Title

Ordinal regression increases statistical power to predict epilepsy surgical outcomes

Journal

Epilepsia Open, 7(2)

ISSN

2470-9239

Authors

Dickey, Adam S
Krafty, Robert T
Pedersen, Nigel P

Publication Date

2022-06-01

DOI

10.1002/epi4.12585

Peer reviewed

Received: 13 October 2021 | Revised: 19 January 2022 | Accepted: 28 January 2022

DOI: 10.1002/epi4.12585

Epilepsia Open® Open Access

SHORT RESEARCH ARTICLE

Ordinal regression increases statistical power to predict epilepsy surgical outcomes

Adam S. Dickey[1] | **Robert T. Krafty**[2] | **Nigel P. Pedersen**[1,3,4]

[1]Department of Neurology, Emory University School of Medicine, Atlanta, Georgia, USA

[2]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia, USA

[3]Department of Biomedical Engineering, Emory University and the Georgia Institute of Technology, Atlanta, Georgia, USA

[4]College of Medicine and Public Health, Flinders University, Adelaide, South Australia, Australia

Correspondence
Adam Dickey, Woodruff Memorial Research Building, Room 6209, 101 Woodruff Circle, 30322 Atlanta, GA, USA.
Email: adam.s.dickey@emory.edu

Funding information
National Institute of General Medical Sciences, Grant/Award Number: R01 GM113243; National Institute of Neurological Disorders and Stroke, Grant/Award Number: K08 NS105929, R01 NS088748 and R21 NS122011; National Center for Advancing Translational Sciences, Grant/Award Number: KL2 TR002381 and UL1 TR002378; Woodruff Foundation; Citizens United for Research in Epilepsy

Abstract

Studies of epilepsy surgery outcomes are often small and thus underpowered to reach statistically valid conclusions. We hypothesized that ordinal logistic regression would have greater statistical power than binary logistic regression when analyzing epilepsy surgery outcomes. We reviewed 10 manuscripts included in a recent meta-analysis which found that mesial temporal sclerosis (MTS) predicted better surgical outcomes after a stereotactic laser amygdalohippocampectomy (SLAH). We extracted data from 239 patients from eight studies that reported four discrete Engel surgical outcomes after SLAH, stratified by the presence or absence of MTS. The rate of freedom from disabling seizures (Engel I) was 64.3% (110/171) for patients with MTS compared to 44.1% (30/68) without MTS. The statistical power to detect MTS as a predictor for better surgical outcome after a SLAH was 29% using ordinal regression, which was significantly more than the 13% power using binary logistic regression (paired *t*-test, $P < .001$). Only 120 patients are needed for this example to achieve 80% power to detect MTS as a predictor using ordinal regression, compared to 210 patients that are needed to achieve 80% power using binary logistic regression. Ordinal regression should be considered when analyzing ordinal outcomes (such as Engel surgical outcomes), especially for datasets with small sample sizes.

KEYWORDS

Engel outcome, epilepsy surgery, logistic regression, seizure freedom, statistical power

1 | INTRODUCTION

Studies of epilepsy surgery often aim to identify predictors of a good outcome to help guide treatment. However, most studies treat seizure outcome as a binary variable of seizure freedom (Engel I) or not, which ignores the difference between rare disabling seizures (Engel II), worthwhile improvement in seizure frequency (Engel III), and no worthwhile improvement (Engel IV).[1] We argue that the distinction between rare seizures and frequent disabling seizures is clinically meaningful to our patients, and therefore statistical analysis should harness the full range of possible outcomes.

We propose using *ordinal logistic regression* rather than traditional binary logistic regression, as the Engel surgical outcome is an ordinal scale. Relying on seizure freedom as a primary outcome amounts to comparing Engel I vs Engel II-IV. There are two additional dichotomous categorizations: Engel I-II vs Engel III-VI and Engel I-III vs Engel IV. Rather than fit three individual binary regressions, ordinal regression fits three separate intercepts and a common, pooled odds ratio.[2] This proportional odds assumption implies the odds ratios are similar between the three comparisons, though non-proportional odds can also be assumed.

It has been shown that binarizing an ordinal categorical outcome can lead to a reduction in statistical power in some settings.[3] We hypothesize that ordinal logistic regression using full Engel outcomes, by discarding less outcome information, will have significantly more statistical power than binary logistic regression. We test this hypothesis using the problem of predicting surgical outcomes after stereotactic laser amygdalohippocampectomy (SLAH). A recent meta-analysis of 10 studies found that the presence of mesial temporal sclerosis (MTS) predicted a better chance (63% vs 42%) of seizure freedom after SLAH when MTS was present than when absent.[4] However, most individual studies did not detect this effect, as their low sample size limited statistical power. We specifically hypothesize that using ordinal logistic regression we will meaningfully increase the statistical power to detect the known relation between MTS and seizure outcome in each study.

