Penn Medicine

Utilizing the ACS NSQIP Database to Develop a Novel Artificial Intelligence Model for Prediction of Reoperation Following Surgical Site Infection for Lumbar Spine Surgeries

Jaskeerat Gujral, MSE, FICS (Jr.)

July 10, 2025

Disclosures

No Disclosures

Surgical Site Infections

Pose a morbidity burden, prolonging hospitalization, increasing healthcare costs, and necessitating unplanned re-operations

Lumbar microdiscectomy carries a comparatively lower risk than instrumented spinal fusions

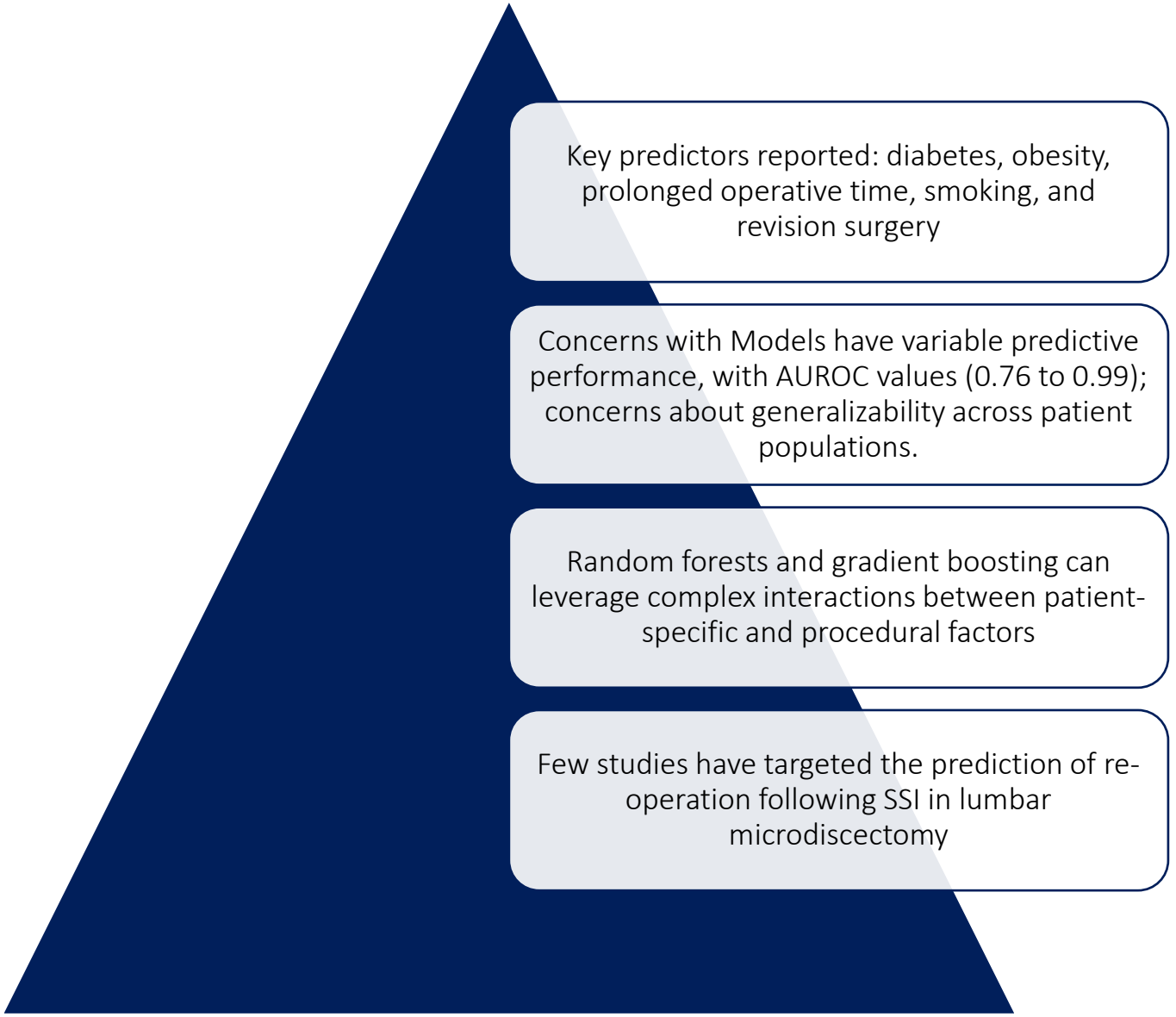
Although re-operation following SSI is a rare event, its consequences are profound

Postoperative SSI remains a clinically significant concern due to its potential to necessitate early re-operation

What is the Need?

There are challenges in predicting rare events, but developing models that accurately differentiate patients at elevated versus lower risk would be crucial for advancing precision medicine in neurosurgery

Past Literature



Goal

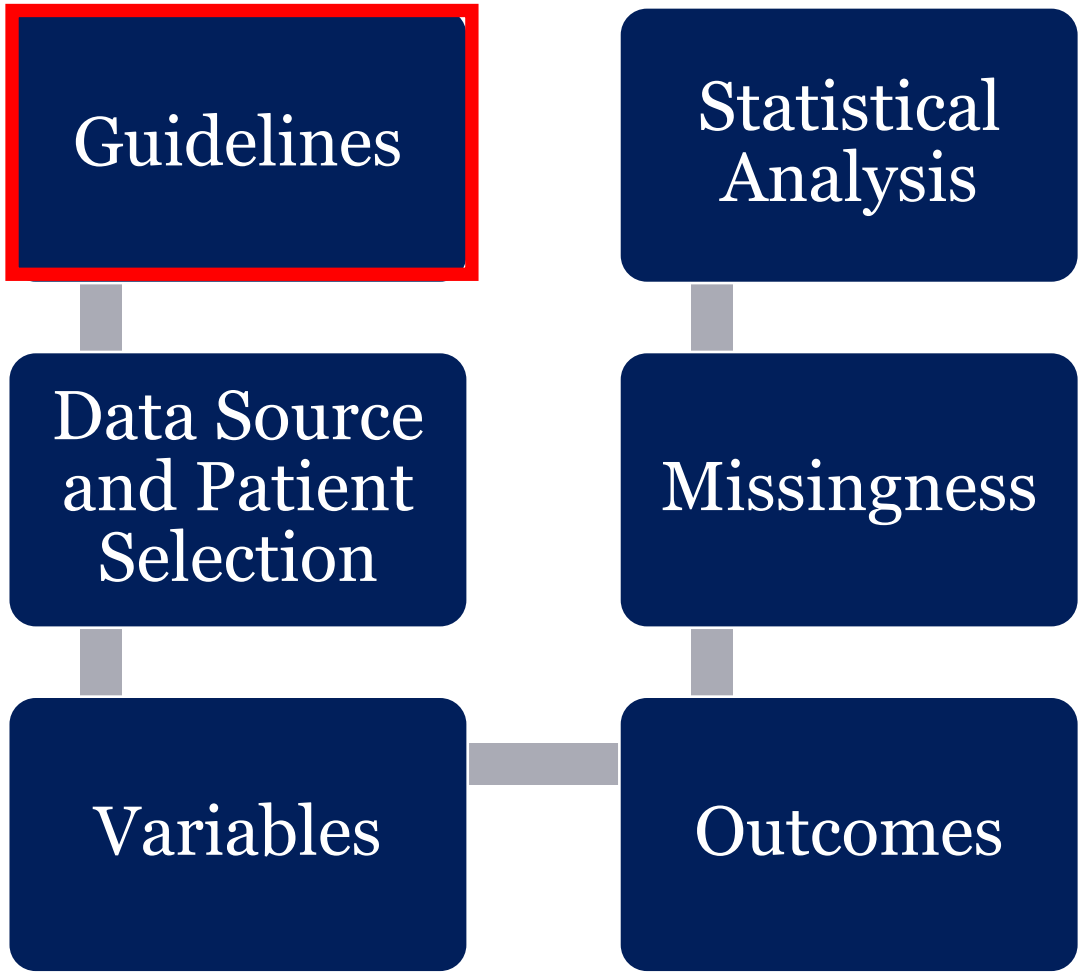
Development and validation of an ML-based predictive model for 30-day re-operation following SSI in lumbar microdiscectomy patients.

