The Impact of Race on Discharge Disposition and Length of Hospitalization Following Craniotomy for Brain Tumor

Whitney E. Muhlestein[a], **Dallin S. Akagi**[b], **Silky Chotai**[a], and **Lola B. Chambless**[a]

[a]Department of Neurological Surgery, Vanderbilt University Vanderbilt University, Medical Center, 1161 21st Avenue South, T4224 Medical Center North, Nashville, TN 37232, United States

[b]DataRobot Inc., One International Place, 5th Floor, Boston, MA 02110, United States

Abstract

Background—Racial disparities exist in healthcare, frequently resulting in unfavorable outcomes for minority patients. Here, we use guided machine learning (ML) ensembles to model the impact of race on discharge disposition and length of stay (LOS) following brain tumor surgery from the HCUP National Inpatient Sample (NIS).

Methods—We performed a retrospective cohort study of 41,222 patients who underwent craniotomies for brain tumors from 2002–2011 and were registered in the NIS. 26 ML algorithms were trained on pre-hospitalization variables to predict non-home discharge and extended LOS (>7 days) following brain tumor resection, and the most predictive algorithms combined to create ensemble models. Partial dependence analysis was performed to measure the independent impact of race on the ensembles.

Results—The guided ML ensembles predicted non-home disposition (AUC = .796) and extended LOS (AUC = .824) with good discrimination. Partial dependence analysis demonstrated that black race increases the risk of non-home discharge and extended LOS over white race by 6.9% and 6.5%, respectively. Other, non-black race increases the risk of extended LOS over white race by 6.0%. The impact of race on these outcomes is not seen when analyzing the general inpatient or general operative population.

Conclusions—Minority race independently increases the risk of extended LOS and black race increases the risk of non-home discharge in patients undergoing brain tumor resection, a finding not mimicked in the general inpatient or operative population. Recognition of the influence of race on discharge and LOS could generate interventions that may improve outcomes in this population.

Corresponding Author Contact Information: Whitney E. Muhlestein, BA, whitney.muhlestein@gmail.com or whitney.e.muhlestein@vanderbilt.edu, Telephone: 323-346-4990.
Other Author Information: Dallin S. Akagi, BS, DataRobot, Inc., Boston, MA, USA, dallin@datarobot.com
Silky Chotai, MD, Department of Neurological Surgery, Vanderbilt University, silky.chotai@vanderbilt.edu
Lola B. Chambless, MD, Department of Neurological Surgery, Vanderbilt University, lola.chambless@vanderbilt.edu

Keywords

Craniotomy for brain tumor; discharge disposition; length of stay; machine learning; outcomes; race

---

Introduction

It has been well established that racial disparities exist in healthcare, frequently resulting in unfavorable outcomes for minority patients.[1,2] Research has demonstrated that these disparities also exist in outcomes for patients undergoing treatment for brain tumor.[3] The persistence of racial disparities in the treatment of neurosurgical patients is directly at odds with one of the key aspects of the Institute of Medicine's definition of high quality care: equitable delivery of care.[4]

In our resource limited health care system, increasing emphasis has been placed on improving quality. This has been made particularly manifest in the introduction of reimbursement schema that incentivize high quality rather than high quantity care. "Quality" alone is an abstract concept, and providers interested in improving quality of care require quantifiable metrics by which their efforts can be assessed. Decreasing racial disparities in specific post-surgical outcomes represents a concrete opportunity for neurosurgeons to improve quality of care for their patients.

Here, we use machine learning (ML) techniques to build a guided ML ensemble to predict two postoperative outcomes for patients undergoing craniotomy for brain tumor (CFBT): discharge disposition and length of hospital stay (length of stay, LOS). We then interrogate the predictive models to investigate the independent impact of race on these outcomes. Previous work has demonstrated the existence of racial disparities for these outcomes following CFBT – Curry and colleagues showed that black patients have more non-home discharge after CFBT,[5] while Dasenbrock et al saw that black or Hispanic patients were more likely to have extended LOS (>7 days).[6] We expand on this work in three important ways: (1) we use a novel guided ML ensemble technique to validate findings ascertained using more traditional statistical methods; (2) we compare findings in the NIS CFBT population to the NIS population as a whole to determine whether our findings are system-wide or specific to the CFBT population; and (3) we explore the interplay between discharge disposition and LOS.

Materials and methods

Data Base

We used the National Inpatient Sample (NIS) in-hospital discharge database for the years 2002–2011. The NIS is the largest all-payer inpatient database publicly available in the United States, containing approximately 80 million hospital stays from ~1000 hospitals, sampled to approximate a 20% stratified sample of US hospitals.[7] The NIS is compiled and maintained by the Agency for Healthcare Research and Quality (AHRQ, Rockville, Maryland, USA). This publicly available, de-identified database was considered exempt from institutional review board (IRB) review.

Patient Selection

All 79,742,743 admissions registered in the NIS between 2002 and 2011 were screened for inclusion in the study. Eligible admissions were first identified by ICD9 diagnosis codes for brain tumor (225.0 - 225.4, 225.8, 225.9, 199.1, 191.0 - 191.9). Admissions in this subset were then screened for ICD9 procedure codes matching craniotomy (01.20 – 01.29, 01.31, 01.32, 01.39, 01.59). We further restricted our cohort to patients 18 years or older. A total of 41,222 admissions met our criteria.

To determine whether trends in the CFBT population as a whole are mimicked within specific brain tumor diagnoses, we derived the following tumor subsets using the appropriate ICD9 diagnosis codes: meningioma (225.2), non-meningioma benign tumor (225.0, 225.1, 225.8, 225.9), and malignant tumor (199.1, 191.0 – 191.9).

To determine whether trends in the CFBT population are also seen in the general inpatient population, we derived a random sample of all NIS admissions. We also used the ICD9 procedure codes 00–86.9 to identify all NIS admissions associated with operations, and from that sample derived a random sample of all NIS operations. The random samples were the same size as the CFBT population (41,222 admissions).

