Original research

Multimodal prognostic features of seizure freedom in epilepsy surgery

Ali Alim-Marvasti, Vejay Niranjan Vakharia, John Sidney Duncan

ABSTRACT

Objective Accurate preoperative predictions of seizure freedom following surgery for focal drug-resistant epilepsy remain elusive. Our objective was to systematically evaluate all meta-analyses of epilepsy surgery with seizure freedom as the primary outcome, to identify clinical features that are consistently prognostic and should be included in the future models.

Methods We searched PubMed and Cochrane using free-text and Medical Subject Heading (MeSH) terms according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses. This study was registered on PROSPERO. We classified features as prognostic, non-prognostic and uncertain and into seven subcategories: ‘clinical’, ‘imaging’, ‘neurophysiology’, ‘multimodal concordance’, ‘genetic’, ‘surgical technique’ and ‘pathology’. We propose a structural causal model based on these features.

Results We found 46 features from 38 meta-analyses over 22 years. The following were consistently prognostic across meta-analyses: febrile convulsions, hippocampal sclerosis, focal abnormal MRI, Single-Photon Emission Computed Tomography (SPECT) coregistered to MRI, focal ictal/interictal EEG, EEG-MRI concordance, temporal lobe resections, complete excision, histopathological lesions, tumours and focal cortical dysplasia type Ilb. Severe learning disability was predictive of poor prognosis. Others, including sex and side of resection, were non-prognostic. There were limited meta-analyses investigating genetic contributions, structural connectivity or multimodal concordance and few adjusted for known confounders or performed corrections for multiple comparisons.

Significance Seizure-free outcomes have not improved over decades of epilepsy surgery and despite a multitude of models, none prognosticate accurately. Our list of multimodal population-invariant prognostic features and proposed structural causal model may serve as an objective foundation for statistical adjustments of plausible confounders for use in high-dimensional models.

PROSPERO registration number CRD42021185232.

INTRODUCTION

Epilepsy surgery can be curative for focal drug-resistant epilepsy, yet in over half of individuals, seizures eventually relapse.1,2 Postsurgical outcomes include seizure freedom, discontinuation of antiseizure medications, neuropsychological and psychiatric outcomes or morbidity. Seizure freedom is the strongest predictor of improved health-related quality of life3 and is classified according to the ILAE or Engel systems.4 These outcomes can be used as ordinal scales, binarised into seizure-free and not seizure-free categories at specified post-operative time points or binarised at each year following surgery to build proportional Hazards models.5,2

Prognostic features can be related to patient characteristics (eg, age, seizure semiology, variability of seizures and genetics), investigation findings (focal lesion on MRI and localising epileptic activity on EEG), surgical factors (resection margins or technique) and combinations of the above (concordance of imaging with neurophysiology). Favourable
clinically relevant prognostic features have been identified from many individual studies, including clearly localising and lateralising semiology, well-circumscribed unilateral, unfocal and temporal lesions, EEG-MRI concordance and complete excision of the evaluated epileptogenic zone.  

Other features are prognostic in some studies but not in others such as focal to bilateral tonic-clonic seizures (FBTCS) and age at seizure onset. A feature may erroneously appear prognostic in a single-centre study due to publication bias or overfitting from investigating many unadjusted variables. Conversely, a feature may appear falsely non-prognostic in small studies due to low statistical power. Most individual studies are small retrospective observational studies from single centres and are prone to such biases.

Meta-analyses aggregate data while accounting for different levels of heterogeneity among patients and between studies. Their strength lies in combining data to achieve greater statistical power while adjusting for heterogeneity and confounders, and attributing weights to studies resulting in summary effect size estimates with wider CIs than unweighted methods.

Nevertheless, accurately predicting seizure freedom prior to surgery has remained elusive. Machine learning models show promise, but have almost entirely been trained on temporal lobe (TL) surgeries. Other recent developments, such as the Epilepsy Surgery Nomogram and the modified Seizure Freedom score, are not better than clinical heuristics which have not resulted in improved surgical outcomes over recent decades.