Refine risk stratification in neurosurgery and contribute to the broader discourse on rare event prediction.

Variables

Demographics	Lifestyle-Related Factors	Pre-Operative Lab Values	Pre-Existing Co-morbidities	Surgery-Related Factors	Post-Operative Complications
• Age • Sex • Hispanic Ethnicity	• BMI • Smoking status • Diabetes	• WBC • HCT • Platelets • PTT • INR • BUN • Creatine • Albumin	• CHF • Disseminated cancer • Steroid use • Bleeding disorder • COPD • Dialysis • Weight loss • Renal failure • Dyspnea	• Operative times • Transfer from home status • Functional status • ASA score • Transfusion • Specialty	• Superficial SSI • Deep incisional SSI • Organ/Space SSI • Re-admission • Post-operative re-operatoin due to SSI

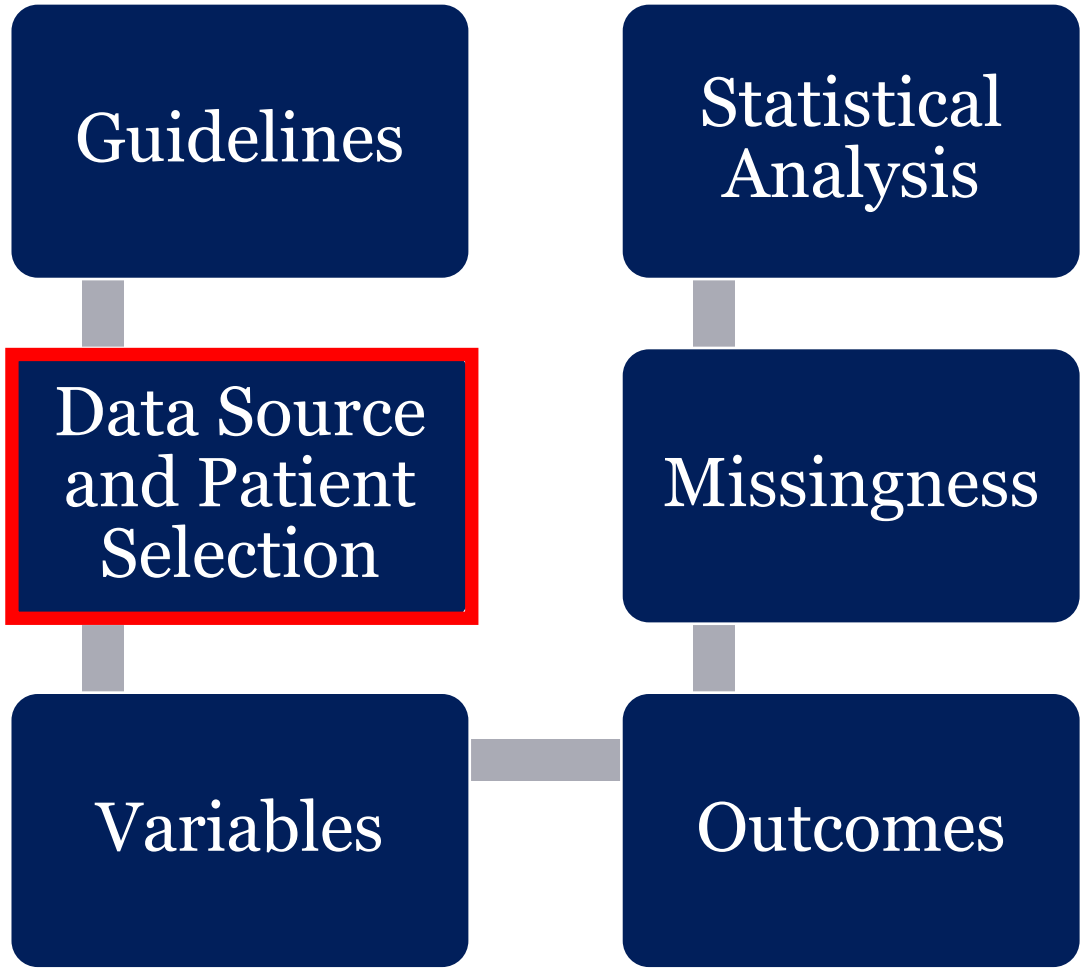
Methodology



Reporting

- Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)
- Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis + Artificial Intelligence (TRIPOD+AI) guidelines