2 | METHODS

2.1 | Data sources

We make use of data from a recently published meta-analysis that included 10 studies and analyzed the outcome of seizures freedom as a function of the exposure of the presence or absence of MTS (full details can be found in the original manuscript).[4] The included studies were retrospective cohort studies comparing the association of the exposure of mesial temporal sclerosis with the outcome of Engel class after surgery. We reviewed the original manuscripts and found eight studies that reported raw counts of patients for each of the four discrete Engel outcomes, stratified by the presence or absence of MTS.[5–12] We excluded one paper because it used an alternative seizure outcome scale and another because it grouped together with the classification of Engel II-IV. This study is thus a meta-analysis of publicly available, published studies that were approved by their local Institutional Review Board or its equivalent.

2.2 | Logistic regression

For each study, we converted the count data into an independent vector with the presence (1) or absence (0) of MTS, and dependent vectors of seizure freedom (1) or not (2), or the full Engel outcomes (1–4). We performed binary logistic regression and ordinal logistic regression on the dependent vectors of seizure freedom or Engel outcome, respectively. We report the *p*-value for the log-odds (or slope) coefficient. Both the binary and ordinal regression were performed using the Matlab function mnrfit in MATLAB (version R2016b, Mathworks) using the ordinal model. The proportional odds assumption is made by setting interactions to "off", while the non-proportional odds assumption is made by setting interactions to "on". When this function is supplied as a vector with two outcomes, it provides the same results as binary logistic regression.

To ensure our results were not sensitive to the assumption of proportional odds, we compared this assumption to non-proportional odds. The latter model has 6 parameters, with a separate slope and intercept for each comparison (Engel I vs II-VI, I-II vs III-VI, and I-III vs IV). In practice, the estimated probabilities from this model are equivalent to the observed percentages. We then use the likelihood-ratio test to evaluate whether any of the slopes are non-zero. The null hypothesis is that the percentage of patients with each Engel outcome is equal to the result obtained when both groups are combined. The alternative is the observed percentage in each Engel outcome for each predictor group separately (MTS or non-MTS). The likelihood ratio test statistic (see Equation 1) follows a chi-square distribution with 3 degrees of freedom (the difference in the number of parameters of the two models).

$$\mathrm{LR} = -2\left[\log(\mathrm{likelihood}(\mathrm{Null})) - \log(\mathrm{likelihood}(\mathrm{Alternative}))\right] \tag{1}$$

We also used the likelihood ratio test to compare the fit to the raw data of the proportional vs non-proportional odds models. Here, the null hypothesis (proportional) is defined by 4 parameters (3 slopes and 1 common slope), and the alternative hypothesis (non-proportional) is again defined by 6 parameters. The likelihood ratio test statistic follows a chi-square distribution with 2 degrees of freedom.

2.3 | Power analysis

We pooled the count data across the eight studies to compute an empiric distribution of the probability of each Engel outcome in the presence or absence of MTS. We performed a power analysis by simulating binomial count data, using the empiric probability distribution and the count data of patients with or without MTS from each study. We performed a bootstrap simulation with 10 000 repetitions and report the power as the percentage of repetitions where the logistic regression (binary or ordinal, with proportional or non-proportional odds), and reported a *P*-value as significant for a log-odds coefficient <.05.

3 | RESULTS

3.1 | Seizure outcomes

The raw and pooled count data are summarized below (Table 1). Across the eight studies, 72% (171/239) of patients had MTS and 28% (68/239) did not. The rate of freedom from disabling seizures (Engel I) was 64% (110/171) for patients with MTS compared to 44% (30/68) without MTS. These results are similar to those from the meta-analysis (63% vs 42%), which included 10 studies.[4]

A comparison of the observed distribution of Engel outcomes to that predicted by ordinal regression is shown, stratified by the presence or absence of MTS (Figure 1). The pooled odds ratio assumption means the predicted difference in Engel I for MTS or not (65% vs 39%) is slightly higher than observed (64% vs 44%), but the predicted difference for Engel IV for MTS or not (5% vs 13%) is smaller than observed (3% vs 18%). A comparison of the pooled odds ratio and the odds ratios of the three possible binary logistic regressions is also shown (Figure 2). The pooled odds ratio of 2.97 is contained within the 95% confidence interval for each binary logistic regression. However, a likelihood ratio test did show that the prediction of ordinal regression with non-proportional odds (Figure 1A) was a significantly better fit ($P = .04$) than the prediction of ordinal regression with proportional odds (Figure 1B).