Variable Selection and Primary Outcomes

Each eligible admission was classified into one of four distinct race categories based on NIS coding: white, black, other non-black race (comprising Hispanic, Asian/Pacific Islander/ Native American, and other race patients), and missing. Because race data was missing for approximately 20% of patients in the sample, "missing" race was included in each model as its own category within the race variable. This proportion of missing patients was consistent across all data samples used for model construction (including the random samples of all NIS admissions and all NIS operations), allowing for comparison between the different models. Additional preoperative data was included as covariates for each admission, including patient age, sex, specific neurosurgical diagnosis, comorbidities, admission type, emergent vs non-emergent surgery, expected payer, and hospital characteristics. The 29 included comorbidity variables were identified using the Elixhauser Comorbidity Software administered by AHRQ. In total, 61 different variables were considered (Table A.1). The primary outcome for the discharge disposition model was home vs other (including death, short-term hospital, nursing facility, AMA, home health, and other). Based on previous studies of LOS in the CFBT population, the primary outcome for the LOS model was defined as greater than 7 inpatient days versus less than or equal to 7 inpatient days.[6] According to this definition, patients with extended LOS made up approximately one third of the CFBT population.

Data Preprocessing

A large number of data preprocessing approaches are represented in the collection of models evaluated in the leaderboard. This section describes the approaches used in the models in the final ensembles: missing numerical data was dealt with by imputing the median of the column, and creating a new binary column to indicate the imputation took place. Numerical data was standardized in each column by subtracting the mean and dividing by the standard

deviation. For linear models (Support Vector Machine, Elastic Net Classifier, Regularized Logistic Regression, Stochastic Gradient Descent Classifier, and Vowpal Wabbit Classifier) categorical data was turned into many binary columns by one-hot encoding. Missing categorical values were treated as their own categorical level and get their own column. For tree-based models categorical data was encoded with integers. The assignment of category values to integers was done randomly.

Leaderboard Construction and Model Validation

Before training 20% of the dataset was randomly selected as the holdout, which was never used in training or cross-validation. The remaining data was divided into 5 mutually exclusive folds of data, four of which were used together as training, with the final fold used for validation.[8] Training was performed five times per algorithm, with each fold used once for validation. Cross-validation scores were calculated by taking the average area under the curve (AUC) of the receiver-operating characteristic (ROC) of the five possible validation folds. 26 ML algorithms were trained and scored and the top performing algorithms selected for use in each ensemble. The algorithms were combined with an Elastic Net Classifier and an Average Blender for the disposition and length of stay ensembles, respectively. Following ensemble model creation, validation was performed on the holdout to demonstrate the ability of the model to generalize to never-before-seen data. The holdout was taken as a single sample, and so no confidence intervals were calculated. Model construction was performed using ML software from DataRobot, Inc.

Permutation Importance

The relative importance of a feature to the final model was assessed using permutation importance, as described by Breiman.[9] Using the training data only, for each variable the model was retrained on data with the values for the variable randomly permuted. The difference in performance in AUC between the model built on the reference data and that of the data with the permuted variable was used to rank and compare the relative importance of the features to the model.

Partial Dependence

To understand the independent impact of race on the disposition and LOS ensembles, we constructed Partial Dependence plots as described by Friedman.[10] A subset of the training data was selected. For any variable we made predictions from the model after having replaced all the values for the variable with a constant test value and computing the mean of those predictions. We tested many values to observe how the model reacts to changes in the variable of interest.

Other Statistical Methods

Traditional statistical analysis was performed on selected patient and hospital characteristics. Continuous variables were compared using the Mann-Whitney U test. Categorical variables were compared using Pearson's χ^2 test. Statistical analysis was performed using the open source statistical tools in SciPy (SciPy ver 0.17).

Results

Patient Characteristics

41,222 admissions for craniotomy for brain tumor were reviewed for analysis. 25,406 resulted in discharge to home and 15,705 admissions did not. 111 admissions had no or unknown discharge disposition recorded and were excluded from the study. Black patients were more likely than white or other non-black minorities to have non-home discharge following CFBT ($P < .001$). (For additional patient characteristics, see Table A.2).

27,314 admissions lasted <= 7 days and 13,907 admissions lasted >7 days. 1 admission had no recorded LOS and was excluded from analysis. Black and non-black minority patients were more likely than white patients to have extended LOS following CFBT ($P < .001$). (For additional patient characteristics, see Table A.3).

ROC Curve and Other Classifier Statistics

An ensemble model including a Nystroem Kernel SVM Classifier, Elastic-Net Classifier, and Extreme Gradient Boosted Trees Classifier was best able to predict discharge disposition, and an ensemble comprising an Elastic Net Classifier, a Vowpal Wabbit Classifier, a Stochastic Gradient Descent Classifier, two Extreme Gradient Boosted Trees Classifiers, a Gradient Boosted Tree Classifier, a Nystroem Kernel SVM, and a Regularized Logistic Regression was best able to predict extended LOS. The disposition ensemble model had an AUC on the validation set of 0.796 (95% CI, 0.790–0.801), and the LOS ensemble had an AUC of 0.824 (95% CI, 0.823–0.826). Validating on the holdout set yielded an AUC of 0.807 for the disposition ensemble and 0.818 for the length of stay ensemble. When optimizing the F1 score, the discharge ensemble had a positive predictive value (PPV) of 60.0% (95% CI, 59.1%–60.9%), a negative predictive value (NPV) of 82.0% (95% CI, 81.5%–82.4%), sensitivity of 75.5% (95% CI, 75.0%–76.0%), and specificity of 68.9% (95% CI, 67.9%–69.8%). The LOS ensemble had a PPV of 58.8% (95% CI, 58.6%–59.0%), a NPV of 86.2% (95% CI, 86.0%–86.4%), sensitivity of 77.2% (95% CI, 76.9%–77.5%), and specificity 72.5% (95% CI, 72.2%–72.7%). All additional ensembles predicted their respective outcomes with similar discrimination (Table 1).

Permutation Importance

We performed permutation and partial dependence analyses to determine which variables are most important to, and how they independently impact, the ensembles. The strongest risk factors for non-home discharge are, in order: increasing age, preoperative paralysis, hospitalization in the Northeast, non-elective surgery, and preoperative electrolyte or fluid abnormalities. The strongest risk factors for extended LOS in the LOS ensemble are: non-elective surgery, preoperative paralysis, preoperative electrolyte or fluid abnormalities, increasing age, and preoperative neurological deficits (Figure 1).

Bivariate Analysis

Bivariate analysis performed on variables that most strongly influenced the disposition ensemble demonstrated that black and non-black minority patients are more likely than white patients to have non-elective procedures, preoperative paralysis, and preoperative

neurological deficits ($P < .001$ for all). White patients tend to be older (average age 55.5 vs 51.8 (black) vs. 50.7 (other non-black), $P < .001$) and seen more frequently at Northeastern hospitals ($P < .001$) (Table 2).