This highlights the need for a review of the evidence in epilepsy surgery, which we present here by evaluating meta-analyses for prognostic variables in exchange for better generalisability. This list should also preclude the need for further meta-analyses on these features, other than to adjust for potential confounders.

Our objectives are to address these questions:

1. Which features are consistently prognostic, and could be used in models of seizure freedom?
   This list should also preclude the need for further meta-analyses on these features, other than to adjust for potential confounders.
2. Which features do not have prognostic value and could be excluded from future machine-learning models and meta-analyses? This would risk the potential loss of only very weak prognostic variables in exchange for better generalisability.
3. What variables have not been evaluated in meta-analyses and how can we improve postsurgical prognostication?

Methods

Search strategy and Criteria
The study was registered on international prospective register of systematic reviews. The search was conducted in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines on PubMed, MEDLINE and Cochrane updated 1 December 2020, using a combination of free-text and Medical Subject Heading (MeSH) terms. We screened titles and abstracts for inclusion criteria and full texts for exclusion criteria for individual prognostic features. Full search strategy and exclusions are in online supplemental methods.

Inclusion criteria
We included studies for full-text review that were meta-analyses of prognostic features for seizure freedom in epilepsy surgery, the neurosurgical resections had to have been performed for patients with drug-resistant focal epilepsy with curative intent.

Data collection
Two neurologists and a neurosurgeon independently screened articles for inclusion criteria, then one collected data and checked against exclusion criteria (AA-M) and the other two checked decisions. Disagreements were resolved through discussion.

The following data, where available, were extracted for each meta-analysis: investigated feature(s) (whether prognostic or not), specified population (resected lobe, adults, specified lesion), numbers of patients and individual studies for each feature or their upper bounds, definition and duration of seizure freedom, effect sizes and method used (univariate, multivariate logistic regression, fixed effect, random effects, network analysis, meta-regression or other). Qualitative evaluation of certainty of evidence was performed using Grading of Recommendations Assessment, Development and Evaluation (GRADE) guidelines (online supplemental references 11–17).

Trial sequence analyses were assessed for bias using an additional checklist. Where possible, we used the current International League against Epilepsy seizure classification.

Data presentation
Features from the same investigation modality were grouped into seven categories (online supplemental table 2).

Features were further split into essential prognostic features (EPF), uncertain prognostic feature (UPF), and non-prognostic feature (NPF) based on consistency of value across meta-analyses such that if all effect sizes were in the same direction (eg, all favoured postsurgical seizure freedom), then this feature was classified as EPF; whereas, UPF included features that in some meta-analyses favoured seizure freedom, while in others showed no effect or worse outcomes. NPFs were non-significant in all meta-analyses.

Statistical analysis
Effect sizes were inverted such that OR and relative risks over 1 indicate better outcomes favoured good outcome. If effect sizes or CIs were not quoted, these were estimated from the raw data (online supplemental methods). When quoting effect sizes across meta-analyses for the same feature, we used range of effect sizes (ROES) for both point estimates and 95% CIs (min, max).

RESULTS

Overview, PRISMA flowchart and meta-analytical methods
From 50 meta-analyses, 12 were excluded on full-text review, leaving 38 from which data were collected (PRISMA flowchart figure 1). Excluded meta-analyses had lower median numbers of individual studies than those from which data were extracted (11 (IQR 7–22) vs 22 (IQR 15–37)), and lower median number of patients (71 (IQR 33–87) vs 1034 (IQR 320–1999)). The largest number of individual studies in any meta-analysis was 258, and the highest number of included patients was 16 855 from the Cochrane review. Two multicentre studies were included, one from eight centres and another from 37 centres.

The main meta-analytical methods and upper bounds on numbers of studies and patients are summarised in table 1. Online supplemental table 1 lists features from each meta-analysis with GRADE scoring, and online supplemental table 2 categorises these under seven modalities.
Table 2 presents EPF that were consistently prognostic in all meta-analyses, and table 3 shows consistently NPF with individual GRADE scores. Online supplemental tables 3,4 provide more details on EPFs, features with UPFs and NPFs.