3.2 | Statistical significance

When the raw count data were analyzed for each study using binary logistic regression, only one study[10] showed a significant difference ($P < .05$) for seizure freedom with vs without MTS. When ordinal logistic regression with proportional odds was used, three studies showed a significant difference in Engel outcome for patients with vs without MTS. Of note, both studies with a significant result for ordinal, but not binary regression, did show a significant difference related to MTS using a Kaplan-Meier analysis. However, those analyses included extended follow-up to 53[8] and 43 months.[11] When ordinal regression with non-proportional odds was used, two studies showed a significant difference in Engel outcome for patients with vs without MTS (Table 1).

TABLE 1 Patient counts by study for Engel outcome for mesial temporal sclerosis or not

	Sample size			Engel MTS				Engel non-MTS				*P*-value			Power (%)		
First author	Tot	MTS	Non	I	II	III	VI	I	II	III	VI	Bin	NP	Prop	Bin	NP	Prop
Gross	58	43	15	26	10	7	0	5	3	4	3	.08	***.02***	***.01***	26	48	48
Donos	43	34	9	23	6	5	0	6	1	2	0	.96	.93	.85	17	35	38
Youngerman	30	18	12	10	1	5	2	7	0	3	2	.88	.75	.94	17	33	32
Le	29	22	7	16	4	2	0	2	2	3	0	***.05***	.16	***.03***	12	27	29
Jermakowicz	23	15	8	11	3	1	0	5	0	1	2	.59	.07	.29	12	27	27
Grewal	23	18	5	13	3	2	0	2	2	1	0	.20	.62	.23	5	22	23
Tao	18	10	8	7	3	0	0	2	0	3	3	.07	***.00***	***.01***	9	21	19
Greenway	15	11	4	4	1	3	3	1	0	1	2	.68	.76	.42	1	17	15
Total (#)	239	171	68	110	31	25	5	30	8	18	12			Avg (%)	12	29	29

Note: *P*-values <.05 are italicized and bolded. The statistical power of ordinal regression with or without proportional odds was 29%, which was significantly higher the 12% power with binary logistic regression (paired *t*-test, $P < .001$).

Abbreviations: Avg, average; Bin, binary logistic regression; Engel, engel surgical outcome; MTS, mesial temporal sclerosis; Non, non-MTS; NP, Non-proportional odds, likelihood ratio; Prop, proportional odds, ordinal regression; Total, total.

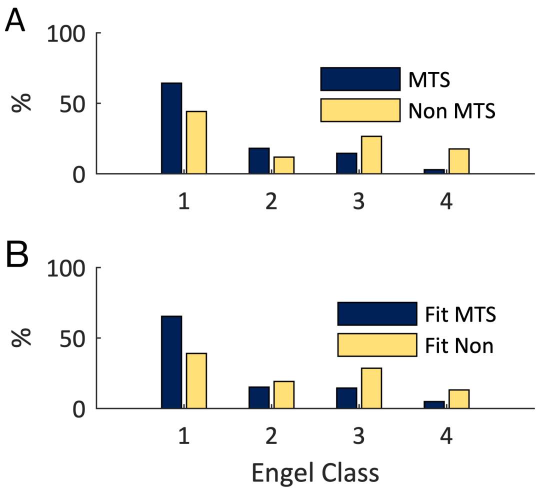

FIGURE 1 Results of ordinal logistic regression. (A) Observed percentages of Engel outcomes I-IV stratified by the presence or absence of mesial temporal sclerosis (MTS), pooled from 8 studies. (B) Predicted percentages of Engel I-IV stratified by MTS or not, using ordinal logistic regression. The pooled odds ratio assumption means the predicted difference in Engel I for MTS or not (65% vs 39%) is slightly higher than observed (64% vs 44%), but the predicted difference for Engel IV for MTS or not (5% vs 13%) is smaller than observed (3% vs 18%)

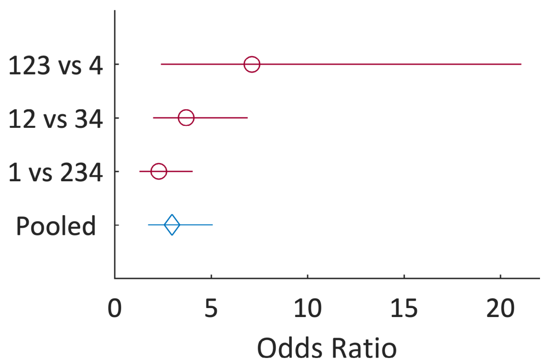

FIGURE 2 Odds ratios for binary vs pooled ordinal logistic regression. Predicted odds ratios (with 95% confidence interval) are displayed for the three possible binary logistic regressions, compared to the pooled odds ratio from ordinal logistic regression. The pooled odds ratio of 2.97 is contained within the 95% confidence interval for each binary logistic regression. However, a log-likelihood ratio test did show modest support ($P = .04$) for ordinal regression with non-proportional odds providing a better fit to the observed data than proportional odds