Methodology

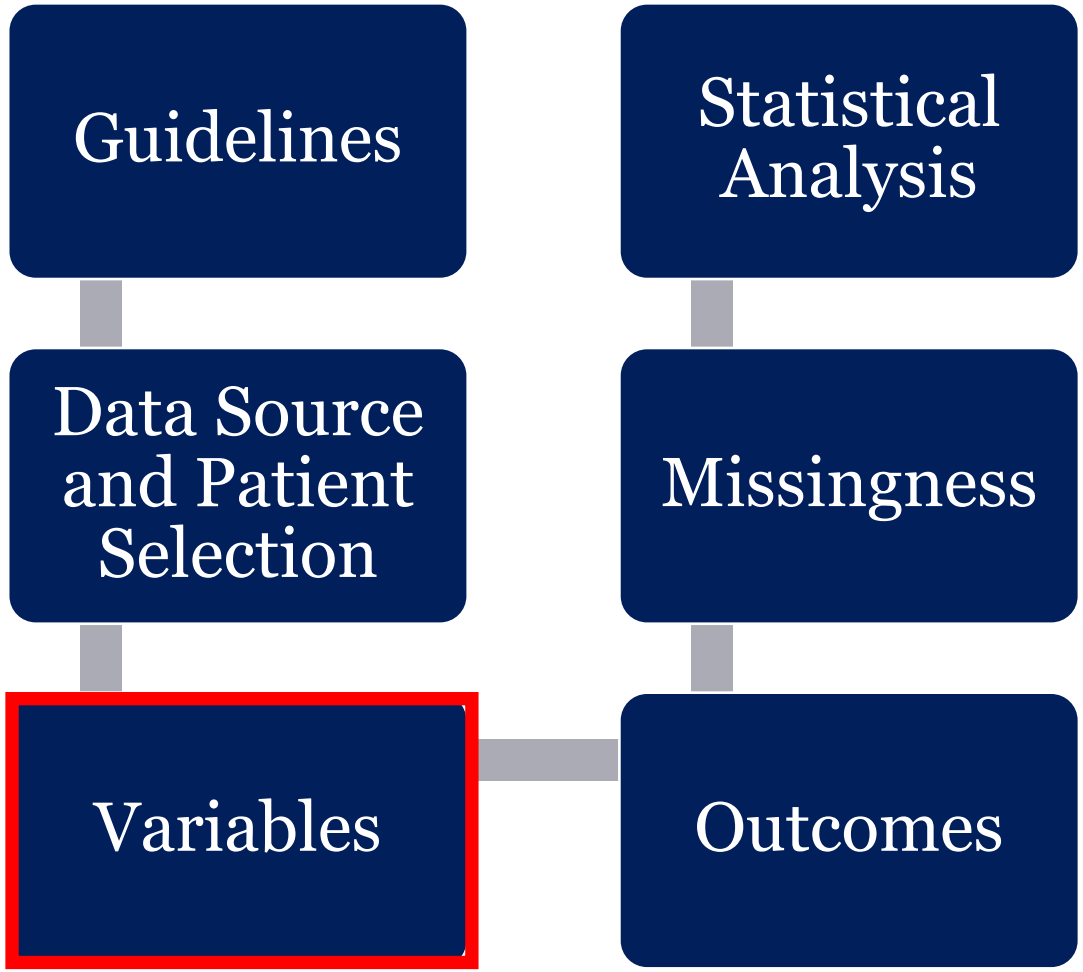


Why Utilize the ACS-NSQIP Database?

- Provides multi-institutional data, which is beneficial for large sample size and increasing generalizability

-Has been validated and shown to have reliable and robust data (Shiloach et al. 2010)

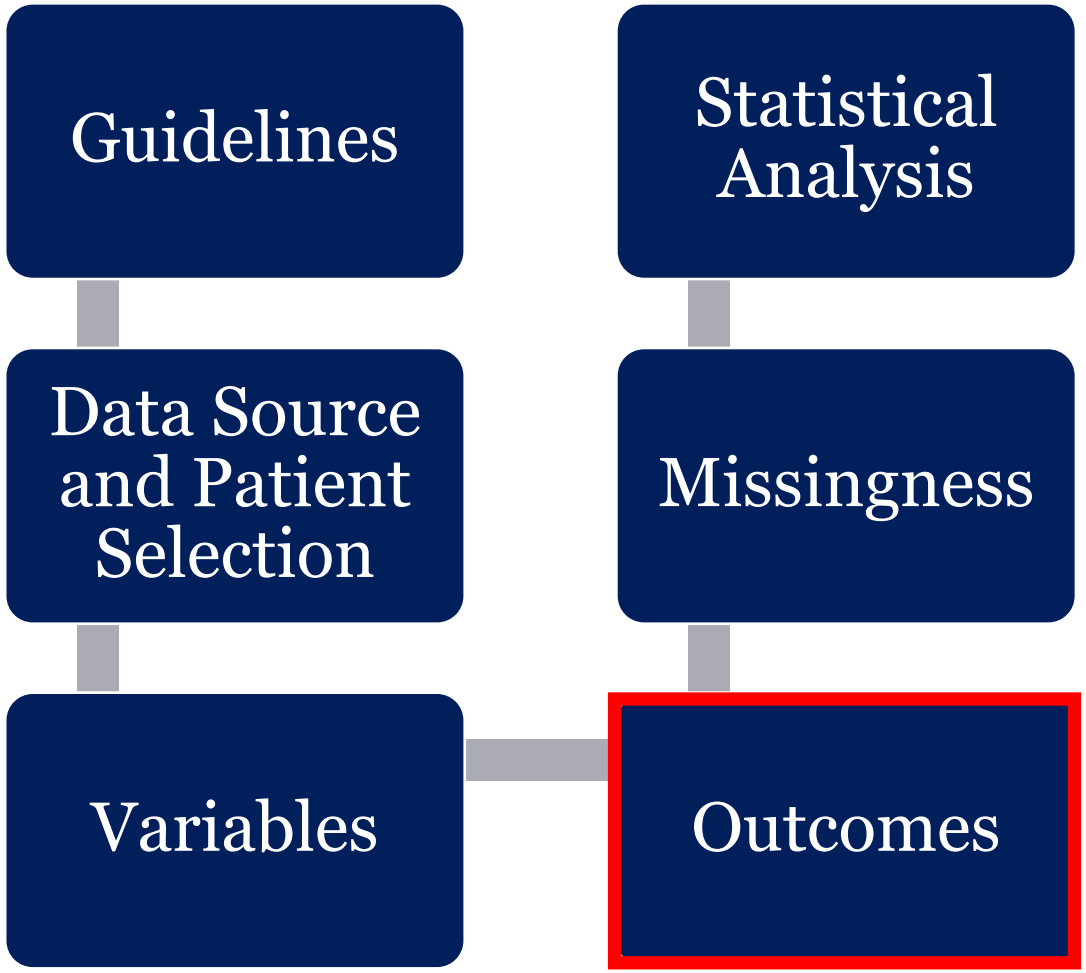
Methodology



Variables

Demographics	Lifestyle-Related Factors	Pre-Operative Lab Values	Pre-Existing Co-morbidities	Surgery-Related Factors
• Age • Sex • Hispanic Ethnicity	• BMI • Smoking status • Diabetes	• WBC • HCT • Platelets • PTT • INR • BUN • Creatine • Albumin	• CHF • Disseminated cancer • Steroid use • Bleeding disorder • COPD • Dialysis • Weight loss • Renal failure • Dyspnea	• Operative times • Transfer from home status • Functional status • ASA score • Transfusion • Specialty

