Preprints.org

Article

Optimizing a Multivariable Logistic Regression for Identification of Perioperative Risk Factors Associated with Deep Brain Stimulator Explantation: A Retrospective Cohort Study

Peyton J Murin , Anagha S Prabhune , Yuri Chaves Martins *

Article

Optimizing a Multivariable Logistic Regression for Identification of Perioperative Risk Factors Associated with Deep Brain Stimulator Explantation: A Retrospective Cohort Study

Peyton J. Murin [1], Anagha S. Prabhune [1] and Yuri Chaves Martins [2,*]

[1] Department of Neurology, Saint Louis University School of Medicine, St. Louis, MO, USA
[2] Department of Anesthesiology, Saint Louis University School of Medicine, St. Louis, MO, USA
[*] Correspondence: ychavesmartins@slu.edu; Tel.: +1 314-577-8750. Fax: +1 314-977-1617

Abstract: Background/Objectives: Deep brain stimulation (DBS) is an effective surgical treatment for Parkinson's disease (PD) and other movement disorders. Despite its benefits, DBS explantation occurs in 5.6% of cases, with costs exceeding $22,000 USD per implant. Traditional statistical methods have struggled to identify reliable risk factors for explantation. We hypothesized that supervised machine learning could better account for complex interactions among perioperative factors, enabling the identification of novel risk factors. **Methods**: The Medical Informatics Operating Room Vitals and Events Repository was queried for patients with DBS, adequate clinical data, and at least two years of follow-up (n = 38). Fisher's exact test assessed demographic and medical history variables. Data was analyzed using Python with pandas, numpy, sklearn, sklearn-extra, matlab, and seaborn. Recursive feature elimination with cross-validation (RFECV) optimized factor selection. A multivariate logistic regression model was trained and evaluated using precision, recall, F1-score, and area under the curve (AUC). **Results**: Fisher's exact test identified chronic pain (p = 0.0108) and tobacco use (p = 0.0026) as risk factors. RFECV selected 24 optimal features. The logistic regression model demonstrated strong performance (precision: 0.89, recall: 0.86, F1-score: 0.86, AUC: 1.0). Significant risk factors included tobacco use (OR: 3.64; CI: 3.60 – 3.68), primary PD (OR: 2.01; CI: 1.99 – 2.02), ASA score (OR: 1.91; CI: 1.90 – 1.92), chronic pain (OR: 1.82; CI: 1.80 – 1.85), and diabetes (OR: 1.63; CI: 1.62 – 1.65). **Conclusions**: Our study suggests that supervised machine learning can identify risk factors for early DBS explantation. Larger studies are needed to validate our findings.

Keywords: Deep Brain Stimulation; Statistical Modeling; Movement Disorders; Surgical Complications

1. Introduction

Deep brain stimulation was first shown to be effective as a therapeutic option in essential tremor and Parkinson's Disease (PD) in 1987 [1]. Since its initial use, the indications for DBS have expanded to include generalized dystonia, tremors, chronic pain, psychiatric disorders, and early PD leading to a growing number of procedures [2-6]. Advancements in DBS technology, such as enhanced battery life, adjustable stimulation parameters, and directional stimulation, have further improved its efficacy [7]. In addition, ongoing developments, including closed-loop adaptive stimulation, continue to refine and expand its clinical applications [7].

However, although DBS is generally considered a safe procedure [8], complications can still occur. Previous studies found a mortality rate of 0.2% and a permanent morbidity rate of 0.6%, with 12.2% of patients having hardware complications [2]. Adverse effects of DBS include cognitive decline [9], ineffective stimulation/device dysfunction [10], infection [10], delirium [11], and

hemorrhage [11] amongst others. In addition, with a cost of over $22,000 USD for the medical care associated with explantation, early explant poses a significant burden to both patients and the healthcare system [8]. Due to these factors, prior literature has identified explant rates of up to 5.6% [12], suggesting a need to improved identification of risk factors for explant. While previous literature has identified risk factors for DBS complication including advanced age, male sex, and comorbidities such as dementia, hypertension, and coagulopathy, only elevated body mass index (BMI) has been demonstrated to be an risk factors for explantation itself [13, 14]. As such, there remains a need to precisely determine which patients are at high risk for early explantation. To this end, we applied a novel approach using supervised machine learning to identify risk factors for DBS explant within the first two-years following placement.

