**Title**: Predicting early postoperative outcomes after pituitary adenoma surgery using a machine learning approach

**Authors**: Todd Hollon, Adish Parikh, Balaji Pandian, Jamaal Tarpeh, Ariel Barkan, Erin McKean, Stephen E. Sullivan

**Abstract**

**Object**: Pituitary adenomas occur in a heterogenous patient population with a diverse set of perioperative risk factors, endocrinopathies, and other tumor-related comorbidities. This heterogeneity makes predicting postoperative outcomes challenging using traditional scoring systems as a predictive method. Modern machine learning algorithms can automatically identify the most predictive risk factors and learn complex risk factor interactions using training data to build a robust predictive model that can generalize to new patient cohorts. We aim to build a predictive model using supervised machine learning to accurately predict early postoperative outcomes after pituitary adenoma surgery.

**Methods**: A retrospective cohort of 400 consecutive pituitary adenoma patients was used as the study population. Patient variables/predictive features were limited to common patients characteristic (e.g. age, gender, tumor type, etc.) to improve model implementation. Univariate and multivariate odds ratio analysis was performed on the full study population to identify individual risk factors for common postoperative complications and to compare risk factors with model predictors. The study population was then split into 300 training/validation patients and 100 testing patients to train and evaluate four machine learning models using binary classification accuracy for predicting early postoperative outcomes after pituitary adenoma surgery.

**Results**: Of the 400 total patients, mean age was 53.9 ± 16.3 years, nonfunctioning adenomas occurred in 59.8% of patients, 84.7% were macroadenomas, and mean body mass index was 32.6 ± 7.8 (58.0% obesity rate). Multivariate odds ratio analysis found age < 40 years was associated with a 2.86 greater odds of postoperative diabetes inspidus; nonobese patients (BMI < 30) were 2.2 times more likely to develop postoperative hyponatremia. Using a broad criteria for a poor early postoperative outcome, which included 1) major medical and 2) early surgical complications, 3) extended length of stay, 4) emergency department admission, 5) inpatient readmission, and 6) death, 31.0% of patient met criteria for a poor early outcome. After model training, a logistic regression model with elastic net regularization (LR-EN) achieved the best performance on the 100 patient testing set with a sensitivity of 68.0%, specificity of 93.3%, and overall accuracy of 87.0% for predicting early postoperative outcomes after pituitary adenoma surgery. The ROC and PR curves for the LR-EN model had an area under the curve of 82.7 and 69.5, respectively. Most important predictive variables were lowest perioperative sodium, age, body mass index, and highest perioperative sodium, and Cushing’s disease.

**Conclusion**: Early postoperative outcomes after pituitary adenomas surgery can be predicted with an accuracy of 87% using a machine learning approach. These results provide insight into how predictive modeling using machine learning can be used to improve the perioperative management of pituitary adenoma patients.

**Introduction**

The ability to predict patient outcomes after a specific treatment is fundamental to providing optimal surgical care. Pituitary adenomas present a unique predictive challenge due to significant heterogeneity among the patient population. This heterogeneity stems from both the diverse at-risk patient population and underlying tumor pathophysiology. Pituitary adenomas can occur at any age, with age-adjusted incidence for 15-75+ year-old patients ranging from 1.5-7.5 tumors per 100,000 people.17 Endocrinopathies that result from functioning adenomas can produce severe preoperative comorbidity, such as obesity, diabetes mellitus, and cardiomyopathies. Complication rates after transphenoidal surgery for Cushing’s disease range up to 42%.19 However, nonfunctioning adenomas are more likely to present in older patients, who may have multiple chronic medical conditions that can increase perioperative surgical risk.9 The clinical diversity of pituitary adenoma patients makes it challenging to use traditional biostatistical techniques or scoring systems to stratifying surgical risk or predict postoperative outcomes given that specific patient characteristics/features (e.g. tumor type, age, body mass index, etc.) are likely to vary in predictive importance across the entire patient population.

Advances in applied predictive modeling using machine learning have provided a novel method for predicting outcomes in healthcare.6 Machine learning models have an advantage over other predictive methods by automatically learning the best predictive features present in training data. Opposed to attempting to manually identifying these features, which is time and labor intensive, machine learning models can automatically identified the most robust predictive features and can potentially generalize to new patient cohorts. Previous studies have used these methods to predict outcomes after stereotacic radiosurgery for brain metastasis15 and arteriovenous malformations16, stratify cardiovascular risk14, predict mortality/readmission/length of stay2, and cancer prognosis.25

In order to improve the perioperative management and risk stratification of pituitary adenoma patients, we aim to predict early postoperative outcomes after pituitary adenoma surgery using a machine learning approach. By analyzing a large cohort of pituitary adenoma patients treated at a tertiary care center, our objective is to develop an accurate predictive model built via modern machine learning methods that will identify patients at the highest risk for poor early postoperative outcomes after pituitary adenoma surgery.

**Methods**

*Study design*

We designed a retrospective analysis of 400 consecutive pituitary adenoma patients treated via endoscopic endonasal approach for surgical resection by the senior authors (E.M., S.E.S.). After IRB approval (HUM00111786), two independent reviewers completed a systematic chart review using a standardized database template. In addition to formal chart review, utilization of the University of Michigan Electronic Medical Research Search Engine (UM-EMERSE) was used to confirm patient details and/or any discrepancy between reviewers.