Methodology

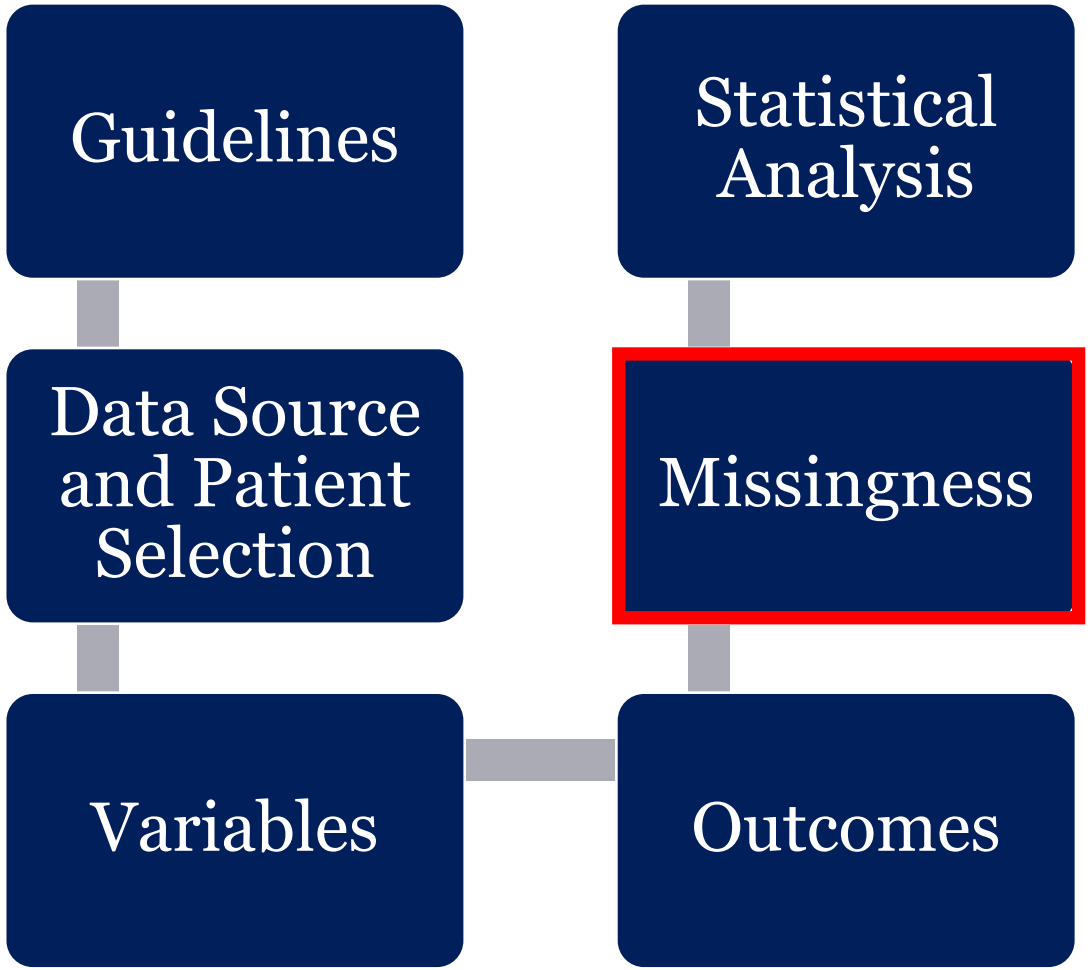


Outcomes

Primary outcome: development and validation of a machine learning algorithm to predict early post-operative re-operation following a post-operative SSI.

Secondary outcomes: interpretability analysis using SHapley Additive Explanations

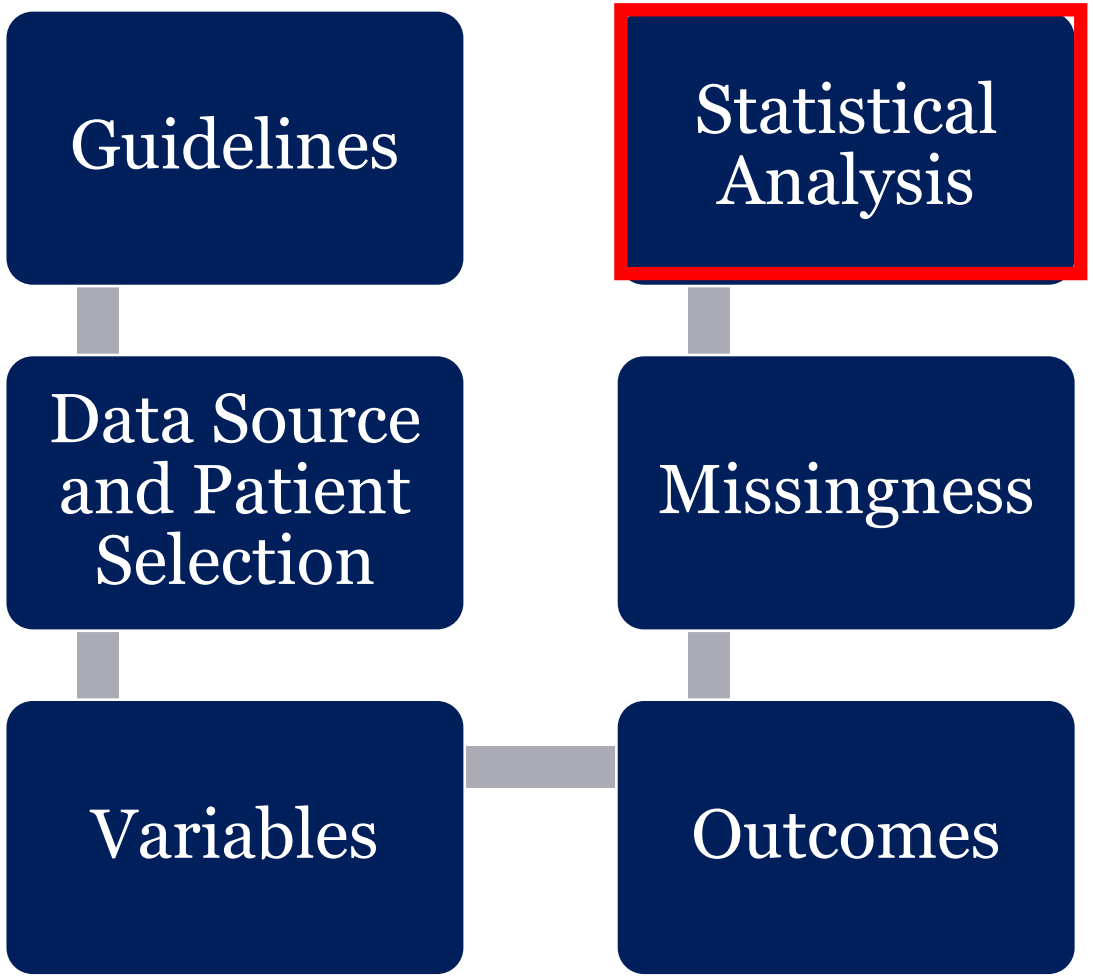
Methodology



Missingness

Prevalence of missing data constituted **under 5% of the final dataset**. Missing values were imputed with multiple imputations using chained equations with 5 imputations.