2. Materials and Methods

The study was designed in accordance with TRIPOD-AI guidelines for reporting clinical prediction models that use regression or machine learning methods.

2.1. Data

Data was obtained via a Data Use Agreement between the authors and University of California Irvine Medical Center. Data was collected from the Medical Informatics Operating room Vitals and Events Repository (MOVER) database [15] which contains de-identified electronic health records data from 58,799 unique patients who underwent surgery at the University of California Irvine Medical Center. Data within the MOVER database is de-identified in accordance with HIPAA Privacy Rule, therefore patient consent was not obtained

2.2. Participants

Records included in the study involved patients with a DBS stimulator placement at a single academic center. Treatment consisted of DBS placement.   Inclusion criteria: adult (≥18 years of age), DBS procedure, and at least two years of follow up. Exclusion Criteria: <18 years of age, <2 years of follow up, no DBS procedure. 38 unique patients were included in the study. (Figure 1)
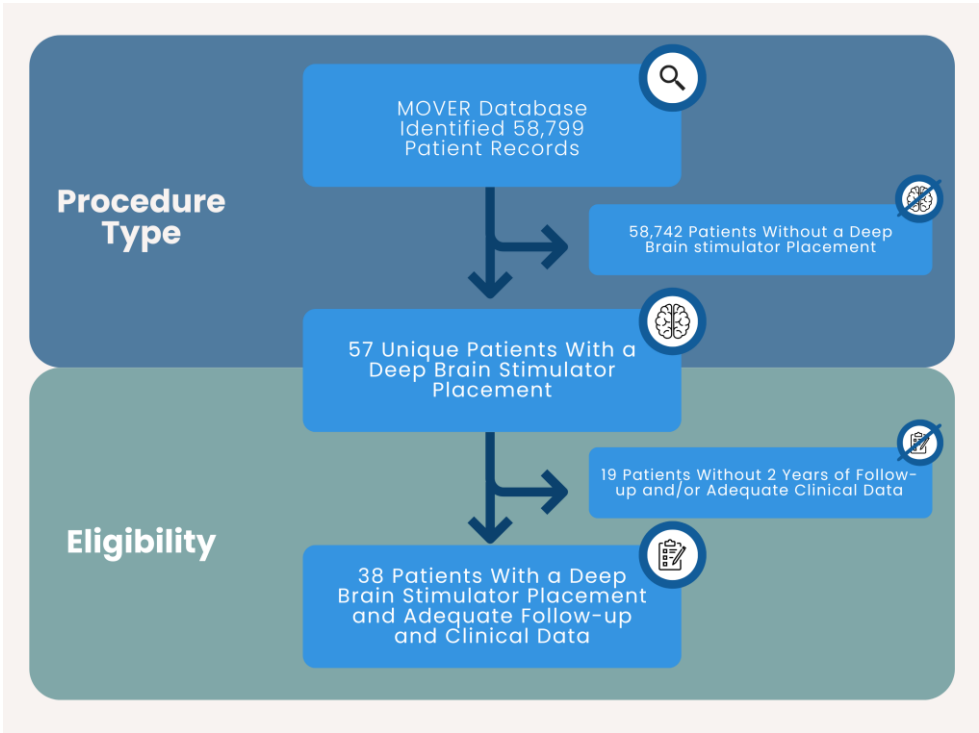

Figure 1. The cohort selection process yielded 38 patients with a DBS procedure and adequate clinical data.