The study aims were to 1) perform exploratory data analysis of a large series of pituitary adenomas patients treated at a high-volume medical center with an integrated neuroendocrine center and 2) develop and validate a supervised machine learning model that can predict early postoperative outcomes. We defined a poor early postoperative outcome using broad and inclusive criterion which included: 1) major adverse medical event within 30 days (including deep vein thrombosis (DVT)/pulmonary embolism (PE), myocardial infarction, severe arrhythmia or stroke), 2) early surgical complication (cerebrospinal fluid leak with or without symptomatic pneumocephalus or postoperative meningitis), 3) expected length of stay (2 days for non-Cushing’s disease, 4 days for Cushing’s disease) exceeded by two days, or any of the following within 30 days of surgery: 4) emergency department (ED) admission, 5) inpatient admission, or 6) death. Because these outcomes can be overlapping (e.g. extended hospital stay due to pulmonary embolism), patients who experienced any or all of these outcomes were assigned to the poor outcome group using a binary classification. Sodium dysregulation (diabetes insipidus or hyponatremia) itself was not considered a poor early postoperative outcome, as it can often be managed effectively as an outpatient without complication. Patient characteristics/model predictors were established prior to initiating chart review. Model predictors were chosen to include only standard clinical information common to all pituitary adenoma patients to improve future model implementation. Disease-specific characteristics (e.g. preoperative ACTH levels) and advanced radiographic features (e.g. Knosp score) were avoided to eliminate missing/not applicable data values and data sparsity.

*Descriptive statistics and data exploration*

All patient characteristics and outcomes were divided into continuous or non-ordered categorical variables for statistical analysis. In addition to the poor early postoperative outcomes defined above, risk factors for common postoperative complications after pituitary adenoma surgery were also explored. Using the full 400 patient dataset, pair-wise odds ratio analysis was performed to explore risk factors for diabetes insipidus, hyponatremia, transient cranial nerve palsy, cerebrospinal fluid leaks, symptomatic pneumocephalus, deep vein thrombosis/pulmonary embolism, postoperative meningitis. For univariate analysis, continuous variables were converted to an indicator variable using a binary encoding (e.g. age > 40 years, body mass index > 30) to allow for odds ratio calculation. Univariate statistical significance was calculated using Fisher exact testing and defined as p-value < 0.05. Multivariate logistic regression was done for postoperative complications with multiple statistically significant predictors to account for covariance among variables. The R Environment for Statistical Computing (version 3.3.1; <http://www.r-project.org>) and Python-based SciPy library (version 0.19.1, <https://www.scipy.org>) were used for statistical analysis.

*Supervised machine learning*

Four supervised machine learning algorithms were trained and tested as binary classifiers to predict early postoperative outcomes in pituitary adenoma patients: naïve Bayes, logistic regression with elastic net regularization (linearly combined L1 and L2 regularization penalties) (LR-EN), support vector machines with linear kernel (SVM), and random forest. These methods were selected for algorithm diversity (i.e. Bayesian model, generalized linear model, margin classifier, and decision trees). A total of 26 patients characteristics were used as predictive variables. Model hyperparameters were selected using a grid search and 10-fold cross-validation was performed for each model. Training/cross-validation and testing sets were selected by random sampling without replacement from the full 400 patient study population using a 75%/25% (300/100 patient) split. To improve clinical relevance and allow patient’s risk to be recalculated in the perioperative setting (“rolling” risk assessment), perioperative lowest and highest sodium levels were used as predictors. Data preprocessing included rescaling continuous variables to between 0-1. Model training and performance was evaluating using prediction accuracy.

To further evaluate the models, both receiver operating characteristic (ROC) and precision-recall (PR) curves were generated and area under the curves (AUC) was calculated. To determine the best performing model, McNemar’s test was used to evaluate marginal homogeneity and determine statistically significant differences between model predictions. Variable importance for the best performing model is reported to improve model interpretability and assessment of clinical relevance. The R “caret” package (<http://caret.r-forge.r-project.org>) was used for model training, hyperparameter search, validation, and testing. The R and Python code can be downloaded at <https://github.com/toddhollon/pituitary_ml>.

**Results**

*Patient population and early postoperative outcomes*

Mean age of the study population was 53.9 ± 16.3 years, ranging from 13 to 91 years, and 54% were male. The majority of patients were white (84%) or black (10%) race. Nonfunctioning pituitary tumors were the most common (59.8%) followed by growth-hormone secreting adenomas (22.8%) and ACTH-secreting adenomas (13.0%). Previous transphenoidal surgery or radiation therapy occurred in 16.5% and 4.0% of tumors, respectively. A listing of patient characteristic can be found in **Table 1**. Differences in sex, tumor size, age, and body mass index with respect to tumor type are shown in **Figure 1**.

Sodium dysregulation was the most common postoperative complication after pituitary adenoma surgery (**Table 2**). Diabetes insipidus and hyponatremia occurred in 14.8% and 14.3% of patients, respectively. Prevalence of cerebrospinal fluid leak was 7% and 2% of patients developed symptomatic pneumocephalus. Acute DVT/PE was found in 1.5% of patients and 1.3% developed postoperative meningitis. Extended length of stay occurred in 20.7% of non-Cushing’s disease patients and 30.8% of Cushing’s disease patients. Thirty-day emergency department admission and subsequent inpatient readmission occurred in 17.0% and 11.8% of patients, respectively. Thirty-day mortality rate was 1% (4/400). Based on the study defined criteria, 31% (124/400) of patients had a poor early postoperative outcome, with the top 4 inclusion criteria being ED admission, extended length of stay, inpatient readmission, and CSF leak (**Figure 2a**). A single inclusion outcome occurred in 13% (52/400) of patients, while 18% (72/400) had two or more (**Figure 2b**).

*Data exploration and odds ratio analysis*

To explore risk factors for specific postoperative complications after pituitary adenoma surgery, we performed a pair-wise univariate odds ratio analysis of patient characteristics and comorbidities (**Figure 3**). Diabetes insipidus was associated with age less than 40 years, Cushing’s disease, microadenomas and no history of anticoagulation/antiplatelet use. On multivariate logistic regression, age was the only predictor that remained statistically significant, with patients less than 40 years having 2.86 greater odds of postoperative DI (95% confidence interval (CI) 1.52 – 5.27, p-value = 0.001). Patients with microadenomas were 1.9X more likely to develop diabetes insipidus, however this trend did not reach statistical significance on multivariate regression (OR 1.93, 95% CI 0.91- 3.98, p-value = 0.076).

