospital admission date on/after January 1, 2025 and hospital discharge date on/after January 1, 2025 but before 60 days from the date of evaluation. The logic behind this stems from the way PC4 sites submit data to AMx³.

³ Sites must submit complete records within 30-60 days of the patient's discharge. A complete record is one with all required data elements present in the record, including hospital admission and discharge dates. Upon data inspection, it appears some sites submit as soon as 2 days after the patient's discharge, thus complying with PC4's standards, the 60 day constraint was added to ensure only data received on the right timeline is included for evaluation.



Evaluation Methods

Logistic models were evaluated using a variety of classification statistics and visualizations, including:

- Classification Metrics
 - Confusion Matrix
 - Predicted Classes
 - Precision
 - Recall
 - F1-score
 - Support
- Classifier Distribution of observed and predicted values
- Prediction Distribution with corresponding observed class
- Receiver Operating Characteristic (ROC) Curve with corresponding area under the curve (AUC)

Linear models were evaluated using a variety of metrics and visualizations, including:

- Mean Absolute Error
- Root Mean Squared Error
- Coefficient of Determination (R-squared / R2 Score)
- Observed vs Predicted Scatterplot (with corresponding R2 score)
- Predicted value vs Residual Sigma Scatterplot
- Residual Sigma Distribution

Metrics and visualizations for both baseline and recalibrated model coefficients were inspected for any drastic performance or prediction changes from 2023 to 2025.

Move to Production

The PC4 team is encouraged to review all deliverables, including baseline and recalibrated coefficients, predictions, and suite of model performance metrics/visualizations. Upon approval, models can be re-trained on the entire corpus of PC4 data currently in production (since no significant changes to model performance are expected, based on evaluation). Model coefficient reference tables can be updated in place with recalibrated coefficients.



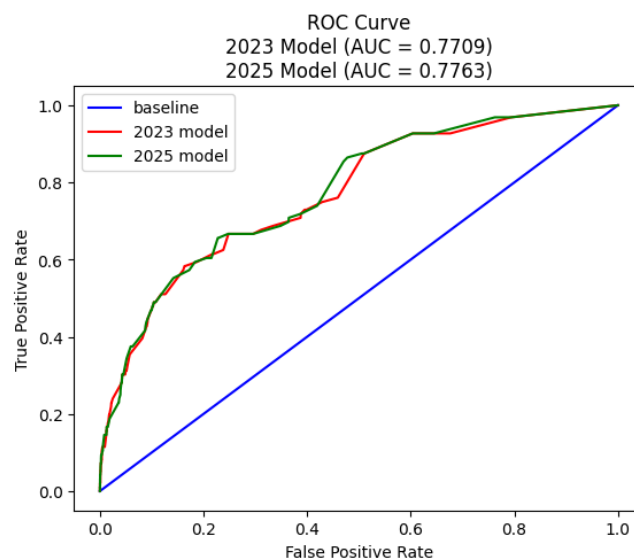
Model Evaluation

Evaluation summaries are provided for each model. Performance is compared to baseline (2023 Recalibration) to determine if any drastic changes are present after recalibration⁴.

Medical Risk Adjusted Metrics

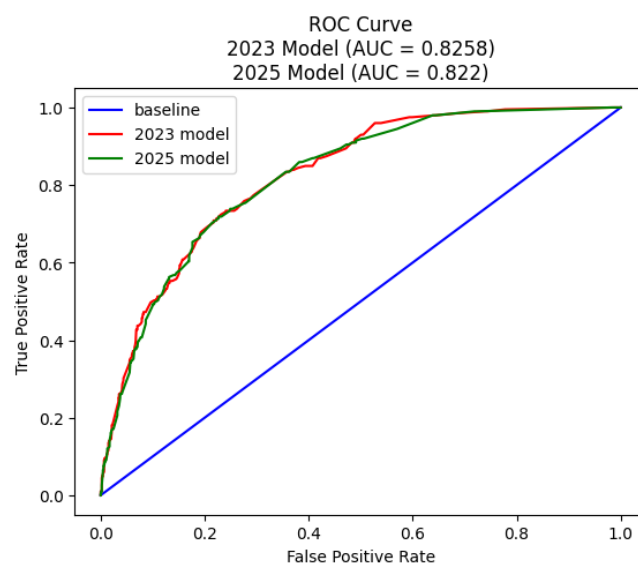
1. FTR (flg_cicu_ftr)

No change to intented variables selected for final model. Minor changes to False Negative/False Positive classifications. Minor changes to positive class precision (-0.01) and f1-scores (-0.01). Minor change in model's overall accuracy (-0.01). Slight decrease in OE ratio (-0.02).



2. FTR-Major (flg_cicu_ftr_major)

Slight change to independent variables selected for final model; **flg_cicustart_normalwt** newly excluded in 2025. Minor changes to False Negative/False Positive classifications. Slight improvement in positive class precision (-0.01) and f1-score (-0.01). Minor change in model's overall accuracy (-0.01). Slight increase in OE Ratio (+0.067).