2.3. Data Preparation

Variables were created using International Classification of Diseases (ICD)-10-CM codes, patient demographics, American Society of Anesthesiologist (ASA) score, anesthesia type, and postoperative events including hospitalizations, and intensive care unit (ICU) admissions. The included codes and/or definitions for each of the study variables can be seen in Table S1. Sex was encoded 1: female, 0: male. Medical comorbidities were encoded 1: present, 0: absent. Anesthesia type was encoded 1: monitored anesthesia care, 0: general anesthesia. ICU admission was encoded 1: yes, 0: no. LOS and ASA score was encoded as the numerical value.

2.4. Predictors

Predictors were chosen broadly, as the methodology using recursive factor elimination with cross validation would remove variables offering limited predictive value in a data-driven manner.

2.5. Sample Size

All patients meeting inclusion/exclusion criteria were included. The study cohort consisted of 38 patient records.

2.6. Missing Data

No records had missing data.

2.7. Analytical Methods

The data analysis plan was written prior to accessing the data. Continuous variables were reported as means ± standard deviation (SD) or medians ± interquartile range (IQR). Categorical variables were reported as n-values and percentages. Initial analysis consisted of comparing patients with DBS explantation to those without explantation using Fisher's exact test for categorical variables and the Mann-Whitney U test for continuous variables. Statistical significance was set at $p < 0.05$. Next, the data was imported into Anaconda Version 2.3.1. (Anaconda Software Distribution. Austin, TX). The following add-ons were used for analysis: pandas [16], numpy, sklearn [17], scikit-learn-extra[17], MatLab (The MathWorks, Inc., Natick, MA), and seaborn [18]. The target variable (explantation or no explantation) was defined, the features were scaled, and Synthetic Minority Oversampling Technique (SMOTE) was applied to balance the classes. To define the optimal features to include in the model, recursive feature elimination with cross validation (RFECV) was imported from sklearn.feature_selection and cross_val_score was imported from sklearn.model_selection [17]. Five iterations of RFECV with different data splits were performed and the ideal number of features (24) was defined using the optimal average area under the receiver operating curve (AUC) (Figure 2). The data was then limited to the 24 selected features and random_state was used to split the data into training (80%) and testing (20%) using deterministic train-test sets. The test data was not used in training of the model. The logistic model was fitted, and performance was assessed using precision, recall, f1-score, and AUC. The model was then used to define odds ratios and 95% confidence intervals for each of the variables. The variables were considered to convey statistically significant increased risk if the odds ratio and entire 95% confidence interval was greater than 1.0.
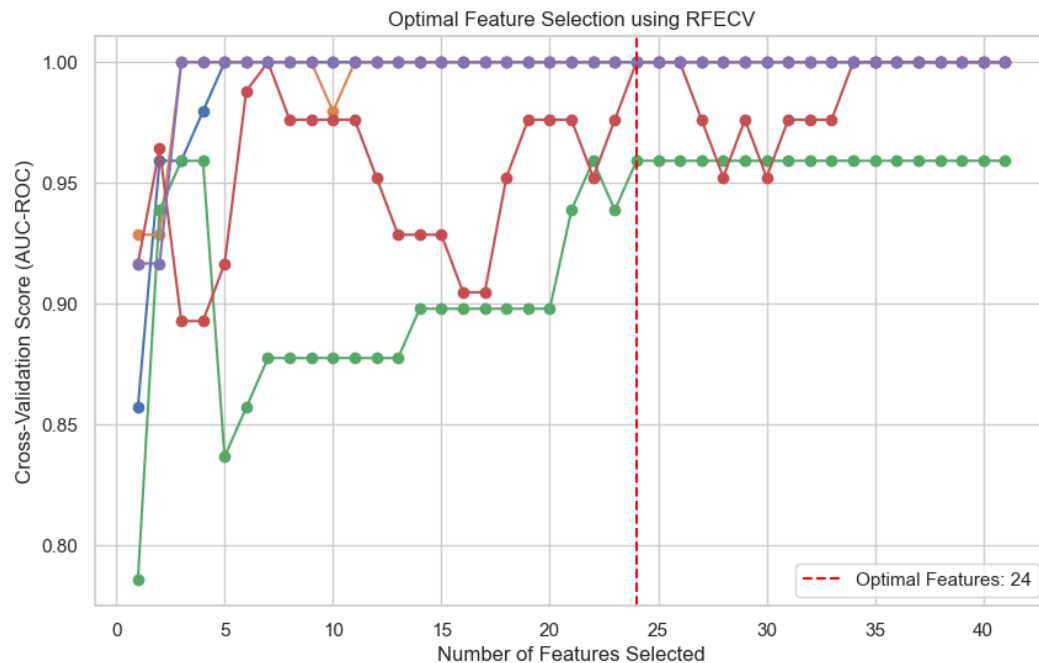
Figure 2. Five unique interations of recursive factor elimination were performed (each delinated by the different colored lines), with model identifying 24 factors as valubale predictors based upon resultant AUC. This approach minimizes risk of overfitting and ensures only factors offering substantial predictive value are included in the model.