**EPF for epilepsy surgery**

Thirteen features were regarded as EPF, as they were consistently prognostic. Three clinical features, from six meta-analyses over 21 years, were severe learning disability including IQ<75, with the largest effect size estimates from the paediatric tuberous sclerosis population (ROES RR 0.26–0.66 (0.14 to 0.94), OR 0.14–0.61 (0.04 to 0.82)), presence of febrile convulsions (RR 1.09 (1.01 to 1.17)) and lack of acute postoperative seizures (OR 4.2 (2.97 to 5.93)) (table 2 and online supplemental table 3).

Prognostic imaging features included the presence of hippocampal sclerosis (RR 1.17 (1.12 to 1.23)) and abnormal Single-Photon Emission Computed Tomography (SPECT) coregistered with MRI (ROES 2.44–3.28 (1.34 to 5.67)). Abnormal MRI was consistently prognostic in 10 meta-analyses with the largest effect sizes from children having hemispherectomies (ROES RR 1.28–1.64 (1.20 to 2.08), OR 1.27–4.6 (1.14 to 16.62)).

**Table 1** The main meta-analytical methods for evaluating prognostic features of epilepsy surgery

<table>
<thead>
<tr>
<th>Type of meta-analysis</th>
<th>Number of meta-analyses</th>
<th>Total number of included individual studies (upper bound)</th>
<th>Total number of patient participants (upper bound)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate (tests of proportions, ANOVA, t-test or crude effect sizes)</td>
<td>9</td>
<td>215</td>
<td>6351</td>
</tr>
<tr>
<td>Proportional Hazards models (Cox regression)</td>
<td>1</td>
<td>19</td>
<td>187</td>
</tr>
<tr>
<td>Fixed or random (mixed) effects models</td>
<td>17</td>
<td>1122</td>
<td>55 502</td>
</tr>
<tr>
<td>Meta-regressions (including logistic regression)</td>
<td>6</td>
<td>372</td>
<td>16 006</td>
</tr>
<tr>
<td>(Bayesian) network analyses (NMA)</td>
<td>4</td>
<td>325</td>
<td>6471</td>
</tr>
<tr>
<td>Hierarchical/multi-level</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other: partial least squares (projection to latent space)</td>
<td>1</td>
<td>20</td>
<td>186</td>
</tr>
</tbody>
</table>

ANOVA, Analysis of Variance; NMA, Network Meta-Analysis.
Neurophysiological features were ictal and interictal (uni-)
focal EEG abnormalities, this effect largely persisted irrespective
of whether the MRI was abnormal or if initial epilepsy surgery had failed (ROES OR 1.55–3.89 (1.14 to 9.08)).

Concordant MRI and EEG abnormalities were consistently
associated with a good prognosis (ROES OR 2.17–4.9 (1.07 to
13.5)). There were no genetics features in EPF.

Surgical technique or anatomic features

Pathological features

Uncertain prognostic features

Eighteen features had mixed results with some meta-analyses
suggesting prognostic value and others suggesting non-
significance; previous head injury, central nervous system (CNS)
infections, focal semiology, infantile spasms, seizure frequency,
age at onset, age at surgery (investigated by 18 separate

The essential prognostic features (EPFs).

See online supplemental table 3 for more details and full list of references.

Multimodal concordance

EEG-MRI concordance

Temporal lobe (vs ET)

Complete excision (vs subtotal
resection)

Presence of tumours

Focal cortical dysplasia (FCD)

Lesional pathology vs non-
lesional

Clinical features

Severe developmental delay
and IQ<75

Febrile convulsions (FC)

Without acute postoperative
seizures (APOS)

Imaging features

Hippocampal sclerosis (HS)

Abnormal or lesional MRI

SPECT: subtraction SPECT-
registered to MRI (SISCOM)

Focal ictal or interictal or
invasive EEG

Neurophysiological features

Table 2 Essential prognostic features for epilepsy surgery (EPF)

