Predicting Future Epilepsy Surgery using Multimodal Electronic Health Record Data

Benjamin D. Wissel, BS[1], Hansel M. Greiner, MD[2,3], Tracy A. Glauser, MD[2,3], Francesco T. Mangano, DO[2,4], Rhonda D. Szczesniak, PhD[2,5], Daniel Santel, PhD[1], Judith W. Dexheimer, PhD[1,2,6]

[1]Department of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; [2]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA; [3]Division of Neurology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; [4]Division of Neurosurgery, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; [5]Division of Biostatistics & Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; [6]Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

Cincinnati Children's — changing the outcome together

University of Cincinnati

Background and Introduction

- Epilepsy affects 3 million adults and 450,000 children in the United States.[1]
- Medications are used to control seizures, but they are ineffective in one-third of patients.[2]
- In patients with drug-resistant epilepsy, neurosurgical resection of the epileptic focus in the brain stops seizures in 67% of patients.[3]

Epilepsy surgery is underutilized and often delayed

- Only 1% of patients receive surgery within the first two years of becoming eligible.[4]

Machine Learning (ML) can identify surgical candidates two years earlier in the disease course

- ML can identify surgical candidates as accurately as epilepsy specialists[5]

Our goal was to develop a generalizable ML modeling process to identify candidates for epilepsy surgery from multi-modal electronic health record (EHR) data.

Methods

Retrospective cohort study in two different health care systems:

Pediatric center (2009 to 2019):
- Two hospitals
- 14 outpatient clinic sites
- 6,000 patients with epilepsy/yr

Adult center (2012 to 2019):
- Two hospitals
- 27 outpatient clinic sites
- 4,000 patients with epilepsy/yr

Modeling process developed using the pediatric dataset, then validated in adults.

Inclusion/exclusion criteria:
- At least two neurology visits for an ICD-9 or -10 diagnosis of epilepsy

Definitions of cases and controls:

Surgical patients (cases):
- Surgeries identified using Current Procedural Terminology (CPT) codes
- Confirmed by chart review
- Used only EHR data from before surgical evaluation; discard data from after referral (minimized label leak)

Non-surgical patients (controls):
- No neurosurgical procedures for epilepsy
- Used all historical EHR data

EHR data:

Neurology notes:
- Normalized abbreviations and drug names
- Tokenized into n-gram features (1:3)

EEG & MRI reports:
- Free-text radiology reports
- Tokenized into n-gram features (1:3)

Structured data:
- Neurology visit patterns
- ED visits and hosps.
- Medication orders
- Labs and procedures

Feature selection:
- Features ranked using a correlation-based filter
- Number included was selected from cross-validation

Statistical analysis:
- Performance estimated from 10-fold cross-validation

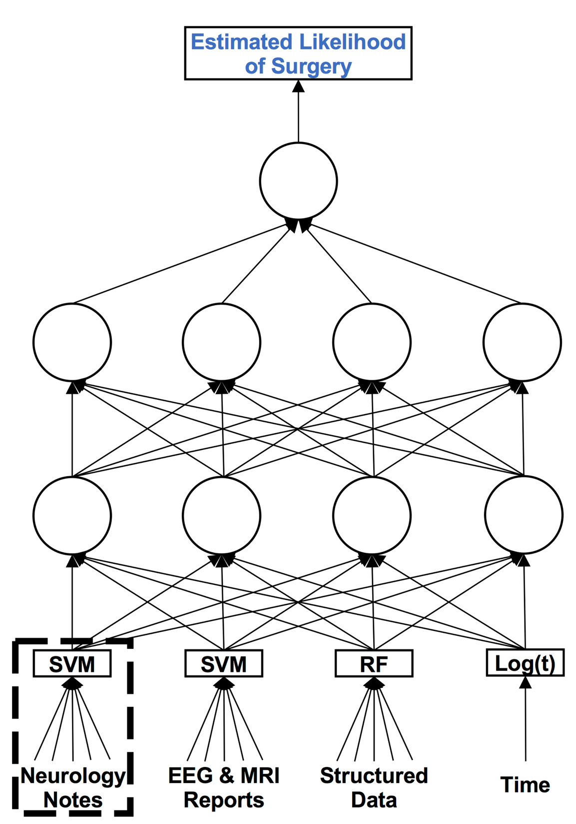

Figure 1 (right): Late fusion multimodal model architecture. A neural network fused unimodal model outputs from two support vector machines (SVM), a random forest (RF), and a log-transformed time component (duration of follow up).