2.8. Class Imbalance

Class imbalance was mitigated using SMOTE.

2.9. Fairness

Five iterations of recursive factor elimination with cross validation were applied to minimize inherent risk of overfitting.

2.10. Model Output

Output consisted of an odds ratio and 95% confidence interval. In rare outcomes, odds ratio approximates relative risk, making this an appropriate measure for the study. Statistical significance was based upon an odds ratio and 95% confidence interval > 1.0 or < 1.0.

3. Results

3.1. Cohort Demographics

Baseline cohort characteristics can be seen in Table 1. The study cohort was 65.8% male with an average age of 64.8 (+/- 11.6) years. 5 of the 38 included patients (13%) had DBS explantation. The most common indication was primary PD (78.9%), followed by dystonia (23.7%). Almost all cases were performed under general anesthesia (97.4%), with patients having a median ASA score of 3.0 (IQR: 3.0 – 3.0). Median length of stay was 1.0 (IQR 1.0 - 1.3) days, with most patients requiring ICU admission (94.7%). No statistically significant differences were noted for any of the demographic or perioperative variables. The most common medical comorbidities seen were hyperlipidemia (18.4%), sleep disorders (15.8%), dysautonomia (15.8%), obesity (15.8%), and personal history of malignancy (15.8%). Chronic pain was more frequently seen in patients with DBS explantation than those without (p = 0.0108), however no other statistically significant differences were noted. Amongst psychiatric and social comorbidities, tobacco use (31.2%), anxiety (21.1%), major depressive disorder (15.8%), and opioid use (5.3%) were present. However, only tobacco use was significantly higher in

patients with explant (p = 0.0026). No other statistically significant differences were seen between the cohorts.

Table 1. Cohort Demographics and Clinical Data.