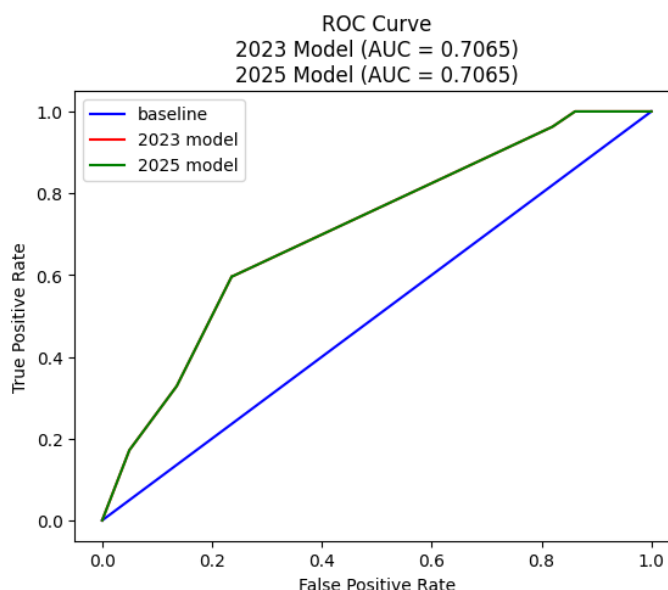


⁴ Additional metrics, plots, and data used to generate evaluations can be found in the project deliverables folder



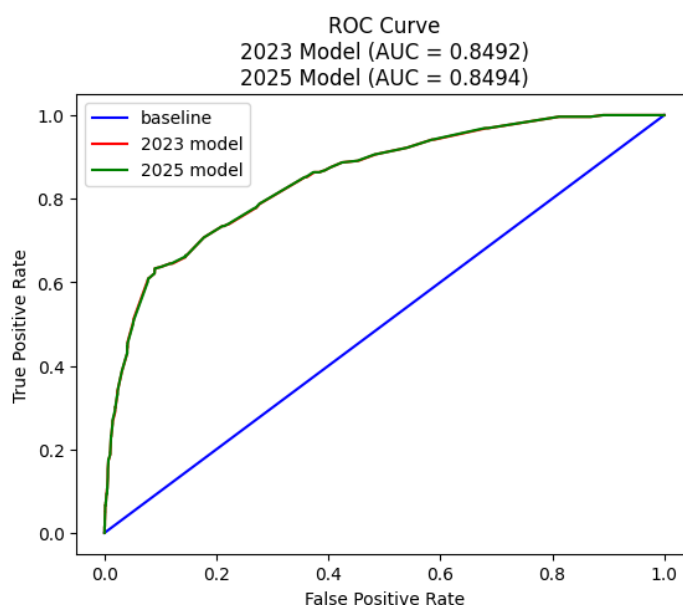
3. Medical Cardiac Arrest (flg_med_any_carrest)

No change to independent variables selected for final model. No changes to classifications. No changes to precision, recall, or f1-scores. No change in model's overall accuracy. Slight increase in OE ratio (+0.02).



4. CICU Medical Mortality (flg_med_cicudeath)

No change to independent variables selected for final model. Minor changes to False Negative/False positive classifications. Slight increase in positive class precision (+0.02), decreased recall (-0.01) and f1-score (-0.02). No change in model's overall accuracy. Slight decrease in OE ratio (-0.015).



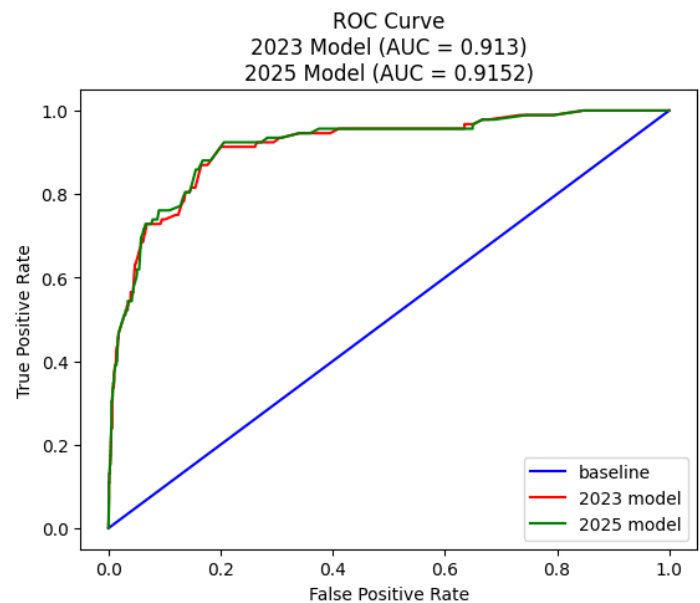


CICU Postoperative Quality Risk Adjusted Metrics

1. CICU Surgical Mortality

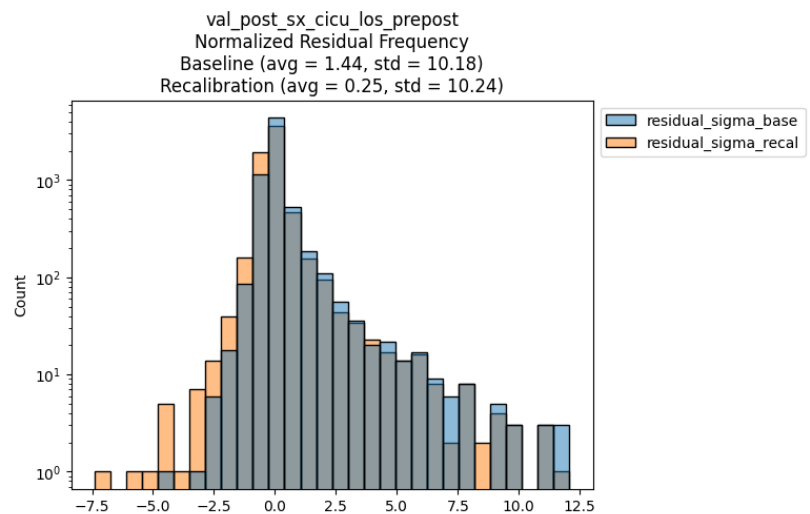
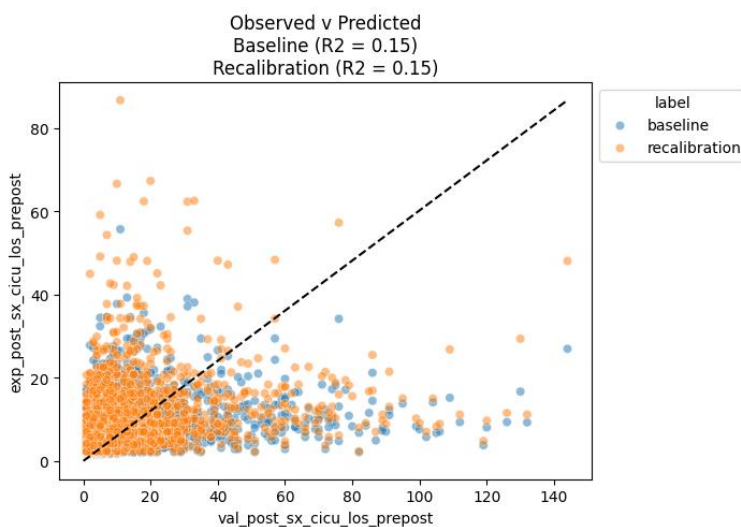
(flg_cicudeathyn)