Methodology



Statistical Analysis

Baseline Characteristics	• Summarized using descriptive statistics
Machine Learning Pipeline	• Nested Cross-Validation • Bayesian Optimization • XGBoost + SMOTE for class imbalance
Model Performance	• Accuracy, Sensitivity (recall), Specificity, PPV, NPV, F1-score, Brier score, AUROC, AU-PRC, and MCC • 95% CI Bootstrapping 10,000 samples
Model Interpretability	• SHAP Analysis

Table 1. Baseline Characteristics

Variable	Total (N=79870)	Control (N=79408)	REOP After SSI (N=462)	*P*-Value
Age (years)	51.4 ± 15.8	51.4 ± 15.8	51.8 ± 15.6	0.636
Gender				
Female	35023 (43.9)	34797 (43.8)	226 (48.9)	0.031
Male	44846 (56.1)	44610 (56.2)	236 (51.1)	0.031
Race				
White	61703 (77.3)	61355 (77.3)	348 (75.3)	0.349
Black or African American	5245 (6.6)	5208 (6.6)	37 (8.0)	0.246
Asian	2041 (2.6)	2035 (2.6)	6 (1.3)	0.117
American Indian or Alaska Native	446 (0.6)	439 (0.6)	7 (1.5)	0.016
Native Hawaiian or Pacific Islander	252 (0.3)	251 (0.3)	1 (0.2)	0.999
Other	11 (0.0)	11 (0.0)		0.999
Hispanic ethnicity	5199 (6.5)	5176 (6.5)	23 (5.0)	0.214
Smoker	17233 (21.6)	17095 (21.5)	138 (29.9)	<0.001
ASA score				
1	8033 (10.1)	8008 (10.1)	25 (5.4)	0.001
2	45714 (57.2)	45492 (57.3)	222 (48.1)	<0.001
3	24924 (31.2)	24730 (31.1)	194 (42.0)	<0.001
4	1114 (1.4)	1094 (1.4)	20 (4.3)	<0.001
5	4 (0.0)	4 (0.0)		0.999
RAI-rev score	15.7 ± 6.8	15.7 ± 6.8	16.1 ± 7.2	0.206
Functional status				
Independent	78587 (98.4)	78144 (98.4)	443 (95.9)	<0.001
Partially Dependent	849 (1.1)	840 (1.1)	9 (1.9)	0.068
Totally Dependent	57 (0.1)	53 (0.1)	4 (0.9)	<0.001
History				
CHF	148 (0.2)	148 (0.2)		0.999
COPD	1950 (2.4)	1927 (2.4)	23 (5.0)	0.001
Medication-requiring hypertension	30458 (38.1)	30236 (38.1)	222 (48.1)	<0.001

Table 1. Baseline Characteristics

Variable	Total (N=79870)	Control (N=79408)	REOP After SSI (N=462)	*P*-Value
Age (years)	51.4 ± 15.8	51.4 ± 15.8	51.8 ± 15.6	0.636
Gender				
Female	35023 (43.9)	34797 (43.8)	226 (48.9)	0.031
Male	44846 (56.1)	44610 (56.2)	236 (51.1)	0.031
Race				
White	61703 (77.3)	61355 (77.3)	348 (75.3)	0.349
Black or African American	5245 (6.6)	5208 (6.6)	37 (8.0)	0.246
Asian	2041 (2.6)	2035 (2.6)	6 (1.3)	0.117
American Indian or Alaska Native	446 (0.6)	439 (0.6)	7 (1.5)	0.016
Native Hawaiian or Pacific Islander	252 (0.3)	251 (0.3)	1 (0.2)	0.999
Other	11 (0.0)	11 (0.0)		0.999
Hispanic ethnicity	5199 (6.5)	5176 (6.5)	23 (5.0)	0.214
Smoker	17233 (21.6)	17095 (21.5)	138 (29.9)	<0.001
ASA score				
1	8033 (10.1)	8008 (10.1)	25 (5.4)	0.001
2	45714 (57.2)	45492 (57.3)	222 (48.1)	<0.001
3	24924 (31.2)	24730 (31.1)	194 (42.0)	<0.001
4	1114 (1.4)	1094 (1.4)	20 (4.3)	<0.001
5	4 (0.0)	4 (0.0)		0.999
RAI-rev score	15.7 ± 6.8	15.7 ± 6.8	16.1 ± 7.2	0.206
Functional status				
Independent	78587 (98.4)	78144 (98.4)	443 (95.9)	<0.001
Partially Dependent	849 (1.1)	840 (1.1)	9 (1.9)	0.068
Totally Dependent	57 (0.1)	53 (0.1)	4 (0.9)	<0.001
History				
CHF	148 (0.2)	148 (0.2)		0.999
COPD	1950 (2.4)	1927 (2.4)	23 (5.0)	0.001
Medication-requiring hypertension	30458 (38.1)	30236 (38.1)	222 (48.1)	<0.001

Table 1. Baseline Characteristics

Steroid use	2808 (3.5)	2778 (3.5)	30 (6.5)	0.001
Bleeding disorder	803 (1.0)	790 (1.0)	13 (2.8)	0.001
Diabetes insulin-dependent	3572 (4.5)	3533 (4.4)	39 (8.4)	<0.001
Diabetes non-insulin	7449 (9.3)	7376 (9.3)	73 (15.8)	<0.001
No diabetes	68849 (86.2)	68499 (86.3)	350 (75.8)	<0.001
Dialysis	121 (0.2)	118 (0.1)	3 (0.6)	0.034
Disseminated cancer	153 (0.2)	150 (0.2)	3 (0.6)	0.060
Ascites	8 (0.0)	7 (0.0)	1 (0.2)	0.045
Pre-operative labs				
Albumin (g/dL)	4.2 ± 0.5	4.2 ± 0.5	4.0 ± 0.6	<0.001
Alkaline phosphatase (U/L)	72.8 ± 27.8	72.8 ± 27.8	79.5 ± 32.0	0.003
Bilirubin (mg/dL)	0.6 ± 0.4	0.6 ± 0.4	0.5 ± 0.3	0.151
BUN (mg/dL)	16.1 ± 6.6	16.1 ± 6.6	16.7 ± 7.8	0.162
Creatinine (mg/dL)	0.9 ± 0.4	0.9 ± 0.4	0.9 ± 0.6	0.474
HCT (%)	42.0 ± 4.3	42.0 ± 4.3	41.4 ± 5.2	0.012
INR	1.0 ± 0.2	1.0 ± 0.2	1.0 ± 0.1	0.853
Platelets ($10^3/\mu L$)	251.0 ± 67.5	251.0 ± 67.5	255.5 ± 79.2	0.247
PTT (seconds)	28.9 ± 4.3	28.9 ± 4.3	29.7 ± 4.5	0.011
SGOT (U/L)	25.2 ± 19.8	25.2 ± 19.8	24.7 ± 14.4	0.608
Sodium (mmol/L)	139.3 ± 5.7	139.4 ± 5.7	139.0 ± 3.3	0.022
WBC ($10^3/\mu L$)	7.8 ± 2.7	7.8 ± 2.7	8.3 ± 3.2	<0.001
Operative time	92.1 ± 54.9	91.9 ± 54.7	120.4 ± 79.6	<0.001

ASA: American Society of Anesthesiologists, BUN: Blood urea nitrogen, CHF: Congestive heart failure, COPD: Chronic obstructive pulmonary disease, HCT: Hematocrit, INR: International normalized ratio, PTT: Partial thromboplastin time, RAI-Rev: Revised Risk Analysis Index, SGOT: Serum glutamic-oxaloacetic transaminase, SSI: Surgical site infection, WBC: White blood cell count.