Bivariate analysis performed on variables that most strongly influenced the LOS ensemble also demonstrated that black and non-black minority patients are more likely than white patients to have non-elective procedures, preoperative paralysis, and preoperative neurological deficits ($P < .001$ for all). Black patients are more likely than white or non-black minority patients to be seen at Southern hospitals ($P < .001$), while white patients tend to be older (average age 55.5 vs 51.8 (black) vs. 50.8 (other non-black), $P < .001$) (Table 3).

Independent Impact of Race

Partial dependence plots were derived to determine the independent impact of race on the disposition and LOS ensembles. Black race increases the risk of non-home discharge over white patients by 6.9% while non-black minority race increases risk of non-home discharge over white patients by only 1.2%. Black race increases the risk of extended LOS over white patients by 6.5%, and non-black minority race increases risk of extended LOS by 6.0%.

Subgroup analysis showed that the effect of minority race is mitigated for both outcomes in patients with benign tumors. Partial dependence of the general inpatient population and general operative population showed very little impact of race for either discharge disposition or length of stay (Figure 2).

Discussion

Using guided ML ensembles, we validated the findings of previous studies that minority race is an independent risk factor for non-home discharge and extended LOS in ensembles that predict discharge disposition and LOS following CFBT. Bivariate analysis demonstrated that minority patients are more likely to have multiple features that put them at higher risk for poor post-operative outcomes, including non-elective surgery, preoperative paralysis, preoperative electrolyte or fluid abnormalities, and preoperative neurological deficits. Importantly, however, the independent impact of race persists in the presence of these and many other covariates that could potentially explain away the importance of race, including age, sex, specific brain tumor diagnosis, primary payer, preoperative comorbidities, admission type, elective vs non-elective surgery, admission day, admission month, and hospital region, size, and type. A critical limitation of our study is the large number of admissions for which a race designation is missing (~20%). Rather than exclude admissions without race, which could bias the data, we chose to include missing race its own category within the race variable. The fact that the independent impact of minority race remains even in the presence of missing race as its own categorical variable suggests that our findings are valid.

Our study employs a novel technique of building guided ML ensembles to predict LOS and discharge disposition. While use of ML in general to model medical outcomes holds great promise for improving patient care,[11] modeling in neurosurgery has tended to rely on logistic regression.[12] Logistic regression is a powerful technique, but is best suited to

datasets with few, independent variables.[13] Given the complexity of disease processes and available patient data, some researchers in neurosurgery have argued that algorithms other than logistic regression may be more adept at answering clinical questions.[14,15] They present no evidence, however, that their chosen algorithms perform better than the dozens of other available ML algorithms. While general recommendations exist for selection of algorithms based on characteristics of a data set or clinical question, the only way to know with certainty that an algorithm is the most predictive is to directly compare it to other algorithms.[16] Using direct comparison of 26 ML algorithms, we identified two ensembles comprised of multiple classes of ML algorithms that best predict our outcomes of interest. Utilizing this guided approach to ML algorithm selection, we objectively build the most predictive ensembles for discharge disposition and LOS from our given data set. By extension, the more predictive our ensembles, the more confident we can be in the interpretation of the impact of race on discharge disposition and LOS derived from the ensemble itself.

Multiple studies have noted that minority patients are more likely to be seen at hospitals that perform fewer craniotomies per year, and suggest that decreased access to specialized care may contribute to racial disparities in post-operative outcomes.[5,17] Yearly hospital case volume is a challenging variable to use in predictive modeling as it must be collected prospectively for each admission. For example, if a patient has a craniotomy in a given hospital in February, the true yearly case volume for craniotomy for that hospital can only be calculated the following January. Using prospectively collected data potentially biases model predictions. To overcome this challenge, researchers may use case volumes from previous years and assume that volume does not vary much year to year. In the NIS database, as in many other national databases, only a subset of hospitals are surveyed each year, making it difficult to make assumptions about the distribution of case volume over multiple years. Due to the challenge of this accurately defining this variable, we did not include hospital case volume in our main analysis. We did build secondary ensembles to predict discharge and LOS that included hospital volume as a covariate and saw no differences between these ensembles and those without the hospital volume variable (Figure A.1). These findings suggest that decreased access of minorities to hospitals that perform many craniotomies per year may not explain the racial disparities we saw in our primary analysis.

What, then, is driving poorer post-operative outcomes for minority patients following CFBT? Minority race can be considered a pseudomarker for low socioeconomic status, and so failure to fully capture socioeconomic status elsewhere in the ensemble may result in disproportionately high impact of race on the ensemble. Mukherjee et al have argued that researchers use poor surrogates for socioeconomic status (such as average income for the admission county), and so looked at the impact of race on postoperative outcomes in a population restricted to Medicaid recipients.[11] Even in this population, they saw that black patients are more likely to have extended LOS following CFBT. We use expected payer (Medicare, Medicaid, private insurance, no pay, and self pay) as our primary surrogate for socioeconomic status, which may not accurately reflect the patient's true socioeconomic status. The findings of Mukherjee and colleagues, however, reinforce our findings that race itself (as opposed to race as a marker for socioeconomic status) does increase the risk of non-home discharge and extended LOS in patients undergoing CFBT.

Other researchers have suggested that minority patients present to their physicians with more progressive disease, resulting in poorer outcomes after surgery than their counterparts who present earlier in the disease process. This could potentially be driven by patient distrust of physicians or the medical system,[18,19] or by provider bias (for example failure to perform diagnostic tests or refer to specialty care early).[2,20–22] We attempt to capture this possibility with several variables, including elective v. non-elective surgery, preoperative paralysis, or other preoperative neurological deficit. These variables are, however, far from comprehensive, and may not fully describe disease severity at time of presentation.

There is some evidence to suggest that minority patients experience more post-operative complications than white patients. One limitation of our study is the fact that we include only preoperative variables in our analysis, and so cannot comment on the effect of race on the discharge disposition or LOS ensembles in the presence of postoperative complications.[23–26] The researchers who noted increased complications in minority patients have proposed that this could be driven in part by worse baseline health for minorities. We do attempt to control for baseline health by including a fairly comprehensive array of 29 preoperative comorbidities in our ensembles.