Slight change to independent variables selected for final model; flg_version3 newly excluded in 2025. No change to False Negative/False Positive classifications. No change in precision, recall, or f1-score. No change in model's overall accuracy. OE ratio decreased (-0.14).



2. Postoperative Severity Adjusted CICU LOS (val_post_sx_cicu_los_prepost)

No change to independent variables selected for final model. No change in R2 score. Slight decrease in RMSE (-0.04) and increase to MAE (+0.39). Based on evaluation plots, there appears to be an increased tendency for overprediction. OE ratio decreased (-0.27).

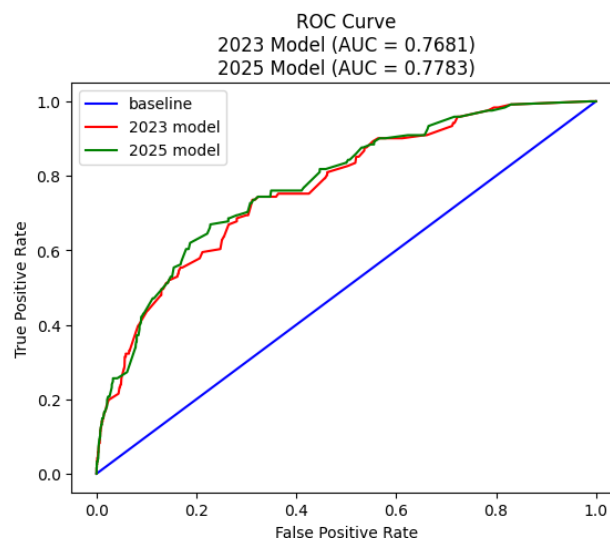




3. Postoperative Cardiac Arrest

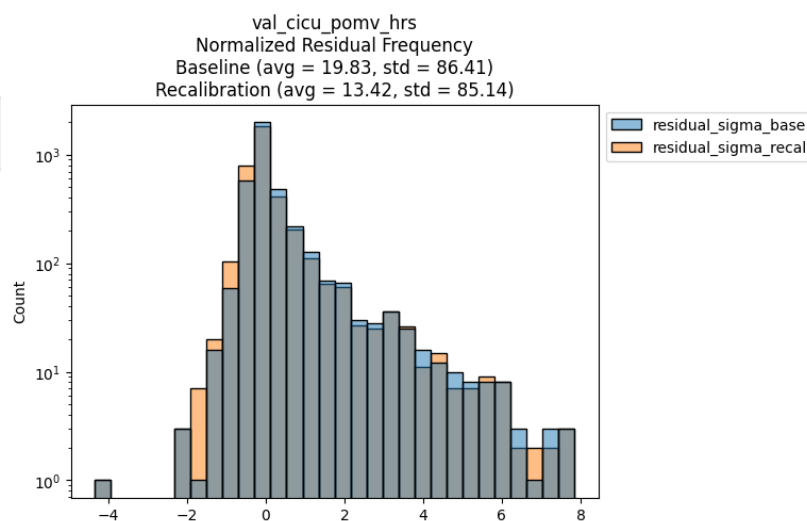
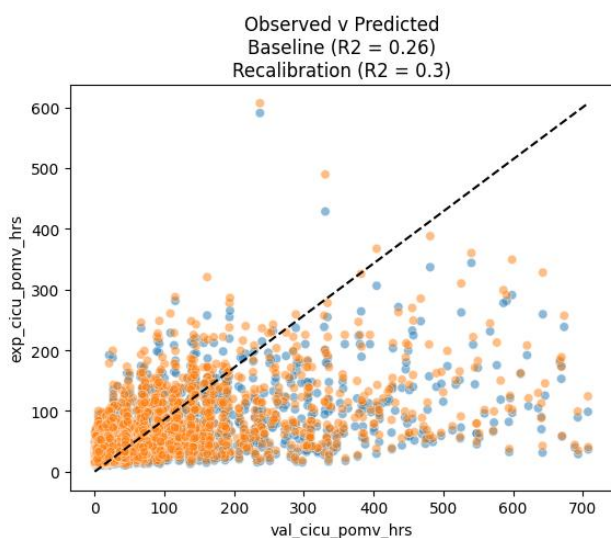
(flg_surg_post2h_carrest)

No change to independent variables selected for final model. No changes to False Negative/False Positive classifications. No change in precision, recall, f1-scores. No change in model's overall accuracy. Slight decrease in **OE ratio (-0.20)**.