3.3 | Power analysis

We used the pooled count data to estimate an empirical distribution of the probability of each Engel outcome given the presence or absence of MTS (Table 1). We assumed this as ground truth and generated simulated data for each study, given the number of subjects with or without MTS. We then estimated the power to detect a difference in Engel's outcome using binary or ordinal logistic regression with proportional odds. The statistical power to detect the true effect was 29% for ordinal regression, which was significantly more than 13% for binary logistic regression (paired *t*-test, $P < .001$). An equivalent result was obtained when using ordinal regression with non-proportional odds (Table 1).

We estimated the sample size needed to achieve 80% power to detect a difference in Engel outcome, assuming a 2:1 allocation of groups (two-thirds with MTS and one-third without). Using binary logistic regression (or a chi-square test), one would need around 210 patients (140 with MTS and 70 without) to achieve 80% power to detect a difference of seizure freedom of 64% vs 44% (bootstrap, 10 000 repetitions).

However, using a pooled empiric probability distribution for Engel outcomes (Table 1), one achieves 80% power to detect the same difference in Engel outcome with only 120 patients (bootstrap, 10 000 repetitions, 80 with MTS, and 40 without). In contrast, binary logistic regression and the chi-squared test only have ~56% power with 120 patients. For this example, switching from binary to ordinal logistic regression almost doubles the effective sample size.

4 | DISCUSSION

To the best of our knowledge, this is the first published report applying ordinal logistic regression to Engel surgical outcomes. Ordinal regression increases statistical power and decreases the sample size need to achieve the desired power. We showed this true both using the proportional and non-proportional odds assumption. We feel the proportional odds are model easier to interpret, as it gives a single pooled odds ratio (see Figure 2). This matches our clinical intuition that a predictor of seizure freedom should also predict seizure improvement. However, a log-likelihood ratio test did show modest support ($P = .04$) for ordinal regression with non-proportional odds providing a better fit to the observed data than proportional odds.

Though we use data from multiple studies, one limitation is that we focus on a specific epilepsy surgery (SLAH), so the generalizability of this finding is not yet clear. Because SLAH is a relatively new technique, case series necessarily have small sample sizes and thus low statistical power. Underpowered studies will miss true effects when present (such as the individual studies which did not find an association between MTS and seizure outcome). Underpowered studies also have inflated false-positive rates[13] and overestimate the magnitude of statistically significant effects when found.[14] Low-powered studies distort our understanding of prognostics factors for epilepsy surgery.

The obvious solution is to acquire larger sample sizes. However, this is difficult in practice. The less obvious solution is to use more sophisticated statistical analysis to avoid discarding relevant information.[2] Here, binary classification of seizure freedom is suboptimal when full Engel outcomes are available. It may also be possible to increase statistical power using continuous predictors. For example, asymmetry scores for MTS could be computed from neuroimaging data, similar to those used for WADA tests.

Our motivation for this paper was not re-demonstrating that MTS predicts seizure freedom after SLAH. Rather, the goal is to identify statistical tools which can detect relevant prognostic factors in small datasets. For example, semiology has been under-analyzed as a predictive factor for good outcomes after SLAH. Signs such as ipsilateral manual automatisms and contralateral dystonia are strongly lateralizing.[15] Well-selected patients with concordant semiology (and other presurgical data) may have better surgical outcomes than patients with MTS but discordant semiology or scalp EEG. Ordinal regression should be an ideal tool to clarify the relative predictive value of neuroimaging (MTS) vs other potential predictors (such as semiology).

5 | CONCLUSION

Ordinal regression increases the statistical power to detect a predictor of good outcomes after epilepsy surgery when compared to binary logistic regression. Ordinal regression also decreases the sample size needed to achieve 80% statistical power, relative to binary logistic regression. Ordinal regression should therefore be considered when analyzing ordinal outcomes (such as Engel surgical outcome), especially for datasets with small sample sizes.