We studied the impact of race on disposition and LOS separately. It is possible, however, that these outcomes are interrelated. It could be the case, for example, that increased LOS seen in minority patients is due in part to increased non-home discharge, which may necessitate increased hospital days as patients await placement. To investigate this possibility, we calculated the percentage of patients with extended LOS for each race group for home vs non-home discharge. We found that, even for home discharge, a larger percentage of minority patients (black = 34.5%, other = 36.9%) than white patients (18.6%) had extended LOS (Figure 3). These data suggest that more non-home discharge in minority patients does not drive increased LOS in this population, and that discharge disposition and LOS are independently impacted by race.

We were surprised to see that our findings of minority race as an independent risk factor for non-home discharge and extended LOS were not mimicked in the general inpatient or general operative population. This is a novel finding of our study not previously reported in the literature. This suggests some specificity of these disparities to the CFBT population. One possible explanation for this finding is that racial disparities in postoperative outcomes in general are exacerbated in such a specialized field of care. Regardless, this does point to a specific point of intervention for quality improvement initiatives in neurosurgery.

While our study has many strengths, we note several limitations not previously addressed: (1) the NIS database relies on ICD-9 coding, which is subject to the individual interpretation of coders and which may contain errors; (2) the NIS database contains missing data, a challenge we have attempted to address by including "missing" as a categorical variable in our analysis, but which nonetheless may impact results; (3) the NIS database contains a finite number of variables, and so we are unable to address all possible explanations for disparities in individual outcomes, including tumor-specific characteristics; (4) the NIS database comprises only hospital admissions in the United States, and so specific conclusions drawn from this study may not be applicable in other countries. We believe,

however, that the general concept that minority groups experience differential outcomes is important, and the disparities in outcome noted here should be assessed in countries outside of the United States.

Conclusions

Minority race independently increases the risk of extended LOS and increases the risk of non-home discharge in patients undergoing brain tumor resection, a finding not mimicked in the general inpatient or operative population. Recognition of the influence of race on discharge and LOS could generate interventions, such as increasing education for neurosurgical providers on the existence of racial disparities or early social work or case management assessment of minority patients at hospital admission, that may improve outcomes for this population and enhance quality of neurosurgical care.

Acknowledgments

Abbreviations

NIS	National Inpatient Sample
ML	machine learning
CFBT	craniotomy for brain tumor
LOS	length of stay
AHRQ	Agency for Healthcare Research and Quality
ROC	receiver-operating characteristic
AUC	area under the curve
CI	confidence interval

References

1. Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. [Accessed 23.11.16] 2015 National Healthcare Disparities Report and Disparities Report and 5th Anniversary Update on the National Quality Strategy. 2016. http://www.ahrq.gov/research/findings/nhqrdr/nhqdr15/index.html

2. Smedley, BD.Stith, AY., Nelson, AR., editors. Unequal treatment: confronting racial and ethnic disparities in health care. Washington, D.C: National Academies Press; 2003.

3. Curry WT Jr, Barker FG 2nd. Racial, ethnic, and socioeconomic disparities in the treatment of brain tumors. J Neurooncol. 2009; 93(1):25–39. [PubMed: 19430880]

4. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, D.C: National Academies Press; 2001.

5. Curry WT Jr, Carter BS, Barker FG 2nd. Racial, ethnic, and socioeconomic disparities in patient outcomes after craniotomy for tumor in adult patients in the United States, 1988–2004. Neurosurgery. 2010; 66(3):427–437. [PubMed: 20124933]

6. Dasenbrock HH, Liu KX, Devine CA, Chavakula V, Smith TR, Gormley WB, et al. Length of hospital stay after craniotomy for tumor: a National Surgical Quality Improvement Program analysis. Neurosurg Focus. 2015; 39(6):E12.

7. Agency for Healthcare Research and Quality. [Accessed 16.11.16.] Healthcare Cost and Utilization Project Databases: Nationwide Inpatient Sample. 2016. http://www.hcup-us.ahrq.gov/nisoverview.jsp

8. Kang J, Schwartz R, Flickinger J, Beriwal S. Machine learning approaches for prediction radiation therapy outcomes: a clinician's perspective. Int J Radiat Oncol Biol Phys. 2015; 93(5):1127–1135. [PubMed: 26581149]

9. Breiman L. Random forests. ML. 2001; 45:5–32.

10. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001; 29(5): 1189–1232.

11. Darcy AM, Louie AK, Roberts LW. ML and the profession of medicine. JAMA. 2016; 15(6):551–552.

12. Deo RC. ML in medicine. Circulation. 2015; 132:1920–1930. [PubMed: 26572668]

13. Bibault JE, Giraud P, Burgun A. Big Data and machine learning in radiation oncology: State of the art and future prospects. Cancer Lett. 2016; 382(1):110–117. [PubMed: 27241666]

14. Oermann EK, Kress MS, 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. 2013; 72(6):944–951. [PubMed: 23467250]

15. Shi HY, Hwang SL, Lee KT, Lin CL. In-hospital mortality after traumatic brain injury surgery: a nationwide population-based comparison of mortality predictors used in artificial neural network and logistic regression models. J Neurosurg. 2013; 118(4):746–752. [PubMed: 23373802]

16. Tanwani, AK., Afridi, J., Shafiq, MZ., Farooq, M. Guidelines to select ML scheme for classification of biomedical datasets. In: Pizzuti, C.Ritchie, MD., Giacobini, M., editors. Evolutionary Computation, ML, and Data Mining in Bioinformatics. Berlin Heidelberg: Springer Verlag; 2009. p. 128-139.

17. Mukherjee D, Patil CG, Todnem N, Ugiliweneza B, Nuno M, Kinsman M, et al. Racial disparities in Medicaid patients after brain tumor surgery. J Clin Neurosci. 2013; 20(1):57–61. [PubMed: 23084348]

18. Armstrong K, McMurphy S, Dean LT, Micco E, Putt M, Halbert CH, et al. Differences in the patterns of health care system distrust between blacks and whites. J Gen Intern Med. 2008; 23(6): 827–833. [PubMed: 18299939]

19. Armstrong K, Rose A, Peters N, Long JA, McMurphy S, Shea JA. Distrust of the health care system and self-reported health in the United States. J Gen Intern Med. 2006; 21(4):292–297. [PubMed: 16686803]

20. Green AR, Carney DR, Pallin DJ, Ngo LH, Raymond KL, Iezzoni LI, et al. Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. J Gen Intern Med. 2007; 22(9):1231–1238. [PubMed: 17594129]

21. Sabin JA, Rivara FP, Greenwald AG. Physician implicit attitudes and stereotypes about race and quality of medical care. Med Care. 2008; 46(7):678–685. [PubMed: 18580386]