4. Severity Adjusted POMV (val_cicu_pomv_hrs)

No change to independent variables selected for final model. R2 score increased (+0.04), and decrease in RMSE (-2.47). Slight increase in MAE (+0.46). Based on the observed v predicted scatterplot, there seems to be a slight increase in overprediction. **OE ratio decreased (+0.40)**.





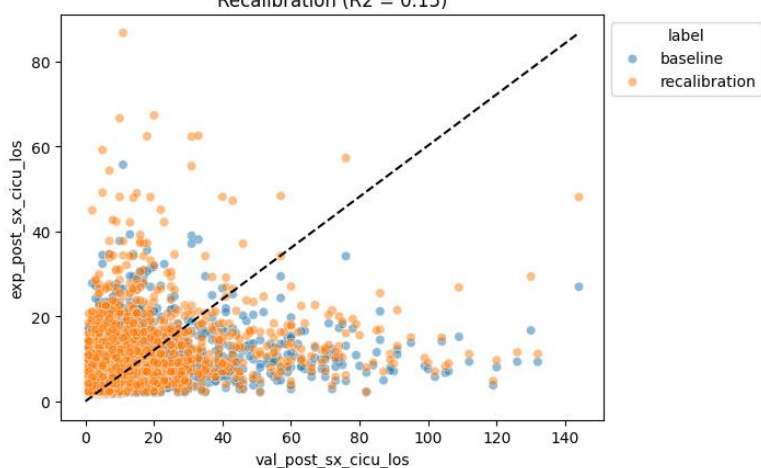
Surgical Program Adjusted Metrics

1. Postop CICU LOS (val_post_sx_cicu_los)

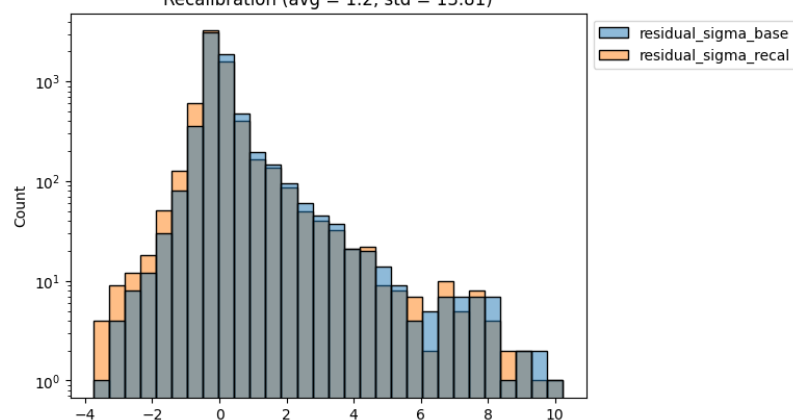
Note that this model uses the same coefficients as the Postop Severity Adjusted CICU LOS model and thus evaluates the same. The summary below is the same as the Postop Severity Adjusted CICU LOS model.

No change to independent variables selected for final model. No change in R2 score. Slight decrease in RMSE (-0.04) and increase to MAE (+0.39). Based on evaluation plots, there appears to be an increased tendency for overprediction. OE ratio decreased (-0.27).

Observed v Predicted
Baseline (R2 = 0.15)
Recalibration (R2 = 0.15)



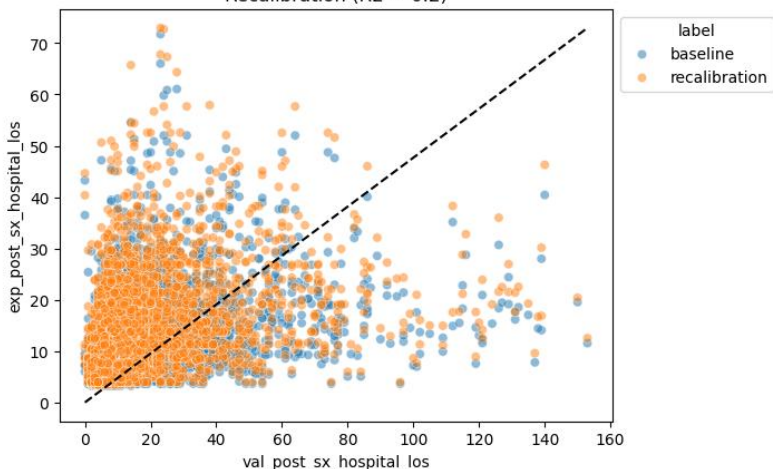
val_post_sx_hospital_los
Normalized Residual Frequency
Baseline (avg = 2.63, std = 13.83)
Recalibration (avg = 1.2, std = 13.81)



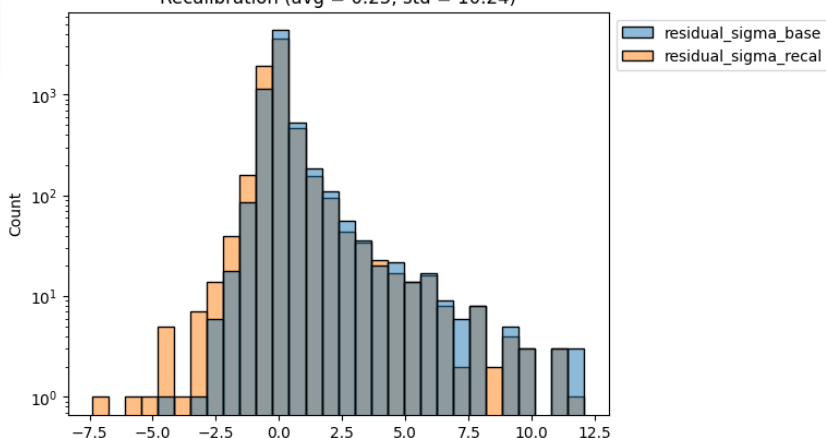
2. Postoperative Hospital LOS (val_post_sx_hospital_los)

No change to independent variables selected for final model. Increase in R2 score (+0.02). Decreased RMSE (-0.22) and increased MAE (+0.29). Predictions appear to be similar after recalibration. OE ratio decreased (-0.197).