Results

Table 1: Cohort demographics and clinical information.

Variable	Pediatric Health System Non-Surgical (n=5,743)	Pediatric Health System Surgery (n=137)	Adult Health System Non-Surgical (n=7,548)	Adult Health System Surgery (n=56)
Age, years	13.3 ± 7.50	9.74 ± 5.86	47.6 ± 16.8	41.6 ± 12.5
Male Gender	2,945 (51.3%)	83 (60.6%)	3,340 (44.3%)	28 (50.0%)
White Race	4,614 (80.3%)	108 (78.8%)	5,929 (78.6%)	51 (91.1%)
Insurance*				
Private	3,064 (53.4%)	63 (46.0%)	3,045 (40.3%)	29 (51.8%)
Public	3,294 (57.4%)	95 (69.3%)	4,262 (56.5%)	25 (44.6%)
Other	55 (0.96%)	3 (2.19%)	241 (3.19%)	2 (3.57%)
Distance from Care				
0-25 miles	2,727 (47.5%)	61 (44.5%)	5,387 (71.4%)	39 (69.6%)
25-50 miles	1,006 (17.5%)	24 (17.5%)	1,375 (18.2%)	7 (12.5%)
50-100 miles	1,004 (17.5%)	30 (21.9%)	586 (7.76%)	8 (14.3%)
>100 miles	1,006 (17.5%)	22 (16.1%)	200 (2.65%)	2 (3.57%)
Number of Neurology Visits	6.32 ± 4.73	7.93 ± 6.04	6.37 ± 5.04	4.21 ± 4.52
Duration of Follow-up, years	2.99 ± 2.58	2.11 ± 2.01	3.04 ± 2.16	1.33 ± 1.69
Anti-epileptic Drugs	1.96 ± 1.52	4.12 ± 2.1	2.09 ± 1.45	1.93 ± 1.56
Procedures and Labs	14.0 ± 11.1	21.0 ± 15.2	23.3 ± 27.7	12.9 ± 16.3
EEG Present	4,046 (70.5%)	98 (71.5%)	3,336 (44.2%)	22 (39.3%)
MRI Present	3,026 (52.7%)	102 (74.5%)	3,093 (41.0%)	18 (32.1%)

Fusing information from multiple sources in the EHR increased performance.

The modeling process generalized from pediatrics to adults.
AUROC = 0.959 in pediatrics vs. 0.930 in adults, respectively; p = 0.16