22. Sabin J, Nosek BA, Greenwald A, Rivara FP. Physicians' implicit and explicit attitudes about race by MD race, ethnicity, and gender. J Health Care Poor Underserved. 2009; 20(3):896–913. [PubMed: 19648715]

23. Vogel TR, Dombrovskiy VY, Carson JL, Graham AM, Lowry SF. Postoperative sepsis in the United States. Ann Surg. 2010; 252:1065–1071. [PubMed: 20571363]

24. Ibrahim SA, Stone RA, Han X, Cohen P, Fine MJ, Henderson WG, et al. Racial/ethnic differences in surgical outcomes in veterans following knee or hip arthroscopy. Arthritis Rheum. 2005; 52:3143–3151. [PubMed: 16200594]

25. Lucas FL, Stukel TA, Morris AM, Siewers AE, Birkmeyer JD. Race and surgical mortality in the United States. Ann Surg. 2006; 243(2):281–286. [PubMed: 16432363]

26. Brooks Carthon JM, Jarrin O, Sloane D, Kutney-Lee A. Variations in postoperative complications according to race, ethnicity, and sex in older adults. J Am Geriatr Soc. 2013; 61(9):1499–1507. [PubMed: 24006851]

Highlights

- A novel machine learning technique models important neurosurgical outcomes.

- Minority race increases risk of non-home discharge after surgery for brain tumor.

- Minority race increases risk of extended length of stay in the same population.

- These findings are not mimicked in the general inpatient or operative populations.
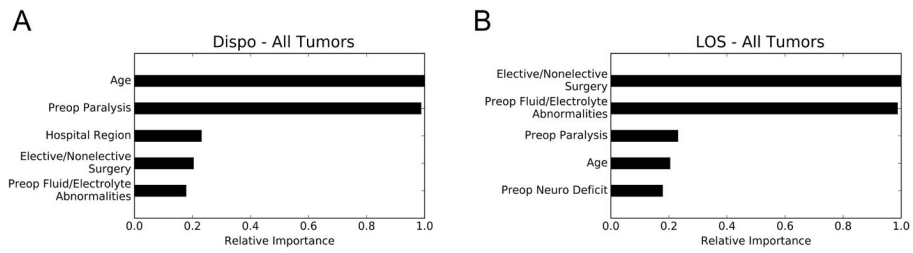
Figure 1. Permutation Importance

Permutation importance demonstrating the relative importance of individual variables to the disposition and LOS ensembles. The most important variable is given an importance value of 1.0 and the importance of other variables is shown relative to 1.0. Dispo, disposition; LOS, length of stay; Preop, preoperative.
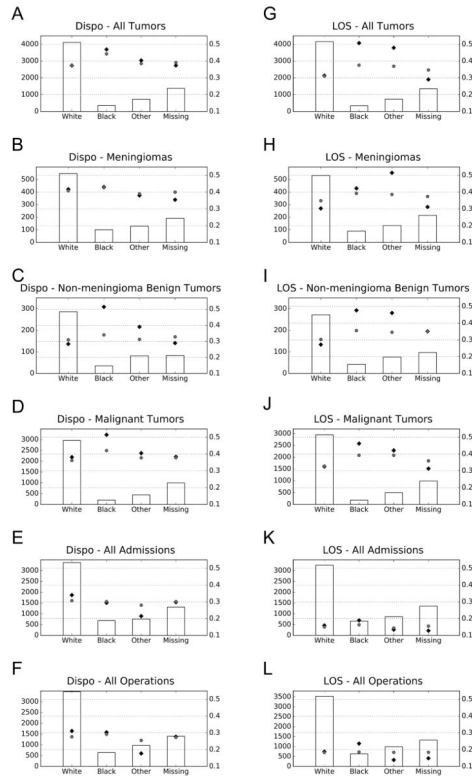
Figure 2. Partial Dependence Plots

Partial dependence analysis demonstrating the independent impact of race on disposition and LOS for different patient populations. Left X-axis represents patient incidence for each patient group and corresponds to bars. Right X-axis represents probability of non-home discharge or extended LOS, with 1 equivalent to 100% likelihood of non-home discharge or extended LOS and 0 equivalent to 0% likelihood of non-home discharge or extended LOS and corresponds to gray and black round and diamond heads. For each group, black diamond heads represents average probability of the outcomes of interest calculated solely from the raw data (e.g. the non-independent impact of each variable); gray round heads represent average probability of the outcome of interest calculated from the partial dependence (e.g. the independent impact of each variable). Panels A–D and G–J represent the impact of race on various populations of patients with intracranial tumors; panels E and K represent the impact of race on a random sample of all admissions from the NIS database; panels F and L represent the impact of race on a random sample of all operations from the NIS database. Dispo, disposition; LOS, length of stay; White, white race; Black, black race; Other, non-white, non-black race; Missing, missing race.
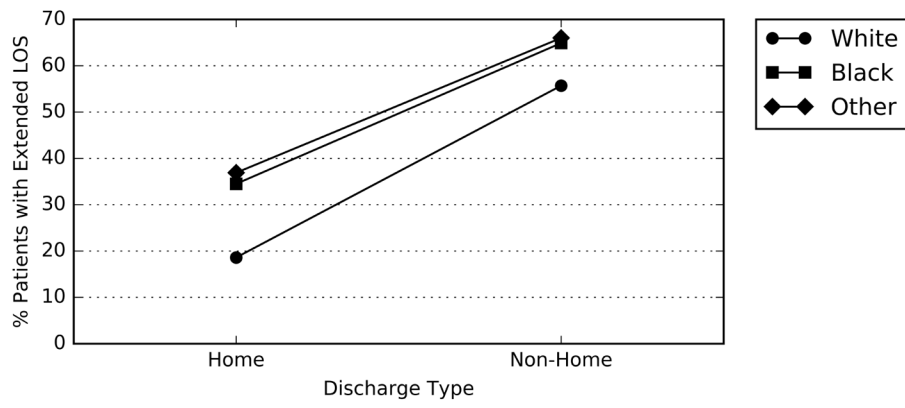
Figure 3. Extended length of stay for different discharge locations by race
Percentage of patients with length of stay for different discharge locations plotted by race.