Observed v Predicted
Baseline (R2 = 0.18)
Recalibration (R2 = 0.2)



val_post_sx_cicu_los
Normalized Residual Frequency
Baseline (avg = 1.44, std = 10.18)
Recalibration (avg = 0.25, std = 10.24)



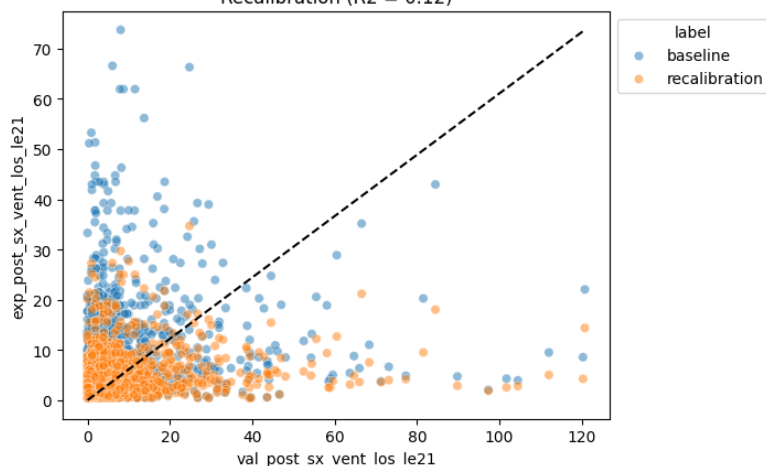


3. Postoperative Duration of Mechanical Ventilation (val_post_sx_vent_los_le21)

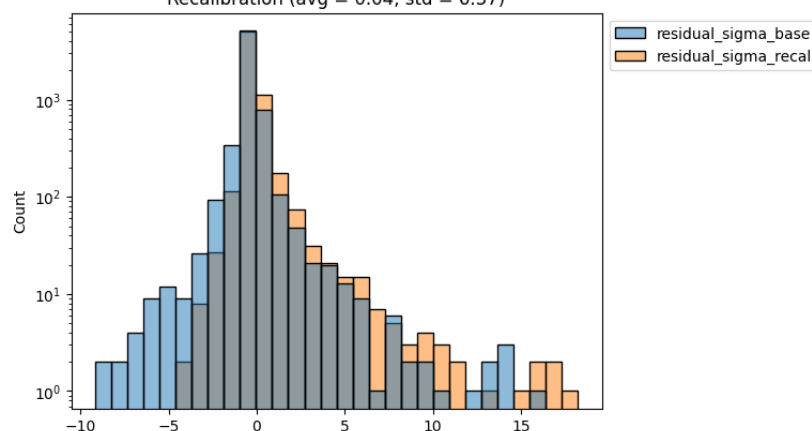
Addition of val_statscore_hosp as candidate and selected independent variable for 2025.

Increased R2 (+0.28), and decreased RMSE (-0.94) and MAE (-0.89). Based on plots, it seems the baseline model has some large overpredictions compared to the recalibrated model. OE ratio increased (+0.55).

Observed v Predicted
Baseline (R2 = -0.16)
Recalibration (R2 = 0.12)



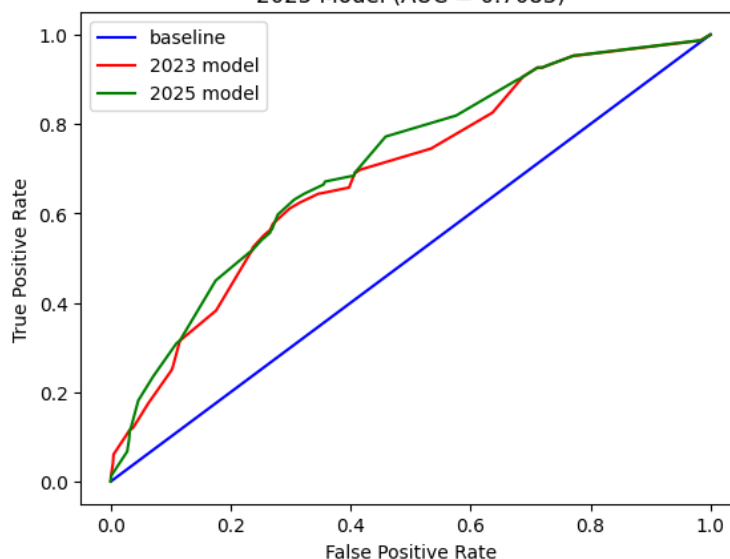
val_post_sx_vent_los_le21
Normalized Residual Frequency
Baseline (avg = -1.39, std = 7.17)
Recalibration (avg = 0.04, std = 6.37)



4. All Surgical Cardiac Arrest (flg_surg_any_carrest)

No change to independent variables selected for final model. No changes to False Negative/False positive classifications. No changes to precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.26).