	DBS	Explant	No Explant	p-value
Number of patients	38	5 (13.2%)	33 (86.8%)	0.1440
Sex				0.1440
Male n (%)	25 (65.8%)	5 (100.0%)	20 (60.6%)	0.4416
Female n (%)	13 (34.2%)	0 (0.0%)	13 (39.4%)	
Age (years +/- SD)	64.8 +/- 11.6	62.4 +/- 6.2	65.2 +/- 12.2	0.1316
Anesthesia Type				0.1316
General Anesthesia n (%)	37 (97.4%)	4 (80.0%)	33 (100.0%)	>0.9999
Monitored Airway n (%)	1 (2.6%)	1 (20.0%)	0 (0.0%)	0.8107
ASA Score (median, IQR)	3.0 (3.0 – 3.0)	3.0 (3.0 – 3.0)	3.0 (3.0 – 3.0)	0.2489
Length of Stay (median, IQR)	1.0 (1.0 – 1.3)	1.0 (1.0 – 1.5)	1.0 (1.0 – 1.5)	
ICU Admission n (%)	36 (94.73%)	4 (80.0%)	32 (97.0%)	>0.9999
Indication for DBS**				>0.9999
Primary PD n (%)	30 (78.94%)	4 (80.0%)	26 (78.8%)	0.1195
Secondary PD n (%)	1 (2.63%)	0 (0.0%)	1 (3.0%)	0.3121
Essential Tremor n (%)	5 (13.15%)	2 (40.0%)	3 (9.1%)	>0.9999
Dystonia n (%)	9 (23.68%)	0 (0.0%)	9 (27.3%)	
Spasticity n (%)	1(2.63%)	0 (0.0%)	1 (3.0%)	0.2440
Surgical Approach				0.2440
Percutaneous n (%)	29 (76.31%)	3 (60.0%)	26 (83.9%)	
Open n (%)	7 (18.42%)	2 (40.0%)	5 (16.1%)	0.1316
Medical Comorbidities				0.4456
Epilepsy (%)	1 (2.63%)	1 (20.0%)	0 (0.0%)	>0.9999
Neuropathy (%)	4 (10.52%)	1 (20.0%)	3 (9.1%)	**0.0108***
Acute Post-Operative Pain (%)	2 (5.26%)	0 (0.0%)	2 (6.1%)	0.1690
Chronic Pain (%)	5 (13.15%)	3 (60.0%)	2 (6.1%)	N/A
Dysautonomia (%)	6 (15.78%)	2 (40.0%)	4 (12.1%)	>0.9999
Chronic Fatigue (%)	0.0%	0.0%	0.0%	0.1316
Cognitive Impairment (%)	5 (13.15%)	0 (0.0%)	5 (15.2%)	>0.9999
Restless Leg Syndrome (%)	1(2.63%)	1 (20.0%)	0 (0.0%)	0.1195
Cerebrovascular Disease (%)	1 (2.63%)	0 (0.0%)	1 (3.0%)	0.5701
Sleep Apnea (%)	5 (13.15%)	2 (40.0%)	3 (9.1%)	0.4456
Sleep Disorder, Any (%)	6 (15.78%)	0 (0.0%)	6 (18.2%)	>0.9999

Chronic Obstructive Pulmonary Disease (%)	4 (10.52%)	1 (20.0%)	3 (9.1%)	0.2227
Hypertension (%)	17 (44.73%)	2 (40.0%)	15 (45.5%)	>0.9999
Hyperlipidemia (%)	7 (18.42%)	2 (40.0%)	5 (15.2%)	0.1195
Atrial Fibrillation (%)	2 (5.26%)	0 (0.0%)	2 (6.1%)	0.5272
Diabetes Mellitus (%)	5 (13.15%)	2 (40.0%)	3 (9.1%)	0.2489
Chronic Kidney Disease (%)	5 (13.15%)	1 (20.0%)	4 (12.1%)	>0.9999
Fibromyalgia (%)	2 (5.26%)	1 (20.0%)	1 (3.0%)	>0.9999
Irritable Bowel Syndrome (%)	1 (2.63%)	0 (0.0%)	1 (3.0%)	0.1690
Underweight/Cachexia (%)	4 (10.52%)	0 (0.0%)	4 (12.1%)	>0.9999
Obesity (%)	6 (15.78%)	2 (40.0%)	4 (12.1%)	>0.9999
Migraine (%)	1 (2.63%)	0 (0.0%)	1 (3.0%)	0.5701
Urinary Incontinence (%)	3 (7.89%)	0 (0.0%)	3 (9.1%)	N/A
Malignancy (%)	6 (15.78%)	0 (0.0%)	6 (18.2%)	
Bowel Incontinence (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0.2489
Substance Use				>0.9999
Opioid Use (%)	2 (5.26%)	1 (20.0%)	1 (3.0%)	0.0026*
Substance Use (%)	1 (2.63%)	0 (0.0%)	1 (3.0%)	
Tobacco Use (%)	13 (34.21%)	5 (100.0%)	8 (24.2%)	>0.9999
Psychiatric Comorbidities				>0.9999
Anxiety (%)	8 (21.05%)	1 (20.0%)	7 (21.2%)	N/A
MDD (%)	6 (15.78%)	1 (20.0%)	5 (15.2%)	N/A
ADHD (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	N/A
ETOH (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	N/A
OCD (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
PTSD (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	

* = p<0.05.