LOS, length of stay.
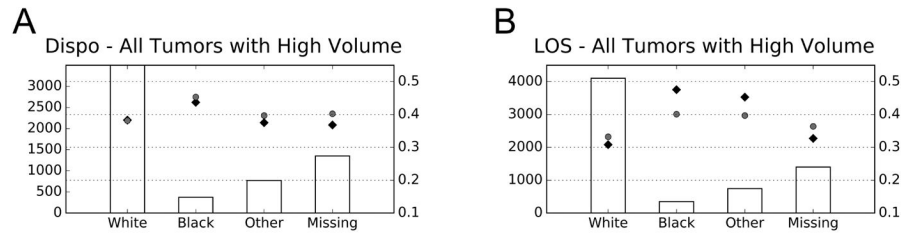
Figure A.1. Partial Dependence Plots for Ensembles Including High Volume Variable
Partial dependence analysis demonstrating the independent impact of race on disposition and LOS for different patient populations. Left X-axis represents patient incidence for each patient group and corresponds to bars. Right X-axis represents probability of non-home discharge or extended LOS, with 1 equivalent to 100% likelihood of non-home discharge or extended LOS and 0 equivalent to 0% likelihood of non-home discharge or extended LOS and corresponds to gray and black round and diamond heads. For each group, black diamond heads represents average probability of the outcomes of interest calculated solely from the raw data (e.g. the non-independent impact of each variable); gray round heads represent average probability of the outcome of interest calculated from the partial dependence (e.g. the independent impact of each variable). Dispo, disposition; LOS, length of stay; White, white race; Black, black race; Other, non-white, non-black race; Missing, missing race.

Table 1

Cross-validation AUC and other metrics for constructed ensembles.

Ensemble and patient population	AUC (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Holdout
Discharge						
CFBT	0.796 (0.790–0.801)	60.0 (59.1–60.9)	82.0 (81.5–82.4)	75.5 (75.0–76.0)	68.9 (67.9–69.8)	0.807
Meningioma	0.789 (0.777–0.799)	57.8 (56.7–58.7)	81.3 (79.7–82.6)	78.7 (76.2–80.8)	61.6 (59.6–63.4)	0.804
Non-meningioma benign tumor	0.761 (0.744–0.776)	50.0 (47.3–51.9)	83.2 (81.9–84.4)	71.4 (69.1–73.5)	66.3 (63.4–68.8)	0.782
Malignant tumor	0.793 (0.788–0.797)	60.2 (59.9–60.4)	82.2 (81.9–82.5)	77.2 (76.5–77.7)	67.3 (66.8–67.8)	0.793
All admissions	0.797 (0.791–0.803)	53.8 (52.9–54.5)	85.3 (84.6–85.9)	72.3 (70.8–73.7)	72.0 (71.2–72.7)	0.801
All operations	0.835 (0.827–0.841)	54.6 (53.4–55.5)	88.6 (88.0–89.2)	75.5 (74.2–76.7)	75.2 (74.4–75.9)	0.827
CFBT with high volume hospital variable	0.800 (0.789–0.809)	62.0 (60.8–63.0)	81.5 (80.7–82.2)	73.5 (72.2–74.6)	72.1 (70.9–73.4)	0.798
Length of Stay						
CFBT	0.824 (0.823–0.826)	58.8 (58.6–59.0)	86.2 (86.0–86.4)	77.2 (76.9–77.5)	72.5 (72.2–72.7)	0.818
Meningioma	0.815 (0.809 – 0.820)	58.7 (60.0–60.3)	86.4 (85.5–87.2)	78.7 (77.0–80.2)	70.9 (68.7–72.8)	0.811
Non-meningioma benign tumor	0.788 (0.761–0.812)	56.0 (52.8–58.8)	84.1 (81.5–86.4)	73.3 (68.1–77.7)	71.0 (68.1–74.5)	0.779
Malignant tumor	0.819 (0.815–0.823)	55.8 (55.1–56.3)	88.2 (87.8–88.5)	82.6 (81.9–83.2)	66.4 (65.5–67.2)	0.817
All admissions	0.751 (0.742–0.758)	33.4 (32.8–34.0)	90.2 (90.0 – 90.4)	50.0 (48.3–51.1)	82.2 (81.9–82.5)	0.751
All operations	0.823 (0.816–0.829)	41.7 (40.6–42.7)	91.7 (91.2–92.1)	66.4 (64.4–68.1)	80.0 (79.6–80.4)	0.828
CFBT with high volume hospital variable	0.823 (0.819–0.826)	56.4 (55.6–57.1)	87.6 (87.1–88.1)	81.1 (80.1–82.0)	68.1 (67.1–68.9)	0.825

Cross-validation receiver operating characteristic and other metrics for each ensemble. Metrics are calculated at the optimized F1 score for each ensemble. The holdout for each ensemble is taken as a single sample, and so no confidence intervals are calculated. AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; CFBT, craniotomy for brain tumor.

Table 2

Patient characteristics by race for most informative variables included in the disposition ensemble.

Variable	All (n=41111)	White (n=25678)	Black (n=2174)	Other race (n=4592)	Missing (n=8667)	P (all) *	P (black v. other) †
Discharge, n (%)							
Home	25406 (61.8)	15935 (62.1)	1182 (54.4)	2831 (61.7)	5458 (63.0)	<.001	<.001
Non-home	15705 (38.2)	9743 (37.1)	992 (45.6)	1761 (38.3)	3209 (37.0)		
Age (SD)	54.4 (15.9)	55.5 (15.9)	51.8 (15.1)	50.7 (16.2)	53.8 (15.8)	<.001	0.01
Admission type, n (%)							
Non-elective	17561 (42.7)	10622 (41.4)	1144 (52.6)	2410 (52.5)	3385 (39.1)	<.001	.95
Elective	23491 (57.1)	15032 (58.5)	1027 (47.3)	2171 (47.3)	5261 (60.7)		
Missing	59 (.2)	24 (.1)	3 (.1)	11 (.2)	21 (.2)		
Hospital Region, n (%)							
Northeast	8382 (20.4)	6506 (25.3)	488 (22.4)	1074 (23.4)	314 (3.6)	<.001	.39
Other	32729 (79.6)	19172 (74.7)	1686 (77.6)	3518 (76.6)	8353 (96.4)		
Preop paralysis, n (%)							
Yes	5274 (12.8)	3220 (12.5)	311 (14.3)	706 (15.4)	1037 (12.0)	<.001	.34
No	35419 (86.2)	22122 (86.2)	1827 (84.0)	3866 (84.2)	7604 (87.7)		
Missing	418 (1.0)	336 (1.3)	36 (1.7)	20 (.4)	26 (.3)		
Preop fluid/electrolyte abnormalities, n (%)							
Yes	5043 (12.8)	2942 (11.5)	335 (15.4)	679 (14.8)	1087 (12.5)	<.001	.38
No	35650 (86.7)	22400 (87.2)	1803 (82.9)	3893 (84.8)	7554 (87.2)		
Missing	418 (1.0)	336 (1.3)	36 (1.7)	20 (.4)	26 (.3)		
Preop neuro deficits, n (%)							
Yes	680 (1.7)	386 (1.5)	62 (2.9)	97 (2.1)	135 (15.6)	<.001	.05
No	40013 (97.3)	24956 (97.2)	2076 (95.5)	4475 (97.5)	8506 (98.1)		
Missing	418 (1.0)	336 (1.3)	36 (1.6)	20 (.4)	26 (.3)		

Patient characteristics by race of the variables that most strongly impact the discharge disposition ensemble.