Table 2. Post-operative Complications and Outcomes

Variable	Total (N=79870)	Control (N=79408)	REOP After SSI (N=462)	*P*-Value
Any post-operative SSI	1059 (1.3)	597 (0.8)	462 (100.0)	<0.001
Post-operative superficial infection	607 (0.8)	473 (0.6)	134 (29.0)	<0.001
Post-operative deep incisional SSI	255 (0.3)	58 (0.1)	197 (42.6)	<0.001
Post-operative organ/space SSI	208 (0.3)	71 (0.1)	137 (29.7)	<0.001
Any re-admission	2613 (3.3)	2197 (2.8)	416 (90.0)	<0.001
Suspected reason superficial SSI	153 (0.2)	60 (0.1)	93 (20.1)	<0.001
Suspected reason deep incisional SSI	169 (0.2)	38 (0.0)	131 (28.4)	<0.001
Suspected reason organ/space SSI	103 (0.1)	24 (0.0)	79 (17.1)	<0.001
Suspected reason SSI	425 (0.5)	122 (0.2)	303 (65.6)	<0.001
Post-operative re-operation	1735 (2.2)	1273 (1.6)	462 (100.0)	<0.001
After post-operative SSI	462 (0.6)	-	462 (100.0)	<0.001

SSI: Surgical site infection, REOP: Reoperation.

Table 3. Performance metrics of the nested cross-validated, feature-engineered, and Bayesian-optimized model evaluated at different classification thresholds, including default, Youden's Index, F1-optimized, and MCC-optimized, with 95% confidence intervals derived from 10,000 bootstrapped resamples

Metric	Default (0.5)	Youden's Index (0.148)	F1-optimized (0.388)	MCC-optimized (0.278)
Accuracy	0.993 (0.992 to 0.993)	0.993 (0.992 to 0.993)	0.993 (0.992 to 0.993)	0.993 (0.992 to 0.993)
Sensitivity	0.924 (0.815 to 0.989)	0.965 (0.931 to 0.996)	0.948 (0.905 to 0.983)	0.985 (0.959 to 1.000)
Specificity	0.993 (0.993 to 0.994)	0.993 (0.993 to 0.993)	0.993 (0.993 to 0.993)	0.993 (0.992 to 0.993)
PPV	0.437 (0.431 to 0.455)	0.436 (0.428 to 0.443)	0.437 (0.429 to 0.445)	0.437 (0.429 to 0.446)
NPW	1.000 (0.999 to 1.000)	1.000 (1.000 to 1.000)	1.000 (0.999 to 1.000)	1.000 (1.000 to 1.000)
F1-score	0.593 (0.564 to 0.619)	0.600 (0.588 to 0.612)	0.598 (0.583 to 0.611)	0.605 (0.596 to 0.615)
Brier score	0.006 (0.006 to 0.007)	0.006 (0.006 to 0.007)	0.006 (0.006 to 0.007)	0.006 (0.006 to 0.007)
AUCROC	0.996 (0.996 to 0.996)	0.996 (0.996 to 0.996)	0.996 (0.996 to 0.996)	0.996 (0.996 to 0.996)
AU-PRC	0.426 (0.385 to 0.471)	0.426 (0.402 to 0.452)	0.426 (0.402 to 0.451)	0.426 (0.402 to 0.452)
MCC	0.632 (0.590 to 0.661)	0.646 (0.630 to 0.660)	0.641 (0.622 to 0.657)	0.654 (0.643 to 0.664)

AUCROC: Area under the receiver operating characteristic curve, AU-PRC: Area under the precision-recall curve, MCC: Matthews correlation coefficient, NPV: Negative predictive value, PPV: Positive predictive value.