ROC Curve
2023 Model (AUC = 0.6887)
2025 Model (AUC = 0.7083)

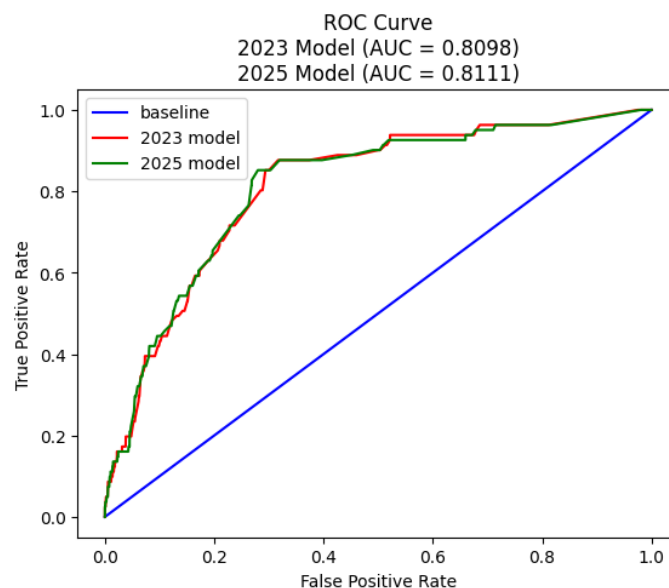




5. Postoperative Infection

(flg_sx_cmp_infection)

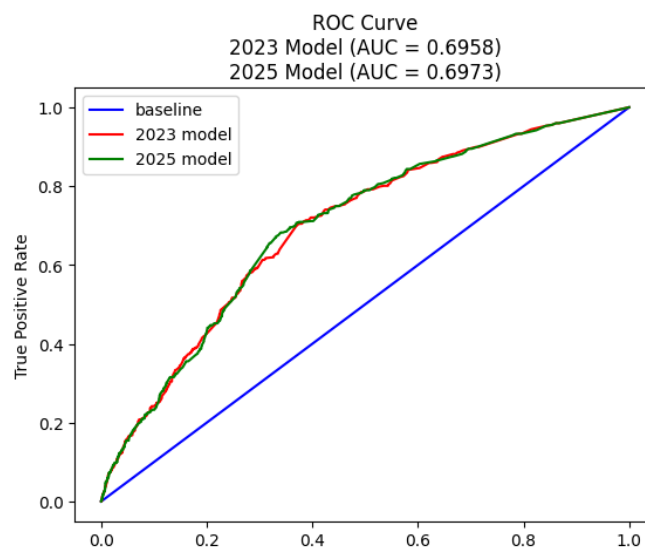
No change to independent variables selected for final model. No changes to predicted classes. No change in precision, recall, f1-scores. No change in model's overall accuracy. Decrease in OE ratio (-0.42).



6. Postop Major Complication

(flg_sx_cmp_major)

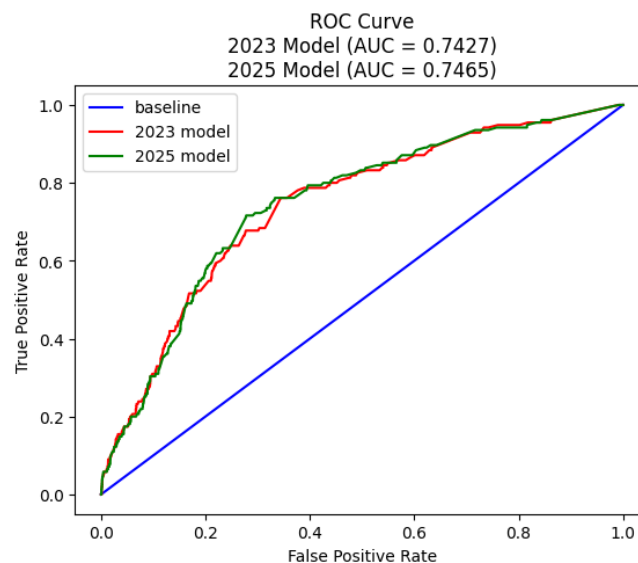
No change to independent variables selected for final model. No changes to False Negative/False positive classifications. No changes in precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.47).



7. Postop ECMO Rate

(flg_sx_cmp_mcs_ecmo)

No change to independent variables selected for final model. No changes to False Negative/False positive classifications. No change in precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.856)

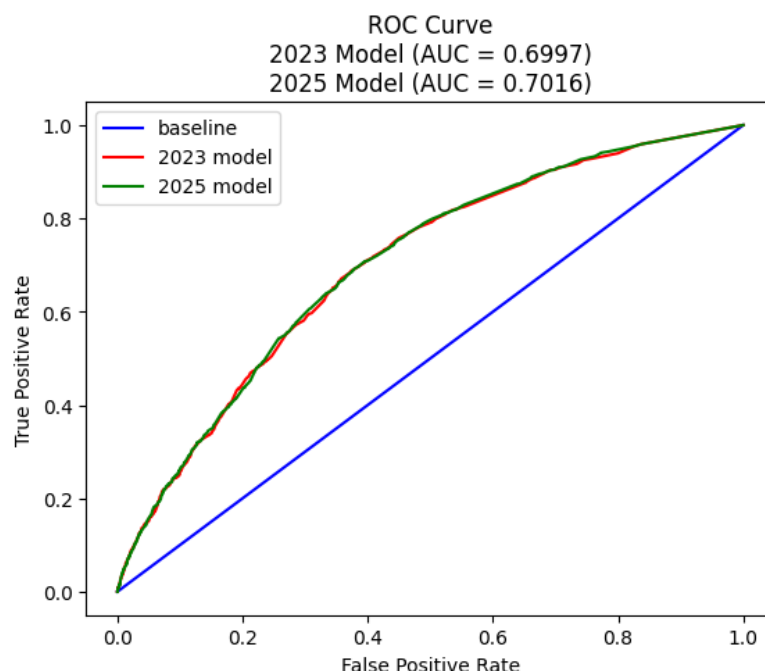




8. Postop Overall Complication

(flg_sx_cmp_overall)

No change to independent variables selected for final model. Changes seen across all positive/negative classifications. Negative class precision increased (+0.01) and decreased in recall (-0.02). Positive class precision decrease (-0.04) and increase in recall (+0.04) and f1-score (+0.06). No change in model's overall accuracy. OE ratio decreased (-0.31).