*
Mann-Whitney U test and Pearson's χ^2 test for significant difference between groups. Chi square analyses do not include missing data, but does include white, black, and other races. SD, standard deviation; y, years; NFP, not for profit.

†Chi square analyses only include black and other races.

Table 3

Patient characteristics by race for most informative variables included in the length of stay ensemble.

Variable	All (n=41221)	White (n=25747)	Black (n=2186)	Other race (n=4613)	Missing (n=8675)	*P*-value (all) [*]	*P*-value (black v. other) [†]
LOS, n (%)							
> 7 days	13907 (33.7)	8019 (31.1)	1034 (47.3)	2079 (45.1)	2775 (32.0)	< .001	.08
<= 7 days	27314 (66.3)	17728 (68.9)	1152 (52.7)	2534 (54.9)	5900 (68.0)		
Age (SD)	54.4 (15.9)	55.5 (15.9)	51.8 (15.2)	50.8 (16.2)	53.8 (15.8)	< .001	.01
Admission type, n (%)							
Non-elective	17616 (42.7)	15063 (58.5)	1033 (47.3)	2184 (47.3)	5264 (60.7)	< .001	.92
Elective	23544 (57.1)	10660 (41.4)	1150 (52.6)	2418 (52.4)	3388 (39.1)		
Missing	61 (.2)	24 (.1)	3 (.1)	11 (.3)	23 (.2)		
Hospital Region, n (%)							
South	15390 (37.3)	9800 (38.1)	1222 (55.9)	1741 (37.7)	2627 (30.3)	< .001	< .001
Other	25832 (62.7)	15947 (61.9)	965 (44.1)	2872 (62.3)	6048 (69.7)		
Preop paralysis, n (%)							
Yes	5304 (12.9)	3242 (12.6)	311 (14.2)	712 (15.4)	1039 (12.0)	< .001	.23
No	35499 (86.1)	22169 (86.1)	1839 (84.1)	3881 (84.1)	7610 (87.7)		
Missing	418 (1)	336 (1.3)	36 (1.7)	20 (.4)	26 (.3)		
Preop fluid/electrolyte abnormalities, n (%)							
Yes	5056 (12.3)	2949 (11.5)	337 (15.4)	682 (14.8)	1088 (12.5)	<.001	.38
No	35747 (86.7)	22462 (87.2)	1813 (82.9)	3911 (84.8)	7561 (87.2)		
Missing	418 (1)	336 (1.3)	36 (1.7)	20 (.4)	26 (.3)		
Preop neuro deficits, n (%)							
Yes	684 (1.7)	388 (1.5)	62 (2.8)	99 (2.1)	135 (1.6)	<.001	.07
No	40119 (97.3)	25023 (97.2)	2088 (95.5)	4494 (97.4)	8514 (98.1)		
Missing	418 (1)	336 (1.3)	36 (1.6)	20 (.4)	26 (.3)		

Patient characteristics by race of the variables that most strongly impact the length of stay ensemble.

[*]
Mann-Whitney U test and Pearson's χ^2 test for significant difference between groups. Chi square analyses do not include missing data, but does include white, black, and other races. LOS, length of stay; SD, standard deviation; y, years; NFP, not for profit.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

†Chi square analyses only include black and other races.

Table A.1

Variables included in ensemble training

Patient Characteristics:	Hypothyroidism	Admission month
Age	Peripheral vascular disease	**ICD9 Diagnoses:**
Sex	Drug abuse	191.0
Race	Valvular disease	191.1
Fluid/electrolyte abnormalities	Liver disease	191.2
Paralysis	Obesity	191.3
Other neurological deficit	Chronic blood loss anemia	191.4
Hypertension	AIDS	191.5
Deficiency anemias	Peptic ulcer disease	191.6
Diabetes mellitus (no complications)	Lymphoma	191.7
Diabetes mellitus (with complications)	Arthritis	191.8
Coagulopathies	**Hospital characteristics:**	191.9
Weight loss	Geographic region	199.1
Chronic lung disease	Ownership (private, government, etc)	225.0
Solid tumor without metastases	Location/teaching status (e.g. urban teaching, rural non-teaching)	225.1
Congestive heart failure	Bed size	225.2
Psychosis	**Admission characteristics:**	225.3
Metastatic cancer	Admission type (emergency, urgent, trauma, etc)	225.4
Depression	Elective/non-elective	225.8
Alcohol abuse	Primary expected payer	225.9
Renal failure	Secondary expected payer	
Pulmonary circulation disorders	Weekend/weekday admission	