3.2. Multivariate Logistic Regression Model

As patient's variables do not exist in isolation, we hypothesized a model which could account for the interaction between variables may more accurately uncover potential perioperative risk factors. To test this hypothesis, we applied supervised machine learning using RFECV and a multivariate logistic regression model. The logistic regression model displayed robust performance with an average precision of 0.89, an average recall of 0.86, an average f1-score of 0.86, and AUC-ROC of 1.0.

Amongst the assessed variables, 12 conveyed increased risk of explant: tobacco use (OR: 3.64; CI: 3.60 – 3.68), primary PD (OR: 2.01; CI: 1.99 – 2.02), ASA score (OR: 1.91; CI: 1.90 – 1.92), chronic pain (OR: 1.82; CI: 1.80 – 1.85), diabetes (OR: 1.63; CI: 1.62 – 1.65) , restless leg syndrome (OR: 1.55; CI: 1.53 – 1.56), neuropathy (OR: 1.40; CI: 1.37 – 1.42), dysautonomia (OR: 1.31; CI: 1.30 – 1.32), epilepsy (OR: 1.21; CI: 1.20 – 1.21), a history of opioid use (OR: 1.14; CI: 1.14 – 1.15), and hyperlipidemia (OR: 1.12; CI: 1.11 – 1.14). Female sex (OR: 0.45; CI: 0.45 – 0.45), an indication of

dystonia (OR: 0.47; CI: 0.46 – 0.47), patient age (OR: 0.55; CI: 0.54 – 0.55), sleep disorders (OR: 0.56; CI: 0.55 – 0.56), history of malignancy (OR: 0.59; CI: 0.59 – 0.60), cognitive impairment (OR: 0.66; CI: 0.66 – 0.67), hypertension (OR: 0.69; CI: 0.68 – 0.69), illicit substance use (OR: 0.71; CI: 0.70 – 0.71), acute post-operative pain (OR: 0.72; CI: 0.71 – 0.72), ICU admission (OR: 0.87; CI: 0.87 – 0.88), comorbid cachexia/underweight (OR: 0.88; CI: 0.87 – 0.88), and atrial fibrillation (OR: 1.00; CI: 1.00 – 1.00) were included in the model, however they all conveyed decreased or unchanged risk. (Figure 3)
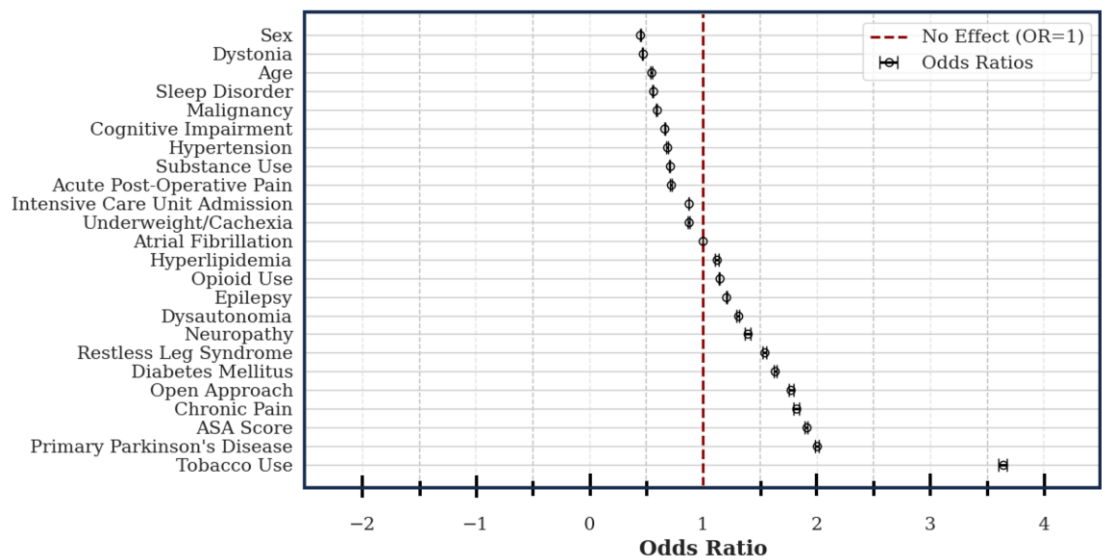
Figure 3. Forest plot illustrating the odds ratio and 95% confidence interval for each of the included variables. Hyperlipidemia, opioid use, epilepsy, dysautonomia, neuropathy, restless leg syndrome, diabetes, an open surgical approach, chronic pain, increased ASA score, an indication of primary PD, and tobacco use all conveyed increased risk of explant.