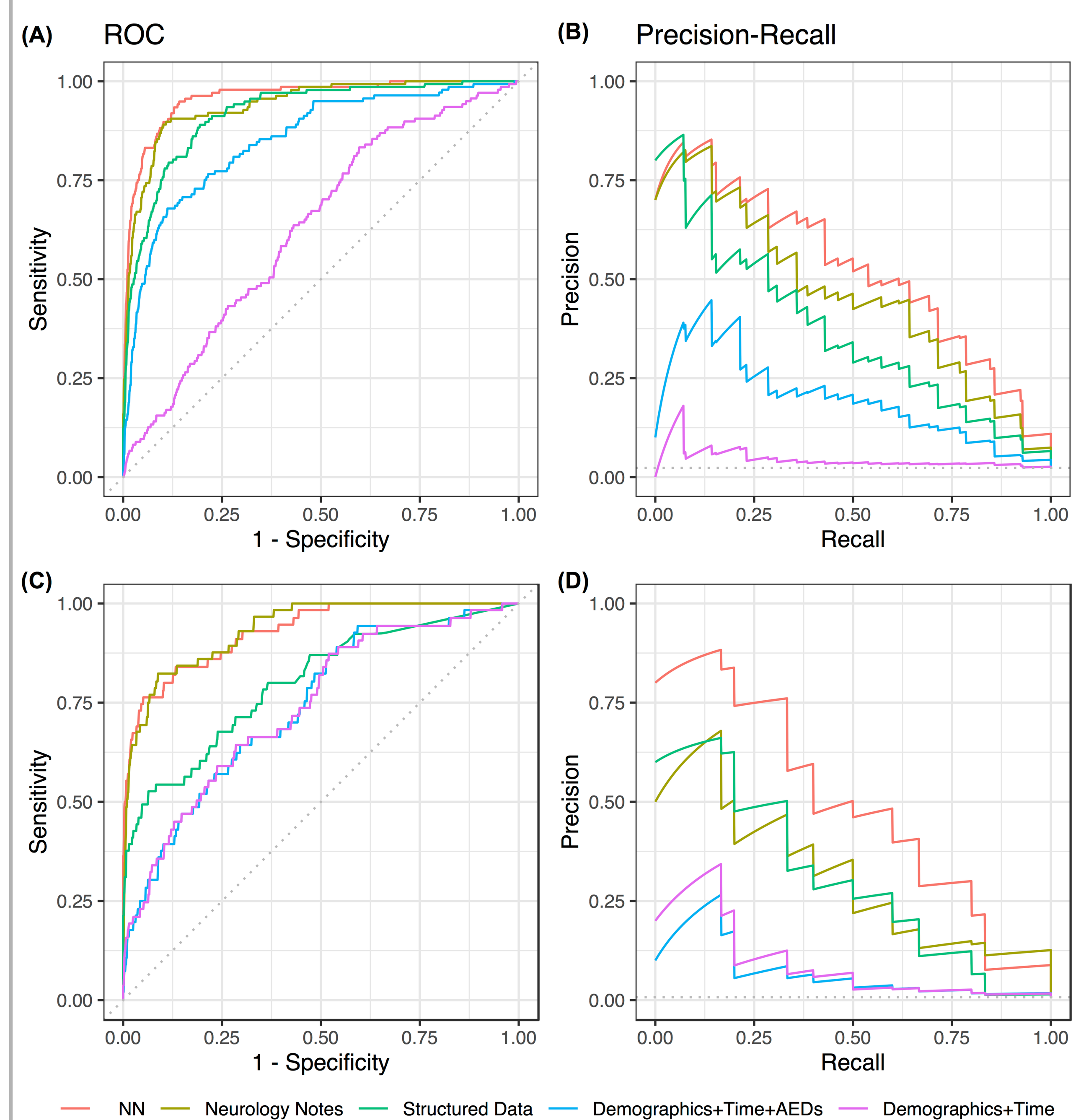

Figure 2: Receiver operating characteristic (ROC) and precision-recall curves for the pediatric (A, B) and adult (C, D) datasets. NN: multimodal neural network shown in Figure 1. "Neurology Notes" and "Structured Data" were unimodal models. "Demographics+Time+AEDs" and "Demographics+Time" were baseline logistic regression models.