9. Postoperative Unplanned

Reintervention (cath or surgery) v1-v2 (flg_sx_cmp_reopunplan)

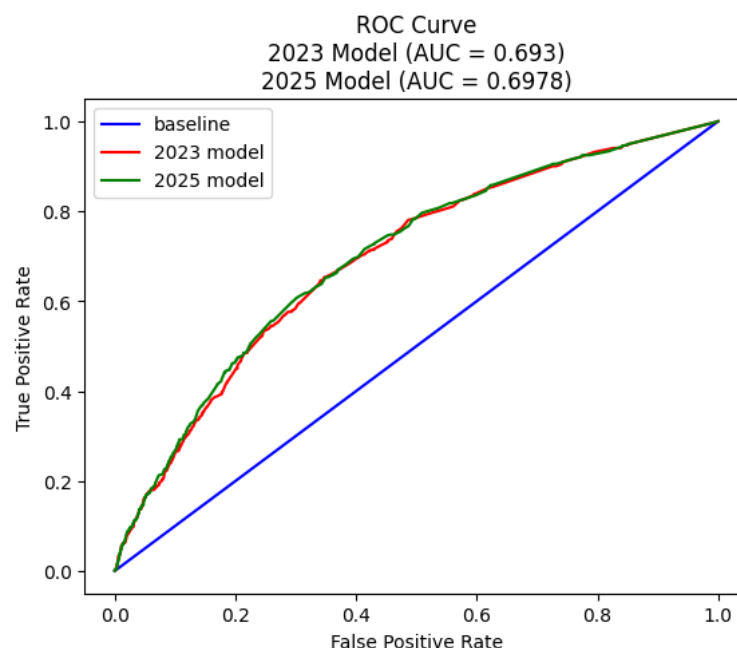
Changes in independent variables selected for final model; e_index_prev_op_cnt and flg_extra_card_anomaly both newly excluded in 2025. There are no observed positive classes present in the test data for this model, thus I was unable to perform evaluation. Coefficient deltas (2023-2025) range from 0.02 to 1. This may or may not be expected due to versioning of dependent variable (flg_sx_cmp_reopunplan). Further data inspection/analysis may be necessary.

10. Postoperative Unplanned

Reintervention (cath or surgery) v3

(flg_sx_cmp_unplan_any_v3)

No changes in independent variables selected for final model. No changes to predicted classes. No changes in precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.50).

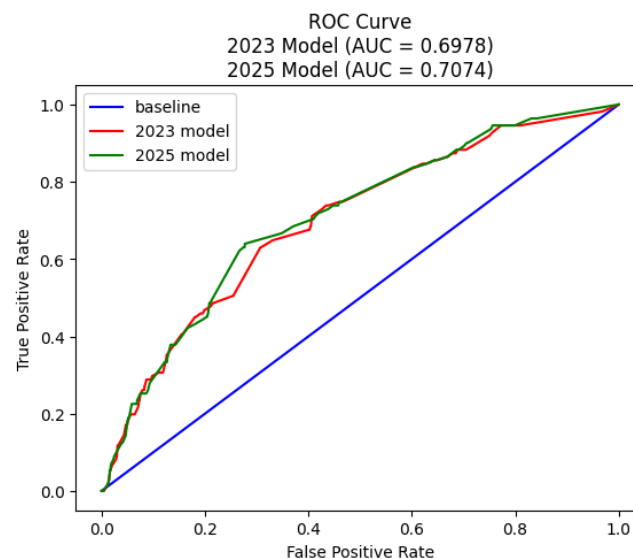




11. Postoperative Seizure

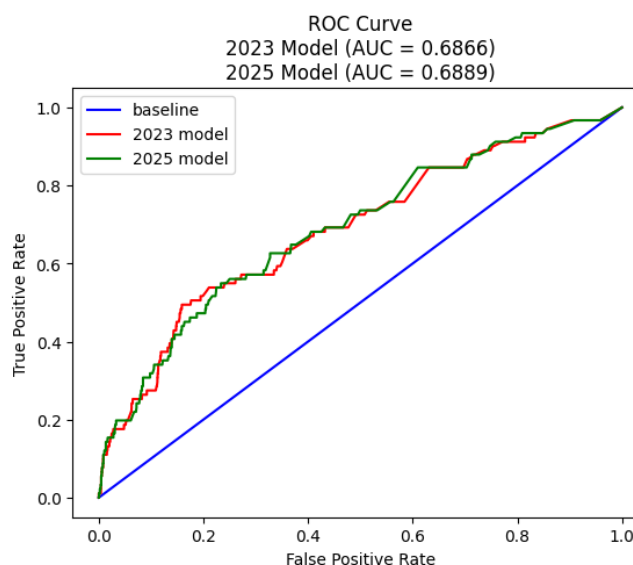
(flg_sx_cmp_seizure)

No change to independent variables selected for final model. No changes to predicted classes. No change in precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.81).