4. Discussion

Since it's advent as a viable treatment for PD, deep brain stimulation has continued to grow in indications and number of procedures [3-6]. While significant advances have been made in increasing the precision of DBS technology itself, rates of hardware-related complications [19] and explantation [12] remain unacceptably high, suggesting a need to more precisely identify patients at risk of adverse outcomes. Here, we sought to address this challenge by taking a unique approach using supervised machine learning in the form of a multivariate logistic regression combined with recursive factor elimination with cross validation. In doing so, we identified a history of tobacco use, an indication of primary PD, increased ASA score, comorbid chronic pain, an open surgical approach, comorbid diabetes, comorbid restless leg syndrome, comorbid neuropathy, dysautonomia, comorbid epilepsy, a history of opioid use, and comorbid hyperlipidemia as potential risk factors for early DBS explantation.

A national study conducted in Korea identified older age, male sex, comorbid dementia, and comorbid fractures as factors predictive of morbidity following DBS placement [14]. Similarly, Ward et al. identified ineffective stimulation (28%), lead dysfunction (30%), and infections (28%) as etiologies of DBS failure [10]. A retrospective, single institution study using basic statistical analysis identified use of topical vancomycin (as opposed to intrawound) as a risk factor for DBS infection [19]. A second single-center retrospective study identified use of general anesthesia, hypertension, heart disease, and depression as risk factors for increased postoperative length of stay [20]. Furthermore, while associations with explant were limited, they did note an association with preoperative body mass index and risk of explant [20]. A retrospective study by Deeb et al. assessing

the association between psychiatric comorbidities and DBS explantation in a Tourette syndrome database found no statistically significant associations [12].

Limitations of our study include the limited number of patients, the retrospective design, the use of ICD codes which limits the ability to capture granular clinical reasoning [21], and the inclusion of a singular center. Furthermore, the rate of explant within the MOVER database was higher (13%) than typically reported in the literature [12]. We hypothesize this may be a result of the small sample size. Strengths of the study include the robust number of assessed variables (40), the use of a supervised machine learning model allowing for the assessment of between variables interactions, and use of recursive factor elimination with cross validation to identify the optimal variables.

5. Conclusions

Identification of risk factors for DBS explant has proven challenging using basic statistical models, not well adapted for rare outcomes. Here, we offer a novel methodology using an optimized multivariable logistic regression which may offer a solution to this challenge. Further study, with larger number of patients are needed to validate the results of this small retrospective, pilot study.

Abbreviations

The following abbreviations are used in this manuscript:

PD	Parkinson's Disease
DBS	Deep brain stimulation

AUC   Area under the curve

RFECV   Recursive factor elimination with cross validation

MOVER   Medical Informatics Operating room Vitals and Events Repository

OR   Odds ratio

CI   95% confidence interval

ASA   American Society of Anesthesiologists

BMI   Body mass index

ICD   International Classification of Diseases

SMOTE   Synthetic Minority Oversampling Technique

SD   Standard deviation

IQR   Interquartile range

ICU   Intensive care unit

LOS   Length of stay

References

1. Benabid, A.L., et al., *Combined (thalamotomy and stimulation) stereotactic surgery of the VIM thalamic nucleus for bilateral Parkinson disease.* Appl Neurophysiol, 1987. **50**(1-6): p. 344-6.
2. Servello, D., et al., *Complications of deep brain stimulation in Parkinson's disease: a single-center experience of 517 consecutive cases.* Acta Neurochir (Wien), 2023. **165**(11): p. 3385-3396.