<table>
<thead>
<tr>
<th>EPF</th>
<th>Prognostic value and supporting evidence base</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature</td>
<td>Population(s) or subgroup(s)</td>
</tr>
<tr>
<td>Clinical features</td>
<td></td>
</tr>
<tr>
<td>Severe developmental delay and IQ&lt;75</td>
<td>Children and adults, TLE, structural lesions, tuberous sclerosis, hemispherectomies</td>
</tr>
<tr>
<td>Febrile convulsions (FC)</td>
<td>TL and ET in children and adults</td>
</tr>
<tr>
<td>Without acute postoperative seizures (APOS)</td>
<td>Children and adults, TLE and ET</td>
</tr>
<tr>
<td>Imaging features</td>
<td></td>
</tr>
<tr>
<td>Hippocampal sclerosis (HS)</td>
<td>Adults and children with TLE</td>
</tr>
<tr>
<td>Abnormal or lesional MRI</td>
<td>Adults and children with TLE and ET, FCD, frontal lobe, occipital lobe and posterior quadrant epilepsies, hemispherectomies</td>
</tr>
<tr>
<td>SPECT: subtraction SPECT-registered to MRI (SISCOM)</td>
<td>TL and ET</td>
</tr>
<tr>
<td>Neurophysiological features</td>
<td></td>
</tr>
<tr>
<td>Focal ictal or interictal or invasive EEG</td>
<td>Adults, children, repeat resections, MRI-negative TLE, tuberous sclerosis, ET</td>
</tr>
<tr>
<td>Multimodal concordance</td>
<td>TL and ET children and adults, tuberous sclerosis, hemispherectomies</td>
</tr>
<tr>
<td>Surgical technique or anatomic features</td>
<td></td>
</tr>
<tr>
<td>Temporal lobe (vs ET) resections</td>
<td>Adults and children with FCD, repeat surgery, low grade gliomas</td>
</tr>
<tr>
<td>Complete excision (vs subtotal resection)</td>
<td>Adults and children with FCD, FLE, repeat resections, TLE, low grade gliomas</td>
</tr>
<tr>
<td>Pathological features</td>
<td></td>
</tr>
<tr>
<td>Presence of tumours</td>
<td>Children and adults, TLE and ET, gangliogliomas, DNET, neuroepithelial tumours</td>
</tr>
<tr>
<td>Focal cortical dysplasia (FCD)</td>
<td>Adults and children, TLE and ET</td>
</tr>
<tr>
<td>Lesional pathology vs non-lesional</td>
<td>Adults and children, TLE, ET, repeat resections, occipital lobe and posterior quadrant.</td>
</tr>
</tbody>
</table>

The essential prognostic features (EPFs).