Obesity was inversely correlated with postoperative hyponatremia on multivariate analysis (OR 0.46 95% CI 0.25- 0.82, p-value = 0.009) and clinically significant trend towards older patients being more likely to develop hyponatremia was observed (OR 2.48 95% CI 1.03- 7.00, p-value = 0.058). History of skull base radiation was associated with postoperative symptomatic pneumocephalus (8.6, 95% CI 1.1 – 42.9, p-value = 0.040), and recurrent pituitary adenomas/previous resection was associated with postoperative meningitis (OR 7.7, 95% CI 1.1 – 67.0, p-value = 0.03). Cushing’s disease (OR 12.2, 95% CI 2.2 – 92.3.0, p-value = 0.006) and a history of congestive heart failure (OR 7.8, 95% CI 0.91 – 49.8, p-value = 0.04) significantly increase the odds of DVT/PE on both univariate and multivariate logistic regression. Multivariate analysis included preoperative antiplatelet/anticoagulant use to account for perioperative cessation of medications. Of the four patients that died within 30 days of surgery, three had Cushing’s disease (p-value = 0.008). To further explore relationship between age, BMI, and sodium dysregulation, the distribution of postoperative sodium values were plotted with respect to age and BMI in **Figure 4**.

*Predicting early postoperative outcomes using machine learning*

Aftertraining and cross-validation of the four machine learning models (**Supplemental Figure 1),** they were tested on an independent testing set of 100 patients. Performance of each model can be found in **Table 3**. LR-ER model achieved the highest accuracy at 87.0% (95% CI 78.8 – 92.9%; optimized hyperparameters: alpha = 0.05, lambda = 0.005), followed by the random forest model (85.0%, 95% CI 76.5 – 91.4: optimized hyperparameter: mtry = 7). A significant improvement in model sensitivity was noted for LR-ER and random forest over naïve Bayes classifier and support vector machines. A statistically significant difference in model prediction accuracy was found between LR-ER versus support vector machines and naïve Bayes, but not random forest. Heatmap of pairwise model comparisons and prediction correlations can be found in **Supplemental Figure 2**. Area under the receiver operating characteristic (AUC-ROC) and precision-recall (AUC-PR) curves are presented in **Table 3**. LR-EN model had the largest AUC-PR (69.5%) and second largest AUROC (82.7%).

ROC and PR curves for each model are presented in **Figure 5A and 5B.** To better understand the output prediction probabilities from the LR-EN classifier, the probability of a poor early postoperative outcome for each test set patient is shown in **Figure 5C.** The majority of patients who did not have a poor outcome had a low prediction probability (mean 0.201 ± 0.189). The LR-EN classifier correctly identified 17/25 (68%) of patients who did have a poor early postoperative outcome and reflects the improvement in LR-ER sensitivity compared the other trained models. The top six most important predictive variables are shown in **Figure 5D**. Lowest perioperative sodium level was the most important predictor, followed patient age and BMI. These findings are concordant with the calculated odds ratios and the relationships identified above between age, BMI and sodium dysregulation.

**Discussion**

Our findings demonstrate that early postoperative outcomes after pituitary adenoma surgery can be accurately predicted using a machine learning approach. Using the full patient cohort, we were first able to identify risk factors for common postoperative complications, including diabetes insipidus, hyponatremia, and DVT/PE, using univariate and multivariate odds ratio anaylsis. By using a large cohort of pituitary adenomas patients to train a machine learning classifier, we were then able to identify patients at high risk for poor postoperative outcomes with an accuracy of 87% and AUC of 83% on ROC analysis on a 100 patient testing set. We identified sodium dysregulation, age, obesity, Cushing’s disease and sex as the most predictive features for stratifying patient’s risk of a poor postoperative outcome. These results provide insight into how predictive modeling using a machine learning approach can improve the surgical management of pituitary tumors.

A major motivation for the study resulted from the high prevalence of pituitary adenomas among central nervous system tumors, coupled with the lack of any system to meaningfully predict postoperative outcomes. Pituitary adenomas represent approximately 16% of all newly diagnosed brain tumors and are among the top three most common primary central nervous system tumors in the US.17 Moreover, they are the second most common non-malignant brain tumor with surgical resection as a potential curative treatment. While scoring systems have been developed to that using radiographic features to classify invasion into adjacent structures5,10,11 and hormone levels to predict treatment response3,24, no scoring system has been developed to comprehensively include patient characteristics and stratify surgical risk. Such scoring systems have been developed for meningiomas21, gliomas (both low-grade4,20 and malignant13,18), brain metastasis7,8, and arteriovenous malformations12,22,23 to predict both early and long-term outcomes. These scoring system help to determine indications for surgery, improve patient counseling, intraoperative decision-making, and postoperative management.

While scoring systems can apply well to homogenous patient populations, such as those seen in glioblastoma, they are not well suited for the clinical heterogeneity found in pituitary adenoma patient populations. Unlike gliomas or meningiomas, pituitary adenomas are unique among brain tumors such that the presence of the tumor can result in severe systemic illness due to the stimulation or suppression of a neuroendocrine axis. As a result, perioperative risk can stem both from tumor morphology and from secondary systemic comorbidities, rather than lesion morphology alone (e.g. eloquent tumor location in gliomas, deep venous drainage in AVMs, etc.). The complex interplay between tumor morphology, patient characteristics, and secondary comorbidities associated with endocrinopathies necessitates a more robust method for applied predictive modeling. Machine learning methods offer the opportunity to improve predictive accuracy by learning the complex interactions between risk factors.

The application of machine learning techniques to healthcare has increased over the last 5 years, mainly due to larger datasets, electronic medical records, and better application programming interfaces.1,6 Leveraging these aforementioned tools, we were able to build a machine learning classifier that captured the complex risk factor interactions of pituitary adenoma patients and provide accurate predictions of early postoperative outcomes. One complex interaction we identified via the odds ratio analysis and model feature importance was that between age, BMI, tumor size and postoperative sodium dysregulation. For example, we found that younger age (<40 years), microadenomas, and Cushing’s disease were associated with postoperative diabetes insipidus. The underlying mechanism for this is unclear, but may be related to microadenoma and Cushing’s disease presenting in younger patients, and resection of these microadenomas can requiring more pituitary gland manipulation, and subsequent diabetes insipidus, compared to nonfunctioning macroadenomas presenting in older patients. Additionally, it is unclear how younger age and obesity, as independent risk factors, could be protective against hyponatremia. This observation may be explained as the inverse of the previous; non-obese older patients with macroadenomas undergo less gland manipulation, therefore less susceptible to diabetes insipidus, but more vulnerable to hyponatremia. While any attempt to interpret these results must be tentative, high quality training data allows the machine learning model to identify these complex interactions and latent variables, which can then be used to accurately predict on new patients.