ACKNOWLEDGMENTS

NPP is supported by the Woodruff Foundation, CURE Epilepsy, and NIH grants K08 NS105929, R01 NS088748, and R21 NS122011. ASD is supported by the National Center for Advancing Translational Sciences of the NIH under award numbers UL1 TR002378 and KL2 TR002381. RTK is supported by R01 GM113243. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We also thank Scott Millis of Wayne State University for help finding relevant references. A version of this manuscript was posted as a preprint to https://www.medrxiv.org/content/10.1101/2021.10.01.21264435. MATLAB code which can be used to reproduce the analyses and figures described here is posted at https://github.com/AdamSDickey/Ordinal_Regression.

CONFLICTS OF INTEREST

NPP has served as a paid consultant for DIXI Medical USA, who manufactures products used in the workup for epilepsy surgery. The terms of this arrangement have been reviewed and approved by Emory University in accordance with its conflict-of-interest policies. ASD and RTK have no conflicts of interest to disclose. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

AUTHOR CONTRIBUTION

ASD and NPP contributed to the conception and design of the study. ASD, RTK, and NPP contributed to the drafting of the text. ASD performed the statistical analysis and prepared the figures and tables.

ORCID

Adam S. Dickey https://orcid.org/0000-0003-4710-1290

REFERENCES

1. Engel J. Update on surgical treatment of the epilepsies: summary of the second international palm desert conference on the surgical treatment of the epilepsies (1992). Neurology. 1993;43(8):1612.
2. Steyerberg EW. Clinical prediction models. Cham: Springer International Publishing; 2019.
3. Whitehead J. Sample size calculations for ordered categorical data. Stat Med. 1993;12:2257–71.
4. Kohlhase K, Zollner JP, Tandon N, Strzelczyk A, Rosenow F. Comparison of minimally invasive and traditional surgical approaches for refractory mesial temporal lobe epilepsy: a systematic review and meta-analysis of outcomes. Epilepsia. 2021;62:831–45.
5. Donos C, Breier J, Friedman E, Rollo P, Johnson J, Moss L, et al. Laser ablation for mesial temporal lobe epilepsy: surgical and cognitive outcomes with and without mesial temporal sclerosis. Epilepsia. 2018;59:1421–32.
6. Greenway MRF, Lucas JA, Feyissa AM, Grewal S, Wharen RE, Tatum WO. Neuropsychological outcomes following stereotactic laser amygdalohippocampectomy. Epilepsy Behav. 2017;75:50–5.
7. Grewal SS, Zimmerman RS, Worrell G, Brinkmann BH, Tatum WO, Crepeau AZ, et al. Laser ablation for mesial temporal epilepsy: a multi-site, single institutional series. J Neurosurg. 2018;1–8.
8. Gross RE, Stern MA, Willie JT, Fasano RE, Saindane AM, Soares BP, et al. Stereotactic laser amygdalohippocampotomy for mesial temporal lobe epilepsy. Ann Neurol. 2018;83:575–87.
9. Jermakowicz WJ, Kanner AM, Sur S, Bermudez C, D'Haese PF, Kolcun JPG, et al. Laser thermal ablation for mesiotemporal

epilepsy: analysis of ablation volumes and trajectories. Epilepsia. 2017;58:801–10.

10. Le S, Ho AL, Fisher RS, Miller KJ, Henderson JM, Grant GA, et al. Laser interstitial thermal therapy (LITT): seizure outcomes for refractory mesial temporal lobe epilepsy. Epilepsy Behav. 2018;89:37–41.
11. Tao JX, Wu S, Lacy M, Rose S, Issa NP, Yang CW, et al. Stereotactic EEG-guided laser interstitial thermal therapy for mesial temporal lobe epilepsy. J Neurol Neurosurg Psychiatry. 2018;89:542–8.
12. Youngerman BE, Oh JY, Anbarasan D, Billakota S, Casadei CH, Corrigan EK, et al. Laser ablation is effective for temporal lobe epilepsy with and without mesial temporal sclerosis if hippocampal seizure onsets are localized by stereoelectroencephalography. Epilepsia. 2018;59(3):595–606.
13. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14:365–76.
14. Dickey AS, Pedersen NP. Low statistical power in a study predicting seizure outcome. Epilepsia. 2021;62(10):2565–6.
15. Kotagal P, Lüders H, Morris H, Dinner D, Wyllie E, Godoy J, et al. Dystonic posturing in complex partial seizures of temporal lobe onset: a new lateralizing sign. Neurology. 1989;39:196.

How to cite this article: Dickey AS, Krafty RT, Pedersen NP. Ordinal regression increases statistical power to predict epilepsy surgical outcomes. Epilepsia Open. 2022;7:344–349. https://doi.org/10.1002/epi4.12585