See online supplemental table 3 for more details and full list of references. ET, extratemporal; FCD, focal cortical dysplasia; FLE, frontal lobe epilepsy; OR/RR, OR and relative risks over 1 indicate better outcomes; PLS, projection to latent space; TL, temporal lobe; TLE, Temporal Lobe Epilepsy.
### Table 3  Non-prognostic features (NPF)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Population(s) or Subgroup(s)</th>
<th>Non-prognostic evidence base</th>
<th>Individual patients*</th>
<th>Individual studies*</th>
<th>Meta-analytical references</th>
<th>Publication years (first, last)</th>
<th>GRADE score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex: male vs female</td>
<td>Adults and children with FLE, TLE, ET, tuberous sclerosis, MRI neg TLE, repeat surgery, hemispherectomies, low grade gliomas</td>
<td>All were non-significant, a large proportion even on weighted univariate tests, which otherwise tend to overestimate significance. Individual unweighted effect sizes ranged from OR 0.83 (0.42, 1.64) in repeat surgery for focal DRE14 to OR 1.44 (0.86, 2.41) in MRI negative TLE.25</td>
<td>5974</td>
<td>148</td>
<td>Englot, Wang; Englot, Rolston; Zhang, Hu; Fallah, Guyatt; Ibrahim, Morgan; Wang; Zhang; Krucoff, Chan; Englot, Breshears; Hu; Zhang; Shan, Fan; Cao, Liu</td>
<td>2012–2018</td>
<td>+++ Moderate</td>
</tr>
<tr>
<td>Epilepsia partialis continua (EPC)</td>
<td>Children undergoing hemispherectomies</td>
<td>Not significant on unweighted univariate testing and result is from only one meta-analysis.</td>
<td>127</td>
<td>7</td>
<td>Cao, Liu</td>
<td>2016</td>
<td>++ Low</td>
</tr>
<tr>
<td><strong>Imaging features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of cortical tubers</td>
<td>Tuberous sclerosis</td>
<td>Numbers of tubers did not predict outcomes. See online supplemental table 5 for observations on methodology, sample sizes, adjustments and heterogeneity.</td>
<td>286</td>
<td>24</td>
<td>Zhang, Hu; Fallah, Guyatt; Ibrahim, Morgan</td>
<td>2013–2015</td>
<td>++ Low</td>
</tr>
<tr>
<td>Magnetic 1H spectroscopy</td>
<td>TLE adults and children</td>
<td>Probably no more valuable than conventional MRI abnormality (online supplemental table 5).</td>
<td>121</td>
<td>22</td>
<td>Willmann, Wennberg</td>
<td>2006</td>
<td>+ Low</td>
</tr>
<tr>
<td>Encephalomalacia</td>
<td>Adults and children</td>
<td>Encephalomalacia was NS in the Cochrane meta-analysis, it was also not significant on subgroup analyses.22</td>
<td>317</td>
<td>5</td>
<td>West, Nevitt</td>
<td>2019</td>
<td>+ Very low</td>
</tr>
<tr>
<td>Enhancement, oedema, and/or mass effect</td>
<td>Low grade gliomas in adults</td>
<td>These combined features are not clinically prognostic of low-grade glioma resection for seizure freedom. Although NS, the point estimate and CI are unavailable.</td>
<td>2641</td>
<td>23</td>
<td>Shan, Fan</td>
<td>2018</td>
<td>+ Very low</td>
</tr>
<tr>
<td><strong>Vascular lesions</strong></td>
<td>Adults and children with TL and ET</td>
<td>Only one meta-analysis investigated this in 2004, comprising only three individual studies, its pathological counterpart was also NS.26</td>
<td>&lt;&lt;3511</td>
<td>3</td>
<td>Tonini, Beghi</td>
<td>2004</td>
<td>+ Very low</td>
</tr>
<tr>
<td>Neurophysiological features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intraoperative invasive EEG</td>
<td>Children and adults with FLE</td>
<td>Electro-corticography did not effect outcomes</td>
<td>1024</td>
<td>21</td>
<td>Englot, Wang</td>
<td>2012</td>
<td>+++ Moderate</td>
</tr>
<tr>
<td>Video telemetry and long-term monitoring</td>
<td>Children and adults with FLE, lesional and non-lesional TLE and ET</td>
<td>Lesional TLE cases do well and this was the only subgroup in which long-term monitoring had a point estimate similar to estimate greater than 1.</td>
<td>1738</td>
<td>65</td>
<td>Englot, Wang; Kobulashvili, Kuchukhidze</td>
<td>2012, 2018</td>
<td>+ Very low</td>
</tr>
<tr>
<td><strong>Surgical technique or anatomic features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mesial vs lateral TL focus</td>
<td>MRI neg TLE</td>
<td>Mesial or lateral TLE, as determined by sEEG, subdural grids, or ATL/SAH vs neocortectomy, are not significant.</td>
<td>92</td>
<td>8</td>
<td>Wang, Zhang</td>
<td>2016</td>
<td>+ Very low</td>
</tr>
<tr>
<td>Side of resection (left vs right)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal, central, or posterior resections vs other</td>
<td>ET, adults, non-lesional</td>
<td>Not prognostic</td>
<td>81</td>
<td>?</td>
<td>Ansari, Tubbs</td>
<td>2010</td>
<td>+ Very low</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>NPF features</th>
<th>Non-prognostic evidence base</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographical location of surgery</td>
<td>Non-prognostic evidence base</td>
</tr>
<tr>
<td>Tuberous sclerosis in children</td>
<td>Only one meta-analysis investigated North America vs Elsewhere, and the GRADE score is from this meta-analysis alone.</td>
</tr>
<tr>
<td>Ibrahim, Morgan</td>
<td>2015</td>
</tr>
<tr>
<td>Pathological features</td>
<td></td>
</tr>
<tr>
<td>Neuro-migrational defects</td>
<td>TL and ET children and adults</td>
</tr>
<tr>
<td>Tonini, Beghi</td>
<td>2004</td>
</tr>
<tr>
<td>Astrocytoma vs non-astrocytoma</td>
<td>Low grade gliomas in adults</td>
</tr>
<tr>
<td>Shan, Fan</td>
<td>2018</td>
</tr>
</tbody>
</table>

See online supplemental table 5 for more details and online supplemental materials for full list of references.