Our study is limited by being completed at a single institution. Patients treated at other institutions and by other surgeons will be needed to further test the generalizability of the predictive model. The current model is designed as a binary classifier. With a larger dataset, a multiclass classifier can be trained that may allow for prediction and risk stratification of multiple outcomes (e.g. medical complications, surgical complications, readmissions, etc.). With long follow up data, the model can be further tailored to include long-term treatment response and predict tumor recurrence. Our study population will be followed longitudinally in preparation for expanding our predictive model and provide additional data for model training using similar machine learning methods described here.

**Conclusion**

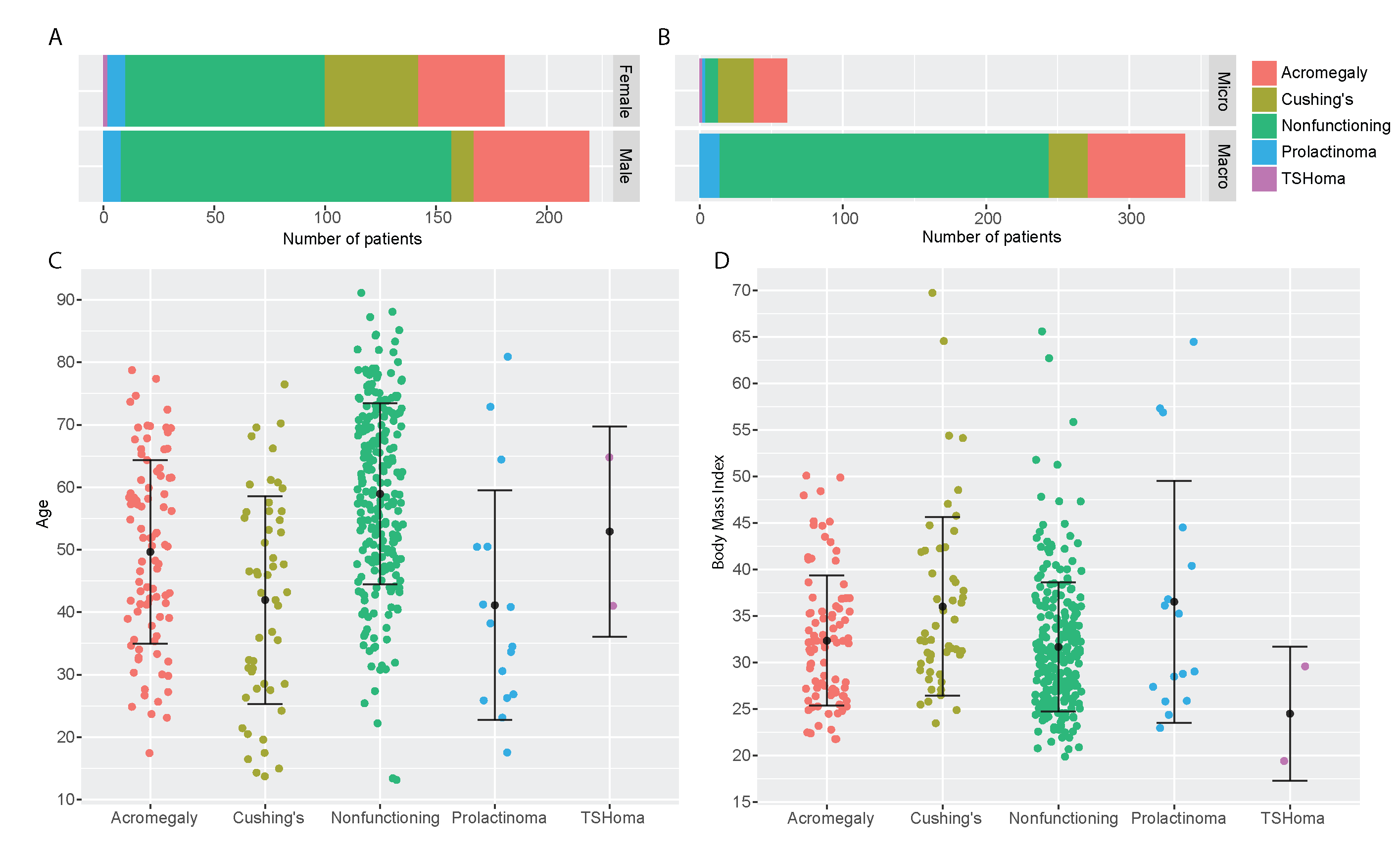
Pituitary adenomas occur in a heterogeneous patient population, which makes predicting postoperative outcomes a challenge. To address this challenge, we analyzed large cohort (400 patients) of consecutive pituitary adenoma patients; using a machine learning approach, early postoperative outcomes after pituitary adenomas surgery can be predicted with an accuracy of 87%. These results provide insight into how machine learning can be used to improve the perioperative management of pituitary adenoma patients.