Results

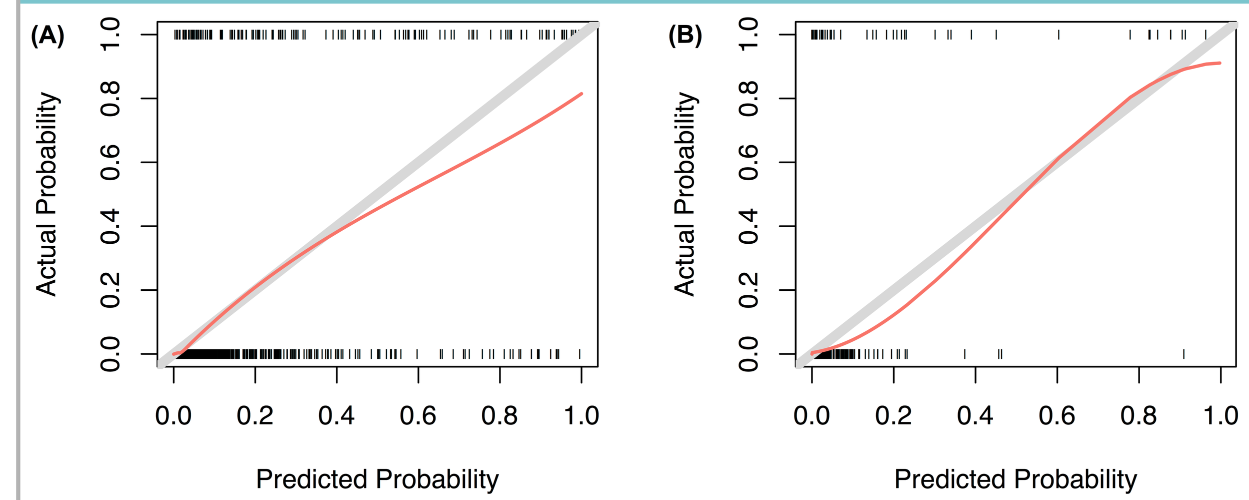

Figure 3: Calibration curves for pediatrics (A) and adults (B).

Scores were calibrated to the observed probability that patients were candidates for surgery.

Clinical workflow:
- If operationalized, the model would screen patients with epilepsy before each neurology visit.
- An alert would be sent to neurology providers before they are scheduled to see a potential candidate for epilepsy surgery.
- **Alerts would facilitate earlier referrals for presurgical evaluations.**

This patient (below) underwent epilepsy surgery after 18 neurology visits. **The model identified them as a surgical candidate 11 visits (three years) earlier than they were referred for a presurgical evaluation.**