*Upper bound of estimate, not including subgroup analyses.

**Multivariate; **univariate; **calculated (usually unweighted) effect size; ET, extratemporal; FLE, frontal lobe epilepsy; MCD, malformations of cortical development; NS, not significant; PLS, projection to latent space; TL, temporal lobe.

---

**DISCUSSION**

We identified 46 features from 38 meta-analyses on prognostic factors for epilepsy surgery, only 15 of which were in the 2019 Cochrane review. We categorised features that were consistently prognostic when unadjusted, and then considered how these might be used in future models, to improve personalised prognosis and outcomes (table 3 and online supplemental table 4).

---

**EFP for epilepsy surgery**

EFP is a minimum essential list based on current best evidence. Our objective was to determine a minimum list of features for use in future models, to improve personalised prognosis and outcomes (EFP). We further propose grouping these features into an a priori SCM to determine if it would be appropriate to adjust for these variables in future studies (see SCM in online supplemental table 5). We propose grouping these features into an a priori SCM to determine if it would be appropriate to adjust for these variables in future studies (see SCM in online supplemental table 5).
When lesions were defined by pathology, MRI abnormalities still had a non-significant trend to higher rates of seizure freedom.14 Another meta-analysis found a prognostic trend for abnormal pathology even in MRI-negative TLE (p=0.06, OR=1.36 (0.7 to 2.63)).25 Within lesional-epilepsies, FCD type IIb is further associated with better outcomes.31 Although it is well established that lesional epilepsies have better postsurgical outcomes,14 21 22 25 30 32 33  and that complete lesionectomy is required15 21 22 31 34 35 the overwhelming majority of studies did not adjust for these.

A meta-analysis of 1999 patients across 35 articles, found outcomes after stereo-electroencephalogram (SEEG), were better than after subdural grids in patients undergoing temporal resections with lesional-MRI (seizure freedom for subdural grid (51.5% to 61.9%) vs SEEG (64.4% to 81.6%)).36 Such a comparison is limited by ascertainment bias and the differing indications for these methods. Interactions between features have not been formally investigated in meta-analyses, except in specific subpopulations and imaging-EEG concordance.

Multimodal concordance
Five meta-analyses attested the value of concordant MRI and EEG,22 28 37–39 but none looked at the value of semiological concordance with other modalities. In our SCM, the prominent causal pathway node is multimodal concordance, which should be further studied as a valuable predictor of seizure freedom (figure 2).

Features of uncertain significance
Even meta-analyses may be underpowered, contributing to lack of statistical significance (online supplemental reference 64). PET results were mixed, but when concordant with EEG, PET could predict good seizure outcomes in non-lesional TLE with a positive predictive value of 71% (online supplemental reference 27).

Most meta-analyses reported non-significance of age at seizure onset, age at surgery and duration of epilepsy; however, there is a mixed picture. For every extra year of duration of epilepsy at time of surgery, one metaregression reported overall odds of seizure freedom reduced by a factor of 0.83 and another analysing data from 1545 patients across 12 studies found shorter duration of epilepsy was associated with higher rates of postsurgical seizure freedom with RR ranging from 1.20 to 1.33 (online supplemental reference 57). Conversely, age at surgery and duration of epilepsy before surgery have been documented as having ‘no association’ with outcomes.27 Longer duration of epilepsy may result in worse surgical outcomes due to selection bias (more difficult cases being deferred) or progressive cerebral damage. Strikingly, these three features of age at onset, age at surgery and duration of epilepsy have not been explored for three-way interactions.