**Table 1: Preoperative patient characteristics**

**Table 2: Summary of early postoperative complications and outcomes**

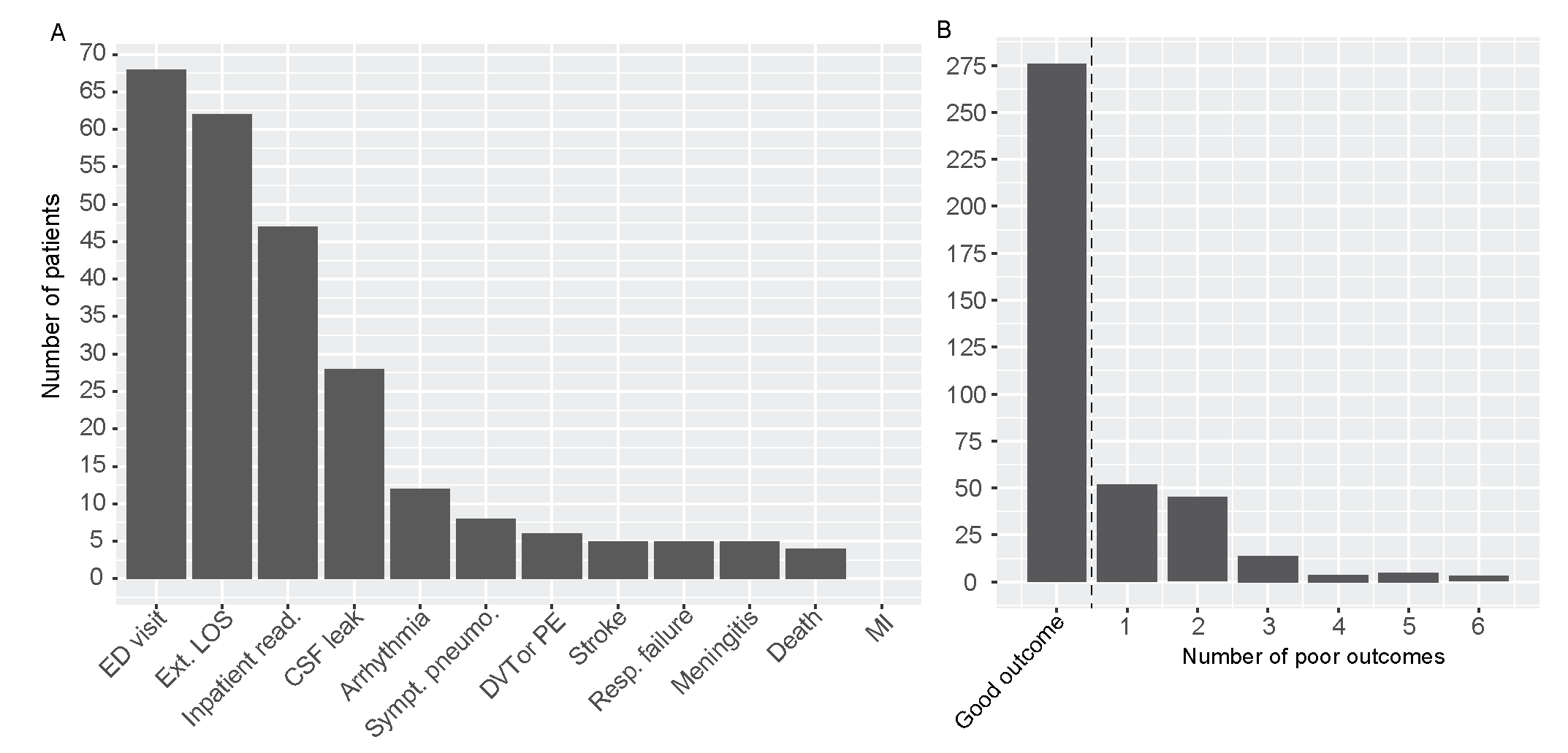
**Table 3: Model performance metrics**

**Figure 1: Patient characteristics by pituitary adenoma diagnosis**

****

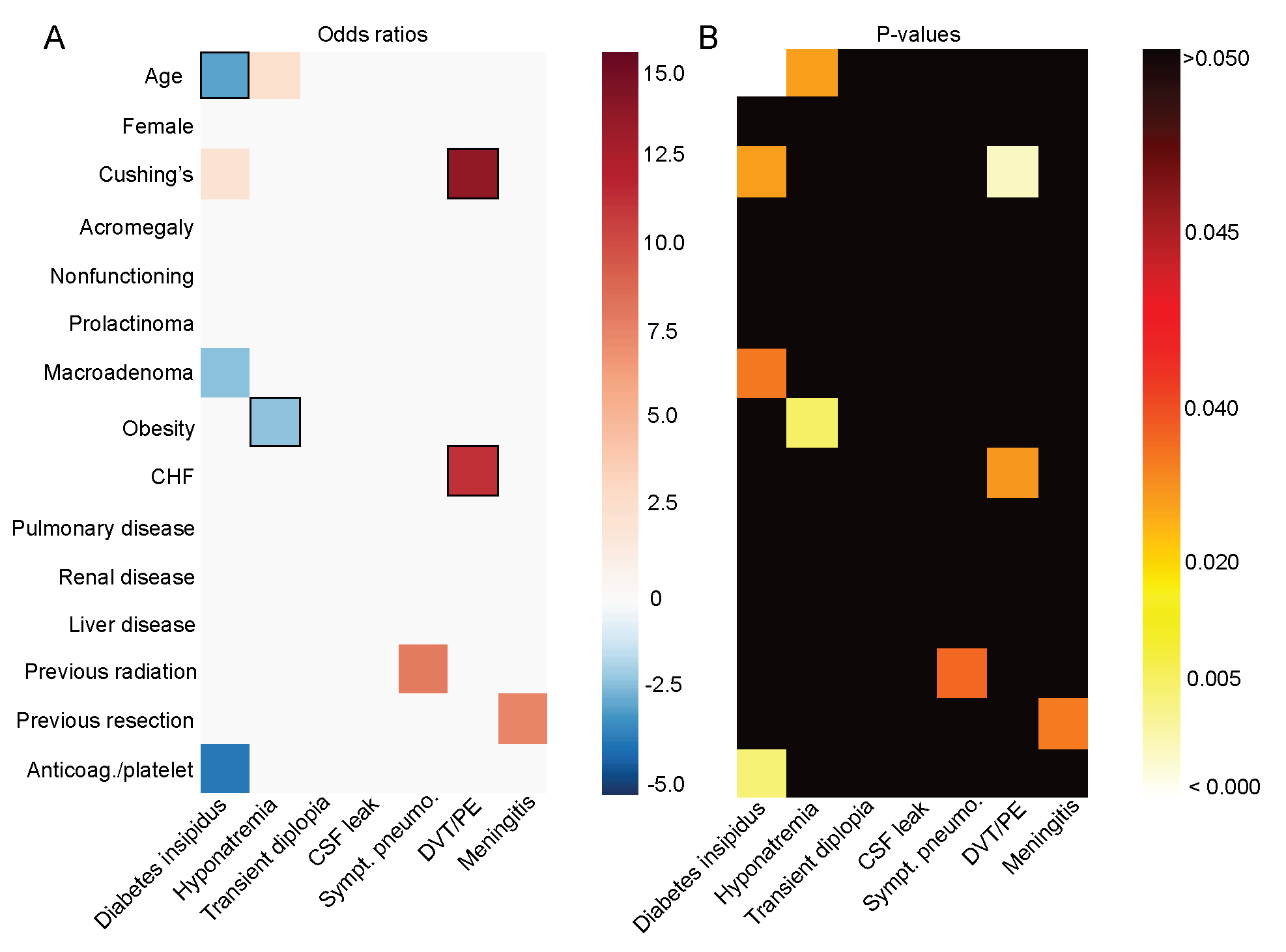
(**A**) Nonfunctioning adenomas (62.3%) and acromegaly (63.4%) were more common in males. Cushing’s disease was more common in females (80.7%). (**B**) Macroadenomas were more common in our study population (84.8%) and the majority were nonfunctioning adenomas (67.8%). Cushing’s disease had almost equal distribution between microadenomas (51.9%) and macroadenomas (48.1%). (**C**) Mean age of nonfunctioning adenoma patients was 58.9 ± 14.4 and was significantly greater than functioning adenomas (mean 46.3 ± 16.1 years, p-value < 0.000). (**D**) Prolactinoma patients had the greatest BMI (36.5 ± 13.0), followed by Cushing’s disease patients (36.0 ± 9.6) and acromegaly patients (32.4 ± 7.0).