10

3. Hacker, M.L., et al., *Deep brain stimulation in early-stage Parkinson disease: Five-year outcomes.* Neurology, 2020. **95**(4): p. e393-e401.
4. Figee, M., et al., *Deep Brain Stimulation for Depression.* Neurotherapeutics, 2022. **19**(4): p. 1229-1245.
5. Knotkova, H., et al., *Neuromodulation for chronic pain.* Lancet, 2021. **397**(10289): p. 2111-2124.
6. Wang, S., et al., *Efficacy and safety of combined deep brain stimulation with capsulotomy for comorbid motor and psychiatric symptoms in Tourette's syndrome: Experience and evidence.* Asian J Psychiatr, 2024. **94**: p. 103960.
7. Krauss, J.K., et al., *Technology of deep brain stimulation: current status and future directions.* Nat Rev Neurol, 2021. **17**(2): p. 75-87.
8. Deng, H., J.K. Yue, and D.D. Wang, *Trends in safety and cost of deep brain stimulation for treatment of movement disorders in the United States: 2002-2014.* Br J Neurosurg, 2021. **35**(1): p. 57-64.
9. Reich, M.M., et al., *A brain network for deep brain stimulation induced cognitive decline in Parkinson's disease.* Brain, 2022. **145**(4): p. 1410-1421.
10. Ward, M., et al., *Complications associated with deep brain stimulation for Parkinson's disease: a MAUDE study.* Br J Neurosurg, 2021. **35**(5): p. 625-628.
11. Olson, M.C., et al., *Deep brain stimulation in PD: risk of complications, morbidity, and hospitalizations: a systematic review.* Front Aging Neurosci, 2023. **15**: p. 1258190.
12. Deeb, W., et al., *Deep brain stimulation lead removal in Tourette syndrome.* Parkinsonism Relat Disord, 2020. **77**: p. 89-93.
13. Jung, I.H., et al., *Complications After Deep Brain Stimulation: A 21-Year Experience in 426 Patients.* Front Aging Neurosci, 2022. **14**: p. 819730.
14. Kim, A., et al., *Mortality of Deep Brain Stimulation and Risk Factors in Patients With Parkinson's Disease: A National Cohort Study in Korea.* J Korean Med Sci, 2023. **38**(3): p. e10.
15. Samad, M., et al., *Medical Informatics Operating Room Vitals and Events Repository (MOVER): a public-access operating room database.* JAMIA Open, 2023. **6**(4): p. ooad084.
16. McKinney, W., *Data Structures for Statistical Computing in Python.* Proceedings of the 9th Python in Science Conference, 2010: p. 51-56.
17. Fabian Pedregosa, G.V., Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay, *Scikit-learn: Machine Learning in Python*, in *Journal of Machine Learning Research*. 2011. p. 2825-2830.
18. Waskom, M.L., *seaborn: statistical data visualization.* The Journal of Open Source Software, 2021. **6**: p. 3021.
19. Abode-Iyamah, K.O., et al., *Deep brain stimulation hardware-related infections: 10-year experience at a single institution.* J Neurosurg, 2019. **130**(2): p. 629-638.
20. Tiefenbach, J., et al., *The Rate and Risk Factors of Deep Brain Stimulation-Associated Complications: A Single-Center Experience.* Oper Neurosurg (Hagerstown), 2024.
21. Hussain, N. and T. Weaver, *Response to the Letter to the Editor Regarding: "Identifying Predictors for Early Percutaneous Spinal Cord Stimulator Explant at One and Two Years: A Retrospective Database Analysis".* Neuromodulation, 2023. **26**(3): p. 710.