Uncertain features can be reclassified into the essential or non-prognostic categories, when future studies that evaluate their value adjust for essential prognostic variables. While such models would clarify to what extent uncertain features may be prognostic over and above the essential features, this may not always be clinically desirable. For example, CNS infections may result in glial scars, and adjusting for imaging lesions may not be clinically desirable. Instead, an SCM could be used (see five-step plan below).

Non-prognostic features
Side of resection and sex were both investigated in 11 meta-analyses but were not prognostic, consistent with a meta-review from 2013.29 Nevertheless, studies have continued to investigate them. Their use in predictive models risks overfitting and compromising generalisability.

Prognostication: common pitfalls and recommendations
Unmodelled features
As there has not been significant improvement in postoperative outcomes, there are likely to be variables that have not been included.14 13 This is problematic for two reasons. First, studies
are unable to adjust for unknown confounders. Second, without these features, individualised predictions will not be accurate. It is, therefore, critical to discuss notable missing features.

No meta-analysis has investigated the role of family history or detailed seizure semiology despite the fact that monitoring seizure semiology is integral to presurgical evaluation. Five meta-analyses reviewed MRI-EEG concordance, but none considered semiological concordance; the closest corollaries were FBTCs, epilepsy partialis continua and epileptic spasms. Future studies should evaluate interactions between semiology, epileptogenic zone, imaging and neurophysiology in patients with both favourable and unfavourable surgical outcomes.

The importance of genetics in seizure-free outcomes is belied by relatively few publications. Individuals with mutations affecting synaptic transmission or ion channels (5 articles, 14 patients) were less likely to benefit from epilepsy surgery than those with mutations in the mechanistic target of rapamycin (mTOR) pathway (10 articles, 30 patients). This was despite six of eight patients with SCN1A mutations having concordant semiology and colocalised MRI lesions. This meta-analysis was the only one to investigate genetics but it met our exclusion criteria as a large proportion of the small samples were lesional and no attempt at adjustments had been made (online supplemental tables). High-frequency oscillations and fast ripples were also excluded in our final synthesis (figure 1) due to lack of appropriate effect sizes (online supplemental references 1, 6 and online supplemental table 1). This should impel us towards multicentre data sharing in comprehensive models (figure 2).

Towards personalised seizure freedom predictions

Meta-analyses have been widely used for over five decades to quantitatively integrate a collection of studies. They are useful to identify important features based on best-available evidence, but cannot identify new features or provide personalised quantitative prognostication. The majority of studies did not statistically correct for multiple comparisons, potentially introducing false positives.

Machine learning models and nomograms have been proposed to predict outcomes, without prospective validation. These models included three features of uncertain significance (duration of epilepsy, frequency of seizures and generalised seizures), one non-prognostic factor (sex) and one EPF (pathological aetiology); it is perhaps unsurprising that the model was not generalisable. We advocate, therefore, that to improve prognostication and outcomes, a five-step plan is adopted:

1. All relevant factors for epilepsy surgery outcome prediction are curated in an agreed international, multicentre endeavour, which include the essential prognostic list curated here. Practically, the preoperative clinical variables should take precedence over postoperative features, for example, severe developmental delay should take priority over acute postoperative seizures and FCD type IIb as the latter two are only known after surgery. The final curated features would then form the starting point for building predictive models.

2. An SCM is devised that links outcomes to prognostic factors, to enable adjusting for EPFs when investigating other variables.

3. Identification of the degree to which polygenic risk scores, family history, seizure semiology and concordance may contribute to outcomes as indirect measures of seizure focality within the SCM.

4. Curation of an international multicentre, high-quality, anonymised retrospective and prospective data set of patients who have undergone epilepsy surgery with features and outcomes, similar to the retrospective collaboration on surgical histopathology.

5. Machine learning models suitable for binary features and outcome classification on the international dataset.

The current study addressed the first two steps including R code to generate and amend SCMs (see online supplemental materials section on SCM for details on R codes for a simplified and complete SCM, and the two online supplemental files: SCM dagitty V.5 super simplified and SCM dagitty V4). We can verify the value of EPFs and the SCM by building high-dimensional predictive models from international collaborations using SCM to adjust for covariates, subsequently showing that the resulting model predicts outcomes better than current methods.