**Figure 2: Study-defined early postoperative outcomes**

****

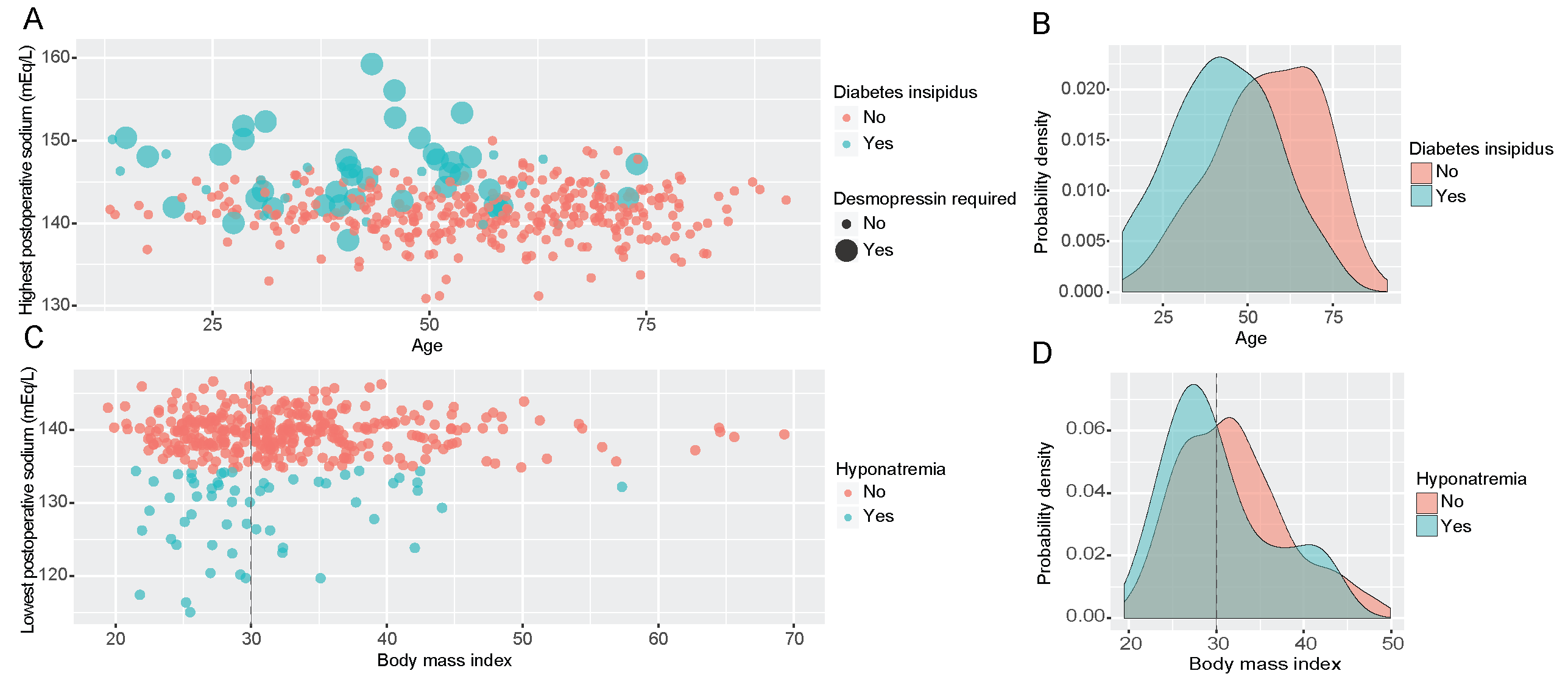
(A) Distribution of inclusion criteria met for the study-defined early postoperative outcomes across the study population (N = 400). Most common criteria met for poor early postoperative outcome were emergency department admission, extended length of stay, inpatient readmission, and CSF leak. No patient suffered myocardial infarction and four patients died within 30 days of surgery. (**B**) Good early postoperative outcome occurred in the majority of patient (69%, 276/400). A single inclusion criterion was met in 13% (52/400) of patients and 18% (72/400) had two or more.