Table A.2

Patient characteristics for admissions included in disposition ensemble

Variable	Total	Discharged to home	Not discharged to home	*P* *
Surgeries, n (%)	41111	25406 (61.8)	15705 (38.2)	...
Sex, n (%)				
Male	21570 (52.5)	13680 (63.4)	7890 (36.6)	< .001
Female	19363 (47.1)	11567 (59.7)	7796 (40.3)	
Missing	178 (.4)	159 (89.3)	19 (.7)	
Age at surgery, mean (SD), y	54.4 (15.9)	50.5 (15.0)	60.7 (15.3)	<.001
Race, n (%)				
White	25678 (62.5)	15935 (62.1)	9743 (37.9)	<.001
Black	2174 (5.3)	1182 (54.4)	992 (45.6)	
Other	4592 (11.1)	2831 (61.7)	1761 (38.3)	
Missing	8667 (21.1)	5458 (63.0)	3209 (37.0)	
Expected payer, n (%)				
Private	22097 (53.7)	15719 (71.1)	6378 (28.9)	<.001
Medicare	12074 (29.4)	4989 (41.3)	7085 (58.7)	
Medicaid	3754 (9.1)	2353 (62.7)	1401 (37.3)	
Self pay	1591 (3.9)	1216 (76.4)	375 (23.6)	
Other	1345 (3.3)	938 (69.7)	407 (30.3)	
No charge	185 (.4)	147 (79.5)	38 (20.5)	
Missing	65 (.2)	44 (67.7)	21 (32.3)	
Admission type, n (%)				
Elective	23491 (57.1)	16525 (70.3)	6966 (29.7)	<.001
Non-elective	17561 (42.7)	8839 (50.3)	8722 (49.7)	
Missing	59 (.2)	42 (71.2)	17 (28.8)	
Hospital Region, n (%)				
South	15324 (37.3)	10135 (66.1)	5189 (33.9)	<.001
West	9152 (22.2)	5746 (62.8)	3406 (37.2)	
Northeast	8382 (20.4)	4601 (54.9)	3781 (45.1)	
Midwest	8253 (20.1)	4924 (59.7)	3329 (40.3)	
Hospital Control, n (%)				
Government or private	34188 (83.2)	21465 (62.8)	12723 (37.2)	<.001
Private NFP	4425 (10.8)	2573 (58.1)	1852 (41.9)	
Private investor	1454 (3.5)	782 (53.7)	672 (46.2)	
Government/non-federal	753 (1.8)	422 (56.0)	331 (44.0)	
Private NFP or investor	291 (.7)	164 (56.4)	127 (43.6)	
Hospital location/teaching, n (%)				
Urban teaching	31584 (76.8)	20046 (63.5)	11538 (36.5)	<.001

Variable	Total	Discharged to home	Not discharged to home	P *
Urban nonteaching	8528 (20.8)	4757 (55.8)	3771 (44.2)	
Rural	999 (2.4)	603 (60.4)	396 (39.6)	
Hospital bed size, n (%)				
Large	32579 (79.3)	20190 (62.0)	12389 (38.0)	.02
Medium	6221 (15.1)	3756 (60.4)	2465 (39.6)	
Small	2311 (5.6)	1460 (63.2)	851 (36.8)	
Preoperative comorbidity [†], n (%)				
Yes	28506 (69.3)	15258 (53.5)	13248 (46.5)	< .001
No	12187 (29.7)	9898 (81.2)	2289 (18.8)	
Missing	418 (1.0)	250 (59.8)	168 (40.2)	

[*] Mann-Whitney U test and Pearson's χ^2 test for significant difference between groups. Chi square analyses do not include missing data.

[†] Indicates presence of at least one preoperative comorbidity. SD, standard deviation; y, years; NFP, not for profit.

Table A.3

Patient characteristics for admissions included in LOS ensemble

Variable	Total	LOS <= 7 days	LOS > 7 days	*P* *
Surgeries, n (%)	41221	27314 (66.3)	13907 (33.7)	…
Sex, n (%)				
Male	21617 (52.4)	14337 (66.3)	7280 (33.7)	0.4
Female	19426 (47.1)	12808 (66.0)	6618 (34.0)	
Missing	179 (.4)	169 (94.4)	9 (5.6)	
Age at surgery, mean (SD), y	54.4 (15.9)	52.7 (15.5)	57.8 (16.1)	< .001
Race, n (%)				
White	25747 (62.5)	17728 (68.9)	8019 (31.1)	<.001
Black	2186 (5.3)	1152 (52.7)	1034 (47.3)	
Other	4613 (11.2)	2534 (54.9)	2079 (45.1)	
Missing	8676 (21.0)	5900 (68.0)	2775 (32.0)	
Expected payer, n (%)				
Private	22136 (53.7)	16418 (74.2)	5718 (25.8)	< .001
Medicare	12128 (29.4)	6912 (57.0)	5216 (43.0)	
Medicaid	3763 (9.1)	2074 (55.1)	1689 (44.9)	
Self pay	1595 (3.9)	896 (56.2)	699 (43.8)	
Other	1349 (3.3)	869 (64.4)	480 (35.6)	
No charge	185 (.4)	104 (56.2)	81 (43.8)	
Missing	65 (.2)	41 (63.1)	24 (36.9)	
Admission type, n (%)				
Elective	23544 (57.1)	19423 (82.5)	4121 (17.5)	< .001
Non-elective	17616 (42.7)	7857 (44.6)	9759 (55.4)	
Missing	61 (.2)	34 (55.7)	27 (44.3)	
Hospital Region, n (%)				
Northeast	8385 (20.3)	5516 (65.8)	2869 (34.2)	< .001
South	15390 (37.3)	9850 (64.0)	5540 (36.0)	
Midwest	8261 (20.1)	5624 (68.1)	2637 (31.9)	
West	9185 (22.3)	6324 (68.9)	2861 (31.1)	
Hospital Control, n (%)				
Government or private	34240 (83.1)	22963 (67.1)	11277 (32.9)	< .001
Private NFP	4457 (10.8)	2895 (65.0)	1562 (35.0)	
Private investor	1472 (3.6)	807 (54.8)	665 (45.2)	
Private NFP or investor	291 (.7)	174 (59.8)	117 (40.2)	
Government nonfederal	761 (1.8)	475 (62.4)	286 (37.6)	
Hospital location/teaching, n (%)				
Urban teaching	31636 (76.7)	21313 (67.4)	10323 (32.6)	< .001

Variable	Total	LOS <= 7 days	LOS > 7 days	P *
Urban nonteaching	8584 (20.8)	5330 (62.1)	3254 (37.9)	
Rural	1001 (2.5)	671 (67.0)	330 (33.0)	
Hospital bed size, n (%)				
Large	32674 (79.3)	21568 (66.0)	11106 (34.0)	< .001
Medium	6231 (15.1)	4025 (64.6)	2206 (35.4)	
Small	2316 (5.6)	1721 (74.3)	595 (25.7)	
Preoperative comorbidity †, n (%)				
Yes	28602 (69.4)	17005 (59.5)	11597 (40.5)	< .001
No	12201 (29.6)	10019 (82.1)	2182 (17.9)	
Missing	418 (1.0)	290 (69.4)	128 (30.6)	

*
Mann-Whitney U test and Pearson's χ^2 test for significant difference between groups. Chi square analyses do not include missing data.

†
Indicates presence of at least one preoperative comorbidity. LOS, length of stay; SD, standard deviation; y, years; NFP, not for profit.