Limitations

Meta-analyses were our unit of analysis, each assuming sufficient homogeneity for estimation of pooled effects. Only English-language articles were searched and we did not check for overlaps between meta-analyses, we, therefore, quote upper limits of numbers of patients and individual studies. We adopted the same definitions of seizure freedom in terms of Engel or ILAE class and duration of follow-up as the meta-analyses, but inconsistent definitions and differing durations meant that we could not adjust for these. Most studies defined seizure freedom as Engel I, potentially compromising results, as this includes patients with ongoing seizures, implying incomplete resection of the epileptogenic zone or multifocal epilepsy.

Meta-analyses improve power, but unless they are hierarchical, lose the granularity of applicability to subgroups. To reduce type I errors, we did not include variables that were significant on unweighted tests, but this can reduce power. Nevertheless, moderate or low-quality evidence from meta-analyses can lead to strong assertions on whether a feature is prognostic (online supplemental reference 13).

Many variables in individual articles of epilepsy surgery outcomes are clinically widely used, contributing to a circular logic, whereby features already considered significant are pooled in meta-analyses. This is why we also discussed unmodelled features.

Whether a feature is of positive or negative prognostic value may be comparable across meta-analyses but due to differing patient populations and seizure-free definitions, diversity of models, unadjusted confounders and unobserved heterogeneity, the magnitude will almost certainly not be, precluding comparisons of effect sizes. Cochrane-Mantel-Haenszel stratification, multinomial logistic regression or projection to latent space attempt to adjust for between-feature correlations; nevertheless, this mitigation is limited if important features are omitted. By not fully adjusting for covariates such as focal MRI abnormality or duration of follow-up, incorrect conclusions may be drawn. This limitation is well known but has not been universally addressed with a definitive set of prognostic features—which was the objective of this study.
As we looked at shared prognostic features across all types of operations and anatomical lobes, our minimum list of EPFs may under-identify variables that may be prognostic for a particular type of operation but not another, such as a selective amygdalohippocampectomy as opposed to anterior TL resection. These variables can be identified by further predictive models that adjust for confounders using this list of EPFs. Ultimately interaction terms (deep machine learning models) could adequately stratify seizure freedom.

**SUMMARY AND CONCLUSIONS**

Personalised prognostication in epilepsy surgery outcomes has remained elusive and outcomes have not improved with time. We curated features into prognostic and uncertain groups and conclude that more meta-analyses on these are not needed; rather, we need predictive models that quantify their relative contributions to outcomes. We proposed a five-step plan towards personalised seizure freedom predictions and addressed the first two steps in this study. EPFs would be particularly useful in machine learning models of a big-data international collaboration to better predict epilepsy surgery outcomes.

**Twitter** Ali Alim-Marvasti @AliMarvasti and Vejay Niranjan Vakharia @vejayvakharia

**Acknowledgements** We thank Narges Marvasti of NMar Graphic Design for help with illustrations.

**Contributors** AA-M conceived the study and conducted the literature search. All three authors screened titles and abstracts for inclusion criteria. AA-M screened for exclusion criteria and VN and JSD checked decisions independently. AA-M wrote the manuscript, VN and JSD edited the manuscript. JSD conceived and supervised the study. AA-M acts as guarantor. All authors approved the final manuscript.

**Funding** This study was funded by Wellcome/EPSCR Centre for Interventional and Surgical Sciences (WEISS) (203 145Z/16/Z).

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** The study has been approved by the Ethical Committee of the University of Birmingham. All patients had undergone informed consent or assent procedures according to local regulations.

**Data availability statement** Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary material. The list of titles and abstracts screened for inclusion criteria are available upon request as excel files.

**Supplemental material** This content has been supplied by the author(s).

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

**ORCID iDs**

Ali Alim-Marvasti http://orcid.org/0000-0002-7811-0344

Vejay Niranjan Vakharia http://orcid.org/0000-0002-9476-4225

**REFERENCES**


16. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016;94:485–514.