Figure 1. STROBE Checklist

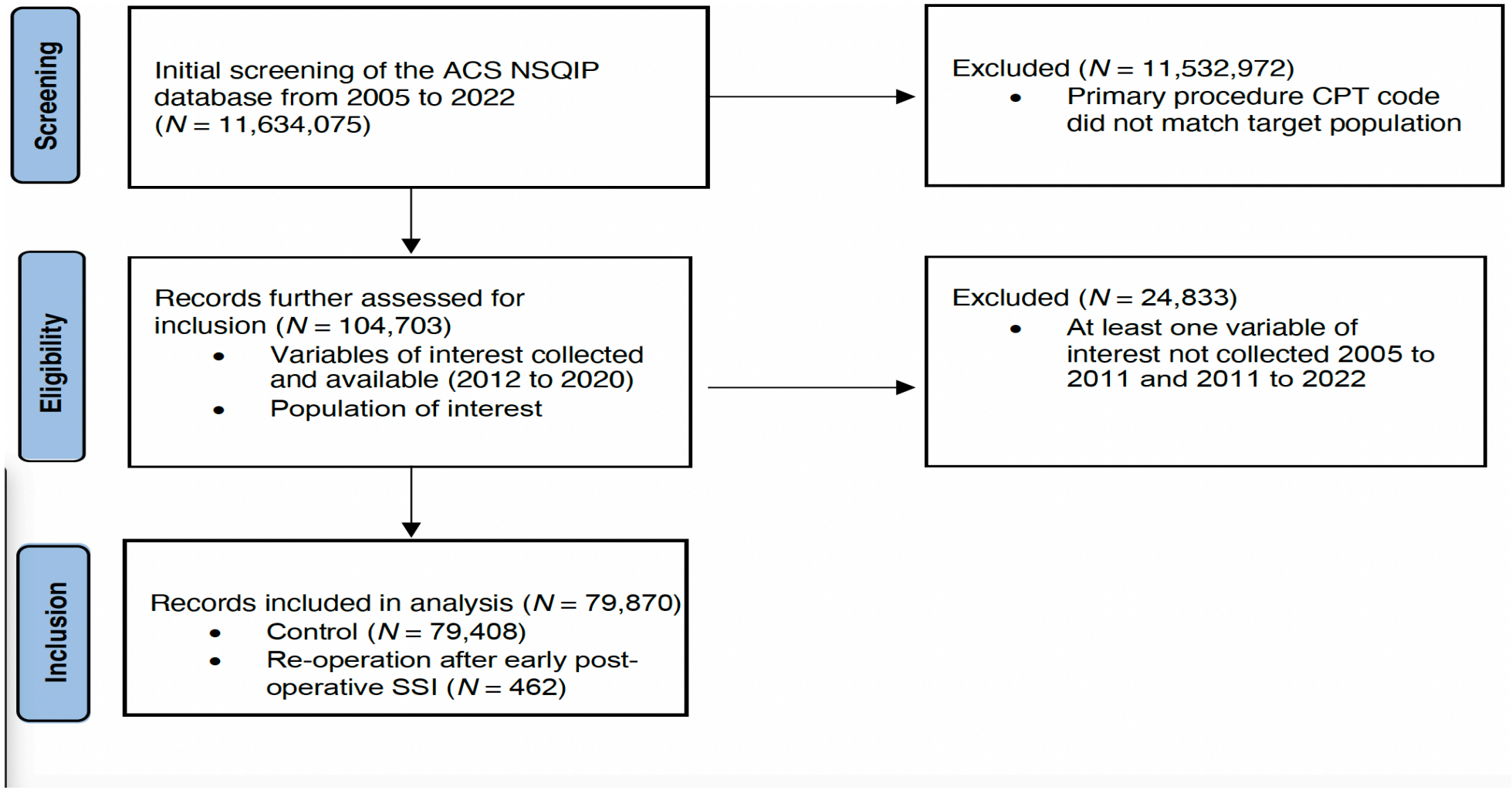


Figure 2. Area under the receiver operating characteristic curve (AUC-ROC; A), area under the precision-recall curve (AUPRC; B), and calibration plot (C) for the final model with 95% confidence intervals from 10,000 bootstrapped resamples

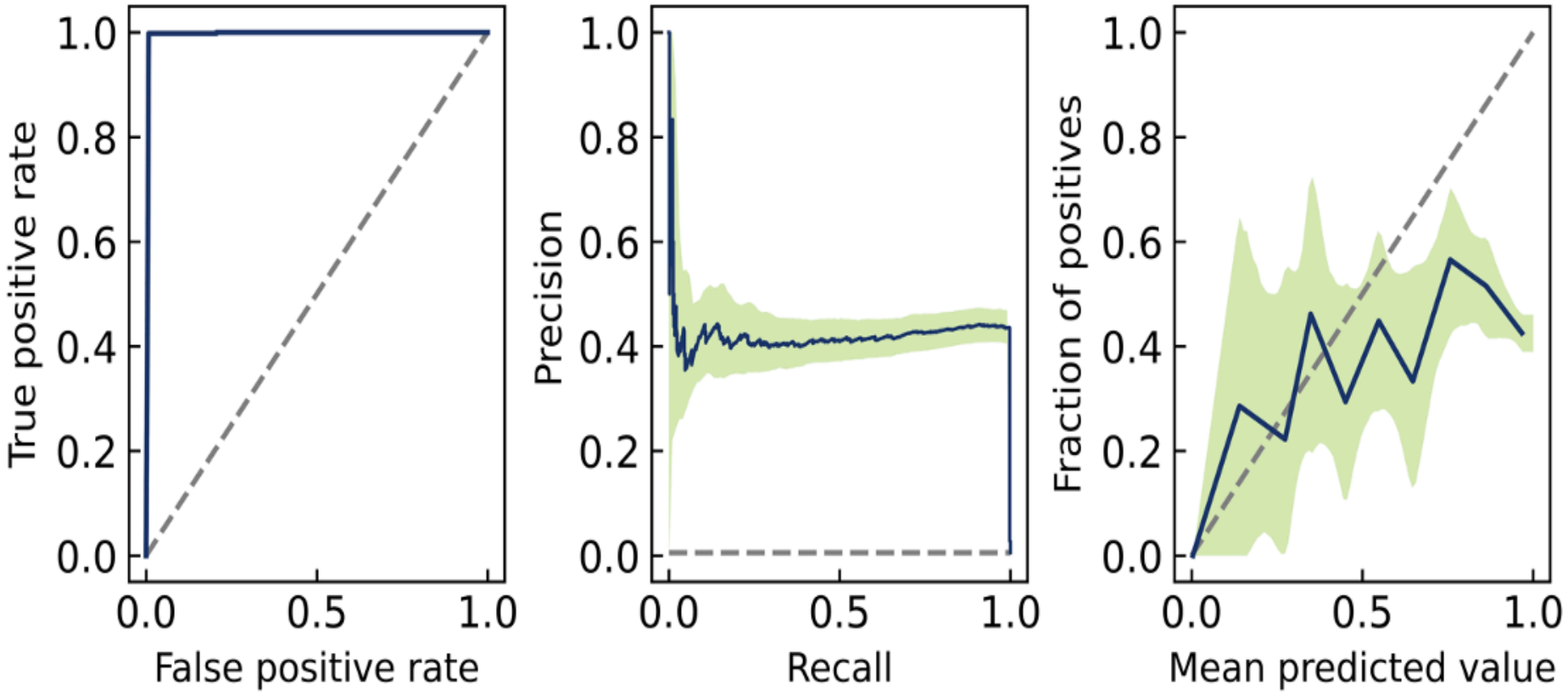
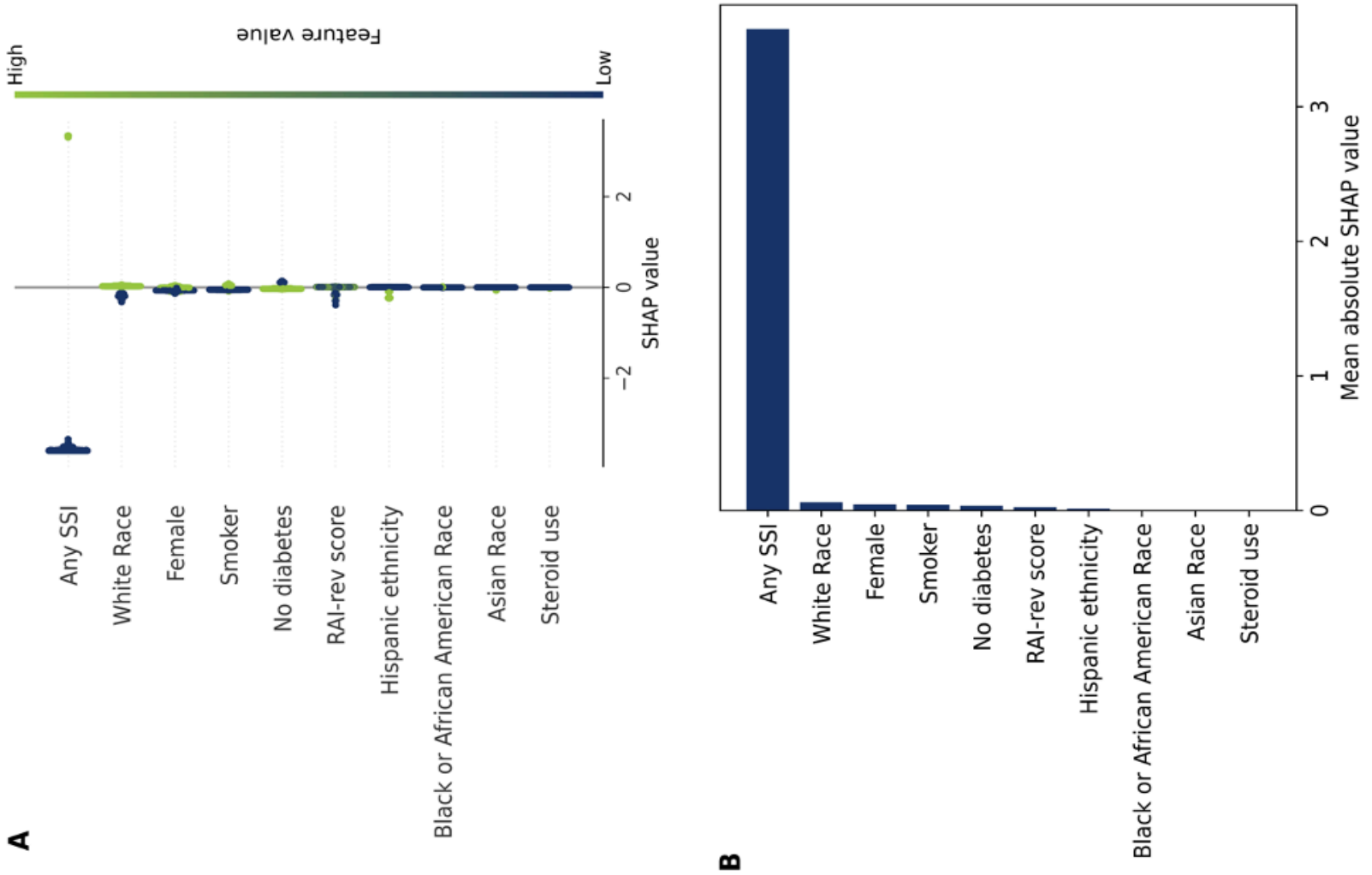