**Figure 3: Univariate and multivariate odds ratio analysis**

****

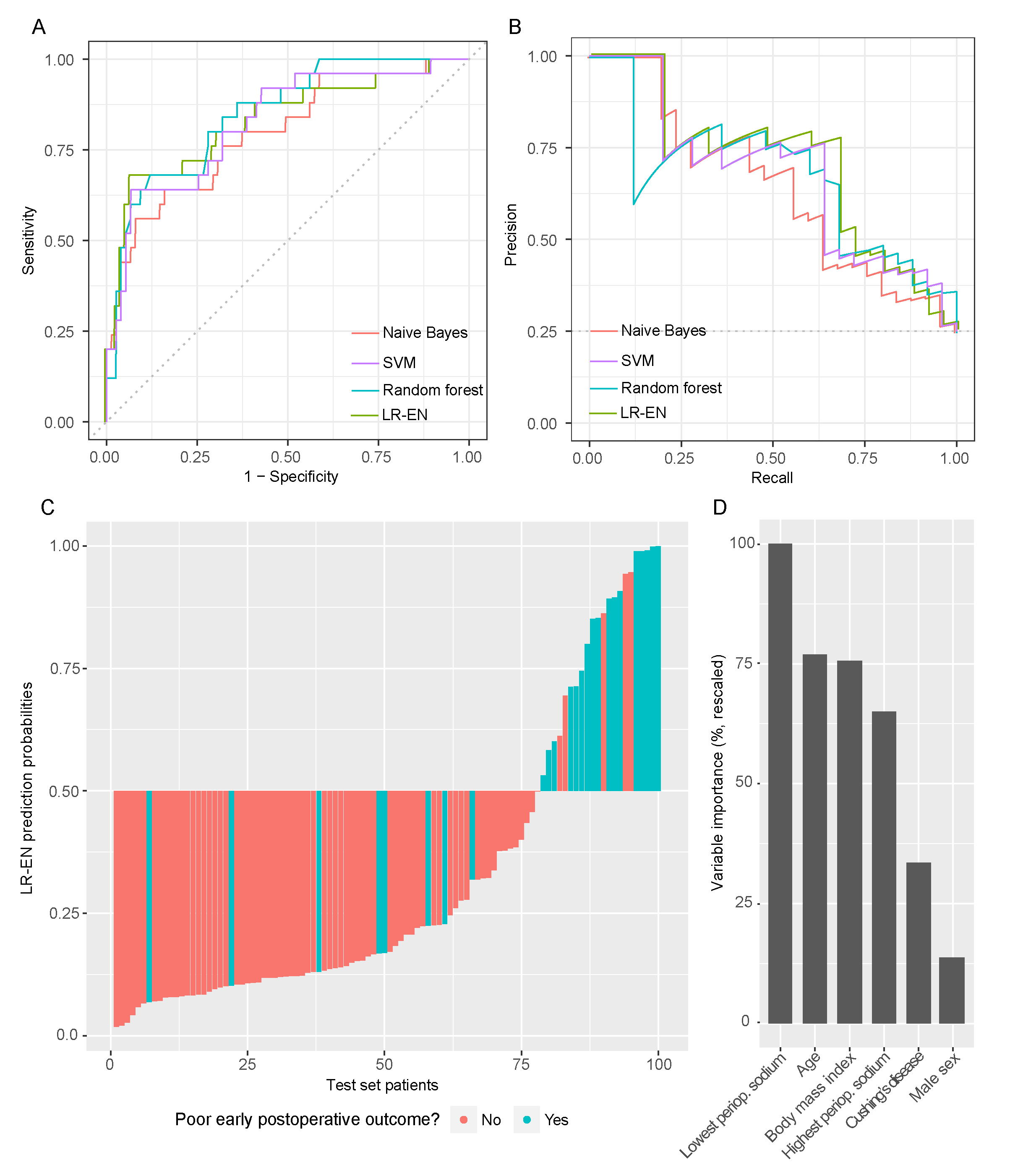
(**A**) Odds ratios and (**B**) p-values are presented as tiled heatmaps comparing patient characteristics with early complications. Odds ratio values are color coded such that red indicates a patient characteristic as a risk factor and blue indicates a protective factor for a given outcome. Black boxes identify patient characteristic-outcome pairs that remained statistically significant on multivariate analysis. P-values are presented as a single continuous variable on a logarithmic scale. Solid black squares are non-statistically significant comparisons.

**Figure 4: Early postoperative sodium dysregulation**

****

(**A**) Scatter plot showing the distribution of highest postoperative sodium levels with respect to patient age. Younger patients had a higher probability of being diagnosed with diabetes insipidus and requiring desmopressin for treatment. (**B**) Probability density function of patient age shows two distinct distributions separable by a diagnosis of diabetes insipidus. (**C**) Scatter plot showing the distribution of postoperative lowest sodium with respect to BMI. (**D**) Probability density function shows unique distributions for patients with BMI less than versus greater than 30 (i.e. obesity diagnosis) when diagnosis of hyponatremia is indicated.

**Figure 5: Machine learning model evaluation, prediction probabilities, and variable importance**

****

(**A**) ROC curves and (**B**) Precision-Recall curves are shown. Random forest and LR-EN had the top 2 AUCs for both curves. (**C**) Because LR-EN model had the highest accuracy on the testing set, prediction probabilities for each patient were plotted in ascending order. Ground truth outcome labels are color-coded and LR-EN output probabilities greater than 50% were predicted as having a poor early postoperative. Of the 100 patients, the LR-EN classifier made 5 false positive and 8 false negative errors. Rescaled variable importance for LR-EN model is shown in (**D**). Lowest perioperative sodium levels, age, and BMI were the top three predictive variables in the trained model.