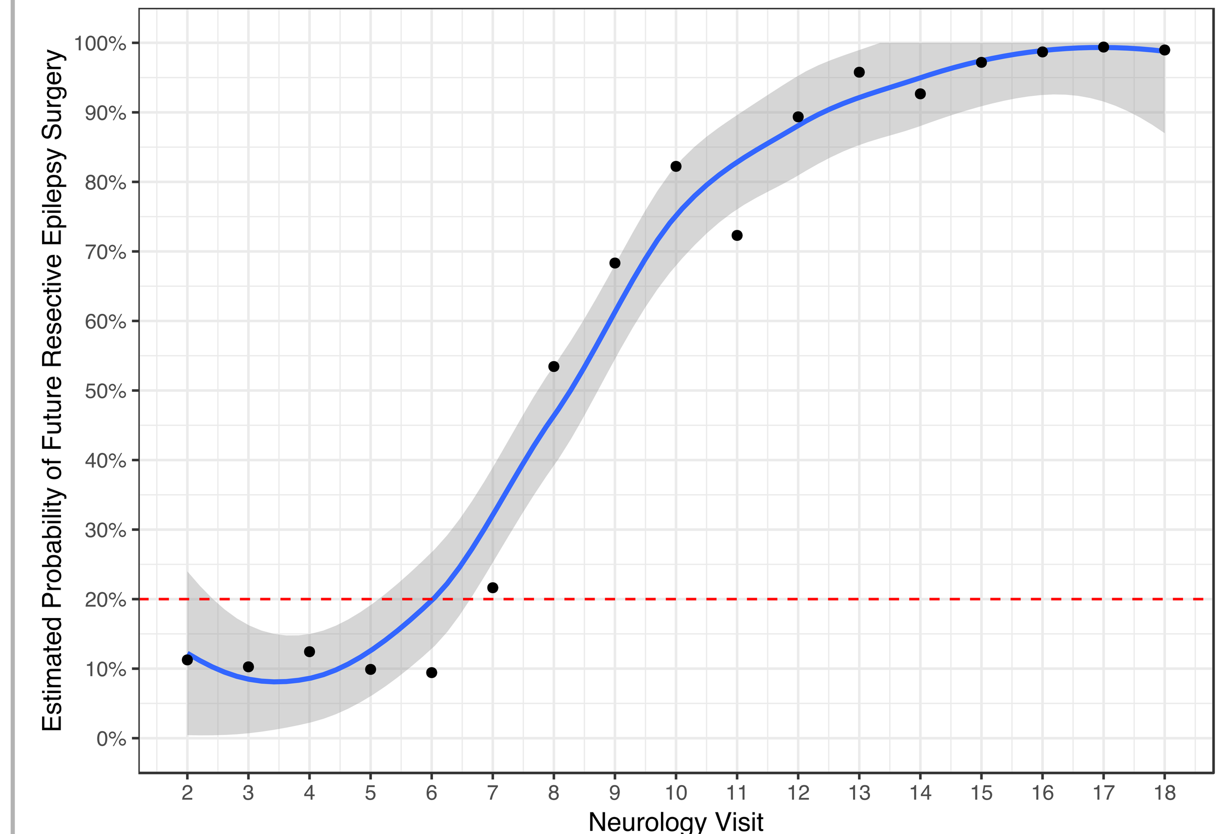

Figure 4: Sample use case. The model retrospectively assigned surgical candidacy scores at every visit. Each dot represents their surgical candidacy score at that neurology visit. The dashed red line represents the threshold surgical candidacy score, above which an alert would be sent to their neurologist.

Discussion and Conclusions

- Fusing the outputs from unimodal ML models increased overall performance.
- Improvements in sensitivity and positive predictive value were in ranges relevant to clinical use.
- Performance was strong despite having few surgical cases to train on.
- Viable alternative to deep learning methods for EHR data that require an order of magnitude more training cases.
- Heavily imbalanced data did not adversely affect this model.

Conclusions:
- Surgical candidates can be identified earlier in the disease course.
- The modeling process, not the model itself, generalized from pediatrics to adults.
- Utilizing multimodal data from the EHR increased performance.

Next Steps:
- Validate prospectively in a clinical setting.
- Determine the optimal methods of sharing the ML predictions with providers (e.g. classifications vs. continuous surgical candidacy scores, longitudinal scores vs. cross-sectional screening)

References

1. Zack MM, Kobau R. National and state estimates of the numbers of adults and children with active epilepsy—United States, 2015. MMWR. Morbidity and mortality weekly report. 2017 Aug 11;66(31):821.
2. Jobst BC, Cascino GD. Resective epilepsy surgery for drug-resistant focal epilepsy: a review. Jama. 2015 Jan 20;313(3):285-93.
3. Lamberink HJ, et al. Seizure outcome and use of antiepileptic drugs after epilepsy surgery according to histopathological diagnosis: a retrospective multicentre cohort study. The Lancet Neurology. 2020 Sep 1;19(9):748-57.
4. Burneo JG, et al. Disparities in surgery among patients with intractable epilepsy in a universal health system. Neurology. 2016 Jan 5;86(1):72-8.
5. Cohen KB, et al. Methodological issues in predicting pediatric epilepsy surgery candidates through natural language processing and machine learning. Biomedical Informatics Insights. 2016 Jan;8:BII-S38308.

Acknowledgements

Research reported in this manuscript was supported by the National Institutes of Health (F31 NS115447 and K25 HL125954) and the Agency for Healthcare Research and Quality (R21 HS024977).

Drs. Greiner and Glauser report a patent pending for the identification of surgery candidates using natural language processing (application num. 16/396,835), licensed to Cincinnati Children's Hospital Medical Center. All other authors report no disclosures.