Figure 3. Summary SHAP plot (A) and mean absolute SHAP value plot (B) showing the impact and relative importance of the selected feature-engineered predictors in the Bayesian-optimized model.



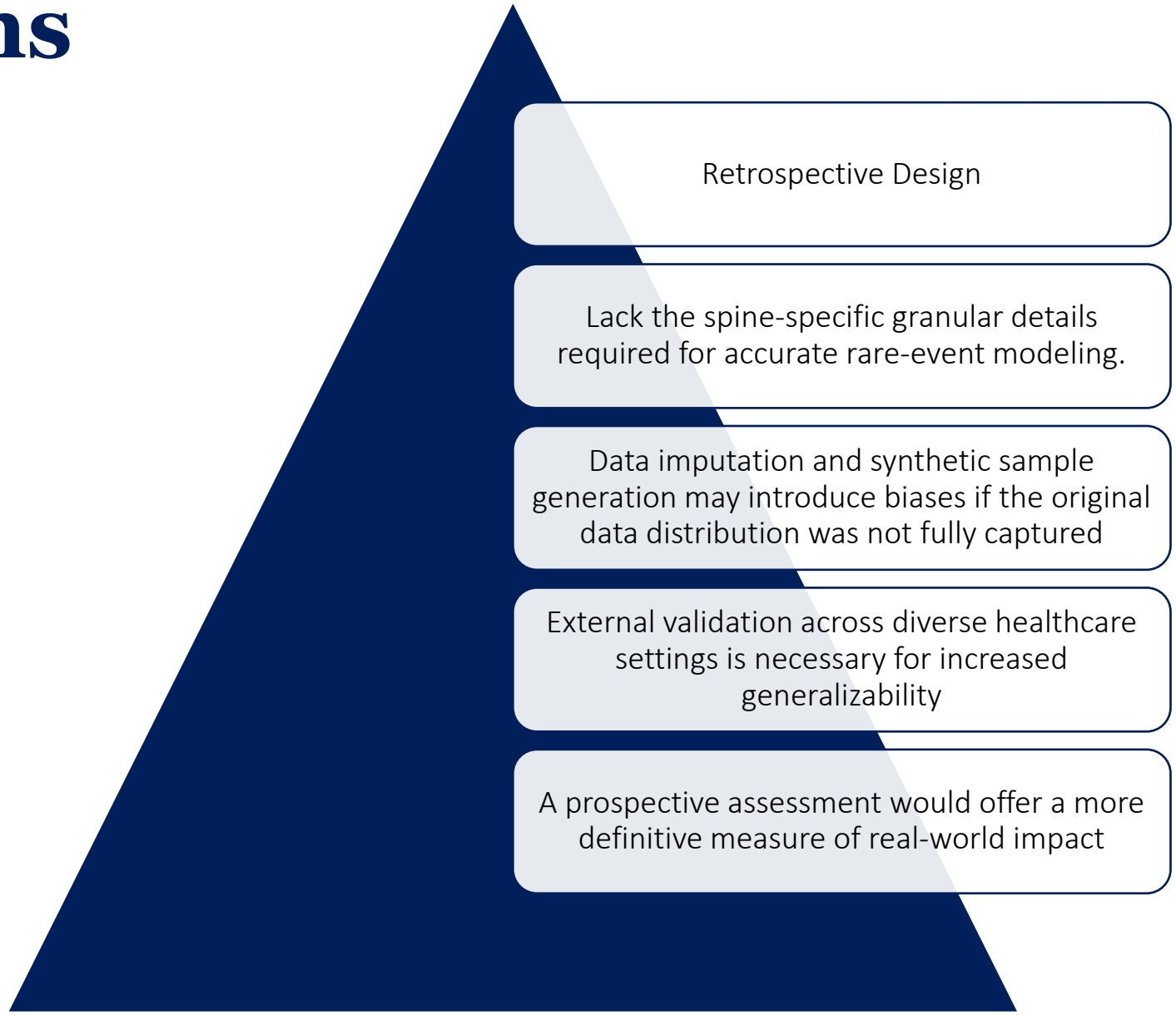
Conclusion

Optimized machine learning approach accurately predicted the rare event of early post-operative re-operation following SSI in lumbar microdiscectomy

Findings underscore the feasibility of data-driven risk stratification to improve outcomes and guide surgical strategies in spine care

External validation and prospective assessment are needed to increase generalizability and further refine individualized surgical decision-making

Limitations



Acknowledgments

-William C. Welch, MD

-Mert Marcel Dagli, MD

-Penn Neurosurgery

Penn Medicine