**Supplemental Figure 1: Training set ROC/PR curves**

**Supplemental Figure 2: McNemar’s P-values and prediction correlations**

1. Alvin Rajkomar EO, Kai Chen, Andrew M. Dai, Nissan Hajaj, PeterJ. Liu, Xiaobing Liu, Mimi Sun, Patrik Sundberg, Hector Yee, Kun Zhang, YiZhang, Gavin E. Duggan, Gerardo Flores, Michaela Hardt, Jamie Irvine, Quoc Le, Kurt Litsch, Jake Marcus, Alexander Mossin, Justin Tansuwan, De Wang, James Wexler, Jimbo Wilson, Dana Ludwig, Samuel L. Volchenboum, Katherine Chou, Michael Pearson, Srinivasan Madabushi, Nigam H. Shah, Atul J. Butte,Michael Howell, Claire Cui, Greg Corrado, Jeff Dean: Scalable and accurate deep learning for electronic health records. **arXiv**, Jan 2018

2. Cai X, Perez-Concha O, Coiera E, Martin-Sanchez F, Day R, Roffe D, et al: Real-time prediction of mortality, readmission, and length of stay using electronic health record data. **J Am Med Inform Assoc 23:**553-561, 2016

3. Chandler WF, Barkan AL, Hollon T, Sakharova A, Sack J, Brahma B, et al: Outcome of Transsphenoidal Surgery for Cushing Disease: A Single-Center Experience Over 32 Years. **Neurosurgery 78:**216-223, 2016

4. Chang EF, Smith JS, Chang SM, Lamborn KR, Prados MD, Butowski N, et al: Preoperative prognostic classification system for hemispheric low-grade gliomas in adults. **J Neurosurg 109:**817-824, 2008

5. Di Ieva A, Rotondo F, Syro LV, Cusimano MD, Kovacs K: Aggressive pituitary adenomas--diagnosis and emerging treatments. **Nat Rev Endocrinol 10:**423-435, 2014

6. Dua S, Acharya U R, Dua P, SpringerLink (Online service): Machine Learning in Healthcare Informatics, in, pp XII, 332 p. 119 illus., 350 illus. in color.

7. Gaspar L, Scott C, Rotman M, Asbell S, Phillips T, Wasserman T, et al: Recursive partitioning analysis (RPA) of prognostic factors in three Radiation Therapy Oncology Group (RTOG) brain metastases trials. **Int J Radiat Oncol Biol Phys 37:**745-751, 1997

8. Gaspar LE, Scott C, Murray K, Curran W: Validation of the RTOG recursive partitioning analysis (RPA) classification for brain metastases. **Int J Radiat Oncol Biol Phys 47:**1001-1006, 2000

9. Gondim JA, Almeida JP, de Albuquerque LA, Gomes E, Schops M, Mota JI: Endoscopic endonasal transsphenoidal surgery in elderly patients with pituitary adenomas. **J Neurosurg 123:**31-38, 2015

10. Hardy J: Transphenoidal microsurgery of the normal and pathological pituitary. **Clinical Neurosurgery 16:**185-217, 1969

11. Knosp E, Steiner E, Kitz K, Matula C: Pituitary adenomas with invasion of the cavernous sinus space: a magnetic resonance imaging classification compared with surgical findings. **Neurosurgery 33:**610-617; discussion 617-618, 1993

12. Lawton MT, Kim H, McCulloch CE, Mikhak B, Young WL: A supplementary grading scale for selecting patients with brain arteriovenous malformations for surgery. **Neurosurgery 66:**702-713; discussion 713, 2010

13. Li J, Wang M, Won M, Shaw EG, Coughlin C, Curran WJ, Jr., et al: Validation and simplification of the Radiation Therapy Oncology Group recursive partitioning analysis classification for glioblastoma. **Int J Radiat Oncol Biol Phys 81:**623-630, 2011

14. Myers PD, Scirica BM, Stultz CM: Machine Learning Improves Risk Stratification After Acute Coronary Syndrome. **Sci Rep 7:**12692, 2017

15. Oermann EK, Kress MA, Collins BT, Collins SP, Morris D, Ahalt SC, et al: Predicting survival in patients with brain metastases treated with radiosurgery using artificial neural networks. **Neurosurgery 72:**944-951; discussion 952, 2013

16. Oermann EK, Rubinsteyn A, Ding D, Mascitelli J, Starke RM, Bederson JB, et al: Using a Machine Learning Approach to Predict Outcomes after Radiosurgery for Cerebral Arteriovenous Malformations. **Sci Rep 6:**21161, 2016

17. Ostrom QT, Gittleman H, Xu J, Kromer C, Wolinsky Y, Kruchko C, et al: CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009-2013. **Neuro Oncol 18:**v1-v75, 2016

18. Park CK, Kim JH, Nam DH, Kim CY, Chung SB, Kim YH, et al: A practical scoring system to determine whether to proceed with surgical resection in recurrent glioblastoma. **Neuro Oncol 15:**1096-1101, 2013

19. Patil CG, Lad SP, Harsh GR, Laws ER, Jr., Boakye M: National trends, complications, and outcomes following transsphenoidal surgery for Cushing's disease from 1993 to 2002. **Neurosurg Focus 23:**E7, 2007

20. Pignatti F, van den Bent M, Curran D, Debruyne C, Sylvester R, Therasse P, et al: Prognostic factors for survival in adult patients with cerebral low-grade glioma. **J Clin Oncol 20:**2076-2084, 2002

21. Simpson D: The recurrence of intracranial meningiomas after surgical treatment. **J Neurol Neurosurg Psychiatry 20:**22-39, 1957

22. Spetzler RF, Martin NA: A proposed grading system for arteriovenous malformations. **J Neurosurg 65:**476-483, 1986

23. Spetzler RF, Ponce FA: A 3-tier classification of cerebral arteriovenous malformations. Clinical article. **J Neurosurg 114:**842-849, 2011

24. Yano S, Shinojima N, Kawashima J, Kondo T, Hide T: Intraoperative Scoring System to Predict Postoperative Remission in Endoscopic Endonasal Transsphenoidal Surgery for Growth Hormone-Secreting Pituitary Adenomas. **World Neurosurg 105:**375-385, 2017

25. Yu KH, Zhang C, Berry GJ, Altman RB, Re C, Rubin DL, et al: Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. **Nat Commun 7:**12474, 2016