12. Failure-to-Rescue (FTR) (flg_sx_ftr)

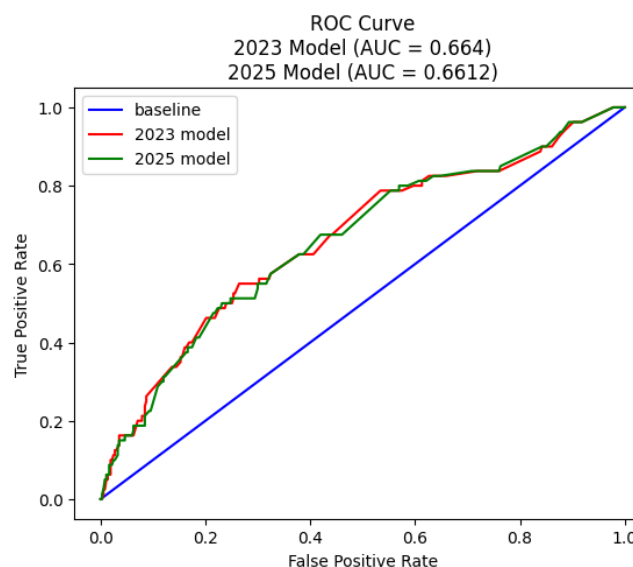
No change to independent variables selected for final model. No changes to predicted classes. No changes to precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.37).



13. Failure-to-Rescue after Major Complication (FTR-Major)

(flg_sx_ftr_major)

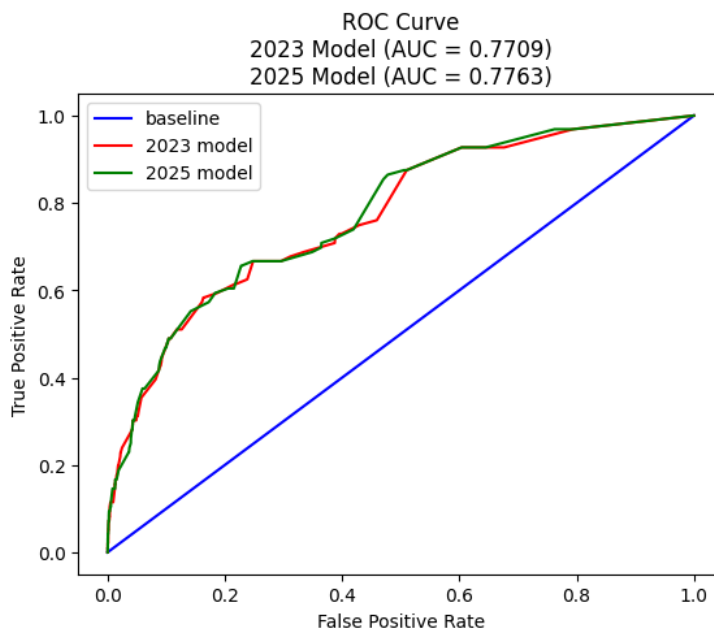
No change to independent variables selected for final model. No changes to predicted classes. No changes to precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.31).





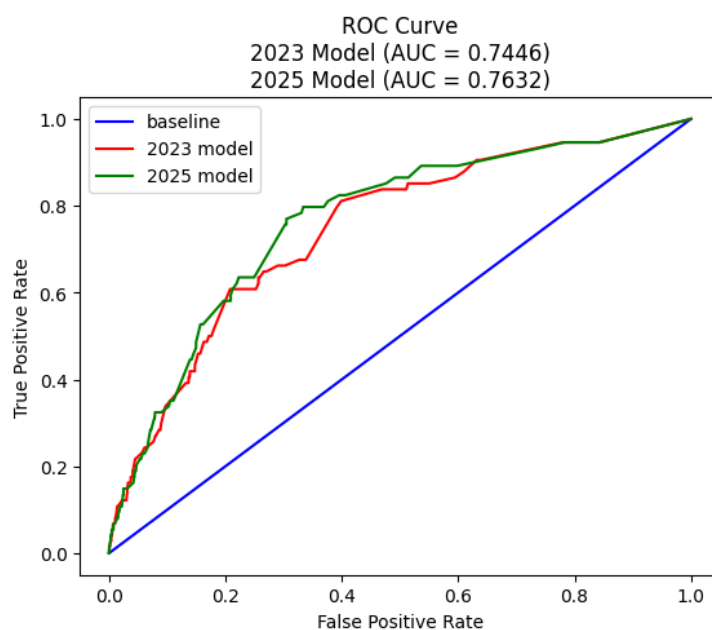
14. Surgical In-Hospital Mortality (flg_sx_hosp_mortality)

No change to independent variables selected for final model. No changes to predicted classes. No changes to precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.58).



15. Postoperative Stroke or Intracranial Hemorrhage (flg_sx_stroke_hem)

Changes in independent variables selected for final model; flg_version3 newly excluded in 2025. No changes to predicted classes. No changes to precision, recall, or f1-scores. No change in model's overall accuracy. OE ratio decreased (-0.98).





Conclusion

2025 recalibration is successful with minimal changes to models and without large, unexpected changes to coefficients ($\Delta < 1$). Models with newly included independent variables include:

- Postoperative Duration of Mechanical Ventilation (val_post_sx_vent_lo21):
val_statscore_hosp

Models with newly excluded independent variables include:

- FTR-Major (flg_cicu_ftr_major): flg_cicustart_normalwt
- CICU Surgical Mortality (flg_cicudeathyn): flg_version3
- Postoperative Unplanned Reintervention (cath or surgery) v1-v2 (flg_sx_cmp_reopunplan):
e_index_prev_op_cnt, flg_extra_card_anomaly
- Postoperative Stroke or Intracranial Hemorrhage (flg_sx_stroke_hem): flg_version3

Comparing predictions resulting from 2023 coefficients versus 2025 coefficients show consistent stability or improvement for all models. One model did not have test data to evaluate on (Postoperative Unplanned Reintervention (cath or surgery) v1-v2). PC4 should review this model's definition and determine if it is appropriate to retire from future recalibration due to versioning. To continue the same cadence in the context of historical recalibration targets the next recalibration in September 2027.

Project deliverables can be found in PC4's sftp folder at [/pc4/dropbox/2025_Model_Recalibration/](#)